С. КАРЛИН

ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

ИЗДАТЕЛЬСТВО · МИР ·

A FIRST COURSE IN STOCHASTIC PROCESSES

Samuel Karlin

Department of Mathematics
Stanford University
Stanford, California

ACADEMIC PRESS
New York and London
1968

С. КАРЛИН

ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Перевод с английского
В. В. КАЛАШНИКОВА

Под редакцией
И. Н. КОВАЛЕНКО

ИЗДАТЕЛЬСТВО «МИР»
Москва · 1971

УДК 519.2

Книга С. Карлина является связующим звеном между элементарным курсом теории вероятностей и специальными курсами теории случайных процессов, которые используют сложный аппарат современной математики. Для чтения книги практически достаточно знания математики в объеме стандартного курса высших учебных заведений. Наряду с изложением математического аппарата книга содержит прекрасный набор приложений к биологии, задачам массового обслуживания и др. вопросам.

Книга представляет интерес как для математиков, интересующихся приложениями, так и для биологов, инженеров и специалистов других областей науки, в которых математика находит свое применение.

Редакция литературы по математическим наукам

Инд. $\frac{2\text{-}2\text{-}3}{27\text{-}71}$

С. Карлин

ОСНОВЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

Редакторы *В. Ф. Пахомов* и *Л. Б. Штейнпресс* Художник *Г. И. Мануйлов*
Художественный редактор *В. И. Шаповалов* Технический редактор *Г. Б. Алюлина*
Корректор *Г. И. Секачева*

Сдано в набор 13/I 1971 г. Подписано к печати 9/VIII 1971 г. Бумага № 1 60×90[1]/[16]=
=16,75 бум. л.; печ. л. 33,5; уч.-изд. л. 31,61. Изд. № 1/5825. Цена 2 р. 52 к. Зак. 939.

ИЗДАТЕЛЬСТВО «МИР» Москва, 1-й Рижский пер., 2

Ордена Трудового Красного Знамени Ленинградская типография № 2 им. Евгении Соколовой
Главполиграфпрома Комитета по печати при Совете Министров СССР,
Измайловский проспект, 29.

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Теория случайных процессов, возникшая в результате построения математических моделей реальных физических процессов, в настоящее время представляет собой наиболее содержательную и более всего используемую в приложениях часть теории вероятностей. Классическим разделом теории случайных процессов является теория стационарных (в широком смысле) процессов, которая дает основу решения многих прикладных задач. Задачи современной науки и техники выдвинули на первый план проблемы, связанные с исследованием случайных процессов, порождаемых последовательностями независимых случайных величин. Это цепи Маркова, марковские процессы со счетным множеством состояний, процессы восстановления, полумарковские процессы. Их роль объясняется в значительной степени тем, что реальные процессы, изучаемые с помощью вероятностных методов, по самой своей природе связаны с чередованием событий случайной продолжительности (например, процесс функционирования резервированной системы с заменой отказавших элементов). Интересно отметить, что и те физические процессы, которые ранее изучались в терминах математического ожидания и корреляционной функции, например флуктуационные процессы в радиотехнических устройствах, теперь в соответствии с новыми задачами рассматриваются с точки зрения случайных последовательностей, отражающих определенные изменения состояний процесса. Из такого рода задач возникает теория потоков однородных событий, порождаемых случайными процессами.

Книга известного ученого С. Карлина, предлагаемая вниманию читателя, посвящена основам теории марковских процессов и процессов, связанных со случайными блужданиями, а также применениям этих процессов к задачам генетики, экологии и массового обслуживания. Наиболее ценное в книге — изложение основных аналитических методов исследования соответствующих процессов. Читатель, стремящийся изучить теорию случайных процессов, найдет здесь аппарат, повседневно применяемый специалистами, а не только собрание готовых результатов. Вместе с тем следует отметить, что рассматриваемые автором классы процессов все же являются достаточно частными: не только в теоретических, но и в прикладных работах исследуются более общие классы (например, классы процессов, описывающих функционирование систем массового обслуживания). Однако это не является недостатком книги: методы, действие которых продемонстрировано на простых примерах, могут служить читателю и в более сложной ситуации; бóльшую же часть интересующих сейчас прикладников математических задач можно решать и в пределах тех моделей, которые даются в книге.

В конце каждой главы автор поместил задачи; читателю, естественно, следует их решать.

Книга С. Карлина может быть рекомендована математикам, физикам, специалистам по исследованию операций, биологам и вообще всем, желающим войти в проблематику теории случайных процессов и овладеть методами этой теории. При переводе и редактировании устранены некоторые погрешности оригинального издания.

И. Н. Коваленко

ПРЕДИСЛОВИЕ

Теория случайных процессов изучает последовательности событий, управляемых вероятностными законами. Она находит многочисленные приложения в физике, технике, биологии, медицине, психологии и других дисциплинах, а также в различных разделах математики. Назначение этой книги — дать введение многочисленным специальным руководствам по случайным процессам. При ее написании я преследовал три цели: во-первых, дать систематическое вводное изложение некоторых основных разделов теории случайных процессов, во-вторых, привлечь внимание тех, кто занимается чистой математикой, к богатому многообразию приложений теории случайных процессов и, в-третьих, для читателя, интересующегося приложениями, подчеркнуть важность «математических тонкостей», показать, что они зачастую связаны с самой природой вероятностных процессов.

Примеры в этой книге в основном заимствованы из биологии и техники; вместе с тем везде делается акцент на тех вероятностных аспектах, которые важны или представляют математический интерес для более чем одной дисциплины. В книге обсуждается и иллюстрируется ряд понятий и проблем, привлекающих в настоящее время внимание исследователей.

Поскольку в элементарной книге невозможно охватить все основные разделы теории случайных процессов, нам пришлось опустить некоторые важные темы и среди них такие, как стационарные случайные процессы и мартингалы. Не предполагалось, что эта книга будет служить исчерпывающим руководством по вопросам, затронутым в ней. Напротив, она должна рассматриваться прежде всего как промежуточное звено между элементарными курсами теории вероятностей и многими превосходными работами по случайным процессам, высокий математический уровень которых делает их недоступными для читателей, знакомых лишь с основами теории вероятностей.

У читателей предполагается знакомство с началами теории вероятностей в объеме первого тома ставшей уже классической книги Феллера «Введение в теорию вероятностей и ее приложения». В § 1 гл. 1 мы даем сводку основных свойств случайных величин и функций распределения, а также вводим важнейшие термины.

Материал, набранный мелким шрифтом, при первом чтении можно опустить. В конце каждой главы приводятся задачи, цель которых — разъяснить, а во многих случаях и развить изложенную теорию.

Книгу можно использовать для полугодового или годового курса в зависимости от потребностей. Логическая взаимозависимость глав отражена на схеме, приведенной ниже.

При написании книги я пользовался обширной литературой по случайным процессам. Каждая глава завершается списком работ, в которых читатель найдет дальнейшую информацию и библиографию.

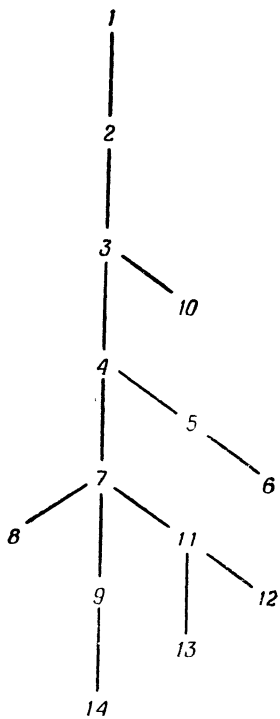
Я благодарен коллегам: профессору Чжун Кай-лаю и профессору Дж. Мак-Грегору (Станфордский университет) за советы и полезные комментарии, профессору Дж. Ламперти (Дартмусский университет), профессору Дж. Киферу (Корнеллский университет) и профессору П. Нею (Висконсинский университет) за конструктивную критику, доктору А. Файнстейну за тщательную проверку значительной части рукописи, а также моим студентам П. Милчу, Б. Сигеру, М. Фелдману и Б. Кришнамурти за их полезные советы и помощь при выборе и составлении задач.

Сэмюэл Карлин

Станфорд, Калифорния

ЛОГИЧЕСКАЯ ЗАВИСИМОСТЬ ГЛАВ

§ 1 гл. 1 можно читать, не вдаваясь в детали. Глава 12 почти не зависит от гл. 11. Первая половина гл. 6 не зависит от гл. 5. Для понимания глав 4—14 не являются необходимыми § 5 и § 6 гл. 3.



Для полугодового курса рекомендуется следующий вариант: § 2 и § 3 гл. 1 с предшествующим кратким изложением материала § 1; гл. 2 полностью; гл. 3, исключая § 5 и, быть может, § 6; гл. 7 без § 3 и, возможно, § 7. Содержание последней части курса составляется по усмотрению лектора, при этом можно использовать гл. 9—12, полностью или частично.

Глава 1

ЭЛЕМЕНТЫ ТЕОРИИ СЛУЧАЙНЫХ ПРОЦЕССОВ

§ 1. СВОДКА ОСНОВНЫХ ТЕРМИНОВ И СВОЙСТВ СЛУЧАЙНЫХ ВЕЛИЧИН И ФУНКЦИЙ РАСПРЕДЕЛЕНИЯ

В этом параграфе приводятся основные термины и элементарные понятия теории вероятностей. В последующих главах мы будем пользоваться ими без каких-либо дальнейших ссылок на литературу. Читателю настоятельно рекомендуется обратиться к упражнениям, помещенным в конце главы; эти упражнения помогут ему вспомнить и закрепить предварительный материал. Детальное изложение этих вопросов можно найти в книгах Феллера, Гнеденко и Парзена (см. литературу в конце этой главы).

Предполагается, что читатель знаком со следующими понятиями:

(1) действительная случайная величина X;
(2) функция распределения F случайной величины X (определяемая как $F(\lambda) = \mathrm{P}\{X \leqslant \lambda\}$) и ее элементарные свойства [1]);
(3) события, связанные со значениями случайной величины X, и их вероятности;
(4) математическое ожидание $\mathrm{M}\{X\}$ случайной величины X и моменты высших порядков $\mathrm{M}\{X^n\}$;
(5) формула полной вероятности и формула Байеса.

Вместо слов «действительная случайная величина» мы будем часто пользоваться сокращением «д. с. в.» [2]).

Д. с. в. X называется *дискретной*, если существует конечное или счетное множество различных чисел $\lambda_1, \lambda_2, \ldots$, такое, что $a_i \equiv \equiv \mathrm{P}\{X = \lambda_i\} > 0$, $i = 1, 2, \ldots$ и $\sum\limits_i a_i = 1$. Если $\mathrm{P}\{X = \lambda\} = 0$ для любого значения λ, д. с. в. X называется *непрерывно распределенной*. Если существует неотрицательная функция $p(t)$, определенная на всей оси $-\infty < t < \infty$ и такая, что функцию распределения F

---

[1]) В отечественной литературе функция распределения, как правило, понимается в смысле $\mathrm{P}\{X < \lambda\}$. Очевидно $\mathrm{P}\{X \leqslant \lambda\} = \lim\limits_{\varepsilon \downarrow 0} \mathrm{P}\{X < \lambda + \varepsilon\}$, $\mathrm{P}\{X < \lambda\} = \lim\limits_{\varepsilon \downarrow 0} \mathrm{P}\{X \leqslant \lambda - \varepsilon\}$, так что между обоими определениями существует простое соответствие. — *Прим. ред.*

[2]) Иногда будет использоваться сокращение «с. в.» — «случайная величина». — *Прим. перев.*

д. с. в. X можно представить в виде

$$F(\lambda)=\int_{-\infty}^{\lambda} p(t)\,dt,$$

то будем говорить, что p является плотностью распределения вероятностей (или короче — плотностью вероятности) случайной величины X. Если д. с. в. X имеет плотность вероятности, то она с необходимостью непрерывно распределена; однако известны примеры непрерывно распределенных случайных величин, не обладающих плотностью вероятности [1].

Если X — дискретная д. с. в., то ее m-й момент (или момент порядка m) определяется так:

$$\mathrm{M}[X^m]=\sum_i \lambda_i^m \mathrm{P}\{X=\lambda_i\}$$

(где λ_i имеют тот же смысл, что и выше) при условии, что ряд сходится абсолютно.

Если X — непрерывно распределенная случайная величина с плотностью вероятности $p(\cdot)$, то ее m-й момент определяется соотношением

$$\mathrm{M}[X^m]=\int_{-\infty}^{\infty} x^m p(x)\,dx$$

при условии, что интеграл сходится абсолютно.

Первый момент (или *среднее значение*) д. с. в. X будем обозначать через m_X или μ_X; m-й центральный момент д. с. в. X определяется как m-й момент д. с. в. $X-m_X$, если m_X существует. Первый центральный момент, очевидно, всегда равен нулю; второй центральный момент называется *дисперсией* σ_X^2 д. с. в. X. *Медианой* д. с. в. X по определению является любое число ν, обладающее тем свойством, что $\mathrm{P}\{X\geqslant\nu\}\geqslant 1/2$ и $\mathrm{P}\{X\leqslant\nu\}\geqslant 1/2$.

А. Совместные функции распределения

Пусть (X, Y) — пара случайных величин; их совместная функция распределения является функцией двух действительных переменных и определяется как

$$F(\lambda_1, \lambda_2)=F_{XY}(\lambda_1, \lambda_2)=\mathrm{P}\{X\leqslant\lambda_1,\ Y\leqslant\lambda_2\}.$$

(Индексы X и Y обычно опускаются, если нет опасности возникновения путаницы.)

---

[1] В отечественной литературе принято называть случайную величину непрерывной именно в случае существования плотности. — *Прим. ред.*

Функция $F(\lambda_1, +\infty) \equiv \lim\limits_{\lambda_2 \to \infty} F(\lambda_1, \lambda_2)$ является функцией распределения и ее называют *частной* (или *маргинальной*) *функцией распределения* д. с. в. X. Аналогично функция $F(+\infty, \lambda_2)$ называется частным распределением д. с. в. Y. Если случится так, что $F(\lambda_1, +\infty) \cdot F(+\infty, \lambda_2) = F(\lambda_1, \lambda_2)$ для любой пары значений параметров λ_1 и λ_2, то случайные величины X и Y называются *независимыми*. Говорят, что совместная функция распределения F_{XY} имеет (совместную) плотность вероятности, если существует функция $p_{XY}(s, t)$ двух действительных переменных, такая, что

$$F_{XY}(\lambda_1, \lambda_2) = \int\limits_{-\infty}^{\lambda_2} \int\limits_{-\infty}^{\lambda_1} p_{XY}(s, t)\, ds\, dt$$

для всех λ_1, λ_2. Если X и Y — независимые д. с. в., то $p_{XY}(s, t)$ с необходимостью представима произведением $p_X(s) p_Y(t)$, где p_X и p_Y — плотности вероятности частных распределений д. с. в. X и Y соответственно.

Совместная функция распределения любого конечного набора $X_1, \ldots, X_n$ д. с. в. определяется как

$$F(\lambda_1, \ldots, \lambda_n) = F_{X_1, \ldots, X_n}(\lambda_1, \ldots, \lambda_n) = \mathrm{P}\{X_1 \leqslant \lambda_1, \ldots, X_n \leqslant \lambda_n\}.$$

Функция распределения

$$F_{X_{i_1}, \ldots, X_{i_k}}(\lambda_{i_1}, \ldots, \lambda_{i_k}) = \lim_{\lambda_i \to \infty,\ i \neq i_1, \ldots, i_k} F(\lambda_1, \ldots, \lambda_n)$$

называется частной функцией распределения случайных величин $X_{i_1}, \ldots, X_{i_k}$.

Если $F(\lambda_1, \ldots, \lambda_n) = F(\lambda_1) F(\lambda_2) \ldots F(\lambda_n)$ для всех значений $\lambda_1, \lambda_2, \ldots, \lambda_n$, то д. с. в. $X_1, \ldots, X_n$ называются независимыми.

Говорят, что совместная функция распределения $F(\lambda_1, \ldots, \lambda_n)$ имеет плотность вероятности, если существует неотрицательная функция $p(t_1, \ldots, t_n)$ от n переменных, такая, что

$$F(\lambda_1, \ldots, \lambda_n) = \int\limits_{-\infty}^{\lambda_n} \ldots \int\limits_{-\infty}^{\lambda_1} p(t_1, \ldots, t_n)\, dt_1 \ldots dt_n$$

для всех действительных $\lambda_1, \ldots, \lambda_n$.

Б. Условные распределения и плотности

Пусть X и Y — д. с. в., принимающие счетные множества значений, скажем $1, 2, \ldots$. *Условная вероятность* $\mathrm{P}\{X = i \mid Y = j\}$ определяется соотношением

$$\mathrm{P}\{X = i \mid Y = j\} = \frac{\mathrm{P}\{X = i,\ Y = j\}}{\mathrm{P}\{Y = j\}}.$$

Здесь предполагается, что $\mathrm{P}\{Y=j\}>0$; в противном случае значение $\mathrm{P}\{X=i \mid Y=j\}$ произвольно (скажем, равно нулю).

Условные вероятности могут быть определены и для случайных величин других типов. Мы рассмотрим только случай, когда X и Y имеют совместную плотность вероятности $p_{XY}(s, t)$. Тогда условное распределение

$$\mathrm{P}\{X \leqslant a \mid Y=b\}$$

определяется формулой

$$\mathrm{P}\{X \leqslant a \mid Y=b\}=(p_Y(b))^{-1} \int_{-\infty}^{a} p_{XY}(s, b)\, ds,$$

если $p_Y(b)>0$, и произвольно в противном случае. Из определения $p_Y(b)$ следует, что

$$\lim_{a \to \infty} \mathrm{P}\{X \leqslant a \mid Y=b\}=1$$

для всех b, таких, что $p_Y(b)>0$. Таким образом, $F^{\{Y=b\}}(a)= =\mathrm{P}\{X \leqslant a \mid Y=b\}$ представляет собой функцию распределения в случае $p_Y(b)>0$. Легко видеть, что эта функция распределения имеет плотность, а именно таковой является функция $p_{XY}(a, b)/p_Y(b)$. Последняя называется условной плотностью вероятности случайной величины X при условии, что $Y=b$, и часто обозначается $p_{X \mid Y}(a; b)$.

Условное математическое ожидание д. с. в. X при условии, что $Y=y$, определяется формулой

$$\mathrm{M}[X \mid Y=y]=\int_{-\infty}^{\infty} x p_{X \mid Y}(x, y)\, dx$$

для всех y, таких, что $p_Y(y)>0$. Аналогичное определение условного математического ожидания можно дать и для дискретного случая. Нетрудно видеть, что $M(X \mid Y)$ является случайной величиной.

Следующее соотношение является очень важным свойством условного математического ожидания:

$$\mathrm{M}[\mathrm{M}[X \mid Y]]=\mathrm{M}[X].$$

В. Бесконечные семейства случайных величин

При рассмотрении бесконечных семейств случайных величин непосредственное обобщение предыдущих определений сопряжено с существенными трудностями. Поэтому мы воспользуемся несколько другим подходом.

Рассмотрим счетное семейство $X_1, X_2, \ldots$ случайных величин. Будем считать их статистические свойства заданными, если для каждого целого числа $n \geqslant 1$ и любого набора $i_1, \ldots, i_n$ из n различных положительных целых чисел задана совместная функция распределения $F_{X_{i_1}, \ldots, X_{i_n}}$ случайных величин $X_{i_1}, \ldots, X_{i_n}$. Разумеется, при этом на бесконечное семейство $\{F_{X_{i_1}, \ldots, X_{i_n}}\}$ должны быть наложены условия согласованности, состоящие в том, что

$$F_{X_{i_1}, \ldots, X_{i_{j-1}}, X_{i_{j+1}}, \ldots, X_{i_n}}(\lambda_1, \ldots, \lambda_{j-1}, \lambda_{j+1}, \ldots, \lambda_n) = \lim_{\lambda_j \to \infty} F_{X_{i_1}, \ldots, X_{i_n}}(\lambda_1, \ldots, \lambda_{j-1}, \lambda_j, \lambda_{j+1}, \ldots, \lambda_n),$$

и в требовании инвариантности функции

$$F_{X_{i_1}, X_{i_2}, \ldots, X_{i_n}}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

относительно перестановки любой пары индексов i_ν и i_μ и соответствующих им λ_ν и λ_μ. Последнее просто означает, что способ нумерации д. с. в. $X_1, X_2, \ldots$ не имеет значения.

Совместные распределения $\{F_{X_{i_1}, \ldots, X_{i_n}}\}$ называются *конечномерными распределениями* семейства случайных величин $\{X_n\}_{n=1}^{\infty}$. В принципе все важные вероятностные свойства этих величин могут быть выражены через конечномерные распределения.

Г. Характеристические функции

С каждой функцией распределения F связана важная функция $\varphi(t)$, где $t \in (-\infty, \infty)$, называемая *характеристической.* По определению имеем [1])

$$\varphi(t) = \int_{-\infty}^{\infty} e^{it\lambda} dF(\lambda), \qquad i^2 = -1. \tag{1.1}$$

Пока читатель может интерпретировать запись интеграла в правой части как символическую. Если F имеет плотность вероятности p, то характеристическая функция представима в виде

$$\varphi(t) = \int_{-\infty}^{\infty} e^{it\lambda} p(\lambda)\, d\lambda.$$

Если F — распределение д. с. в. X со счетным множеством возможных значений $\{\lambda_k\}_{k=0}^{\infty}$ и $\mathrm{P}\{X = \lambda_k\} = a_k$ $(k = 0, 1, \ldots)$, то

[1]) Интеграл в правой части (1.1) является интегралом Лебега — Стильтьеса. У читателей не предполагается знания интегралов этого типа.

правая часть (1.1) сводится к ряду

$$\varphi(t)=\sum_{k=0}^{\infty} e^{it\lambda_k} a_k.$$

Весьма полезным фактом является следующее легко доказываемое утверждение: *если $\{X_1, \ldots, X_n\}$ — семейство независимых д. с. в., то характеристическая функция их суммы представляет собой произведение характеристических функций этих д. с. в.*

Соответствие между функциями распределения и характеристическими функциями — взаимно однозначное. Представление функции распределения через ее характеристическую функцию известно как формула обращения Леви. Поскольку она нам в дальнейшем не понадобится, мы рекомендуем интересующемуся читателю обратиться к литературе, где эти вопросы рассматриваются подробно.

Взаимно однозначное соответствие между функциями распределения и характеристическими функциями сохраняется и при различных предельных переходах. В частности, если $F, F_1, F_2, \ldots$ — функции распределения, такие, что $\lim\limits_{n\to\infty} F_n(\lambda)=F(\lambda)$ для всех λ, где F непрерывна, и $\varphi_n(t)$ — характеристические функции членов последовательности $\{F_n\}$, то

$$\varphi_n(t)=\int_{-\infty}^{\infty} e^{it\lambda}\, dF_n(\lambda) \to \varphi(t)=\int_{-\infty}^{\infty} e^{it\lambda}\, dF(\lambda)$$

равномерно по t в каждом конечном интервале. Обратно, если $\varphi_1, \varphi_2, \ldots$ — характеристические функции членов некоторой последовательности $F_1, F_2, \ldots$ функций распределения и $\lim\limits_{n\to\infty} \varphi_n(t)=\varphi(t)$ для любого t, а $\varphi(t)$ непрерывна в точке $t=0$, то $\varphi(t)$ является характеристической функцией функции распределения F и $\lim\limits_{n\to\infty} F_n(\lambda)=F(\lambda)$ для любого λ, где F непрерывна. Этот результат известен под названием критерия сходимости Леви.

Д. Производящие функции

Для случайных величин, которые могут принимать только неотрицательные целочисленные значения, вместо характеристической функции используется функция

$$g(s)=\sum_{k=0}^{\infty} p_k s^k, \qquad p_k=\mathrm{P}\{X=k\},$$

называемая производящей. Так как по предположению $p_k \geqslant 0$ и $\sum\limits_{k=0}^{\infty} p_k=1$, функция $g(s)$ определена по крайней мере для s, таких, что $|s| \leqslant 1$ (s — комплексная переменная), и бесконечно дифференцируема при $|s|<1$.

Отметим следующие элементарные свойства производящих функций.

(а) Если $Y=X_1+X_2$, где д. с. в. X_1 и X_2 независимы и могут принимать лишь неотрицательные целочисленные значения, то производящая функция д. с. в. Y[1]) является произведением производящих функций слагаемых.

(б) Пусть $X_1, X_2, \ldots$ — независимые и одинаково распределенные неотрицательные целочисленные случайные величины, и пусть N — неотрицательная целочисленная д. с. в., не зависящая от X_i. Мы хотим найти производящую функцию д. с. в. $R = X_1 + \ldots + X_N$ (суммы случайного числа слагаемых, каждое из которых является д. с. в.). Пусть $g_N(s)$ — производящая функция д. с. в. N, а $g(s)$ — производящая функция, общая для всех X_i. Тогда

$$g_R(s) = g_N(g(s)).$$

Имеет место следующее обобщение предыдущего результата. Пусть $X_1, X_2, \ldots$ — произвольные независимые одинаково распределенные случайные величины (необязательно целочисленные), а д. с. в. N, как и ранее, неотрицательная целочисленная и не зависящая от X_i. Тогда

$$\varphi_R(t) = g_N(\varphi(t)),$$

где φ_R и g_N — характеристическая и производящая функции д. с. в. $R = X_1 + \ldots + X_N$ и д. с. в. N соответственно, а φ — общая характеристическая функция д. с. в. X_i.

Рассматривая неотрицательные д. с. в., полезно заменить характеристические функции преобразованиями Лапласа функций распределения. Если распределение F_X имеет плотность p, преобразование Лапласа определяется формулой

$$\psi_X(s) = \int\limits_0^\infty e^{-sx} p(x)\, dx.$$

Интеграл в правой части существует для значений комплексной переменной $s = \sigma + it$, таких, что $\sigma \geqslant 0$. Для s чисто мнимых $\psi_X(s)$ сводится к характеристической функции $\varphi_X(-t)$. Для дискретных неотрицательных д. с. в. преобразование Лапласа определяется по формуле

$$\psi_X(s) = \sum_{n=0}^\infty e^{-s\lambda_n} a_n.$$

Как и для характеристических функций, если $X_1, X_2, \ldots, X_n$ — неотрицательные независимые д. с. в., то

$$\psi_{X_1 + \ldots + X_n}(s) = \prod_{k=1}^n \psi_{X_k}(s).$$

---

[1]) В дальнейшем для краткости мы будем пользоваться термином «производящая функция случайной величины», хотя более правильно было бы говорить «производящая функция распределения д. с. в.». В оригинале часто используется термин «probability generation function», т. е. «вероятностная производящая функция». — *Прим. перев.*

В случае общих функций распределения для преобразования Лапласа имеем формулу

$$\psi_X(s) = \int_0^\infty e^{-s\xi}\, dF_X(\xi).$$

Как и характеристическая функция, преобразование Лапласа однозначно определяет функцию распределения.

Е. Примеры функций распределения

Мы предполагаем известными элементарные свойства функций распределения, которые приведены в таблицах I и II.

Следующие два многомерных распределения имеют принципиальное значение.

(а) *Многомерное нормальное распределение*

Пусть $\|a_{ij}\|$, $i, j = 1, 2$ — симметрическая положительно определенная матрица порядка 2×2 (т. е. $a_{12} = a_{21}$ и $a_{11} > 0$, $a_{11}a_{22} - a_{12}^2 > 0$) и m_1, m_2 — любые действительные постоянные. Тогда

$$f(x_1, x_2) = \frac{\sqrt{B}}{2\pi} \exp\left[-\frac{1}{2}\sum_{i, j=1}^{2} b_{ij}(x_i - m_i)(x_j - m_j)\right],$$

где $\|b_{ij}\|$ — матрица, обратная к $\|a_{ij}\|$, и $B = \det\|b_{ij}\|$, является плотностью распределения, называемого невырожденным двумерным нормальным распределением. Если X_1, X_2 — д. с. в., совместная функция распределения которых равна

$$F(\lambda_1, \lambda_2) = \int_{-\infty}^{\lambda_2}\int_{-\infty}^{\lambda_1} f(x_1, x_2)\, dx_1\, dx_2,$$

то $\mathrm{M}(X_i) = m_i$, $\sigma^2(X_i) = a_{ii}$ при $i = 1, 2$ и $\mathrm{M}(X_1X_2) - m_1m_2 = a_{12}$.

Переход к случаю n-мерного нормального распределения очевиден.

(б) *Полиномиальное распределение*

Совместное дискретное распределение r величин, принимающих только неотрицательные целочисленные значения $0, \ldots, n$, называется полиномиальным. Оно определяется выражением

$$\mathrm{P}(X_1 = k_1, \ldots, X_r = k_r) = \begin{cases} \dfrac{n!}{k_1! \ldots k_r!} p_1^{k_1} \ldots p_r^{k_r}, & \text{если } k_1 + \ldots + k_r = n, \\ 0 & \text{в противном случае,} \end{cases}$$

где $p_i \geqslant 0$, $i = 1, \ldots, r$ и $\sum_{i=1}^{r} p_i = 1$.

Таблица I

Непрерывная функция распределения	*Плотность*	*Допустимые значения параметров*	*Характеристическая функция*	*Среднее*	*Дисперсия*
Нормальное распределение с параметрами m и σ^2	$\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]$ для $-\infty < x < \infty$	m — любое действительное число $\sigma > 0$	$\exp\left[-\frac{\sigma^2 t^2}{2} + imt\right]$	m	σ^2
Гамма-распределение [1]) с параметрами α, λ	$\begin{cases} \frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{\alpha-1} e^{-\lambda x} & \text{для } x > 0, \\ 0 & \text{для } x \leqslant 0 \end{cases}$	$\lambda > 0$ $\alpha > 0$	$\frac{\lambda^\alpha}{(\lambda - it)^\alpha}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
Бета-распределение с параметрами p, q [2])	$\begin{cases} \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)} x^{p-1}(1-x)^{q-1} & \text{для } 0 < x < 1, \\ 0 & \text{в противном случае} \end{cases}$	$p > 0$ $q > 0$	$\frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)} \int\limits_0^1 e^{itx} x^{p-1} \times$ $\times (1-x)^{q-1}\,dx$	$\frac{p}{p+q}$	$\frac{qp}{(p+q)^2(p+q+1)}$

[1]) Гамма-распределение при $\alpha = 1$ называется экспоненциальным; параметр λ называется масштабным.
[2]) Бета-распределение с параметрами $p = q = 1$ обычно называется равномерным распределением в $(0,1)$.

Таблица II

Дискретная функция распределения	*Возможные значения случайных величин*	*Вероятность состояния*	*Производящая функция*	*Среднее*	*Дисперсия*
Пуассоновское распределение с параметром $\lambda > 0$	$n = 0, 1, 2, \ldots$	$\frac{e^{-\lambda}\lambda^n}{n!}$	$e^{-\lambda+\lambda s}$	λ	λ
Биномиальное распределение с параметрами N и p (N — целое положительное, $0 \leqslant p \leqslant 1$), $q = 1 - p$	$n = 0, 1, \ldots, N$	$\binom{N}{n} p^n q^{N-n}$	$(1 - p + ps)^N$	Np	Npq
Отрицательное биномиальное распределение (Паскаля) [1]) с параметрами $\alpha > 0$, $0 < p < 1$	$n = 0, 1, 2, \ldots$	$\binom{\alpha + n - 1}{n} p^\alpha q^n$	$\left(\frac{p}{1 - qs}\right)^\alpha$	$\frac{\alpha q}{p}$	$\frac{\alpha q}{p^2}$

[1]) Геометрическое распределение является важным частным случаем распределения Паскаля, получающимся при $\alpha = 1$.

Ж. Предельные теоремы

Многие из основных результатов теории вероятностей имеют вид предельных теорем. (Мы не будем формулировать эти результаты при наиболее слабых предположениях.) Из них упомянем следующие.

Закон больших чисел (*слабый*). Пусть $X_1, X_2, \ldots$ — независимые одинаково распределенные случайные величины со средним m и конечной дисперсией. Тогда для любого $\varepsilon > 0$

$$\lim_{n \to \infty} \mathrm{P}\left\{\left|\frac{X_1 + \ldots + X_n}{n} - m\right| \leqslant \varepsilon\right\} = 1.$$

Усиленный закон больших чисел. Пусть случайные величины $X_1, X_2, \ldots$ те же, что и выше, и пусть $S_n = (X_1 + \ldots + X_n)/n$; тогда для любых $\varepsilon, \delta > 0$ существует целое число $N(\varepsilon, \delta)$, такое, что

$$\mathrm{P}\{|S_n - m| \geqslant \varepsilon \text{ для всех } n \geqslant N(\varepsilon, \delta)\} < \delta \text{ [1])}.$$

[1]) Оба последних утверждения остаются в силе и при отсутствии конечной дисперсии у случайных величин X_n. — *Прим. ред.*

Центральная предельная теорема. Пусть случайные величины $X_1, X_2, \ldots$ те же, что и выше; тогда

$$\lim_{n\to\infty} \mathrm{P}\left\{ a \leqslant \frac{X_1 + \ldots + X_n - nm}{\sigma\sqrt{n}} \leqslant b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{x^2}{2}\right) dx,$$

где σ^2 — общая дисперсия величин X_i.

Лемма Бореля — Кантелли. Пусть $A_1, A_2, \ldots$ — бесконечная последовательность независимых событий. Тогда событие A_∞, которое означает осуществление бесконечного числа событий из A_1, $A_2, \ldots$, имеет вид

$$A_\infty = \bigcap_{k=1}^{\infty} \bigcup_{i=k}^{\infty} A_i.$$

Лемма Бореля — Кантелли устанавливает, что вероятность события A_∞ равна нулю или единице, если соответственно $\sum_{i=1}^{\infty} \mathrm{P}(A_i) < \infty$ или $\sum_{i=1}^{\infty} \mathrm{P}(A_i) = \infty$.

§ 2. ДВА ПРОСТЫХ ПРИМЕРА СЛУЧАЙНЫХ ПРОЦЕССОВ

Данная книга может служить введением в различные направления случайных процессов. Теория случайных процессов имеет дело с исследованием структуры семейств случайных величин X_t, где t — параметр, принадлежащий соответствующему множеству T. Иногда, когда это не приводит к недоразумениям, мы будем писать $X(t)$ вместо X_t.

Реализацией, или *выборочной функцией*, случайного процесса $\{X_t,\ t \in T\}$ является функция, ставящая в соответствие каждому $t \in T$ одно из возможных значений X_t. Множество параметров T может быть дискретным, т. е. $T = \{0, 1, 2, 3, \ldots\}$, а $\{X_t\}$ может при этом представлять исходы последовательных испытаний, таких, как результаты бросаний монеты, последовательные реакции объекта на изучающий эксперимент или последовательные наблюдения некоторой характеристики популяции и т. д.

Величина X_t может быть одномерной, двумерной, n-мерной или иметь более общую природу. В случае, когда X_n является исходом n-го бросания кости, возможные ее значения образуют множество $\{1, 2, 3, 4, 5, 6\}$, а одной из типичных реализаций процесса является последовательность 5, 1, 3, 2, 2, 4, 1, 6, 3, 6, ... Она показана на рис. 1, где ординатой при $t = n$ является значение X_n. В этом примере случайные величины X_n взаимно независимы, но в общем случае величины X_n являются зависимыми.

Случайные процессы, у которых $T = [0, \infty)$, особенно важны в приложениях. В этом случае t обычно интерпретируется как время.

Здесь мы ограничимся весьма кратким обсуждением некоторых понятий, относящихся к случайным процессам, и двумя примерами.

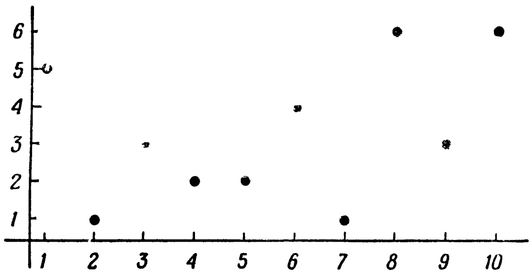


Рис. 1.

В конце главы дана сводка различных типов случайных процессов. Некоторые из них будут подробно изучаться в последующих главах.

Пример 1. Весьма важным примером является известный *процесс броуновского движения.* Этот процесс характеризуется следующими свойствами:

(а) Пусть $t_0 < t_1 < \ldots < t_n$. Тогда приращения $X_{t_1} - X_{t_0}, \ldots$ $\ldots, X_{t_n} - X_{t_{n-1}}$ — независимые в совокупности д. с. в. Говорят, что процесс, обладающий таким свойством, является процессом с независимыми приращениями, выражая тем самым тот факт, что изменения X_t на неперекрывающихся интервалах времени являются независимыми д. с. в.

(б) Вероятностное распределение величин $X_{t_2} - X_{t_1}$, если $t_2 > t_1$, зависит только от $t_2 - t_1$ (и не зависит, например, от t_1).

(в) $$P\{X_t - X_s \leqslant x\} = [2\pi B^2 (t-s)]^{-1/2} \int_{-\infty}^{x} \exp\left[-\frac{u^2}{2B^2(t-s)}\right] du, \quad t > s,$$
B — положительная постоянная.

Пусть $X_0 = 0$. Заметим, что $M(X_t) = 0$, $\sigma^2(X_t) = Bt$, где B — фиксированное положительное число. Можно доказать, что если $0 < t_1 < t_2 < \ldots < t_n < t$, то условное вероятностное распределение величины X_t при известных значениях $X_{t_1}, \ldots, X_{t_n}$ равно (см. гл. 10)

$$P\{X_t \leqslant x \mid X_{t_1} = x_1, \ldots, X_{t_n} = x_n\} =$$
$$= [2\pi B(t - t_n)]^{-1/2} \int_{-\infty}^{x - x_n} \exp\left[-\frac{u^2}{2B(t - t_n)}\right] du.$$

История возникновения этого случайного процесса началась с наблюдений Р. Броуна в 1827 г., заметившего, что маленькие частицы, помещенные в жидкость, совершают непрерывное беспорядочное движение. В 1905 г. А. Эйнштейн объяснил это явление тем, что наблюдаемые частицы подвержены непрерывным соударениям с молекулами окружающей среды. Выведенные Эйнштейном аналитические результаты были позднее проверены экспериментально и обобщены другими физиками и математиками.

Пусть X_t — расстояние броуновской частицы от начальной точки в момент t. Смещение $X_t - X_s$ за интервал времени (s, t) можно рассматривать как сумму большого числа малых смещений. В этой ситуации применима центральная предельная теорема, поэтому естественно ожидать, что $X_t - X_s$ имеет нормальное распределение. Аналогично естественно предположить, что распределения величин $X_t - X_s$ и $X_{t+h} - X_{s+h}$ совпадают при любом $h > 0$, если среда находится в равновесии. Наконец, интуитивно ясно, что смещение $X_t - X_s$ должно зависеть только от $t - s$, а не от момента начала наблюдения.

Процесс броуновского движения (называемый также винеровским процессом) играет фундаментальную роль при изучении множества случайных процессов других типов. Одномерный процесс броуновского движения будет более полно изучен в гл. 10.

Пример 2. Другим важным примером случайного процесса, непрерывного по времени ($T = [0, \infty)$), является пуассоновский процесс. Его выборочная функция X_t представляет собой число регистрации наступления некоторого события за период от 0 до текущего момента времени t. Очевидно, всякая возможная реализация X_t есть неубывающая ступенчатая функция.

Реализация, изображенная на рис. 2, соответствует первому наступлению события в момент t_1, второму — в момент t_2, третьему — в момент t_3 и т. д. Общее число наступлений события возрастает только единичными скачками, а $X_0 = 0$. Конкретными примерами наблюдаемых величин, образующих подобного рода процессы, являются: число фотонов, испускаемых веществом при радиоактивном распаде, число телефонных вызовов из данного района, число происшествий на данном перекрестке, число ошибок на странице машинописного текста, число выходов из строя некоторого механизма и число поступивших заявок на обслуживание.

Рассмотрение этих процессов как пуассоновских основывается на законе редких событий. Представим себе большое число испытаний Бернулли с малой вероятностью успеха (наступления события) и постоянным средним числом успехов. При этих условиях известная теорема утверждает, что число наступлений события подчиняется закону Пуассона. В случае радиоактивного распада пуассоновское приближение дает превосходное согласие с экспериментом, если время наблюдения много меньше периода полураспада радиоактивного вещества.

Мы постулируем, что число наступлений события в некотором интервале не зависит от числа наступлений этого события в любом

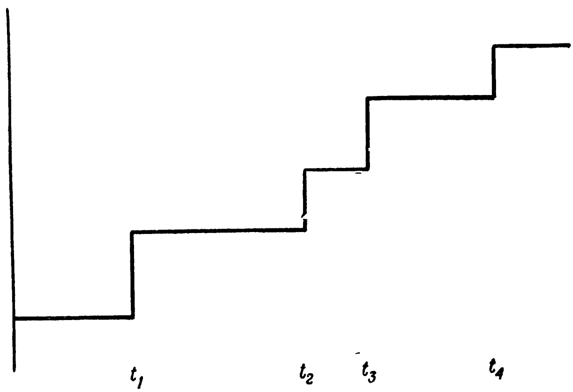


Рис. 2.

другом не пересекающемся с ним интервале (см. свойство (a) предыдущего примера). Предположим также, что с. в. $X_{t_0+t} - X_{t_0}$ зависит только от t и не зависит от t_0 или от значения X_{t_0}. Сформулируем дальнейшие постулаты, согласующиеся с интуитивным описанием, данным выше:

I. Вероятность того, что за период времени продолжительности h произойдет по меньшей мере одно событие, есть

$$p(h) = ah + o(h), \qquad h \to 0, \qquad a > 0$$

[$g(t) = o(t)$ при $t \to 0$ означает, что $\lim_{t \to 0} g(t)/t = 0$].

II. Вероятность того, что за время h произойдет два или более события, есть $o(h)$.

Постулат II равнозначен исключению возможности более чем однократного одновременного наступления события. В приведенных нами физических процессах это условие обычно выполняется.

Пусть $P_m(t)$ обозначает вероятность того, что за время t произойдет ровно m событий, т. е.

$$P_m(t) = \mathrm{P}\{X_t = m\}, \qquad m = 0,\ 1,\ 2,\ \ldots.$$

Условие II можно записать в виде

$$\sum_{m=2}^{\infty} P_m(h) = o(h),$$

и, очевидно,

$$p(h) = P_1(h) + P_2(h) + \ldots.$$

В силу предположения о независимости имеем

$$P_0(t+h) = P_0(t)\,P_0(h) = P_0(t)\,(1 - p(h))$$

и поэтому

$$\frac{P_0(t+h) - P_0(t)}{h} = -\,P_0(t)\,\frac{p(h)}{h}.$$

Из постулата I следует, что $p(h)/h \to a$. Поэтому вероятность $P_0(t)$ того, что событие не наступит в интервале $(0, t)$, удовлетворяет дифференциальному уравнению

$$P_0'(t) = -\,aP_0(t),$$

общее решение которого имеет вид $P_0(t) = ce^{-at}$. Константа c определяется из начального условия $P_0(0) = 1$, из которого следует, что $c = 1$. Итак, $P_0(t) = e^{-at}$.

Найдем теперь $P_m(t)$ для любого m. Легко видеть, что

$$P_m(t+h) = P_m(t)\,P_0(h) + P_{m-1}(t)\,P_1(h) + \sum_{i=2}^{m} P_{m-i}(t)\,P_i(h). \qquad (2.1)$$

По определению $P_0(h) = 1 - p(h)$. Из постулата I следует, что

$$\begin{gathered} P_1(h) = p(h) + o(h), \\ \sum_{i=2}^{m} P_{m-i}(t)\,P_i(h) \leqslant \sum_{i=2}^{m} P_i(h) = o(h), \end{gathered} \qquad (2.2)$$

поскольку, очевидно, $P_k(t) \leqslant 1$. С помощью (2.2) мы можем переписать (2. 1) в виде

$$\begin{aligned} P_m(t+h) - P_m(t) &= P_m(t)\,[P_0(h) - 1] + P_{m-1}(t)\,P_1(h) + \\ &+ \sum_{i=2}^{m} P_{m-i}(t)\,P_i(h) = -\,P_m(t)\,p(h) + P_{m-1}(t)\,P_1(h) + \\ &+ \sum_{i=2}^{m} P_{m-i}(t)\,P_i(h) = -\,aP_m(t)\,h + aP_{m-1}(t)\,h + o(h). \end{aligned}$$

Отсюда следует, что

$$\frac{P_m(t+h) - P_m(t)}{h} \to -\,aP_m(t) + aP_{m-1}(t) \quad \text{при} \quad h \to 0$$

или, формально,

$$P_m'(t) = -\,aP_m(t) + aP_{m-1}(t), \quad m = 1, 2, \ldots, \qquad (2.3)$$

при начальных условиях

$$P_m(0) = 0, \quad m = 1, 2, \ldots.$$

Для решения уравнений (2.3) введем функции

$$Q_m(t) = P_m(t)\, e^{at}, \quad m = 0,\ 1,\ 2,\ \ldots.$$

Подставляя $Q_m(t)$ в (2.3), получаем

$$Q'_m(t) = aQ_{m-1}(t), \quad m = 1,\ 2,\ \ldots, \tag{2.4}$$

где $Q_0(t) \equiv 1$; начальные условия остаются теми же: $Q_m(0) = 0$, $m = 1, 2, \ldots$. Последовательно решая уравнения (2.4), получаем

$$\begin{aligned}
&Q'_1(t) = a, \quad \text{или} \quad Q_1(t) = at + c, && \text{так что } Q_1(t) = at;\\
&Q_2(t) = \frac{a^2t^2}{2} + c, && \text{так что } Q_2(t) = \frac{a^2t^2}{2!};\\
&\quad\vdots && \quad\vdots\\
& && Q_m(t) = \frac{a^mt^m}{m!}.
\end{aligned}$$

Следовательно, имеем

$$P_m(t) = \frac{a^mt^m}{m!}\, e^{-at}.$$

Другими словами, при каждом t с. в. X_t подчиняется распределению Пуассона с параметром at. В частности, среднее число наступлений события за время t равно at.

Часто пуассоновский процесс возникает в форме, где временнóй параметр заменяется соответствующим пространственным параметром. Следующий формальный пример иллюстрирует эту ситуацию. Рассмотрим систему точек, распределенных в пространстве E (E обозначает евклидово пространство размерности $d \geqslant 1$). Пусть N_R обозначает число точек, содержащихся в области R пространства E. Предположим, что N_R является случайной величиной. Совокупность случайных величин $\{N_R\}$, где область изменения индекса R состоит из всех возможных подмножеств пространства E, представляет собой однородный пуассоновский процесс, если удовлетворяются следующие условия:

(i) количества точек в неперекрывающихся областях являются независимыми с. в.;

(ii) для любой области R конечного объема с. в. N_R подчиняется распределению Пуассона со средним $\lambda V(R)$, где $V(R)$ — объем области R. Параметр λ фиксирован и в некотором смысле служит мерой интенсивности распределения, не зависящей от размера и формы.

Пространственные пуассоновские процессы возникают при рассмотрении распределений звезд или галактик, распределений растений и животных, бактерий на предметном стекле и т. д. Мы вернемся к этим вопросам в гл. 12 для более глубокого их изучения.

§ 3. КЛАССИФИКАЦИЯ ОБЩИХ СЛУЧАЙНЫХ ПРОЦЕССОВ

Основные признаки, по которым различаются случайные процессы, касаются природы пространства состояний S, временно́го параметра T и отношений зависимости между случайными величинами X_t.

Пространство состояний S

Это пространство, которому принадлежат все возможные значения, принимаемые всеми с. в. X_t. В случае, если $S=(0, 1, 2, \ldots)$, мы относим процесс к классу целочисленных процессов, или процессов с дискретным пространством состояний. Если S совпадает со всей действительной осью $(-\infty, \infty)$, то мы называем процесс $\{X_t\}$ действительным случайным процессом. Если S — евклидово k-мерное пространство, то говорят, что процесс $\{X_t\}$ является k-мерным.

Как и в случае отдельной с. в., выбор пространства состояний не определяется однозначно описываемым физическим процессом, хотя во многих случаях выбор наиболее подходящего пространства состояний очевиден.

Временно́й параметр T

Если $T=(0, 1, \ldots)$, то мы будем говорить, что $\{X_t\}$ является случайным процессом с дискретным временем. В этом случае мы будем часто писать X_n вместо X_t. Если $T=[0, \infty)$, то случайный процесс $\{X_t\}$ будем называть процессом с непрерывным временем.

Мы уже приводили примеры временно́го параметра T более чем одного измерения (пространственные пуассоновские процессы). Другим примером могут служить волны в океане. Географические долготу и широту можно рассматривать как t, а высоту волны в данном месте — как X_t.

Отношения зависимости

Важной чертой случайного процесса $\{X_t\}$ является зависимость между случайными величинами X_t, $t \in T$.

Характер этой зависимости определяется заданием совместных функций распределения для каждого конечного семейства $X_{t_1}, \ldots$ $\ldots, X_{t_n}$ с. в. процесса. Как мы увидим из примеров (а) и (б), приводимых ниже, совместные функции распределения часто могут быть выражены через другие распределения, связанные с процессом.

Для целей настоящей книги случайный процесс можно считать полностью заданным, если определены его пространство состояний, временно́й параметр и семейство конечномерных распределений. Однако при рассмотрении случайных процессов с непрерывным параметром возникают некоторые трудности, которые мы иллюстрируем на следующем примере.

Пусть U — с. в., равномерно распределенная на $[0, 1]$; определим X_t и Y_t следующим образом:

$$X_t = \begin{cases} 1 & \text{при} \quad U = t, \\ 0 & \text{в противном случае;} \end{cases} \qquad Y_t \equiv 0 \quad (t > 0).$$

Непосредственным вычислением легко убедиться, что $\{X_t\}$ и $\{Y_t\}$ имеют одинаковые конечномерные распределения. Вместе с тем

$$\mathrm{P}\left\{ X_t \leqslant \frac{1}{2} \text{ при всех } 0 \leqslant t \leqslant 1 \right\} = 0,$$

а

$$\mathrm{P}\left\{ Y_t \leqslant \frac{1}{2} \text{ при всех } 0 \leqslant t \leqslant 1 \right\} = 1$$

— положение вещей просто обескураживающее. Чтобы объяснить, чем вызваны эти затруднения, рассмотрим следующий пример.

Предположим, что $\{X_t, 0 \leqslant t \leqslant \infty\}$ является случайным процессом с непрерывным временем. Наша задача — найти $\mathrm{P}\{X_t \geqslant 0, 0 \leqslant t \leqslant 1\}$.

Рассмотрим последовательность сужающихся событий:

$$A_n = \{X_{t_i} \geqslant 0, \ t_i = i/2^n, \ i = 0, 1, \ldots, 2^n\}, \quad n = 1, 2, \ldots.$$

Вероятность каждого события A_n можно выразить через совместную функцию распределения соответствующих с. в. X_{t_i}, $i = 0, \ldots, 2^n$. На первый взгляд кажется естественным положить $\mathrm{P}\{X_t \geqslant 0, 0 \leqslant t \leqslant 1\}$ равным значению предела $\lim\limits_{n \to \infty} \mathrm{P}\{A_n\}$. Однако не все, что кажется естественным, свободно от противоречий. В такой же степени естественно считать, что $\mathrm{P}\{X_t \geqslant 0, 0 \leqslant t \leqslant 1\} = \lim\limits_{n \to \infty} \mathrm{P}\{A'_n\}$, где $A'_n = \{X_{t_i} \geqslant 0, \ t_i = i/3^n, \ i = 0, 1, \ldots, 3^n\}$, но отнюдь не очевидно, что пределы $\lim\limits_{n \to \infty} \mathrm{P}\{A_n\}$ и $\lim\limits_{n \to \infty} \mathrm{P}\{A'_n\}$ равны между собой. Более того, без дополнительных ограничений, касающихся гладкости выборочных функций процесса, эти пределы, вообще говоря, могут иметь различные числовые значения. Известны различные достаточные условия их равенства; одно из этих условий состоит в том, что $\lim\limits_{\tau \to t} \mathrm{P}\{|X_t - X_\tau| > \varepsilon\} = 0$ для всех $\varepsilon > 0$ и всех t. В этом случае можно показать, что никаких противоречий не возникает, если определить $\mathrm{P}\{X_t \geqslant 0, 0 \leqslant t \leqslant 1\}$ как общее значение указанных пределов. Более того, если $t_1, t_2, \ldots$ — любое множество точек, всюду плотное в интервале $(0, 1)$, то $\mathrm{P}\{X_t \geqslant 0, 0 \leqslant t \leqslant 1\} = \lim_{n\to\infty} \mathrm{P}\{X_{t_i} \geqslant 0,\ i = 1, 2, \ldots, n\}$.

Суть дела состоит в следующем: в то время как формула полной вероятности позволяет вычислять вероятности событий, включающих последовательности с. в., через вероятности событий, связанных с любым конечным, а следовательно, и счетным подмножеством этой последовательности, событие $\{X_t \geqslant 0,\ 0 \leqslant t \leqslant 1\}$ зависит от несчетного числа с. в. Мы не можем подробно исследовать этот вопрос в нашей книге и рекомендуем читателю обратиться к монографии Дж. Дуба [4]. Некоторые вопросы, касающиеся оснований теории случайных процессов, мы затронем в гл. 8.

Опишем теперь некоторые классические типы случайных процессов, характеризующиеся различными видами зависимости между X_t. В этих примерах мы будем считать, если не оговаривается противное, что $T = [0, \infty)$. Для простоты изложения мы предполагаем, что с. в. X_t — действительные.

(а) *Процесс с независимыми приращениями*

Если с. в.

$$X_{t_2} - X_{t_1},\quad X_{t_3} - X_{t_2},\ \ldots,\ X_{t_n} - X_{t_{n-1}}$$

независимы для всех $t_1, \ldots, t_n$, таких, что

$$t_1 < t_2 < t_3 < \ldots < t_n,$$

то мы будем говорить, что $\{X_t\}$ является процессом с *независимыми приращениями*. Если множество индексов содержит наименьший индекс t_0, то предполагается также, что

$$X_{t_0},\quad X_{t_1} - X_{t_0},\quad X_{t_2} - X_{t_1},\ \ldots,\ X_{t_n} - X_{t_{n-1}}$$

независимы. Если $T = (0, 1, \ldots)$, то процесс с независимыми приращениями сводится к последовательности независимых с. в. $Z_0 = X_0$, $Z_i = X_i - X_{i-1}$ $(i = 1, 2, 3, \ldots)$ в том смысле, что, зная распределения с. в. $Z_0, Z_1, \ldots$, мы можем определить (и это должно быть ясно читателю) совместное распределение любого конечного множества с. в. X_i. В самом деле,

$$X_i = Z_0 + Z_1 + \ldots + Z_i,\qquad i = 0, 1, 2, \ldots.$$

(б) *Мартингалы*

Пусть $\{X_t\}$ — действительный случайный процесс с дискретным или непрерывным пространством индексов. Мы назовем $\{X_t\}$ *мартингалом*, если для любых $t_1 < t_2 < \ldots < t_{n+1}$ математическое ожидание $\mathrm{M}(X_{t_{n+1}} \mid X_{t_1} = a_1, \ldots, X_{t_n} = a_n)$ равно a_n для всех допустимых значений $a_1, \ldots, a_n$. Мартингалы можно рассматривать как модели безобидных игр. В самом деле, если X_t описывает состояние капитала игрока в момент t, то по определению мартингала средняя величина его капитала в момент t_{n+1} при условии, что в момент t_n он располагал капиталом a_n, равна a_n независимо от того, каков был его капитал в предшествующие моменты времени. Легко убедиться в том, что процесс $X_n = Z_1 + \ldots + Z_n$, $n = 1, 2, \ldots$, является мартингалом (с дискретным временем), если с. в. Z_i — независимые, с нулевыми средними значениями. Аналогично если процесс $\{X_t, 0 \leqslant t < \infty\}$ имеет независимые приращения, средние значения которых равны нулю, то $\{X_t\}$ является мартингалом с непрерывным временем (см. упр. 18).

(в) *Марковские процессы*

Марковский процесс — это процесс, обладающий тем свойством, что если известно значение с. в. X_t, то значения X_s, $s > t$, не зависят от X_u, $u < t$; другими словами, вероятность любого события, связанного с будущим поведением процесса, при условии, что его настоящее состояние точно известно, не изменится, если учесть дополнительную информацию относительно прошлого этого процесса. Подчеркнем однако, что если наше знание настоящего состояния процесса не точно, то вероятность будущего поведения процесса будет, вообще говоря, зависеть от того, что мы знаем о прошлом процесса. Формально, процесс является марковским, если

$$\mathrm{P}\left\{a < X_t \leqslant b \mid X_{t_1} = x_1,\ X_{t_2} = x_2,\ \ldots,\ X_{t_n} = x_n\right\} = \\ = \mathrm{P}\left\{a < X_t \leqslant b \mid X_{t_n} = x_n\right\} \qquad (3.1)$$

при $t_1 < t_2 < \ldots < t_n < t$. Пусть A — интервал на действительной оси. Функция

$$P(x, s; t, A) = \mathrm{P}\{X_t \in A \mid X_s = x\}, \quad t > s, \qquad (3.2)$$

называется *функцией переходных вероятностей*. Эта функция играет важную роль при изучении марковских процессов. Условие (3.1) можно выразить следующим образом:

$$\mathrm{P}\left\{a < X_t \leqslant b \mid X_{t_1} = x_1,\ X_{t_2} = x_2,\ \ldots,\ X_{t_n} = x_n\right\} = P(x_n, t_n; t, A), \qquad (3.3)$$

где $A = \{\xi \mid a < \xi \leqslant b\}$. Можно доказать, что распределение набора

$$(X_{t_1},\ X_{t_2},\ \ldots,\ X_{t_n})$$

можно выразить через функцию переходных вероятностей (3.2) и распределение начальной с. в. X_{t_1}. Мы остановимся подробнее на этих понятиях при рассмотрении дискретных во времени и пространстве марковских процессов (гл. 2).

(г) *Стационарные процессы*

Случайный процесс $\{X_t,\ t \in T\}$ [здесь T может быть одним из следующих множеств: $(-\infty, \infty)$, $[0, \infty)$, множество всех целых чисел, множество положительных чисел] называется *стационарным в узком смысле*, если совместные распределения семейств с. в.

$$\left(X_{t_1+h},\ X_{t_2+h},\ \ldots,\ X_{t_n+h}\right) \quad \text{и} \quad \left(X_{t_1},\ X_{t_2},\ \ldots,\ X_{t_n}\right)$$

одинаковы при всех $h > 0$ и всех $t_1, t_2, \ldots, t_n$ из T. Это условие означает, по существу, что процесс находится в вероятностном равновесии и момент начала нашего наблюдения не имеет значения. В частности, распределение с. в. X_t одно и то же при всех t.

Случайный процесс $\{X_t,\ t \in T\}$ называется *стационарным в широком смысле*, или *ковариационно стационарным*, если он обладает конечными вторыми моментами и $\operatorname{cov}(X_t,\ X_{t+h}) = \mathrm{M}(X_t X_{t+h}) - \mathrm{M}(X_t)\mathrm{M}(X_{t+h})$ зависит только от h при всех $t \in T$.

Стационарные процессы служат для описания многих явлений в теории связи, астрономии, биологии, а иногда и экономики.

Говорят, что марковский процесс имеет стационарные переходные вероятности, если $P(s,\ x;\ t,\ A)$, определенная формулой (3.2), является функцией лишь разности $t - s$ [1]. Вспомним, что $P(s,\ x;\ t,\ A)$ есть условная вероятность — настоящее состояние процесса считается известным. Поэтому нет никаких оснований ожидать, что марковский процесс со стационарными переходными вероятностями является стационарным процессом, что и соответствует действительному положению вещей.

ЗАДАЧИ

1. Пусть a, b и c — независимые с. в., равномерно распределенные в $(0, 1)$. Какова вероятность того, что уравнение $ax^2 + bx + c = 0$ имеет действительные корни?

Ответ: $(5 + 3 \ln 4)/36$.

2. Пусть при каждом фиксированном $\lambda > 0$ с. в. X имеет пуассоновское распределение с параметром λ. Предположим, что сам параметр λ является с. в., подчиняющейся гамма-распределению (т. е. имеющей плотность распределения

$$f(\lambda) = \begin{cases} \dfrac{1}{\Gamma(n)} \lambda^{n-1} e^{-\lambda}, & \lambda \geqslant 0, \\ 0, & \lambda < 0, \end{cases}$$

где n — фиксированная положительная константа). Показать, что

$$\mathrm{P}\{X = k\} = \frac{\Gamma(k+n)}{\Gamma(n)\,\Gamma(k+1)} \cdot \left(\frac{1}{2}\right)^{k+n}, \quad k = 0,\ 1,\ \ldots.$$

Когда n — целое, эта формула есть не что иное, как отрицательное биномиальное распределение с $p = 1/2$.

3. Пусть при каждом заданном p с. в. X имеет биномиальное распределение с параметрами p и N. Предположим, что параметр N сам подчиняется биномиальному распределению с параметрами q и M, причем $M \geqslant N$.

---

[1]) Такой процесс называется также однородным марковским процессом. — *Прим. ред.*

(а) Показать аналитически, что X имеет биномиальное распределение с параметрами pq и M.

(б) Дать вероятностное обоснование этому результату.

4. Пусть с. в. X при каждом заданном p имеет биномиальное распределение с параметрами p и N. Предположим, что p само подчиняется бета-распределению с параметрами r и s. Найти результирующее распределение с. в. X. Когда это распределение является равномерным на множестве $\{0, 1, \ldots, N\}$?

Ответ:

$$P\{X=k\} = \binom{N}{k} \frac{\Gamma(r+s)\,\Gamma(k+r)\,\Gamma(n-k+s)}{\Gamma(r)\,\Gamma(s)\,\Gamma(N+r+s)};$$

$$P\{X=k\} = 1/(N+1), \quad \text{когда} \quad r=s=1.$$

5. Проводится следующий эксперимент. Наблюдается с. в. X, подчиняющаяся пуассоновскому распределению с параметром λ. Затем проводится X испытаний Бернулли с вероятностью успеха p. Каково распределение числа успехов σ?

Ответ: пуассоновское с параметром λp.

6. (а) Предположим, что с. в. X подчиняется пуассоновскому распределению с параметром λ. Параметр λ сам является с. в. с экспоненциальным распределением и средним, равным $1/c$. Найти распределение с. в. X.

(б) То же для случая, когда λ подчиняется гамма-распределению порядка α с масштабным параметром c, т. е. плотность распределения с. в. λ равна

$$c^{\alpha+1} \frac{\lambda^{\alpha}}{\Gamma(\alpha+1)} e^{-\lambda c}$$

при $\lambda > 0$ и равна 0 при $\lambda \leqslant 0$.

Ответ: (а) $P\{X=k\} = \dfrac{c}{(c+1)^{k+1}}$;

(б) $P\{X=k\} = \dfrac{\Gamma(k+\alpha+1)}{k!\,\Gamma(\alpha+1)} \left(\dfrac{1}{1+c}\right)^{k+\alpha+1} c^{\alpha+1}$.

7. Число происшествий за неделю на некотором производстве является с. в. со средним μ и дисперсией $\varkappa^2$. Количества травм, полученных в результате различных происшествий, представляют собой независимые с. в. с одинаковыми средними ν и дисперсиями τ^2. Найти среднее и дисперсию числа травм за неделю.

Указание: Выразить производящую функцию распределения числа травм за неделю через производящие функции распределений числа происшествий за неделю и числа травм за одно происшествие.

Ответ: среднее число травм за неделю равно $\mu\nu$; дисперсия числа травм за неделю равна $\nu^2\varkappa^2 + \mu\tau^2$.

8. Предположим, что имеется N фишек, помеченных числами $1, 2, \ldots, N$ соответственно. Выберем без возвращений и случайным образом $2n+1$ фишек. Пусть Y — медиана полученной случайной выборки. Показать, что Y имеет распределение

$$P\{Y=k\} = \frac{\binom{k-1}{n}\binom{N-k}{n}}{\binom{N}{2n+1}} \qquad \text{при} \quad k = n+1,\ n+2,\ \ldots,\ N-n.$$

Проверить, что

$$M(Y) = \frac{N+1}{2} \qquad \text{и} \qquad \sigma^2(Y) = \frac{(N-2n-1)(N+1)}{8n+12},$$

9. Предположим, что имеется N фишек, занумерованных числами 1, 2, ..., N. Извлекается без возвращений случайная выборка объема n. Пусть X — наибольший номер в случайной выборке. Показать, что X имеет распределение

$$P\{X=k\} = \frac{\binom{k-1}{n-1}}{\binom{N}{n}} \quad \text{при} \quad k=n,\ n+1,\ \dots,\ N$$

и что

$$M(X) = \frac{n}{n+1}(N+1), \qquad \sigma^2(X) = \frac{n(N-n)(N+1)}{(n+1)^2(n+2)}.$$

10. Предположим, что в урне находится n фишек, занумерованных числами 1, 2, ..., n. Производится последовательное извлечение фишек до тех пор, пока одна и та же фишка не появится дважды подряд. Пусть X — число извлечений до наступления этого события. Найти распределение вероятностей с. в. X.

Ответ: $p(k) = (k-1)!\binom{n}{k-1}\frac{k-1}{n^k}$ при $k=2,\ 3,\ \dots,\ n+1$.

11 (продолжение). Показать, что среднее значение с. в. X есть

$$M(X) = 2 + \left(1-\frac{1}{n}\right) + \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \dots + \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{n-1}{n}\right).$$

12. Пусть X_1 и X_2 — независимые с. в., равномерно распределенные на отрезке $[\theta - 1/2,\ \theta + 1/2]$. Показать, что разность $X_1 - X_2$ имеет распределение, не зависящее от θ, и найти ее плотность вероятности.

Ответ:

$$f_{X_1-X_2}(y) = \begin{cases} 1+y, & -1 \leqslant y \leqslant 0, \\ 1-y, & 0 \leqslant y < 1, \\ 0, & |y| > 1. \end{cases}$$

***13.** С. в. X и Y обладают следующими свойствами: X положительна, т. е. $P\{X > 0\} = 1$, и имеет непрерывную плотность вероятности $f(x)$; Y при фиксированном X имеет равномерное распределение в $(0, X)$. Доказать, что если Y и $X - Y$ независимы, то

$$f(x) = a^2xe^{-ax}, \qquad x > 0,\ a > 0.$$

***14.** Пусть с. в. U подчиняется гамма-распределению порядка p, и пусть с. в. V подчиняется бета-распределению с параметрами q и $p-q$ $(0 < q < p)$. Предположим, что U и V независимы. Показать, что UV имеет гамма-распределение с параметром q.

Указание.

$$P\{UV \leqslant x\} = \int_0^1 \left(\int_0^{x/\xi} e^{-\lambda}\lambda^{p-1}\,d\lambda\right) \frac{\xi^{q-1}(1-\xi)^{p-q-1}\,d\xi}{\Gamma(q)\,\Gamma(p-q)}.$$

Перейти в этом соотношении к преобразованию Лапласа, изменить порядок интегрирования и разложить найденное выражение в ряд вида

$$(1+y)^{-\alpha-1} = \sum_{k=0}^{\infty}\binom{\alpha+k}{k}y^k.$$

*15. Пусть X и Y — независимые одинаково распределенные положительные с. в. с непрерывной плотностью распределения $f(x)$. Предположим, далее, что $U = X - Y$ и $V = \min(X, Y)$ являются независимыми с. в. Показать, что

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{при } x \geqslant 0, \\ 0 & \text{в противном случае,} \end{cases}$$

при некотором $\lambda > 0$.

Указание. Показать сначала, что совместная плотность вероятности с. в. U и V имеет вид

$$f_{U, V}(u, v) = f(v) f(v + |u|).$$

Затем показать, что частные распределения с. в. U и V равны соответственно

$$h_U(u) = \int\limits_{\max(0, u)}^{\infty} f(\xi) f(\xi - u)\, d\xi, \qquad h_V(v) = 2[1 - F(v)] f(v),$$

где $F(v) = \int\limits_0^v f(\xi)\, d\xi$. Приравнивая произведение частных распределений совместному распределению, получить соотношение

$$f(v + |u|) = f(|u|)[1 - F(v)]$$

и с его помощью вывести требуемый результат.

16. С. в. X_n принимает значения k/n, $k = 1, 2, \ldots, n$, каждое с вероятностью $1/n$. Найти ее характеристическую функцию и предел последней при $n \to \infty$. Каков вид с. в., которой соответствует предельная характеристическая функция?

Ответ: (а) $\varphi_n(t) = (1 - e^{it}) \dfrac{1}{n} \cdot \dfrac{1}{\exp(-in^{-1}t) - 1}$,

(б) равномерно распределенная на (0, 1).

17. Применив центральную предельную теорему к соответствующим пуассоновским с. в., показать, что

$$\lim_{n \to \infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

18. Показать, что суммы $S_n = X_1 + \ldots + X_n$ независимых с. в. с нулевыми средними образуют мартингал.

19. Показать, что всякий случайный процесс с независимыми приращениями является марковским процессом.

20. Пусть $A_0, A_1, \ldots, A_r$ — исходы некоторого эксперимента. Пусть p_i — вероятность исхода A_i $(i = 0, 1, 2, \ldots, r)$. Предположим, что эксперимент повторяется до тех пор, пока исход A_0 не будет отмечен k раз. Пусть X_i — число

появлений исхода A_i. Показать, что

$$P\left\{X_1=x_1, \ldots, X_r=x_r;\ A_0 \text{ появляется в } k\text{-й раз при } \left(k+\sum_{i=1}^{r} x_i\right)\text{-м испытании}\right\}=$$

$$=\frac{\Gamma\left(k+\sum_{i=1}^{r} x_i\right)}{\Gamma(k)\prod_{i=1}^{r} x_i!}\, p_0^k \prod_{i=1}^{r} p_i^{x_i}. \tag{I}$$

21. Показать, что производящая функция *отрицательного полиномиального распределения* (I) с параметрами $(k;\ p_0, p_1, \ldots, p_r)$ имеет вид

$$\varphi(t_1, \ldots, t_r)=p_0^k\left(1-\sum_{i=1}^{r} t_i p_i\right)^{-k}.$$

22. Рассмотрим векторную случайную величину $\{X_0, X_1, \ldots, X_r\}$, подчиняющуюся полиномиальному распределению с параметрами $(n;\ p_0, p_1, \ldots, p_r)$, и предположим, что n является случайной величиной с отрицательным биномиальным распределением с параметрами $(k;\rho)$. Показать, что распределение вектора $\{X_1, X_2, \ldots, X_r\}$ при условии $X_0=k$ является отрицательным полиномиальным распределением с параметрами $(k;\ p_0(1-\rho),\ p_1(1-\rho),\ \ldots,\ p_r(1-\rho))$ $(0<\rho<1)$.

23. Предположим, что некоторая совокупность состоит из $m, n_1, \ldots, n_r$ предметов, принадлежащих нулевому, первому, ..., r-му классам соответственно. Предметы извлекаются один за другим без возвращения до тех пор, пока не будет извлечено k предметов нулевого класса. Показать, что совместное распределение частот $X_1, \ldots, X_r$ первого, ..., r-го классов имеет вид

$$P\{X_1=x_1, \ldots, X_r=x_r\}=\left\{\binom{m}{k-1}\prod_{i=1}^{r}\binom{n_i}{x_i}\Big/\binom{m+n}{k+y-1}\right\}\frac{m-(k-1)}{m+n-(k+y-1)},$$

где

$$y=\sum_{i=1}^{r} x_i \quad \text{и} \quad n=\sum_{i=1}^{r} n_i.$$

24 (продолжение). Показать, что если $m\to\infty$ и $n\to\infty$ таким образом, что $m/(m+n)\to p_0$ и $n_i/(m+n)\to p_i$, $i=1, 2, \ldots, r$, то это распределение стремится к отрицательному полиномиальному.

ЗАМЕЧАНИЯ

Колоритное и богатое разнообразным материалом введение в теорию вероятностей и ее приложения читатель найдет в книге Феллера [1]. В этой книге автор ограничивается изложением дискретных распределений [1]).

[1]) Распределениям общего вида посвящен второй том «Введения в теорию вероятностей и ее приложения» В. Феллера, переведенный на русский язык. — *Прим. ред.*

Прекрасным введением может служить учебник Гнеденко [2].

Другим полезным элементарным учебником является книга Парзена [3].

Классической монографией, посвященной теории случайных процессов, является книга Дуба [4]. Ее можно считать обязательной для всех, занимающихся случайными процессами.

Другой выдающейся монографией по теории случайных процессов является книга Дынкина [5].

ЛИТЕРАТУРА

1. Феллер В., Введение в теорию вероятностей и ее приложения, т. 1, «Мир», М., 1964.
2. Гнеденко Б. В., Курс теории вероятностей, Физматгиз, М., 1961.
3. Parzen E., Modern Probability Theory and Its Applications, Wiley, New York, 1960.
4. Дуб Дж. Л., Вероятностные процессы, ИЛ, М., 1956.
5. Дынкин Е. Б., Марковские процессы, Физматгиз, М., 1963.

Глава 2

МАРКОВСКИЕ ЦЕПИ

§ 1. ОПРЕДЕЛЕНИЯ

Дискретная марковская цепь $\{X_n\}$ представляет собой марковский случайный процесс, пространство состояний которого счетно или конечно, а множество индексов $T = (0, 1, 2, \ldots)$. Мы можем говорить об X_n как об исходе n-го испытания.

Часто пространство состояний процесса удобно отождествить с множеством неотрицательных целых чисел $(0, 1, 2, \ldots)$ и говорить, что X_n находится в состоянии i, если $X_n = i$.

Вероятность с. в. X_{n+1} попасть в состояние j, если известно, что X_n находится в состоянии i (называемая одношаговой переходной вероятностью), обозначается $P_{ij}^{n,\, n+1}$, т. е.

$$P_{ij}^{n,\, n+1} = \mathrm{P}\{X_{n+1} = j \mid X_n = i\}. \tag{1.1}$$

В таком обозначении подчеркивается, что в общем случае переходные вероятности зависят не только от начального и конечного состояний, но и от момента осуществления перехода. Когда одношаговые переходные вероятности не зависят от временной переменной (т. е. от значения n), мы говорим, что марковский процесс обладает *стационарными переходными вероятностями* (см. § 3 гл. 1). Именно на этом классе марковских цепей мы и сосредоточим свое внимание.

Итак, отметим, что $P_{ij}^{n,\, n+1} = P_{ij}$ не зависит от n и P_{ij} есть вероятность перехода из состояния i в состояние j за одно испытание. Обычно вероятности P_{ij} объединяют в матрицу

$$\mathbf{P} = \left\| \begin{matrix} P_{00} & P_{01} & P_{02} & P_{03} & \ldots \\ P_{10} & P_{11} & P_{12} & P_{13} & \ldots \\ P_{20} & P_{21} & P_{22} & P_{23} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & P_{i3} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix} \right\|,$$

которую называют марковской матрицей, или *матрицей переходных вероятностей* марковской цепи.

В матрице $\mathbf{P}$ $(i+1)$-я строка представляет собой распределение вероятностей с. в. X_{n+1} при условии, что $X_n = i$. Если число состояний конечно, то $\mathbf{P}$ — конечная квадратная матрица, порядок которой (число строк) равен числу состояний. Очевидно, вероятности P_{ij} удовлетворяют следующим двум условиям:

$$P_{ij} \geqslant 0, \qquad i, j = 0, 1, 2, \ldots, \tag{1.2}$$

$$\sum_{j=0}^{\infty} P_{ij} = 1, \qquad i = 0, 1, 2, \ldots. \tag{1.3}$$

Условие (1.3) отражает тот факт, что каждое испытание вызывает некоторый переход из состояний в состояние. (Для удобства мы говорим о переходе и в том случае, когда состояние остается неизменным.)

Процесс полностью определен, коль скоро заданы вероятности (1.1) и состояние (или, более общó, распределение вероятностей) с. в. X_0. Докажем это утверждение.

Пусть $\mathrm{P}\{X_0 = i\} = p_i$. Достаточно показать, как вычисляются вероятности

$$\mathrm{P}\{X_0 = i_0,\ X_1 = i_1,\ X_2 = i_2,\ \ldots,\ X_n = i_n\} \tag{1.4}$$

для любого конечного n, так как по формуле полной вероятности любые другие вероятности, касающиеся с. в. $X_{j_1}, \ldots, X_{j_k}$, $j_1 < j_2 < \ldots < j_k$, могут быть получены суммированием членов вида (1.4).

По определению условной вероятности имеем

$$\begin{aligned}\mathrm{P}\{X_0 = i_0,\ X_1 = i_1,\ X_2 = i_2,\ \ldots,\ X_n = i_n\} &= \\ = \mathrm{P}\{X_n = i_n \mid X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_{n-1} = i_{n-1}\} &\times \\ \times \mathrm{P}\{X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_{n-1} = i_{n-1}\}&.\end{aligned} \tag{1.5}$$

Но по определению марковского процесса имеем

$$\begin{aligned}\mathrm{P}\{X_n = i_n \mid X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_{n-1} = i_{n-1}\} &= \\ = \mathrm{P}\{X_n = i_n \mid X_{n-1} = i_{n-1}\} &= P_{i_{n-1},\, i_n}.\end{aligned} \tag{1.6}$$

Подстановка (1.6) в (1.5) дает

$$\begin{aligned}\mathrm{P}\{X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_n = i_n\} &= \\ = P_{i_{n-1},\, i_n} \mathrm{P}\{X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_{n-1} = i_{n-1}\}&.\end{aligned} \tag{1.7}$$

Продолжая по индукции, получаем

$$\mathrm{P}\{X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_n = i_n\} = P_{i_{n-1},\, i_n} P_{i_{n-2},\, i_{n-1}} \ldots P_{i_0,\, i_1} p_{i_0}. \tag{1.8}$$

§ 2. ПРИМЕРЫ МАРКОВСКИХ ЦЕПЕЙ

Большое число физических, биологических и экономических явлений описываются марковскими цепями. Ниже приводится несколько примеров.

А. Пространственно однородные марковские цепи

Пусть дискретная с. в. ξ принимает неотрицательные целочисленные значения, причем $P\{\xi = i\} = a_i$, $a_i \geqslant 0$ и $\sum_{i=1}^{\infty} a_i = 1$. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ представляют результаты независимых наблюдений с. в. ξ.

Мы опишем две различные марковские цепи, связанные с последовательностью $\{\xi_i\}$. В обоих случаях пространство состояний совпадает с множеством неотрицательных целых чисел.

(I) Определим процесс $\{X_n,\ n = 0, 1, 2, \ldots\}$, положив $X_n = \xi_n$ (начальное значение $X_0 = \xi_0$ задано). Матрица переходных вероятностей этого процесса имеет вид

$$\mathbf{P} = \begin{Vmatrix} a_0 & a_1 & a_2 & a_3 & \ldots \\ a_0 & a_1 & a_2 & a_3 & \ldots \\ a_0 & a_1 & a_2 & a_3 & \ldots \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \end{Vmatrix}.$$

Тот факт, что у матрицы $\mathbf{P}$ все строки одинаковы, означает, что с. в. X_{n+1} не зависит от с. в. X_n.

(II) Другой важный класс марковских цепей возникает при рассмотрении последовательных частичных сумм η_n с. в. ξ_i, т. е.

$$\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n, \qquad n = 1, 2, \ldots .$$

Считаем по определению, что $\eta_0 = 0$. Процесс $\{X_n = \eta_n\}$, как нетрудно видеть, является марковским. Легко найти его матрицу переходных вероятностей; именно:

$$P\{X_{n+1} = j \mid X_n = i\} = P\{\xi_1 + \ldots + \xi_{n+1} = j \mid \xi_1 + \ldots + \xi_n = i\} =$$
$$= P\{\xi_{n+1} = j - i\} = \begin{cases} a_{j-i}, & j \geqslant i, \\ 0, & j < i. \end{cases}$$

Здесь мы, очевидно, воспользовались независимостью ξ_i. В данном случае матрица $\mathbf{P}$ имеет вид

$$\mathbf{P} = \begin{Vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & \ldots \\ 0 & a_0 & a_1 & a_2 & a_3 & \ldots \\ 0 & 0 & a_0 & a_1 & a_2 & \ldots \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \end{Vmatrix}. \tag{2.1}$$

Если с. в. ξ может принимать как положительные, так и отрицательные целочисленные значения, то возможные значения η_n для каждого n содержатся в множестве всех целых чисел. В данном случае пространство состояний удобнее отождествить со всеми целыми числами (а не преобразовывать в множество неотрицательных чисел), так как тогда матрица переходных вероятностей имеет более симметричную форму

$$\mathbf{P} = \left\| \begin{array}{ccccccc} & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \ldots & a_{-1} & a_0 & a_1 & a_2 & a_3 & \ldots \\ \ldots & a_{-2} & a_{-1} & a_0 & a_1 & a_2 & \ldots \\ \ldots & a_{-3} & a_{-2} & a_{-1} & a_0 & a_1 & \ldots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \end{array} \right\|,$$

где $\mathrm{P}\{\xi = k\} = a_k$, $k = 0, \pm 1, \pm 2, \ldots$, и $a_k \geqslant 0$, $\sum\limits_{k=-\infty}^{\infty} a_k = 1$.

Б. Одномерные случайные блуждания

При рассмотрении случайных блужданий состояние системы для наглядности интерпретируют как положение движущейся «частицы».

Одномерное случайное блуждание представляет собой марковскую цепь, пространство состояний которой состоит из конечного или бесконечного множества целых чисел; если частица находится в состоянии i, то за один шаг она может либо перейти в одно из своих соседних состояний ($i - 1$ или $i + 1$), либо остаться в состоянии i. Если пространством состояний служит множество неотрицательных целых чисел, то матрица переходных вероятностей случайного блуждания имеет вид

$$\mathbf{P} = \left\| \begin{array}{cccccccc} r_0 & p_0 & 0 & 0 & \ldots & & & \\ q_1 & r_1 & p_1 & 0 & \ldots & & & \\ 0 & q_2 & r_2 & p_2 & \ldots & & & \\ & \ddots & & & & & & \\ & & 0 & q_i & r_i & p_i & 0 & \\ & & & \ddots & & & \ddots & \end{array} \right\|, \tag{2.2}$$

где $p_i > 0$, $q_i > 0$, $r_i \geqslant 0$ и $q_i + r_i + p_i = 1$, $i = 1, 2, \ldots$ $(i \geqslant 1)$, $p_0 \geqslant 0$, $r_0 \geqslant 0$, $r_0 + p_0 = 1$. Числа q_i, p_i, r_i имеют следующий смысл: если $X_n = i$, то при $i \geqslant 1$

$$\mathrm{P}\{X_{n+1} = i + 1 \mid X_n = i\} = p_i; \quad \mathrm{P}\{X_{n+1} = i - 1 \mid X_n = i\} = q_i;$$
$$\mathrm{P}\{X_{n+1} = i \mid X_n = i\} = r_i;$$

изменения для $i = 0$ очевидны.

В пользу названия «случайное блуждание» для процесса такого типа говорит тот факт, что его реализация описывает путь «абсолютно пьяного» человека, делающего случайным образом шаг вперед или шаг назад.

Капитал игрока, участвующего в серии партий азартной игры, часто описывают процессом случайного блуждания. Предположим, что игрок А, имеющий капитал k, играет с бесконечно богатым партнером, при этом вероятность того, что он выиграет партию и увеличит свой капитал на единицу, равна p_k, а вероятность того, что он проиграет и тем самым уменьшит свой капитал на единицу, равна $q_k = 1 - p_k (k \geqslant 1)$. Зависимость вероятностей выигрыша и проигрыша от k отражает возможную зависимость условий игры от капитала. Так, можно условиться, что, оказавшись в состоянии 0 (соответствующем разорению игрока А), процесс остается в этом состоянии, т. е. $r_0 = 1$. Процесс $\{X_n\}$, где X_n — размер капитала игрока А после n партий, является процессом случайного блуждания. Этот процесс известен под названием «задачи о разорении игрока».

Случайное блуждание с $p_k = p$, $q_k = 1 - p = q (k \geqslant 1)$ и $r_0 = 1$ соответствует одинаковым повторяющимся партиям; если $p > q$, то в каждой партии шансы игрока А явно предпочтительнее. В гл. 3 мы покажем, что в этом случае с вероятностью $(q/p)^{x_0}$, где x_0 — его начальный капитал, игрок А разоряется (теряет свой капитал) и с вероятностью $1 - (q/p)^{x_0}$ его капитал будет беспредельно возрастать. Если же $p < q$, то игра явно выгодна хозяевам игорного заведения, и почти наверное (с вероятностью 1) игрок А разорится, если будет играть достаточно долго. Игрок А обречен на разорение (с вероятностью 1) и в том случае, когда игра безобидна, т. е. когда $p_k = q_k = 1/2$.

Если партнер, игрок Б, тоже начинает игру, располагая ограниченным капиталом y, то капитал игрока А снова описывается марковской цепью $\{X_n\}$. Однако эта цепь имеет конечное множество состояний $0, 1, \ldots, a$, где $a = x + y$, x и y — начальные состояния игроков А и Б соответственно. Разность $a - X_n$ интерпретируется как капитал игрока Б после n партий. Если среди исходов каждой партии допускается ничья, то матрица переходных вероятностей цепи $\{X_n\}$ имеет вид

$$\mathbf{P} = \left\| \begin{matrix} 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ q_1 & r_1 & p_1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & q_2 & r_2 & p_2 & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cdot & & & & & \\ & & & q_{a-1} & r_{a-1} & p_{a-1} & & & \\ 0 & \ldots & \ldots & 0 & 0 & 1 & & & \end{matrix} \right\|. \tag{2.3}$$

Как и ранее, $p_i(q_i)$, $i = 1, 2, \ldots, a-1$, есть вероятность того, что игрок А, имея капитал i, увеличит (уменьшит) его на единицу в следующей партии. Отметим, что в соответствии с матрицей переходных вероятностей (2.3) капитал игрока А (состояние процесса), достигнув величины a или обратившись в 0, остается в этих состояниях навсегда. Мы говорим, что игрок А разорен, если процесс достиг состояния 0; если же процесс попадает в состояние a, то мы говорим, что разорен игрок Б.

Случайные блуждания оказываются полезными не только для описания игровых ситуаций, но и служат неплохими моделями физических процессов, в частности диффузии частиц. Если частица претерпевает случайные столкновения, то ее положение подвержено случайным флуктуациям, хотя описываемая ею траектория непрерывна. Если будущее положение (точнее, его распределение вероятностей) частицы зависит только от ее настоящего положения, то процесс $\{X_t\}$, где X_t — положение частицы в момент t, является марковским. Дискретная аппроксимация такого непрерывного движения соответствует случайному блужданию. Симметричное случайное блуждание представляет собой классический дискретный аналог броуновского движения (см. § 2 гл. 1). Под симметричным случайным блужданием на множестве всех целых чисел подразумевается марковская цепь с пространством состояний, являющимся множеством всех целых чисел, с элементами матрицы переходных вероятностей вида

$$P_{ij} = \begin{cases} p, & \text{если } j = i+1, \\ p, & \text{если } j = i-1, \\ r, & \text{если } j = i, \\ 0 & \text{во всех остальных случаях,} \end{cases}$$

где $p > 0$, $r \geqslant 0$, $2p + r = 1$ и $i, j = 0, \pm 1, \pm 2, \ldots$. Обычно симметричным случайным блужданием называют марковскую цепь с $r = 0$ и $p = {}^1/_2$.

Исследование некоторых физических моделей приводит нас к рассмотрению случайных блужданий на множестве неотрицательных целых чисел. Можно дать классификацию таких процессов на основе свойств нулевого состояния. Пусть случайное блуждание описывается матрицей (2.2). Если $p_0 = 1$ (а значит, и $r_0 = 0$), то нулевое состояние обладает свойствами отражающего экрана. Всякий раз, когда частица достигает нулевого состояния, в результате следующего перехода она оказывается в состоянии 1. Это соответствует ситуации, когда в нуле находится упругая стенка и частица отскакивает от нее без каких-либо остаточных явлений.

Если $p_0 = 0$ и $r_0 = 1$, то нулевое состояние ведет себя как поглощающий экран. Попав в нулевое состояние, частица остается в нем навсегда. Если $p_0 > 0$ и $r_0 > 0$, то нулевое состояние является частично отражающим экраном.

Если случайное блуждание ограничено конечным числом состояний, скажем $0, 1, 2, \ldots, a$, оба крайних состояния 0 и a независимо и в любой комбинации могут быть отражающими, поглощающими или частично отражающими экранами. Мы уже имели дело со случаем, когда состояния 0 и a были поглощающими [см. (2.3)].

Классическую модель диффузии через мембрану представляет собой модель Эренфестов. Модель описывается как процесс случайного блуждания с конечным числом состояний ($i = -a, -a+1, \ldots, -1, 0, 1, \ldots, a$), причем крайние состояния $-a$ и a являются отражающими экранами. Матрица переходных вероятностей задается следующим образом:

$$P_{ij} = \begin{cases} \dfrac{a-i}{2a}, & \text{если } j = i+1, \\ \dfrac{a+i}{2a}, & \text{если } j = i-1, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Физическая интерпретация этой модели такова. Имеется две урны А и Б, содержащие вместе $2a$ шаров. Предположим, что урна А содержит k шаров. При каждом испытании случайным образом выбирается один шар и перекладывается в другую урну; при этом каждый из шаров имеет равную со всеми остальными вероятность быть переложенным независимо от того, в какой урне он находится. Каждое испытание приводит к изменению состояния [1] системы. Характерным для перемещения шаров будет преимущественное направление от урны с большей концентрацией к урне с меньшей концентрацией. Модель Эренфестов в некоторых случаях можно использовать для исследования физических систем, находящихся под действием восстанавливающих сил, величина которых пропорциональна расстоянию от положения равновесия.

Классическое симметричное n-мерное случайное блуждание определяется следующим образом. Пространством состояний процесса является целочисленная решетка в E^n (n-мерном евклидовом пространстве), точки которой суть наборы из n целых чисел вида $\mathbf{k} = (k_1, k_2, \ldots, k_n)$. Переходные вероятности определяются следующим образом:

$$\mathbf{P}_{\mathbf{kl}} = \begin{cases} \dfrac{1}{2n}, & \text{если } \sum\limits_{i=1}^{n} |l_i - k_i| = 1, \\ 0 & \text{в противном случае.} \end{cases}$$

---

[1] Состоянием здесь является величина разности между числом шаров в урне А и урне Б с учетом знака. — *Прим. перев.*

Аналогично одномерному случаю n-мерное симметричное случайное блуждание представляет собой дискретный аналог n-мерного броуновского движения.

В. Дискретная марковская цепь, описывающая очередь

Заявки поступают к месту обслуживания и становятся в очередь. Допустим, что обслуживание одной заявки занимает фиксированное время (период), продолжительность которого принимается за единицу. Если к моменту окончания обслуживания заявки очередь отсутствует, в следующий период обслуживания не происходит. (Можно представить себе, например, стоянку такси, на которую через одинаковые промежутки времени прибывают машины одна за другой. Если в очереди нет пассажиров, машина сразу же уезжает.) Во время обслуживания некоторой заявки могут поступить новые заявки. Предположим, что число заявок ξ_n, поступающих в течение n-го периода, является случайной величиной, функция распределения которой не зависит от номера периода и имеет вид

$$P\{\text{за один период поступило } k \text{ заявок}\} = P\{\xi_n = k\} = a_k, \tag{2.4}$$

$$k = 0,\ 1,\ \ldots,\ a_k > 0, \quad \sum_{k=0}^{\infty} a_k = 1.$$

Мы предположим также, что с. в. ξ_n независимы. Состояние системы в n-й момент времени определяется как число заявок, ждущих обслуживания к началу n-го периода. Если система находится в состоянии i, то по прошествии одного периода она перейдет в состояние

$$j = \begin{cases} i - 1 + \xi, & \text{если} \quad i \geqslant 1, \\ \xi, & \text{если} \quad i = 0, \end{cases} \tag{2.5}$$

где ξ — число поступивших за этот период заявок. В терминах значений случайного процесса мы можем выразить (2.5) так:

$$X_{n+1} = (X_n - 1)^+ + \xi_n, \tag{2.6}$$

где $Y^+ = \max(Y, 0)$. В силу (2.4) и (2.5) матрица переходных вероятностей имеет вид

$$\| P_{ij} \| = \begin{Vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & \ldots \\ a_0 & a_1 & a_2 & a_3 & a_4 & \ldots \\ 0 & a_0 & a_1 & a_2 & a_3 & \ldots \\ 0 & 0 & a_0 & a_1 & a_2 & \ldots \\ 0 & 0 & 0 & a_0 & a_1 & \ldots \\ \vdots & & & & & \end{Vmatrix}. \tag{2.7}$$

Интуитивно ясно, что если среднее число новых заявок $\sum\limits_{k=0}^{\infty} ka_k$, прибывающих за период, превышает единицу, то со временем очередь будет беспредельно увеличиваться.

С другой стороны, если $\sum ka_k < 1$, то, как мы увидим, распределение длины очереди приближается к некоторому равновесному (стационарному) распределению. Если $\sum ka_k = 1$, то возникает существенно неустойчивая ситуация. Все эти утверждения будут строго доказаны после того, как будет изложена теория рекуррентных событий (см. § 5 гл. 3).

Г. Модель из теории запасов

Рассмотрим систему, в которой запасается некоторый товар с целью удовлетворения постоянного спроса. Предположим, что восполнение запаса осуществляется в моменты времени $t_1, t_2, \ldots$, а суммарный спрос ξ_n на товар в интервале (t_{n-1}, t_n) представляет собой случайную величину с распределением

$$P\{\xi_n = k\} = a_k, \qquad k = 0, 1, 2, \ldots, \tag{2.8}$$

одинаковым для всех интервалов, причем как обычно, $a_k \geqslant 0$ и $\sum\limits_{k=0}^{\infty} a_k = 1$. Уровень запаса фиксируется в начале каждого периода. Стратегия запасания такова: если имеющееся количество товара не превышает некоторого критического уровня s, то производится немедленное пополнение запаса до уровня $S > s$. Если же имеющееся количество товара больше s, то пополнения не производится. Пусть X_n обозначает уровень наличного запаса непосредственно перед моментом t_n. Пространство состояний процесса $\{X_n\}$ складывается из возможных значений уровня запаса

$$S, S-1, \ldots, +1, 0, -1, -2, \ldots,$$

где отрицательные значения интерпретируются как неудовлетворенный спрос (эти заказы подлежат немедленному исполнению при пополнении запаса). Согласно описанной стратегии, уровни запаса двух последовательных периодов связаны соотношением

$$X_{n+1} = \begin{cases} X_n - \xi_{n+1}, & \text{если } s < X_n \leqslant S, \\ S - \xi_{n+1}, & \text{если } X_n \leqslant s, \end{cases} \tag{2.9}$$

где ξ_n — суммарный спрос за n-й период с распределением вероятностей (2.8). Если предположить, что с. в. ξ_n независимы, то уровни запаса $X_0, X_1, X_2, \ldots$ образуют марковскую цепь, матрицу переходных вероятностей которой можно вычислить, отправляясь от соотношения (2.9).

Д. Серии успехов

Рассмотрим марковскую цепь с неотрицательными целыми значениями и матрицей переходных вероятностей вида

$$\| P_{ij} \| = \left\| \begin{matrix} p_0 & q_0 & 0 & 0 & \dots \\ p_1 & 0 & q_1 & 0 & \dots \\ p_2 & 0 & 0 & q_2 & \dots \\ p_3 & 0 & 0 & 0 & \dots \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \end{matrix} \right\|, \tag{2.10}$$

где $q_i > 0$, $p_i > 0$ и $q_i + p_i = 1$, $i = 0, 1, 2, \dots$. Нулевое состояние отличается от других тем, что оно может быть достигнуто из всех остальных за один переход, тогда как состояние $i + 1$ достижимо только из состояния i.

Этот пример очень прост с вычислительной точки зрения, и мы будем часто к нему обращаться для иллюстрации вводимых понятий и получаемых результатов.

Частный случай матрицы переходных вероятностей (2.10) описывает серию повторяющихся испытаний, имеющих два возможных исхода — успех (У) и неудача (Н). В каждом испытании вероятность успеха равна α, а вероятность неудачи $\beta = 1 - \alpha$. Мы будем говорить, что на n-м испытании реализовалась серия успехов длины r, если последние $r + 1$ испытаний, включая n-е, имели своими исходами последовательность НУУ...У. Мы будем отмечать состояние процесса длиной серии успехов, реализовавшейся на последнем испытании. В частности, если исходом последнего испытания была неудача, то процесс находится в нулевом состоянии. Если же исходы предшествующих $r + 1$ испытаний дали последовательность НУУ...У, то переменная состояния будет помечена индексом r. Такой процесс, очевидно, является марковским (поскольку каждое испытание было независимо от всех остальных), а его матрица переходных вероятностей имеет вид (2.10), где

$$p_n = \beta, \qquad n = 0, 1, 2, \dots .$$

Е. Ветвящиеся процессы

Предположим, что организм в конце своего времени жизни производит случайное число ξ потомков согласно распределению вероятностей

$$\mathrm{P}\{\xi = k\} = a_k, \qquad k = 0, 1, 2, \dots, \tag{2.11}$$

где, как обычно, $a_k \geqslant 0$ и $\sum\limits_{k=0}^{\infty} a_k = 1$. В свою очередь потомки независимо друг от друга в конце своего времени жизни (для простоты продолжительность жизни предполагается одинаковой для всех организмов) производят потомство, каждый в соответствии с распределением (2.11), продолжая таким образом свой биологический вид. Процесс $\{X_n\}$, где X_n — численность популяции в n-м поколении, представляет собой марковскую цепь.

В самом деле, в конечномерном распределении случайных величин $X_{n_1}, X_{n_2}, \ldots, X_{n_r}, X_n$, $n_1 < n_2 < \ldots < n_r < n$, существенны лишь данные последней переписи популяции, поскольку число потомков является функцией числа непосредственных предков и только. Матрица переходных вероятностей, очевидно, определяется соотношением

$$P_{ij} = \mathrm{P}\{X_{n+1} = j \mid X_n = i\} = \mathrm{P}\{\xi_1 + \ldots + \xi_i = j\}, \tag{2.12}$$

где $\xi_1, \ldots, \xi_i$ — независимые наблюдения с. в., имеющей распределение (2.11). Формулу (2.12) можно объяснить следующим образом. Независимо друг от друга i индивидуумов, составляющих n-е поколение популяции, производят потомство (l-й — в количестве ξ_l). Следовательно, общее число потомков равно $\xi_1 + \xi_2 + \ldots + \xi_i$.

Мы знаем, что производящая функция суммы $\xi_1 + \xi_2 + \ldots + \xi_i$ равна $[g(s)]^i$, где g — производящая функция распределения вероятностей с. в. ξ. (Мы опираемся здесь на свойство композиции производящих функций для случая суммы независимых с. в. См. § 1 гл. 1.) Следовательно, P_{ij} является коэффициентом при j-й степени в степенном разложении функции $[g(s)]^i$.

Ж. Марковские цепи в генетике

Следующая идеализированная генетическая модель была предложена С. Райтом для исследования флуктуаций частот генотипов под влиянием мутаций и отбора. Начнем с так называемой простой гаплоидной модели случайного воспроизведения, в которой не учитываются действие мутационного процесса и силы отбора. Предположим, что мы имеем дело с популяцией фиксированного объема, состоящей из $2N$ индивидуумов типов a и A. Состав следующего поколения определяется результатом $2N$ испытаний Бернулли; а именно: если родительская популяция состоит из j генов типа a и $2N - j$ генов типа A, то вероятности появления гена a или гена A при каждом испытании соответственно равны

$$p_j = \frac{j}{2N}, \qquad q_j = 1 - \frac{j}{2N},$$

что эквивалентно случайному выбору с возвращением. Эта процедура описывается марковской цепью $\{X_n\}$, где X_n — число a-генов в n-м поколении популяции, размер которой остается неизменным от поколения к поколению. Пространством состояний цепи является набор из $2N+1$ чисел $\{0, 1, 2, \ldots, 2N\}$. Элементы матрицы переходных вероятностей вычисляются в соответствии с биномиальным распределением:

$$\mathrm{P}\{X_{n+1}=k \mid X_n=j\}=P_{jk}=\binom{2N}{k} p_j^k q_j^{2N-k}, \quad j,\ k=0,\ 1,\ \ldots,\ 2N. \tag{2.13}$$

Биологическое обоснование этой модели обсуждается в книге Р. Фишера [1]) (см. также гл. 13).

Отметим, что состояния 0 и $2N$ являются поглощающими в том смысле, что если $X_n=0$ или $2N$, то и $X_{n+k}=0$ или $2N$ соответственно для всех $k \geqslant 0$. Один из интересующих нас вопросов — это определение вероятности того, что популяция в каком-либо поколении, а значит и во всех последующих, будет целиком состоять из индивидуумов одного типа (произойдет «фиксация») a или A при условии, что $X_0=i$. Представляет интерес также скорость приближения к такому состоянию. Эти и связанные с ними вопросы будут изучены при общем анализе вероятностей поглощения.

Более реалистичная модель учитывает мутационное давление. Предположим, что прежде, чем будет образовано новое поколение, каждый ген имеет возможность мутировать, т. е. превратиться в ген другого типа. Точнее, мы предположим, что для всех генов типа a вероятность мутации $a \to A$ равна α_1, а для генов типа A вероятность мутации $A \to a$ равна α_2. Как и ранее, будем считать, что состав каждого следующего поколения определяется результатом $2N$ испытаний Бернулли. В том случае, когда родительская популяция содержит j генов типа a, соответствующие величины p_i и q_i имеют вид:

$$\begin{aligned} p_j &= \frac{j}{2N}(1-\alpha_1)+\left(1-\frac{j}{2N}\right)\alpha_2, \\ q_j &= \frac{j}{2N}\alpha_1+\left(1-\frac{j}{2N}\right)(1-\alpha_2). \end{aligned} \tag{2.14}$$

Суть дела состоит в следующем: согласно этой модели, сперва действует мутационный процесс, после чего тип каждого гена следующего поколения определяется случайным выбором (с возвращением) из родительской популяции. Вероятность выбора a-гена из популяции, подвергшейся действию мутационного процесса, равна числу a-генов в этой популяции, умноженному на $1/2N$; поэтому средняя вероятность (усредненная по возможным мутациям) есть произведение среднего числа a-генов после мутаций на $1/2N$. Но это число, очевидно, равно $j(1-\alpha_1)+(2N-j)\alpha_2$, что и приводит к (2.14).

---

[1]) R. A. Fischer, The Genetical Theory of Natural Selection, Oxford (Clarendon) Press, London — New York, 1962.

Переходные вероятности соответствующей марковской цепи вычисляются по формулам (2.13) с использованием величин p_j и q_j из (2.14). Если $\alpha_1, \alpha_2 > 0$, то система не фиксируется ни в одном из состояний. При $n \to \infty$ распределение вероятностей с. в. X_n стремится к распределению некоторой случайной величины ξ:

$$\mathrm{P}\{\xi = k\} = \pi_k, \quad k = 0, 1, \ldots, 2N, \quad \sum_{k=0}^{2N} \pi_k = 1, \quad \pi_k > 0.$$

Это распределение называется стационарным распределением частот генотипа в популяции.

Рассмотрим простую модель случайного скрещивания и введем понятие силы отбора, действующей, скажем, в пользу a-генотипа. Предположим, что мы хотим определить селективное преимущество a-гена над A-геном так, чтобы математические ожидания относительного числа потомков были пропорциональны $1 + s$ и 1 соответственно. Заменим величины $p_j = j/2N$ и $q_j = 1 - j/2N$ на

$$p_j = \frac{(1+s)\, j}{2N + sj} \quad \text{и} \quad q_j = 1 - p_j$$

и определим состав следующего поколения, как и прежде, с помощью испытаний Бернулли. Если родительская популяция содержала j генов типа a, то в следующем поколении средние значения числа генов типа a и числа генов типа A равны соответственно

$$2N \frac{(1+s)\, j}{2N + sj} \quad \text{и} \quad 2N \frac{(2N - j)}{2N + sj}.$$

Отношение среднего числа a-генов к среднему числу A-генов в $(n + 1)$-м поколении равно

$$\frac{1+s}{1} \cdot \frac{j}{2N - j} = \left(\frac{1+s}{1}\right) \left(\frac{\text{число } a\text{-генов в } n\text{-м поколении}}{\text{число } A\text{-генов в } n\text{-м поколении}}\right).$$

Это соотношение поясняет сущность действия отбора.

3. Генетическая модель II

Ген состоит из некоторого числа, скажем, N элементарных единиц. Прежде чем клетка, содержащая ген, начнет делиться, каждая из этих единиц удваивается, и каждая из двух «дочерних» клеток получает ген, состоящий из того же количества элементарных единиц, что и ген «материнской» клетки. Одна или несколько единиц могут быть в мутантной форме. Предполагается, что в процессе удвоения гена каждая мутантная единица превращается в две мутантные единицы, а каждая немутантная единица — в две немутантные. Кроме того, мы предполагаем, что после удвоения эти единицы распределяются между двумя новыми генами случайным образом, как если бы они извлекались из урны. Мы будем следить за судьбой потомков только по одной линии, а не за всей популяцией. Чтобы описать генетическую историю этой линии, рассмотрим марковскую цепь, пространство состояний которой образовано числами 0, 1, ..., N. Мы будем говорить, что ген находится в состоянии i, если он содержит i мутантных и $N-i$ немутантных единиц. Переходные вероятности цепи вычисляются по формуле

$$P_{ij} = \frac{\binom{2i}{j}\binom{2N-2i}{N-j}}{\binom{2N}{N}}. \tag{2.15}$$

Это выражение для P_{ij} получено следующим образом. Предположим, что родительский ген находится в состоянии i; тогда после удвоения мы получим набор из $2i$ мутантных и $2N-2i$ немутантных единиц. Для того чтобы образовать дочерний ген, из этого набора извлекаются N произвольных единиц. В соответствии с гипергеометрическим распределением вероятность того, что дочерний ген окажется в состоянии j, определяется формулой (2.15).

Состояния $j = 1, 2, \ldots, N-1$ называются смешанными, а состояния 0 и N — чистыми. Состояние N интересно тем, что ген, все элементарные единицы которого мутантны, может принести смерть своему обладателю, в то время как ген, находящийся в состоянии 0, при делении не порождает мутантных генов. Позже мы вычислим в явном виде вероятности того, что ген из состояния i попадает в состояние 0 или N.

§ 3. МАТРИЦЫ ПЕРЕХОДНЫХ ВЕРОЯТНОСТЕЙ МАРКОВСКОЙ ЦЕПИ

Марковская цепь полностью определяется своей матрицей одношаговых переходных вероятностей и заданием распределения вероятностей состояния процесса в момент времени 0. Анализ марковской цепи связан главным образом с вычислением вероятностей возможных ее реализаций, важнейшей характеристикой которых является матрица вероятностей переходов за n шагов $\mathbf{P}^{(n)} = \| P_{ij}^{(n)} \|$. $P_{ij}^{(n)}$ обозначает вероятность того, что процесс перейдет из состояния i в состояние j за n переходов, или, в принятых ранее обозначениях,

$$P_{ij}^{n} = \mathrm{P}\{X_{n+m} = j \mid X_m = i\}. \tag{3.1}$$

Заметим, что мы имеем дело только с процессами, однородными во времени, т. е. с процессами, имеющими стационарные переходные вероятности; в противном случае правая часть в (3.1) зависела бы от m.

Марковское свойство процесса позволяет выразить (3.1) непосредственно через $\|P_{ij}\|$, как это видно из следующей теоремы.

Теорема 3.1. *Если* $\mathbf{P} = \| P_{ij} \|$ — *матрица одношаговых переходных вероятностей марковской цепи, то*

$$P^n_{ij} = \sum_{k=0}^{\infty} P^r_{ik} P^s_{kj} \tag{3.2}$$

для любой фиксированной пары неотрицательных целых чисел r *и* s, *такой, что* $r + s = n$; *при этом по определению*

$$P^0_{ij} = \begin{cases} 1, & \text{если} \quad i = j, \\ 0, & \text{если} \quad i \neq j. \end{cases}$$

В формуле (3.2) нетрудно узнать формулу умножения матриц (см. приложение). Отсюда следует, что $\mathbf{P}^{(n)} = \mathbf{P}^n$; другими словами, вероятности P^n_{ij} можно рассматривать как элементы матрицы $\mathbf{P}^n$ — n-й степени матрицы $\mathbf{P}$.

Доказательство. Рассуждения проведем для случая $n = 2$. Событие, состоящее в переходе из состояния i в состояние j за 2 шага, может произойти любым из следующих взаимно исключающих друг друга путей: на первом шаге — переход в некоторое промежуточное состояние k ($k = 0, 1, 2, \ldots$), затем, на втором шаге, переход из состояния k в состояние j.

Так как процесс марковский, то вероятность второго перехода равна P_{kj}, а первого — P_{ik}. Воспользовавшись формулой полной вероятности, приходим к (3.2). Рассуждения в общем случае точно такие же.

Если вероятность того, что процесс в начальный момент находится в состоянии j, равна p_j, то вероятность оказаться в состоянии k в момент n равна

$$p^{(n)}_k = \sum_{j=0}^{\infty} p_j P^n_{jk} = \mathrm{P}\{X_n = k\}. \tag{3.3}$$

Помимо определения совместных распределений вероятностей процесса для всех моментов времени, что, кстати, обычно является очень трудной задачей, часто интерес представляет выяснение асимптотического поведения вероятностей P^n_{ij} при $n \to \infty$. Можно ожидать, что влияние начального состояния со временем уменьшается, а, следовательно, P^n_{ij} стремится при $n \to \infty$ к пределу, не зависящему от i. Для того чтобы дать точный анализ асимптотического поведения марковской цепи, мы введем классификацию ее состояний.

§ 4. КЛАССИФИКАЦИЯ СОСТОЯНИЙ МАРКОВСКОЙ ЦЕПИ

Говорят, что состояние j *достижимо* из состояния i, если $P_{ij}^n > 0$ для некоторого целого числа $n \geqslant 0$, т. е. вероятность того, что процесс за конечное число шагов попадает в состояние j, отправляясь из состояния i, положительна. Состояния i и j, достижимые друг из друга, называют *сообщающимися*; этот факт обозначают $i \leftrightarrow j$. Если два состояния i и j не сообщаются, то либо $P_{ij}^n = 0$ для всех $n \geqslant 0$, либо $P_{ij}^n = 0$ для всех $n \geqslant 0$, либо оба условия выполняются одновременно. Свойство сообщаемости представляет собой отношение эквивалентности. Действительно:

(i) $i \leftrightarrow i$ (рефлексивность), так как по определению имеем

$$P_{ii}^0 = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases};$$

(ii) если $i \leftrightarrow j$, то $j \leftrightarrow i$ (симметричность), согласно определению сообщающихся состояний;

(iii) если $i \leftrightarrow j$ и $j \leftrightarrow k$, то $i \leftrightarrow k$ (транзитивность).

Транзитивность доказывается следующим образом. Из $i \leftrightarrow j$ и $j \leftrightarrow k$ следует, что существует пара неотрицательных целых чисел n и m, таких, что $P_{ij}^n > 0$ и $P_{jk}^m > 0$. Следовательно, в силу (3.2) и неотрицательности всех P_{rs}^t имеем

$$P_{ik}^{n+m} = \sum_{r=0}^{\infty} P_{ir}^n P_{rk}^m \geqslant P_{ij}^n P_{jk}^m > 0.$$

Аналогично показывается существование целого числа ν, такого, что $P_{ki}^{\nu} > 0$.

Из этого следует, что все множество состояний можно разбить на классы эквивалентности. Состояния объединяются в один класс, если они сообщаются друг с другом. Возможно, что, отправляясь из состояния, принадлежащего одному классу, мы с положительной вероятностью попадаем в другой класс, но тогда, очевидно, возврат в исходный класс уже невозможен, так как иначе оба упомянутых класса входили бы в один класс эквивалентности. Мы будем говорить, что марковская цепь *неприводима*, если введенное нами соотношение эквивалентности порождает только один класс состояний. Другими словами, процесс неприводим, если все его состояния сообщаются друг с другом.

Для иллюстрации рассмотрим матрицу переходных вероятностей вида

$$\mathbf{P} = \left\| \begin{array}{cc:ccc} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{array} \right\| = \left\| \begin{array}{cc} \mathbf{P}_1 & 0 \\ 0 & \mathbf{P}_2 \end{array} \right\|.$$

Состояния марковской цепи, имеющей такую матрицу переходных вероятностей, распадаются на два класса сообщающихся состояний: $\{1, 2\}$ и $\{3, 4, 5\}$. В зависимости от начального состояния процесс развертывается либо только в первом классе состояний и его переходы описываются подматрицей $\mathbf{P}_1$, либо только во втором классе и его переходы описываются подматрицей $\mathbf{P}_2$.

В марковской цепи процесса случайного блуждания с матрицей переходных вероятностей

$$\mathbf{P} = \begin{array}{l} \\ \left\| \begin{array}{cccccccc} 1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ q & 0 & p & 0 & \ldots & 0 & 0 & 0 \\ 0 & q & 0 & p & \ldots & 0 & 0 & 0 \\ \cdot & & & & & \cdot & \cdot & \cdot \\ \cdot & & & & & \cdot & \cdot & \cdot \\ \cdot & & & & & \cdot & \cdot & \cdot \\ 0 & \ldots & & & \ldots & q & 0 & p \\ 0 & \ldots & & & \ldots & 0 & 0 & 1 \end{array} \right\| \end{array} \begin{array}{c} \text{состояния} \\ 0 \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ a-1 \\ a \end{array}$$

мы имеем три класса состояний: $\{0\}$, $\{1, 2, \ldots, a-1\}$ и $\{a\}$. Очевидно, что из второго класса можно попасть и в первый и в третий классы, но возвратиться из них во второй класс невозможно.

Марковская цепь, описывающая процесс образования очереди в примере В из § 2, неприводима, если $a_k > 0$ при всех k. Легко проверить, что при этом же условии неприводима и марковская цепь из примера Г. Если $q_i > 0$, $p_i > 0$ $(i = 0, 1, \ldots)$, то марковская цепь серий успехов (пример Д) также неприводима.

Периодичность марковской цепи

Определим период состояния i, далее обозначаемый $d(i)$, как наибольший общий делитель (н. о. д.) всех целых чисел $n \geqslant 1$, для которых $P_{ii}^n > 0$. (Если $P_{ii}^n = 0$ при всех $n \geqslant 1$, то по определению

$d(i) = 0$.) Если в матрице переходных вероятностей (2.2) случайного блуждания все $r_i = 0$, то период каждого из состояний цепи равен двум. Если же хотя бы для одного состояния i_0 величина r_{i_0} больше нуля, то все состояния цепи имеют период 1, так как независимо от начального состояния j система может попасть в состояние i_0 и оставаться в нем в течение произвольного времени прежде, чем вернуться в состояние j.

В конечной марковской цепи с n состояниями и матрицей переходных вероятностей

$$\begin{array}{c} \longleftarrow n \longrightarrow \\ \mathbf{P} = \left\| \begin{matrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ 0 & 0 & & \dots & & 1 \\ 1 & 0 & 0 & \dots & & 0 \end{matrix} \right\| \end{array}$$

каждое состояние имеет период n.

Приведем без доказательства три основных свойства периода состояния (см. задачи 5—7 в конце главы).

Теорема 4.1. *Если $i \leftrightarrow j$, то $d(i) = d(j)$.*

Это утверждение определяет период как характеристику класса сообщающихся состояний.

Теорема 4.2. *Если состояние i имеет период $d(i)$, то существует целое число N, зависящее от i, такое, что для всех целых чисел $n \geqslant N$ вероятность $P_{ii}^{nd(i)} > 0$.*

Этим утверждается, что возвращение в состояние i может происходить во все достаточно далекие моменты времени, кратные периоду $d(i)$.

Следствие 4.1. *Если $P_{ji}^m > 0$, то $P_{ji}^{m+nd(i)} > 0$ для всех достаточно больших положительных целых чисел n.*

Марковская цепь, каждое состояние которой имеет период 1, называется *непериодической*. Большинство марковских цепей, рассматриваемых нами, относятся к классу непериодических. Результаты будут в основном доказываться именно для этого случая; для общего же случая мы обычно будем ограничиваться их формулировками. Трудолюбивый читатель сможет легко провести доказательство самостоятельно

§ 5. ВОЗВРАТНОСТЬ

Рассмотрим произвольное, но фиксированное состояние i. Положим для каждого целого числа $n \geqslant 1$

$$f_{ii}^n = \mathrm{P}\left\{X_n = i, \quad X_\nu \neq i, \quad \nu = 1, 2, \ldots n-1 \mid X_0 = i\right\}.$$

Другими словами, f_{ii}^n есть вероятность того, что, отправляясь из состояния i, система впервые возвратится в это состояние через n переходов. Ясно, что $f_{ii}^1 = P_{ii}$, а f_{ii}^n можно вычислить рекуррентно в соответствии с формулой

$$P_{ii}^n = \sum_{k=0}^{n} f_{ii}^k P_{ii}^{n-k}, \quad n \geqslant 1, \tag{5.1}$$

где, согласно определению, $f_{ii}^0 = 0$ для всех i. Соотношение (5.1) выводится следующим образом. Рассмотрим все возможные реализации процесса, для которых $X_0 = i$, $X_n = i$, а первое возвращение в состояние i происходит на k-м шаге. Обозначим это событие символом E_k. События E_k $(k = 1, 2, \ldots, n)$, очевидно, являются несовместными. Вероятность события, состоящего в том, что первое возвращение происходит на k-м шаге, есть, согласно определению, f_{ii}^k. Рассмотрим теперь те реализации, которые в течение оставшихся $n-k$ шагов ведут себя так, что $X_n = i$. Используя марковское свойство, имеем

$$\mathrm{P}\{E_k\} = \mathrm{P}\{\text{первое возвращение происходит на } k\text{-м шаге} \mid X_0 = i\} \times$$
$$\times \mathrm{P}\left\{X_n = i \mid X_k = i\right\} = f_{ii}^k P_{ii}^{n-k}, \quad 1 \leqslant k \leqslant n$$

(напомним, что $P_{ii}^0 = 1$). Следовательно,

$$\mathrm{P}\left\{X_n = i \mid X_0 = i\right\} = \sum_{k=1}^{n} \mathrm{P}\left\{E_k\right\} = \sum_{k=1}^{n} f_{ii}^k P_{ii}^{n-k} = \sum_{k=0}^{n} f_{ii}^k P_{ii}^{n-k},$$

поскольку по определению $f_{ii}^0 = 0$.

Введем теперь *производящие функции*.

Определение. Производящая функция $P_{ij}(s)$ последовательности $\left\{P_{ij}^n, n = 0, 1, 2, \ldots\right\}$ задается формулой

$$P_{ij}(s) = \sum_{n=0}^{\infty} P_{ij}^n s^n, \quad |s| < 1. \tag{5.2}$$

Аналогично определяется производящая функция последовательности $\left\{f_{ij}^n, n = 0, 1, 2, \ldots\right\}$ (определение вероятностей f_{ij}^n для случая, когда $i \neq j$, следует ниже непосредственно за формулой (5.9)):

$$F_{ij}(s) = \sum_{n=0}^{\infty} f_{ij}^n s^n, \quad |s| < 1. \tag{5.3}$$

Мы уже приводили (см. стр. 16 гл. 1) [1]) следующее свойство производящих функций:
если

$$A(s) = \sum_{k=0}^{\infty} a_k s^k \quad \text{и} \quad B(s) = \sum_{l=0}^{\infty} b_l s^l, \tag{5.4}$$

то

$$A(s) B(s) = C(s) = \sum_{r=0}^{\infty} c_r s^r, \quad |s| < 1, \tag{5.5}$$

где

$$c_r = a_0 b_r + a_1 b_{r-1} + \ldots + a_r b_0. \tag{5.6}$$

Если a_k положить равными f_{ii}^k, а b_l равными P_{ii}^l, то из (5.1) и (5.6) следует, что

$$F_{ii}(s) P_{ii}(s) = P_{ii}(s) - 1, \quad |s| < 1, \tag{5.7}$$

или

$$P_{ii}(s) = \frac{1}{1 - F_{ii}(s)}, \quad |s| < 1. \tag{5.8}$$

В (5.7) единица вычитается потому, что (5.1) не выполняется при $n = 0$.

Точно так же, как мы получили соотношение (5.1), приходим к соотношению

$$P_{ij}^n = \sum_{k=0}^{n} f_{ij}^k P_{jj}^{n-k}, \quad i \neq j, \quad n \geqslant 0, \tag{5.9}$$

где f_{ij}^k есть вероятность того, что процесс впервые достигнет состояния j из состояния i на k-м шаге. Как и ранее, полагаем $f_{ij}^0 = 0$ для всех i и j. Из (5.9) и (5.5) получаем

$$P_{ij}(s) = F_{ij}(s) P_{jj}(s), \quad |s| < 1. \tag{5.10}$$

---

[1]) $A(s) B(s) = \left(\sum_{k=0}^{\infty} a_k s^k \right) \left(\sum_{l=0}^{\infty} b_l s^l \right) =$

$$= a_0 b_0 + (a_1 b_0 + b_1 a_0) s + (a_2 b_0 + a_1 b_1 - a_0 b_2) s^2 + \ldots =$$

$$= \sum_{k=0}^{\infty} s^k \left(\sum_{j=0}^{k} a_j b_{k-j} \right) = \sum_{k=0}^{\infty} c_k s^k.$$

Будем называть состояние i *возвратным*, если $\sum\limits_{n=1}^{\infty} f_{ii}^n = 1$. Это означает, что состояние i является возвратным тогда и только тогда, когда вероятность вернуться в исходное состояние i после некоторого конечного числа шагов равна единице. Если $\sum\limits_{n=1}^{\infty} f_{ii}^n < 1$, то состояние i будем называть *невозвратным*. Ниже мы докажем теорему, связывающую свойство возвратности состояния с поведением n-шаговых переходных вероятностей P_{ij}^n. Для доказательства этой теоремы нам понадобится следующая

Лемма 5.1 (Абель).

(а) *Если ряд* $\sum\limits_{k=0}^{\infty} a_k$ *сходится, то*

$$\lim_{s \uparrow 1} \sum_{k=0}^{\infty} a_k s^k = \sum_{k=0}^{\infty} a_k = a \tag{5.11}$$

($\lim\limits_{s \uparrow 1}$ *означает, что* s *стремится к* 1 *слева, т. е. по значениям, меньшим* 1).

(б) *Если* $a_k \geqslant 0$ *и* $\lim\limits_{s \uparrow 1} \sum\limits_{k=0}^{\infty} a_k s^k = a < \infty$, *то*

$$\sum_{k=0}^{\infty} a_k = \lim_{N \to \infty} \sum_{k=0}^{N} a_k = a.$$

Доказательство. (а) Мы покажем, что

$$\lim_{s \uparrow 1} \left| \sum_{k=0}^{\infty} a_k (s^k - 1) \right| = 0. \tag{5.12}$$

Поскольку ряд $\sum\limits_{k=0}^{\infty} a_k$ сходится, то для любого $\varepsilon > 0$ можно найти $N(\varepsilon)$, такое, что $\left| \sum\limits_{k=N}^{N'} a_k \right| < \varepsilon/4$ для всех $N' > N$. Выберем такое N. Далее,

$$\left| \sum_{k=0}^{\infty} a_k (s^k - 1) \right| = \left| \sum_{k=0}^{N} a_k (s^k - 1) + \sum_{k=N+1}^{\infty} a_k (s^k - 1) \right| \leqslant$$

$$\leqslant \left| \sum_{k=0}^{N} a_k (s^k - 1) \right| + \left| \sum_{k=N+1}^{\infty} a_k (s^k - 1) \right|. \tag{5.13}$$

Для $0 \leqslant s < 1$

$$\left|\sum_{k=0}^{N} a_k (s^k - 1)\right| \leqslant MN \mid s^N - 1 \mid, \tag{5.14}$$

где $M = \max\limits_{0 \leqslant k \leqslant N} \mid a_k \mid < \infty$, так что для s, достаточно близких к 1, имеем

$$\left|\sum_{k=0}^{N} a_k (s^k - 1)\right| < \varepsilon/2.$$

Для оценки $\sum\limits_{k=N+1}^{\infty} a_k (s^k - 1)$ просуммируем по частям:

$$\left|\sum_{k=N+1}^{\infty} a_k (s^k - 1)\right| = \left|\sum_{k=N+1}^{\infty} (A_k - A_{k+1})(s^k - 1)\right| =$$

$$= \left| A_{N+1} (s^{N+1} - 1) - \sum_{k=N+2}^{\infty} A_k (s^k - s^{k+1})\right|, \tag{5.15}$$

где $A_k = \sum\limits_{r=k}^{\infty} a_r$. Легко видеть, что для (5.15) имеет место оценка

$$\frac{\varepsilon}{4} \mid (s^{N+1} - 1) \mid + \frac{\varepsilon}{4} s^{N+1} \leqslant \frac{\varepsilon}{2}.$$

Вместе с предыдущей оценкой это дает

$$\left|\sum_{k=0}^{\infty} a_k (s^k - 1)\right| < \varepsilon$$

при условии, что s достаточно близко к 1. ■

(б) Поскольку $\sum\limits_{k=0}^{\infty} a_k s^k \leqslant \sum\limits_{k=0}^{\infty} a_k$ для $0 < s < 1$, случай $a = \infty$ очевиден. Если $a < \infty$, то в силу исходного предположения имеем

$$\sum_{k=0}^{\infty} a_k s^k < a < \infty \quad \text{для} \quad 0 < s < 1,$$

а значит, и

$$\sum_{k=0}^{\infty} a_k \leqslant a \quad \text{для всех } n.$$

Кроме того, $\sum\limits_{k=0}^{n} a_k$ является монотонно возрастающей функцией n.

Таким образом, ряд $\sum\limits_{k=0}^{\infty} a_k$ сходится. Пусть сумма ряда равна a'. Обращаясь к первой части леммы, получаем, что $a' = a$. ■

С помощью этой леммы легко доказывается

Теорема 5.1. *Состояние i является возвратным тогда и только тогда, когда*

$$\sum_{n=1}^{\infty} P_{ii}^{n} = \infty.$$

Доказательство. Пусть i — возвратное состояние, т. е. $\sum_{n=0}^{\infty} f_{ii}^{n} = 1$. В силу леммы 5.1 (а) имеем

$$\lim_{s \uparrow 1} \sum_{n=0}^{\infty} f_{ii}^{n} s^{n} = \lim_{s \uparrow 1} F_{ii}(s) = 1;$$

тогда из (5.8) следует, что

$$\lim_{s \uparrow 1} P_{ii}(s) = \lim_{s \uparrow 1} \sum_{n=0}^{\infty} P_{ii}^{n} s^{n} = \infty.$$

Обращаясь к лемме 5.1 (б), заключаем, что

$$\sum_{n=0}^{\infty} P_{ii}^{n} = \infty.$$

Чтобы доказать достаточность, предположим, что состояние i невозвратное, т. е. $\sum_{n=1}^{\infty} f_{ii}^{n} < 1$. Используя лемму 5.1 (а) и соотношение (5.8), получаем

$$\lim_{s \uparrow 1} P_{ii}(s) < \infty.$$

Отсюда в силу леммы 5.1 (б) следует, что $\sum_{n=1}^{\infty} P_{ii}^{n} < \infty$. ■

Из теоремы 5.1 непосредственно вытекает

Следствие 5.1. *Если $i \leftrightarrow j$ и i — возвратное состояние, то состояние j также является возвратным.*

Доказательство. Так как $i \leftrightarrow j$, то существуют такие целые числа $m, n \geqslant 1$, что

$$P_{ij}^{n} > 0 \quad \text{и} \quad P_{ji}^{m} > 0.$$

Пусть ν — положительное целое число. С помощью аргументов, к которым мы уже прибегали (см. стр. 52), нетрудно получить неравенство

$$P_{jj}^{m+n+\nu} \geqslant P_{ji}^{m} P_{ii}^{\nu} P_{ij}^{n}.$$

Суммирование по ν дает

$$\sum_{\nu=0}^{\infty} P_{jj}^{m+n+\nu} \geqslant \sum_{\nu=0}^{\infty} P_{ji}^{m} P_{ii}^{\nu} P_{ij}^{n} = P_{ji}^{m} P_{ij}^{n} \sum_{\nu=0}^{\infty} P_{ii}^{\nu}.$$

Таким образом, если $\sum\limits_{\nu=0}^{\infty} P_{ii}^{\nu}$ расходится, то расходится и $\sum\limits_{\nu=0}^{\infty} P_{jj}^{\nu}$. ■

Доказав это следствие, мы установили, что возвратность, как и периодичность, является свойством класса эквивалентности, т. е. все состояния в классе эквивалентности либо возвратны, либо невозвратны одновременно.

§ 6. ПРИМЕРЫ ВОЗВРАТНЫХ МАРКОВСКИХ ЦЕПЕЙ

Пример 1. Рассмотрим одномерное случайное блуждание по одномерной целочисленной решетке. За каждый переход частица с вероятностью p перемещается на единицу вправо и с вероятностью q — на единицу влево $(p+q=1)$; следовательно, имеем

$$P_{00}^{2n+1}=0, \quad P_{00}^{2n}=\binom{2n}{n} p^n q^n = \frac{(2n)!}{n!\,n!} p^n q^n, \quad n=0,\ 1,\ 2,\ \ldots. \tag{6.1}$$

Воспользуемся формулой Стирлинга

$$n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \tag{6.2}$$

и запишем (6.1) в виде

$$P_{00}^{2n} \sim \frac{(pq)^n\, 2^{2n}}{\sqrt{\pi n}} = \frac{(4pq)^n}{\sqrt{\pi n}}.$$

Легко проверить, что $p(1-p)=pq \leqslant 1/4$, причем равенство имеет место тогда и только тогда, когда $p=q=1/2$. Следовательно, $\sum\limits_{n=0}^{\infty} P_{00}^{n}=\infty$ тогда и только тогда, когда $p=1/2$. Таким образом, одномерное случайное блуждание возвратно тогда и только тогда, когда $p=q=1/2$. (Вспомним, что возвратность является свойством класса!) Интуитивно ясно, что если $p \neq q$, то положительна вероятность того, что частица, отправляясь из начала координат, будет смещаться к $+\infty$, если $p>q$ (к $-\infty$, если $p<q$), ни разу не возвращаясь в исходное состояние.

Пример 2. Обратимся теперь к двумерному случайному блужданию по двумерной целочисленной решетке. Пусть вероятности смещения на единицу влево, вправо, вверх, вниз — все равны $^1/_4$. Как и ранее, будем исследовать возвратность состояния, представляемого началом координат. Рассмотрим все траектории, состоящие из i перемещений вправо, i перемещений влево, j перемещений вниз и j перемещений вверх, $2i+2j=2n$. Легко убедиться, воспользовавшись полиномиальным распределением, что

$$\begin{aligned} P_{00}^{2n} &= \sum_{i,\,j,\,i+j=n} \frac{(2n)!}{i!\,i!\,j!\,j!}\left(\frac{1}{4}\right)^{2n}, \quad && n=1,\ 2,\ 3,\ \ldots, \\ P_{00}^{2n+1} &= 0, && n=0,\ 1,\ 2,\ \ldots. \end{aligned} \tag{6.3}$$

Умножая числитель и знаменатель в правой части (6.3) на $(n!)^2$, получаем

$$P_{00}^{2n} = \left(\frac{1}{4}\right)^{2n}\binom{2n}{n}\sum_{i=0}^{n}\binom{n}{i}\binom{n}{n-i},$$

но

$$\sum_{i=0}^{n}\binom{n}{i}\binom{n}{n-i} = \binom{2n}{n}.$$

Следовательно,

$$P_{00}^{2n} = \left(\frac{1}{4}\right)^{2n}\binom{2n}{n}^2.$$

Формула Стирлинга (6.2) дает

$$P_{00}^{2n} \sim \frac{1}{\pi n}.$$

Таким образом, $\sum\limits_{n=0}^{\infty} P_{00}^{n} = \infty$ и состояние, представляемое началом координат, является возвратным.

Пример 3. Рассмотрим теперь симметричное случайное блуждание по трем измерениям. Аналогично предыдущему легко убедиться, что

$$\begin{gathered} P_{00}^{2n+1} = 0, \quad n=0,\ 1,\ 2,\ \ldots, \\ P_{00}^{2n} = \sum_{i,\,j,\,0\leqslant i+j\leqslant n} \frac{(2n)!}{i!\,i!\,j!\,j!\,(n-i-j)!\,(n-i-j)!}\left(\frac{1}{6}\right)^{2n}, \quad n=0,\ 1,\ \ldots. \end{gathered} \tag{6.4}$$

Умножая числитель и знаменатель на $(n!)^2$ и выделяя сомножитель $(^1/_2)^{2n}$, получаем

$$P_{00}^{2n} = \frac{1}{2^{2n}}\binom{2n}{n}\sum_{i,\,j,\,0\leqslant i+j\leqslant n}\left[\frac{n!}{i!\,j!\,(n-i-j)!}\right]^2\left(\frac{1}{3}\right)^{2n}, \tag{6.5}$$

$$P_{00}^{2n} \leqslant c_n\frac{1}{2^{2n}}\binom{2n}{n}\frac{1}{3^n}, \tag{6.6}$$

где

$$c_n = \max_{i,\ j,\ 0 \leqslant i+j \leqslant n} \left[\frac{n!}{i!\, j!\, (n-i-j)!} \right]. \tag{6.7}$$

Здесь мы воспользовались тем фактом, что

$$\sum_{i,\ j,\ 0 \leqslant i+j \leqslant n} \frac{n!}{i!\, j!\, (n-i-j)!} \left(\frac{1}{3} \right)^n = 1. \tag{6.8}$$

Для больших n значение c_n достигается при $i = j \sim n/3$. Действительно, пусть i_0 и j_0 — те значения i и j, на которых достигается максимум выражения

$$\frac{n!}{i!\, j!\, (n-i-j)!}$$

при $0 \leqslant i + j \leqslant n$. Сразу же можно выписать следующие четыре неравенства:

$$\frac{n!}{j_0!\, (i_0-1)!\, (n-j_0-i_0+1)!} \leqslant \frac{n!}{j_0!\, i_0!\, (n-j_0-i_0)!};$$

$$\frac{n!}{j_0!\, (i_0+1)!\, (n-j_0-i_0-1)!} \leqslant \frac{n!}{j_0!\, i_0!\, (n-j_0-i_0)!};$$

$$\frac{n!}{(j_0-1)!\, i_0!\, (n-j_0-i_0+1)!} \leqslant \frac{n!}{j_0!\, i_0!\, (n-j_0-i_0)!};$$

$$\frac{n!}{(j_0+1)!\, i_0!\, (n-j_0-i_0-1)!} \leqslant \frac{n!}{j_0!\, i_0!\, (n-j_0-i_0)!}.$$

Эти неравенства сводятся к следующим двум:

$$n - i_0 - 1 \leqslant 2j_0 \leqslant n - i_0 + 1,$$
$$n - j_0 - 1 \leqslant 2i_0 \leqslant n - j_0 + 1.$$

Следовательно, $i_0 \sim n/3$ и $j_0 \sim n/3$ при больших n. Подставляя в (6.7), преобразуем (6.6) к виду

$$P_{00}^{2n} \leqslant \frac{n!}{(n/3)!\, (n/3)!\, (n/3)!\, 2^{2n} 3^n} \binom{2n}{n}. \tag{6.9}$$

Воспользовавшись формулой Стирлинга, для правой части (6.9) получаем следующее асимптотическое выражение:

$$\frac{3\sqrt{3}}{2\pi^{3/2} n^{3/2}}.$$

Суммируя такие члены, получаем

$$\sum_{n=1}^{\infty} \frac{3\sqrt{3}}{2\pi^{3/2} n^{3/2}} < \infty.$$

Следовательно, $\sum_{n=1}^{\infty} P_{00}^n < \infty$, и в силу теоремы 5.1 состояние, представляемое началом координат, является невозвратным. Поскольку возвратность — свойство класса, а все состояния сообщаются, частица, совершающая одно- или двумерное симметричное случайное блуждание, с достоверностью вернется во всякое состояние, в котором она когда-либо пребывала. В трехмерном симметричном случайном блуждании частица, покинув состояние, с положительной вероятностью не вернется в него никогда.

Пример 4. Рассмотрим марковскую цепь, описывающую серии успехов при биномиальных испытаниях. Матрица переходных вероятностей этой цепи имеет вид

$$\left\| \begin{array}{cccccc} p_0 & 1-p_0 & 0 & 0 & \ldots & \ldots \\ p_1 & 0 & 1-p_1 & 0 & \ldots & \ldots \\ p_2 & 0 & 0 & 1-p_2 & & \\ \vdots & \vdots & & & \ddots & \\ p_r & 0 & \ldots & \ldots & 1-p_r & 0 \ldots \\ \vdots & & & & \vdots & \vdots \end{array} \right\| \qquad (0 < p_i < 1).$$

Состояния цепи образуют единственный класс эквивалентности (всякое состояние цепи можно достичь из любого другого состояния). Поэтому в силу следствия 5.1 нам достаточно исследовать возвратность одного состояния, например нулевого. Имеем

$$\begin{aligned} f_{00}^1 &= p_0 = 1 - (1 - p_0), \\ f_{00}^n &= \left(\prod_{i=0}^{n-2} (1 - p_i) \right) p_{n-1}, \quad n > 1. \end{aligned} \tag{6.10}$$

Уравнения (6.10) можно переписать в виде

$$f_{00}^n = \prod_{i=0}^{n-2} (1 - p_i) [1 - (1 - p_{n-1})], \quad n > 1.$$

$$f_{00}^n = (1 - p_0)(1 - p_1) \ldots (1 - p_{n-2}) - (1 - p_0)(1 - p_1) \ldots (1 - p_{n-1}), \quad n > 1.$$

Положим

$$u_n = \begin{cases} (1 - p_0)(1 - p_1) \ldots (1 - p_n), & n \geqslant 0, \\ 1, & n = -1. \end{cases}$$

Теперь если мы просуммируем вероятности f_{00}^n, то получим

$$\sum_{n=1}^{m+1} f_{00}^n = \sum_{n=1}^{m+1} (u_{n-2} - u_{n-1}) = (1 - u_0) + (u_0 - u_1) + \ldots + (u_{m-1} - u_m),$$

или

$$\sum_{n=1}^{m+1} f_{00}^{n} = 1 - u_m. \tag{6.11}$$

Для завершения наших рассуждений полезной окажется следующая

Лемма 6.1. *Если* $0 < p_i < 1$ $(i = 0, 1, 2, \ldots)$, *то* $u_m = \prod_{i=0}^{m} (1 - p_i) \to 0$ *при* $m \to \infty$ *тогда и только тогда, когда* $\sum_{i=0}^{\infty} p_i = \infty$.

Доказательство. Предположим, что $\sum_{i=0}^{\infty} p_i = \infty$. Так как степенное разложение функции $\exp(-p_i)$ представляет собой знакопеременный ряд с уменьшающимися по абсолютной величине членами, то справедливо неравенство

$$1 - p_i < 1 - p_i + \frac{p_i^2}{2!} - \frac{p_i^3}{3!} + \ldots = \exp(-p_i), \quad i = 0, 1, 2, \ldots \tag{6.12}$$

Так как (6.12) выполняется при всех i, то $\prod_{i=0}^{m} (1 - p_i) < \exp\left(-\sum_{i=0}^{m} p_i\right)$.

Но по предположению $\lim_{m \to \infty} \sum_{i=0}^{m} p_i = \infty$; следовательно,

$$\lim_{m \to \infty} \prod_{i=0}^{m} (1 - p_i) = 0.$$

Для доказательства необходимости воспользуемся следующим легко выводимым неравенством:

$$\prod_{i=j}^{m} (1 - p_i) > (1 - p_j - p_{j+1} - \ldots - p_m),$$

справедливым при всех j и $m = j+1, j+2, \ldots$. Предположим теперь, что $\sum_{i=1}^{\infty} p_i < \infty$; тогда $0 < \sum_{i=j}^{\infty} p_i < 1$ при некотором j. Отсюда следует, что

$$\lim_{m \to \infty} \prod_{i=j}^{m} (1 - p_i) > \lim_{m \to \infty} \left(1 - \sum_{i=j}^{m} p_i\right) > 0.$$

Это противоречит предположению о том, что $u_m \to 0$. ■

Возвращаясь к (6.11) и применяя лемму 6.1, заключаем, что $\sum_{n=1}^{\infty} f_{00}^{n} = 1$ тогда и только тогда, когда $\sum_{i=0}^{\infty} p_i = \infty$, или, что то же самое, состояние 0 возвратно тогда и только тогда, когда ряд $\sum_{i=0}^{\infty} p_i$ расходится.

Заметим попутно, что для любого набора положительных чисел $\{a_1, a_2, \ldots\}$, такого, что $a_i > 0$, $\sum_{i=1}^{\infty} a_i \leqslant 1$, можно построить марковскую цепь только что рассмотренного вида, причем величины p_i будут такими, что $f_{00}^n = a_n$. Действительно, положим

$$f_{00}^1 = p_0 = a_1, \quad f_{00}^2 = (1 - p_0)\, p_1 = a_2,$$

откуда получаем

$$p_1 = \frac{a_2}{1 - a_1}.$$

Пусть

$$f_{00}^3 = (1 - p_0)(1 - p_1)\, p_2 = a_3,$$

тогда

$$p_2 = \frac{a_3}{1 - a_1 - a_2}.$$

Продолжая аналогичным образом, мы получим явные выражения для величин p_i, причем $0 < p_i < 1$.

§ 7. ЕЩЕ О ВОЗВРАТНОСТИ

Теорема, которую мы сейчас докажем, утверждает, что если некоторое состояние возвратно, то это состояние с вероятностью 1 встречается в процессе бесконечное число раз. Пусть

$$Q_{ii} = \mathrm{P}\left\{ \begin{matrix} \text{частица, отправляясь из состояния } i, \\ \text{возвращается в него бесконечно часто} \end{matrix} \right\}.$$

Теорема 7.1. *Состояние i возвратно или невозвратно в зависимости от того, $Q_{ii} = 1$ или $Q_{ii} = 0$.*

Доказательство. Положим

$$Q_{ii}^N = \mathrm{P}\left\{ \begin{matrix} \text{частица, отправляясь из состояния } i, \\ \text{возвращается в него по крайней мере } N \text{ раз} \end{matrix} \right\}.$$

Имеет место соотношение

$$Q_{ii}^N = \sum_{k=1}^{\infty} f_{ii}^k Q_{ii}^{N-1} = Q_{ii}^{N-1} f_{ii}^*, \quad \text{где} \quad f_{ii}^* = \sum_{k=1}^{\infty} f_{ii}^k,$$

в справедливости которого нетрудно убедиться, представляя событие, фигурирующее в правой части предыдущего соотношения, в виде суммы несовместных событий, определяемых временем первого возвращения. Последовательно применяя последнюю формулу, получаем

$$Q_{ii}^N - f_{ii}^* Q_{ii}^{N-1} = (f_{ii}^*)^2 Q_{ii}^{N-2} = \ldots = [f_{ii}^*]^{N-1} Q_{ii}^1.$$

Но, очевидно, $Q_{ii}^1 = f_{ii}^*$; следовательно,

$$Q_{ii}^N = [f_{ii}^*]^N.$$

Поскольку $\lim\limits_{N\to\infty} Q_{ii}^N = Q_{ii}$, то $Q_{ii} = 1$ или 0 при $f_{ii}^* = 1$ или < 1 соответственно, или, что эквивалентно, в зависимости от того, является ли состояние i возвратным или невозвратным.

Теорема 7.2. *Если $i \leftrightarrow j$ и оба состояния принадлежат возвратному классу, то*

$$f_{ij}^* = \sum_{n=1}^{\infty} f_{ij}^n = 1.$$

Мы опускаем простое доказательство этого факта.

Пусть

$$Q_{ij} = \mathrm{P}\left\{\begin{array}{c}\text{частица, отправляясь из состояния } i, \text{ будет} \\ \text{бесконечное число раз находиться в состоянии } j\end{array}\right\}$$

Из теоремы 7.2 непосредственно вытекает

Следствие 7.1. *Если $i \leftrightarrow j$ и оба состояния принадлежат возвратному классу, то $Q_{ij} = 1$.*

Доказательство. Нетрудно видеть, что

$$Q_{ij} = f_{ij}^* Q_{jj}.$$

Поскольку состояние j возвратно, то $Q_{jj} = 1$ по теореме 7.1. По теореме 7.2 имеем $f_{ij}^* = 1$, следовательно, $Q_{ij} = 1$.

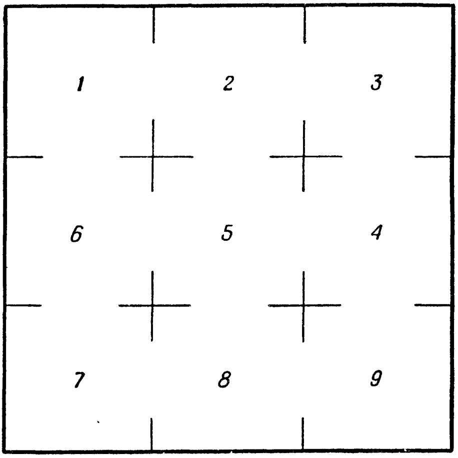
ЗАДАЧИ

1. Найти матрицы переходных вероятностей для марковских цепей, описывающих следующие процессы.

(а) Рассмотрим серию бросаний монеты с вероятностью выпадения решетки, равной p. Состояние процесса после n переходов (бросаний монеты) определим как разность между числом выпадений решетки и числом выпадений герба.

(б) В двух урнах размещены N черных и N белых шаров так, что каждая содержит по N шаров. Состоянием системы является число белых шаров в первой урне.

(в) Белую крысу помещают в лабиринт, изображенный на рисунке. Крыса передвигается из ячейки в ячейку случайным образом, т. е. если ячейка имеет k выходов, то крыса выбирает каждый из них с вероятностью $1/k$. В каждый момент времени крыса обязательно переходит в одну из соседних ячеек. Состояние системы — номер ячейки, в которой находится крыса.



(г) Рассмотрим производственную линию, где каждая единица выпускаемой продукции с вероятностью p идет в брак. Качество каждого отдельного изделия (годно или дефектно) предполагается не зависимым от качества других изделий. Процедура контроля качества изделий состоит в следующем. Сначала проверяется каждое выпускаемое изделие. Так продолжается до тех пор, пока не появятся i небракованных изделий подряд. В этом случае из каждых r последующих изделий для проверки равновероятно выбирается лишь одно. Если теперь будет обнаружено бракованное изделие, то процедура предписывает возвращение к исходному правилу: проверять каждое изделие впредь до появления r небракованных изделий подряд и т. д.

Состояние E_k $(k = 0, 1, \ldots, i)$ означает, что при проверке согласно первой части процедуры контроля (проверяется каждое выпускаемое изделие) последовательно появились k небракованных изделий. Состояние же E_{i+1} означает, что проверка осуществляется согласно второй части процедуры (проверяется одно изделие из r) и появилось одно, или более, небракованное изделие. (Предполагается, что время m отсчитывается вместе с появлением каждого изделия при проверке по первому правилу и с появлением серии из r изделий — по второму.)

Ответы:

(а)

$$P_{jk} = \mathrm{P}\left\{\begin{matrix}\text{разность между числом выпадений решетки и числом} \\ \text{выпадений герба } = k \text{ после } n+1 \text{ бросаний} \mid \text{эта раз-} \\ \text{ность } = j \text{ после } n \text{ бросаний}\end{matrix}\right\} =$$

$$= \begin{cases} p, & \text{если } k = j+1, \\ 1-p, & \text{если } k = j-1. \\ 0 & \text{во всех остальных случаях.} \end{cases}$$

P_{jk} не зависит от n.

(б)

$$P_{jk}=\mathrm{P}\left\{\begin{matrix} k \text{ белых шаров в первой урне после } n+1 \text{ перекладываний} \\ \mid j \text{ белых шаров в первой урне после } n \text{ перекладываний}\end{matrix}\right\}=$$

$$=\begin{cases} \left(\dfrac{j}{N}\right)^2, & \text{если } k=j-1, \quad j=1, 2, \ldots, N, \\ 2\left(\dfrac{j}{N}\right)\left(\dfrac{N-j}{N}\right), & \text{если } k=j, \quad j=0, 1, \ldots, N, \\ \left(1-\dfrac{j}{N}\right)^2, & \text{если } k=j+1, \quad j=0, 1, \ldots, N-1, \\ 0 & \text{во всех остальных случаях.}\end{cases}$$

P_{jk} не зависит от n.

(г)

$$P_{jk}=\mathrm{P}\left\{\begin{matrix} \text{пребывание в состоянии } E_k \text{ после } m+1 \text{ проверок} \mid \text{пре-} \\ \text{бывание в состоянии } E_j \text{ после } m \text{ проверок}\end{matrix}\right\}=$$

$$=\begin{cases} p, & \text{если } k=0, \quad j=0, 1, \ldots, i, i+1, \\ 1-p, & \text{если } k=j+1, \quad j=0, 1, \ldots, i \quad \text{или} \quad k=j=i+1, \\ 0 & \text{во всех остальных случаях}\end{cases}$$

при всех m.

2. (а) N шаров размещены в двух урнах А и Б. В момент времени t ($t = 1, 2, \ldots$) из общего числа N шаров случайно (все выборы равновероятны) выбирается один шар и помещается с вероятностью p в урну А и с вероятностью q в урну Б. Состояние системы при каждом испытании представляется числом шаров в урне А. Определить матрицу переходных вероятностей марковской цепи, описывающей серию таких испытаний.

(б) Предположим, что в момент t в урне А лежит k шаров. В момент $t+1$ с вероятностью, пропорциональной числу содержащихся в ней шаров, выбирается одна из урн [т. е. урна А выбирается с вероятностью k/N, а урна Б — с вероятностью $(N-k)/N$]. В выбранную урну кладется шар, который предварительно извлекается из урны А или из урны Б с вероятностями p и q ($p+q=1$) соответственно. Определить матрицу переходных вероятностей этой марковской цепи.

(в) Предположим теперь, что урна, из которой извлекается шар, также выбирается с вероятностью, пропорциональной числу содержащихся в ней шаров [т. е. шар извлекается из урны А с вероятностью k/N или из урны Б с вероятностью $(N-k)/N$ и возвращается в урну А с вероятностью k/N или в урну Б с вероятностью $(N-k)/N$]. Найти матрицу переходных вероятностей марковской цепи, состояния которой соответствуют числу шаров в урне А.

(г) Определить классы эквивалентности (классы сообщающихся состояний) в (а), (б) и (в).

Ответы:

(а)

$$P_{ik}=\begin{cases} \dfrac{N-i}{N}p, & \text{если } k=i+1, \\ \dfrac{i}{N}p+\dfrac{N-i}{N}q, & \text{если } k=i, \\ \dfrac{i}{N}q, & \text{если } k=i-1, \\ 0 & \text{во всех остальных случаях.}\end{cases} \qquad (i=0, 1, 2, \ldots, N)$$

Один класс эквивалентности: $\{0, 1, 2, \ldots, N\}$.

(б)

$$P_{ik}=\begin{cases}\dfrac{i}{N}q, & \text{если } k=i+1,\\[2mm] \dfrac{i}{N}p+\dfrac{N-1}{N}q, & \text{если } k=i, \qquad (i=1,\ 2,\ \ldots,\ N-1)\\[2mm] \dfrac{N-i}{N}p, & \text{если } k=i-1,\\[2mm] 0 & \text{во всех остальных случаях,}\end{cases}$$

$P_{ii}=1$, если $i=0$ и $i=N$. Классы эквивалентности: $\{0\}$, $\{N\}$, $\{1, 2, \ldots, N-1\}$.

(в)

$$P_{ik}=\begin{cases}\dfrac{i^2}{N^2}+\dfrac{(N-i)^2}{N^2}, & \text{если } k=i,\\[2mm] \dfrac{i(N-i)}{N^2}, & \text{если } k=i+1 \text{ и } k=i-1,\\[2mm] 0 & \text{во всех остальных случаях.}\end{cases}$$

Классы эквивалентности: $\{0\}$, $\{1, 2, \ldots, N-1\}$, $\{N\}$.

3. (а) В серии психологических экспериментов испытуемый реагирует на воздействия S_1, S_2, ..., S_N одним из двух возможных способов, A_1 или A_2. Каждое воздействие вызывает одну из этих реакций. В каждом эксперименте испытуемый подвергается случайно выбираемому воздействию (все воздействия имеют одинаковую вероятность быть выбранными) и реагирует на него в зависимости от того, с какой реакцией (A_1 или A_2) данное воздействие связано в настоящий момент. После проявления реакции испытуемый с вероятностью π $(0<\pi<1)$ получает поощрение независимо от предшествующей части опыта. Если поощрение реализовалось, то воздействие остается связанным с той же реакцией, в противном случае это воздействие вызовет у испытуемого другую реакцию в следующий раз, когда оно будет выбрано. Рассмотреть марковскую цепь, состояния которой отождествляются с числом воздействий, связанных в данный момент с реакцией A_1, и найти ее матрицу переходных вероятностей.

(б) Испытуемый S реагирует одним из трех возможных способов: A_0, A_1 или A_2. Реакция A_0 соответствует состоянию, в котором может произойти смена реакции. За реакцию A_1 испытуемый получает поощрение с вероятностью π_1. Если поощрение реализуется, то в следующем эксперименте S реагирует тем же способом. В противном случае (его вероятность равна $1-\pi_1$) S переходит в состояние смены реакции. Аналогично вероятность получить поощрение за реакцию A_2 равна π_2; S остается в том же состоянии (реагирует прежним образом), если он получает поощрение, и переходит в состояние смены реакции в противном случае. Находясь в состоянии смены реакции, испытуемый остается в нем до следующего эксперимента с вероятностью $1-c$ или реагирует способами A_1 или A_2 с одинаковыми вероятностями, равными $c/2$. Рассмотреть марковскую цепь, состояниями которой являются состояния испытуемого, и определить ее матрицу переходных вероятностей.

Ответы:

(а) $P_{ii}=\pi$; $P_{i,\,i+1}=[(N-i)/N](1-\pi)$, $P_{i,\,i-1}=(i/N)(1-\pi)$, $P_{ij}=0$ во всех остальных случаях $(i,\ j=1,\ 2,\ \ldots,\ N)$.

(б) $P_{00}=1-c$, $P_{01}=P_{02}=c/2$; $P_{10}=1-\pi_1$, $P_{11}=\pi_1$; $P_{20}=1-\pi_2$, $P_{22}=\pi_2$, $P_{12}=P_{21}=0$.

4. Каждой стохастической матрице соответствует марковская цепь, для которой она является матрицей одношаговых переходных вероятностей. (Под «стохастической матрицей» мы понимаем матрицу $\mathbf{P} = \|P_{ij}\|$, элементы которой удовлетворяют условиям $0 \leqslant P_{ij} \leqslant 1$ и $\sum_j P_{ij} = 1$.) Однако не всякая стохастическая матрица может служить в качестве матрицы двухшаговых переходных вероятностей марковской цепи. В частности, показать, что стохастическая матрица второго порядка является матрицей двухшаговых переходных вероятностей тогда и только тогда, когда сумма ее диагональных элементов больше или равна единице.

5. Пусть $n_1, n_2, \ldots, n_k$ — положительные целые числа, наибольший общий делитель которых равен d. Показать, что существует такое положительное целое число M, для которого из неравенства $m \geqslant M$ следует существование неотрицательных целых чисел $\{c_j\}_{j=1}^{k}$, таких, что

$$md = \sum_{j=1}^{k} c_j n_j.$$

(Этот результат потребуется ниже в задаче 7.)

Указание: Пусть

$$A = \{n \mid n = c_1 n_1 + \ldots + c_k n_k,\ \{c_i\} \text{ — неотрицательные целые}\},$$

$$B = \left\{ \begin{matrix} b_1 n_1 + \ldots + b_j n_j \mid n_1, n_2, \ldots, n_j \in A \text{ и } b_1, \ldots, b_j \\ \text{суть положительные или отрицательные целые} \end{matrix} \right\}.$$

Пусть d' — наименьшее положительное целое число в множестве B. Показать, что d' является общим делителем всех чисел множества A. Затем показать, что d' является наибольшим общим делителем чисел множества A. Следовательно, $d' = d$. Перенумеруем слагаемые в представлении $d = a_1 n_1 + \ldots + a_l n_l$ так, чтобы члены с положительными коэффициентами были записаны сначала. Тогда $d = N_1 - N_2$, где $N_1 \in A$ и $N_2 \in A$. Пусть $M = N_2^2/d$. Каждое целое число $m \geqslant M$ может быть записано в виде $m = M + k = N_2^2/d + k$ $(k = 0, 1, \ldots)$, причем $k = \delta N_2/d + b$, где $0 \leqslant b \leqslant N_2/d$ и $\delta = j$, если $j(N_2/d) \leqslant k \leqslant (j+1)N_2/d$ $(j = 0, 1, 2, \ldots)$. Таким образом, $md = N_2^2 + (\delta N_2/d + b)d = N_2(N_2 + \delta - b) + bN_1$

6. Доказать теорему 4.1.

Указание: Пусть $P_{ii}^s > 0$, $P_{jj}^{n+s+m} \geqslant P_{ji}^n P_{ii}^s P_{ij}^m > 0$ для некоторых $m > 0$ и $n > 0$. Поскольку $P_{ii}^{2s} > 0$, имеем $P_{jj}^{n+2s+m} > 0$; таким образом, $d(j)$ является делителем числа $(n + 2s + m) - (n + s + m) = s$.

7. Доказать теорему 4.2 и следствие 4.1.

Указание: В задаче 5 мы видели, что существует такое N, что если $n \geqslant N$, то

$$P_{ii}^{nd(i)} = P_{ii}^{(c_1 n_1 + \ldots + c_k n_k)}.$$

8. Доказать, что для непериодической неприводимой конечной цепи Маркова все элементы матрицы $\mathbf{P}^n$ положительны для некоторого n.

9. Доказать, что если j — невозвратное состояние, то для всех i имеем

$$\sum_{n=1}^{\infty} P_{ij}^n < \infty.$$

Указание: Использовать соотношение (5.9).

10. Пусть марковская цепь имеет r состояний. Показать, что

(а) если состояние k достижимо из состояния j, то оно может быть достигнуто меньше чем за r шагов;

(б) если j — возвратное состояние, то существует число α $(0 < \alpha < 1)$, такое, что для $n > r$ вероятность того, что первое возвращение в состояние j произойдет после n переходов, меньше или равна α^n.

11. Рассмотрим последовательность испытаний Бернулли, исходы которых описываются случайными величинами $X_1, X_2, X_3, \ldots$, где $X_n = 1$ или 0. Предположим, что

$$\mathrm{P}\{X_n = 1 \mid X_1, X_2, \ldots, X_{n-1}\} \geqslant \alpha > 0, \quad n = 1, 2, \ldots.$$

Доказать, что

(а) $\mathrm{P}\{X_n = 1$ для некоторого $n\} = 1$,

(б) $\mathrm{P}\{X_n = 1$ бесконечно много раз$\} = 1$.

НЕКОТОРЫЕ ЭЛЕМЕНТАРНЫЕ ЗАДАЧИ

1. Для марковских цепей с матрицами переходных вероятностей

$$\begin{Vmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{Vmatrix} \quad \text{и} \quad \begin{Vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \end{Vmatrix}$$

определить классы эквивалентности и периодичность различных состояний.

2. Рассмотрим случайное блуждание по одномерной целочисленной решетке, где $P_{i,\,i+1} = p$, $P_{i,\,i-1} = q$ для всех целых $i\,(0 < p < 1,\ p + q = 1)$. Определить $P_{00}^{(n)}$.

Ответ: $P_{00}^{2m} = \binom{2m}{m} p^m q^m$; $P_{00}^{(2m+1)} = 0$.

3 (продолжение). Найти производящую функцию вероятностей $u_n = P_{00}^n$, т. е. определить $P(x) = \sum\limits_{n=0}^{\infty} u_n x^n$.

Указание. Использовать тождество $\binom{2n}{n} = (-1)^n \binom{-\frac{1}{2}}{n} 2^{2n}$, где

$$\binom{a}{n} = \frac{a(a-1)\ldots(a-n+1)}{n!} \quad \text{для всех действительных } a.$$

Ответ:

$$P(x) = (1 - 4pqx^2)^{-1/2}.$$

4 (продолжение). Определить производящую функцию времени возвращения для состояния 0.

Ответ: $F(x) = 1 - \sqrt{(1 - 4pqx^2)}$.

5 (продолжение). Какова вероятность когда-нибудь возвратиться в начало координат?

6. Рассмотрим повторяющиеся независимые испытания, для каждого из которых возможны два исхода, У (успех) или Н (неудача). Определить распределение числа испытаний, требуемых для наступления события УН (т. е. успех и следующая за ним неудача). То же самое для событий УУН и УНУ.

7. Предположим, что две правильные различимые монеты бросаются одновременно и многократно. Ведется счет числу выпадений герба и решетки для каждой монеты. Рассмотрим событие E_n, заключающееся в том, что при n-м бросании суммарное число выпадений герба у одной монеты равно суммарному числу выпадений герба у другой. Установить связь между событием E_n и временем возвращения в заданное состояние при симметричном случайном блуждании на целочисленной решетке.

ЗАМЕЧАНИЯ

Некоторые аспекты теории марковских цепей освещены в книге Феллера [1].

Книга Кемени и Снелла [2] содержит много увлекательных примеров марковских цепей, встречающихся в психологии, социологии, экономике, биологии и других областях.

Наиболее полное и глубокое рассмотрение марковских цепей дано в книге Чжун Кай-лая [3].

ЛИТЕРАТУРА

1. Феллер В., Введение в теорию вероятностей и ее приложения, т. 1, «Мир», М., 1964.
2. Kemeny J. G., Snell J. L., Finite Markov Chains, Princeton, New Jersey, 1960.
3. Чжун Кай-лай, Однородные цепи Маркова, «Мир», М., 1964.

Глава 3

ОСНОВНЫЕ ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ МАРКОВСКИХ ЦЕПЕЙ И ИХ ПРИЛОЖЕНИЯ

§ 1. ДИСКРЕТНОЕ УРАВНЕНИЕ ВОССТАНОВЛЕНИЯ

Ключевым в анализе марковских цепей является результат, формулируемый в следующей теореме:

Теорема 1.1. *Пусть* $\{a_k\}$, $\{u_k\}$, $\{b_k\}$ — *числовые последовательности с индексом* k, *принимающим значения* $0, \pm 1, \pm 2, \ldots$. *Предположим, что* $a_k \geqslant 0$, $\sum a_k = 1$, $\sum |k| a_k < \infty$, $\sum k a_k > 0$, $\sum |b_k| < \infty$ *и что наибольший общий делитель индексов* k, *для которых* $a_k > 0$, *равен* 1. *Если уравнение восстановления*

$$u_n - \sum_{k=-\infty}^{\infty} a_{n-k} u_k = b_n, \qquad n = 0, \pm 1, \pm 2, \ldots,$$

имеет своим решением ограниченную последовательность $\{u_n\}$ *действительных чисел, то пределы* $\lim\limits_{n \to \infty} u_n$ *и* $\lim\limits_{n \to -\infty} u_n$ *существуют. Более того, если*

$$\lim_{n \to -\infty} u_n = 0, \quad \text{то} \quad \lim_{n \to \infty} u_n = \frac{\sum\limits_{k=-\infty}^{\infty} b_k}{\sum\limits_{k=-\infty}^{\infty} k a_k}. \tag{1.1}$$

В случае, если $\sum\limits_{k=-\infty}^{\infty} k a_k = \infty$, *предельное соотношение остается справедливым, если положить*

$$\frac{\sum\limits_{k=-\infty}^{\infty} b_k}{\sum\limits_{k=-\infty}^{\infty} k a_k} = 0.$$

Доказательство этой теоремы в том общем виде, в каком она сформулирована выше, выходит за рамки этой книги. Нам понадобится ее частный случай для последовательностей $\{a_k\}$, $\{u_k\}$, $\{b_k\}$, обращающихся в нуль при отрицательных значениях k, и $b_k \geqslant 0$. Доказательство теоремы для этого случая мы дадим в § 2.

Замечание 1.1. В случае, когда $a_{-k}=0$, $b_{-k}=0$ и $u_{-k}=0$ для $k>0$, уравнение восстановления принимает вид

$$u_n=\sum_{k=0}^{n} a_{n-k}u_k=b_n, \quad n=0,\ 1,\ 2,\ \ldots .$$

Замечание 1.2. (Довод в пользу названия «уравнение восстановления».)

Пусть «время жизни» электрической лампочки представляет собой случайную величину ξ и измеряется в дискретных единицах, причем

$$\mathrm{P}\{\xi=k\}=a_k, \quad k=0,\ 1,\ 2,\ \ldots, \quad a_k\geqslant 0, \quad \sum_{k=0}^{\infty} a_k=1.$$

Как только лампочка перегорает, ее сразу же заменяют новой. Пусть первая лампочка перегорает в момент ξ_1, вторая — в момент $\xi_1+\xi_2$ и n-я — в момент $\sum_{i=1}^{n}\xi_i$, где ξ_i — взаимно независимые, одинаково распределенные случайные величины (распределение каждой из них совпадает с распределением с. в. ξ). Пусть u_n обозначает среднее число замен, произведенных к моменту n. Если первая замена имела место в момент k, то среднее число замен в оставшееся до момента n время есть u_{n-k}. Суммируя по всем возможным значениям k, получаем

$$u_n=\sum_{k=0}^{n}(1+u_{n-k})a_k+0\cdot\sum_{k=n+1}^{\infty}a_k=$$
$$=\sum_{k=0}^{n}u_{n-k}a_k+\sum_{k=0}^{n}a_k=\sum_{k=0}^{n}a_{n-k}u_k+b_n, \tag{1.2}$$

где

$$\sum_{k=0}^{n}a_k=b_n.$$

Обоснование соотношения (1.2) таково. Величина $1+u_{n-k}$ представляет собой среднее число замен за время n при условии, что первая лампочка перегорела в момент k $(0\leqslant k\leqslant n)$; вероятность этого события равна a_k. Вторая сумма есть вероятность того, что первая лампочка будет служить более чем n единиц времени. Учитывая повторяющийся характер процесса, мы получаем выражение для u_n с помощью разложения возможных реализаций по моменту первой замены.

Следующая теорема, так называемая «эргодическая» теорема для данного частного случая, описывает предельное поведение вероятностей P_{ij}^n при $n\to\infty$ для всех i и j для непериодической возвратной марковской цепи.

Теорема 1.2. (Основная предельная теорема для марковских цепей.)

(а) *Рассмотрим возвратную неприводимую непериодическую марковскую цепь. Пусть* P_{ii}^n *есть вероятность оказаться в* i*-м состоянии на* n*-м шаге,* $n = 0, 1, 2, \ldots,$ *при условии, что* $X(0) = i$ *(т. е. состояние* i *— начальное). Пусть, как и ранее,* $P_{ii}^0 = 1$*. Пусть* f_{ii}^n *есть вероятность впервые возвратиться в состояние* i *на* n*-м шаге, причем* $f_{ii}^0 = 0$*. По доказанному ранее* [*см. формулу* (5.1) *гл.* 2], *имеем*

$$P_{ii}^n - \sum_{k=0}^{n} f_{ii}^{n-k} P_{ii}^k = \begin{cases} 1, & \text{если} \quad n = 0, \\ 0, & \text{если} \quad n > 0. \end{cases}$$

В этом случае справедливо равенство

$$\lim_{n\to\infty} P_{ii}^n = \frac{1}{\sum\limits_{n=0}^{\infty} n f_{ii}^n}.$$

(б) *При этих же условиях* $\lim\limits_{n\to\infty} P_{ji}^n = \lim\limits_{n\to\infty} P_{ii}^n$.

Доказательство. (а) Положим

$$u_n = P_{ii}^n, \quad n \geqslant 0; \quad u_n = 0, \quad n < 0;$$

$$a_n = f_{ii}^n, \quad n \geqslant 0; \quad a_n = 0, \quad n < 0;$$

$$b_n = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

Воспользовавшись теперь теоремой 1.1, получаем доказываемый результат.

(б) Пусть

$$y_n = \sum_{k=0}^{n} a_{n-k} x_k,$$

где

$$a_m \geqslant 0, \quad \sum_{m=0}^{\infty} a_m = 1, \quad \lim_{k\to\infty} x_k = c.$$

Докажем, что при этих условиях $\lim\limits_{n\to\infty} y_n = c$. В самом деле, имеем

$$y_n - c = \sum_{k=0}^{n} a_{n-k} x_k - c \sum_{m=0}^{\infty} a_m = \sum_{k=0}^{n} a_{n-k}(x_k - c) - c \sum_{m=n+1}^{\infty} a_m.$$

Для любого заданного $\varepsilon > 0$ существует $K(\varepsilon)$, такое, что $|x_k - c| < \varepsilon/3$ для всех $k \geqslant K(\varepsilon)$. Следовательно,

$$y_n - c = \sum_{k=0}^{K(\varepsilon)} a_{n-k}(x_k - c) + \sum_{k=K(\varepsilon)+1}^{n} a_{n-k}(x_k - c) - c \sum_{m=n+1}^{\infty} a_m,$$

откуда

$$|y_n - c| \leqslant M \sum_{k=0}^{K(\varepsilon)} a_{n-k} + \frac{\varepsilon}{3} \sum_{h=K(\varepsilon)+1}^{n} a_{n-h} + |c| \sum_{m=n+1}^{\infty} a_m,$$

где

$$M = \max_{k \geqslant 0} |x_k - c|.$$

Выберем теперь $N(\varepsilon)$, такое, что $|c| \sum\limits_{m=n+1}^{\infty} a_m < \varepsilon/3$ и

$$\sum_{k=0}^{K(\varepsilon)} a_{n-k} \equiv \sum_{m=n-K(\varepsilon)}^{n} a_m < \frac{\varepsilon}{3M} \quad \text{для} \quad n \geqslant N(\varepsilon);$$

мы видим, что

$$|y_n - c| \leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \text{для} \quad n \geqslant N(\varepsilon).$$

Воспользуемся теперь ранее доказанным нами соотношением (см. формулу (5.9) гл. 2)

$$P_{ji}^n = \sum_{v=0}^{n} f_{ji}^v P_{ii}^{n-v}, \quad i \neq j, \quad n \geqslant 0.$$

Полагая

$$y_n = P_{ji}^n, \quad a_n = f_{ji}^n, \quad x_n = P_{ii}^n,$$

получаем доказываемый результат.

Замечание 1.3. Пусть C — возвратный класс. Тогда $P_{ij}^n = 0$ при $i \in C$ и $j \notin C$ для всех n. Следовательно, попав в C, выйти из него невозможно. Таким образом, подматрица $\|P_{ij}\|$, $i, j \in C$, является матрицей переходных вероятностей, а соответствующая марковская цепь неприводима и возвратна. Это означает, что предельная теорема справедлива дословно для любого непериодического возвратного класса.

Замечание 1.4. Если $a_n \to a$ при $n \to \infty$, то, как легко показать, выполняется равенство

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} a_k = a. \tag{1.3}$$

Значит, если состояние i входит в возвратный непериодический класс, то

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ii}^{m} = \frac{1}{\sum\limits_{n=0}^{\infty} n f_{ii}^{n}} = \frac{1}{m_i}, \tag{1.4}$$

где m_i — среднее время возвращения.

Если состояние i входит в возвратный периодический класс, то, как можно показать (см. задачу 7 гл. 2), $P_{ii}^{m}=0$, если m не кратно периоду d (т. е. если $m \neq nd$ для какого-либо n), и

$$\lim_{n\to\infty} P_{ii}^{nd} = \frac{d}{m_i}.$$

Эти два последних результата вместе с (1.3) показывают, что соотношение (1.4) справедливо и для периодического случая.

Если $\lim\limits_{n\to\infty} P_{ii}^{n} = \pi_i > 0$ для некоторого состояния i из непериодического возвратного класса, то $\pi_j > 0$ для всех j из этого класса. (Доказательство этого факта аналогично доказательству следствия 5.1, и мы его опускаем.) В этом случае мы называем класс *возвратным положительным*, или *сильно эргодическим*. Если все $\pi_i = 0$ и класс возвратный, то будем говорить, что класс *возвратный нулевой*, или *слабо эргодический*.

Теорема 1.3. *Для непериодического возвратного положительного класса с состояниями* $j = 0, 1, 2, \ldots$

$$\lim_{n\to\infty} P_{jj}^{n} = \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \qquad \sum_{i=0}^{\infty} \pi_i = 1,$$

и величины $\{\pi_i\}$ *однозначно определяются условиями*

$$\pi_i \geqslant 0, \quad \sum_{i=0}^{\infty} \pi_i = 1, \quad \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}. \tag{1.5}$$

Набор $\{\pi_i\}$, удовлетворяющий условиям (1.5), называется *стационарным распределением* марковской цепи. Подробнее об этом речь пойдет в гл. 5.

Доказательство. Для любых n и M

$$1 = \sum_{j=0}^{\infty} P_{ij}^{n} \geqslant \sum_{j=0}^{M} P_{ij}^{n}.$$

Устремляя $n \to \infty$ и используя теорему 1.2, получаем

$$1 \geqslant \sum_{j=0}^{M} \pi_j$$

для любого M, откуда $\sum_{j=0}^{M} \pi_j \leqslant 1$. Далее, $P_{ij}^{n+1} \geqslant \sum_{k=0}^{M} P_{ik}^n P_{kj}$; при $n \to \infty$ это дает $\pi_j \geqslant \sum_{k=0}^{M} \pi_k P_{kj}$. Поскольку левая часть этого неравенства не зависит от M, при $M \to \infty$ получаем

$$\pi_j \geqslant \sum_{k=0}^{\infty} \pi_k P_{kj}. \tag{1.6}$$

Умножая обе части (1.6) на P_{ji} и суммируя по j, получаем неравенство $\pi_j \geqslant \sum_{k=0}^{\infty} \pi_k P_{kj}^2$. Точно так же убеждаемся, что это неравенство справедливо для любого n: $\pi_j \geqslant \sum_{k=0}^{\infty} \pi_k P_{kj}^n$. Предположим, что для некоторого j_0 имеет место строгое неравенство. Суммируя по j, имеем

$$\sum_{j=0}^{\infty} \pi_j > \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \pi_k P_{kj}^n = \sum_{k=0}^{\infty} \pi_k \sum_{j=0}^{\infty} P_{kj}^n = \sum_{k=0}^{\infty} \pi_k;$$

таким образом, $\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}^n$ для всех n. Поскольку ряд $\sum \pi_k$ сходится, а P_{kj}^n равномерно ограничены, при $n \to \infty$

$$\pi_j = \sum_{k=0}^{\infty} \pi_k \lim_{n \to \infty} P_{kj}^n = \pi_j \sum_{k=0}^{\infty} \pi_k \text{ для всех } j.$$

В силу того что класс возвратный положительный, имеем $\pi_j > 0$, и поэтому $\sum_{k=0}^{\infty} \pi_k = 1$.

Предположим теперь, что последовательность $x = \{x_n\}$ удовлетворяет соотношениям (1.5), тогда

$$x_k = \sum_{j=0}^{\infty} x_j P_{jk} = \sum_{j=0}^{\infty} x_j P_{jk}^n,$$

откуда, устремляя $n \to \infty$, получаем

$$x_k = \sum_{j=0}^{\infty} x_j \lim_{n \to \infty} P_{jk}^n = \pi_k \sum_{j=0}^{\infty} x_j = \pi_k. \blacksquare$$

Пример. Рассмотрим класс процессов случайного блуждания, матрицы переходных вероятностей которых имеют вид

$$\mathbf{P} = \| P_{ij} \| = \begin{Vmatrix} 0 & 1 & 0 & \ldots & \\ q_1 & 0 & p_1 & \ldots & \\ 0 & q_2 & 0 & p_2 & \ldots \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \end{Vmatrix}$$

(см. пример Б гл. 2).

Мы исследуем существование стационарного распределения, т. е. найдем положительное решение уравнений

$$x_i = \sum_{j=0}^{\infty} x_j P_{ji} = p_{i-1} x_{i-1} + q_{i+1} x_{i+1}, \qquad i = 0,\ 1,\ \ldots, \tag{1.7}$$

при условии «нормировки»

$$\sum_{i=0}^{\infty} x_i = 1,$$

где $p_{-1} = 0$ и $p_0 = 1$, а значит, $x_0 = q_1 x_1$. Уравнение (1.7) при $i = 1$ позволяет выразить x_2 через x_0, при $i = 2$ выразить x_3 через x_0 и т. д. Легко проверить, что

$$x_i = \frac{p_{i-1} p_{i-2} \cdots p_1}{q_i q_{i-1} \cdots q_1} x_0 = x_0 \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}}, \qquad i \geqslant 1,$$

удовлетворяют (1.7), причем x_0 еще надлежит определить. Используя условие нормировки, получаем

$$1 = x_0 + \sum_{i=1}^{\infty} x_0 \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}},$$

или

$$x_0 = \frac{1}{1 + \sum\limits_{i=1}^{\infty} \prod\limits_{k=0}^{i-1} \frac{p_k}{q_{k+1}}}.$$

Таким образом, $x_0 > 0$ тогда и только тогда, когда

$$\sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} < \infty.$$

В частности, если $p_k = p$ и $q_k = q = 1 - p$ при $k \geqslant 1$, ряд

$$\sum_{i=1}^{\infty} \prod_{k=0}^{i-1} \frac{p_k}{q_{k+1}} = \frac{1}{p} \sum_{i=1}^{\infty} \left(\frac{p}{q}\right)^i$$

сходится, только когда $p < q$.

§ 2. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1.1

Мы докажем теорему 1.1 при следующих дополнительных предположениях: при $k < 0$ все члены последовательностей $\{a_k\}$, $\{b_k\}$, $\{u_k\}$ равны нулю; $b_k \geqslant 0$ и $a_1 > 0$. Уравнение восстановления в этом случае принимает вид

$$u_n - \sum_{k=0}^{n} a_{n-k} u_k = b_n, \qquad n = 0,\ 1,\ 2,\ \ldots,$$

или, что то же самое,

$$u_n - \sum_{k=0}^{n} a_k u_{n-k} = b_n, \qquad n = 0,\ 1,\ 2,\ \ldots. \tag{2.1}$$

Применяя индукцию (рассматривая уравнения последовательно), легко убедиться, что $u_k > 0$ при всех k. Поскольку, согласно предположению, $\{u_n\}$ — ограниченная последовательность, величина $\lambda = \overline{\lim\limits_{n\to\infty}}\, u_n$ конечна. Пусть $\{u_{n_j}\}$ такая подпоследовательность, что $\lim\limits_{j\to\infty} u_{n_j} = \lambda$. Используя условие $a_1 > 0$, мы докажем, что $\lim\limits_{j\to\infty} u_{n_j-1} = \lambda$. Предположим противное; тогда из определения λ следует, что существует такое $\lambda' < \lambda$, что $u_{n_j-1} < \lambda'$ для бесконечного числа значений индекса j. Положим $\varepsilon = [a_1(\lambda - \lambda')]/4$ и $M = \sup\limits_{n\geqslant 0} u_n$ и найдем N, такое, что

$$\sum_{k=0}^{n} a_k > 1 - \frac{\varepsilon}{M}, \quad \text{если } n \geqslant N. \tag{2.2}$$

Пусть j настолько велико, что $n_j \geqslant N$ и

$$u_{n_j} > \lambda - \varepsilon, \quad u_{n_j-1} < \lambda' < \lambda, \quad 0 \leqslant b_{n_j} < \varepsilon$$

и

$$u_n < \lambda + \varepsilon \quad \text{для всех } n \geqslant n_j - N. \tag{2.3}$$

Такое j существует по определению λ и λ'.

Из (2.1), (2.2) и (2.3) имеем

$$\begin{aligned}u_{n_j} &\leqslant \sum_{k=0}^{n_j} a_k u_{n_j-k} + \varepsilon < \sum_{k=0}^{N} a_k u_{n_j-k} + M \sum_{k=N+1}^{n_j} a_k + \varepsilon < \\ &< \sum_{k=0}^{N} a_k u_{n_j-k} + 2\varepsilon < \\ &< (a_0 + a_2 + a_3 + \ldots + a_N)(\lambda + \varepsilon) + a_1\lambda' + 2\varepsilon \leqslant \\ &\leqslant (1 - a_1)(\lambda + \varepsilon) + a_1\lambda' + 2\varepsilon < \lambda + 3\varepsilon - a_1(\lambda - \lambda') = \lambda - \varepsilon.\end{aligned}$$

Но это противоречит первому из неравенств (2.3), и, следовательно, $\lim\limits_{j\to\infty} u_{n_j-1} = \lambda$. Повторением предыдущих рассуждений убеждаемся, что для любого целого числа $d \geqslant 0$

$$\lim_{j\to\infty} u_{n_j-d} = \lambda. \tag{2.4}$$

Далее, положим $r_n = a_{n+1} + a_{n+2} + \ldots$; очевидно, $\sum\limits_{k=0}^{\infty} k a_k = \sum\limits_{n=0}^{\infty} r_n$. (Заметим, что сходимости ряда $\sum r_n$ не требуется.) Подставляя $a_1 = r_0 - r_1$, $a_2 = r_1 - r_2$, ... в (2.1), получаем

$$r_0 u_n + r_1 u_{n-1} + \ldots + r_n u_0 = r_0 u_{n-1} + r_1 u_{n-2} + \ldots + r_{n-1} u_0 + b_n$$
$$(n = 1, 2, \ldots).$$

Полагая $A_n = r_0 u_n + \ldots + r_n u_0$, мы можем записать последнее равенство в виде

$$A_n = A_{n-1} + b_n, \qquad n = 1, 2, \ldots,$$

где $A_0 = r_0 u_0 = (1 - a_0) u_0 = b_0$. Отсюда следует, что $A_n = \sum\limits_{i=0}^{n} b_i$. Так как $r_n \geqslant 0$ и $u_n \geqslant 0$ при всех n, то для любых фиксированных $N > 0$ и $j > 0$ имеет место неравенство

$$r_0 u_{n_j} + r_1 u_{n_j-1} + \ldots + r_N u_{n_j-N} \leqslant A_{n_j} = \sum_{n=0}^{n_j} b_n.$$

Устремляя $j \to \infty$, получаем неравенство $(r_0 + \ldots + r_N)\lambda \leqslant \sum\limits_{n=0}^{\infty} b_n$, которое можно записать в виде $\lambda \leqslant \sum\limits_0^{\infty} b_n \left(\sum\limits_0^{N} r_n\right)^{-1}$. Поскольку $N > 0$ произвольно, отсюда следует неравенство

$$\lambda \leqslant \frac{\sum\limits_{n=0}^{\infty} b_n}{\sum\limits_{n=0}^{\infty} r_n}. \tag{2.5}$$

Так как $u_k \geqslant 0$ при всех k, из неравенства (2.5) следует утверждение теоремы для случая, когда $\sum\limits_{n=0}^{\infty} r_n = \infty$, поскольку [как это следует из (2.5)] $\lambda = \lim\limits_{n\to\infty} u_n = 0$.

Если $\sum\limits_{n=0}^{\infty} r_n < \infty$, положим $\mu = \varliminf\limits_{n\to\infty} u_n$. Те же рассуждения, что и для верхнего предела, показывают, что если $\lim\limits_{j\to\infty} u_{n_j} = \mu$, то $\lim\limits_{j\to\infty} u_{n_j - d} = \mu$ для любого целого числа $d \geqslant 0$. Положим $\sum\limits_{n=N+1}^{\infty} r_n = g(N)$; ясно, что $\lim\limits_{N\to\infty} g(N) = 0$ и

$$\sum_{n=0}^{n_j} b_n \leqslant r_0 u_{n_j} + r_1 u_{n_j - 1} + \ldots + r_N u_{n_j - N} + g(N) M.$$

Устремляя $j \to \infty$, получаем

$$\sum_{n=0}^{\infty} b_n \leqslant (r_0 + \ldots + r_N) \mu + g(N) M.$$

Переходя теперь к пределу при $N \to \infty$, получаем неравенство

$$\sum_{n=0}^{\infty} b_n \leqslant \mu \sum_{n=0}^{\infty} r_n, \quad \text{или} \quad \mu \geqslant \frac{\sum\limits_{n=0}^{\infty} b_n}{\sum\limits_{n=0}^{\infty} r_n}. \tag{2.6}$$

Из (2.5) и (2.6) следует, что $\mu \geqslant \lambda$. С другой стороны, $\mu \leqslant \lambda$ по самому их смыслу. Следовательно, $\mu = \lambda$, что означает, что предел $\lim\limits_{n\to\infty} u_n$ существует и, более того,

$$\lim_{n\to\infty} u_n = \frac{\sum\limits_{n=0}^{\infty} b_n}{\sum\limits_{n=0}^{\infty} r_n}.$$

Для случая, когда $a_1 = 0$, но наибольший общий делитель целых чисел m, для которых $a_m > 0$, равен 1, теорему 1.1 можно доказать аналогичным способом, воспользовавшись при этом следствием 4.1 гл. 2.

§ 3. ВЕРОЯТНОСТИ ПОГЛОЩЕНИЯ

Ранее мы установили (см. задачу 9 гл. 2), что если состояние j невозвратное, то $P_{ij}^n \to 0$, и что если состояния i и j принадлежат одному и тому же непериодическому возвратному классу, то

$P_{ij}^n \to \pi_j \geqslant 0$. Если состояния i и j входят в один и тот же возвратный периодический класс, то последнее утверждение сохраняет силу, если P_{ij}^n заменить в нем на $n^{-1}\sum_{m=1}^{n} P_{ij}^m$. Для того, чтобы завершить рассмотрение предельного поведения вероятностей P_{ij}^n, остается рассмотреть случай, когда состояние i невозвратное, а состояние j возвратное.

Пусть T — множество всех невозвратных состояний; введем величины x_i^n с помощью следующей рекуррентной формулы:

$$x_i^1 = \sum_{j \in T} P_{ij},$$
$$x_i^n = \sum_{j \in T} P_{ij} x_j^{n-1},$$

где $i \in T$ и $n \geqslant 2$. Заметим, что x_i^n есть вероятность того, что, отправившись из состояния i, процесс не выйдет из класса T в течение следующих n шагов. Покажем с помощью индукции, что последовательность $\{x_i^n,\ n = 1, 2, \ldots\}$ является невозрастающей. Действительно, так как $x_i^n \leqslant 1$ при всех n, то

$$x_i^2 = \sum_{j \in T} P_{ij} x_j^1 \leqslant \sum_{j \in T} P_{ij} = x_i^1.$$

Предположим теперь, что $x_j^n \leqslant x_j^{n-1}$ при всех $j \in T$, тогда

$$0 \leqslant x_i^{n+1} = \sum_{j \in T} P_{ij} x_j^n \leqslant \sum_{j \in T} P_{ij} x_j^{n-1} = x_i^n.$$

Это означает, что ограниченная последовательность $\{x_i^n,\ n = 1, 2, \ldots\}$ не возрастает и, следовательно, стремится к некоторому пределу x_i, причем

$$x_i = \sum_{j \in T} P_{ij} x_j, \qquad i \in T. \tag{3.1}$$

Таким образом, если единственным ограниченным решением уравнений (3.1) является нулевой вектор $(0, 0, \ldots)$, то, отправляясь из любого невозвратного состояния, процесс с вероятностью 1 будет поглощен некоторым классом возвратных состояний. В самом деле, x_i $(i \in T)$ есть вероятность никогда не попасть в возвратный класс, если i было начальным состоянием процесса. Поскольку $\{x_i,\ i \in T\}$ является ограниченным решением уравнений (3.1), то $x_i = 0$ при всех i.

Замечание 3.1. Если марковская цепь имеет лишь конечное число состояний, скажем M, то среди них нет возвратных нулевых состояний, а все состояния не могут быть невозвратными. Действительно, так как $\sum_{j=0}^{M-1} P_{ij}^n = 1$ для всех n, то $\lim_{n\to\infty} P_{ij}^n = 0$ не для всех j. По этой же причине отсутствуют нулевые возвратные состояния. Пусть C, C_1, C_2, ... — возвратные классы. Определим $\pi_i(C)$ как вероятность того, что, отправляясь из невозвратного состояния i, процесс рано или поздно войдет в класс C. (Вспомним, что, однажды попав в возвратный класс, процесс уже никогда его не покидает.)

Пусть $\pi_i^n(C)$ есть вероятность того, что процесс достигнет класса C и, следовательно, будет им поглощен, впервые на n-м шаге, при условии, что начальным состоянием было $i \in T$; тогда

$$\pi_i(C) = \sum_{n=1}^{\infty} \pi_i^n(C) \leqslant 1,$$

$$\pi_i^1(C) = \sum_{j \in C} P_{ij}, \tag{3.2}$$

$$\pi_i^n(C) = \sum_{j \in T} P_{ij} \pi_j^{n-1}(C), \qquad n \geqslant 2. \tag{3.3}$$

Используя (3.3), (3.2) можно переписать в виде

$$\pi_i(C) = \pi_i^1(C) + \sum_{n=2}^{\infty} \pi_i^n(C) = \pi_i^1(C) + \sum_{n=2}^{\infty} \sum_{j \in T} P_{ij} \pi_j^{n-1}(C) =$$

$$= \pi_i^1(C) + \sum_{j \in T} P_{ij} \sum_{n=2}^{\infty} \pi_j^{n-1}(C),$$

$$\pi_i(C) = \pi_i^1(C) + \sum_{j \in T} P_{ij} \pi_j(C), \qquad i \in T. \tag{3.4}$$

Если предположить, что единственным *ограниченным* решением однородной системы уравнений

$$w_i = \sum_{j \in T} P_{ij} w_j, \qquad i \in T,$$

является тривиальное решение (нулевой вектор), то $\{\pi_i(C)\}$ является единственным ограниченным решением системы уравнений (3.4). Более того, либо $\pi_i^1(C) > 0$ для некоторого $i \in T$, либо $\pi_i(C) = 0$ для всех $i \in T$ и, следовательно, $\pi_i^n(C) = 0$ для всех n.

Теорема 3.1. *Пусть $j \in C$ (C — непериодический возвратный класс), тогда для $i \in T$ имеем*

$$\lim_{n\to\infty} P_{ij}^n = \pi_i(C) \lim_{n\to\infty} P_{jj}^n = \pi_i(C)\, \pi_j.$$

Доказательство. Легко видеть, что $\pi_i^n(C) = \sum_{k \in C} \pi_{ik}^n(C)$, где $\pi_{ik}^n(C)$ есть вероятность того, что, отправляясь из состояния $i \in T$, процесс на n-м шаге войдет в класс C через состояние k. Имеем

$$\pi_i(C) = \sum_{\nu=1}^{\infty} \sum_{k \in C} \pi_{ik}^{\nu}(C) \leqslant 1.$$

Следовательно, для любого $\varepsilon > 0$ существуют конечное число состояний $C' \subset C$ и целое число $N(\varepsilon)$, такие, что

$$\left| \pi_i(C) - \sum_{\nu=1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu}(C) \right| < \varepsilon, \quad \text{т. е.} \quad \left| \sum_{\nu=1}^{\infty} \sum_{k \in C} \pi_{ik}^{\nu} - \sum_{\nu=1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} \right| < \varepsilon \tag{3.5}$$

для всех $n > N(\varepsilon)$. (Здесь мы пишем π_{ik}^{ν} вместо $\pi_{ik}^{\nu}(C)$.)

Для $j \in C$ рассмотрим разность

$$P_{ij}^n - \sum_{\nu=1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} \pi_j.$$

С помощью знакомых нам рассуждений получаем

$$P_{ij}^n = \sum_{\nu=1}^{n} \sum_{k \in C} \pi_{ik}^{\nu} P_{kj}^{n-\nu}, \quad i \in T, \quad j \in C.$$

Опираясь на эти соотношения, получаем неравенство

$$\left| P_{ij}^n - \left(\sum_{\nu=1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} \right) \pi_j \right| =$$

$$= \left| \sum_{\nu=1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} (P_{kj}^{n-\nu} - \pi_j) + \sum_{\nu=1}^{n} \sum_{k \in C,\ k \notin C'} \pi_{ik}^{\nu} P_{kj}^{n-\nu} \right| \leqslant$$

$$\leqslant \left| \sum_{\nu=1}^{N} \sum_{k \in C'} \pi_{ik}^{\nu} (P_{kj}^{n-\nu} - \pi_j) \right| +$$

$$+ \left| \sum_{\nu=N+1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} (P_{kj}^{n-\nu} - \pi_j) \right| + \sum_{\nu=1}^{n} \sum_{k \in C,\ k \notin C'} \pi_{ik}^{\nu} P_{kj}^{n-\nu}.$$

Но $P_{kj}^{n-\nu} \leqslant 1$, $|P_{kj}^{n-\nu} - \pi_j| \leqslant 2$ и $\lim_{n \to \infty} P_{kj}^{n-\nu} = \pi_j$, если C — непериодический класс и $k \in C'$. Поэтому существует такое $N' > N$, что при $n > N'$ имеем $|P_{kj}^{n-N} - \pi_j| < \varepsilon$ $(k \in C')$, так что при $n > N'$ выполняется неравенство

$$\left| P_{ij}^n - \left(\sum_{\nu=1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} \right) \pi_j \right| \leqslant \varepsilon + 2 \sum_{\nu=N+1}^{n} \sum_{k \in C'} \pi_{ik}^{\nu} + \sum_{\nu=1}^{n} \sum_{k \in C,\ k \in C} \pi_{ik}^{\nu}.$$

Однако выбор N и C' гарантирует нам, что правая часть последнего неравенства не больше, чем 4ε. Отсюда, воспользовавшись (3.5), получаем

$$|P_{ij}^n - \pi_i(C)\pi_j| \leqslant 4\varepsilon + \varepsilon\pi_j \quad \text{при} \quad n > N'(\varepsilon)$$

и, следовательно,

$$\lim_{n\to\infty} P_{ij}^n = \pi_i(C)\pi_j. \ \blacksquare$$

Если C — периодический класс и $j \in C$, то точно так же можно показать, что

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = \pi_i(C)\pi_j.$$

В заключение заметим, что если i — невозвратное состояние, а j — возвратное, то предельное значение вероятности зависит от обоих состояний i и j. В этом состоит существенное отличие от случая, когда i и j принадлежат одному и тому же возвратному классу.

Пример. (Задача о разорении игрока, играющего с партнером, капитал которого ограничен.)

Как мы видели, марковская цепь, описывающая игру, имеет конечное число состояний, скажем $n+1$, а ее матрица переходных вероятностей имеет вид

$$\left\| \begin{matrix} 1 & 0 & 0 & 0 & \dots & & & \\ q & 0 & p & 0 & \dots & & & \\ 0 & q & 0 & p & \dots & & & \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ \cdot & & & & & & & \\ & & & & \dots & q & 0 & p \\ & & & & \dots & 0 & 0 & 1 \end{matrix} \right\|.$$

Мы найдем $u_i = \pi_i(C_0)$ и $v_i = \pi_i(C_n)$ — вероятности, отправляясь из состояния i, рано или поздно попасть в поглощающие (и, следовательно, возвратные) состояния 0 и n соответственно. Система уравнений (3.4) для рассматриваемой задачи имеет вид

$$\begin{aligned} u_1 &= q + pu_2, \\ u_i &= qu_{i-1} + pu_{i+1} \qquad (2 \leqslant i \leqslant n-2), \\ u_{n-1} &= qu_{n-2}. \end{aligned} \tag{3.6}$$

Система состоит из $n-1$ неоднородных уравнений с $n-1$ неизвестными. Будем искать решение в виде $u_r = x^r$. Подставляя это выражение в средние из уравнений (3.6), получаем

$$px^2 + q = x.$$

Последнее уравнение имеет два решения, $x=1$ и $x=q/p$. Таким образом, величины $u_r = A + B(q/p)^r$, $r=2, 3, \ldots, n-1$, удовлетворяют средним уравнениям из (3.6) при любых значениях A и B. Определим A и B так, чтобы первое и последнее уравнения также удовлетворялись. (Если $q=p$, то решение $x=1$ является двукратным корнем уравнения $px^2+q=x$, в этом случае $(q/p)^r$ следует заменить на r.) Подставляя соответствующие выражения в первое уравнение, получаем

$$A + B\frac{q}{p} = q + p\left(A + B\frac{q^2}{p^2}\right)$$

или, упрощая,

$$A = 1 - B.$$

Последнее уравнение дает

$$A + B\left(\frac{q}{p}\right)^{n-1} = q\left(A + B\left(\frac{q}{p}\right)^{n-2}\right), \text{ или } p^nA + q^nB = 0.$$

Отсюда

$$A = \frac{q^n}{q^n - p^n}, \quad B = \frac{-p^n}{q^n - p^n}$$

и

$$u_r = \frac{(q/p)^n - (q/p)^r}{(q/p)^n - 1}, \text{ если } \frac{q}{p} \neq 1.$$

Если $q=p$, то точно так же находим, что $A=1$, $B=-1/n$, так что

$$u_r = \frac{n-r}{n}, \text{ если } p=q.$$

Аналогичные выкладки показывают, что

$$v_i = 1 - u_i,$$

чего и следовало ожидать, поскольку поглощение одним из классов C_0 или C_n есть событие достоверное.

Рассмотрим теперь игру с бесконечно богатым партнером. Уравнения для вероятности разорения игрока (поглощение состоянием 0) имеют вид

$$\begin{aligned} u_1 &= q + pu_2, \\ u_i &= qu_{i-1} + pu_{i+1}, \qquad i \geqslant 2. \end{aligned} \tag{3.7}$$

Так же как и раньше, находим, что

$$u_i = A + B(q/p)^i \quad (q \neq p) \quad \text{и} \quad u_i = A + Bi \quad (q = p = 1/2).$$

Если $q \geqslant p$, то из условия ограниченности u_i следует, что $B = 0$, а из первого из уравнений (3.7) следует, что $u_i \equiv 1$. Если $q < p$, то мы находим, что $u_i = (q/p)^i$, для чего нужно лишь перейти к пределу при $n \to \infty$ в решении предыдущей задачи — задачи о разорении с конечным числом состояний.

§ 4. КРИТЕРИИ ВОЗВРАТНОСТИ

Мы докажем две теоремы, которые окажутся полезными при определении возвратности или невозвратности марковских цепей, а затем применим их к нескольким примерам.

Теорема 4.1. *Пусть $\mathfrak{P}$ — неприводимая марковская цепь, состояния которой отождествлены с неотрицательными целыми числами. Для того чтобы цепь $\mathfrak{P}$ была невозвратна, необходимо и достаточно, чтобы система уравнений*

$$\sum_{j=0}^{\infty} P_{ij} y_j = y_i, \qquad i \neq 0, \tag{4.1}$$

имела ограниченное решение, отличное от $y_i = \text{const}$.

Доказательство. Пусть

$$\mathbf{P} = \| P_{ij} \| = \left\| \begin{matrix} P_{00} & P_{01} & \dots \\ P_{10} & P_{11} & \dots \\ \dots & \dots & \dots \end{matrix} \right\|$$

— матрица переходных вероятностей цепи $\mathfrak{P}$. Сопоставим ей матрицу переходных вероятностей

$$\tilde{\mathbf{P}} = \| \tilde{P}_{ij} \| = \left\| \begin{matrix} 1 & 0 & 0 & \dots \\ P_{10} & P_{11} & P_{12} & \dots \\ P_{20} & P_{21} & P_{22} & \dots \\ \dots & \dots & \dots & \dots \end{matrix} \right\|, \tag{4.2}$$

которая обращает нулевое состояние в поглощающий экран, оставляя вероятности переходов между другими состояниями без изменения. Обозначим марковскую цепь, матрица переходных вероятностей которой имеет вид (4.2), символом $\tilde{\mathfrak{P}}$.

Для доказательства необходимости предположим, что процесс невозвратный, а затем покажем, что в этом случае система (4.1) имеет ограниченное решение, отличное от константы.

Пусть f_{i0}^* есть вероятность рано или поздно попасть в состояние 0, выйдя из состояния i. Поскольку процесс $\mathfrak{P}$ невозвратен, то $f_{j0}^* < 1$ для некоторого $j \neq 0$, так как в противном случае состояние было бы возвратным. (Докажите это! Напомним, что состояния неприводимой марковской цепи либо все одновременно возвратны, либо невозвратны.) Для процесса $\tilde{\mathfrak{P}}$, очевидно, имеем $\tilde{\pi}_0(C_0)=1$, $\tilde{\pi}_j(C_0)=f^*_{j0}<1$ для некоторого $j\neq 0$ и $\tilde{\pi}_i(C_0)=$ $=\sum\limits_{j=0}^{\infty}\tilde{P}_{ij}\tilde{\pi}_j(C_0)$ для всех i. Следовательно, $\tilde{\pi}_i(C_0)=\sum\limits_{j=0}^{\infty}P_{ij}\tilde{\pi}_j(C_0)$ для $i\neq 0$; таким образом, $y_j=\tilde{\pi}_j(C_0)$ $(j=0,\ 1,\ 2,\ \ldots)$ есть искомое ограниченное решение, отличное от константы.

Предположим теперь, что система (4.1) имеет ограниченное отличное от константы решение $\{y_i\}$. Поскольку постоянный вектор [1]) является решением системы (4.1), то $z_i = ay_i + b$ также есть решение этой системы, которое при подходящем выборе a и b будет удовлетворять условиям $z_0 = 1$, $0 \leqslant z_i \leqslant 2$. Поэтому можно предположить, что $0 \leqslant y_i \leqslant 2$ и $y_0 = 1$; в таком случае $\sum\limits_{j=0}^{\infty}\tilde{P}_{ij}y_j=y_i$ для всех $i\geqslant 0$ и, значит, для всех $n\geqslant 1$ имеем

$$\sum_{j=0}^{\infty}\tilde{P}^n_{ij}y_j=y_i,\qquad i\geqslant 0.$$

Поскольку цепь $\mathfrak{P}$ неприводима, $P^n_{i0}>0$ для любого i и некоторого n, поэтому каждое из состояний $j = 1, 2, \ldots$ должно быть невозвратным в цепи $\tilde{\mathfrak{P}}$, так что $\tilde{P}^n_{ij}\to 0$ для $j\neq 0$ и по теореме 3.1 имеем $\tilde{P}^n_{i0}\to\tilde{\pi}_i(C_0)$, где $\tilde{\pi}_i(C_0)$ — вероятность (относительно $\tilde{\mathfrak{P}}$), отправившись из состояния i, быть поглощенным состоянием 0. Следовательно, так как при всех $i\geqslant 0$ выполняется неравенство

$$\tilde{P}^n_{i0}=\tilde{P}^n_{i0}y_0\leqslant\sum_{j=0}^{\infty}\tilde{P}^n_{ij}y_j=y_i,$$

то, устремляя $n\to\infty$, мы приходим к неравенству $\tilde{\pi}_i(C_0)\leqslant y_i$. Возможны два случая: либо существует такое $y_k\,(k\neq 0)$, что $y_k < 1$, либо такое, что $y_k > 1$ $(k\neq 0)$. В первом случае $f^*_{k0}=$ $=\tilde{\pi}_k(C_0)\leqslant y_k<1$, откуда следует, что цепь $\mathfrak{P}$ невозвратна, так как состояние k достижимо из состояния 0 по предположению, в то время как вероятность возвращения меньше, чем 1. Во втором случае эти же рассуждения следует применить к решению $z_i = 2 - y_i$ системы (4.1), что даст неравенство $\tilde{\pi}_k(C_0)\leqslant z_k<1$, откуда опять же следует, что цепь $\mathfrak{P}$ невозвратна.

Теорема 4.2. *Для того чтобы неприводимая марковская цепь была возвратной, достаточно, чтобы существовала последовательность* $\{y_i\}$, *такая, что*

$$\sum_{j=0}^{\infty}P_{ij}y_j\leqslant y_i,\qquad i\neq 0,\qquad y_i\to\infty. \tag{4.3}$$

[1]) Здесь и далее под «постоянным вектором» понимается вектор с одинаковыми компонентами. — *Прим. ред.*

Доказательство. Используя обозначения предыдущей теоремы, имеем

$$\sum_{j=0}^{\infty} \tilde{P}_{ij} y_j \leqslant y_i \quad \text{для всех } i.$$

Поскольку $z_i = y_i + b$ удовлетворяет неравенствам (4.3), мы можем предположить, что $y_i > 0$ для всех $i \geqslant 0$. Из предыдущего неравенства получаем

$$\sum_{j=0}^{\infty} \tilde{P}_{ij}^m y_j \leqslant y_i.$$

Пусть задано $\varepsilon > 0$. Мы можем выбрать такое $M(\varepsilon)$, что $1/y_i \leqslant \varepsilon$ для $i \geqslant M(\varepsilon)$. Далее, имеем

$$\sum_{j=0}^{M-1} \tilde{P}_{ij}^m y_j + \sum_{j=M}^{\infty} \tilde{P}_{ij}^m y_j \leqslant y_i,$$

откуда

$$\sum_{j=0}^{M-1} \tilde{P}_{ij}^m y_j + \min_{r \geqslant m} \{y_r\} \sum_{j=M}^{\infty} \tilde{P}_{ij}^m \leqslant y_i,$$

и поскольку

$$\sum_{j=0}^{\infty} \tilde{P}_{ij}^m = 1,$$

то

$$\sum_{j=0}^{M-1} \tilde{P}_{ij}^m y_j + \min_{r \geqslant M} \{y_r\} \left(1 - \sum_{j=0}^{M-1} \tilde{P}_{ij}^m\right) \leqslant y_i.$$

Как было отмечено в доказательстве предыдущей теоремы,

$$\lim_{n \to \infty} \tilde{P}_{ij}^n = 0 \text{ при } j \neq 0.$$

Переходя к пределу при $m \to \infty$, получаем для каждого фиксированного i

$$\tilde{\pi}_i(C_0) y_0 + \min_{r \geqslant M} \{y_r\} (1 - \tilde{\pi}_i(C_0)) \leqslant y_i,$$

или

$$1 - \tilde{\pi}_i(C_0) \leqslant \frac{1}{\min\limits_{r \geqslant M} \{y_r\}} (y_i - \tilde{\pi}_i(C_0) y_0) \leqslant \varepsilon K,$$

где

$$K = y_i - \tilde{\pi}_i(C_0) y_0.$$

Поскольку ε произвольно, а $\tilde{\pi}_i(C_0) \leqslant 1$, то $\tilde{\pi}_i(C_0) = 1$ для каждого i, что и означает возвратность исходной марковской цепи.

§ 5. ПРИМЕР ИЗ ТЕОРИИ ОЧЕРЕДЕЙ

Рассмотрим модель процесса обслуживания из гл. 2 (пример В). Матрица переходных вероятностей соответствующей марковской цепи имеет вид

$$\| P_{ij} \| = \begin{Vmatrix} a_0 & a_1 & a_2 & a_3 & \ldots \\ a_0 & a_1 & a_2 & a_3 & \ldots \\ 0 & a_0 & a_1 & a_2 & \ldots \\ 0 & 0 & a_0 & a_1 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix}, \quad \text{где } a_k > 0 \text{ и } \sum_{k=0}^{\infty} a_k = 1.$$

(На самом деле при дальнейшем анализе нам потребуются только два условия: $0 < a_0 < 1$ и $a_0 + a_1 < 1$, обеспечивающие неприводимость марковской цепи.) Мы покажем, что если $\sum\limits_{k=0}^{\infty} k a_k > 1$, то система уравнений $\sum\limits_{j=0}^{\infty} P_{ij} y_j = y_i$, $i \neq 0$, имеет ограниченное решение, отличное от константы, что, согласно теореме 4.1, означает невозвратность процесса. Положим $y_j = \xi^j$, тогда упомянутая система уравнений принимает вид

$$\sum_{j=0}^{\infty} P_{ij} \xi^j = \sum_{j=i-1}^{\infty} a_{j-i+1} \xi^j = \xi^i,$$

или

$$\sum_{j=i-1}^{\infty} a_{j-i+1} \xi^{j-i+1} = \xi = \sum_{k=0}^{\infty} a_k \xi^k = f(\xi), \qquad i \neq 0.$$

Так как $f(0) = a_0 > 0$ и $f(1) = \sum\limits_{k=0}^{\infty} a_k = 1$, то из условия $f'(1) = \sum\limits_{k=0}^{\infty} k a_k > 1$ следует, что существует точка ξ_0, $0 < \xi_0 < 1$, такая, что $f(\xi_0) = \xi_0$. Это легко видеть на рис. 1. Вектор $y_i = \xi_0^i$, $i = 0, 1, \ldots$, и представляет собой искомое ограниченное решение, очевидно, не являющееся постоянным.

Пусть теперь $\sum k a_k \leqslant 1$. Тогда, полагая $y_j = j$, имеем

$$\sum_{j=0}^{\infty} P_{ij} j = \sum_{j=i-1}^{\infty} a_{j-i+1} j = \sum_{j=i-1}^{\infty} a_{j-i+1} (j - i + 1) + i - 1 =$$

$$= \sum_{k=0}^{\infty} k a_k - 1 + i \leqslant i \qquad (i \neq 0).$$

Таким образом, в силу теоремы 4.2 процесс является возвратным, если $\sum ka_k \leqslant 1$.

Прежде чем обратиться к вопросу о том, является ли процесс $\mathfrak{P}$ возвратным нулевым или возвратным положительным, рассмотрим следующую вспомогательную задачу, представляющую самостоятельный интерес.

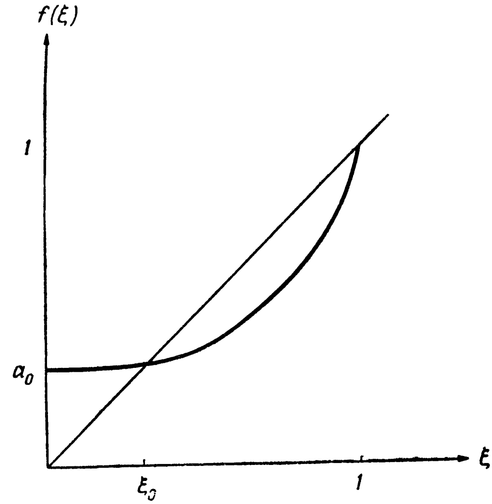


Рис. 1.

Пусть $X_1, X_2, X_3, \ldots$ — последовательность независимых, одинаково распределенных случайных величин, принимающих значения $-1, 0, 1, 2, \ldots$ с вероятностями

$$P\{X_i = k\} = b_k, \qquad k = -1, 0, 1, 2, \ldots, \; b_{-1} > 0,$$

и пусть $S_n = X_1 + X_2 + \ldots + X_n$. Определим Z как значение параметра n, для которого S_n впервые становится отрицательным, и пусть

$$P\{Z = k\} = \gamma_k, \qquad k = 1, 2, 3, \ldots. \tag{5.1}$$

Пусть

$$U(s) = \sum_{k=0}^{\infty} \gamma_k s^k \qquad (\gamma_0 = 0) \tag{5.2}$$

есть производящая функция распределения (5.1). Пусть $T_n^{(r)}$ есть случайная величина, равная первому значению параметра n, для которого $T_n^{(r)} < 0$, где $T_n^{(r)} = r + S_n$ (r — неотрицательное целое число). Поскольку каждая из с. в. $X_i > -1$, нетрудно убедиться в том, что $Z^{(r)} = Z_1 + Z_2 + \ldots + Z_{r+1}$, где с. в. Z_i — независимые и одинаково распределенные в соответствии с (5.1). Производящая функция с. в. $Z^{(r)}$, очевидно, равна $[U(s)]^{r+1}$. Пусть γ_m^{r+1} есть коэффициент при S^m в разложении функции $[U(s)]^{r+1}$. Наконец, положим

$$G(s)=\frac{b_{-1}}{s}+b_0+b_1s+b_2s^2+\ \ldots .$$

Наша цель — выразить $U(s)$ через $G(s)$. Для этого запишем следующие соотношения типа уравнений восстановления:

$$\gamma_1=b_{-1},\quad \gamma_k=\sum_{j=0}^{\infty}b_j\gamma_{k-1}^{(j+1)},\qquad k\geqslant 2. \tag{5.3}$$

Первое из этих соотношений очевидно. Что касается второго, то событие $\{S_n\geqslant 0,\ n=1,\ \ldots,\ k-1;\ S_k=-1\}$ представляет собой объединение следующих несовместных событий: $\{X_1=j;\ X_2+\ldots+X_n+j>0,\ n=2,\ \ldots,\ k-1;\ X_2+\ldots+X_k+j=-1\}$, $j=0,\ 1,\ \ldots$. Так как с. в. X_i независимы и одинаково распределены, вероятности этих событий, как легко видеть, равны $b_j\gamma_{k-1}^{(j+1)}$. Формула полных вероятностей дает (5.3). Переходя к производящим функциям, с помощью (5.3) получаем

$$\begin{aligned}
U(s)&=b_{-1}s+\sum_{n=2}^{\infty}\left(\sum_{j=0}^{\infty}b_j\gamma_{n-1}^{(j+1)}\right)s^n=\\
&=b_{-1}s+s\sum_{j=0}^{\infty}b_j\left(\sum_{n=2}^{\infty}\gamma_{n-1}^{(j+1)}s^{n-1}\right)=\\
&=b_{-1}s+s\sum_{j=0}^{\infty}b_j[U(s)]^{j+1}=\\
&=b_{-1}s+sU(s)\left[G(U(s))-\frac{b_{-1}}{U(s)}\right]=\\
&=sU(s)\,G(U(s)),\qquad 0<s\leqslant 1.
\end{aligned}$$

Далее, $U(s)$ непрерывна и строго возрастает при $s\in[0,\ 1]$, причем $U(0)=0$. Следовательно, $U(s)$ удовлетворяет уравнению $G(U(s))=1/s$ при $0<s\leqslant 1$. Но

$$G''(s)=\frac{2b_{-1}}{s^3}+2b_2+6b_3s+12b_4s^2+\ \ldots>0\ \text{при}\ s>0,$$

так что функция $G(s)$ является выпуклой; к тому же, по определению $G(s)$, $\lim\limits_{s\downarrow 0}G(s)=+\infty$ и $G(1)=1$. Из рис. 2 легко заключить, что уравнение $G(x)=1/s$ может иметь самое большее два положительных решения при каждом фиксированном $s\in[0,\ 1]$.

Поскольку $\lim\limits_{s\downarrow 0} U(s)=0$ и $U(s)$ строго возрастает на интервале $[0, 1]$, то $U(s)$ должно быть меньшим из двух решений уравнения $G(x) = 1/s$, если таковых два.

Исследуем теперь условия, при которых $\sum\limits_{k=0}^{\infty}\gamma_k=1$ или <1. Возникают следующие две возможности:

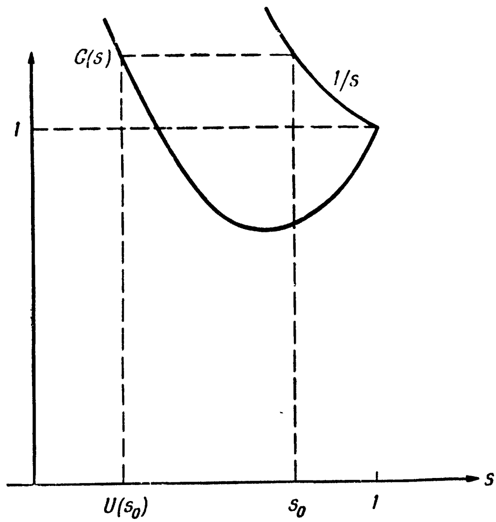


Рис. 2.

Случай 1. $G'(1) > 0$. Условие $G'(1) > 0$ эквивалентно тому, что $b_{-1}<\sum\limits_{n=0}^{\infty} nb_n$. Из рис. 2 видно, что $U(1)=\sum\limits_{k=0}^{\infty}\gamma_k=\xi_0<1$. Следовательно, вероятность события $\{S_n \geqslant 0$ при всех $n\}$ строго положительна.

Случай 2. $G'(1) \leqslant 0$. Условие $G'(1) \leqslant 0$ эквивалентно тому, что $b_{-1}\geqslant\sum\limits_{n=0}^{\infty} nb_n$, и в этом случае мы имеем $\sum\limits_{k=0}^{\infty}\gamma_k=U(1)=1$. Далее, $G'(U(s))\,U'(s)=-1/s^2$ при $0<s\leqslant 1$, так что в рассматриваемом случае $U(s)\to 1$ при $s\to 1$ (рис. 3). Отсюда следует, что если $G'(1) < 0$, т. е. если

$$b_{-1}>\sum_{n=0}^{\infty} nb_n,$$

то

$$\mathrm{M}(Z)=\sum_{n=0}^{\infty} n\gamma_n=U'(1)=\frac{-1}{G'(1)}<\infty,$$

а если $G'(1) = 0$, т. е.

$$b_{-1} = \sum_{n=0}^{\infty} nb_n,$$

то

$$M(Z) = \sum_{n=0}^{\infty} n\gamma_n = U'(1) = \infty.$$

Возвращаясь к процессу обслуживания, поставим в соответствие распределению $\{b_k\}$ распределение $\{a_k\}$ количества посту-

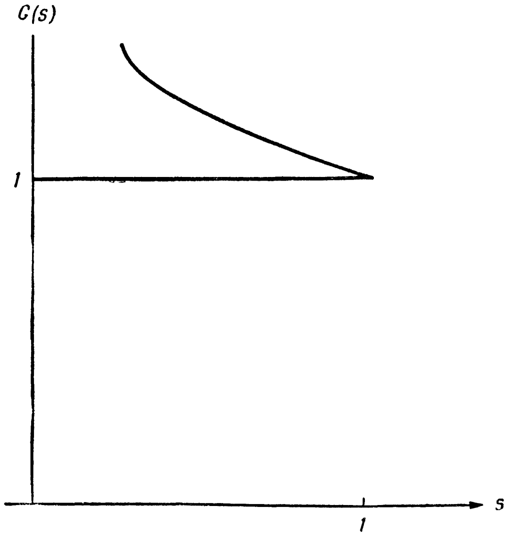


Рис. 3.

пающих заявок за период следующим образом: $a_k = b_{k-1}$. Определим Z_{ij} как число переходов (время), требующееся для того, чтобы впервые попасть в состояние $j < i$ из состояния i. Нетрудно видеть, что $Z_{i,\,i-1}$ есть в точности с. в. Z, производящую функцию $U(s)$ которой мы только что рассматривали. Так как

$$\sum_{i=0}^{\infty} a_i = 1,$$

то мы имеем

$$\left(b_{-1} > \sum_{n=0}^{\infty} nb_n\right) \leftrightarrow \left(a_0 > \sum_{n=0}^{\infty} na_{n+1}\right) \leftrightarrow \left(1 > \sum_{n=0}^{\infty} na_n\right)$$

и, аналогично,

$$\left(b_{-1}=\sum_{n=0}^{\infty} nb_n\right) \leftrightarrow \left(a_0=\sum_{n=0}^{\infty} na_{n+1}\right) \leftrightarrow \left(1=\sum_{n=0}^{\infty} na_n\right).$$

Следовательно, $M(Z_{i,\,i-1})=\mu<\infty$, если $\sum\limits_{n=0}^{\infty} na_n<1$, и $M(Z_{i,\,i-1}) = = \mu=\infty$, если $\sum na_n=1$. Ясно, что

$$Z_{i,\,j}=Z_{i,\,i-1}+Z_{i-1,\,i-2}+\ \dots\ +Z_{j+1,\,j}, \qquad j<i,$$

и поэтому $M(Z_{i,\,j}) = (i-j)\mu$; в частности, $M(Z_{i,\,0}) = i\mu$.

Рассмотрим теперь среднее время возвращения в состояние 0. Отметим прежде всего, что вероятность времени возвращения быть равным 1 есть просто a_0, т. е. вероятность перехода P_{00}. Далее, траектории, которые выходят из состояния 0 и возвращаются в это состояние впервые за два или более переходов, могут быть разбиты на группы в зависимости от состояния i, занимаемого после первого перехода. Такое разложение в совокупности с марковским свойством процесса позволяет получить для среднего времени возвращения следующее выражение:

$$\begin{aligned}\sum_{n=0}^{\infty} nf_{00}^{n} &= \text{M (время возвращения)} = \\ &= a_0+\sum_{i=1}^{\infty} a_i\left[\mathrm{E}\left(Z_{i,\,0}\right)+1\right]=1+\sum_{i=1}^{\infty} a_i \mathrm{E}\left(Z_{i,\,0}\right)= \\ &= 1+\sum_{i=1}^{\infty} i\mu a_i=1+\mu \sum_{i=0}^{\infty} ia_i.\end{aligned}$$

Таким образом,

$$\sum_{n=0}^{\infty} nf_{00}^{n}<\infty, \text{ если } \mu<\infty,$$

т. е. при условии

$$\sum_{i=0}^{\infty} ia_i<1,$$

и

$$\sum_{n=0}^{\infty} nf_{00}^{n}=\infty, \text{ если } \mu=\infty,$$

или, что то же, если

$$\sum_{i=0}^{\infty} ia_i=1.$$

Резюмируя полученные результаты, имеем

$$\sum_{n=0}^{\infty} na_n < 1 \Rightarrow \text{возвратный положительный},$$

$$\sum_{n=0}^{\infty} na_n = 1 \Rightarrow \text{возвратный нулевой} \tag{5.4}$$

и

$$\sum_{n=0}^{\infty} na_n > 1 \Rightarrow \text{невозвратный}.$$

Эти результаты представляются довольно естественными. Выражение $\sum\limits_{n=0}^{\infty} na_n$ есть среднее число требований, прибывающих за один период обслуживания. Тогда если $\sum\limits_{n=0}^{\infty} na_n > 1$, то в среднем больше заявок поступает, чем обслуживается за каждый период. Следовательно, можно ожидать, что очередь будет расти беспредельно. С другой стороны, если $\sum\limits_{n=0}^{\infty} na_n < 1$, то процесс стремится к некоторому стационарному состоянию. Нахождение стационарного распределения связано со значительными трудностями (см. гл. 14).

§ 6. ЕЩЕ ОДИН ПРИМЕР ИЗ ТЕОРИИ ОЧЕРЕДЕЙ

Состоянием процесса, как и ранее, является длина очереди; за каждую единицу времени прибывает одна заявка, а обслуживается k заявок в соответствии с распределением $\{a_k > 0, k = 0, 1, 2, \ldots\}$, если в очереди столько заявок окажется. Матрица переходных вероятностей, как нетрудно убедиться, в этом случае имеет вид

$$\|P_{ij}\| = \begin{Vmatrix} \sum\limits_{i=1}^{\infty} a_i & a_0 & 0 & 0 & \ldots \\ \sum\limits_{i=2}^{\infty} a_i & a_1 & a_0 & 0 & \ldots \\ \sum\limits_{i=3}^{\infty} a_i & a_2 & a_1 & a_0 & \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix}.$$

Мы покажем, что если $\sum_{k=0}^{\infty} ka_k > 1$, то существует стационарное распределение, так что в этом случае процесс возвратный положительный. Так как $\sum ka_k$ есть среднее число обслуживаемых за период заявок, тогда как за это же время поступает только одна заявка, то существование стационарного распределения при указанном условии не является неожиданным.

Рассмотрим уравнения $\sum_{i=0}^{\infty} \xi_i P_{ij} = \xi_j$ и положим $\xi_i = \xi^i$. Тогда $\sum_{i=j-1}^{\infty} \xi^i a_{i-j+1} = \xi^j$ при $j \geqslant 1$, т. е. $\sum_{i=j-1}^{\infty} \xi^{i-j+1} a_{i-j+1} = \xi$, откуда заменой индекса суммирования получаем

$$\sum_{k=0}^{\infty} a_k \xi^k = \xi.$$

Если ξ $(0 < \xi < 1)$ удовлетворяет этому уравнению, то для $j = 0$ получаем

$$\begin{aligned}\sum_{i=0}^{\infty} P_{i0}\xi^i &= \sum_{i=0}^{\infty}\left(\sum_{k=i+1}^{\infty} a_k\right)\xi^i = \\ &= \sum_{k=1}^{\infty}\sum_{i=0}^{k-1} a_k \xi^i = \\ &= \sum_{k=1}^{\infty} a_k\left(\frac{1-\xi^k}{1-\xi}\right) = \frac{1}{1-\xi}\left(1 - a_0 - \sum_{k=1}^{\infty} a_k\xi^k\right) = \\ &= \frac{1}{1-\xi}\left(1 - a_0 - (\xi - a_0)\right) = 1,\end{aligned}$$

т. е. уравнение удовлетворяется и для $j = 0$.

Рассмотрим производящую функцию $f(\xi) = \sum_{k=0}^{\infty} a_k \xi^k$. Так как $f(0) = a_0 > 0$ и $f(1) = 1$, то при условии $f'(1) = \sum_{k=0}^{\infty} ka_k > 1$ существует точка ξ_0, $0 < \xi_0 < 1$, такая, что $f(\xi_0) = \xi_0$ (см. рис. 4). Величины $\pi_i = (1 - \xi_0)\xi_0^i$, $i = 0, 1, 2, \ldots$, сумма которых равна 1, представляют собой стационарное распределение вероятностей исследуемой марковской цепи. В частности, финальная вероятность отсутствия очереди равна $1 - \xi_0$.

Система уравнений $\sum_{i=0}^{\infty} \xi_i P_{ij} = \xi_j$, $j \neq 0$, совпадает с системой $\sum_{j=0}^{\infty} \tilde{P}_{ij}\xi_j = \xi_i$, $i \neq 0$, из предыдущего примера. Как мы видели,

$\tilde{P}$-процесс является возвратным, если $\sum_{k=0}^{\infty} ka_k \leqslant 1$. В этом случае последняя система не имеет ограниченного непостоянного решения. Следовательно, если $\sum ka_k \leqslant 1$, то система $\sum_{j=0}^{\infty} \eta_i P_{ij} = \eta_j$ не имеет ограниченного непостоянного решения, и поэтому,

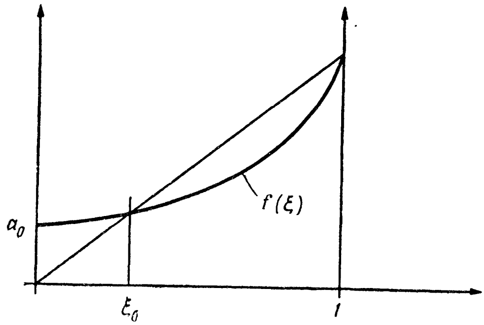


Рис. 4.

в частности, не существует стационарного распределения и процесс является либо возвратным нулевым, либо невозвратным. Мы докажем сейчас, что система

$$\sum_{j=0}^{\infty} P_{ij} y_j = y_i, \qquad i \neq 0, \tag{6.1}$$

имеет непостоянное ограниченное решение тогда и только тогда, когда $\sum ka_k < 1$. Следовательно, процесс является невозвратным тогда и только тогда, когда $\sum ka_k < 1$, и возвратным нулевым, когда $\sum_{k=0}^{\infty} ka_k = 1$. Так как любая последовательность с одинаковыми членами удовлетворяет системе (6.1), мы можем считать, что $y_0 = 0$. Тогда (6.1) сводится к уравнениям

$$\begin{aligned} a_2 y_0 + a_1 y_1 + a_0 y_2 &= y_1, \\ a_3 y_0 + a_2 y_1 + a_1 y_2 + a_0 y_3 &= y_2, \\ &\dots\dots\dots\dots \\ a_{n+1} y_0 + a_n y_1 + \ldots + a_1 y_n + a_0 y_{n+1} &= y_n, \\ &\dots\dots\dots\dots \end{aligned}$$

Умножая i-е уравнение на s^{i+1}, суммируя и пользуясь формулой для преобразования свертки, получаем

$$Y(s) A(s) - s a_0 y_1 = sY(s), \text{ или } Y(s) = \frac{s a_0 y_1}{A(s) - s}, \tag{6.2}$$

при условии, что $A(s) \neq s$. В (6.2) мы положили

$$Y(s) = \sum_{k=0}^{\infty} y_k s^k, \quad A(s) = \sum_{k=0}^{\infty} a_k s^k.$$

Так как $A(0) = a_0$ и $A(1) = 1$, то $A(s) = s$ для некоторого s, такого, что $0 < s < 1$, если $A'(1) = \sum_{k=0}^{\infty} ka_k > 1$. Следовательно, $Y(s)$ не может иметь ограниченных коэффициентов в этом случае, так как это означало бы, что $Y(s)$ сходится для каждого $s \in [0, 1]$. Таким образом, если $\sum_{k=0}^{\infty} ka_k > 1$, то процесс возвратен.

Из строгой выпуклости функции $A(s)$, т. е. из того факта, что $A''(s) > 0$, следует, что $A(s) \neq s$ при $0 \leqslant s < 1$, если $A'(1) = \sum_{k=0}^{\infty} ka_k \leqslant 1$ (см. рис. 5). При условии $\sum ka_k \leqslant 1$ имеем:

$$\begin{aligned} A(s) - s &= (1-s)\left[1 - \frac{1-A(s)}{1-s}\right] = \\ &= (1-s)\left[1 - (A(1) - A(s))\sum_{k=0}^{\infty} s^k\right] = \\ &= (1-s)\left[1 - \sum_{n=0}^{\infty}\left(A(1) - \sum_{i=0}^{n} a_i\right)s^n\right] = \\ &= (1-s)\left[1 - \sum_{n=0}^{\infty}\left(\sum_{i=n+1}^{\infty} a_i\right)s^n\right] = \\ &= (1-s)[1 - W(s)], \quad W(s) = \sum_{n=0}^{\infty} w_n s^n, \end{aligned}$$

где

$$w_n = \sum_{i=n+1}^{\infty} a_i > 0$$

и

$$\sum_{n=0}^{\infty} w_n = \sum_{n=0}^{\infty}\left(\sum_{i=n+1}^{\infty} a_i\right) = \sum_{k=0}^{\infty} ka_k \leqslant 1.$$

Тогда

$$Y(s)=\frac{sa_0y_1}{(1-s)[1-W(s)]}=$$
$$=\frac{sa_0y_1}{1-s}(1+W(s)+(W(s))^2+\ \ldots)=$$
$$=sa_0y_1\frac{U(s)}{1-s},\qquad \text{где } u_n\geqslant 0,\ U(s)=\sum_{n=0}^{\infty}u_ns^n=\sum_{k=0}^{\infty}[W(s)]^k,$$
$$Y(s)=sa_0y_1V(s),\qquad \text{где } v_n=\sum_{k=0}^{n}u_k,\ V(s)=\sum_{n=0}^{\infty}v_ns^n,$$

т. е.

$$V(s)=\frac{U(s)}{1-s}.$$

Далее, следующие условия эквивалентны:

$$\left(W(1)=\sum_{k=0}^{\infty}ka_k<1\right)\longleftrightarrow\left(U(1)=\sum_{k=0}^{\infty}u_k<\infty\right),$$

так как $U(1)=1+W(1)+(W(1))^2+\ \ldots$ есть сходящийся ряд типа геометрической прогрессии. Ясно, что $v_1<v_2<\ \ldots\rightarrow U(1)$,

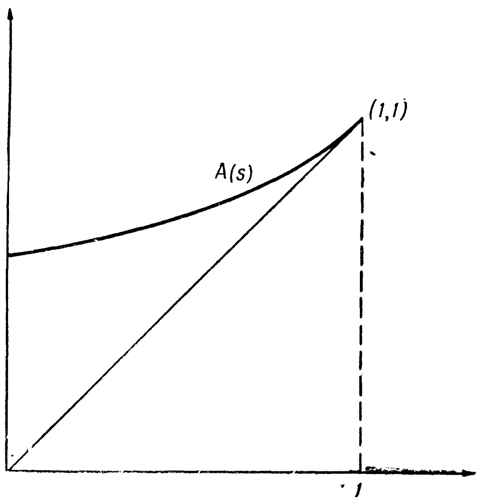


Рис. 5.

так что $Y(s)=sa_0y_1V(s)$, будучи разложенной в степенной ряд, имеет ограниченные коэффициенты тогда и только тогда, когда $\sum ka_k<1$. Поэтому если $\sum ka_k<1$, можно взять $y_1\neq 0$, $y_k=a_0y_1v_{k-1}$ и получить ограниченное непостоянное решение уравнений (6.1), последовательно возвращаясь к уравнению (6.2) и приравнивая коэффициенты. Это означает, что процесс невозвратный. Если $\sum ka_k = 1$, любое решение системы (6.1) с необходимостью неограничено, откуда следует, что в этом случае процесс возвратен. Итак,

$$\begin{aligned} &\text{процесс невозвратный, если } \sum ka_k < 1,\\ &\text{процесс возвратный нулевой, если } \sum ka_k = 1,\\ &\text{процесс возвратный положительный, если } \sum ka_k > 1. \end{aligned}$$

§ 7. СЛУЧАЙНОЕ БЛУЖДАНИЕ

Мы приложим теперь критерии возвратности из § 4 к исследованию процесса случайного блуждания с матрицей переходных вероятностей

$$\| P_{ij} \| = \left\| \begin{matrix} r_0 & p_0 & 0 & 0 & \dots \\ q_1 & r_1 & p_1 & 0 & \dots \\ 0 & q_2 & r_2 & p_2 & \dots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix} \right\|.$$

Пусть

$$\pi_0 = 1, \quad \pi_n = \frac{p_0 p_1 \dots p_{n-1}}{q_1 q_2 \dots q_n}.$$

Для случая, когда $r_i \equiv 0$, было показано (см. пример в § 1), что процесс случайного блуждания обладает стационарным распределением тогда и только тогда, когда $\sum_{n=0}^{\infty} \pi_n < \infty$. Рассмотрим систему уравнений

$$\sum_{j=0}^{\infty} P_{ij} y_j = y_i, \quad i \neq 0,$$

или

$$\begin{gathered} q_1 y_0 + r_1 y_1 + p_1 y_2 = y_1, \\ \cdots\cdots\cdots\cdots \\ q_n y_{n-1} + r_n y_n + p_n y_{n+1} = y_n, \\ \cdots\cdots\cdots\cdots \end{gathered}$$

Легко видеть, что решения этой системы образуют двумерное линейное пространство. Мы можем задать y_0 и y_1 произвольно, и тогда все остальные y_i определяются из системы. Очевидно, $y_i \equiv 1$ является решением. Покажем, что $y_0 = 0$, $y_n = \sum_{i=0}^{n-1} 1/p_i \pi_i$, $n \geqslant 1$, также является решением. Для первого уравнения имеем

$$q_1y_0+r_1y_1+p_1y_2=r_1\left(\frac{1}{p_0}\right)+p_1\left(\frac{1}{p_0}+\frac{q_1}{p_1p_0}\right)=\frac{1}{p_0}=y_1.$$

Проверяя выполнение n-го уравнения, мы должны показать, что

$$q_n\left(\sum_{i=0}^{n-2}\frac{1}{p_i\pi_i}\right)+r_n\sum_{i=0}^{n-1}\frac{1}{p_i\pi_i}+p_n\sum_{i=0}^{n}\frac{1}{p_i\pi_i}=\sum_{i=0}^{n-1}\frac{1}{p_i\pi_i}.$$

Поскольку $p_n+r_n+q_n=1$, достаточно убедиться в том, что

$$q_n\sum_{i=0}^{n-2}\frac{1}{p_i\pi_i}+p_n\sum_{i=0}^{n}\frac{1}{p_i\pi_i}=(p_n+q_n)\sum_{i=0}^{n-1}\frac{1}{p_i\pi_i}.$$

Но левая часть этого равенства есть не что иное, как

$$(q_n+p_n)\sum_{i=0}^{n-1}\frac{1}{p_i\pi_i}-q_n\frac{1}{p_{n-1}\pi_{n-1}}+p_n\frac{1}{p_n\pi_n},$$

$$-q_n\frac{1}{p_{n-1}\pi_{n-1}}=\frac{-1}{(p_{n-1}/q_n)\,\pi_{n-1}}=\frac{-1}{\pi_n}$$

по определению величин π_n. Этим проверка и завершается. Поскольку два решения $y_i\equiv 1$ и $y_n=\sum_{i=1}^{n-1}1/p_i\pi_i$ линейно независимы, общее решение системы $\sum_{j=0}^{\infty}P_{ij}z_j=z_i, i\neq 0$, имеет вид $z_n=\alpha+$ $+\beta y_n$, и ограниченное непостоянное решение существует у этой системы тогда и только тогда, когда ограничены y_n, т. е. когда $\sum_{i=0}^{\infty}1/p_i\pi_i<\infty$. Итак, мы установили, что

процесс возвратный, если $\sum_{i=0}^{\infty}\frac{1}{p_i\pi_i}=\infty$,

процесс возвратный нулевой, если $\sum_{i=0}^{\infty}\frac{1}{p_i\pi_i}=\infty$ и $\sum_{i=0}^{\infty}\pi_i=\infty$,

процесс возвратный положительный, если $\sum_{i=0}^{\infty}\frac{1}{p_i\pi_i}=\infty$ и $\sum_{i=0}^{\infty}\pi_i<\infty$,

процесс невозвратный, если $\sum_{i=0}^{\infty}\frac{1}{p_i\pi_i}<\infty$.

ЗАДАЧИ

1. Рассмотреть процесс случайного блуждания, где

$$P_{i,\,i+1}=p,\ 0<p<1,$$
$$P_{i,\,i-1}=q=1-q \text{ при } i=1,\,2,\,\ldots,\,r-1,$$
$$P_{0,\,0}=P_{r,\,r}=1,$$

и найти $d(k)=\mathrm{M}$ [время до поглощения состояниями 0 или $r\,|$ начальным состоянием является k].

Ответ:

$$d(k)=\begin{cases}\dfrac{k}{q-p}-\dfrac{r}{q-p}\,\dfrac{(1-(q/p)^k)}{1-(q/p)^r}, & \text{если } p\neq\dfrac{1}{2},\\ k(r-k), \text{ если } p=\dfrac{1}{2}.\end{cases}$$

2. Матрица $\mathbf{P}=\|P_{ij}\|_{i,\,j=1}^{\infty}$ называется стохастической, если

(i) $P_{ij}\geqslant 0$ при всех $i,\ j=1,\,2,\,\ldots$

(ii) $\sum\limits_{j=1}^{\infty}P_{ij}=1$ при всех $i=1,\,2,\,\ldots.$

Матрица $\mathbf{P}$ называется дважды стохастической, если помимо условий (i) и (ii) выполняется следующее условие:

$$\sum_{i=1}^{\infty}P_{ij}=1 \text{ при всех } j=1,\,2,\,\ldots.$$

Доказать, что если матрица переходных вероятностей конечной неприводимой марковской цепи дважды стохастическая, то все стационарные вероятности равны между собой.

3. Пусть $\|P_{ij}\|_{i,\,j=1}^{N}$ — матрица переходных вероятностей неприводимой марковской цепи с конечным числом состояний, и пусть $\{\pi_j\}$ — стационарное распределение этой цепи. Пусть, далее, $\varphi(x)$ — выпуклая функция, определенная на положительной полуоси $x\geqslant 0$, $\|P_{ij}^{(m)}\|$ — матрица вероятностей переходов за n шагов и

$$E_m=\sum_{j=1}^{N}\pi_j\varphi\left(P_{jl}^{(m)}\right),\quad l \text{ фиксировано.}$$

Доказать, что E_m является неубывающей функцией аргумента m, т. е. $E_{m+1}\geqslant E_m$ при всех $m\geqslant 1$.

4. Пусть $\mathbf{P}=\|P_{ij}\|$ — матрица переходных вероятностей неприводимой марковской цепи. Доказать, что если матрица $\mathbf{P}$ идемпотентна (т. е. $\mathbf{P}^2=\mathbf{P}$), то $P_{ij}=P_{jj}$ для всех i и j и марковская цепь непериодична.

Указание: Использовать теорему 1.2 для средних $(1/m)\sum\limits_{m=1}^{n}P_{ij}^{m}$.

5. Предположим, что состояние 0 — возвратное положительное, и обозначим через $\{W_n\}$ $(n=1,\,2,\,\ldots)$ последовательные времена возвращения в состояние 0. Очевидно, с. в. $\{W_n\}$ — независимые, одинаково распределенные, с конечным средним. Пусть $F(t)=\sum\limits_{k=1}^{\infty}t^k\mathrm{P}\{W_1=k\}\ (|t|<1)$ — производящая функция их общего распределения. Определим Y_n как момент последнего пребывания в состоянии 0 перед моментом n. Показать, что

$$\sum_{n=0}^{\infty} t^n \sum_{j=0}^{n} x^j \mathrm{P}\{Y_n = j\} = \frac{(1-F(t))}{(1-t)(1-F(xt))}.$$

Указание: Доказать и воспользоваться соотношением $\mathrm{P}\{Y_n = j\} = \mathrm{P}\{W_1 + \ldots \ldots + W_{N_n} = j\} \cdot q_{n-j}$, где $q_i = \mathrm{P}\{W_1 > i\}$ и N_n — число «визитов» в состояние 0 за первые n испытаний.

6. Рассмотрим марковскую цепь $\mathfrak{M}$ с конечным числом состояний и матрицей переходных вероятностей $\mathbf{P} = \| P_{ij} \|_{i,\, j=0}^{N}$, порождающей три класса: $\{0\}$, $\{1, 2, \ldots, N-1\}$ и $\{N\}$. Пусть первый и третий из названных классов поглощающие, а второй — невозвратный, и пусть состояние k принадлежит второму классу. Определим вспомогательный процесс $\tilde{\mathfrak{M}}$, называемый «процессом возвращения», изменив первую и последнюю строки матрицы $\mathbf{P}$ таким образом, что $P_{0k} = P_{Nk} = 1$, оставив все другие строки без изменения. Очевидно, что процесс возвращения неприводим. Доказать, что среднее время u_k до того момента, когда процесс $\mathfrak{M}$ будет поглощен состояниями 0 или N, отправившись из состояния k, равно $1/(\pi_0 + \pi_N) - 1$, где $\pi_0 + \pi_N$ есть стационарная вероятность быть в состоянии 0 или N для процесса $\tilde{\mathfrak{M}}$.

Указание: Воспользоваться соотношением между стационарными вероятностями и средними временами возвращения для состояний.

7. Рассмотрим марковскую цепь с состояниями $0, 1, \ldots, N$ и матрицей переходных вероятностей с элементами

$$P_{ij} = \begin{cases} \mu_i, & j = i-1, \\ \lambda_i, & j = i+1, \\ 1 - \lambda_i - \mu_i, & j = i, \\ 0, & |j-i| > 1, \end{cases} \qquad i, j = 0, 1, \ldots, N.$$

Предположим, что $\mu_0 = \lambda_0 = \mu_N = \lambda_N = 0$, а все остальные μ_i и λ_i положительны. Пусть k — начальное состояние процесса. Определить вероятности поглощения для состояний 0 и N.

Ответ:

P {поглощение в состоянии 0} = 1 — P {поглощение в состоянии N} =

$$= \frac{\sum_{i=k}^{N-1} \rho_i}{\sum_{i=0}^{N-1} \rho_i},$$

где

$$\rho_0 = 1, \quad \rho_i = \frac{\mu_1 \mu_2 \ldots \mu_i}{\lambda_1 \lambda_2 \ldots \lambda_i}.$$

*8. В предыдущем упражнении определить среднее время до поглощения.

Указание: Использовать метод процесса возвращения, описанный в задаче 6. Показать, что система уравнений для стационарных вероятностей $\{\pi_i\}_{i=0}^{N}$ сводится к уравнениям

$$\mu_j \pi_j - \lambda_{j-1} \pi_{j-1} = \pi_0, \qquad 1 \leqslant j \leqslant k,$$
$$\mu_j \pi_j - \lambda_{j-1} \pi_{j-1} = -\pi_N, \qquad k+1 \leqslant j \leqslant N,$$

Ответ: Среднее время до поглощения $=\sum_{i=1}^{N-1}\theta_i$, где

$$\theta_i=\begin{cases}\left(\sum_{j=k}^{N-1}\rho_j\right)\left(\sum_{j=0}^{i-1}\rho_j\right)\Big/\left\{\sum_{j=0}^{N-1}\rho_j\mu_i\rho_{i-1}\right\}, & i=1, 2, \ldots, k,\\ \left(\sum_{j=0}^{k-1}\rho_j\right)\left(\sum_{j=i}^{N-1}\rho_j\right)\Big/\left\{\sum_{j=0}^{N-1}\rho_j\mu_i\rho_{i-1}\right\}, & i=k+1, \ldots, N-1.\end{cases}$$

9. Рассмотрим марковскую цепь с состояниями $0, 1, \ldots, N$ и матрицей переходных вероятностей с элементами

$$P_{ij}=\binom{N}{j}\pi_i^j(1-\pi_i)^{N-j}, \qquad 0\leqslant i, j\leqslant N,$$

где

$$\pi_i=\frac{1-e^{-2ai/N}}{1-e^{-2a}}, \qquad a>0.$$

Отметим, что состояния 0 и N — поглощающие. Проверить, что $\exp(-2aX_t)$ является мартингалом, т. е. $M(\exp(-2aX_{t+1})\mid X_t)=\exp(-2aX_t)$, где X_t — состояние процесса в момент t $(t=0, 1, 2, \ldots)$. Используя это свойство, показать, что вероятность $P_N(k)$ поглощения состоянием N равна

$$P_N(k)=\frac{1-e^{-2ak}}{1-e^{-2aN}},$$

где k — начальное состояние.

Указание: Использовать тот факт, что поглощение одним из состояний 0 или N происходит с достоверностью; доказать соотношение

$$M(\exp(-2aX_0))=M(\exp(-2aX_n))=P_N(k)\exp(-2aN)+(1-P_N(k))$$

и воспользоваться им.

10. Рассмотрим следующий процесс роста конечной популяции (фиксированного размера N), состоящей из индивидуумов двух типов А и а. В моменты времени $t_1<t_2<t_3\ldots$ один индивидуум умирает и заменяется другим одного из возможных двух типов. Если непосредственно перед моментом замены t_n популяция состоит из j индивидуумов типа А и $N-j$ индивидуумов типа а, то вероятность того, что умрет индивидуум типа А, равна $j\mu_1/B_j$ и того, что умрет индивидуум типа а, равна $(N-j)\mu_2/B_j$, где $B_j=\mu_1 j+\mu_2(N-j)$. Логическая основа этой модели такова. Вообще говоря, вероятность смерти в момент t_n равна $\mu_1/(\mu_1+\mu_2)$ для индивидуума типа А и $\mu_2/(\mu_1+\mu_2)$ для индивидуума типа а (μ_1/μ_2 можно интерпретировать как преимущество при отборе типа А над типом а). Принимая во внимание состав популяции, естественно приписать вероятность $\mu_1 j/B_j$ событию, состоящему в том, что будет заменен индивидуум типа А, и вероятность $\mu_2(N-j)/B_j$ — замене индивидуума типа а. Предположим, что характер рождения одинаков для обоих типов: вероятность того, что будет рожден индивидуум типа А, равна j/N, и типа а — $(N-j)/N$. Рассмотреть марковскую цепь $\{X_n\}$, где X_n есть число индивидуумов типа А в момент t_n $(n=1, 2, \ldots)$, с вероятностями перехода

$$P_{j,j-1}=\frac{\mu_1 j(N-j)}{B_jN}, \quad P_{j,j+1}=\frac{\mu_2(N-j)j}{B_jN},$$

$$P_{jj}=1-P_{j,j-1}-P_{j,j+1}, \quad P_{ij}=0, \quad |i-j|>1.$$

Найти вероятность того, что популяция рано или поздно будет состоять только из индивидуумов типа а, при условии, что в начале популяция состояла из k индивидуумов типа А и $(N-k)$ индивидуумов типа а.

Указание: Показать, что уравнения, определяющие вероятности поглощения, могут быть сведены к соответствующей системе уравнений для вероятностей поглощения в задаче о разорении игрока, где используется представление процессом случайного блуждания.

Ответ:

$$\mathrm{P}\{\text{популяция состоит только из индивидуумов типа а}\} = \begin{cases} \dfrac{(\mu_1/\mu_2)^N - (\mu_1/\mu_2)^k}{(\mu_1/\mu_2)^N - 1}, & \mu_1 \neq \mu_2, \\ 1 - k/N, & \mu_1 = \mu_2. \end{cases}$$

*11. Пусть A — конечная марковская цепь и $\mathfrak{L}$ — множество всех возможных предельных матриц для подпоследовательностей последовательности $\{A^k,\ k = 1, 2, \ldots\}$. Доказать, что $\mathfrak{L}$ обладает следующим свойством: если $\Gamma_1, \Gamma_2 \in \mathfrak{L}$, то $\Gamma_1\Gamma_2 \in \mathfrak{L}$ и $\Gamma_1\Gamma_2^{-1} \in \mathfrak{L}$ (при условии, что Γ_2^{-1} существует).

12. Пусть $\mathbf{P}$ — марковская матрица третьего порядка и $\mu(\mathbf{P}) = \max\limits_{i_1,\, i_2,\, j} [P_{i_1,\, j} - P_{i_2,\, j}]$. Показать, что $\mu(\mathbf{P}) = 1$ тогда и только тогда, когда $\mathbf{P}$ имеет вид

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & p & q \\ r & s & t \end{Vmatrix} \quad (p,\ q \geqslant 0,\ p+q=1;\ r,\ s,\ t \geqslant 0,\ r+s+t=1)$$

либо получена из этой матрицы перестановкой местами строк и/или столбцов.

*13. Пусть $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_k$ — матрицы переходных вероятностей неприводимых непериодических марковских цепей, каждая из которых имеет по три состояния, и пусть $\mu(\mathbf{P}) = \max\limits_{i_1,\, i_2,\, j} (P_{i_1,\, j} - P_{i_2,\, j})$. Предположим, что для каждого набора целых чисел $\alpha_i\ (1 \leqslant \alpha_i \leqslant k)$, $i = 1, 2, \ldots$, матрица $\prod\limits_{i=1}^{m} \mathbf{P}_{\alpha_i}$ также представляет собой матрицу переходных вероятностей неприводимой непериодической марковской цепи. Доказать, что для любого $\varepsilon > 0$ существует $M(\varepsilon)$ такое, что при $m > M$

$$\mu\left(\prod_{i=1}^{m} \mathbf{P}_{\alpha_i}\right) < \varepsilon \text{ для любого набора } \alpha_i\ (1 \leqslant \alpha_i \leqslant k),\ i = 1, 2, \ldots, m.$$

Указание: Воспользоваться интерпретацией матриц $\mathbf{P}_\alpha$ как линейных отображений симплекса Δ, задаваемого условиями $x_1 + x_2 + x_3 = 1$, $1 \geqslant x_i \geqslant 0$, в себя. (Показать, что такая интерпретация имеет смысл.) Показать, что

(а) $\mu(\mathbf{AB}) \leqslant \mu(\mathbf{A})\mu(\mathbf{B})$ для любой пары стохастических матриц $\mathbf{A}$ и $\mathbf{B}$;

(б) если $\mathbf{A}$ — стохастическая матрица, то $\mu(\mathbf{A}) = 1$ тогда и только тогда, когда $\mathbf{A}$ преобразует некоторую вершину симплекса Δ в одну из других вершин, а вторую вершину в точку на ребре, противоположном образу первой вершины, или когда **A** преобразует одну из вершин в вершину, а противоположное ей ребро в ребро (или на ребро), противоположное образу этой вершины. Показать, что последний случай невозможен в условиях теоремы;

(в) показать, что $\mu(\mathbf{P}_{\alpha_1}\mathbf{P}_{\alpha_2}\mathbf{P}_{\alpha_3}) < 1$ для любых α_1, α_2, α_3, для чего убедиться в том, что если это неравенство не выполняется, то ребро отображается на себя произведением матриц $\mathbf{P}_\alpha$, что противоречит исходному предположению.

14. Если i — возвратное состояние, а X_k представляет состояние марковской цепи в момент k, то показать, что

$$\lim_{N\to\infty} \mathrm{P}\{X_k \neq i \quad \text{при} \quad n+1 \leqslant k \leqslant n+N \mid X_0 = i\} = 0.$$

Показать, что если i — возвратное положительное состояние, то стремление к пределу равномерно по n.

15. Обобщенная урновая схема Пойа. Из урны, содержащей a белых и b черных шаров, случайно извлекается шар. Если шар оказывается белым, то его возвращают, добавив еще α белых и β черных шаров, а если извлеченный шар оказывается черным, его возвращают, добавив γ белых и δ черных шаров, причем $\alpha + \beta = \gamma + \delta$. Процесс многократно повторяется. Пусть X_n — число извлечений белого шара после первых n испытаний.

(i) Если $P_{n,k} = \mathrm{P}\{X_n = k\}$ и $\varphi_n(x) = \sum_{k=0}^{n} P_{n,k}x^k$, то показать, что имеет место соотношение

$$\varphi_n(x) = \frac{(\alpha-\gamma)(x^2-x)}{(n-1)(\alpha+\beta)+a+b}\varphi'_{n-1}(x) + \\ + \frac{\{x[(n-1)\gamma+a]+b+(n-1)\delta\}}{(n-1)(\alpha+\beta)+a+b}\varphi_{n-1}(x).$$

(ii) Показать, что $\mathrm{M}(X_n/n) \to \gamma/(\beta+\gamma)$ при $n \to \infty$.

Указание: Показать, что

$$\varphi'_n(1) = (\alpha-\gamma)\sum_{k=1}^{n} \frac{\varphi'_{k-1}(1)}{(k-1)(\alpha+\beta)+a+b} + \sum_{k=1}^{n} \frac{a+(k-1)\gamma}{(k-1)(\alpha+\beta)+a+b},$$

и отсюда вывести предельное соотношение $\varphi'_n(1)/n \to \gamma/(\beta+\gamma)$.

16. В условиях предыдущей задачи показать, что

$$\lim_{n\to\infty} \mathrm{M}\left[\left(\frac{X_n}{n}\right)^2\right] = \left(\frac{\gamma}{\beta+\gamma}\right)^2 \quad \text{при } n\to\infty.$$

Указание: Найти рекуррентное соотношение для $\varphi''_n(1)$, подобно тому как это делается в (ii).

17. При тех же условиях показать, что $X_n/n \to \gamma/(\beta+\gamma)$ по вероятности при $n \to \infty$.

НЕКОТОРЫЕ ЭЛЕМЕНТАРНЫЕ ЗАДАЧИ

1. Марковская цепь на состояниях $\{0, 1, 2, 3, 4, 5\}$ имеет матрицу переходных вероятностей вида

$$
\text{(a)} \begin{Vmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{5} & \frac{4}{5} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{Vmatrix}, \quad \text{(b)} \begin{Vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{7}{8} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{8} & \frac{3}{8} & 0 \\ \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.
$$

Выделить классы состояний и найти предельные вероятности $\lim\limits_{n \to \infty} P^n_{5,\,i}$ для $i = 0, 1, 2, 3, 4, 5$.

2. Рассмотреть задачу о разорении с начальными капиталами a и b ($a > 10$, $b > 10$) у игрока I и игрока II соответственно. Пусть $p(1-p)$ есть вероятность игроку I выиграть (проиграть) единицу у игрока II в каждой партии. Какова вероятность того, что размер капитала игрока I достигнет величины $a+b-3$ раньше, чем уменьшится до 5.

ЗАМЕЧАНИЯ

Содержание § 1—4 является стандартным аппаратом марковских цепей и имеется в большинстве руководств по этому предмету.

Примеры из § 5 являются классическими для теории очередей. Последовательное изложение теории читатель найдет, например, в книге Такача [1].

ЛИТЕРАТУРА

1. Takács L., Introduction to the Theory of Queues, Oxford Univ. Press, London and New York, 1962.

Глава 4

АЛГЕБРАИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЯ МАРКОВСКИХ ЦЕПЕЙ

Алгебраические методы либо их сочетание с вероятностными позволяют получить многие важные результаты теории марковских цепей. Мы остановимся на ряде таких методов в настоящей главе. Для того чтобы не уходить далеко от основного предмета, мы дадим лишь краткую сводку основных результатов из теории матриц, которые потребуются нам в дальнейшем. Более полное изложение этих результатов читатель найдет в приложении.

§ 1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

При рассмотрении марковских цепей вычисление вероятностей перехода за n шагов занимает очень важное место. Для этой цели мы разовьем специальный аппарат, основывающийся на теории собственных значений и собственных векторов[1]). (Специальные методы для исследования марковских цепей, описывающих процессы случайного блуждания, будут изложены в § 4—6.)

(а) Спектральное представление

Пусть $\mathbf{A}$ — квадратная матрица порядка n. Ненулевой вектор $\mathbf{x}$, удовлетворяющий соотношению $\mathbf{Ax} = \lambda\mathbf{x}$ для некоторого комплексного числа λ, называется правым собственным вектором матрицы $\mathbf{A}$, принадлежащим (или соответствующим) собственному значению λ. Если $\mathbf{xA} = \lambda\mathbf{x}$, то мы назовем вектор $\mathbf{x}$ левым собственным вектором матрицы $\mathbf{A}$. Если правые (или же левые) собственные векторы образуют базис в n-мерном линейном пространстве, то существуют линейно независимое семейство $\boldsymbol{\varphi}^{(1)}, \ldots, \boldsymbol{\varphi}^{(n)}$ правых собственных векторов и линейно независимое семейство $\boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(n)}$ левых собственных векторов матрицы $\mathbf{A}$, являющиеся биортогональными системами. Это значит, что

$$(\boldsymbol{\varphi}^{(i)}, \boldsymbol{\psi}^{(j)}) \equiv \sum_{k=1}^{n} \varphi_{ik}\bar{\psi}_{jk} = \delta_{ij} = \begin{cases} 0, & \text{если } i \neq j, \\ 1, & \text{если } i = j, \end{cases}$$

---

[1]) Читателю, не знакомому с основами теории собственных значений и собственных векторов матриц, рекомендуем обратиться к приложению.

где $\varphi^{(i)} = (\varphi_{i1}, \ldots, \varphi_{in})$, $\psi^{(j)} = (\psi_{j1}, \ldots, \psi_{jn})$ и $\bar{\psi}_{jh}$ — число, комплексно сопряженное числу ψ_{jh}. В этом случае говорят, что матрица **A** диагонализируемая. Пусть

$$\mathbf{\Phi} = \begin{Vmatrix} \varphi_{11} & \cdots & \varphi_{n1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \varphi_{1n} & \cdots & \varphi_{nn} \end{Vmatrix}, \quad \mathbf{\Psi} = \begin{Vmatrix} \psi_{11} & \cdots & \psi_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \psi_{n1} & \cdots & \psi_{nn} \end{Vmatrix}, \quad \mathbf{\Lambda} = \begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{Vmatrix},$$

где $\lambda_1, \ldots, \lambda_n$ — собственные числа (не обязательно различные) матрицы **A**, а $\varphi^{(1)}, \ldots, \varphi^{(n)}$ — соответствующие им собственные векторы. (Отметим, что нумерация элементов матрицы **Φ** отличается от обычной.) Тогда матрица **A** обладает представлением, называемым спектральным, в виде произведения трех специальных матриц:

$$\mathbf{A} = \mathbf{\Phi\Lambda\Psi}.$$

Используя соотношение $(\varphi^{(i)}, \psi^{(j)}) = \delta_{ij}$, можно непосредственно убедиться, что $\mathbf{\Psi\Phi} = \mathbf{\Phi\Psi} = \mathbf{I}$ (**I** — единичная матрица). Отсюда $\mathbf{A}^2 = \mathbf{\Phi\Lambda\Psi\Phi\Lambda\Psi} = \mathbf{\Phi\Lambda}^2\mathbf{\Psi}$, и вообще

$$\mathbf{A}^m = \mathbf{\Phi\Lambda}^m\mathbf{\Psi}, \tag{1.1}$$

где, очевидно,

$$\mathbf{\Lambda}^m = \begin{Vmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n^m \end{Vmatrix}.$$

В случае если **A** — марковская матрица, формула (1.1) дает удобное представление матрицы m-шаговых переходных вероятностей. Правда, его эффективное использование требует вычисления всех левых и правых собственных векторов.

(б) Положительные матрицы

Пусть **A** — действительная матрица, которая имеет по крайней мере один положительный элемент, но не имеет ни одного отрицательного элемента. В этом случае мы будем писать $\mathbf{A} > \mathbf{0}$ и называть матрицу **A** положительной. Если все элементы матрицы **A** положительны, то мы будем писать $\mathbf{A} \gg \mathbf{0}$ и называть матрицу **A** строго положительной. Нам понадобятся следующие результаты (их доказательства вынесены в приложение).

Каждой матрице $\mathbf{A} > \mathbf{0}$ соответствует число $r(\mathbf{A}) \geqslant 0$, называемое спектральным радиусом матрицы **A**, которое равно нулю тогда и только тогда, когда $\mathbf{A}^m = \mathbf{0}$ для некоторого целого числа

$m > 0$. Если $\mathbf{A} > \mathbf{0}$, то существуют положительные векторы $\mathbf{f}$, $\mathbf{x} > \mathbf{0}$, такие, что $\mathbf{Ax} = r(\mathbf{A})\mathbf{x}$, $\mathbf{fA} = r(\mathbf{A})\mathbf{f}$. Если λ — любое собственное значение матрицы $\mathbf{A}$, то $|\lambda| \leqslant r(\mathbf{A})$; если $|\lambda| = r(\mathbf{A})$, то $\eta = \lambda/r(\mathbf{A})$ есть корень из единицы, т. е. $\eta^k = 1$ для некоторого положительного целого числа k, и $\eta^m r(\mathbf{A})$ является собственным значением матрицы $\mathbf{A}$ при $m = 1, 2, \ldots$. Наконец, предположим, что $\mathbf{A}^m \gg \mathbf{0}$ для некоторого $m > 0$; тогда векторы $\mathbf{x}$ и $\mathbf{f}$ строго положительны и определены однозначно с точностью до постоянного скалярного множителя. Более того, если λ — собственное значение матрицы $\mathbf{A}$, отличное от $r(\mathbf{A})$, то $|\lambda| < r(\mathbf{A})$.

§ 2. СВЯЗЬ МЕЖДУ СОБСТВЕННЫМИ ЗНАЧЕНИЯМИ И КЛАССАМИ ВОЗВРАТНЫХ СОСТОЯНИЙ

Предыдущие результаты имеют непосредственные приложения в теории марковских цепей с конечным числом состояний.

Пусть $\mathbf{P} = \|P_{ij}\|$, $i, j = 1, \ldots, n$, — матрица переходных вероятностей. Очевидно, $\mathbf{P} > \mathbf{0}$. Пусть $\mathbf{x}$ — произвольный вектор, удовлетворяющий условию $\sum_{i=1}^{n} x_i = 1$. Тогда

$$\mathbf{xP} = \left(\sum_{i=1}^{n} x_i P_{i1}, \ \sum_{i=1}^{n} x_i P_{i2}, \ \ldots, \ \sum_{i=1}^{n} x_i P_{in}\right).$$

Далее,

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{n} x_i P_{ij}\right) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} P_{ij} = \sum_{j=1}^{n} x_i = 1. \tag{2.1}$$

Мы утверждаем, что неравенство $\mathbf{xP} \geqslant \lambda\mathbf{x}$ не может выполняться для вектора $\mathbf{x} > \mathbf{0}$ при любом $\lambda > 1$ (откуда следует, что $r(\mathbf{P}) \leqslant 1$). В самом деле, суммируя компоненты в обеих частях соотношения $\mathbf{xP} \geqslant \lambda\mathbf{x}$ и учитывая (2.1), получаем $\sum_{i=1}^{n} x_i \geqslant \lambda \sum_{i=1}^{n} x_i$. Так как $\sum_{i=1}^{n} x_i > 0$, то отсюда следует, что $\lambda \leqslant 1$.

С другой стороны, вектор $(1, \ldots, 1)$, как легко видеть, является собственным вектором матрицы $\mathbf{P}$, принадлежащим собственному значению 1. Таким образом, $r(\mathbf{P}) = 1$.

То, что 1 — собственное значение любой конечной марковской матрицы и ему соответствует положительный левый собственный вектор, можно также вывести из теоремы 1.3 гл. 3. Действительно, мы знаем, что в конечной марковской цепи по крайней мере одно состояние (а следовательно, по крайней мере один класс) является возвратным положительным. Перенумеровав состояния, если это необходимо, мы можем считать, что состояния $i = 1, \ldots, r$ составляют возвратный положительный класс. Следовательно, $P_{ij}=0$ для всякой пары состояний i, j, такой, что $i \in \{1, \ldots, r\}$, а $j \in \{r+1, \ldots, n\}$. Таким образом, $\mathbf{P}$ имеет вид

$$\mathbf{P}=\begin{Vmatrix} \mathbf{P}_1 & \mathbf{0} \\ \mathbf{B} & \mathbf{C} \end{Vmatrix}, \tag{2.2}$$

где $\mathbf{P}_1$ — марковская матрица порядка $r \times r$. В силу основной предельной теоремы для марковских цепей (см. теорему 1.3 гл. 3) существуют положительные числа $\pi_1, \ldots, \pi_r$, такие, что

$$\sum_{i=1}^{r} \pi_i P_{ij} = \pi_j, \qquad j=1, \ldots, r,$$

и

$$\sum_{i=1}^{r} \pi_i = 1.$$

Пусть $\mathbf{x}^0=(\pi_1, \ldots, \pi_r, 0, \ldots, 0)$; опираясь на специальную структуру (2.2) матрицы $\mathbf{P}$, легко проверить, что $\mathbf{x}^0\mathbf{P}=\mathbf{x}^0$. Несколько более подробный анализ показывает, что справедлива следующая

Теорема 2.1. *Если $\mathbf{P}$ — матрица переходных вероятностей марковской цепи с конечным числом состояний, то кратность собственного значения* 1 *равняется числу возвратных классов в цепи.*

Доказательство. Мы видели, что если C_1 — возвратный класс состояний, то существует левый собственный вектор $\mathbf{x}^{(1)}>0$, отвечающий собственному значению 1, такой, что $x_i^{(1)}=0$ при $i \notin C_1$. Точно так же для каждого возвратного класса C_h, $h = 2, 3, \ldots$, существует положительный собственный вектор $\mathbf{x}^{(h)}$, $h=2, 3, \ldots$, отвечающий собственному значению 1, такой, что $x_i^{(h)}=0$ при $i \notin C_h$. Так как различные классы не пересекаются, то векторы $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$ линейно независимы. Отсюда следует, что кратность собственного значения 1 не меньше, чем число различных возвратных классов. Покажем теперь, что она и не больше этого числа. Пусть $\mathbf{xP}=\mathbf{x}$, тогда $\mathbf{xP}^m=\mathbf{x}$ при $m=1, 2, \ldots$, т. е.

$$\sum_{i=1}^{n} x_i P_{ij}^{m} = x_j, \qquad j=1, \ldots, n; \quad m=1, 2, \ldots .$$

Но если j — невозвратное состояние, то, как мы знаем, $\lim\limits_{m\to\infty} P_{ij}^{(m)}=0$ для всех i. Отсюда следует, что $x_j=0$ для каждого невозвратного состояния j, так что мы можем написать

$$\sum_{h=1}^{r} \sum_{i \in C_h} x_i P_{ij} = x_j, \qquad j \in \bigcup_{h=1}^{r} C_h,$$

где $C_1, \ldots, C_r$ — возвратные классы. Так как $P_{ij}=0$, если i и j принадлежат разным возвратным классам, то

$$\sum_{i \in C_h} x_i P_{ij} = x_j, \qquad j \in C_h, \quad h = 1, \ldots, r.$$

Если $x_i \neq 0$ для некоторого $i \in C_h$, то по теореме 1.3 гл. 3 существует константа a_h, такая, что

$$x_i = a_h x_i^{(h)}, \qquad i \in C_h.$$

Таким образом,

$$\mathbf{x} = \sum_{h=1}^{r} a_h \mathbf{x}^{(h)},$$

откуда заключаем, что векторы $\mathbf{x}^{(h)}$ образуют базис в подпространстве левых собственных векторов, соответствующих собственному значению 1.

Вероятностная интерпретация собственных значений и собственных векторов

Рассмотрим подпространство правых собственных векторов матрицы $\mathbf{P}$, отвечающих собственному значению 1. Оказывается, в этом подпространстве существует базис, имеющий очень простую вероятностную интерпретацию.

Пусть $C_1, \ldots, C_r$ — возвратные классы марковской цепи с матрицей переходных вероятностей $\mathbf{P}$. Определим $p_i^{(h)}$ как вероятность рано или поздно попасть в класс C_h из состояния i, т. е.

$$p_i^{(h)} = \mathrm{P}\left\{X_n \in C_h \text{ для некоторого } n \,\middle|\, X_0 = i\right\}.$$

Ясно, что

$$p_i^{(h)} = \begin{cases} 1 & \text{при} \quad i \in C_h, \\ 0 & \text{при} \quad i \in C_j, \quad j \neq h, \end{cases} \tag{2.3}$$

так как возвратный класс покинуть невозможно. Пусть $\mathbf{p}^{(h)} = \left(p_1^{(h)}, \ldots, p_n^{(h)}\right)$, $h = 1, \ldots, r$, тогда из (2.3) следует, что векторы $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(r)}$ линейно независимы. Кроме того, вероятности $p_i^{(h)}$ удовлетворяют уравнениям

$$p_i^{(h)} = \sum_{j=1}^{n} P_{ij} p_j^{(h)}, \qquad i = 1, \ldots, n; \quad h = 1, \ldots, r$$

(см. уравнения (3.1) гл. 3), из которых следует, что $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(r)}$ являются правыми собственными векторами матрицы $\mathbf{P}$, соответствующими собственному значению 1. Так как этих векторов r и они линейно независимы, то в правом собственном подпространстве, отвечающем собственному значению 1, векторы $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(r)}$ образуют базис. Непосредственной проверкой нетрудно убедиться в том, что

$$(\mathbf{x}^{(i)}, \mathbf{p}^{(j)}) = \begin{cases} 1, & \text{если } i=j, \\ 0, & \text{если } i \neq j, \end{cases}$$

где $\mathbf{x}^{(i)}$ — левые собственные векторы, отвечающие собственному значению 1 (см. доказательство предыдущей теоремы).

Предположим теперь, что матрица $\mathbf{P}$ обладает спектральным представлением и что собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_m$ занумерованы так, что $1 = \lambda_1 = \ldots = \lambda_r \geqslant |\lambda_{r+1}| \geqslant \ldots$ и $\lambda_{r+1} \neq 1$. Мы можем положить $\boldsymbol{\varphi}^{(1)} = \mathbf{p}^{(1)}, \ldots, \boldsymbol{\varphi}^{(r)} = \mathbf{p}^{(r)}$ и $\boldsymbol{\psi}^{(1)} = \mathbf{x}^{(1)}, \ldots, \boldsymbol{\psi}^{(r)} = \mathbf{x}^{(r)}$ (см. приложение). Из представления

$$\mathbf{P}^m = \boldsymbol{\Phi}\boldsymbol{\Lambda}^m\boldsymbol{\Psi}$$

получаем

$$P_{ij}^m = \sum_{h=1}^{n} \varphi_{hi}\lambda_h^m\psi_{hj} = \varphi_{1i}\psi_{1j} + \ldots + \varphi_{ri}\psi_{rj} + \sum_{h=r+1}^{n} \varphi_{hi}\lambda_h^m\psi_{hj}.$$

Предположим, что матрица $\mathbf{P}$ не имеет собственных значений, отличных от 1, модуль которых был бы равен 1. Тогда $|\lambda_h| < 1$, $h = r+1, \ldots, n$, и при $m \to \infty$

$$\sum_{h=r+1}^{n} \varphi_{hi}\lambda_h^m\psi_{hj} \to 0,$$

причем скорость сходимости имеет порядок самое меньшее $|\lambda_{r+1}|^m$. Скоро будет показано, что $|\lambda_h| < 1$, $h = r+1, \ldots, n$, тогда и только тогда, когда $\mathbf{P}$ не имеет периодических возвратных классов (см. теорему 3.1 следующего параграфа). Предполагая, что $\mathbf{P}$ не имеет таких классов, и вспоминая, что $x_j^{(h)} = \psi_{hj}$, $h = 1, 2, \ldots, r$, $j = 1, \ldots, n$, отлично от нуля тогда и только тогда, когда $j \in C_h$, мы видим, что

$$\varphi_{1i}\psi_{1j} = \varphi_{2i}\psi_{2j} = \ldots = \varphi_{ri}\psi_{rj} = 0 \text{ для невозвратных состояний } j.$$

Таким образом, если состояние j невозвратно, то $P_{ij}^m = \sum_{h=r+1}^{n} \varphi_{hi}\lambda_h^m\psi_{hj}$ и эта величина стремится к нулю со скоростью $|\lambda_{r+1}|^m$ при $m \to \infty$. Если же $j, i \in C_h$, то среди первых r слагаемых в выражении для P_{ij}^m отлично от нуля только $\varphi_{hi}\psi_{hj}$; но $\varphi_{hi} = 1$ (вспомним, что $\varphi_{hi} = p_i^{(h)}$) и $\psi_{hj} = \pi_j = \lim_{m\to\infty} P_{ij}^m$. Мы видим, что вообще для всех состояний j разность $\pi_j - P_{ij}^m$ стремится к нулю при $m \to \infty$ со скоростью самое меньшее $|\lambda_{r+1}|^m$.

Теперь предположим, что, кроме $|\lambda_{r+1}| < 1$, имеет место еще и неравенство $|\lambda_{r+2}| < |\lambda_{r+1}|$. Пусть, как обычно, T обозначает множество всех невозвратных состояний, $i, j \in T$; мы хотим найти следующий предел:

$$\lim_{m \to \infty} \mathrm{P}\{X_m = j \mid X_0 = i, \ X_m \in T\},$$

т. е. предельное значение (при $m \to \infty$) вероятности того, что, исходя из $i \in T$, процесс будет находиться в невозвратном состоянии j в момент m. Имеем

$$\mathrm{P}\{X_m = j \mid X_0 = i, \ X_m \in T\} = \frac{P_{ij}^m}{\sum\limits_{j \in T} P_{ij}^m}.$$

Как мы видели, для невозвратного состояния j вероятность $P_{ij}^m = \sum\limits_{h=r+1}^{n} \varphi_{hi} \lambda_h^m \psi_{hj}$. Так как $|\lambda_{r+1}| > |\lambda_{r+2}|$, то

$$\lim_{m \to \infty} \frac{P_{ij}^m}{\sum\limits_{j \in T} P_{ij}^m} = \frac{\varphi_{r+1,\,i} \psi_{r+1,\,j}}{\sum\limits_{j \in T} \varphi_{r+1,\,i} \psi_{r+1,\,j}} = \frac{\psi_{r+1,\,j}}{\sum\limits_{j \in T} \psi_{r+1,\,j}}$$

при условии, что знаменатель не равен нулю. Если же знаменатель равен нулю, то нам придется исследовать члены в сумме $\sum\limits_{k=r+1}^{n} \varphi_{hi} \lambda_h^m \psi_{hj}$, содержащие λ_{r+2} и другие собственные значения, модуль которых равен $|\lambda_{r+2}|$, и т. д. (См. § 8 гл. 13, где эти идеи развиваются далее.)

§ 3. ПЕРИОДИЧЕСКИЕ КЛАССЫ

В этом параграфе мы дадим более полное описание структуры периодической цепи. Простейший пример периодического класса с периодом d дает цепь с d состояниями $1, \ldots, d$ и матрицей переходных вероятностей

$$P_{12} = P_{23} = \ldots = P_{d-1,\,d} = P_{d1} = 1, \quad \mathbf{P} = \begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 1 & 0 & 0 & \ldots & 0 \end{Vmatrix}.$$

Менее тривиальный пример возникает, если каждое из состояний $1, \ldots, d$ заменить группой состояний, причем группы состояний $C_1, \ldots, C_d$ не пересекаются, а P_{ij} определить таким образом, что

$P_{ij} \neq 0$ только тогда, когда $i \in C_1$, $j \in C_2$, или $i \in C_2$, $j \in C_3$, ..., или $i \in C_d$, $j \in C_1$. Матрица $\mathbf{P}$ принимает в этом случае вид

$$\mathbf{P} = \left\| \begin{matrix} \mathbf{0} & \mathbf{P}_1 & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_2 & \dots \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \mathbf{P}_d & \mathbf{0} & \mathbf{0} & \dots \end{matrix} \right\|.$$

При этом мы можем определить P_{ij} так, чтобы любые два состояния сообщались. Мы покажем сейчас, что всякий периодический класс имеет такую структуру. Пусть d — период класса состояний W, которые помечены целыми числами $1, \dots, M$. Обозначим через C_1 множество всех состояний, достижимых из состояния 1 за какое-либо число шагов, кратное d, т. е. $j \in C_1$ тогда и только тогда, когда $P_{1j}^{nd} > 0$ для некоторого целого $n > 0$. Далее, для каждого $r = 1, \dots, d-1$ определим C_{r+1} как множество тех состояний, которые могут быть достигнуты из состояния 1 за r плюс кратное d число шагов, т. е. $j \in C_{r+1}$ тогда и только тогда, когда $P_{1j}^{nd+r} > 0$ для некоторого целого $n > 0$.

Покажем сначала, что если $j \in C_1$, то из $P_{j1}^h > 0$ следует, что $h = md$ для некоторого целого $m > 0$. Действительно, так как $j \in C_1$, то $P_{1j}^{nd} > 0$ для некоторого $n > 0$. Следовательно, так как $P_{jj}^{nd+h} \geqslant P_{j1}^h P_{1j}^{nd} > 0$ (см. следствие 4.1 гл. 2), то по определению периода $nd + h$ должно делиться на d, а потому и h делится на d. Теперь мы покажем, что если $i \in C_1$ и $j \in C_{r+1}$, то из $P_{ij}^h > 0$ следует, что $h = nd + r$ для некоторого целого $n \geqslant 0$. В самом деле, пусть $P_{ji}^s > 0$ для некоторого $s > 0$, $P_{i1}^{qd} > 0$ для некоторого $q > 0$ и $P_{1j}^{md+r} > 0$ для некоторого $m \geqslant 0$; тогда если $w = s + dq + md + r$, то $P_{11}^w \geqslant P_{1j}^{md+r} P_{ji}^s P_{i1}^{qd} > 0$, и, значит, w кратно d; следовательно, $s + r$ кратно d. Но $P_{jj}^{h+s} \geqslant P_{ji}^s P_{ij}^h > 0$, так что $h + s$ делится на d. Эти два результата позволяют утверждать, что $h - r$ делится на d, т. е. $h = nd + r$ для некоторого $n \geqslant 0$.

Мы оставляем читателю проверить, что из доказанного следует, что множества $C_1, \dots, C_d$ не пересекаются и непусты, что $\bigcup_{i=1}^{d} C_i = W$ и что $i \in C_r$ гарантирует равенство $P_{ij} = 0$ для каждого $j \notin C_{r+1}$, где $C_{d+1} = C_1$. Проанализировав матрицу периодического класса, мы можем теперь доказать ранее высказанное утверждение относительно появления собственных значений, модуль которых равен 1, у матрицы переходных вероятностей марковской цепи.

Теорема 3.1. *Если* $\mathbf{P}$ — *матрица переходных вероятностей конечной неприводимой периодической марковской цепи с периодом* d, *то корни* d-*й степени из* 1 *являются собственными значениями матрицы* $\mathbf{P}$, *кратность каждого из них равна единице, и не существует других собственных значений, модуль которых равен* 1.

Доказательство. Пусть $D_1, \ldots, D_d$ — «циклические классы» процесса в том смысле, в каком мы их только что определили, т. е. из $i \in D_r$ следует, что $P_{ij} = 0$ для каждого $j \notin D_{r+1}$. Не теряя общности, предположим, что $D_1 = \{1, \ldots, n_1\}$, $D_2 = \{n_1 + 1, \ldots \ldots, n_1 + n_2\}, \ldots, D_d = \{M - n_d + 1, \ldots, M\}$. Из определения циклических классов следует, что

$$\mathbf{P}^d = \begin{Vmatrix} \mathbf{A}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{A}_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ & \cdot & & \cdot \\ 0 & 0 & \ldots & \mathbf{A}_d \end{Vmatrix},$$

где $\mathbf{A}_i$ — марковская $n_i \times n_i$-матрица. Кроме того, для каждого i существует целое число $m > 0$, такое, что $\mathbf{A}_i^m \gg 0$ (см. упр. 8 гл. 2). Следовательно, $\mathbf{A}_i$ имеет строго положительный левый собственный вектор $\boldsymbol{\mu}^{(i)}$, принадлежащий собственному значению 1 алгебраической кратности 1. Благодаря структуре матрицы $\mathbf{P}^d$ мы можем, дописав нужное количество нулей к $\boldsymbol{\mu}^{(i)}$, получить линейно независимые векторы $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}$, такие, что

$$\mathbf{x}^{(i)} = \mathbf{x}^{(i)}\mathbf{P}^d, \qquad i = 1, \ldots, d.$$

Рассмотрим векторы $\mathbf{y}^{(1)} = \mathbf{x}^{(1)}$, $\mathbf{y}^{(2)} = \mathbf{x}^{(1)}\mathbf{P}, \ldots, \mathbf{y}^{(d)} = \mathbf{x}^{(1)}\mathbf{P}^{d-1}$. Единственные отличные от нуля компоненты вектора $\mathbf{x}^{(1)}$ — это те, что имеют индексы $1, 2, \ldots, n_1$; вероятность $P_{jk}^{(h)}$ отлична от нуля только в том случае, если класс, которому принадлежит состояние j, переходит за h шагов в циклический класс, которому принадлежит состояние k. Таким образом, ненулевые компоненты вектора $\mathbf{y}^{(i)}$ имеют индексы $n_1 + \ldots + n_{i-1} + 1, \ldots, n_1 + \ldots + n_i$. Это значит, что векторы $\mathbf{y}^{(i)}$ $(i = 1, \ldots, d)$ линейно независимы. Кроме того,

$$\mathbf{y}^{(i)}\mathbf{P}^d = \mathbf{x}^{(1)}\mathbf{P}^{i-1}\mathbf{P}^d = \mathbf{x}^{(1)}\mathbf{P}^d\mathbf{P}^{i-1} = \mathbf{x}^{(1)}\mathbf{P}^{i-1} = \mathbf{y}^{(i)}.$$

Следовательно, если мы ограничимся рассмотрением n_i-мерного линейного пространства, включающего только те компоненты вектора $\mathbf{y}^{(i)}$, что лежат в D_i, то получим левый собственный вектор матрицы $\mathbf{A}_i$, соответствующий собственному значению 1.

Так как собственное значение 1 матрицы $\mathbf{A}_i$ простое (имеет кратность 1), то каждый $\mathbf{y}^{(i)}$ совпадает с $\mathbf{x}^{(i)}$ с точностью до скалярного множителя. На самом деле, если нормировать векторы $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}$ условием $\sum_{i=1}^{n} x_i^{(h)} = 1$, $h = 1, \ldots, d$, то получим, что $\mathbf{y}^{(h)} = \mathbf{x}^{(h)}$, $h = 1, \ldots, d$. Соответственно мы можем теперь записать $\mathbf{x}^{(2)} = \mathbf{x}^{(1)}\mathbf{P}$, $\mathbf{x}^{(3)} = \mathbf{x}^{(2)}\mathbf{P}, \ldots, \mathbf{x}^{(1)} = \mathbf{x}^{(d)}\mathbf{P}$.

Пусть $\omega = e^{2\pi i/d}$; тогда с помощью последних соотношений получаем

$$\left(\mathbf{x}^{(1)} + \mathbf{x}^{(2)} + \mathbf{x}^{(3)} + \ldots + \mathbf{x}^{(d)}\right) \mathbf{P} = \mathbf{x}^{(1)} + \mathbf{x}^{(2)} + \ldots + \mathbf{x}^{(d)},$$

$$\left(\mathbf{x}^{(1)} + \omega\mathbf{x}^{(2)} + \omega^2\mathbf{x}^{(3)} + \ldots + \omega^{d-1}\mathbf{x}^{(d)}\right) \mathbf{P} = \omega^{-1}\left(\mathbf{x}^{(1)} + \omega\mathbf{x}^{(2)} + \ldots + \omega^{d-1}\mathbf{x}^{(d)}\right),$$

$$\left(\mathbf{x}^{(1)} + \omega^2\mathbf{x}^{(2)} + \omega^4\mathbf{x}^{(3)} + \ldots + \omega^{2(d-1)}\mathbf{x}^{(d)}\right) \mathbf{P} = \omega^{-2}\left(\mathbf{x}^{(1)} + \omega^2\mathbf{x}^{(2)} + \ldots + \omega^{2(d-1)}\mathbf{x}^{(d)}\right),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\left(\mathbf{x}^{(1)} + \omega^{(d-1)}\mathbf{x}^{(2)} + \omega^{2(d-1)}\mathbf{x}^{(3)} + \ldots + \omega^{(d-1)^2}\mathbf{x}^{(d)}\right) \mathbf{P} = \omega^{-(d-1)}\left(\mathbf{x}^{(1)} + \omega^{(d-1)}\mathbf{x}^{(2)} + \ldots + \omega^{(d-1)^2}\mathbf{x}^{(d)}\right).$$

Линейная независимость векторов $\mathbf{x}^{(i)}$ гарантирует, что ни один из векторов в этих соотношениях не равен нулю. Кроме того, последние соотношения означают, что все корни d-й степени из 1 являются собственными значениями матрицы $\mathbf{P}$.

Предположим далее, что $\mathbf{xP} = \lambda\mathbf{x}$ для некоторого ненулевого вектора $\mathbf{x}$. Тогда $\mathbf{xP}^d = \lambda^d\mathbf{x}$. Разбивая этот вектор на векторы $\mathbf{z}^{(1)} = \left(x_1, \ldots, x_{n_1}\right)$, $\mathbf{z}^{(2)} = \left(x_{n_1+1}, \ldots, x_{n_1+n_2}\right), \ldots, \mathbf{z}^{(d)} = \left(x_{M-n_d+1}, \ldots, x_M\right)$, мы видим, что

$$\mathbf{z}^{(i)}\mathbf{A}_i = \lambda^d\mathbf{z}^{(i)}, \qquad i = 1, \ldots, d.$$

Поскольку по крайней мере один из $\mathbf{z}^{(i)}$ не равен нулю, а для каждой матрицы $\mathbf{A}_i$ существует целое число m, такое, что $\mathbf{A}_i^m \gg \mathbf{0}$, то либо $\lambda^d = 1$, либо $|\lambda^d| < 1$. Если $\lambda^d = 1$, то существуют константы $c_1, \ldots, c_d$, такие, что

$$\mathbf{z}^{(i)} = c_i\mathbf{x}^{(i)}, \qquad i = 1, \ldots, d,$$

а следовательно, $\mathbf{x} = c_1\mathbf{x}^{(1)} + \ldots + c_d\mathbf{x}^{(d)}$.

Итак,

$$\lambda\mathbf{x} = \mathbf{xP} = c_1\mathbf{x}^{(2)} + c_2\mathbf{x}^{(3)} + \ldots + c_d\mathbf{x}^{(1)},$$

или

$$\lambda c_1\mathbf{x}^{(1)} + \ldots + \lambda c_d\mathbf{x}^{(d)} = c_d\mathbf{x}^{(1)} + c_1\mathbf{x}^{(2)} + \ldots + c_{d-1}\mathbf{x}^{(d)}.$$

Так как векторы $\mathbf{x}^{(i)}$ линейно независимы, то

$$\lambda c_1 = c_d, \quad \lambda c_2 = c_1, \ \ldots, \ \lambda c_d = c_{d-1},$$

или

$$c_{d-1} = \lambda c_d = \left(\lambda^{-1}\right)^{d-1} c_d,$$

$$c_{d-2} = \lambda^2 c_d = \left(\lambda^{-1}\right)^{d-2} c_d, \ \ldots, \ c_1 = \lambda^{d-1} c_d = \lambda^{-1} c_d,$$

поскольку $\lambda^d = 1$. Это означает, что $\mathbf{x}$ с точностью до скалярного множителя совпадает с одним из уже построенных собственных векторов матрицы $\mathbf{P}$. ■

Перейдем к случаю произвольной марковской матрицы.

Теорема 3.2. *Если* $\mathbf{P}$ — *матрица переходных вероятностей конечной марковской цепи, то все ее собственные значения, по модулю равные* 1, *являются корнями из* 1. *Корни* d-*й степени из* 1 *являются собственными значениями матрицы* $\mathbf{P}$ *тогда и только тогда, когда множество состояний марковской цепи, которой соответствует* $\mathbf{P}$, *включает возвратный периодический класс периода* d. *Кратность каждого корня* d-*й степени совпадает с числом возвратных классов периода* d.

Доказательство этой теоремы по существу идентично доказательству теоремы 3.1. Поскольку из соотношения $\lambda\mathbf{x} = \mathbf{x}\mathbf{P}$ следует, что

$$\lambda^m\mathbf{x} = \mathbf{x}\mathbf{P}^m,$$

или, в координатной форме,

$$\lambda^m x_j = \sum_{i=1}^{n} x_i P_{ij}^m,$$

то, переходя к пределу при $m \to \infty$, мы видим, что $x_j = 0$, если состояние j невозвратно. Таким образом, мы можем рассматривать только возвратные состояния, что немедленно сводит рассматриваемый общий случай к предыдущей теореме.

§ 4. СПЕЦИАЛЬНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ ДЛЯ МАРКОВСКИХ ЦЕПЕЙ

Пусть $\mathbf{P}$ — матрица переходных вероятностей случайного блуждания по неотрицательным целым числам, вероятность перехода в одно из соседних состояний из состояния $k (k \geqslant 1)$ равна $^1/_2$, а нулевое состояние является отражающим экраном, т. е.

$$\mathbf{P} = \begin{Vmatrix} 0 & 1 & 0 & 0 & \ldots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ldots \\ \vdots & \vdots & \vdots & & \end{Vmatrix}.$$

Чтобы найти вероятность перехода из состояния k в состояние l за n шагов, мы могли бы получить n-ю степень матрицы $\mathbf{P}$ и выделить элемент P_{kl}^n. Этот путь, однако, слишком громоздок.

Другой возможный путь — это попытаться обобщить метод собственных значений и собственных векторов, развитый в § 2. В случае бесконечных матриц это не всегда можно сделать.

Однако для матриц только что описанного вида и, более того, для матриц переходных вероятностей, соответствующих процессам случайного блуждания, имеет место бесконечномерный аналог представления (1.1).

Мы сейчас получим выражение для P_{kl}^n способом, который иллюстрирует общий подход, применимый к произвольным процессам случайного блуждания.

Складывая два тригонометрических тождества

$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta,$$

приходим к следующему тождеству:

$$\cos\alpha\cos\beta = \frac{1}{2}\cos(\alpha+\beta) + \frac{1}{2}\cos(\alpha-\beta). \qquad (*)$$

Пусть $\alpha = \theta$, а $\beta = k\theta\,(k = 1, 2, \ldots)$; в этом случае имеем

$$\cos\theta\cos k\theta = \frac{1}{2}\cos(k+1)\theta + \frac{1}{2}\cos(k-1)\theta. \qquad (4.1)$$

Так как элементы k-й строки матрицы $\mathbf{P}$ имеют вид

$$P_{k,0} = P_{k,1} = 0, \ldots, \qquad P_{k,k-2} = 0, \quad P_{k,k-1} = \frac{1}{2}, \quad P_{k,k} = 0,$$
$$P_{k,k+1} = \frac{1}{2}, \quad P_{k,k+2} = 0, \ldots, \quad k = 2, 3, \ldots,$$
$$P_{1,0} = \frac{1}{2}, \quad P_{1,1} = 0, \quad P_{1,2} = \frac{1}{2}, \quad P_{1,3} = 0, \ldots,$$
$$P_{0,0} = 0, \quad P_{0,1} = 1, \quad P_{0,2} = 0, \ldots,$$

тождества (4.1) можно записать как

$$\cos\theta\cos k\theta = \sum_{r=0}^{\infty} P_{kr}\cos r\theta, \qquad k = 0, 1, \ldots. \qquad (4.2)$$

Умножая обе части (4.2) на $\cos\theta$, получаем

$$\cos^2\theta\cos k\theta = \sum_{r=0}^{\infty} P_{kr}\cos\theta\cos r\theta. \qquad (4.3)$$

Представим произведение $\cos\theta\cos r\theta$ с помощью (4.2) в виде

$$\cos\theta\cos r\theta = \sum_{s=0}^{\infty} P_{rs}\cos s\theta.$$

Подставляя это в (4.3), получаем

$$\cos^2\theta\cos k\theta = \sum_{r=0}^{\infty}\left(P_{kr}\sum_{s=0}^{\infty}P_{rs}\cos s\theta\right) = \sum_{s=0}^{\infty}\left(\cos s\theta\sum_{r=0}^{\infty}P_{kr}P_{rs}\right) = \sum_{s=0}^{\infty}P_{ks}^2\cos s\theta.$$

Отметим, что суммирование от 0 до ∞ введено лишь для упрощения обозначений; на самом деле лишь конечное число слагаемых отлично от нуля. После $n-1$ повторений процедуры, состоящей в умножении на $\cos\theta$ и изменении порядка суммирования, получаем

$$\cos^n\theta\cos k\theta = \sum_{r=0}^{\infty}P_{kr}^n\cos r\theta. \tag{4.4}$$

Умножим обе части этого уравнения на $\cos s\theta$ и проинтегрируем по θ от 0 до 2π:

$$\int_0^{2\pi}\cos^n\theta\cos k\theta\cos s\theta\,d\theta = \int_0^{2\pi}\sum_{r=0}^{\infty}P_{kr}^n\cos r\theta\cos s\theta\,d\theta = \sum_{r=0}^{\infty}P_{kr}^n\int_0^{2\pi}\cos r\theta\cos s\theta\,d\theta. \tag{4.5}$$

Используя тождество (*) с $\alpha = r\theta$ и $\beta = s\theta$, легко показать, что

$$\int_0^{2\pi}\cos r\theta\cos s\theta\,d\theta = \begin{cases} 0 & \text{при } r\neq s, \\ \pi & \text{при } r = s\geqslant 1, \\ 2\pi & \text{при } r = s = 0. \end{cases} \tag{4.6}$$

Из (4.5) и (4.6) сразу же получаем

$$P_{ks}^n = \begin{cases} \dfrac{1}{\pi}\displaystyle\int_0^{2\pi}\cos^n\theta\cos k\theta\cos s\theta\,d\theta & \text{при } s\neq 0, \\ \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}\cos^n\theta\cos k\theta\,d\theta & \text{при } s = 0. \end{cases} \tag{4.7}$$

Эти интегралы легко вычислить для любых заданных n, k и s.

Общий же метод, частным случаем которого мы только что воспользовались, состоит в следующем.

Пусть нам задан процесс случайного блуждания на множестве неотрицательных целых чисел с матрицей одношаговых переходных вероятностей вида

$$\mathbf{P}=\left\|\begin{matrix} r_0 & p_0 & 0 & 0 & \dots \\ q_1 & r_1 & p_1 & 0 & \dots \\ 0 & q_2 & r_2 & p_2 & \dots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix}\right\|, \tag{4.8}$$

где $q_n+r_n+p_n=1$, $q_n>0$, $p_n>0$, $r_n\geqslant 0$ при $n=1, 2, \dots$ и $r_0+p_0=1$, $p_0>0$, $r_0\geqslant 0$. (Отметим для будущих ссылок, что ни один из результатов, которые мы получим ниже в этом параграфе, не зависит от условий $q_n+r_n+p_n=1$, $n=1, 2, \dots$, и $r_0+p_0=1$.) Рассмотрим следующую систему уравнений:

$$xQ_k(x)=q_kQ_{k-1}(x)+r_kQ_k(x)+p_kQ_{k+1}(x), \qquad k=1,\ 2,\ \dots, \tag{4.9}$$

с начальными условиями $Q_0(x)\equiv 1$ и $Q_1(x)=(x-r_0)/p_0$. Поскольку $p_n>0$ при всех $n=0, 1, 2, \dots$, ясно, что $Q_n(x)$, $n\geqslant 2$, последовательно определяются из формулы (4.9) и что $Q_n(x)$ является полиномом от x степени в точности n. Далее мы воспользуемся теоремой, доказательство которой выходит за рамки этой книги и которая устанавливает следующий факт. Существует неубывающая и непостоянная функция $\sigma(x)$, определенная на отрезке $[-1, 1]$, такая, что

$$\int_{-1}^{1} Q_k(x)\,Q_s(x)\,d\sigma(x)\begin{cases} =0 & \text{при} \quad k\neq s, \\ >0 & \text{при} \quad k=s, \end{cases} \qquad k=0,\ 1,\ 2,\ \dots. \tag{4.10}$$

О функциях $Q_k(x)$, $k=0, 1, \dots$, обладающих свойством (4.10), говорят, что они являются ортогональными полиномами относительно распределения $\sigma(x)$ на отрезке $[-1, 1]$. Функция $\sigma(x)$ единственна с точностью до аддитивной постоянной. Эта общая теорема позволяет получить явные выражения для вероятностей P^n_{ks}. В самом деле, учитывая то, как мы задали $Q_0(x)$ и $Q_1(x)$, уравнения (4.9) можно переписать в виде

$$xQ_k(x)=\sum_{r=0}^{\infty} P_{kr}Q_r(x), \qquad k=0,\ 1,\ \dots. \tag{4.11}$$

Умножая обе части на x и подставляя это в (4.11), получаем

$$x^2Q_k(x)=\sum_{r=0}^{\infty} P_{kr}\sum_{s=0}^{\infty} P_{rs}Q_s(x)=\sum_{s=0}^{\infty} P^2_{ks}Q_s(x), \qquad k=0,\ 1,\ \dots. \tag{4.12}$$

Продолжая таким же образом, переходим к соотношениям

$$x^n Q_k(x) = \sum_{r=0}^{\infty} P_{kr}^n Q_r(x), \qquad k = 0,\ 1,\ \ldots, \quad n = 1,\ 2,\ \ldots . \tag{4.13}$$

Умножая обе части на $Q_s(x)$ и интегрируя на интервале $[-1, 1]$ по $d\sigma(x)$, мы находим, воспользовавшись соотношением ортогональности (4.10), что

$$\int_{-1}^{1} x^n Q_k(x)\, Q_s(x)\, d\sigma(x) = \sum_{r=0}^{\infty} P_{kr}^n \int_{-1}^{1} Q_r(x)\, Q_s(x)\, d\sigma(x) = P_{ks}^n \int_{-1}^{1} Q_s^2(x)\, d\sigma(x),$$

откуда следует доказываемая формула

$$P_{ks}^n = \frac{\int\limits_{-1}^{1} x^n Q_k(x)\, Q_s(x)\, d\sigma(x)}{\int\limits_{-1}^{1} Q_s^2(x)\, d\sigma(x)} . \tag{4.14}$$

Как мы уже отмечали, приведенная процедура напоминает метод диагонализации из § 1. Действительно, соотношения (4.9) просто означают, что для каждого x бесконечномерный вектор $(Q_0(x),\ Q_1(x),\ \ldots)$ является, формально, собственным вектором матрицы (4.8), соответствующим собственному значению x. Поскольку мы имеем дело с континуумом собственных значений, естественно предположить, что представление P_{ij}^n дискретной суммой, аналогичное полученному в § 1, 2, не имеет места. Оказывается, однако, что соответствующее обобщение спектрального представления (1.1) дается формулой (4.14). В его основе лежит существование «непрерывного спектра» в дополнение к дискретному (возможно, пустому), что является характерным для бесконечных матриц. Строгое математическое рассмотрение этих вопросов выходит за рамки настоящей книги. Тем не менее мы рассмотрим еще несколько иллюстративных примеров, связанных с этой теорией.

Может показаться, что изложенный метод, сколь ни элегантна его теория, не имеет практической ценности. Действительно, чтобы найти P_{ks}^n, необходимо определить полиномы $\{Q_k(x)\}_{k=0}^{\infty}$ и распределение $\sigma(x)$, относительно которого нами установлен до сих пор лишь факт существования. В действительности же положение дел много лучше, чем это намечается указанными трудностями. Прежде всего мы располагаем хорошо разработанной теорией ортогональных полиномов, которая позволяет получить важные теоретические результаты, касающиеся поведения вероятностей P_{ks}^n, и в частности их отношений, при $n \to \infty$.

С другой стороны, процессы случайного блуждания, возникающие в конкретных задачах, имеют матрицы переходных вероятностей более специального вида, чем (4.8). Например, может быть, что $p_0 = p_1 = p_2 = \ldots$, $q_0 = q_1 = \ldots$, или же $p_n = p_{n+1} = \ldots$, $q_n = q_{n+1} = \ldots$ для некоторого n. В этих случаях, как впрочем и в других, можно показать, что полиномы $Q_n(x)$ являются комбинациями различных классических систем полиномов, теория которых хорошо развита.

§ 5. ПРИМЕРЫ

(а) *Симметричное случайное блуждание с отражающим экраном.* Для того чтобы подвести вычисления, проведенные в § 4, под общую схему, основывающуюся на ортогональных полиномах, достаточно положить

$$Q_k(x) = \cos k(\arc\cos x), \qquad k = 0,\ 1,\ \ldots .[1]$$

Полиномы $Q_k(x)$ ортогональны на отрезке $[-1, 1]$ по отношению к распределению $d\sigma(x) = \rho(x)\,dx$, где $\rho(x) = (^1/_2\,\pi)(1-x^2)^{-^1/_2}$, так как

$$\int\limits_{-1}^{1} Q_k(x)\,Q_l(x)\,\rho(x)\,dx = C\int\limits_{0}^{\pi}\cos k\theta\cos l\theta\,d\theta = 0, \quad \text{если} \quad k \neq l,$$

в чем убеждаемся заменой переменной $x = \cos\theta$.

(б) *Еще один процесс случайного блуждания с отражающим экраном.* Рассмотрим случайное блуждание во множестве неотрицательных целых чисел матрицей переходных вероятностей вида

$$\mathbf{P} = \begin{Vmatrix} 0 & 1 & 0 & 0 & 0 & \ldots \\ q & 0 & p & 0 & 0 & \ldots \\ 0 & q & 0 & p & 0 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \end{Vmatrix},$$

где $q,\ p > 0$, $q + p = 1$.

Умножая обе части тождества (4.1) на $2\sqrt{pq}\,(\sqrt{p/q})^k$, получаем

$$2\sqrt{pq}\cos\theta\,(\sqrt{q/p})^k\cos k\theta = p(\sqrt{q/p})^{k+1}\cos(k+1)\theta + \\ + q(\sqrt{q/p})^{k-1}\cos(k-1)\theta, \qquad k = 1,\ 2,\ \ldots . \tag{5.1}$$

Таким образом, полиномы

$$Q_k(x) = (\sqrt{q/p})^k\cos k\theta, \quad 2\sqrt{pq}\cos\theta = x, \qquad k = 0,\ 1,\ 2,\ \ldots,$$

---

[1]) $Q_k(x)$ называются полиномами Чебышева. — *Прим. ред.*

удовлетворяют системе уравнений (4.9), соответствующей рассматриваемой матрице $\mathbf{P}$, за исключением уравнения для $k=0$. Здесь мы имеем $Q_0(x)=1$ и $Q_1(x)=x/2p$, тогда как нам нужны начальные условия $Q_0(1)=1$, $Q_1(x)=x$.

Чтобы исправить положение, начнем с тождества

$$\cos\theta\sin(k+1)\theta=\frac{1}{2}\sin k\theta+\frac{1}{2}\sin(k+2)\theta,\quad k=0,\ 1,\ 2,\ \ldots. \tag{5.2}$$

Умножив обе его части на $2\sqrt{pq}\left(\sqrt{q/p}\right)^k$ и разделив на $\sin\theta$, перепишем (5.2) в виде

$$2\sqrt{pq}\,(\cos\theta)\left(\sqrt{q/p}\right)^k\frac{\sin(k+1)\theta}{\sin\theta}=$$
$$=q\left(\sqrt{q/p}\right)^{k-1}\frac{\sin k\theta}{\sin\theta}+p\left(\sqrt{q/p}\right)^{k+1}\frac{\sin(k+2)\theta}{\sin\theta},\qquad k=1,\ 2,\ \ldots.$$

Пусть

$$Z_k(\theta)=\left(\sqrt{q/p}\right)^k\frac{\sin(k+1)\theta}{\sin\theta},\qquad k=0,\ 1,\ \ldots;$$

тогда

$$Z_0(\theta)\equiv 1\quad\text{и}\quad Z_r(\theta)=\sqrt{q/p}\,\frac{\sin 2\theta}{\sin\theta},$$

и при этом

$$2\sqrt{pq}\,(\cos\theta)\,Z_k(\theta)=qZ_{k-1}(\theta)+pZ_{k+1}(\theta),\qquad k=1,\ 2,\ \ldots.$$

Пусть

$$R_k(x)=Z_k(\theta);\quad x=2\sqrt{pq}\cos\theta.$$

Отметим, что $R_0(x)=1$ и $R_1(x)=x/p$, тогда как

$$xR_k(x)=qR_{k-1}(x)+pR_{k+1}(x),\qquad k=1,\ 2,\ \ldots,$$

причем $R_k(x)$ является полиномом k-й степени. Наконец, положим $P_k(x)=(2p-1)R_k(x)+(2-2p)Q_k(x)$, $k=0,\ 1,\ 2,\ \ldots$; тогда $P_0(x)\equiv 1$, $P_1(x)=x$ и

$$xP_k(x)=qP_{k-1}(x)+pP_{k+1}(x),\qquad k=1,\ 2,\ \ldots,$$

так как $R_k(x)$ и $Q_k(x)$ удовлетворяют одним и тем же соотношениям. Таким образом, полиномы $P_k(x)$ отвечают матрице переходных вероятностей $\mathbf{P}$.

Детальная процедура нахождения распределения $\sigma(x)$, ортогонализирующего полиномы $P_k(x)$ на отрезке $[-1,\ 1]$, нами не рассматривается. Мы просто приведем соответствующее распределение, оставив читателю проверить, что оно обладает всеми нужными свойствами.

Если $p \geqslant 1/2$, то $\sigma(x)$ постоянна вне $[-\sqrt{4pq}, \sqrt{4pq}]$, а в самом интервале

$$d\sigma(x) = \frac{C\sqrt{4pq - x^2}}{1 - x^2}\,dx.$$

Если $p < 1/2$, то $\sigma(x)$ сохраняет свой вид внутри отрезка $[-\sqrt{4pq}, \sqrt{4pq}]$, а в точках -1 и $+1$ появляются скачки величины $1/2(1-2p)/q$. Константа C служит в качестве нормирующего множителя, обеспечивающего равенство единице интеграла $\int\limits_{-1}^{1} d\sigma(x)$.

(в) *Случайное блуждание с поглощающим экраном.* В качестве следующего примера мы рассмотрим процесс случайного блуждания по целым числам $-1, 0, 1, 2, 3, \ldots$ с вероятностью перехода в одно из соседних состояний из состояния $k\,(k \geqslant 0)$, равной $1/2$, и с поглощающим экраном, расположенным в состоянии -1. Матрица переходных вероятностей этого процесса имеет вид

$$\mathbf{P} = \left\| \begin{matrix} 1 & 0 & 0 & 0 & \ldots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ldots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix} \right\|.$$

Хотя эта матрица и отличается от тех, для которых был развит общий метод, мы будем следовать, по существу, процедуре, изложенной в § 4.

Ключевым в нашем анализе будет тождество

$$\cos\theta \sin(k+1)\theta = \frac{1}{2}\sin k\theta + \frac{1}{2}\sin(k+2)\theta, \quad k = 0, 1, 2, \ldots. \tag{5.3}$$

Так как k-я строка матрицы $\mathbf{P}$ состоит из элементов

$$P_{k,-1} = 0, \quad P_{k,0} = 0, \ldots, P_{k,k-1} = \frac{1}{2}, \quad P_{k,k} = 0, \quad P_{k,k+1} = \frac{1}{2},$$
$$P_{k,k+2} = 0, \ldots, \quad k = 1, 2, \ldots,$$
$$P_{0,-1} = \frac{1}{2}, \quad P_{0,0} = 0, \quad P_{0,1} = \frac{1}{2}, \quad P_{0,2} = 0, \ldots,$$

соотношение (5.3) можно записать для $k = 0, 1, \ldots$ в виде

$$\cos\theta \sin(k+1)\theta = \sum_{r=-1}^{\infty} P_{kr} \sin(r+1)\theta.$$

Умножая обе части на $\cos\theta$ и подставляя

$$\cos\theta\sin(r+1)\theta=\sum_{s=-1}^{\infty}P_{rs}\sin(s+1)\theta$$

в правую часть получающегося соотношения, имеем

$$\begin{aligned}\cos^2\theta\sin(k+1)\theta&=\sum_{r=-1}^{\infty}P_{kr}\sum_{s=-1}^{\infty}P_{rs}\sin(s+1)\theta=\\&=\sum_{s=-1}^{\infty}\sin(s+1)\theta\sum_{r=-1}^{\infty}P_{kr}P_{rs}=\\&=\sum_{s=-1}^{\infty}P_{ks}^{2}\sin(s+1)\theta.\end{aligned}$$

Повторив эти действия $n-1$ раз, приходим к формуле

$$\cos^n\theta\sin(k+1)\theta=\sum_{r=-1}^{\infty}P_{kr}^{n}\sin(r+1)\theta.$$

Умножим обе ее части на $\sin(s+1)\theta$ и проинтегрируем по θ на отрезке $[0,2\pi]$:

$$\begin{aligned}&\int_0^{2\pi}\cos^n\theta\sin(k+1)\theta\sin(s+1)\theta\,d\theta=\\&\qquad=\int_0^{2\pi}\sum_{r=-1}^{\infty}P_{kr}^{n}\sin(r+1)\theta\sin(s+1)\theta\,d\theta=\\&=\sum_{r=-1}^{\infty}P_{kr}^{n}\int_0^{2\pi}\sin(r+1)\theta\sin(s+1)\theta\,d\theta\qquad(s=0,\ 1,\ \ldots).\end{aligned}\tag{5.4}$$

Легко показать, используя элементарные тригонометрические тождества, что

$$\int_0^{2\pi}\sin(r+1)\theta\sin(s+1)\theta\,d\theta=\begin{cases}0, & \text{если}\quad r\neq s,\\ \pi, & \text{если}\quad r=s,\end{cases}\tag{5.5}$$

$$r,\ s=0,\ 1,\ 2,\ \ldots.$$

Из (5.4) и (5.5) следует, что вероятности перехода за n шагов выражаются формулой

$$P_{ks}^{n}=\frac{1}{\pi}\int_0^{2\pi}\cos^n\theta\sin(k+1)\theta\sin(s+1)\theta\,d\theta,\tag{5.6}$$

$$k,\ s=0,\ 1,\ 2,\ \ldots,\quad n=0,\ 1,\ \ldots.$$

До сих пор все, что мы сделали, — это получили применение общего метода к матрице

$$\mathbf{P}' = \begin{Vmatrix} 0 & \frac{1}{2} & 0 & 0 & \dots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \dots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix},$$

получаемой вычеркиванием первой строки и первого столбца из матрицы

$$\mathbf{P} = \begin{Vmatrix} 1 & 0 & 0 & 0 & \dots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \dots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix}.$$

Возможность такого сведе́ния при вычислении вероятностей P_{ks}^n перехода за n шагов для $k, s = 0, 1, \dots$ основывается на том факте, что нам нет необходимости рассматривать те траектории, которые ведут в состояние -1, так как эти траектории не могут выйти из него.

Ортогональные полиномы в рассматриваемом случае таковы:

$$Q_k(x) = \frac{\sin(k+1)\theta}{\sin\theta}, \qquad k = 0, 1, 2, \dots,$$

где $x = \cos\theta$ и, как это легко проверить, полиномы $Q_k(x)$ ортогональны по отношению к $d\sigma(x) = \pi^{-1}(1-x^2)^{+1/2}\,dx$ на интервале $[-1, 1]$.

Как приложение полученного результата вычислим вероятность события, состоящего в том, что поглощение состоянием -1 произойдет точно на n-м шаге, если исходным состоянием было состояние k. Поглощение состоянием -1 на n-м шаге может произойти, очевидно, только в том случае, если на $(n-1)$-м шаге процесс пребывал в состоянии 0. Но вероятность попасть в состояние 0 на $(n-1)$-м шаге, отправляясь из состояния k, задается формулой (5.6), т. е.

$$P_{k0}^{n-1} = \frac{1}{\pi}\int_0^{2\pi} \cos^{n-1}\theta \sin(k+1)\theta \sin\theta\, d\theta,$$

5 Зак. 939

а вероятность $P_{0,-1}=\frac{1}{2}$. Следовательно, вероятность поглощения состоянием -1 на n-м шаге при начальном состоянии k есть

$$A_k^n=\frac{1}{2\pi}\int_0^{2\pi}\cos^{n-1}\theta\sin(k+1)\theta\sin\theta\,d\theta. \tag{5.7}$$

§ 6. ПРИЛОЖЕНИЯ К БРОСАНИЯМ МОНЕТЫ

Рассмотренные процессы случайного блуждания связаны с задачами о бросании монеты. Предположим, что два игрока договариваются провести серию бросаний симметричной монеты на следующих условиях: если выпадает герб, то игрок I выигрывает единицу у игрока II, в противном случае он проигрывает единицу игроку II. Пусть

$$X_i=\begin{cases}+1, & \text{если игрок I выигрывает,}\\ -1, & \text{если игрок I проигрывает}\end{cases}$$

при i-м бросании монеты. Тогда $\mathrm{P}\{X_i=+1\}=\mathrm{P}\{X_i=-1\}={}^1/_2$ и $S_n=\sum_{i=1}^{n}X_i$ $(n\geqslant 1)$ есть суммарный выигрыш игрока I после n бросаний монеты. Положим также $S_0=0$. Один из самых простых вопросов, касающихся этой игры, таков: какова вероятность того, что после n бросаний монеты суммарный выигрыш игрока I будет равен нулю? Очевидно, что выигрыш игрока I не может быть равным нулю, если n нечетно. Пусть $n=2m$, тогда если выигрыш равен нулю после $2m$ бросаний, то игрок I выиграл m партий и столько же проиграл. Искомая вероятность, очевидно, равна

$$\mu_{2m}=2^{-2m}\binom{2m}{m},\qquad m=1,\ 2,\ \ldots.$$

Далее, чему равна вероятность того, что после $n=2m$ бросаний выигрыш игрока I будет равняться нулю в первый раз? Очевидно, S_n описывают симметричное случайное блуждание на множестве всех целых чисел. Поэтому наш вопрос может быть сформулирован так: чему равна вероятность $f_{0,0}^n$ первого возвращения в состояние 0 на n-м шаге? Первый переход из состояния 0 может произойти в одно из двух состояний -1 или $+1$ с одинаковыми вероятностями обоих исходов, равными ${}^1/_2$. В силу очевидной симметрии относительно нулевого состояния вероятность первого достижения состояния 0 из состояния $+1$ должна равняться вероятности первого достижения состояния 0 из состояния -1, поэтому мы ответим на вопрос, если найдем вероятность первого достижения состояния 0 из состояния $+1$ за $2m-1$ шагов. Но эта вероятность равна вероятности первого достижения состояния -1 из состояния 0 в силу однородного характера процесса. Последняя же вероятность равна вероятности A_0^{2m-1} поглощения состоянием -1 за $2m-1$ шагов при условии, что исходным состоянием было состояние 0, в процессе случайного блуждания на множестве целых чисел $\{-1, 0, 1, 2, \ldots\}$ с поглощающим экраном, расположенным в состоянии -1. Из формулы (5.7) при $k=0$ и $n=2m-1$ получаем

$$A_0^{2m-1} = \frac{1}{2\pi}\int_0^{2\pi} \cos^{2m-2}\theta \sin^2\theta\, d\theta = \frac{1}{\pi}\int_0^{\pi} \cos^{2m-2}\theta \sin^2\theta\, d\theta.$$

В последнем интеграле сделаем замену $x = \cos\theta$; тогда

$$A_0^{2m-1} = \frac{1}{\pi}\int_{-1}^{1} x^{2m-2}(1-x^2)^{\frac{1}{2}}\, dx = \frac{2}{\pi}\cdot\frac{1}{2}\int_0^1 (x^2)^{m-\frac{3}{2}}(1-x^2)^{\frac{1}{2}}\, 2x\, dx.$$

Еще одна замена $t = x^2$ дает

$$A_0^{2m-1} = \frac{1}{\pi}\int_0^1 t^{m-\frac{3}{2}}(1-t)^{\frac{3}{2}}\, dt = \frac{1}{\pi} B\left(m-\frac{1}{2},\ \frac{3}{2}\right), \tag{6.1}$$

где

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

есть бета-функция, которую можно выразить через гамма-функцию:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Воспользовавшись известными свойствами гамма-функции, получаем

$$A_0^{2m-1} = \frac{1}{\pi}\,\frac{\Gamma\left(m-\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)}{\Gamma(m+1)} =$$

$$= \frac{1}{\pi}\,\frac{\left(m-\frac{3}{2}\right)\left(m-\frac{5}{2}\right)\cdot\ \ldots\ \cdot\frac{3}{2}\cdot\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right)\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right)}{m(m-1)\cdot\ \ldots\ \cdot 2\cdot 1}.$$

Так как $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, мы находим, что вероятность равенства нулю суммарного выигрыша игрока I после $2m$ бросаний монеты задается формулой

$$f_{0,0}^{2m}=\begin{cases}\dfrac{1}{2} & \text{при} \quad m=1,\\ \dfrac{(2m-3)(2m-5)\cdot\ \ldots\ \cdot 3\cdot 1}{2^m\cdot m!} & \text{при} \quad m\geqslant 2.\end{cases} \tag{6.2}$$

Простой подсчет приводит к следующему интересному результату:

$$f_{0,0}^{2m}=\mu_{2m-2}-\mu_{2m}, \qquad m=1,\ 2,\ \ldots,$$

где по определению $\mu_0 = 1$. Далее,

$$\sum_{k=m+1}^{\infty} f_{0,0}^{2k}=\sum_{k=m+1}^{\infty}\left[\mu_{2k-2}-\mu_{2k}\right]=\mu_{2m}-\lim_{n\to\infty}\mu_{2n}=$$

$$=\mu_{2m}-\lim_{n\to\infty}2^{-2n}\binom{2n}{n}=\mu_{2m}, \tag{6.3}$$

а это значит, что

$$\mu_{2m}=\mathrm{P}\{S_{2m}=0\}=\mathrm{P}\{S_1\neq 0,\ S_2\neq 0,\ \ldots,\ S_{2m}\neq 0\}. \tag{6.4}$$

Чему равна вероятность того, что выигрыш игрока I обратится в нуль в k-й раз ($k>1$) после $2m$-го бросания монеты? Мы ответим на этот вопрос, сформулировав его в терминах процесса случайного блуждания: «какова вероятность того, что k-е возвращение в нулевое состояние произойдет на $2m$-м шаге?» Как и ранее, можно считать, что из состояния 0 на первом шаге процесс попадает в состояние $+1$, и, более того, это же происходит после каждого из $k-1$ первых возвращений в состояние 0. Далее, поскольку наше случайное блуждание однородно, мы можем «менять местами» во времени промежуточные шаги, не изменяя при этом вероятности достижения одного состояния из другого. Так, мы можем считать, что непосредственно за первым шагом «вправо» (в состояние $+1$) происходят все переходы в состояние $+1$, следующие за каждым из $k-1$ первых возвращений в состояние 0, и, таким образом, за первые k шагов процесс оказывается в состоянии k. Тогда искомая вероятность равна вероятности достижения состояния 0 в первый раз за $2m-k$ шагов при начальном состоянии k, или, что то же, вероятности достижения состояния -1 из состояния $k-1$ в первый раз за $2m-k$ шагов. Последняя же вероятность есть вероятность A_{k-1}^{2m-k} поглощения состоянием -1 на $(2m-k)$-м шаге при начальном состоянии $k-1$ в процессе случайного блуждания по множеству целых чисел $\{-1,\ 0,\ 1,\ 2,\ \ldots\}$

с поглощающим экраном, расположенным в состоянии -1. Формула (5.7) дает

$$A_{k-1}^{2m-k} = \frac{1}{2\pi} \int_0^{2\pi} \cos^{2m-k-1}\theta \sin k\theta \sin\theta \, d\theta. \tag{6.5}$$

Значение этого интеграла, как можно подсчитать (см. задачу 7 в конце главы), таково:

$$A_{k-1}^{2m-k} = \frac{1}{2^{2m-k}} \binom{2m-k}{m} \frac{k}{2m-k}. \tag{6.6}$$

Рассмотрим теперь последовательность $S_1, S_2, \ldots, S_{2m}$. Нас будет интересовать следующий вопрос: чему равна вероятность того, что ровно k членов этой последовательности обращаются в нуль. Это соответствует вероятности $z_{k,2m}$ того, что за $2m$ бросаний монеты выигрыш игрока I обратится в нуль ровно k раз.

В силу (6.4) имеем $z_{0,2m} = \mu_{2m}$. Вычислим теперь $z_{1,2m}$. Пусть B_r — событие, состоящее в том, что среди $S_1, \ldots, S_{2m}$ только S_{2r} равно нулю. Тогда для $r < m$

$$B_r = \{S_1 \neq 0, \ldots, S_{2r-1} \neq 0, S_{2r} = 0, S_{2r+1} \neq 0, \ldots, S_{2m} \neq 0\} =$$
$$= \{S_1 \neq 0, \ldots, S_{2r-1} \neq 0, S_{2r} = 0\} \cap \{S_{2r+1} - S_{2r} \neq 0, \ldots, S_{2m} - S_{2r} \neq 0\}.$$

Ясно, что события $\{S_1 \neq 0, \ldots, S_{2r-1} \neq 0, S_{2r} = 0\}$ и $\{S_{2r+1} - S_{2r} \neq \neq 0, \ldots, S_{2m} - S_{2r} \neq 0\}$ независимы и вероятность последнего есть просто μ_{2m-2r}. Тогда

$$P\{B_r\} = f_{0,0}^{2r} z_{0,2m-2r}, \qquad 1 \leqslant r \leqslant m,$$

где по определению $z_{0,0} = 1$. Таким образом,

$$z_{1,2m} = \sum_{r=1}^{m} P\{B_r\} = \sum_{r=1}^{m} f_{0,0}^{2r} z_{0,2m-2r}. \tag{6.7}$$

Однако μ_{2m-2r} есть вероятность того, что выигрыш игрока I будет равен нулю после $2m - 2r$ бросаний монеты; эта вероятность, как мы уже отмечали, равна вероятности $P\{S_{2m} - S_{2r} = 0\}$. События $\{S_1 \neq 0, \ldots, S_{2r-1} \neq 0, S_{2r} = 0\}$ и $\{S_{2m} - S_{2r} = 0\}$ также независимы, и, таким образом, мы получаем соотношение

$$P\{B_r\} = P\{S_1 \neq 0, \ldots, S_{2r-1} \neq 0, S_{2r} = 0, S_{2m} - S_{2r} = 0\}.$$

События в правой части этого соотношения несовместны, а их объединение составляет событие $\{S_{2m} = 0\}$; следовательно,

$$\sum_{r=1}^{m} P\{B_r\} = P\{S_{2m} = 0\} = \mu_{2m}.$$

Таким образом, $z_{1,2m} = \mu_{2m} = z_{0,2m}$ при $m \geqslant 1$. Точно так же можно показать, что

$$z_{k,2m} = \sum_{r=1}^{m-1} f_{0,0}^{2r} z_{k-1,\,2m-2r}, \qquad k \geqslant 2, \quad m \geqslant 1. \tag{6.8}$$

Сравнивая (6.8) и (6.7) и используя равенство $z_{1,2m} = z_{0,2m}$, а также тот факт, что $f_{0,0}^{2m} = \mu_{2m-2} - \mu_{2m} = z_{0,\,2m-2} - z_{1,\,2m}$, получаем

$$z_{2,\,2m} = z_{1,\,2m} - f_{0,0}^{2m} = 2z_{1,\,2m} - z_{0,\,2m-2}, \qquad m \geqslant 2.$$

Подставляя это в (6.8), индукцией получаем следующее рекуррентное соотношение:

$$z_{k,2m} = 2z_{k-1,2m} - z_{k-2,\,2m-2}. \tag{6.9}$$

Если положить в нем $z_{k,2m} = 2^{k-2m} a_{k,2m}$, то оно сведется к соотношению

$$a_{k-1,2m} = a_{k,2m} + a_{k-2,\,2m-2}. \tag{6.10}$$

Этому рекуррентному соотношению удовлетворяют величины

$$a_{k,2m} = \binom{2m-k}{m}. \tag{6.11}$$

Итак, нам известны $z_{0,2m}$ и $z_{1,2m}$, а прямая подстановка показывает, что $a_{0,2m}$ и $a_{1,2m}$ задаются формулой (6.11). Очевидно, соотношение (6.10) однозначно определяет $a_{k,2m}$ при $k \geqslant 2$; следовательно, формула (6.11) дает нам выражение для величин $a_{k,2m}$, и поэтому

$$z_{k,2m} = 2^{k-2m} \binom{2m-k}{m}. \tag{6.12}$$

С тем чтобы ответить на следующий вопрос, касающийся последовательности $S_1, \ldots, S_{2n}$, определим понятие перемены знака в последовательности. Мы скажем, что в момент k произошла перемена знака, если $S_k = 0$ и $S_{k-1}S_{k+1} = -1$. Вопрос же состоит в следующем. Чему равна вероятность того, что в последовательности $S_1, \ldots, S_{2n}$ имеется ровно r перемен знака? Событие, о котором идет речь, произойдет, если среди $S_2, S_4, \ldots, S_{2n-2}$ имеется ровно k нулей ($k = r, r+1, \ldots, n-1$) и ровно в r из них процесс меняет направление[1]). Пусть

$$C_{r,2n} = \sum_{k=r}^{n-1} \mathrm{P}\{\text{среди } S_2, S_4, \ldots, S_{2n-2} \text{ ровно } k \text{ нулей}\} \times$$
$$\times \mathrm{P}\{\text{находясь в нуле } k \text{ раз, процесс меняет направление } r \text{ раз}\}.$$

---

[1]) Под сменой направления следует понимать переход на другую полуось. — *Прим. перев.*

Но

$$\mathsf{P}\{\text{среди } S_2, S_4, \ldots, S_{2n-2} \text{ ровно } k \text{ нулей}\} = \\ = \mathsf{P}\{\text{среди } S_1, S_2, S_3, S_4, \ldots, S_{2n-2} \text{ ровно } k \text{ нулей}\} = \\ = 2^{k-2n+2}\binom{2n-k-2}{n-1},$$

как это следует из формулы (6.12). Далее, смена направления в нуле происходит с вероятностью $^1/_2$. Следовательно,

$$\mathsf{P}\{\text{находясь в нуле } k \text{ раз, процесс меняет направление } r \text{ раз}\} = \\ = \binom{k}{r}\frac{1}{2^r}\frac{1}{2^{k-r}} = \binom{k}{r}2^{-k},$$

откуда

$$C_{r,2n} = \sum_{k=r}^{n-1} 2^{k-2n+2}\binom{2n-k-2}{n-1}\binom{k}{r}2^{-k} = \\ = 2^{2-2n}\sum_{j=0}^{n-r-1}\binom{2n-r-j-2}{n-1}\binom{r+j}{r} \\ = 2^{2-2n}\sum_{j=0}^{n-r-1}\binom{2n-r-j-2}{n-r-j-1}\binom{r+j}{j}.$$

Используя равенства

$$\binom{-a}{i} = (-1)^i\binom{a+i-1}{i} \quad \text{и} \quad \sum_{i=0}^{s}\binom{a}{i}\binom{b}{s-i} = \binom{a+b}{s},$$

получаем

$$C_{r,2n} = 2^{2-2n}\sum_{j=0}^{n-r-1}(-1)^{n-r-j-1}\binom{-n}{n-r-j-1}(-1)^j\binom{-(r+1)}{j} = \\ = 2^{2-2n}(-1)^{n-r-1}\binom{-(n+r+1)}{n-r-1} = 2^{2-2n}\binom{2n-1}{n-r-1}.$$

ЗАДАЧИ

1. Пусть

$$\mathbf{P} = \left\| \begin{matrix} 1-a & a \\ b & 1-b \end{matrix} \right\|, \qquad 0 < a,\ b < 1.$$

Доказать, что

$$\mathbf{P}^n = \frac{1}{a+b}\left\| \begin{matrix} b & a \\ b & a \end{matrix} \right\| + \frac{(1-a-b)^n}{a+b}\left\| \begin{matrix} a & -a \\ -b & b \end{matrix} \right\|.$$

2. Рассмотреть конечное случайное блуждание по множеству чисел 0, 1, 2, ..., N, матрица одношаговых переходных вероятностей которого имеет вид

$$\mathbf{P} = \left\| \begin{matrix} 0 & 1 & 0 & 0 & \dots & & \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \dots & & \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \dots & & \\ & \cdot & & \cdot & \cdot & & \\ & & \cdot & & \cdot & \cdot & \\ & & & \cdot & & \cdot & \cdot \\ \dots & & & & \frac{1}{2} & 0 & \frac{1}{2} \\ \dots & & & & & 1 & 0 \end{matrix} \right\|.$$

Найти формулу для вероятностей перехода за r шагов с помощью метода ортогональных полиномов.

Ответ: Тригонометрические полиномы

$$Q_n(x) = \cos n\theta, \qquad x = \cos\theta,$$

удовлетворяют рекуррентным соотношениям

$$xQ_n(x) = \frac{1}{2} Q_{n-1}(x) + \frac{1}{2} Q_{n+1}(x) \quad \text{при} \quad n = 1, 2, \dots, N-1.$$

Кроме того,

$$xQ_0(x) = Q_1(x).$$

Для того чтобы удовлетворялось уравнение

$$xQ_N(x) = Q_{N-1}(x),$$

должно выполняться равенство

$$\cos\theta \cos N\theta = \cos(N-1)\theta,$$

т. е.

$$\sin N\theta \sin\theta = 0.$$

Это означает, что $\theta = k\pi/N$, $k = 0, 1, 2, \dots, 2N-1$. Итак, при $\theta = k\pi/N$ мы имеем

$$xQ_n(x) = \sum_{m=0}^{N} P_{nm} Q_m(x),$$

где $\mathbf{P} = \|P_{nm}\|$. Из этого уравнения также следует, что при $\theta = k\pi/N$

$$x^r Q_n(x) = \sum_{m=0}^{N} P_{nm}^{(r)} Q_m(x), \tag{*}$$

где $\mathbf{P}^r = \| P_{nm}^{(r)} \|$. Покажем, что $Q_n(x)$ образуют ортогональную систему на конечном множестве $x_k = \cos k\pi/N$, $k = 0, 1, 2, \ldots, 2N-1$. Действительно,

$$\sum_{k=0}^{2N-1} Q_n(x_k) Q_m(x_k) = \sum_{k=0}^{2N-1} \cos\frac{nk\pi}{N} \cos\frac{mk\pi}{N} =$$

$$= \frac{1}{2} \sum_{k=0}^{2N} \left[\cos\frac{(n-m)k\pi}{N} + \cos\frac{(n+m)k\pi}{N}\right] =$$

$$= \frac{1}{2} \operatorname{Re} \sum_{k=0}^{2N-1} \left[\exp\left\{\frac{(n-m)k\pi i}{N}\right\} + \exp\left\{\frac{(n+m)k\pi i}{N}\right\}\right] = \qquad (\text{здесь } i = \sqrt{-1})$$

$$= \frac{1}{2} \operatorname{Re}\left\{\frac{1-\exp[2(n-m)\pi i]}{1-\exp[(n-m)\pi i/N]} + \frac{1-[\exp 2(n+m)\pi i]}{1-[\exp(n+m)\pi i/N]}\right\} = 0, \quad \text{если } n \neq m.$$

Кроме того,

$$\sum_{k=0}^{2N-1} Q_n^2(x_k) = \frac{1}{2} \sum_{k=0}^{2N-1} \left(1 + \cos\frac{2nk\pi}{N}\right) = N.$$

Умножая обе части соотношения (*) на $Q_{m_0}(x)$, суммируя по множеству $x_k = \cos k\pi/N$, $k = 0, 1, \ldots, 2N-1$, и используя соотношения ортогональности, получаем

$$\sum_{k=0}^{2N-1} x_k^r Q_n(x_k) Q_{m_0}(x_0) = P_{nm_0}^{(r)} \cdot N,$$

откуда, опустив нуль в индексе, имеем

$$P_{nm}^{(r)} = \frac{1}{N} \sum_{k=0}^{2N-1} \cos^r\frac{k\pi}{N} \cos\frac{nk\pi}{N} \cos\frac{mk\pi}{N}, \qquad n, m = 0, 1, \ldots, N.$$

3. Рассмотреть процесс, описанный в предыдущем упражнении, но отличающийся от последнего тем, что состояния 0 и N являются поглощающими экранами. В этом случае $(N-1) \times (N-1)$-матрица одношаговых переходных вероятностей, соответствующая невозвратным состояниям, имеет вид

$$\mathbf{P} = \begin{Vmatrix} 0 & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{2} & 0 \end{Vmatrix}.$$

Найти формулы вероятностей перехода за r шагов при условии, что поглощение не имело места.

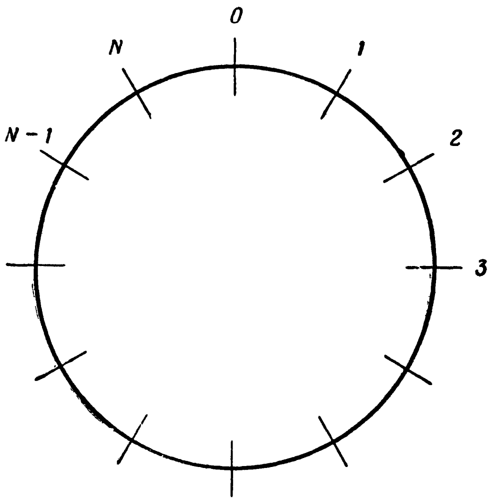
Указание: В качестве ортогональных полиномов взять $Q_n(x) = \sin n\theta$, $x = \cos\theta$, где θ пробегает значения $\theta = k\pi/N$, $k = 0, 1, \ldots, 2N-1$.
Ответ:

$$P_{nm}^{(r)} = \frac{1}{N}\sum_{k=0}^{2N-1}[\cos k\pi/N]^r \sin nk\pi/N \sin mk\pi/N \qquad (n,\ m = 1, 2, \ldots, N-1).$$

4. Рассмотреть случайное блуждание на окружности, имеющее $N+1$ состояний, симметрично расположенных на этой окружности. Матрица одношаговых переходных вероятностей процесса имеет вид

$$\mathbf{P} = \begin{Vmatrix} 0 & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \ldots & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \ldots & 0 & \frac{1}{2} & 0 \end{Vmatrix}.$$

Найти выражение для r-шаговых переходных вероятностей.



Ответ: Пусть

$$Z(\theta) = \frac{1}{2}e^{i\theta} + \frac{1}{2}e^{-i\theta} = \cos\theta$$

и

$$Q_n(\theta) = e^{in\theta},$$

Эти величины удовлетворяют рекуррентному соотношению

$$Z(\theta)\,Q_n(\theta)=\frac{1}{2}Q_{n+1}(\theta)+\frac{1}{2}Q_{n-1}(\theta)\qquad \text{при}\qquad n=1,\,2,\,\ldots,\,N-1.$$

Кроме этого, требуется выполнение еще двух соотношений:

$$Z(\theta)\,Q_0(\theta)=\frac{1}{2}Q_1(\theta)+\frac{1}{2}Q_N(\theta)$$

и

$$Z(\theta)\,Q_N(\theta)=\frac{1}{2}Q_{N-1}(\theta)+\frac{1}{2}Q_0(\theta).$$

Они выполняются при одном лишь условии

$$1=e^{i(N+1)\theta},$$

т. е. при

$$\theta=\frac{2\pi k}{N+1},\qquad k=0,\,1,\,\ldots,\,N.$$

Итак, при $\theta=2\pi k/(N+1)$

$$Z(\theta)\,Q_n(\theta)=\sum_{m=0}^{N}P_{nm}Q_m(\theta),\qquad n,\,m=0,\,1,\,\ldots,\,N,$$

откуда получаем

$$Z^r(\theta)\,Q_n(\theta)=\sum_{m=0}^{N}P_{nm}^{(r)}Q_m(\theta),\qquad \theta=\frac{2\pi k}{N+1}. \qquad (*)$$

Далее, функции $Q_n(\theta)$ образуют ортогональную систему на конечном множестве $\theta_k=2\pi k/(N+1)$ $(k=0,\,1,\,\ldots,\,N;\ n=0,\,1,\,\ldots,\,N)$, так как при $n\neq m$

$$\sum_{k=0}^{N}Q_n(\theta)\,\overline{Q_m(\theta)}=\sum_{k=0}^{N}\exp\left[\frac{i2\pi k\,(n-m)}{N+1}\right]=\frac{1-\exp\,[2\pi i\,(n-m)]}{1-\exp\,[2\pi i\,(n-m)/(N+1)]}=0.$$

Кроме того,

$$\sum_{k=0}^{N}|\,Q_n(\theta_k)\,|^2=N+1.$$

Умножая обе части соотношения (*) на $\overline{Q_{m_0}(\theta)}$, суммируя по всем $\theta_k=2\pi k/(N+1)$, $k=0,\,1,\,\ldots,\,N$, и пользуясь свойством ортогональности, получаем

$$\sum_{k=0}^{N}Z^r(\theta_k)\,Q_n(\theta_k)\,\overline{Q_{m_0}(\theta_k)}=(N+1)\,P_{nm_0}^{(r)},$$

откуда

$$P_{nm}^{(r)}=\frac{1}{N+1}\sum_{k=0}^{N}\cos^r\frac{2\pi k}{N+1}\exp\left[\frac{2\pi k\,(n-m)\,i}{N+1}\right],\qquad n,\,m=0,\,1,\,2,\,\ldots,\,N.$$

5. Рассмотреть случайное блуждание на окружности (см. задачу 4) с матрицей одношаговых переходных вероятностей вида

$$\mathbf{P} = \left\| \begin{matrix} 0 & p & 0 \dots 0 & q \\ q & 0 & p \dots 0 & 0 \\ 0 & q & 0 \dots 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \dots 0 & p \\ p & 0 & 0 \dots q & 0 \end{matrix} \right\|.$$

Найти выражение для r-шаговых переходных вероятностей.

Указание: В решении к задаче 4 положить $Z(\theta) = re^{i\theta} + qe^{-i\theta}$.

Ответ:

$$P_{nm}^{(r)} = \frac{1}{N+1} \sum_{k=0}^{N} \left(p \exp\left(\frac{2\pi ik}{N+1}\right) + q \exp\left(-\frac{2\pi ik}{N+1}\right)\right)^r \exp\left[\frac{2\pi ik(n-m)}{N+1}\right],$$

$$n, m = 0, 1, 2, \dots, N.$$

6. Доказать, что

$$\frac{1}{\pi}\int_0^{2\pi} \cos^n\theta \cos k\theta \cos l\theta \, d\theta =$$

$$= \begin{cases} \left[\binom{n}{(n+k-l)/2} + \binom{n}{(n+k+l)/2}\right]\dfrac{1}{2^{n+1}}, & \text{если } n+k+l \text{ четно,} \\ 0 & \text{в противном случае.} \end{cases}$$

(Этот интеграл есть выражение для P_{kl}^n при $l \neq 0$ из формулы (4.7).) Найти предел $\lim\limits_{n\to\infty} \sqrt{n}\, P_{kl}^n$.

Указание: Пусть k — неотрицательное целое число. Доказать, что

$$\frac{1}{2\pi}\int_0^{2\pi} \cos^n\theta \cos k\theta \, d\theta = \begin{cases} 0, & \text{если } n+k \text{ нечетно,} \\ 2^{-n}\dbinom{n}{(n+k)/2} & \text{в противном случае.} \end{cases}$$

Для этого воспользоваться тождеством

$$\cos^{n+1}\theta \cos(k-1)\theta = \cos^n\theta\cos k\theta + \cos^n\theta \sin(k-1)\theta \sin\theta,$$

проинтегрировать по частям и получить рекуррентное соотношение

$$\int \cos^n\theta\cos k\theta\, d\theta = \frac{n-k+2}{n+1}\int \cos^{n+1}\theta\cos(k-1)\theta\, d\theta.$$

Ответ: $\sqrt{2/\pi}$.

7. Найти значение интеграла

$$P_{kl}^n = \frac{1}{\pi}\int_0^{2\pi} \cos^n\theta \sin(k+1)\theta \sin(l+1)\theta\, d\theta.$$

Указание: Воспользоваться решением задачи 6.
Ответ:

$$2^{-n-1}\left[\binom{n}{(n-k+l)/2}-\binom{n}{(n+k+l+2)/2}\right].$$

8. Пусть $\mathbf{P}=\|P_{ij}\|$ обозначает матрицу переходных вероятностей конечной марковской цепи $\{X_n\}_0^\infty$, имеющей три класса состояний $\{0\}$, $\{1, 2, \ldots, N-1\}$ и $\{N\}$, из которых 0 и N — поглощающие состояния, а остальные — невозвратные состояния. Введем семейство матриц

$$\mathbf{P}(\theta)=\|P_{ij}e^{\theta(j-i)}\|=\|P_{ij}(\theta)\| \qquad (\theta - \text{любое действительное число})$$

и производящую функцию моментов $M^{(t)}(\theta\mid k)$ по формуле

$$M^{(t)}(\theta\mid X_0=k)=\mathrm{M}\left[\exp\left(\theta\left(X_t-X_0\right)\right)\mid X_0=k\right]=\mathbf{e}'_k\mathbf{P}^t(\theta)\,\mathbf{e}=\sum_{j=0}^{N}P^t_{kj}(\theta),$$

где $\mathbf{e}'_k$ обозначает вектор-строку $(0, \ldots, 1, 0, \ldots, 0)$ с единицей на k-м месте, $\mathbf{e}$ обозначает $N+1$-мерный вектор-столбец, все элементы которого равны единице, и $\mathbf{P}^t(\theta)$ — t-ю степень матрицы $\mathbf{P}(\theta)$. Пусть π_{k0} и $\pi_{kN}=1-\pi_{k0}$ $(1\leqslant k\leqslant N-1)$ обозначают вероятности поглощения состояниями 0 и N соответственно при начальном состоянии k. Доказать, что

$$\lim_{t\to\infty}M^{(t)}(\theta\mid k)=\pi_{k0}e^{-k\theta}+\pi_{kN}e^{(N-k)\theta}.$$

9. В условиях предыдущей задачи предположить, что существуют действительные числа a и b, такие, что

$$M(a\mid k)\leqslant 1\leqslant M(b\mid k) \qquad (k=1, 2, \ldots, N-1), \qquad (+)$$

где $M(\theta\mid k)=M^{(1)}(\theta\mid k)$. Доказать, что

$$\frac{e^{kb}-1}{e^{Nb}-1}\leqslant\pi_{kN}\leqslant\frac{e^{ka}-1}{e^{Na}-1}.$$

Указание: Показать, что $M^{(t)}(a\mid k)\leqslant 1\leqslant M^{(t)}(b\mid k)$ $(k=1, 2, \ldots, N-1)$, и затем воспользоваться результатом предыдущей задачи.

10. Рассмотреть марковскую цепь с состояниями $(0, 1, \ldots, N)$ и переходными вероятностями

$$P_{ij}=\binom{N}{j}p_i^j(1-p_i)^{N-j}, \quad \text{где} \quad p_i=\frac{(1+\sigma)\,i}{N+\sigma i}, \qquad 0<\sigma<1$$

(см. пример Ж из § 2 гл. 2). Показать, что

$$M(\theta\mid k)=e^{-k\theta}\left(p_ke^{\theta}+1-p_k\right)^N$$

(определение $M(\theta/k)$ см. в задаче 8).

11. Для марковской цепи из предыдущей задачи показать, что числа

$$a=\ln\frac{1}{1+\sigma}, \qquad b=\ln\frac{1-\sigma}{1+\sigma}$$

удовлетворяют условию (+) задачи 9, и затем получить оценки

$$\frac{[(1-\sigma)/(1+\sigma)]^k-1}{[(1-\sigma)/(1+\sigma)]^N-1}\leqslant\pi_{kN}\leqslant\frac{1/(1+\sigma)^k-1}{1/(1+\sigma)^N-1}.$$

12. Рассмотреть конечную марковскую цепь $\{X_n\}$, $n = 0, 1, \ldots$, с двумя классами состояний, один из которых представляет собой поглощающее состояние. Для определенности пусть поглощающий экран расположен в нуле, а $i = 1, 2, \ldots, N$ представляют состояния другого класса. Пусть собственные значения матрицы переходных вероятностей занумерованы в порядке убывания их абсолютных значений: $\lambda_0 = 1 > |\lambda_2| > |\lambda_3| \geqslant |\lambda_4| \geqslant \ldots$ (Подчеркнем, что мы предполагаем условие $1 > |\lambda_2| > |\lambda_3|$.) Пусть b_j — предельная вероятность пребывания в состоянии j при условиях, что поглощения состоянием 0 не произошло, а начальным состоянием было i, т. е. $b_j = \lim\limits_{n\to\infty} \mathrm{P}\{X(n) = j \mid X(n) \neq 0,\ X(0) = i\}$. (См. стр. 116.) Оценить скорость сходимости к нулю величин

$$\frac{P_{ij}^n}{1 - P_{i0}^n} - b_j, \qquad j = 1, 2, \ldots, N \quad (i \geqslant 1).$$

Ответ: Скорость сходимости имеет порядок $|\lambda_3/\lambda_2|^n$.

13. Некоторые соотношения для бросаний монеты. Пусть с. в. $\{X_i\}$, $1 \leqslant i < \infty$, независимы, одинаково распределены и

$$\mathrm{P}\{X_i = 1\} = \mathrm{P}\{X_i = -1\} = {}^1/_2.$$

Пусть $S_n = \sum\limits_{i=1}^{n} X_i$ при $1 \leqslant n < \infty$ и

$$P(m, n) = \mathrm{P}\{S_{2j} = 0 \text{ для некоторого } j,\ m \leqslant j < m + n\}.$$

Доказать, что $P(m, n) + P(n, m) = 1$ при $m \geqslant 1$, $n \geqslant 1$.

Указание: Воспользоваться равенством $\mathrm{P}\{S_{2n} = 0\} = \mathrm{P}\{S_1 \neq 0,\ S_2 \neq 0, \ldots \ldots, S_{2n} = 0\}$ и, предположив, что доказываемый результат справедлив при $m = k$ и произвольном $n \geqslant 1$, обосновать следующие равенства:

$$\begin{aligned}
1 - P(k+1, n) &= \mathrm{P}\{S_{2j} \neq 0 \text{ при } k \leqslant j < k+n+1\} + \\
&\qquad + \mathrm{P}\{S_{2k} = 0 \text{ и } S_{2j} \neq 0 \text{ при } k+1 \leqslant j < k+n+1\} = \\
&= \mathrm{P}\{S_{2j} \neq 0 \text{ при } k \leqslant j < k+n+1\} + \\
&\qquad + \mathrm{P}\{S_{2k} = 0\}\,\mathrm{P}\{S_{2j} \neq 0 \text{ при } 1 \leqslant j < n+1\} = \\
&= \mathrm{P}\{S_{2j} = 0 \text{ при некотором } j,\ n+1 \leqslant j < k+n+1\} + \\
&\qquad + \mathrm{P}\{S_{2j} \neq 0 \text{ при } 1 \leqslant j < k+1\}\,\mathrm{P}\{S_{2n} = 0\} = \\
&= \mathrm{P}\{S_{2j} = 0 \text{ при некотором } j,\ n+1 \leqslant j < k+n+1\} + \\
&\qquad + \mathrm{P}\{S_{2n} = 0 \text{ и } S_{2j} \neq 0 \text{ при } n+1 \leqslant j < k+n+1\} = \\
&= P(n, k+1).
\end{aligned}$$

НЕКОТОРЫЕ ЭЛЕМЕНТАРНЫЕ ЗАДАЧИ

1. Проверить утверждение теоремы 2.1 на марковской матрице $\begin{Vmatrix} p & 1-p \\ q & 1-q \end{Vmatrix}$, где $0 \leqslant p \leqslant 1$ и $0 \leqslant q \leqslant 1$. При каких значениях p и q марковская цепь с двумя состояниями, которой соответствует эта матрица, имеет ровно один возвратный класс?

2. Пусть $\mathbf{P}$ и $\mathbf{Q}$ — конечные матрицы переходных вероятностей порядка n, такие, что $\mathbf{PQ} = \mathbf{I} = \mathbf{QP}$. Показать, что матрицы $\mathbf{P}$ и $\mathbf{Q}$ являются матрицами перестановок, т. е. матрицами с одним отличным от нуля элементом в каждом столбце и каждой строке.

3. Пусть $\{X_n, n \geqslant 0\}$ — марковская цепь с двумя состояниями 0 и 1, причем $P_{00} = 1 - \alpha$, $P_{01} = \alpha$, $P_{11} = 1 - \beta$, $P_{10} = \beta (0 < \alpha, \beta < 1)$. Пусть N есть значение индекса $n \geqslant 1$, для которого $X_{n-1} = X_n = 0$, и пусть $d_0 = \mathrm{M}(N \mid X_0 = 0)$. Показать, что

$$d_0 = 1 + \frac{\alpha}{1-\alpha} \frac{1+\beta}{\beta}.$$

Указание: Установить связь между d_0 и $d_1 = \mathrm{M}(N \mid X_0 = 1)$.

ЗАМЕЧАНИЯ

Алгебраическим методам исследования марковских цепей посвящена гл. 16 книги Феллера [1]. Эти же вопросы освещены в книге Кемени и Снелла [2].

ЛИТЕРАТУРА

1. Феллер В., Введение в теорию вероятностей и ее приложения, т. 1, «Мир», 1964.
2. Kemeny J. G., Snell J. L., Finite Markov Chains, Van Nostrand, Princeton, New Jersey, 1960.

Глава 5

ТЕОРЕМЫ ОБ ОТНОШЕНИЯХ ПЕРЕХОДНЫХ ВЕРОЯТНОСТЕЙ И ИХ ПРИЛОЖЕНИЯ

§ 1. ВЕРОЯТНОСТИ ПЕРЕХОДА С ЗАПРЕЩЕНИЕМ

В возвратной нулевой неприводимой марковской цепи среднее переходных вероятностей $\frac{1}{n}\sum_{m=0}^{n-1} P_{ij}^m$ стремится к нулю при стремлении n к ∞. Но и при этом возвращение в любое данное состояние происходит с достоверностью. Другими словами, относительная частота посещений любого данного состояния стремится к нулю с ростом времени, но тем не менее процесс находится в каждом состоянии бесконечное число раз с вероятностью 1. Имеет смысл рассмотреть число пребываний в данном состоянии i в отношении к числу пребываний в некотором другом состоянии j при бесконечном возрастании числа переходов. Для этой цели полезно ввести «вероятности перехода с запрещением»:

$${}_kP_{ij}^n = \mathrm{P}\{X_n = j,\ X_\nu \neq k,\ \nu = 1,\ 2,\ \ldots,\ n-1 \mid X_0 = i\} \text{ при } k \neq j,\ n \geqslant 1.$$

Здесь в правой части стоит событие, состоящее в том, что процесс перейдет из состояния i в состояние j за n шагов, ни разу не попав при переходе в состояние k. Состояние k в этом смысле называют запрещенным. Аналогично при $k \neq j$, $n \geqslant 1$ определим

$${}_kf_{ij}^n = \mathrm{P}\left\{X_n = j,\ X_\nu \neq j,\ X_\nu \neq k, \quad \nu = 1,\ 2,\ \ldots,\ n-1 \mid X_0 = i\right\},$$

вероятность того, что процесс, исходя из состояния i, на n-м шаге впервые достигнет состояния j, не попав при переходе в состояние k. Для удобства положим при $k \neq i$

$${}_kP_{ij}^0 = \begin{cases} 0, & \text{если } i \neq j, \\ 1, & \text{если } i = j, \end{cases}$$

и

$${}_kf_{ij}^0 \equiv 0 \quad \text{для всех} \quad i,\ j.$$

Далее, при $i \neq j$ и $n \geqslant 0$ имеет место следующая важная формула:

$$f_{ij}^n = \sum_{\nu=0}^{n} {}_jP_{ii}^\nu f_{ij}^{n-\nu}, \tag{1.1}$$

основывающаяся на разложении события, состоящего в первом достижении состояния i из состояния j на n-м шаге, на n несовместных событий, состоящих в возвращении в состояние i на ν-м шаге при запрещенном состоянии j и последующем первом достижении состояния j за $n-\nu$ шагов при запрещенном состоянии i, $\nu = 0, 1, \ldots, n-1$.

При выводе (1.1) ключевую роль играет классификация траекторий по последнему моменту нахождения процесса в состоянии i, предшествующему моменту n. Упомянем здесь, что при выводе формулы (5.1) гл. 2 траектории подразделялись по моменту первого наступления этого события.

Вообще соотношения, связанные с запрещенными состояниями, чаще всего устанавливаются с помощью рассмотрения первого или последнего появления некоторого события. Эта двойственность между первым и последним играет важную роль во многих разделах теории вероятностей. Наиболее яркой иллюстрацией этого служит теория сумм независимых одинаково распределенных случайных величин, где имеет место полная эквивалентность между понятием первого и понятием последнего момента.

Соотношение (1.1) можно заменить эквивалентным ему уравнением для производящих функций. Последнее можно вывести методом, аналогичным примененному нами при выводе соотношения (5.10) из соотношения (5.9). Итак, сначала определим производящие функции:

$$ {}_if_{ij}(s) = \sum_{n=1}^{\infty} {}_if_{ij}^n s^n, $$

$$ {}_jP_{ii}(s) = \sum_{n=0}^{\infty} {}_jP_{ii}^n s^n. $$

Затем, так как соотношение (1.1) является сверткой, при $i \neq j$ мы получаем

$$ f_{ij}(s) = {}_jP_{ii}(s)\, {}_if_{ij}(s). \tag{1.2} $$

Поскольку

$$ \sum_{n=1}^{\infty} f_{ij}^n \leqslant 1 \quad \text{и} \quad \sum_{n=1}^{\infty} {}_if_{ij}^n \leqslant 1, $$

в силу части (а) леммы Абеля (лемма 5.1 гл. 2) имеют место следующие соотношения:

$$ \lim_{s\to 1-} f_{ij}(s) = \sum_{n=1}^{\infty} f_{ij}^n, \quad \lim_{s\to 1-} {}_if_{ij}(s) = \sum_{n=1}^{\infty} {}_if_{ij}^n. $$

Если состояния i и j сообщаются, то существует такое целое число $n \geqslant 1$, что $f_{ij}^n > 0$, и тогда в силу (1.1)

$$ {}_if_{ij}^{n-\nu} > 0 \quad \text{для некоторого} \quad \nu = 0, 1, \ldots, n-1. $$

Последнее неравенство позволяет утверждать, что

$$\sum_{n=1}^{\infty} {}_i f_{ij}^n > 0,$$

а следовательно, и

$$\lim_{s \to 1-} {}_j P_{ii}(s) = \frac{\sum\limits_{n=1}^{\infty} f_{ij}^n}{\sum\limits_{n=1}^{\infty} {}_i f_{ij}^n} < \infty.$$

Наконец, в силу части (б) леммы Абеля имеем

$${}_j P_{ii}^* = \sum_{n=0}^{\infty} {}_j P_{ii}^n = \lim_{s \to 1-} {}_j P_{ii}(s) < \infty.$$

§ 2. ТЕОРЕМЫ ОБ ОТНОШЕНИЯХ

Для доказательства двух теорем этого параграфа нам понадобится следующая

Лемма 2.1. *Пусть*

$$c_n = \sum_{\nu=0}^{n} a_{n-\nu} b_\nu, \qquad n = 0,\ 1,\ 2,\ \ldots,$$

где $0 \leqslant a_n \leqslant K$ *(K — положительная константа) и* $\sum\limits_{n=0}^{\infty} a_n = \infty$. *Тогда если предел* $\lim\limits_{n\to\infty} b_n = b$ *конечен, то*

$$\lim_{n \to \infty} \frac{c_n}{\sum\limits_{\nu=0}^{n} a_\nu} = b.$$

Доказательство. Заметим, что $\sum\limits_{\nu=0}^{n} a_{n-\nu} = \sum\limits_{\nu=0}^{n} a_\nu$. Следовательно,

$$\frac{c_n}{\sum\limits_{\nu=0}^{n} a_\nu} - b = \frac{\sum\limits_{\nu=0}^{n} a_{n-\nu} b_\nu}{\sum\limits_{\nu=0}^{n} a_\nu} - b = \frac{\sum\limits_{\nu=0}^{n} a_{n-\nu}(b_\nu - b)}{\sum\limits_{\nu=0}^{n} a_{n-\nu}}.$$

Поскольку b_n стремятся к конечному пределу b, существует такое $M > 0$, что $|b_n| < M$ при всех $n \geqslant 0$. Выберем теперь такое $N = N(\varepsilon)$, что

$$|b_n - b| < \varepsilon \quad \text{для всех} \quad n \geqslant N,$$

Тогда при $n \geqslant N$ имеем

$$\left|\frac{c_n}{\sum\limits_{\nu=0}^{n} a_\nu} - b\right| \leqslant \left|\frac{\sum\limits_{\nu=0}^{N-1} a_{n-\nu}(b_\nu - b)}{\sum\limits_{\nu=0}^{n} a_\nu}\right| + \left|\frac{\sum\limits_{\nu=N}^{n} a_{n-\nu}(b_\nu - b)}{\sum\limits_{\nu=0}^{n} a_\nu}\right| \leqslant$$

$$\leqslant 2M \left|\frac{\sum\limits_{\nu=0}^{N-1} a_{n-\nu}}{\sum\limits_{\nu=0}^{n} a_\nu}\right| + \varepsilon \left|\frac{\sum\limits_{\nu=N}^{n} a_{n-\nu}}{\sum\limits_{\nu=0}^{n} a_\nu}\right| \leqslant \frac{2MNK}{\sum\limits_{\nu=0}^{n} a_\nu} + \varepsilon.$$

Так как $\varepsilon > 0$ произвольно, то, устремляя n к ∞, мы получаем доказываемый результат.

Приведем теперь три соотношения, подобных соотношению (1.1):

$$_kP_{ij}^n = \sum_{\nu=0}^{n} {_kf_{ij}^\nu}\, {_kP_{jj}^{n-\nu}}, \tag{2.1}$$

$$P_{ij}^n = \sum_{\nu=0}^{n} P_{ii}^\nu\, {_iP_{ij}^{n-\nu}}, \tag{2.2}$$

$$P_{ij}^n = \sum_{\nu=0}^{n} f_{ij}^\nu P_{jj}^{n-\nu} \tag{2.3}$$

при $k \neq j$, $i \neq j$ и $n \geqslant 0$. Последнее соотношение совпадает с соотношением (5.9) гл. 2. Вывод их аналогичен выводу соотношения (1.1).

Теорема 2.1. *Пусть i и j — произвольные состояния, причем состояние j — возвратное; тогда*

$$\lim_{m\to\infty} \frac{\sum\limits_{n=0}^{m} P_{ij}^n}{\sum\limits_{n=0}^{m} P_{jj}^n} = f_{ij}^*,$$

где

$$f_{ij}^* = \sum_{n=1}^{\infty} f_{ij}^n. \tag{2.4}$$

Доказательство. Из (2.3) имеем

$$\sum_{n=0}^{m} P_{ij}^n = \sum_{n=0}^{m} \sum_{\nu=0}^{n} f_{ij}^\nu P_{jj}^{n-\nu} = \sum_{n=0}^{m} \sum_{\mu=0}^{n} f_{ij}^{n-\mu} P_{jj}^\mu = \sum_{n=0}^{m} \sum_{\mu=0}^{\infty} f_{ij}^{n-\mu} P_{jj}^\mu,$$

поскольку $f_{ij}^{n-\mu} = 0$ при $\mu > n$ (по определению), Поскольку на самом деле обе суммы конечны, порядок суммирования можно поменять:

$$\sum_{n=0}^{m} P_{ij}^{n} = \sum_{\mu=0}^{\infty} P_{jj}^{\mu} \sum_{n=0}^{m} f_{ij}^{n-\mu} = \sum_{\mu=0}^{\infty} P_{jj}^{\mu} \sum_{r=0}^{m-\mu} f_{ij}^{r}.$$

Положим

$$F_{ij}^{m} = \sum_{r=0}^{m} f_{ij}^{r} \quad \text{при} \quad m = 0,\ 1,\ 2,\ \ldots$$

и

$$F_{ij}^{m} = 0 \qquad \text{при} \quad m = -1,\ -2,\ \ldots,$$

тогда

$$\sum_{n=0}^{m} P_{ij}^{n} = \sum_{\mu=0}^{\infty} F_{ij}^{m-\mu} P_{jj}^{\mu} = \sum_{\mu=0}^{m} F_{ij}^{m-\mu} P_{jj}^{\mu} = \sum_{\nu=0}^{m} P_{jj}^{m-\nu} F_{ij}^{\nu}.$$

Применим теперь лемму 2.1, положив

$$a_\nu = P_{jj}^{\nu}, \quad b_\nu = F_{ij}^{\nu} \quad \text{и} \quad c_m = \sum_{n=0}^{m} P_{ij}^{n}.$$

Условия леммы выполнены, так как

$$0 \leqslant P_{jj}^{\nu} \leqslant 1 \quad \text{и} \quad \sum_{\nu=0}^{\infty} P_{jj}^{\nu} = \infty;$$

последнее в силу того, что состояние j — возвратное. Поскольку

$$\lim_{m\to\infty} b_m = \lim_{m\to\infty} F_{ij}^{m} = \sum_{r=0}^{\infty} f_{ij}^{r} = f_{ij}^{*},$$

то в силу леммы имеем

$$\lim_{m\to\infty} \frac{\sum_{n=0}^{m} P_{ij}^{n}}{\sum_{\nu=0}^{m} P_{jj}^{\nu}} = f_{ij}^{*}$$

и теорема доказана. ■

Перепишем теперь соотношение (2.1) для $k = i \neq j$:

$$ {}_{i}P_{ij}^{n} = \sum_{\nu=0}^{n} {}_{i}f_{ij}^{\nu}\, {}_{i}P_{jj}^{n-\nu}, \qquad n \geqslant 0. \tag{2.5}$$

Переходя к соответствующим производящим функциям, получаем

$${}_{i}P_{ij}(s) = {}_{i}f_{ij}(s)\, {}_{i}P_{jj}(s).$$

Ранее мы уже показали, что ${}_iP^*_{jj}=\sum_{n=0}^{\infty}{}_iP^n_{jj}<\infty$, если состояния i и j — сообщающиеся. Поэтому по лемме Абеля

$$\lim_{s\to 1-}{}_iP_{ij}(s)=\lim_{s\to 1-}{}_if_{ij}(s)\cdot\lim_{s\to 1-}{}_iP_{jj}(s)<\infty$$

и окончательно имеем

$${}_iP^*_{ij}=\sum_{n=0}^{\infty}{}_iP^n_{ij}=\lim_{s\to 1-}{}_iP_{ij}(s)<\infty. \tag{2.6}$$

Теорема 2.2. *Если i и j принадлежат одному и тому же возвратному классу состояний, то*

$$\lim_{m\to\infty}\frac{\sum\limits_{n=0}^{m}P^n_{ij}}{\sum\limits_{n=0}^{m}P^n_{ii}}={}_iP^*_{ij}.$$

Замечание. Для $i\neq j$ введем случайные величины:

$$U_n=U(i,j,n)=\begin{cases}1, & \text{если процесс, исходя из состояния } i, \text{ за } n \text{ шагов попадет в состояние } j, \text{ при переходе не возвратившись в состояние } i;\\ 0 & \text{в противном случае.}\end{cases}$$

Тогда $M(U_n)={}_iP^n_{ij}$ и

$$M\left(\sum_{n=1}^{\infty}U_n\right)=\sum_{n=1}^{\infty}{}_iP^n_{ij}={}_iP^*_{ij}.$$

Таким образом, в условиях теоремы 2.2 вероятность ${}_iP^*_{ij}$ есть среднее число попаданий в состояние j между последовательными возвращениями в состояние i.

Доказательство. В силу соотношения (2.2) имеем

$$\sum_{n=0}^{m}P^n_{ij}=\sum_{n=0}^{m}\sum_{\nu=0}^{n}P^{\nu}_{ii}\,{}_iP^{n-\nu}_{ij}=\sum_{n=0}^{m}\sum_{\nu=0}^{\infty}P^{\nu}_{ii}\,{}_iP^{n-\nu}_{ij},$$

так как ${}_iP^{n-\nu}_{ij}=0$ при $\nu>n$. Меняя порядок суммирования, получаем

$$\sum_{n=0}^{m}P^n_{ij}=\sum_{\nu=0}^{\infty}P^{\nu}_{ii}\sum_{n=0}^{m}{}_iP^{n-\nu}_{ij}=\sum_{\nu=0}^{\infty}P^{\nu}_{ii}\,{}_i\tilde{P}^{m-\nu}_{ij},$$

где

$${}_i\tilde{P}^m_{ij} = \sum_{\nu=0}^{m} {}_iP^{\nu}_{ij}, \qquad m = 0,\ 1,\ 2,\ \ldots$$

и

$${}_i\tilde{P}^m_{ij} = 0, \qquad m = -1,\ -2,\ \ldots.$$

Тогда

$$\sum_{n=0}^{m} P^n_{ij} = \sum_{\nu=0}^{m} P^{\nu}_{ii}\, {}_i\tilde{P}^{m-\nu}_{ij} = \sum_{\nu=0}^{m} P^{m-\nu}_{ii}\, {}_i\tilde{P}^{\nu}_{ij}.$$

Теперь мы можем применить лемму 2.1, положив $a_\nu = P^{\nu}_{ii}$, $b_\nu = {}_i\tilde{P}^{\nu}_{ij}$ и $c_m = \sum_{n=0}^{m} P^n_{ij}$, так как $|P^{\nu}_{ii}| \leqslant 1$ и $\sum_{\nu=0}^{\infty} P^{\nu}_{ii} = \infty$ (состояние i возвратное). Но

$$\lim_{m\to\infty} b_m = \lim_{m\to\infty} {}_i\tilde{P}^m_{ij} = \sum_{\nu=0}^{\infty} {}_iP^{\nu}_{ij} = {}_iP^*_{ij},$$

а, следовательно, в силу леммы 2.1 имеет место равенство

$$\lim_{m\to\infty} \frac{\sum_{n=0}^{m} P^n_{ij}}{\sum_{n=0}^{m} P^n_{ii}} = {}_iP^*_{ij}.$$

Теорема доказана. ■

Если состояния i и j — сообщающиеся, то мы можем записать

$$\lim_{m\to\infty} \frac{\sum_{n=0}^{m} P^n_{jj}}{\sum_{n=0}^{m} P^n_{ii}} = \lim_{m\to\infty} \frac{\sum_{n=0}^{m} P^n_{jj}}{\sum_{n=0}^{m} P^n_{ij}} \frac{\sum_{n=0}^{m} P^n_{ij}}{\sum_{n=0}^{m} P^n_{ii}},$$

так как $\sum_{n=0}^{m} P^n_{ij} > 0$ для достаточно больших m. Далее, если оба состояния i и j возвратны и принадлежат одному и тому же классу, то, согласно последним двум теоремам, первое отношение в правой части стремится к $1/f^*_{ij} = 1$, а второе отношение — к ${}_iP^*_{ij}$; следовательно,

$$\lim_{m\to\infty} \frac{\sum_{n=0}^{m} P^n_{jj}}{\sum_{n=0}^{m} P^n_{ii}} = {}_iP^*_{ij}.$$

Из (1.1) и (2.5) получаем соответственно

$$f_{ij}(s) = {}_jP_{ii}(s)\,{}_if_{ij}(s),$$
$${}_iP_{ij}(s) = {}_if_{ij}(s)\,{}_iP_{jj}(s).$$

Отсюда, воспользовавшись леммой Абеля, получаем тождества

$$f^*_{ij} = {}_jP^*_{ii}\,{}_if^*_{ij},$$
$${}_iP^*_{ij} = {}_if^*_{ij}\,{}_iP^*_{jj}.$$

Если i и j принадлежат одному и тому же классу возвратных состояний, то $f^*_{ij} = 1$, откуда

$${}_iP^*_{ij} = \frac{{}_iP^*_{jj}}{{}_jP^*_{ii}}.$$

§ 3. СУЩЕСТВОВАНИЕ ОБОБЩЕННЫХ СТАЦИОНАРНЫХ РАСПРЕДЕЛЕНИЙ

В случае неприводимого возвратного положительного класса стационарное распределение $\{\pi_i\}_{i=0}^{\infty}$ представляет собой сходящееся $\left(\text{такое, что } \sum\limits_{i=0}^{\infty} \pi_i < \infty\right)$ положительное решение системы уравнений

$$\sum_{i=0}^{\infty} x_i P_{ij} = x_j, \qquad j = 0,\ 1,\ 2,\ \ldots .$$

Это утверждение доказывается в теореме 1.3 гл. 3. В следующей теореме доказывается, что это свойство является достаточным условием для положительной возвратности.

Теорема 3.1. *Предположим, что марковская цепь неприводима. Если система уравнений*

$$\sum_{j=0}^{\infty} x_j P_{ji} = x_i, \qquad i = 0,\ 1,\ 2,\ \ldots, \tag{3.1}$$

имеет решение, у которого

$$\sum_{j=0}^{\infty} |x_j| < \infty,$$

причем не все x_j равны нулю, то марковская цепь является возвратной положительной.

Доказательство. Из (3.1) последовательно получаем

$$\sum_{j=0}^{\infty} x_j P^n_{ji} = x_i, \qquad n \geqslant 1,$$

Пусть $\tilde{P}_{ij}^m = (1/m)\sum_{n=1}^{m} P_{ij}^n$; тогда

$$\sum_{j=0}^{\infty} x_j \tilde{P}_{ji}^m = x_i.$$

Перейдем теперь к пределу при $m \to \infty$. Так как

$$\sum_{j=0}^{\infty} |x_j \tilde{P}_{ji}^m| \leqslant \sum_{j=0}^{\infty} |x_j| < \infty,$$

мы можем перейти к пределу в каждом слагаемом (т. е. поменять местами суммирование и переход к пределу); тогда

$$x_i = \lim_{m\to\infty} \sum_{j=0}^{\infty} x_j \tilde{P}_{ji}^m = \sum_{j=0}^{\infty} x_j \lim_{m\to\infty} \tilde{P}_{ji}^m.$$

Но

$$\lim_{m\to\infty} \tilde{P}_{ji}^m = \lim_{m\to\infty} \frac{1}{m} \sum_{n=0}^{m} P_{ji}^n = \pi_i \geqslant 0.$$

Следовательно,

$$x_i = \pi_i \sum_{j=0}^{\infty} x_j.$$

Так как $\sum_{j=0}^{\infty} |x_j| < \infty$ и, согласно условию, существует i, такое, что $x_i \neq 0$, то последнее равенство гарантирует, что для некоторого i $\pi_i \neq 0$, а значит, $\pi_i > 0$ для всех $i = 0, 1, 2, \ldots$. ■

Сейчас мы докажем теорему, обратную к только что доказанной, в усиленном варианте, опирающемся на систему неравенств.

Теорема 3.2. *Если неприводимая марковская цепь — возвратная положительная и* $\{x_j \geqslant 0,\ j = 0, 1, 2, \ldots\}$ *есть решение системы неравенств*

$$\sum_{j=0}^{\infty} x_j P_{ji} \leqslant x_i, \qquad i = 0, 1, 2, \ldots,$$

то

$$\sum_{j=0}^{\infty} x_j < \infty.$$

Доказательство. Так же как и в предыдущем доказательстве, мы имеем

$$\sum_{j=0}^{\infty} x_j P_{ji}^n \leqslant x_i, \qquad n \geqslant 1,$$

а при $m \geqslant 1$

$$\sum_{j=0}^{\infty} x_j \tilde{P}_{ji}^m \leqslant x_i, \quad \text{где} \quad \tilde{P}_{ji}^m = \frac{1}{m} \sum_{n=0}^{m-1} P_{ji}^n.$$

Так как $x_j \geqslant 0$ и $\tilde{P}_{ji}^m \geqslant 0$, то $\sum\limits_{j=0}^{M} x_j \tilde{P}_{ji}^m \leqslant x_i$ при любом $M > 0$.

Переходя к пределу при $m \to \infty$, получаем

$$\lim_{m \to \infty} \sum_{j=0}^{M} x_j \tilde{P}_{ji}^m = \pi_i \sum_{j=0}^{M} x_j \leqslant x_i.$$

Так как $\pi_i > 0$, частичные суммы $\sum\limits_{j=0}^{M} x_j$ равномерно ограничены при всех $M > 0$; следовательно, имеет место

$$\sum_{j=0}^{\infty} x_j < \infty. \; ■$$

Согласно теореме 3.1, в случае неприводимой возвратной нулевой марковской цепи система (3.1) не может иметь нетривиального сходящегося решения. Вместе с тем существуют положительные решения, представляющие значительный интерес; о них пойдет речь в следующей теореме.

Теорема 3.3. *Если марковская цепь неприводима и возвратна, то положительная последовательность*

$$v_0 = 1, \quad v_i = {}_0P_{0i}^*, \qquad i = 1, 2, \ldots,$$

является решением системы уравнений

$$v_i = \sum_{j=0}^{\infty} v_j P_{ji}, \qquad i = 0, 1, 2, \ldots \tag{3.2}$$

(определение ${}_0P_{0i}^*$ дается в формуле (2.6)).

Доказательство. По определению ${}_0P_{0i}^*$ имеем

$$\sum_{j=0}^{\infty} v_j P_{ji} = \sum_{j=1}^{\infty} {}_0P_{0j}^* P_{ji} + P_{0i} = P_{0i} + \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} {}_0P_{0j}^n P_{ji}.$$

Поскольку ${}_0P_{0j}^* < \infty$ [1]), повторный ряд в правой части, все члены которого неотрицательны, сходится, и поэтому порядок суммирования можно изменить, что дает нам

$$\sum_{j=0}^{\infty} v_j P_{ji} = P_{0i} + \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} {}_0P_{0j}^n P_{ji}.$$

[1]) Это условие требует пояснения. Мы можем формально изменить порядок суммирования, и если получающийся ряд сходится абсолютно, то сходится и исходный ряд, причем суммы обоих рядов одинаковы. — *Прим. перев.*

Но

$$\sum_{j=1}^{\infty} {}_0P_{0j}^{n} P_{ji} = \begin{cases} {}_0P_{0i}^{n+1}, & \text{если} \quad i \neq 0, \\ f_{00}^{n+1}, & \text{если} \quad i = 0. \end{cases}$$

Следовательно, при $i \neq 0$ имеем

$$\sum_{j=0}^{\infty} v_j P_{ji} = P_{0i} + \sum_{n=1}^{\infty} {}_0P_{0i}^{n+1} = \sum_{n=0}^{\infty} {}_0P_{0i}^{n+1} = {}_0P_{0i}^{*} = v_i,$$

так как

$$P_{0i} = {}_0P_{0i} \quad \text{при} \quad i \neq 0.$$

При $i = 0$

$$\sum_{j=0}^{\infty} v_j P_{j0} = P_{00} + \sum_{n=1}^{\infty} f_{00}^{n+1} = \sum_{n=0}^{\infty} f_{00}^{n+1} = f_{00}^{*} = 1 = v_0,$$

чем и завершается доказательство теоремы. ■

Теорема 3.4. *Для неприводимой возвратной марковской цепи система*

$$v_i = \sum_{j=0}^{\infty} v_j P_{ji}, \qquad i = 0,\ 1,\ 2,\ \ldots, \tag{3.3}$$

$$v_0 = 1, \quad v_i \geqslant 0, \qquad i = 1,\ 2,\ \ldots, \tag{3.4}$$

имеет единственное решение.

Доказательство. В силу предыдущей теоремы последовательность $v_0 = 1$, $v_i = {}_0P_{0i}^{*}$, $i = 1,2, \ldots$, является решением системы уравнений (3.3), удовлетворяющим условиям (3.4). Мы докажем нашу теорему, если покажем, что не существует никакого другого решения системы (3.3), которое удовлетворяло бы условиям (3.4).

Пусть $\{a_i\}$ — последовательность, удовлетворяющая (3.3) и (3.4), тогда

$$a_i = \sum_{j=0}^{\infty} a_j P_{ji}, \qquad i = 0,\ 1,\ 2,\ \ldots .$$

Умножая обе части последнего равенства на P_{ik} и суммируя по i, получаем

$$a_k = \sum_{i=0}^{\infty} a_i P_{ik} = \sum_{i=0}^{\infty} P_{ik} \sum_{j=0}^{\infty} a_j P_{ji} =$$

$$= \sum_{j=0}^{\infty} a_j \sum_{i=0}^{\infty} P_{ji} P_{ik} = \sum_{j=0}^{\infty} a_j P_{jk}^{2}.$$

Изменение порядка суммирования правомерно, так как все члены ряда неотрицательны. Повторяя эту процедуру, получаем для любого $n \geqslant 1$

$$a_i = \sum_{j=0}^{\infty} a_j P_{ji}^n, \qquad i = 0,\ 1,\ 2,\ \ldots .$$

Так как рассматриваемая марковская цепь неприводима и возвратна, для каждого i существует $n \geqslant 1$, такое, что $P_{0i}^n \geqslant 0$. Следовательно,

$$a_i = \sum_{j=0}^{\infty} a_j P_{ji}^n \geqslant a_0 P_{0i}^n > 0,$$

так как $a_0 > 0$. Итак, $a_i > 0$ при всех i.

Введем в рассмотрение следующие величины:

$$Q_{ij} = \frac{a_j}{a_i} P_{ji}. \tag{3.5}$$

Очевидно,

$$Q_{ij} \geqslant 0 \quad \text{и} \quad \sum_{j=0}^{\infty} Q_{ij} = \frac{\sum\limits_{j=0}^{\infty} a_j P_{ji}}{a_i} = 1.$$

Таким образом, мы можем рассматривать величины Q_{ij} как элементы матрицы переходных вероятностей некоторой марковской цепи Q. Соответствующие вероятности перехода за два шага задаются формулой

$$Q_{ij}^2 = \sum_{k=0}^{\infty} Q_{ik} Q_{kj} = \sum_{k=0}^{\infty} \frac{a_k}{a_i} P_{ki} \frac{a_j}{a_k} P_{jk} = \frac{a_j}{a_i} \sum_{k=0}^{\infty} P_{jk} P_{ki} = \frac{a_j}{a_i} P_{ji}^2.$$

Вероятности перехода за n шагов имеют вид

$$Q_{ij}^n = \frac{a_j}{a_i} P_{ji}^n, \qquad n \geqslant 1.$$

Поэтому

$$\sum_{n=0}^{\infty} Q_{ii}^n = \sum_{n=0}^{\infty} P_{ii}^n = \infty$$

и Q_{ij} являются переходными вероятностями возвратной марковской цепи. Воспользуемся теперь теоремами об отношениях. В силу теоремы 2.1 имеем

$$\lim_{m \to \infty} \frac{\sum\limits_{n=0}^{\infty} Q_{i0}^n}{\sum\limits_{n=0}^{\infty} Q_{00}^n} = f_{i0}^*(Q) = 1,$$

где $f^*_{i0}(Q)$ определена по отношению к цепи Q обычным образом. Ее значение равно 1, так как Q — возвратная неприводимая марковская цепь. Но в силу теоремы 2.2

$$\lim_{m\to\infty} \frac{\sum_{n=0}^{m} Q^n_{i0}}{\sum_{n=0}^{m} Q^n_{00}} = \frac{a_0}{a_i} \lim_{m\to\infty} \frac{\sum_{n=0}^{m} P^n_{0i}}{\sum_{n=0}^{m} P^n_{00}} = \frac{a_0}{a_i}\, {}_0P^*_{0i}.$$

Поскольку $a_0 = 1$, то тем самым мы показали, что

$$a_i = {}_0P^*_{0i}, \qquad i = 1, 2, \ldots .$$

Единственность доказана. ■

§ 4. ИНТЕРПРЕТАЦИЯ ОБОБЩЕННЫХ СТАЦИОНАРНЫХ РАСПРЕДЕЛЕНИЙ

Марковскую цепь с матрицей переходных вероятностей $\|Q_{ij}\|$, связанной с матрицей $\|P_{ij}\|$ по формуле (3.5) через некоторое положительное решение системы (3.3), называют обратной марковской цепью к цепи P. В возвратном положительном случае, когда $v_i = c\pi_i$ (c — константа), Q_{ij} можно интерпретировать следующим образом. Предположим, что начальное распределение совпадает с $\{\pi_i\}$, т. е. марковская цепь $X(n)$ в начальный момент находится в состоянии i с вероятностью π_i. Вычислим условную вероятность того, что начальным состоянием было j, если известно, что после одного перехода процесс находился в состоянии i. По формуле Байеса имеем

$$Q_{ij} = \mathrm{P}\{X(0) = j \mid X(1) = i\} = \frac{\mathrm{P}\{X(1) = i \mid X(0) = j\}\,\mathrm{P}\{X(0) = j\}}{\mathrm{P}\{X(1) = i\}}. \tag{4.1}$$

Так как процесс стационарен, то

$$\mathrm{P}\{X(1) = i\} = \pi_i, \qquad \mathrm{P}\{X(0) = j\} = \pi_j$$

и (4.1) принимает вид

$$Q_{ij} = \frac{P_{ji}\pi_j}{\pi_i}. \tag{4.2}$$

Последовательное применение соотношения (4.2) приводит, по существу, к «обращению» времени. Легко видеть, что если с. в. $X(0)$ имеет своим распределением $\{\pi_j\}$, то

$$Q^n_{ij} = \mathrm{P}\{X(0) = j \mid X(n) = i\}.$$

Название «обратный процесс», данное марковской цепи с матрицей переходных вероятностей $\|Q_{ij}\|$, таким образом, отражает сущность дела.

Метод введения обратного процесса применим всякий раз, когда имеется положительное (не обязательно сходящееся) решение системы (3.3). Этим приемом мы воспользуемся ниже в теореме 5.2.

Как следует из теоремы 3.1, решение системы (3.3) расходится в случае возвратного нулевого класса и сходится в случае возвратного положительного класса. Интересную интерпретацию можно дать как сходящемуся, так и расходящемуся решениям системы (3.3). Для возвратного положительного класса значения $\{v_i\}_{i=0}^{\infty}$ пропорциональны стационарным вероятностям пребывания в соответствующих состояниях. В общем случае значения v_i можно интерпретировать как стационарное среднее число частиц, находящихся в состоянии i при соответствующих условиях равновесия. Точный смысл имеющегося в виду равновесия выявляется в следующей теореме:

Теорема 4.1. *Предположим, что счетное число частиц независимо подчиняются марковскому процессу, заданному матрицей* $\mathbf{P} = \|P_{ij}\|$. *Пусть д. с. в.* $A_i(n)$ *представляет число частиц, находящихся в состоянии* i *в момент* n. *Если* $A_i(0)$, $i = 1, 2, \ldots$, *являются независимыми д. с. в., подчиняющимися распределению Пуассона со средними* v_i *соответственно, где* $\sum\limits_k v_k P_{ki} = v_i$, $v_i > 0$, *то* $A_i(n)$, $i = 1, 2, \ldots$, *также являются независимыми д. с. в. с теми же распределениями, что и соответствующие* $A_i(0)$, $i = 1, 2, \ldots$.

Замечание. Когда мы говорим о бесконечном числе случайных величин, что они независимы и подчиняются распределению Пуассона, то это означает, что любая их конечная совокупность обладает этим свойством.

Доказательство. Пусть $A_k(n;\, i)$ есть число частиц, находящихся в состоянии k в момент n, из общего числа частиц, пребывавших в состоянии i в момент $n-1$. Определим векторы $\mathbf{A}(n)$ и $\mathbf{A}(n;\, i)$ следующим образом:

$$\mathbf{A}(n) = (A_1(n),\ A_2(n),\ \ldots), \qquad \mathbf{A}(n;\, i) = (A_1(n;\, i),\ A_2(n;\, i),\ \ldots),$$

тогда $\mathbf{A}(n) = \sum\limits_i \mathbf{A}(n;\, i)$. Доказательство теоремы проведем по индукции. В силу индуктивного предположения (что д. с. в. $A_i(n-1)$ независимы), а также потому, что частицы ведут себя независимо друг от друга, д. с. в. $A_k(n;\, i)$, $i = 1, 2, \ldots$, независимы при каждом фиксированном k. Мы покажем, что компоненты $A_k(n;\, i)$, $k = 1, 2, \ldots$, каждого вектора $\mathbf{A}(n;\, i)$ также являются независимыми д. с. в., откуда следует, что $A_k(n) = \sum\limits_i A_k(n;\, i)$, $k = 1, 2, \ldots$, — независимые д. с. в.

Для любого конечного числа компонент вектора $\mathbf{A}(n;\ i)$ и целых чисел $a_1, a_2, \ldots, a_r$ имеем

$$\begin{aligned} &\mathrm{P}\left\{A_{k_1}(n;\ i)=a_1,\ A_{k_2}(n;\ i)=a_2,\ \ldots,\ A_{k_r}(n;\ i)=a_r\right\}= \\ &=\sum_{a=0}^{\infty}\mathrm{P}\left\{A_i(n-1)=a\right\}\mathrm{P}\left\{A_{k_1}(n;i)=a_1,\ \ldots,\ A_{k_r}(n;i)=a_r \mid A_i(n-1)=a\right\}. \end{aligned} \tag{4.3}$$

По предположению индукции первый сомножитель общего члена суммы равен

$$\exp(-v_i)\frac{v_i^a}{a!}. \tag{4.4}$$

Ввиду обусловленности и независимого характера поведения частиц второй сомножитель представляет собой полиномиальное распределение, т. е.

$$\begin{aligned} &\mathrm{P}\left\{A_{k_1}(n;\ i)=a_1,\ \ldots,\ A_{k_r}(n;\ i)=a_r \mid A_i(n-1)=a\right\}= \\ &=\frac{a!}{a_1!\,a_2!\,\ldots\,a_r!}\left(\prod_{\nu=1}^{r}(P_{ik_\nu})^{a_\nu}\right)\frac{\left(1-\sum\limits_{\nu=1}^{r}P_{ik_\nu}\right)^{a-\sum\limits_{\nu=1}^{r}a_\nu}}{\left(a-\sum\limits_{\nu=1}^{r}a_\nu\right)!}. \end{aligned} \tag{4.5}$$

Это выражение полагается равным нулю, если $\sum\limits_{\nu=1}^{r}a_\nu>a$. Подставляя (4.4) и (4.5) в (4.3), получаем

$$\begin{aligned} &\mathrm{P}\left\{A_{k_1}(n;\ i)=a_1,\ A_{k_2}(n;\ i)=a_2,\ \ldots,\ A_{k_r}(n;\ i)=a_r\right\}= \\ &=\sum_{a=\sum\limits_{\nu=1}^{r}a_\nu}^{\infty}\ \prod_{m=1}^{r}\frac{(v_iP_{ik_m})^{a_m}}{a_m!}\exp(-v_iP_{ik_m})\times \\ &\times\frac{\left[\left(1-\sum\limits_{m=1}^{r}P_{ik_m}\right)v_i\right]^{a-\sum\limits_{m=1}^{r}a_m}}{\left(a-\sum\limits_{m=1}^{r}a_m\right)!}\exp\left(-v_i\left[1-\sum_{m=1}^{r}P_{ik_m}\right]\right)= \\ &=\prod_{m=1}^{r}\left[\frac{(v_iP_{ik_m})^{a_m}}{a_m!}\exp(-v_iP_{ik_m})\right]\sum_{a=a^*}^{\infty}\frac{\left[\left(1-\sum\limits_{m=1}^{r}P_{ik_m}\right)v_i\right]^{a-a^*}}{(a-a^*)!}\times \\ &\times\exp\left[-v_i\left(1-\sum_{m=1}^{r}P_{ik_m}\right)\right]. \end{aligned} \tag{4.6}$$

Но данная сумма равна 1 (достаточно положить $a = a^* + \alpha$ и просуммировать по α), так как она является суммой вероятностей частных значений для пуассоновского распределения с параметром

$$\left(1 - \sum_{m=1}^{r} P_{ik_m}\right) \cdot v_i.$$

Полученное в формуле (4.6) разложение показывает, что $A_k(n;\ i)$, $k = 1, 2, \ldots$, являются независимыми с. в., подчиняющимися распределению Пуассона со средними $\{v_i P_{ik}\}$ соответственно.

Следовательно, $A_k(n) = \sum_{i=0}^{\infty} A_k(n;\ i)$ — независимые с. в. с пуассоновским распределением и средними $\sum_{i=0}^{\infty} v_i P_{ik} = v_k$ соответственно.

Поскольку по условию $A_i(0)$, $i = 1, 2, \ldots$, независимы и имеют пуассоновское распределение со средними v_i, то тем самым доказательство по индукции закончено. ■

§ 5. РЕГУЛЯРНЫЕ, СУПЕРРЕГУЛЯРНЫЕ И СУБРЕГУЛЯРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ МАРКОВСКИХ ЦЕПЕЙ

Мы уже познакомились с несколькими критериями для определения возвратности, невозвратности и положительной возвратности марковской цепи (см. теоремы 4.1—4.2 гл. 3) и применили их при изучении некоторых моделей из теории очередей. Условия этих критериев связаны с характером решений системы уравнений

$$\sum_{i=0}^{\infty} \xi_i P_{ik} = \xi_k, \qquad k = 0,\ 1,\ 2,\ \ldots, \tag{5.1}$$

либо системы уравнений

$$\sum_{k=0}^{\infty} P_{ik} \eta_k = \eta_i, \qquad i = 0,\ 1,\ 2,\ \ldots.$$

Современный подход к этой проблеме состоит в применении теории регулярных, суперрегулярных и субрегулярных последовательностей. Мы остановимся на наиболее простых аспектах этой элегантной теории, которая основывается на теории потенциала марковских матричных операторов. Классическая теория потенциала привлекается при рассмотрении этих же идей в исследовании броуновского движения. Такое взаимопроникновение теории потенциала и теории вероятностей чрезвычайно плодотворно и в последнее время привлекло внимание исследователей.

Пусть $\mathbf{P}$ — заданная матрица переходных вероятностей. Будем говорить, что неотрицательный вектор (неотрицательная последовательность) $\mathbf{u} = \{u(j)\}_{j=0}^{\infty}$ является по отношению к $\mathbf{P}$

правым регулярным (сокращенно r-регулярным),

$$\text{если} \quad \sum_j P_{ij} u(j) = u(i),$$

правым суперрегулярным (r-суперрегулярным),

$$\text{если} \quad \sum_j P_{ij} u(j) \leqslant u(i),$$

правым субрегулярным (r-субрегулярным),

$$\text{если} \quad \sum_j P_{ij} u(j) \geqslant u(i).$$

Правую суперрегулярную последовательность $\{u(i)\}$ будем называть минимальной, если из условия $0 \leqslant \xi(i) \leqslant u(i)$ $(i \geqslant 0)$, где $\xi(i)$ — регулярная последовательность, следует, что $\xi(i) = cu(i)$ при некоторой константе c. Неотрицательный вектор $\{v(i)\}_{i=0}^{\infty}$ будем называть

левым регулярным (l-регулярным),

$$\text{если} \quad \sum_i v(i) P_{ij} = v(j),$$

левым суперрегулярным (l-суперрегулярным),

$$\text{если} \quad \sum_i v(i) P_{ij} \leqslant v(j),$$

левым субрегулярным (l-субрегулярным),

$$\text{если} \quad \sum_i v(i) P_{ij} \geqslant v(j).$$

Сначала мы докажем теорему о представлении правых регулярных последовательностей минимальными регулярными последовательностями.

Теорема 5.1. *Пусть* $\mathbf{u}$ *есть r-суперрегулярный вектор по отношению к* $\mathbf{P}$. *Тогда предел*

$$a(i) = \lim_{n \to \infty} \sum_j P_{ij}^{(n)} u(j)$$

существует для всех i и вектор $\mathbf{a}$ *является r-регулярным вектором по отношению к* $\mathbf{P}$. *Более того, если* $\mathbf{b}$ *есть r-регулярный вектор по отношению к* $\mathbf{P}$ *и* $b(i) \leqslant u(i)$ *при всех i, то* $b(i) \leqslant a(i)$ *для всех i. Если мы представим компоненты вектора* $\mathbf{u}$ *в виде*

$$u(i) = a(i) + c(i), \qquad i \geqslant 0, \tag{5.2}$$

где

$$c(i) = u(i) - a(i),$$

то $c(i)$ образуют минимальный r-суперрегулярный вектор.

Доказательство. По определению r-суперрегулярного вектора имеем

$$\sum_j P_{ij}^{(n)} u(j) = \sum_j \sum_k P_{ik}^{(n-1)} P_{kj} u(j) = \sum_k P_{ik}^{(n-1)} \sum_j P_{kj} u(j) \leqslant \sum_k P_{ik}^{(n-1)} u(k).$$

В векторно-матричных обозначениях мы можем записать это соотношение в виде $\mathbf{P}^n \mathbf{u} \leqslant \mathbf{P}^{n-1} \mathbf{u}$, подразумевая под этим покомпонентные неравенства.

Итак, для каждого i

$$u(i) \geqslant \sum_j P_{ij} u(j) \geqslant \sum_j P_{ij}^{(2)} u(j) \geqslant \ldots .$$

Поскольку все члены этой цепочки неравенств неотрицательны, то $a(i)$ существует и $a(i) \leqslant u(i)$ при всех i. Далее,

$$\sum_j P_{ij} \sum_k P_{jk}^{(n)} u(k) = \sum_k P_{ik}^{(n+1)} u(k).$$

В пределе при $n \to \infty$ выражение в правой части сходится к $a(i)$, тогда как левая часть, если формально перейти к пределу под знаком суммы, стремится к $\sum_j P_{ij} a(j)$, т. е.

$$\sum_j P_{ij} a(j) = a(i).$$

Для доказательства правомерности предельного перехода под знаком суммы зафиксируем некоторое значение индекса i. Для любого $\varepsilon > 0$ существует $N(\varepsilon)$, такое, что

$$\sum_{j > N(\varepsilon)} P_{ij} u(j) \leqslant \varepsilon;$$

тогда

$$\sum_{j > N(\varepsilon)} P_{ij} \sum_k P_{jk}^{(n)} u(k) \leqslant \sum_{j > N(\varepsilon)} P_{ij} u(j) \leqslant \varepsilon \quad \text{при всех } n.$$

Далее, воспользуемся представлением

$$\sum_j P_{ij} \sum_k P_{jk}^{(n)} u(k) = \sum_{j \leqslant N(\varepsilon)} P_{ij} \sum_k P_{jk}^{(n)} u(k) + \sum_{j > N(\varepsilon)} P_{ij} \sum_k P_{jk}^{(n)} u(k).$$

Как мы уже видели, при $n \to \infty$ левая часть стремится к $a(i)$. Так как первый член в правой части является конечной суммой по j, то его предел есть просто

$$\sum_{j \leqslant N(\varepsilon)} P_{ij} a(j).$$

Величина второго члена не превышает ε. Итак, имеем

$$a(i) = \sum_{j \leqslant N(\varepsilon)} P_{ij} a(j) + d(i), \quad \text{где } 0 \leqslant d(i) \leqslant \varepsilon,$$

6 Зак. 939

откуда следует, что $a(i)=\sum_j P_{ij}a(j)$, т. е. $\mathbf{a}=\{a(i)\}$ является r-регулярным вектором по отношению к $\mathbf{P}$.

Предположим, наконец, что

$$b(i)=\sum_j P_{ij}b(j)\leqslant u(i) \quad \text{при всех } i.$$

Тогда с помощью индукции получаем

$$b(i)=\sum_j P_{ij}^{(n)}b(j)\leqslant\sum_j P_{ij}^{(n)}u(j) \quad \text{при всех } n \text{ и } i.$$

Следовательно,

$$b(i)\leqslant\lim_{n\to\infty}\sum_j P_{ij}^{(n)}u(j)=a(i) \quad \text{при всех } i.$$

Не составляет никакого труда проверить, что последовательность $c(i)=u(i)-a(i)$ является r-суперрегулярной. Остается только установить минимальность последовательности $c(i)$. Предположим, что

$$0\leqslant\xi(i)\leqslant c(i), \quad i\geqslant 0, \tag{5.3}$$

где $\{\xi(i)\}$ есть r-регулярная последовательность. Применяя n раз $\mathbf{P}$ к обеим частям (5.3), получаем

$$\xi=\mathbf{P}^n\xi\leqslant\mathbf{P}^n\mathbf{c}.$$

Но из определения вектора $\mathbf{c}$ следует, что $\mathbf{P}^n\mathbf{c}=\mathbf{P}^n\mathbf{u}-\mathbf{P}^n\mathbf{a}=\mathbf{P}^n\mathbf{u}-\mathbf{a}$ стремится к нулевому вектору. Этим завершается доказательство минимальности вектора $\mathbf{c}$. ■

Представление (5.2) в случае броуновского движения является не чем иным, как классическим представлением Рисса, связанным с гармоническими, супергармоническими и потенциальными функциями. Изложение этой элегантной теории выходит за рамки нашей книги.

Для невозвратных марковских цепей очень легко построить r-суперрегулярные последовательности. Напомним, что в невозвратном случае $\sum_{n=0}^{\infty} P_{ik}^n<\infty$ при всех i, $k\geqslant 0$. Мы утверждаем, что при фиксированном k_0

$$u_i=\sum_{n=0}^{\infty} P_{ik_0}^n, \qquad i=0,\ 1,\ \ldots, \tag{5.4}$$

является r-суперрегулярной последовательностью. Действительно,

$$\sum_{j=0}^{\infty} P_{ij}u_j=\sum_{n=0}^{\infty} P_{ik_0}^{n+1}=u_i-P_{ik_0}^0\leqslant u_i. \tag{5.5}$$

Приведенное построение позволяет сделать вывод о том, что существует достаточно обширное множество непостоянных положительных r-суперрегулярных векторов. То, что последовательности вида (5.4) не являются постоянными, следует из соотношения (5.5), которое при $i = k_0$ является строгим неравенством. Совершенно иная картина имеет место в возвратном случае. Следующая теорема утверждает, что единственной r-суперрегулярной последовательностью является постоянный вектор. Она обобщает критерий, полученный в теореме 4.1 гл. 3.

Теорема 5.2. *Неприводимая марковская цепь с матрицей переходных вероятностей* $\mathbf{P}$ *возвратна тогда и только тогда, когда всякий неотрицательный вектор* $\mathbf{v}$, *r-суперрегулярный по отношению к* $\mathbf{P}$, *у которого хотя бы одна компонента положительна, является постоянным.*

Доказательство. Пусть марковская цепь возвратна. Рассмотрим систему неравенств

$$u_i \geqslant \sum_j P_{ij} u_j, \quad u_i \geqslant 0, \qquad i = 0, 1, \ldots .$$

Мы покажем сначала, что если $u_{j_0} > 0$ для некоторого j_0, то $u_j > 0$ при всех j. В самом деле, для любых заданных j_0 и k существует такое n, что $P_{kj_0}^{(n)} > 0$. Тогда, как и в предыдущей теореме, имеем

$$u_k \geqslant \sum_j P_{kj}^{(n)} u_j \geqslant P_{kj_0}^{(n)} u_{j_0} > 0.$$

Таким образом, если $\mathbf{u} \neq \mathbf{0}$, то $u_i > 0$ при всех i. Пусть теперь k произвольно, но фиксировано и положим $\xi_i = u_i/u_k$. Тогда

$$\xi_i \geqslant \sum_j P_{ij} \xi_j = \sum_{j \neq k} P_{ij} \xi_j + P_{ik}. \tag{5.6}$$

Итерируя это неравенство, получаем

$$\begin{aligned} \xi_i &\geqslant \sum_{j \neq k} P_{ij} \Big[\sum_{s \neq k} P_{js} \xi_s + P_{jk} \Big] + P_{ik} = \\ &= \sum_{j, s \neq k} P_{ij} P_{js} \xi_s + \sum_{j \neq k} P_{ij} P_{jk} + P_{ik} = \\ &= \sum_{j, s \neq k} P_{ij} P_{js} \xi_s + f_{ik}^{(2)} + f_{ik}^{(1)}, \end{aligned}$$

где последние два члена по ранее данному определению являются вероятностями первого достижения. Опять подставляя (5.6) в полученное неравенство, получаем

$$\begin{aligned} \xi_i &\geqslant \sum_{j, s \neq k} P_{ij} P_{js} \Big[\sum_{r \neq k} P_{sr} \xi_r + P_{sk} \Big] + f_{ik}^{(2)} + f_{ik}^{(1)} = \\ &= \sum_{j, s, r \neq k} P_{ij} P_{js} P_{sr} \xi_r + \sum_{j, s \neq k} P_{ij} P_{js} P_{sk} + f_{ik}^{(2)} + f_{ik}^{(1)} = \\ &= \sum_{j, s, r \neq k} P_{ij} P_{js} P_{sr} \xi_r + f_{ik}^{(3)} + f_{ik}^{(2)} + f_{ik}^{(1)}. \end{aligned}$$

Продолжая таким же образом, приходим к неравенству

$$\xi_i \geqslant \sum_{n=1}^{m} f_{ik}^{(n)},$$

справедливому при любом m. Отсюда следует, что

$$\xi_i \geqslant \sum_{n=1}^{\infty} f_{ik}^{(n)} = f_{ik}^{*} = 1,$$

так как цепь возвратна и неприводима.

Таким образом, $\xi_i = u_i/u_k \geqslant 1$, или $u_i \geqslant u_k$. Но i и k произвольны, значит, $u_i = u_k$ при всех i, k.

Докажем теперь достаточность. Пусть марковская цепь невозвратна. Положим

$$u_i = \begin{cases} f_{ik}^{*}, & \text{если} \quad i \neq k, \\ 1, & \text{если} \quad i = k, \end{cases}$$

где k произвольно, но фиксировано. (Напомним, что f_{ik}^{*} есть вероятность попасть в состояние k за конечное число шагов, исходя из состояния i.) Отсюда следует

$$u_i = f_{ik}^{*} = \sum_{j \neq k} P_{ij} f_{jk}^{*} + P_{ik} = \sum_{j} P_{ij} u_j, \quad \text{если } i \neq k,$$

и

$$u_k = 1 \geqslant f_{kk}^{*} = \sum_{j \neq k} P_{kj} f_{jk}^{*} + P_{kk} = \sum_{j} P_{kj} u_j.$$

Таким образом, вектор $\mathbf{u}$ является r-суперрегулярным. Но если $f_{jk}^{*} = 1$ при $j \neq k$, то

$$f_{kk}^{*} = \sum_{j \neq k} P_{kj} f_{jk}^{*} + P_{kk} = \sum_{j \neq k} P_{kj} + P_{kk} = \sum_{j} P_{kj} = 1,$$

что противоречит предположению о невозвратности цепи. Следовательно, $\mathbf{u}$ — непостоянный ограниченный r-суперрегулярный вектор. ■

С помощью теоремы 5.2 легко доказать усиленный вариант теоремы 3.4, допускающий неравенства.

Теорема 5.3. *Для возвратной неприводимой марковской цепи система*

$$\begin{aligned} v_i &\geqslant \sum_{j=0}^{\infty} v_j P_{ji}, \qquad i = 0, 1, \ldots, \\ v_0 &= 1, \ v_i \geqslant 0, \qquad i = 1, 2, \ldots, \end{aligned} \tag{5.7}$$

имеет единственное решение.

Метод, принятый в доказательстве теоремы 3.4, состоит в переходе к обратному процессу, что позволяет свести рассмотрение l-суперрегулярных векторных последовательностей к рассмотрению r-суперрегулярных числовых последовательностей. Этот прием сведе́ния задач о левых регулярных объектах к задачам о правых регулярных объектах с помощью обратного процесса является довольно распространенным и продуктивным.

Доказательство. Как мы уже видели, $v_0 = 1$, $v_i = {}_0P^*_{0i}$ является решением системы (5.7). Действительно, эта последовательность удовлетворяет условиям $v_i \geqslant \sum\limits_{j=0}^{\infty} v_j P_{ji}$, $i = 0, 1, \ldots$, если в них заменить знак $\geqslant$ знаком равенства, и, кроме того, $v_i > 0$ при всех i. Полагая

$$Q_{ij} = \frac{v_j P_{ji}}{v_i}, \qquad i, j = 0, 1, \ldots,$$

мы получаем некоторую матрицу переходных вероятностей $\mathbf{Q} = \|Q_{ij}\|$, поскольку $Q_{ij} \geqslant 0$ и

$$\sum_{j=0}^{\infty} Q_{ij} = \frac{1}{v_i} \sum_{j=0}^{\infty} v_j P_{ji} = \frac{v_i}{v_i} = 1.$$

К тому же, как и в теореме 3.4,

$$Q_{ik}^{(n)} = \frac{v_k}{v_i} P_{ki}^{(n)}$$

и неприводимость матрицы $\mathbf{Q}$ следует из неприводимости матрицы $\mathbf{P}$. Далее, если

$$c_i \geqslant \sum_{j=0}^{\infty} c_j P_{ji}, \qquad i = 0, 1, \ldots,$$

и

$$c_0 = 1, \quad c_j \geqslant 0, \qquad i = 1, 2, \ldots,$$

то

$$\sum_{j=0}^{\infty} Q_{ij} \frac{c_j}{c_i} = \frac{1}{v_i} \sum_{j=0}^{\infty} c_j P_{ji} \leqslant \frac{c_i}{v_i},$$

т. е. вектор $\{c_j/v_j\}$ является r-суперрегулярным по отношению к матрице $\mathbf{Q}$. Но в силу предыдущей теоремы $\{c_j/v_j\}$ должен быть постоянным вектором. Так как $c_0 = v_0 = 1$, то $c_i = v_i$, $i = 0, 1, \ldots$. Теорема доказана. ■

ЗАДАЧИ

1. Рассмотрим неприводимую возвратную положительную марковскую цепь, начальное состояние которой $X_0 = i$. Пусть $N_n(i)$ — число возвращений в состояние i за первые n переходов. Доказать, что

$$\lim_{n\to\infty} \frac{\mathrm{M}(N_n(i))}{n} = \frac{1}{\mu_i},$$

где μ_i — среднее время возвращения в состояние i, т. е. $\mu_i = \sum_{n=1}^{\infty} n f_{ii}^n$.

2 (продолжение). Пусть $T_m(i)$ обозначает число переходов до m-го возвращения в состояние i $(X_0 = i)$. Показать, что

$$\mathrm{P}\{T_m(i) > n\} = \mathrm{P}\{N_n(i) < m\}.$$

***3** (продолжение). Предположим, что $\sum_{n=1}^{\infty} n^2 f_{ii}^n < \infty$. $T_m(i)$ является суммой m независимых одинаково распределенных с. в. со средним μ_i и дисперсией σ_i. С помощью центральной предельной теоремы и соотношения, полученного в предыдущей задаче, найти предельное распределение с. в. $N_n(i)$, должным образом нормированной, т. е. найти такие $a_n > 0$ и $b_n > 0$, что $(N_n(i) - a_n)/b_n$ имеет предельное нормальное распределение.

4. Пусть $\{X_n, n \geqslant 0\}$ — неприводимая возвратная марковская цепь с матрицей переходных вероятностей $\mathbf{P} = \|P_{ij}\|$ и обобщенной инвариантной мерой $\{v_i\}$, т. е. $\sum v_i P_{ij} = v_j$, $v_j > 0$, $j = 0, 1, \ldots$. Определим вложенный процесс $\{Y_n, n \geqslant 0\}$. Он определяется предыдущим, если рассматривать только те моменты времени, когда $X_n = 0$ или 1 (т. е. $Y_0 = X_{n_0}$, где n_0 — момент первого достижения состояний 0 или 1; $Y_m = X_{n_m}$, где n_m — момент m-го возвращения в состояния 0 или 1). Процесс $\{Y_n, n \geqslant 0\}$ является неприводимой возвратной марковской цепью. Пусть w_0, w_1 обозначают стационарные вероятности вложенной марковской цепи. Показать, что $w_1/w_0 = v_1/v_0$.

Указание: Воспользоваться интерпретацией v_i, данной в теоремах 3.3 и 3.4.

5. Пусть $\mathbf{P} = \|P_{ij}\|$ и $P_n = \|P_{ij}(n)\|$, $n = 1, 2, \ldots$, — матрицы переходных вероятностей неприводимых марковских цепей, и пусть $\{v_i\}$ и $\{v_i^{(n)}\}$, $n = 1, 2, \ldots$, — соответствующие инвариантные меры, нормированные таким образом, что $v_0 = v_0^{(n)} = 1$ при всех n. Доказать, что если $P_{ij}(n) \to P_{ij}$ для всех i и j при $n \to \infty$, то $v_i^{(n)} \to v_i$ для всех i.

Указание: Доказать, что $\lim_{n\to\infty} v_j^{(n)} = w_j$ существует и удовлетворяет системе

$$\sum_{i=0}^{\infty} w_i P_{ij} \leqslant w_j, \qquad j = 0, 1, 2, \ldots \qquad w_0 = 1.$$

Затем воспользоваться свойством единственности, установленным в теореме 5.3.

6. Доказать, что для неприводимой марковской цепи любая неотрицательная r-суперрегулярная последовательность $\{u(i)\}$ обладает следующим свойством:

$$u(i) \geqslant f_{ik}^{*} u(k) \quad \text{при всех } i \text{ и } k,$$

где f_{ik}^{*} — вероятность достижения состояния k из состояния i.

Указание: См. доказательство теоремы 5.2.

7. Пусть $\{u(i)\}$ — конечная неотрицательная r-суперрегулярная последовательность. Положим

$$w(i) = u(i) - \sum_j P_{ij} u(j).$$

Показать, что множество A всех состояний i, для которых $w(i) > 0$, совпадает с множеством невозвратных состояний.

Указание: Обобщить результат задачи 6 на рассматриваемый случай; точнее, получить строгие неравенства и воспользоваться ими.

8. В неприводимой марковской цепи зафиксируем некоторое состояние, обозначим его 0. Показать, что если $f^*_{i0} \geqslant \alpha > 0$ для всех $i \neq 0$, то цепь возвратна.

Указание: Показать, что вероятность возвратиться в состояние 0 лишь конечное число раз равна нулю (см. теоремы 7.1 и 7.2 гл. 2).

9. Доказать, что неприводимая марковская цепь невозвратна тогда и только тогда, когда существует ограниченная l-суперрегулярная последовательность $\{\mu(i)\}$, такая, что

$$\mu(k) > \sum_i \mu(i) P_{ik} \quad \text{для некоторого состояния } k.$$

Указание: Воспользоваться теоремами 5.3, 3.3 и 3.4 (достаточность); воспользоваться соотношением (5.5) (необходимость).

*10. Доказать следующие тождества для тройки состояний из одного и того же положительного возвратного класса $(j \neq k)$:

(а) $$m_{ik} + m_{kj} - m_{ij} = {}_kf^*_{ij}(m_{jk} + m_{kj}),$$

(б) $$\frac{m_{jj}}{m_{kk}} = \frac{{}_jf^*_{jk}}{{}_kf^*_{kj}},$$

где $m_{jk} = \sum_{n=1}^{\infty} n f^n_{jk}$.

11. Пусть для неприводимой, но не обязательно возвратной марковской цепи ${}_0P^*_{00} = 1$ и ${}_0P^*_{0i} < \infty$ $(i \geqslant 1)$. Показать, что последовательность ${}_0P^*_{0i}$ является l-суперрегулярной.

12. Рассмотрим электрическую цепь с m граничными и n внутренними контактами; всего $n + m$ контактов. Пусть T_{ij} — ток, текущий от контакта i к контакту j, и R_{ij} — сопротивление между этими контактами, причем i и j не являются граничными контактами одновременно. Пусть V_j — потенциал контакта j. Предположим, что сопротивления известны и заданы потенциалы в граничных контактах. Согласно закону Ома, имеем

(i) $$V_i - V_j = R_{ij} T_{ij},$$

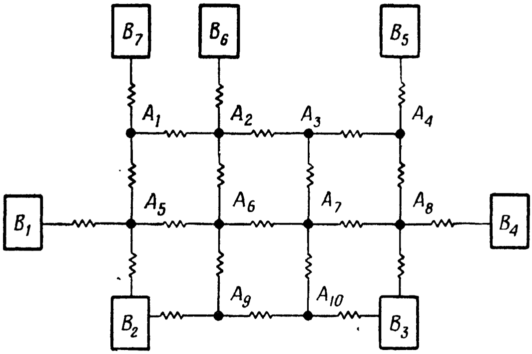
а по первому закону Кирхгофа

(ii) $$\sum_j T_{ij} = 0.$$

С помощью (i) и (ii) показать, что V_i удовлетворяет соотношениям

(iii) $$V_i = \left(\sum_j R_{ij}^{-1}\right)^{-1} \sum_k V_k R_{ik}^{-1}.$$

Дать интерпретацию (iii), согласно которой мы можем рассмотреть случайное блуждание по контактам электрической цепи, считая граничные контакты поглощающими состояниями и определив переходные вероятности в виде соответствующих выражений для R_{ij}.



13 (продолжение). Занумеруем состояния процесса случайного блуждания так, что состояния 1, 2, ..., m соответствуют граничным контактам, а состояния $m+1$, $m+2$, ..., $m+n$ — внутренним контактам. Показать, что потенциал внутреннего контакта A_i можно представить в виде

$$V_{m+i} = \sum_{k=1}^{m} b_{ik} V_k \quad \text{при} \quad i = 1, 2, \ldots,$$

где b_{ik} — вероятность поглощения граничным контактом B_k при начальном состоянии, соответствующем внутреннему контакту A_k.

14. Пусть $\{U_j\}_{j=0}^{\infty}$ — действительное решение системы уравнений

$$\sum_{j=0}^{\infty} P_{ij} U_j - U_i = 1, \qquad i = 0, 1, 2, \ldots,$$

и пусть $M\left(U_{X_n} \mid X_0 = i\right)$ обозначает среднее значение с. в. U_{X_n} при начальном условии $X_0 = i$. Доказать, что

$$\lim_{n \to \infty} \frac{M\left(U_{X_n} \mid X_0 = i\right)}{n} = 1.$$

Указание: Воспользоваться соотношением $\sum_{j=0}^{\infty} P_{ij}^{(n)} U_j - U_i = n, \ i = 0, 1, \ldots.$

15 (продолжение). Доказать, что $U_{X_n} - n$ является мартингалом.

16. Пусть $\|P_{ij}\|$ — матрица переходных вероятностей возвратной нулевой или невозвратной марковской цепи, и пусть $\{u_i\}$ — l-регулярный положительный вектор, а A — некоторое множество состояний. Введем обозначение $P_{iA}^{(n)} = \sum_{j \in A} P_{ij}^{(n)}$. Показать, что если $\mu(A) = \sum_{i \in A} u_i < \infty$, то $P_{iA}^{(n)} \to 0$ при $n \to \infty$.

Указание: Пусть B — конечное подмножество A, такое, что $\sum_{i \in A-B} u_i < \varepsilon$. Показать, что $P_{i,A-B}^{(n)} < \varepsilon/u_i$, и воспользоваться этим фактом. ($A—B$ обозначает множество состояний, входящих в A и не входящих в B.)

17. Рассмотреть задачу о разорении игрока с $N+1$ состояниями и матрицей переходных вероятностей, приведенной в гл. 3. Найти r-регулярные векторы относительно этой матрицы.

Ответ: $u(i) = a\pi_i(c_0) + b\pi_i(c_n)$, где a и b — произвольные числа (обозначения см. в гл. 3).

18. Пусть $\|P_{ij}\|$ — матрица переходных вероятностей марковской цепи с бесконечным числом состояний. Предположим, что $P_{ij} = 0$ при $j > i+1$ и $P_{i,i+1} > 0$ при всех i. Зададим систему полиномов $Q_i(z)$ рекуррентными соотношениями

$$Q_0(z) \equiv 1,$$

$$zQ_i(z) = P_{i,\,i+1}Q_{i+1}(z) + \sum_{j=0}^{i} P_{ij}Q_j(z), \qquad i = 0, 1, 2, \ldots .$$

Пусть T_j — время первого достижения состояния $j+1$ из состояния j, и пусть $f_j(z) = \sum_{n=1}^{\infty} \mathrm{P}\{T_j = n\} z^n$ — производящая функция с. в. T_j. Показать, что

$$f_j(z) = Q_j(1/z)/Q_{j+1}(1/z).$$

Указание: Получить рекуррентную формулу для $\mathrm{P}\{T_j = n\}$, рассматривая возможные исходы первого перехода, затем перейти к производящим функциям.

ЗАМЕЧАНИЯ

Материал этой главы почерпнут в основном из книги Чжун Кай-лая [1].

Наше изложение представляет собой лишь введение в эту важную и развивающуюся область теории вероятностей.

Готовящаяся к изданию книга Кемени и Снелла [2] содержит подробное изложение теории потенциалов для марковских цепей.

ЛИТЕРАТУРА

1. Чжун Кай-лай, Однородные цепи Маркова, «Мир», М., 1964.
2. Kemeny J. G., Snell J. L., Potential Theory for Markov Chains (готовится к печати).

Глава 6

ПОСЛЕДОВАТЕЛЬНОСТЬ СУММ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН КАК МАРКОВСКАЯ ЦЕПЬ

§ 1. СВОЙСТВА ВОЗВРАТНОСТИ СУММ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Пусть X_1, $X_2, \ldots$ — последовательность целочисленных независимых одинаково распределенных случайных величин. Положим $S_n = X_1 + X_2 + \ldots + X_n$, $n = 1, 2, \ldots$. Для полноты будем считать, что $S_0 = 0$.

В этой главе мы остановимся на некоторых свойствах сумм S_n, $n = 0, 1, 2, \ldots$, независимых случайных величин. Сами суммы мы будем рассматривать как последовательные значения марковской цепи специальной структуры с дискретным пространством состояний. В пределах нашего элементарного изложения мы лишь поверхностно коснемся теории сумм независимых случайных величин. Полное и элегантное изложение этой богатой и красивой теории читатель найдет в книге Спицера [1].

В примере А § 1 гл. 2 мы упоминали о последовательности S_n (где X_i были неотрицательными целочисленными случайными величинами) как о примере марковской цепи. В этой главе пространство состояний соответствующей марковской цепи состоит из всех целых чисел: положительных, отрицательных и нуля. Так как мы положили $S_0 \equiv 0$, начальным состоянием является нуль. Характерной чертой марковской цепи $\{S_n\}$ является пространственная однородность, т. е. ее одношаговые переходные вероятности обладают следующим свойством: $\mathrm{P}\{S_n = j \mid S_{n-1} = i\} = P_{ij} = P_{0,\, j-i} = P_{i-j,\, 0}$. Простой индукцией легко убедиться, что этим же свойством обладают n-шаговые переходные вероятности:

$$P_{ij}^n = P_{0,\, j-i}^n = P_{i-j,\, 0}^n = \mathrm{P}\{S_{n+k} = j \mid S_k = i\}, \quad n \geqslant 1.$$

В этой главе мы будем предполагать, что случайная величина X_1 «неприводима». Под этим подразумевается, что марковская цепь с матрицей переходных вероятностей $P_{ij} = \mathrm{P}\{S_n = j \mid S_{n-1} = i\}$ неприводима. Ниже мы дадим простой критерий неприводимости для рассматриваемого случая.

Условимся с самого начала, чтобы не повторять это каждый раз, что с. в. X_1 является невырожденной, т. е. она может принимать по крайней мере два значения с положительными вероятностями.

Перейдем теперь к условиям возвратности и невозвратности марковской цепи, порожденной последовательностью $\{S_n\}$. Для того чтобы получить эти условия, введем следующие величины:

$$G_{ij}^n = \sum_{m=0}^{n} P_{ij}^m, \quad G_{ij} = \sum_{m=0}^{\infty} P_{ij}^m \leqslant \infty,$$

определенные для всех i, j и $n = 0, 1, 2, \ldots$. Величина G_{ij} является аналогом функции Грина и связана с элементами теории потенциала, о которой мы упоминали в гл. 5.

Лемма 1.1.

$$G_{ij}^n \leqslant G_{00}^n, \qquad n = 0,\ 1,\ 2,\ \ldots, \tag{1.1}$$

для всех целых чисел i и j. В частности, при $n \to \infty$

$$G_{ij} \leqslant G_{00} \tag{1.2}$$

для всех целых чисел i и j.

Доказательство. В силу пространственной однородности процесса $G_{ij}^n = G_{i-j,0}^n$. Поэтому достаточно показать, что $G_{i0}^n \leqslant G_{00}^n$ при всех $i = 0, \pm 1, \pm 2, \ldots$ и $n = 0, 1, 2, \ldots$. Но

$$G_{i0}^n = \sum_{m=0}^{n} P_{i0}^m = \sum_{m=0}^{n} \sum_{l=0}^{m} f_{i0}^{m-l} P_{00}^l = \sum_{l=0}^{n} P_{00}^l \sum_{m=l}^{n} f_{i0}^{m-l} = \sum_{l=0}^{n} P_{00}^l \sum_{r=0}^{n-l} f_{i0}^r,$$

где f_{i0}^r есть вероятность достичь 0 из i в первый раз на r-м шаге. Так как $\sum_{r=0}^{n-l} f_{i0}^r \leqslant 1$, то

$$G_{i0}^n \leqslant \sum_{l=0}^{n} P_{00}^l = G_{00}^n. \ \blacksquare$$

Изящный и полезный критерий возвратности составляет содержание следующей теоремы.

Теорема 1.1. *Если*

$$\mathrm{M}|X_k| = \mathrm{M}|X_1| = \sum_{j=-\infty}^{\infty} |j| P_{0j} < \infty, \qquad k = 2,\ 3,\ \ldots, \tag{1.3}$$

и

$$\mu = \mathrm{M}(X_k) = \mathrm{M}(X_1) = \sum_{j=-\infty}^{\infty} j P_{0j} = 0, \tag{1.4}$$

то марковская цепь $\{S_n\}$ является возвратной.

Замечание. Поскольку $\mathrm{M}(X_1) = 0$ и X_1 — невырожденная случайная величина, то она принимает как положительные, так и отрицательные значения с положительными вероятностями. Условие неприводимости гарантирует, что марковская цепь, порождаемая последовательностью $\{S_n,\ n = 0, 1, 2, \ldots\}$, неприводима (состоит из одного класса). Пространство состояний цепи состоит из всех целых чисел: положительных, отрицательных и нуля. Наконец, в соответствии со следствием 5 гл. 2 для доказательства возвратности рассматриваемой цепи достаточно установить возвратность хотя бы одного (скажем, нулевого) состояния.

Доказательство. В силу леммы 1.1 $G_{0j}^n \leqslant G_{00}^n$ при всех целых j. Это неравенство сохраняется и при осреднении:

$$\frac{1}{2M+1}\sum_{j=-M}^{M} G_{0j}^n \leqslant G_{00}^n. \tag{1.5}$$

Но

$$\sum_{j=-M}^{M} G_{0j}^n = \sum_{|j|\leqslant M}\sum_{m=0}^{n} P_{0j}^m = \sum_{m=0}^{n}\sum_{|j|\leqslant M} P_{0j}^m \geqslant \sum_{m=0}^{n}\sum_{|j/m|\leqslant M/n} P_{0j}^m; \tag{1.6}$$

последнее неравенство имеет место просто потому, что $m \leqslant n$. Сравнив (1.5) и (1.6), мы видим, что

$$G_{00}^n \geqslant \frac{1}{2M+1}\sum_{m=0}^{n}\sum_{|j/M|\leqslant M/n} P_{0j}^m. \tag{1.7}$$

Далее, по определению

$$P_{0j}^m = \mathrm{P}\{S_m = j \mid S_0 = 0\}. \tag{1.8}$$

Поскольку S_k представляет собой сумму k независимых одинаково распределенных случайных величин с конечным средним $\mu = \mathrm{M}(X_1) = \mathrm{M}(S_k) = 0$, можно применить закон больших чисел (см. § 1 гл. 1, стр. 20), согласно которому для любого $\varepsilon > 0$

$$\mathrm{P}\left\{\left|\frac{S_m - m\mu}{m}\right| \leqslant \varepsilon\right\} = \mathrm{P}\left\{\left|\frac{S_m}{m}\right| \leqslant \varepsilon\right\} \to 1 \quad \text{при} \quad m \to \infty. \tag{1.9}$$

Из (1.8) находим

$$\mathrm{P}\{|S_m| \leqslant m\varepsilon\} = \sum_{|j|\leqslant [m\varepsilon]} P_{0j}^n.$$

(Здесь $[h]$ обозначает наибольшее целое число, не превышающее h; поэтому $h-1 < [h] \leqslant h$.) Предельное соотношение (1.9) можно представить в следующем эквивалентном ему виде:

$$H_m(\varepsilon) = \sum_{|j|\leqslant [m\varepsilon]} P_{0j}^m \to 1 \quad \text{при} \quad m \to \infty. \tag{1.10}$$

Положим теперь в (1.7) $M = [n\varepsilon]$, где $\varepsilon > 0$. Тогда

$$G_{00}^n \geqslant \frac{1}{2[n\varepsilon]+1}\sum_{m=0}^{n}\sum_{|j|\leqslant m[n\varepsilon]/n} P_{0j}^m = \frac{1}{2[n\varepsilon]+1}\sum_{m=0}^{n}\sum_{|j|\leqslant [m\varepsilon]} P_{0j}^m =$$

$$= \frac{n+1}{2[n\varepsilon]+1}\,\frac{1}{n+1}\sum_{m=0}^{n} H_m(\varepsilon). \tag{1.11}$$

Из (1.10) следует, что

$$\frac{1}{n+1}\sum_{m=0}^{n} H_m(\varepsilon)\to 1 \quad \text{при } n\to\infty. \tag{1.12}$$

Кроме того,

$$\lim_{n\to\infty}\frac{n+1}{2[n\varepsilon]+1}=\lim_{n\to\infty}\frac{n+1}{2n\varepsilon+1}=\frac{1}{2\varepsilon}$$

Из (1.11) и (1.12) заключаем, что

$$\lim_{n\to\infty} G_{00}^{n}\geqslant\frac{1}{2\varepsilon}.$$

Так как $\varepsilon>0$ может быть выбрано сколь угодно малым, то из последнего соотношения следует, что

$$\sum_{k=0}^{\infty} P_{00}^{k}=\lim_{n\to\infty} G_{00}^{n}=\infty.$$

Наконец, обращаясь к теореме 5.1 гл. 2, вспоминаем, что равенство $G_{00}=\infty$ эквивалентно утверждению, что нулевое состояние возвратно. ■

Отметим, что мы не использовали всей силы предположения о том, что X_i имеет конечное среднее значение. Нам понадобился лишь слабый закон больших чисел, выполнение которого в форме соотношения (1.9) *достаточно для справедливости* утверждения теоремы.

В следующей теореме, являющейся частичным обращением теоремы 1.1, существование конечного среднего играет более важную роль.

Теорема 1.2. *Если*

$$\mathrm{M}(|X_i|)=\sum_{j=-\infty}^{\infty}|j|P_{0j}<\infty, \qquad i=1,\ 2,\ \ldots, \tag{1.13}$$

и

$$\mu=\mathrm{M}(X_i)=\sum_{j=-\infty}^{\infty} jP_{0j}\neq 0,$$

то марковская цепь $\{S_n\}$ является невозвратной.

Доказательство. Пусть A обозначает событие $\{S_n=0\}$.

Воспроизведем известный нам критерий возвратности (см. теорему 7.1 гл. 2) в форме

$$\mathrm{P}\left\{\begin{matrix} A_n \text{ наступает при беско-} \\ \text{нечно многих } n\end{matrix}\right\}=\begin{cases} 1 & \text{тогда и только тогда, когда марковская цепь } \{S_n\} \text{ возвратна,} \\ 0 & \text{тогда и только тогда, когда марковская цепь } \{S_n\} \text{ невозвратна.}\end{cases} \tag{1.14}$$

Усиленный закон больших чисел утверждает, что

$$P\left\{\lim_{n\to\infty} \frac{S_n}{n} = \mu\right\} = 1. \tag{1.15}$$

Поскольку $\mu \neq 0$, рассмотрим события

$$C_n = \left\{\left|\frac{S_n}{n} - \mu\right| > \frac{|\mu|}{2}\right\}, \quad n = 1, 2, \ldots .$$

Пусть C — событие, состоящее в том, что C_n наступает при бесконечно многих n. Мы найдем вероятность $P\{C\}$. Всякая реализация процесса, для которой $\lim\limits_{n\to\infty} S_n/n = \mu$, очевидно, не может принадлежать событию C. Но, согласно (1.15), реализации процесса, для которых $\lim\limits_{n\to\infty} S_n/n = \mu$, имеют вероятность 1. Следовательно, вероятность события, дополнительного к C, равна 1, или $P\{C_n$ происходит при бесконечно многих $n\} = 0$. Ясно, что наступление события A_n влечет наступление события C_n, т. е. $A_n \subset C_n$. Поэтому

$$P\{A_n \text{ наступает при бесконечно многих } n\} \leqslant P\{C\} = 0.$$

Принимая во внимание критерий (1.14), заключаем, что марковская цепь $\{S_n\}$ невозвратна.

§ 2. ЛОКАЛЬНЫЕ ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Заметим, что если марковская цепь $\{S_n\}$ возвратна, то она может быть только возвратной нулевой. Это так, потому что в силу пространственной однородности

$$\pi_i = \lim_{n\to\infty} P_{ii}^n = \lim_{n\to\infty} P_{00}^n = \pi_0 \quad \text{при всех } i.$$

Если же $\pi_0 > 0$, то $\sum\limits_{i=-\infty}^{\infty} \pi_i = \infty$, что невозможно. Следовательно, $\pi_i = 0$ при всех i.

Итак, марковская цепь $\{S_n\}$ является либо возвратной нулевой, либо невозвратной; следовательно, при всех j

$$P_{0j}^n \to 0 \quad \text{при} \quad n \to \infty.$$

Представляет интерес оценка скорости сходимости к нулю. Результат подобного рода относится к так называемым «локальным предельным теоремам». Приступая к этой задаче, введем характеристическую функцию

$$\phi_{X_k}(\theta) = \sum_{\nu=-\infty}^{\infty} P_{0\nu} \exp(i\nu\theta) = M[\exp(iX_k\theta)], \; k = 1, 2, \ldots, -\pi \leqslant \theta < \pi, \tag{2.1}$$

где ряд сходится абсолютно и равномерно. Мы утверждаем, что

$$[\phi(\theta)]^n = \sum_{\nu=-\infty}^{\infty} P_{0\nu}^n \exp(i\nu\theta), \quad -\pi \leqslant \theta < \pi. \tag{2.2}$$

Действительно, X_k, $k = 1, 2, \ldots, n$, — независимые одинаково распределенные случайные величины, так что

$$P_{0j}^n = \mathrm{P}\{S_n = j\},$$

$$\sum_{\nu=-\infty}^{\infty} P_{0\nu}^n \exp(i\nu\theta) = \mathrm{M}[\exp(iS_n\theta)] = \mathrm{M}[\exp[i\theta(X_1 + \ldots + X_n)]] =$$

$$= \prod_{k=1}^{n} \mathrm{M}[\exp(i\theta X_n)] = \prod_{k=1}^{n} [\phi_{X_k}(\theta)] = [\phi_{X_1}(\theta)]^n$$

(см. стр. 15). Отметим, далее, что

$$\int_{-\pi}^{\pi} e^{i(j-k)\theta}\, d\theta = \begin{cases} 0, & \text{если } j \neq k, \\ 2\pi, & \text{если } j = k; \end{cases} \tag{2.3}$$

здесь $i = \sqrt{-1}$. Таким образом, функции $(2\pi)^{-\frac{1}{2}} e^{ij\theta}$ (j — любое целое) образуют ортонормированную систему. Поэтому, умножив обе части (2.2) на $(2\pi)^{-1} e^{-ik\theta}$ и проинтегрировав по θ на отрезке $[-\pi, \pi]$, получим

$$P_{0k}^{(n)} = (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-ik\theta} [\phi(\theta)]^n\, d\theta, \tag{2.4}$$

так как в правой части остается лишь член, соответствующий $\nu = k$.

Прежде чем сформулировать и доказать результат, касающийся скорости сходимости P_{0j}^n к нулю при $n \to \infty$, введем некоторые понятия, которые нам понадобятся для этого, и обсудим их свойства. Будем говорить, что X является периодической случайной величиной, если все значения, которые X может принимать с положительной вероятностью, содержатся в множестве

$$X = \omega + rc, \qquad r = 0, \pm 1, \pm 2, \ldots,$$

где ω и c — целые и $|c| \neq 1$.

Отметим, что из утверждения «марковская цепь $\{S_n\}$ является периодической» следует, что X_k — периодические случайные величины, но не обратно. (Доказать это.) Напомним еще, что если марковская цепь $\{S_n\}$ — непериодическая, то наименьшая аддитивная группа, порождаемая целыми числами i, для которых $\mathrm{P}\{X_1 = i\} > 0$, есть группа всех целых чисел.

Случайная величина

$$X_k = \begin{cases} +1 & \text{с вероятностью } p, \\ -1 & \text{с вероятностью } q \end{cases}$$

дает пример периодической случайной величины. В самом деле, ее возможные значения можно представить в виде

$$X = 1 + 2r, \qquad r = 0, \pm 1, \pm 2, \ldots .$$

(Здесь $c = 2$, $\omega = 1$.)

Лемма 2.1. *Случайные величины X_k являются периодическими тогда и только тогда, когда их характеристическая функция $\phi(\theta)$ обладает следующим свойством:*

$$|\phi(\theta_0)| = 1 \tag{2.5}$$

для некоторого $\theta_0 \neq 0$, $-\pi \leqslant \theta_0 \leqslant \pi$.

Доказательство. Предположим, что при $\theta_0 = h \neq 0$ $(-\pi \leqslant h \leqslant \pi)$

$$|\phi(h)| = 1.$$

Тогда существует такое действительное число w, что $\phi(h) = e^{iwh}$, а следовательно,

$$1 = e^{-ihw}\phi(h) = \sum_{j=-\infty}^{\infty} P_{0j} e^{i(j-w)h} =$$

$$= \sum_{j=-\infty}^{\infty} P_{0j} \cos(j-w)h + i \sum_{j=-\infty}^{\infty} P_{0j} \sin(j-w)h.$$

Отсюда

$$1 = \sum_{j=-\infty}^{\infty} P_{0j} \cos(j-w)h.$$

Так как $|\cos x| \leqslant 1$ при всех x, то для всех состояний j, достижимых из нулевого, т. е. для тех j, для которых $P_{0j} > 0$, с необходимостью

$$\cos(j-w)h = 1.$$

Решения последнего уравнения задаются формулой

$$(j-w)h = 2\pi r, \qquad r = 0, \pm 1, \pm 2, \ldots .$$

Это означает, что всякое j, достижимое из нуля, можно представить в виде $j = w + (2\pi/h)r$, $r = 0, \pm 1, \pm 2, \ldots$, где, очевидно, $|c| = |2\pi/h| \neq 1$. Ясно, что X_k может принимать только эти значения j, т. е. X_k — периодическая случайная величина.

Докажем теперь необходимость. Если все возможные значения содержатся в множестве

$$X_k = \omega + rc, \qquad r = 0, \pm 1, \ldots$$

(ω и c — целые, $|c| \neq 1$), то

$$\phi(\theta) = \sum_{r=-\infty}^{\infty} P_{0,\,\omega+rc} e^{i(\omega+rc)\theta}$$

и

$$\sum_{r=-\infty}^{\infty} P_{0,\,\omega+rc} = 1.$$

Положим $\theta_0 = (2\pi/c)$. Так как c — целое и $|c| \neq 1$, то $\theta_0 \neq 0$, $-\pi \leqslant \theta_0 \leqslant \pi$ и

$$\phi\left(\frac{2\pi}{c}\right) = \sum_{r=-\infty}^{\infty} P_{0,\,\omega+rc} e^{i\omega(2\pi/c)} e^{i2\pi r} = e^{i\omega(2\pi/c)} \sum_{r=-\infty}^{\infty} P_{0,\,\omega+rc} = e^{i\omega(2\pi/c)}.$$

Таким образом, (2.5) выполняется для $\theta_0 = (2\pi/c) \neq 0$, $-\pi \leqslant \theta_0 < \pi$. ■

Далее мы будем предполагать, если не оговорено противное, что X_k является непериодической случайной величиной.

Лемма 2.2. *Существует константа $\lambda > 0$, такая, что*

$$1 - \operatorname{Re}\phi(\theta) \geqslant \lambda\theta^2, \qquad -\pi \leqslant \theta \leqslant \pi. \tag{2.6}$$

Доказательство. Заметим, что

$$1 - \operatorname{Re}\phi(\theta) = 1 - \sum_{j=-\infty}^{\infty} P_{0j} \cos j\theta = \sum_{j=-\infty}^{\infty} (1 - \cos j\theta) P_{0j}.$$

Нам понадобится тождество

$$1 - \cos\alpha = 2\sin^2\frac{\alpha}{2},$$

имеющее место при любом действительном α. С помощью этого тождества приходим к неравенству

$$1 - \operatorname{Re}\phi(\theta) = 2\sum_{j=-\infty}^{\infty}\left(\sin^2\frac{j\theta}{2}\right)P_{0j} \geqslant 2\sum_{j=-L}^{L}\left(\sin^2\frac{j\theta}{2}\right)P_{0j}, \tag{2.7}$$

справедливому для любого положительного L.

Мы воспользуемся хорошо известным неравенством

$$|\sin x| \geqslant \frac{2|x|}{\pi}, \qquad -\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2}. \tag{2.8}$$

Доказать это неравенство можно, например, так. Функция $(\sin x)/x$ является убывающей при $0 \leqslant x \leqslant \pi/2$. Действительно,

$$\frac{d}{dx}\left(\frac{\sin x}{x}\right) = \frac{x\cos x - \sin x}{x^2}.$$

Но $1 \leqslant \sec^2 x$, и интегрирование обеих частей этого неравенства от 0 до x дает $x \leqslant \operatorname{tg} x$. Так как $\cos x > 0$ при $0 \leqslant x < \pi/2$, то $(d/dx)[(\sin x)/x] < 0$; значит, $(\sin x)/x$ убывает при $0 \leqslant x \leqslant \pi/2$. Следовательно, при $0 \leqslant x < \frac{\pi}{2}$

$$\frac{\sin x}{x} \geqslant \frac{\sin(\pi/2)}{\pi/2} = \frac{2}{\pi}, \qquad \text{или} \qquad \sin x \geqslant \frac{2}{\pi} x.$$

Поскольку обе части последнего неравенства являются нечетными функциями, ясно, что (2.8) выполняется.

Учитывая (2.8), из неравенства (2.7) получаем

$$1 - \operatorname{Re}\phi(\theta) \geqslant 2 \sum_{|j| \leqslant L} \left(\frac{j\theta}{\pi}\right)^2 P_{0j} = \frac{2}{\pi^2}\theta^2 \sum_{|j| \leqslant L} j^2 P_{0j}. \tag{2.9}$$

Это неравенство справедливо для значений θ, таких, что $|j\theta| \leqslant \pi$. Но если $|j| \leqslant L$, то неравенство $|j\theta| \leqslant \pi$ будет выполнено, если

$$|\theta| \leqslant \frac{\pi}{L}. \tag{2.10}$$

При достаточно большом L должно существовать по крайней мере одно j, $|j| \leqslant L$, для которого $P_{0j} > 0$. Выбирая L именно таким имеем

$$C = 2\pi^{-2} \sum_{|j| \leqslant L} j^2 P_{0j} > 0$$

и

$$1 - \operatorname{Re}\phi(\theta) \geqslant C\theta^2 \tag{2.11}$$

при всех $|\theta| \leqslant \pi/L$.

До сих пор мы не пользовались непериодичностью случайной величины X_k. Это допущение нам потребуется для оценки $1 - \operatorname{Re}\phi(\theta)$ при $|\theta| > \pi/L$. Мы знаем, что непериодичность X_k эквивалентна тому, что равенство $|\phi(\theta)| = 1$ выполняется на отрезке $[-\pi, \pi]$ только при $\theta = 0$ (лемма 2.1). Но $|\phi(\theta)| \leqslant 1$ при всех θ, как характеристическая функция. Следовательно,

$$1 - \operatorname{Re}\phi(\theta) \geqslant 1 - |\phi(\theta)| > 0 \tag{2.12}$$

при $\theta \neq 0$, $-\pi \leqslant \theta \leqslant \pi$. Так как разность $1 - \operatorname{Re}\phi(\theta)$ является непрерывной функцией аргумента θ на отрезке $[-\pi, \pi]$, то

$$m = \min_{\pi \geqslant |\theta| \geqslant \pi/L} \{1 - \operatorname{Re}\phi(\theta)\}$$

существует и в силу (2.12) положителен. По самому смыслу величины m неравенство

$$1 - \operatorname{Re} \phi(\theta) \geqslant m \frac{\theta^2}{\pi^2} \tag{2.13}$$

выполняется при всех θ, таких, что $\pi \geqslant |\theta| \geqslant \pi/L$. Положим теперь

$$\lambda = \min(C, \; m/\pi^2).$$

Тогда (2.6) выполняется при всех $|\theta| \leqslant \pi$. ■

Мы подготовили все, что требуется для того, чтобы сформулировать и доказать теорему об оценке скорости сходимости вероятностей P_{0j}^n к нулю при $n \to \infty$.

Теорема 2.1. *Если случайные величины $\{X_k\}$ непериодические, то при некоторой константе $A > 0$ (не зависящей от j и n)*

$$P_{0k}^n \leqslant \frac{A}{\sqrt{n}} \tag{2.14}$$

для всех целых j и $n \geqslant 1$.

Доказательство. Из (2.4) следует, что

$$P_{0k}^{2n} \leqslant (2\pi)^{-1} \int_{-\pi}^{\pi} |\phi(\theta)|^{2n} \, d\theta. \tag{2.15}$$

Покажем, что $|\phi(\theta)|^2$ является характеристической функцией некоторой целочисленной случайной величины. В самом деле, так как

$$\phi(\theta) = \mathrm{M}[\exp(iX_k\theta)] \quad \text{для любого} \quad k = 1, 2, \ldots$$

и

$$\overline{\phi(\theta)} = \mathrm{M}[\exp(-iX_l\theta)] \quad \text{для любого} \quad l = 1, 2, \ldots,$$

то при $k \neq l$

$$|\phi(\theta)|^2 = \phi(\theta)\overline{\phi(\theta)} = \mathrm{M}[\exp(iX_k\theta)]\,\mathrm{M}[\exp(-iX_l\theta)] = \\ = \mathrm{M}[\exp(i(X_k - X_l)\theta)],$$

так что $|\phi(\theta)|^2$ является характеристической функцией целочисленной случайной величины $(X_k - X_l)$, где X_k и X_l — независимые и одинаково распределенные случайные величины. Непериодичность X_k эквивалентна непериодичности разности $X_k - X_l$. Это является непосредственным следствием леммы 2.1. Пусть $\psi(\theta) = |\phi(\theta)|^2$. Применим теперь результат леммы 2.2 к действительной характеристической функции $\psi(\theta)$. Это даст неравенство $1 - \psi(\theta) \geqslant \lambda\theta^2$, справедливое при всех θ из отрезка $[-\pi, \pi]$ и некотором $\lambda > 0$. Перепишем это соотношение в виде

$$\psi(\theta) \leqslant 1 - \lambda\theta^2 \leqslant \exp(-\lambda\theta^2), \tag{2.16}$$

где последнее неравенство следует из соотношения $1-y \leqslant e^{-y}$, $y \geqslant 0$, которое в свою очередь следует из тривиального неравенства $e^{-\xi} \leqslant 1$ как результат интегрирования последнего в пределах от 0 до y.

Из (2.16) интегрированием получаем

$$\int_{-\pi}^{\pi} \psi(\theta)^n \, d\theta \leqslant \int_{-\pi}^{\pi} \exp(-n\lambda\theta^2) \, d\theta = \frac{1}{\sqrt{n}} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp(-\lambda\alpha^2) \, d\alpha \leqslant$$

$$\leqslant \left(\sqrt{n}\right)^{-1} \int_{-\infty}^{\infty} \exp(-\lambda\alpha^2) \, d\alpha. \tag{2.17}$$

Сравнение соотношений (2.15) и (2.17) дает

$$P_{0k}^{2n} \leqslant \frac{1}{\sqrt{n}} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-\lambda\alpha^2) \, d\alpha = \frac{A_1}{\sqrt{2n}}, \tag{2.18}$$

где

$$A_1 = \sqrt{2} \, (2\pi)^{-1} \int_{-\infty}^{\infty} \exp(-\lambda\alpha^2) \, d\alpha.$$

Поскольку $|\phi(\theta)| \leqslant 1$, мы также получаем

$$P_{0k}^{2n+1} \leqslant (2\pi)^{-1} \int_{-\pi}^{\pi} |\phi(\theta)|^{2n+1} \, d\theta \leqslant (2\pi)^{-1} \int_{-\pi}^{\pi} |\phi(\theta)|^{2n} \, d\theta =$$

$$= (2\pi)^{-1} \int_{-\pi}^{\pi} \{\psi(\theta)\}^n \, d\theta \leqslant \frac{A_1}{\sqrt{2n}} \leqslant \frac{\sqrt{2} A_1}{\sqrt{2n+1}}. \tag{2.19}$$

Положим теперь $A = \sqrt{2} A_1$; (2.18) и (2.19) совместно дают (2.14), что и требовалось. ■

Нужно подчеркнуть, что оценка (2.14) справедлива как для возвратной марковской цепи $\{S_n\}$, так и для случая, когда эта цепь невозвратна. В этой связи полезно еще раз обратиться к классической вероятностной модели бросаний монеты. В этой модели

$$X_k = \begin{cases} 1 & \text{с вероятностью } p, \\ -1 & \text{с вероятностью } q, \end{cases}$$

а последовательность $\{S_n\}$ представляет собой марковскую цепь с переходными вероятностями специального вида[1]):

$$P_{ij} = \begin{cases} p, & \text{если } j = i+1, \\ q, & \text{если } j = i-1, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Как мы видели (формулы (6.1)—(6.2) в гл. 2),

$$P_{00}^{2n} = \binom{2n}{n} p^n q^n,$$

а асимптотическая формула для P_{00}^{2n} имеет вид

$$P_{00}^{2n} \sim \frac{1}{\sqrt{n\pi}} (4pq)^n. \qquad (2.20)$$

Из (2.20) видно, что если $p \neq \frac{1}{2}$, то P_{00}^{2n} экспоненциально (со скоростью геометрической прогрессии) стремится к нулю. Формула (2.14) дает точную оценку, если $p = q = \frac{1}{2}$.

Приведенный нами пример является типичным для общей ситуации. Имеются значительные уточнения оценки (2.14) при дополнительных ограничениях $M(X_k^2) < \infty$ и $M(X_k) = \mu = 0$. В этом случае, согласно теореме 1.1, марковская цепь $\{S_n\}$ является возвратной, а точнее, возвратной нулевой. Следовательно, $\sum\limits_{n=0}^{\infty} P_{00}^n = \infty$, а $P_{0j}^n \to 0$. С помощью центральной предельной теоремы доказывается, что

$$\lim_{n\to\infty} \sqrt{n}\, P_{0j}^n = B, \qquad (2.21)$$

где B — конечная положительная константа, не зависящая от j. Доказательство этого результата выходит за рамки нашей книги. Мы отсылаем читателя к монографиям [1, 2], где подробно излагаются результаты этого характера. Если $M(X_k^2) = \infty$ и $M(|X_k|^{1+\delta}) < \infty$ при $0 < \delta < 1$, но $M(|X_k|^{1+\xi}) = \infty$ при $\xi > \delta$, то часто (2.21) заменяется другой точной асимптотической формулой

$$\lim_{n\to\infty} n^{1/(1+\delta)} P_{0j}^n = B. \qquad (2.22)$$

Последний результат справедлив в случае, когда центральная предельная теорема неприменима, но имеет место притяжение к соответствующему устойчивому закону. Теория устойчивых законов

---

[1]) Заметим, что здесь X_k — периодическая случайная величина. — *Прим. перев.*

представляет собой довольно сложную область теории вероятностей, играющую важную роль в приложениях к физике и астрономии. Элементарный характер нашей книги не позволяет нам даже поверхностно затронуть эти вопросы. Мы ограничимся лишь констатацией существования этой теории и рекомендуем читателю обратиться к ней при дальнейшем изучении теории вероятностей.

§ 3. ПРАВЫЕ РЕГУЛЯРНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ МАРКОВСКИХ ЦЕПЕЙ

Для специальных классов марковских цепей обычно удается получить более сильные результаты, касающиеся свойства возвратности, задач о времени пребывания, нахождения законов распределения различных функционалов от процессов и др. В этом параграфе мы получим уточнение свойств правых регулярных последовательностей (теорема 5.2 гл. 5) для марковской цепи $\{S_n\}$.

Если марковская цепь является неприводимой и возвратной, то единственным правым регулярным вектором является постоянный вектор (теорема 5.2 гл. 5). Если же марковская цепь описывает последовательность сумм независимых случайных величин, то этот результат можно распространить на непериодический невозвратный случай, если мы потребуем, чтобы правый регулярный вектор был ограниченным. Более точно, мы докажем следующее.

Теорема 3.1. *Если марковская цепь $\{S_n\}$ с матрицей переходных вероятностей* **P** *неприводима, и если последовательность $\{y_j\}$— правая регулярная, т. е. $\{y_j\}$ удовлетворяет условиям*

$$y_j \geqslant 0, \quad j = 0, \pm 1, \pm 2, \ldots, \tag{3.1}$$

и

$$\sum_{j=-\infty}^{\infty} P_{ij} y_j = y_i, \quad i = 0, \pm 1, \pm 2, \ldots, \tag{3.2}$$

причем $\{y_j\}$ ограничена, то $y_j \equiv \text{const}$ при всех j.

Доказательство. Пусть ограниченная последовательность $\{y_j\}$ удовлетворяет условиям (3.1) и (3.2). Пусть k_0 — любое состояние, отличное от нулевого и достижимое из него. Тогда существует такое n, что $P_{0k_0}^n > 0$. Зафиксируем k_0 и положим

$$z_j = y_j - y_{j-k_0}, \quad j = 0, \pm 1, \pm 2, \ldots.$$

Используя пространственную однородность процесса, мы можем написать

$$\sum_{j=-\infty}^{\infty} P_{i,j} y_{j-k_0} = \sum_{k=-\infty}^{\infty} P_{i,k_0+k} y_k = \sum_{k=-\infty}^{\infty} P_{i-k_0,k} y_k = y_{i-k_0},$$

т. е. последовательность $\{u_j\}=\{y_{j-k_0}\}_{j=-\infty}^{\infty}$ также удовлетворяет (3.2). Следовательно, $\{z_j\}$ удовлетворяет условиям (3.2). Очевидно, $\{z_j\}$ ограничена, поскольку ограничена $\{y_j\}$. Пусть

$$M=\sup_j z_j<\infty \quad \text{и} \quad \sup_j |z_j|=M'. \tag{3.3}$$

Так как $\{z_j\}$ ограничена, мы можем выбрать последовательность целых чисел $\{r_n\}$, для которой

$$\lim_{n\to\infty} z_{r_n}=M. \tag{3.4}$$

Опять же в силу ограниченности $\{z_j\}$ из последовательности $\{r_n\}$ можно выбрать подпоследовательность $\{r_n^{(1)}\}$, такую, что существует $\lim\limits_{n\to\infty} z_{1+r_n^{(1)}}$. В свою очередь из $\{r_n^{(1)}\}$ можно выбрать подпоследовательность $\{r_n^{(-1)}\}$, для которой существует предел $\lim\limits_{n\to\infty} z_{-1+r_n^{(-1)}}$. Затем из $\{r_n^{(-1)}\}$ выбираем подпоследовательность $\{r_n^{(2)}\}$, такую, что существует $\lim\limits_{n\to\infty} z_{2+r_n^{(2)}}$. Продолжая таким образом, мы получим ряд последовательностей $\{r_n^{(-2)}\}$, $\{r_n^{(3)}\}$, $\{r_n^{(-3)}\}$, ..., каждая из которых является подпоследовательностью предыдущей. Существует последовательность, а именно $\{s_n=r_n^{(-n)}\}$, которая является подпоследовательностью каждой из последовательностей

$$\{r_n\},\ \{r_n^{(1)}\},\ \{r_n^{(-1)}\},\ \{r_n^{(2)}\},\ \{r_n^{(-2)}\},\ \{r_n^{(3)}\},\ \{r_n^{(-3)}\},\ \ldots,$$

начиная с некоторого члена. Действительно, $\{s_n\}$ является подпоследовательностью последовательности $\{r_n^{(\rho)}\}$, начиная по крайней мере с $n\geqslant|\rho|$. (Построение последовательности $\{s_n\}$ носит название процесса диагонализации.) В силу этого свойства $\{s_n\}$ и факта существования соответствующих пределов мы можем утверждать, что

$$\lim_{n\to\infty} z_{j+s_n}=z_j^* \tag{3.5}$$

существует для каждого j.

По построению $\{r_n\}$ (см. (3.4)) имеем

$$z_0^*=\lim_{n\to\infty} z_{s_n}=M, \tag{3.6}$$

а в силу (3.3)

$$z_j^*=\lim_{n\to\infty} z_{j+s_n}\leqslant M, \qquad j=0,\ \pm 1,\ \pm 2,\ \ldots.$$

Как мы уже видели,

$$\sum_{j=-\infty}^{\infty} P_{i+s_n,j}\, z_j=z_{i+s_n}.$$

Переходя к пределу в обеих частях при $n \to \infty$, получаем

$$\lim_{n\to\infty} z_{i+s_n} = \lim_{n\to\infty} \sum_{j=-\infty}^{\infty} P_{i+s_n,\, j} z_j = \lim_{n\to\infty} \sum_{j=-\infty}^{\infty} P_{i,\, j-s_n} z_j =$$

$$= \lim_{n\to\infty} \sum_{l=-\infty}^{\infty} P_{il} z_{l+s_n}. \qquad (3.7)$$

Мы сейчас покажем, что в правой части соотношения (3.7) можно перейти к пределу под знаком суммы. Для этого нам нужно показать, что для любого заданного $\varepsilon > 0$ существует целое число $n(\varepsilon)$, такое, что

$$\left| \sum_{l=-\infty}^{\infty} P_{il} \left(z_{l+s_n} - z_l^* \right) \right| \leqslant \varepsilon$$

при условии, что $n \geqslant n(\varepsilon)$. Выберем L настолько большим, что

$$\sum_{|l|>L} P_{il} < \frac{\varepsilon}{4M'}.$$

Затем выберем n_0 таким, что

$$\left| z_{l+s_n} - z_l^* \right| < \frac{\varepsilon}{2}$$

при $-L \leqslant t \leqslant L$ и всех $n \geqslant n_0$. Тогда при $n \geqslant n_0(\varepsilon)$

$$\left| \sum_{l=-\infty}^{\infty} P_{il} \left(z_{l+s_n} - z_l^* \right) \right| \leqslant$$

$$\leqslant 2M' \sum_{|l|>L} P_{il} + \sum_{|l|\leqslant L} P_{il} \left| z_{l+s_n} - z_l^* \right| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Установив правомерность перехода к пределу под знаком суммы, из (3.7) получаем

$$z_i^* = \sum_{l=-\infty}^{\infty} P_{il} z_l^*, \qquad (3.8)$$

т. е. z_j^* удовлетворяет условию (3.2), хотя не обязательно условию (3.1). Последовательно применяя соотношение (3.8), получаем

$$z_i^* = \sum_{j=-\infty}^{\infty} P_{ij}^n z_j^* \quad \text{при всех } n \geqslant 0.$$

Обращаясь к (3.6), мы видим, что предыдущее соотношение при $i = 0$ дает

$$\sum_{j=-\infty}^{\infty} P_{0j}^n z_j^* = z_0^* = M \quad \text{при всех } n \geqslant 0. \qquad (3.9)$$

Левая часть в (3.9) представляет собой взвешенное среднее чисел, каждое из которых $\leqslant M$. Поэтому равенство (3.9) может иметь место только в том случае, если для всех j, таких, что $P_{0j}^n > 0$ при некотором $n \geqslant 0$, величина z_j^* равна M. В частности, по определению k_0

$$z_{k_0}^* = M, \quad z_{2k_0}^* = M, \ \ldots, \ z_{tk_0}^* = M, \ \ldots$$

для любого положительного целого t. В силу (3.5) n можно выбрать таким, чтобы все неравенства

$$\begin{gathered} z_{k_0+s_n} > M - \varepsilon, \\ z_{2k_0+s_n} > M - \varepsilon, \\ \vdots \\ z_{tk_0+s_n} > M - \varepsilon \end{gathered}$$

выполнялись одновременно. Складывая эти неравенства, получаем

$$\begin{aligned} t(M-\varepsilon) < z_{k_0+s_n} + z_{2k_0+s_n} + \ldots + z_{tk_0+s_n} = \\ = \left(y_{k_0+s_n} - y_{s_n}\right) + \left(y_{2k_0+s_n} - y_{k_0+s_n}\right) + \\ + \left(y_{3k_0+s_n} - y_{2k_0+s_n}\right) + \ldots + \left(y_{tk_0+s_n} - y_{(t-1)k_0+s_n}\right) = \\ = y_{tk_0+s_n} - y_{s_n}. \end{aligned}$$

Так как y_j ограничены, то существует такое $K > 0$, что $y_j < K$ для всех j. Тогда, очевидно,

$$t(M-\varepsilon) < 2K.$$

Поскольку это неравенство должно выполняться при любом целом $t > 0$ и любом $\varepsilon > 0$, M с необходимостью должно быть отрицательным либо равным нулю. Таким образом, $y_j - y_{j-k_0} = z_j \leqslant 0$, или

$$y_j \leqslant y_{j-k_0} \tag{3.10}$$

для всех j и любого k_0, достижимого из нуля.

До сих пор мы нигде не использовали условия (3.1), т. е. того факта, что $y_j \geqslant 0$. Нам потребовалось только, чтобы $|y_j|$ были ограничены. Следовательно, мы можем провести все предыдущие рассуждения для последовательности $\{y_j'\}$, у которой $y_j' = -y_j$. Разумеется, $\{y_j'\}$ остается ограниченной и удовлетворяет (3.2). В этом случае

$$y_j \geqslant y_{j-k_0} \tag{3.11}$$

для всех j и всех состояний k_0, достижимых из нулевого состояния. Сравнивая (3.10) и (3.11), получаем

$$y_j = y_{j-k_0}$$

для всех j и всех состояний k_0, достижимых из нулевого состояния. В частности,

$$y_0 = y_{-k_0} \tag{3.12}$$

для всех состояний k_0, достижимых из нулевого состояния. Теперь мы впервые за все доказательство воспользуемся предположением о неприводимости марковской цепи. Это предположение гарантирует, что все состояния достижимы из нулевого состояния. Следовательно, (3.12) выполняется для всех $k_0 = 0, \pm 1, \pm 2, \ldots$ и все члены последовательности $\{y_j\}$ равны константе y_0. ■

В заключение этой главы мы применим теорему 3.1 для доказательства обобщенной теоремы восстановления для сумм независимых одинаково распределенных целочисленных случайных величин. Мы будем далее предполагать, что суммы S_n являются непериодическими д. с. в., т. е. что наименьшая аддитивная группа, порождаемая целыми числами i, для которых $P\{X_1 = i\} > 0$, есть группа всех целых чисел.

Мы дадим прямое доказательство теоремы восстановления; однако несколько вспомогательных результатов, получаемых попутно, представляют самостоятельный интерес, а методы их доказательства типичны для задач о флуктуациях в теории сумм независимых д. с. в.

Теорема 3.2. *Если $S_n = X_1 + X_2 + \ldots + X_n$ являются непериодическими д. с. в. и $M(|X_1|) < \infty$, а $M(X_1) > 0$, то*

$$\lim_{j\to\infty} G_{ij} = \lim_{j\to\infty} \sum_{n=0}^{\infty} P^n_{0,\,j-i} = \frac{1}{M(X_1)}, \tag{3.13}$$

$$\lim_{j\to-\infty} G_{ij} = 0. \tag{3.14}$$

Для удобства разобьем доказательство на несколько этапов. Для любой величины a положим $a^+ = \max(a, 0)$ и $a^- = \min(a, 0)$. Пусть

$$M_n = \min(S_1, S_2, \ldots, S_n).$$

Поскольку последовательность M_n не возрастает, предел

$$\lim_{n\to\infty} M_n = \inf(S_1, S_2, \ldots) = M$$

существует, хотя возможно, что $M = -\infty$. Однако, поскольку $M(X_1) > 0$, усиленный закон больших чисел гарантирует нам, что

$$P\{M = -\infty\} \leqslant P\{S_n \leqslant 0 \text{ при бесконечно многих } n\} = 0.$$

Следовательно, M конечно с вероятностью 1.

Далее,

$$\begin{aligned} M(M_n) &= M\{\min(S_1, \ldots, S_n)\} = \\ &= M\{X_1 + \min(0, X_2, X_2 + X_3, \ldots, X_2 + \ldots + X_n)\} = \\ &= M(X_1) + M\{\min(0, X_2, X_2 + X_3, \ldots, X_2 + \ldots + X_n)\}. \end{aligned} \tag{3.15}$$

Так как X_i независимы и одинаково распределены, то последнее слагаемое есть не что иное, как $\mathrm{M}\left(M^-_{n-1}\right)$.

Устремляя n к ∞, получаем

$$\mathrm{M}(M)=\mathrm{M}(X_1)+\mathrm{M}(M^-). \tag{3.16}$$

(Читателю рекомендуется доказать возможность перехода к пределу под знаком математического ожидания.) Очевидно,

$$\mathrm{M}(M)=\mathrm{M}(M^+)+\mathrm{M}(M^-),$$

откуда с учетом (3.16) получаем

$$\mathrm{M}(M^+)=\mathrm{M}(X_1). \tag{3.17}$$

Наши рассуждения имеют одно слабое место: мы не знаем a priori, что $\mathrm{M}(M^-)>-\infty$, а только в этом случае соотношение (3.16) имеет смысл.

Чтобы восполнить этот пробел, нам понадобится следующая теорема, имеющая и самостоятельный интерес.

Теорема 3.3. *Пусть Z — неотрицательная целочисленная с. в., т. е. $\mathrm{P}\{Z=n\}=p_n$, $n=0,1,\ldots$, $\sum\limits_{n=0}^{\infty}p_n=1$, и $\varphi(\theta)$ — ее характеристическая функция. Предположим, что $[\varphi(\theta)-1]/i\theta=(1/i\theta)\sum\limits_{n=0}^{\infty}p_n\left(e^{in\theta}-1\right)$ сходится к α $(0<\alpha<\infty)$ при $\theta\downarrow 0$. Тогда $\mathrm{M}(Z)=\sum\limits_{n=0}^{\infty}np_n=\alpha$.*

Доказательство. Выделяя действительную и мнимую части выражения $(\varphi(\theta)-1)/i\theta$, мы заключаем, что

$$\lim_{\theta\downarrow 0}\sum_{n=0}^{\infty}p_n\frac{(1-\cos n\theta)}{\theta}=0 \quad \text{и} \quad \lim_{\theta\downarrow 0}\sum_{n=0}^{\infty}p_n\frac{\sin n\theta}{\theta}=\alpha. \tag{3.18}$$

Для любого фиксированного $\theta>0$ найдем наибольшее целое число $k=k(\theta)$, удовлетворяющее условию $\pi/2\geqslant k\theta>0$.

В силу (3.18) сумма $\sum\limits_{n=0}^{\infty}p_n(\sin n\theta)/\theta$ равномерно ограничена при достаточно малых θ. Представим эту сумму в виде

$$\sum_{n=0}^{\infty}p_n\frac{\sin n\theta}{\theta}=\sum_{n=0}^{k}p_n\frac{\sin n\theta}{\theta}+\sum_{n=k+1}^{\infty}p_n\frac{\sin n\theta}{\theta}. \tag{3.19}$$

Оценим снизу $(\sin n\theta)/\theta$, используя тот факт, что $(\sin\theta)/\theta$ является убывающей функцией при $0<\theta<\pi/2$ (см. стр. 178):

$$\frac{\sin\theta}{\theta}\geqslant\frac{\sin\pi/2}{\pi/2}=\frac{2}{\pi}, \qquad 0<\theta<\frac{\pi}{2}.$$

Из определения k следует, что $0<n\theta\leqslant\pi/2$ при $0\leqslant n\leqslant k$. Тогда предыдущее неравенство дает

$$\sum_{n=0}^{k}p_n\frac{\sin n\theta}{\theta}\geqslant\frac{2}{\pi}\sum_{n=0}^{k}np_n.$$

С другой стороны, для второй суммы в (3.19) имеем

$$\left|\sum_{n=\bar{k}+1}^{\infty} p_n \frac{\sin n\theta}{\theta}\right| \leqslant \frac{1}{\theta} \sum_{n=\bar{k}+1}^{\infty} p_n.$$

Но $1-(\sin n\theta)/n\theta \geqslant 1-(1/n\theta) > 1-2/\pi = b > 0$ для всех n, таких, что $n\theta > \pi/2$. Следовательно,

$$\frac{1}{\theta}\sum_{n=\bar{k}+1}^{\infty} p_n \leqslant \frac{1}{b\theta}\sum_{n=\bar{k}+1}^{\infty} p_n\left(1-\frac{\sin n\theta}{n\theta}\right) =$$

$$= \frac{1}{b\theta}\sum_{n=\bar{k}+1}^{\infty} p_n \frac{1}{\theta}\int_0^{\theta}(1-\cos n\xi)\,d\xi \leqslant$$

$$\leqslant \frac{1}{b\theta}\sum_{n=\bar{k}+1}^{\infty} p_n \int_0^{\theta}\frac{1-\cos n\xi}{\xi}\,d\xi =$$

$$= \frac{1}{b\theta}\int_0^{\theta}\sum_{n=\bar{k}+1}^{\infty} p_n\left[\frac{1-\cos n\xi}{\xi}\right]d\xi \leqslant$$ (перемена местами суммирования и интегрирования возможна, так как все члены неотрицательны)

$$\leqslant \frac{1}{b\theta}\int_0^{\theta}\sum_{n=0}^{\infty} p_n \frac{1-\cos n\xi}{\xi}\,d\xi =$$

$$= \frac{1}{b\theta}\int_0^{\theta}\operatorname{Im}\left[\frac{\varphi(\xi)-1}{i\xi}\right]d\xi.$$

Но $\operatorname{Im}([\varphi(\xi)-1]/i\xi) \to 0$ при $\xi \downarrow 0$ в соответствии с (3.18), и поэтому ее среднее значение стремится к нулю:

$$\lim_{\theta\downarrow 0}\frac{1}{\theta}\int_0^{\theta}\operatorname{Im}\left[\frac{\varphi(\xi)-1}{i\xi}\right]d\xi = 0.$$

(Читателю следует доказать это.)

Из полученных оценок следует, что вторая сумма в (3.19) стремится к нулю при $\theta \downarrow 0$. Поэтому сумма

$$\sum_{n=0}^{\bar{k}} np_n \qquad (\bar{k} = k(\theta))$$

равномерно ограничена при $\theta > 0$. Очевидно, $\bar{k} = k(\theta)$ возрастает до бесконечности при $\theta \downarrow 0$, а отсюда и следует сходимость ряда $\sum_{n=1}^{\infty} np_n = \mathrm{M}(Z)$. Чтобы завершить доказательство теоремы 3.3, мы должны показать, что

$$\sum_{n=0}^{\infty} np_n = \lim_{\theta\to 0} \frac{\varphi(\theta)-1}{i\theta}. \tag{3.20}$$

Для этого зададим $\varepsilon > 0$ и выберем $N(\varepsilon)$ таким, что $\sum\limits_{n=N+1}^{\infty} np_n < \varepsilon/2$. Последнее возможно, поскольку, как мы уже показали, ряд $\sum\limits_{n=0}^{\infty} np_n$ сходится. Далее,

$$\frac{\varphi(\theta)-1}{i\theta} - \sum_{n=0}^{\infty} np_n = \sum_{n=0}^{N} p_n\left(\frac{e^{in\theta}-1}{i\theta}-n\right) + \sum_{n=N+1}^{\infty} p_n\left(\frac{e^{in\theta}-1}{i\theta}\right) - \sum_{n=N+1}^{\infty} np_n.$$

Так как $|(e^{in\theta}-1)/i\theta| \leqslant n$, то вторая и третья суммы ограничены величиной $\varepsilon/2$. Отсюда

$$\overline{\lim_{\theta\downarrow 0}}\left|\frac{\varphi(\theta)-1}{i\theta} - \mathrm{M}(Z)\right| \leqslant \overline{\lim_{\theta\downarrow 0}}\left|\sum_{n=0}^{N} p_n\left[\frac{e^{in\theta}-1}{i\theta}-n\right]\right| + \varepsilon.$$

При фиксированном N сумма в правой части стремится к нулю, так как каждое ее слагаемое стремится к нулю. Поэтому

$$\overline{\lim_{\theta\downarrow 0}}\left|\frac{\varphi(\theta)-1}{i\theta} - \mathrm{M}(Z)\right| \leqslant \varepsilon.$$

Выражение в левой части представляет собой фиксированное неотрицательное число; $\varepsilon > 0$ можно выбрать произвольно малым. Отсюда вытекает справедливость соотношения (3.20). ■

Теперь мы можем дать строгое доказательство соотношения (3.17). Введем характеристические функции случайных величин M_n, X_1 и M_{n-1}^-:

$$\varphi_{M_n}(\theta) = \sum_{k=-\infty}^{\infty} e^{ik\theta}\mathrm{P}\{M_n = k\},$$

$$\varphi_{X_1}(\theta) = \sum_{k=-\infty}^{\infty} e^{ik\theta}\mathrm{P}\{X_1 = k\}$$

и

$$\varphi_{M_{n-1}^-}(\theta) = \sum_{k=-\infty}^{0} e^{ik\theta}\mathrm{P}\{M_{n-1}^- = k\}.$$

Отметим, что случайная величина M_{n-1}^- может принимать только неотрицательные целые значения. Очевидно, с. в. X_1 и $\widetilde{M}_{n-1}^- = \min(0, X_2, X_2+X_3, \ldots, X_2+X_3+\ldots+X_n)$ независимы. Кроме того, $\widetilde{M}_{n-1}^-$ и M_{n-1}^- одинаково распределены, так как $\widetilde{M}_{n-1}^-$ определяется через $X_2, X_3, \ldots, X_n$ точно так же, как M_{n-1}^- определяется через $X_1, \ldots, X_{n-1}$.

Так как $M_n = X_1 + \widetilde{M}_{n-1}^-$ (это соотношение неявно фигурирует в (3.15)), то

$$\varphi_{M_n}(\theta) = \varphi_{X_1}(\theta)\,\varphi_{\widetilde{M}_{n-1}^-}(\theta).$$

Из определений следует, что $M_n \to M$ и $M_{n-1}^- \to M^-$ при $n \to \infty$ (сходимость понимается как сходимость в смысле распределений). В силу критерия сходимости П. Леви (стр. 16) имеем

$$\varphi_M(\theta) = \varphi_{X_1}(\theta)\,\varphi_{M^-}(\theta). \tag{3.21}$$

Но

$$\varphi_M(\theta) = \sum_{k=-\infty}^{\infty} e^{ik\theta}\mathrm{P}\{M = k\} =$$

$$= \sum_{k=1}^{\infty} e^{ik\theta}\mathrm{P}\{M^+ = k\} + \sum_{k=-\infty}^{-1} e^{ik\theta}\mathrm{P}\{M^- = k\} + \mathrm{P}\{M = 0\}.$$

Так как

$$\mathrm{P}\{M = 0\} = \mathrm{P}\{M \geqslant 0\} + \mathrm{P}\{M \leqslant 0\} - 1 =$$

$$= \mathrm{P}\{M^- = 0\} + \mathrm{P}\{M^+ = 0\} - 1,$$

мы можем записать предыдущее соотношение в виде

$$\varphi_M(\theta) = \sum_{k=0}^{\infty} e^{ik\theta}\mathrm{P}\{M^+ = k\} + \sum_{-\infty}^{k=0} e^{ik\theta}\mathrm{P}\{M^- = k\} - 1 =$$

$$= \varphi_{M^+}(\theta) + \varphi_{M^-}(\theta) - 1.$$

Подставляя это соотношение в (3.21), получаем

$$\varphi_{M^+}(\theta) - 1 = \varphi_{M^-}(\theta)\left(\varphi_{X_1}(\theta) - 1\right). \tag{3.22}$$

Поделив обе части на $i\theta$ и устремив $\theta \to 0$, получаем

$$\lim_{\theta \to 0} \frac{\varphi_{X_1}(\theta) - 1}{i\theta} = \mathrm{M}(X_1),$$

так как $\mathrm{M}(|X_1|) < \infty$. (Формальное доказательство этого предельного соотношения проводится так же, как и доказательство соотношения (3.20) в теореме 3.3.) Очевидно, $\lim\limits_{\theta \to 0} \varphi_{M^-}(\theta) = 1$, и из (3.22) следует, что

$$0 \leqslant \lim_{\theta \to 0} \frac{\varphi_{M^+}(\theta) - 1}{i\theta} = \mathrm{M}(X_1) < \infty. \tag{3.23}$$

Так как M^+ — неотрицательная д. с. в., мы можем воспользоваться теоремой 3.3, из которой следует, что предел в (3.23) равен $\mathrm{M}(M^+)$ и, таким образом, $\mathrm{M}(M^+) = \mathrm{M}(X_1)$.

Нам потребуется следующая лемма.

Лемма 3.1. *Пусть e_i ($i \leqslant 0$) есть вероятность того, что процесс $\{S_n\}$, отправляясь из состояния i, на первом шаге окажется в одном из положительных* *состояний и после этого никогда не окажется ни в одном отрицательном состоянии. При $i > 0$ вероятности e_i полагаются равными нулю. Тогда*

$$\sum_{i=-\infty}^{0} e_i = \mathrm{M}(X_1).$$

Доказательство. Из самого определения вероятностей e_i следует, что

$$e_i = \begin{cases} \mathrm{P}\{\inf(S_1, S_2, \ldots) > -i\} = \mathrm{P}\{M^+ > -i\}, & i \leqslant 0, \\ 0, & i > 0. \end{cases}$$

Далее,

$$\sum_{i=-\infty}^{0} e_i = \sum_{i=-\infty}^{0} \mathrm{P}\{M^+ > -i\} = \sum_{j=0}^{\infty} \mathrm{P}\{M^+ > j\} =$$

$$= \sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} \mathrm{P}\{M^+ = k\} = \sum_{k=0}^{\infty} \mathrm{P}\{M^+ = k\} \sum_{j=0}^{k-1} 1 = \mathrm{M}(M^+) = \mathrm{M}(X_1).$$

Лемма доказана. ■

Пусть

$$V(i) = \mathrm{P}\{S_n > 0 \quad \text{для некоторого} \quad n \geqslant 1 \mid S_0 = i\},$$

т. е. $V(i)$ есть вероятность, отправляясь из состояния i, оказаться на положительной оси.

Рассмотрим те реализации процесса, которые выходят из состояния i и включают в себя неположительные состояния. По условию теоремы 3.2 $\mathrm{M}(X_1) = \mu > 0$; поэтому в силу закона больших чисел такая реализация с вероятностью 1 может быть в неположительных состояниях лишь конечное число раз. Вероятность того, что процесс в последний раз находился в неположительном состоянии на n-м шаге, равна $\sum_{k=-\infty}^{0} P_{ik}^{n} e_k = \sum_{-\infty}^{\infty} P_{ik}^{n} e_k$; последнее равенство следует из определения e_k для $k > 0$. Перебирая все возможности, связанные с последним моментом пребывания в неположительных состояниях, получаем важное соотношение:

$$V(i) = \sum_{n=1}^{\infty} \sum_{k=-\infty}^{\infty} P_{ik}^{(n)} e_k = \sum_{k=-\infty}^{\infty} e_k \left(\sum_{n=1}^{\infty} P_{ik}^{(n)} \right) =$$

$$= \sum_{k=-\infty}^{\infty} e_k \left(\sum_{n=0}^{\infty} P_{ik}^{(n)} - \delta_{ik} \right) = \sum_{k=-\infty}^{\infty} G_{ik} e_k - e_i \tag{3.24}$$

(δ_{ij} — символ Кронекера). Теперь мы можем доказать теорему восстановления.

Доказательство теоремы 3.2. Из определения G_{ij} следует, что

$$\sum_k P_{ik} G_{kj} = \sum_k P_{ik} \sum_{n=0}^{\infty} P_{kj}^{(n)} = \sum_{n=0}^{\infty} P_{ij}^{(n+1)} = G_{ij} - \delta_{ij}. \tag{3.25}$$

Так как $G_{ij} \leqslant G_{00}$ (лемма 1.1), то с помощью диагонализации (см. доказательство теоремы 3.1) мы можем выделить подпоследовательность $\{j_m\}$, такую, что предел

$$\lim_{m \to \infty} G_{ij_m} = \varphi_i$$

существует для каждого i.

В соотношении (3.25) перейдем к пределу при j, стремящемся к бесконечности по значениям подпоследовательности $\{j_m\}$; это даст

$$\sum_{k=-\infty}^{\infty} P_{ik}\varphi_k = \varphi_i, \tag{3.26}$$

причем $\{\varphi_i\}$ — ограниченная последовательность. (Читателю рекомендуется самостоятельно убедиться в возможности предельного перехода под знаком суммы.) В силу теоремы 3.1 ограниченная регулярная последовательность $\{\varphi_i\}$ является постоянной; пусть все ее члены равны константе α. Из (3.24) имеем

$$V(-j_m) = \sum_k G_{-j_m, k} e_k - e_{j_m}.$$

Так как $G_{-j_m, k} = G_{-k, j_m}$ и ряд $\sum\limits_k e_k$ сходится, то

$$\lim_{m \to \infty} V(-j_m) = \alpha \sum_{k=-\infty}^{\infty} e_k = \alpha \mathrm{M}(X_1) \quad \text{(в силу леммы 3.1).}$$

Но, очевидно, $\lim\limits_{m \to \infty} V(-j_m) = 1$. Следовательно, $\alpha = 1/\mathrm{M}(X_1)$. Если имеется другая последовательность $\tilde{j}_m \to \infty$, обладающая тем свойством, что $G_{i\tilde{j}_m}$ сходится при всех i, то точно так же можно показать, что предел $\lim\limits_{m \to \infty} G_{i\tilde{j}_m} = (\mathrm{M}(X_1)^{-1})$. Таким образом, мы установили справедливость соотношения (3.13). Предельное соотношение (3.14) доказывается аналогичным образом.

ЗАДАЧИ

1. Рассмотрим выборочное пространство, состоящее из n циклических перестановок набора $(a_1, a_2, \ldots, a_n)$; вероятность каждой перестановки равна $1/n$. Для каждой точки $\mathbf{x} = (a_k, a_{k+1}, \ldots, a_{k+n-1})$, где по определению $a_{n+l} = a_l$, $l = 1, \ldots, n-1$, пусть $N(\mathbf{x})$ обозначает число частичных сумм среди $\{a_k, a_k + a_{k+1}, a_k + a_{k+1} + a_{k+2}, \ldots, a_k + \ldots + a_{k+n}\}$, равных нулю, а $\tilde{M}(\mathbf{x})$ — число различных частичных сумм. Показать, что если $\sum\limits_{i=1}^{n} a_i = 0$, то $\mathrm{M}(1/N(\mathbf{x})) = \tilde{M}(\mathbf{x})/n$.

2. Пусть $X_1, X_2, \ldots, X_n, \ldots$ — независимые случайные величины, равномерно распределенные на отрезке $[0, 1]$. Показать, что при $0 \leqslant a \leqslant n$

$$\mathrm{P}\{X_1 + \ldots + X_n \leqslant a\} = \sum_{j=0}^{[a]} (-1)^j \binom{n}{j} \frac{(a-j)^n}{n!}.$$

$[a]$ обозначает, как обычно, наибольшее целое число, меньшее или равное a.

Указание: Применить индукцию по n.

3 (продолжение). Пусть r — такой индекс, что $X_1 + \ldots + X_{r-1} \leqslant a$, $X_1 + \ldots + X_r > a$. Показать, что

$$\mathrm{M}(r) = \sum_{j=0}^{[a]} (-1)^j \frac{(a-j)^j}{j!} e^{a-j}.$$

Указание: Доказать соотношение $M(r) = \sum_{n=0}^{\infty} P\{r > n\} = \sum_{n=0}^{\infty} P\{X_1 + \ldots + X_n \leqslant a\}$ и затем воспользоваться результатом задачи 2.

4. Пусть $X_1, X_2, \ldots$ — последовательность независимых одинаково распределенных целочисленных случайных величин, а $S_k = X_1 + X_2 + \ldots + X_k$, $k = 1, 2, \ldots$, — частичные суммы. Как обычно, $S_0 = 0$. Индекс $n \geqslant 1$ назовем лестничным, если $S_n > S_j$ при $j = 0, 1, \ldots, n-1$. Событие, состоящее в том, что n является лестничным индексом, обозначим через $\mathfrak{E}$. Определим Y_N как момент (т. е. индекс) последнего наступления события $\mathfrak{E}$, где N — номер текущего испытания. Пусть W — момент первого наступления $\mathfrak{E}$. Предположим, что $\mathfrak{E}$ происходит на n-м испытании. Показать, что число испытаний до следующего наступления $\mathfrak{E}$ не зависит от n и распределено, как W.

5 (продолжение). Доказать соотношение

$$\sum_{n=0}^{\infty} t^n M\left(x^{Y_n}\right) = \frac{1 - F(t)}{1 - t} \cdot \frac{1}{1 - F(xt)},$$

где $F(t) = \sum_{n=1}^{\infty} P\{W = n\} t^n$.

Указание: Использовать соотношение

$$P\{Y_n = k\} = P\{Y_k = k\} P\{Y_{n-k} = 0\}.$$

6 (продолжение). Пусть $U(t) = 1/(1 - F(t))$. Показать, что $U(t)$ можно представить в виде

$$U(t) = \exp\left[\sum_{k=1}^{\infty} \frac{t^k}{k} \{M(Y_k) - M(Y_{k-1})\}\right].$$

Указание: С помощью результата предыдущего упражнения вывести дифференциальное уравнение

$$\frac{U'(t)}{U(t)} = \sum_{n=1}^{\infty} M(Y_n - Y_{n-1}) t^{n-1}$$

и решить его.

7. Пусть $X_1, X_2, \ldots$ — независимые одинаково распределенные целочисленные с. в. Положим

$$S_0 = 0 \quad \text{и} \quad S_k = X_1 + X_2 + \ldots + X_k \quad \text{при} \quad k = 1, 2, \ldots.$$

Пусть f_n — вероятность того, что процесс $\{S_k\}$ впервые вернется в нулевое состояние на n-м шаге. Пусть γ_n — вероятность того, что на n-м шаге процесс будет занимать новое состояние, т. е. $\gamma_n = P\{S_n \neq S_0,\ S_n \neq S_1,\ S_n \neq S_2, \ldots, S_n \neq S_{n-1}\}$. Доказать, что

$$\gamma_n = P\{S_1 \neq 0,\ S_2 \neq 0,\ \ldots,\ S_n \neq 0\}.$$

Выразить γ_n через $\{f_i\}_{i=1}^{\infty}$ для случая, когда $M(|X_i|) < \infty$ и $M(X_i) = 0$.

8 (продолжение). Пусть R_n — число различных состояний, которые процесс $\{S_k\}_{k=0}^{\infty}$ занимал за первые n шагов. Выразить среднее значение д. с. в. R_n через $\{\gamma_k\}$.

Ответ: $\mathrm{M}(R_n) = \sum_{i=1}^{n} \gamma_i$.

9 (продолжение). Показать, что

$$\lim_{n\to\infty} \frac{\mathrm{M}(R_n)}{n} = 1 - f_{00}^*,$$

где

$$f_{00}^* = \sum_{k=1}^{\infty} f_{00}^k.$$

10 (продолжение). Пусть $L_n = j$ тогда и только тогда, когда $S_j > S_i$ при $0 \leqslant i < j$ и $S_j \geqslant S_i$ при $j < i \leqslant n$, т. е. L_n есть первый индекс, на котором достигается $\max\limits_{0\leqslant i\leqslant n} S_i$. Доказать соотношение

$$\mathrm{P}\{L_n = j\} = \mathrm{P}\{L_j = j\}\,\mathrm{P}\{L_{n-j} = 0\}.$$

11. Определим процесс $\{X_t,\ t = 0, 1, 2, \ldots\}$, где X_t — действительные числа, следующим образом: $X_0 = 0$,

$$X_t = \begin{cases} X_{t-1} + \left(\frac{1}{2}\right)^t \text{ с вероятностью } \frac{1}{2}, \\ X_{t-1} - \left(\frac{1}{2}\right)^t \text{ с вероятностью } \frac{1}{2}. \end{cases}$$

Показать, что в пределе при $t\to\infty$ распределение с. в. X_t стремится к равномерному распределению на интервале $(-1, 1)$.

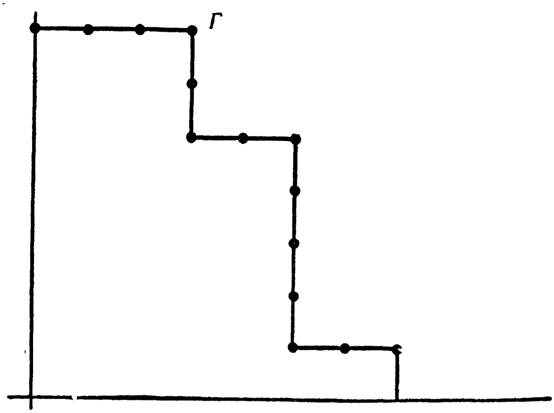
Указание: Распределение с. в. X_t стремится к распределению с. в. $Y = \sum_{k=1}^{\infty} Y_k \left(\frac{1}{2}\right)^k$, где Y_k — одинаково и независимо распределенные с. в. со значениями ± 1, принимаемыми равновероятно. Пусть $f(s) = \prod_{k=1}^{\infty} \cos s/2^k$; $f(s)$ удовлетворяет функциональному уравнению

$$f(2s) = (\cos s)\, f(s).$$

Показать, что единственным непрерывным решением этого уравнения, удовлетворяющим условию $f(0) = 1$, является функция $(\sin s)/s$, т. е. характеристическая функция равномерного распределения на интервале $(-1, 1)$.

12. Рассмотрим двумерное случайное блуждание по целочисленной решетке в положительном квадранте. Если на некотором шаге процесс находится в состоянии (m, n), то на следующем шаге процесс с одинаковой вероятностью $1/2$ перейдет либо в состояние $(m+1, n)$, либо в $(m, n+1)$. Пусть Γ — ломаная, соединяющая соседние точки решетки (и простирающаяся от оси Y до оси X) в положительном квадранте. Показать, что $M(Y_1) = M(Y_2)$ где Y_1 и Y_2 обо-

значают число шагов вправо и число шагов вверх соответственно перед попаданием на границу Г. Предполагается, что процесс выходит из начального состояния (0, 0). Пример ломаной Г приведен на рисунке.



Указание: (а) Сначала рассмотреть случай, когда ломаная Г состоит из двух сегментов AB и BC. AB — горизонтальный отрезок, соединяющий точки $(0, 1)$ и $(N, 1)$, а BC — вертикальный отрезок, соединяющий точки $(N, 1)$ и $(N, 0)$. Показать, что для этого случая

$$\mathrm{M}(Y_2) = \frac{1}{2} + \sum_{k=1}^{N-1} \left(\frac{1}{2}\right)^{k+1}, \qquad \mathrm{M}(Y_1) = \sum_{k=1}^{N-1} k \left(\frac{1}{2}\right)^{k+1} + \left(\frac{1}{2}\right)^N N$$

и эти величины равны между собой.

(б) Всякую область, ограниченную осями X и Y и ломаной Г, можно разбить на блоки вида, описанного в (а). Использовать соответствующее правило сложения средних и результат пункта (а) настоящего указания.

Задачи 13—18 основаны на следующей модели.

Рассмотрим случайное блуждание $X_n = (A_n, B_n)$ на плоскости, возможными состояниями которого являются все точки с целочисленными координатами в двумерном пространстве. Предположим, что вероятность перехода в любое из четырех соседних состояний равна $^1/_4$. Пусть T — время первого попадания случайного блуждания, начинающегося из начала координат, на биссектрису положительного квадранта, а X_T — точка на биссектрисе, в которую попадает случайное блуждание. Пусть

$$Q_{0j} = \mathrm{P}\{X_T = (j,\ j) \mid X_0 = (0;\ 0)\}.$$

Этим соотношением определяется переходная вероятность одномерного случайного блуждания, порождаемого исходным двумерным случайным блужданием, наблюдаемым лишь в моменты его попадания на биссектрису. Определим характеристическую функцию указанного одномерного случайного блуждания:

$$\psi(\theta) = \sum_{m=-\infty}^{\infty} Q_{0m} e^{im\theta}, \quad -\infty < \theta < \infty.$$

13. Положим $U_0 = V_0 = 0$ и $U_n = A_n + B_n$, $V_n = A_n - B_n$, $n = 1, 2, \ldots$. Показать, что последовательность с. в. $\{U_n\}$ не зависит от последовательности $\{V_n\}$. (Переход к переменным $\{U_n, V_n\}$ соответствует замене исходной системы координат на систему, в которой одной из осей является биссектриса положительного квадранта.)

Указание: (а) Показать сначала, что

$$P\{U_m - U_{m-1} = \pm r \mid U_0 = V_0 = 0\} = \begin{cases} \frac{1}{2}, & |r| = 1, \\ 0, & |r| \neq 1; \end{cases} \quad (m = 1, 2, \ldots)$$

и

$$P\{V_n - V_{n-1} = \pm s \mid U_0 = V_0 = 0\} = \begin{cases} \frac{1}{2}, & |s| = 1, \\ 0, & |s| \neq 0; \end{cases} \quad (n = 1, 2, \ldots).$$

(б) Далее показать, что

$$P\{U_m - U_{m-1} = \pm 1, \ V_n - V_{n-1} = \pm 1 \mid U_0 = V_0 = 0\} = \frac{1}{4}$$

и

$$P\{U_m - U_{m-1} = r, \ V_n - V_{n-1} = s \mid U_0 = V_0 = 0\} = 0,$$

если $|r| \neq 1$ либо $|s| \neq 1$ $(n, m = 1, 2, \ldots)$.

Воспользоваться формулами пунктов (а) и (б) и показать, что последовательность $\{U_m - U_{m-1}\}$ не зависит от последовательности $\{V_n - V_{n-1}\}$; вывести отсюда, что $\{U_m\}$ не зависит от $\{V_n\}$.

14. Доказать, что с. в. T не зависит от последовательности $\{U_n\}$.

15. Убедившись предварительно, что из $T = k$ следует, что $V_k = 0$, установить формулу

$$\psi(\theta) = \sum_{k=1}^{\infty} P\{T = k \mid X_0 = (0, 0)\} \sum_{l=-\infty}^{\infty} e^{il\theta} P\{U_k = 2l \mid X_0 = (0, 0)\}.$$

16. На основе того факта, что $\{U_n\}$ описывает одномерное случайное блуждание, показать, что

$$\psi(\theta) = \sum_{k=1}^{\infty} P\{T = k \mid X_0 = (0, 0)\} (\cos \theta/2)^k.$$

17. Доказать, что производящая функция с. в. T имеет вид

$$M(s_T) = 1 - \sqrt{1 - s^2}.$$

18. Доказать формулу

$$\psi(\theta) = 1 - |\sin \theta/2|, \quad -\infty < \theta < \infty.$$

ЗАМЕЧАНИЯ

В значительной степени источником материала этой главы послужила элегантная книга Спицера [1]. Превосходными руководствами по теории вероятностей, содержащими изложение предельных теорем для сумм независимых случайных величин на значительно более высоком уровне, являются монографии Гнеденко и

Колмогорова [2], Лоэва [3] и Реньи [4]. Современное полное изложение предельных теорем теории вероятностей в различных аспектах читатель найдет во втором томе книги Феллера [5].

ЛИТЕРАТУРА

1. Спицер Ф., Принципы случайного блуждания, «Мир», М., 1968.
2. Гнеденко Б. В., Колмогоров А. Н., Предельные распределения для сумм независимых случайных величин, Гостехиздат, М., 1949.
3. Лоэв М., Теория вероятностей, ИЛ, М., 1962.
4. Rényi A., Wahrscheinlichkeitsrechnung, mit einem Anhang über Informationstheorie. Deutscher Verlag der Wissenschaften, Berlin, 1962.
5. Феллер В., Введение в теорию вероятностей и ее приложения, т. 2, «Мир», 1967.

Глава 7

КЛАССИЧЕСКИЕ ПРИМЕРЫ ЦЕПЕЙ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

§ 1. ОБЩИЕ ПРОЦЕССЫ ЧИСТОГО РОЖДЕНИЯ (РАЗМНОЖЕНИЯ) И ПУАССОНОВСКИЕ ПРОЦЕССЫ

В предыдущих главах были введены основные понятия и рассмотрены методы анализа цепей Маркова с дискретным временем. В этой главе дается краткое обсуждение некоторых важных примеров марковских процессов с дискретным множеством состояний и непрерывным временем.

Точнее, здесь мы будем иметь дело с семейством случайных величин $\{X(t),\ 0 \leqslant t \leqslant \infty\}$, принимающих неотрицательные целочисленные значения. Мы ограничимся случаем, когда $\{X(t)\}$ — марковский процесс со стационарными переходными вероятностями. Таким образом, переходная вероятностная функция при $t > 0$

$$P_{ij}(t) = \mathrm{P}\{X(t+u) = j \mid X(u) = i\}, \quad i, j = 0, 1, 2, \ldots, \tag{1.1}$$

не зависит от $u \geqslant 0$.

Обычно при исследовании частных вероятностных моделей физических явлений более естественно описать так называемые инфинитезимальные вероятности, связанные с процессом, а затем вывести из них точное выражение для переходной функции.

В рассматриваемом случае мы будем постулировать вид $P_{ij}(h)$ для малых h и, используя марковское свойство, выведем систему дифференциальных уравнений, которой удовлетворяют $P_{ij}(t)$ при всех $t > 0$. $P_{ij}(t)$ являются решением этих уравнений при соответствующих начальных условиях. Напомним, что пуассоновский процесс, введенный в § 2 гл. 1, рассматривался именно таким образом.

Перед тем как перейти к общему процессу чистого рождения, напомним кратко аксиомы, характеризующие пуассоновский процесс.

А. Постулаты пуассоновского процесса

Пуассоновский процесс был рассмотрен в § 2 гл. 1, где было показано, что его можно определить с помощью нескольких простых постулатов. Для того чтобы определить более общие процессы подобного рода, укажем на некоторые свойства, которыми обладает пуассоновский процесс. Пуассоновский процесс — это марковский процесс, принимающий неотрицательные целочисленные значения и обладающий следующими свойствами:

(1) $\mathrm{P}\{X(t+h)-X(t)=1 \mid X(t)=x\}=\lambda h+o(h)$ при $h \downarrow 0$
$(x=0,\ 1,\ 2,\ \ldots)$

Свойство (1) можно записать еще так:

$$\lim_{h\downarrow 0}\frac{1}{h}\mathrm{P}\{X(t+h)-X(t)=1 \mid X(t)=x\}=\lambda.$$

Символ $o(h)$ обозначает такую величину, которая, будучи деленной на h, стремится к 0 при $h\to 0$. Заметим, что правая часть не зависит от x.

(2) $\mathrm{P}\{X(t+h)-X(t)=0 \mid X(t)=x\}=1-\lambda h+o(h)$ при $h \downarrow 0$.

(3) $X(0)=0$.

Эти свойства легко проверить непосредственным вычислением, поскольку имеются точные формулы для всех рассматриваемых вероятностей.

Б. Примеры пуассоновских процессов

(а) Можно проиллюстрировать пуассоновский процесс на примере процесса ловли рыбы. Пусть случайная величина $X(t)$ обозначает число пойманных рыб за временной интервал $[0, t]$. Предположим, что общее число рыбы, имеющейся в наличии, весьма велико, так что «завзятые» рыбаки имеют не больше шансов поймать рыбу, чем все остальные, и что интенсивность клева не зависит от времени. При этих «идеальных» условиях процесс $\{X(t),\ t\geqslant 0\}$ можно считать пуассоновским. Этот пример поясняет марковское свойство (вероятность поймать рыбу не зависит от числа уже пойманных) и свойство независимости от времени ожидания[1]), которое является наиболее характерным свойством пуассоновского процесса; оно означает, что рыболов, только что прибывший на рыбалку, имеет такой же шанс поймать рыбу за определенный отрезок времени, как и тот, который уже просидел четыре часа безуспешно.

(б) Менее искусственный пример связан с задачами, возникающими в теории счетчиков. Если $X(t)$ — число радиоактивных частиц, зарегистрированных счетчиком Гейгера за временной интервал $[0, t]$, то этот процесс является пуассоновским при t, много меньших полупериода распада вещества. Отсюда получается, что вероятность распада за единицу времени можно считать не меняющейся с течением времени.

---

[1]) В оригинале «no premium for waiting property», т. е. дословно «свойство отсутствия поощрения за ожидание». — *Прим. перев.*

(в) Пуассоновские процессы естественным образом возникают во многих моделях образования очередей (массового обслуживания). В этих примерах наибольшее внимание уделяется моментам времени, когда $X(t)$ (длина очереди в момент t) изменяется, а не самим значениям $X(t)$ [1]. Процесс ловли рыбы (см. пример (а)) является, конечно, частным случаем подобных моделей.

В. Процесс чистого рождения

Естественное обобщение пуассоновского процесса получим, допустив зависимость вероятности осуществления события в данный момент времени от числа событий, которые уже произошли; примером такого явления может служить воспроизведение живых организмов (отсюда и название процесса), когда при соответствующих условиях — избытке пищи, отсутствии смертности, отсутствии миграции и т. д. — вероятность рождения в данный момент прямо пропорциональна размеру популяции. Этот пример известен как процесс Юла.

Рассмотрим последовательность положительных чисел $\{\lambda_k\}$. Определим процесс чистого рождения как марковский процесс, удовлетворяющий постулатам:

(1) $\mathrm{P}\{X(t+h)-X(t)=1 \mid X(t)=k\}=\lambda_k h+o_{1,k}(h), \quad h \downarrow 0,$

(2) $\mathrm{P}\{X(t+h)-X(t)=0 \mid X(t)=k\}=1-\lambda_k h+o_{2,k}(h),$

(3) $X(0)=0,$

(4) $\mathrm{P}\{X(t+h)-X(t)<0 \mid X(t)=k\}=0, \quad k \geqslant 0.$

Постулат (3) введен только для удобства. В нем под $X(t)$ подразумевается не размер популяции, а число рождений в интервале $[0, t]$.

Заметим, что левые части (1) и (2) равны $P_{k,k+1}(h)$ и $P_{k,k}(h)$ соответственно (вследствие стационарности), так что $o_{1,k}(h)$ и $o_{2,k}(h)$ не зависят от t.

Определим $P_n(t)=\mathrm{P}\{X(t)=n\}$.

Точно так же, как и для пуассоновского процесса [2], можно вывести систему дифференциальных уравнений относительно $P_n(t)$ при $t \geqslant 0$ вида

$$\begin{aligned} P_0'(t) &= -\lambda_0 P_0(t), \\ P_n'(t) &= -\lambda_n P_n(t)+\lambda_{n-1} P_{n-1}(t), \quad n \geqslant 1, \end{aligned} \tag{1.2}$$

с начальными условиями

$$P_0(0)=1, \quad P_n(0)=0, \quad n>0.$$

[1] Для пуассоновского процесса такие моменты времени образуют *пуассоновский поток*. — *Прим. перев.*

[2] См. гл. 1, § 2. — *Прим. перев.*

В самом деле, если $h > 0$, $n \geqslant 1$, то, используя формулу полной вероятности, марковское свойство и постулат (4), получим

$$P_n(t+h) = \sum_{k=0}^{\infty} P_k(t) \mathrm{P}\{X(t+h) = n \mid X(t) = k\} =$$
$$= \sum_{k=0}^{\infty} P_k(t) \mathrm{P}\{X(t+h) - X(t) = n-k \mid X(t) = k\} =$$
$$= \sum_{k=0}^{n} P_k(t) \mathrm{P}\{X(t+h) - X(t) = n-k \mid X(t) = k\}.$$

Далее, при $k = 0, 1, \ldots, n-2$ имеем

$$\mathrm{P}\{X(t+h) - X(t) = n-k \mid X(t) = k\} \leqslant$$
$$\leqslant \mathrm{P}\{X(t+h) - X(t) \geqslant 2 \mid X(t) = k\} = o_{1,k}(h) + o_{2,k}(h),$$

или

$$\mathrm{P}\{X(t+h) - X(t) = n-k \mid X(t) = k\} = o_{3,n,k}(h), \quad k = 0, \ldots, n-2.$$

Таким образом,

$$P_n(t+h) = P_n(t)[1 - \lambda_n h + o_{2,n}(h)] +$$
$$+ P_{n-1}(t)[\lambda_{n-1}h + o_{1,n-1}(h)] + \sum_{k=0}^{n-2} P_k(t)\, o_{3,n,k}(h),$$

или

$$P_n(t+h) - P_n(t) = P_n(t)[-\lambda_n h + o_{2,n}(h)] +$$
$$+ P_{n-1}(t)[\lambda_{n-1}h + o_{1,n-1}(h)] + o_n(h), \qquad (1.3)$$

где, очевидно, $\lim_{h \downarrow 0} o_n(h)/h = 0$ равномерно по $t \geqslant 0$, поскольку $o_n(h)$ ограничена конечной суммой $\sum_{k=0}^{n-2} o_{3,n,k}(h)$, не зависящей от t.

Деля на h и переходя к пределу при $h \downarrow 0$, получим соотношение (1.2), в котором, если быть точным, в левой части нужно писать правостороннюю производную. Однако, используя несколько более тонкие рассуждения, можно вывести такое же соотношение для левосторонней производной. В самом деле, из (1.3) сразу следует, что $P_n(t)$ — непрерывная функция t. Заменяя в соотношении (1.3) t на $t-h$, деля на h и переходя к пределу при $h \downarrow 0$, найдем, что каждая функция $P_n(t)$ имеет левостороннюю производную, которая также удовлетворяет уравнению (1.2).

Первое уравнение в (1.2) можно решить сразу и получить

$$P_0(t) = \exp(-\lambda_0 t) > 0.$$

Обозначим через T_k время между k-м и $(k+1)$-м рождениями. Тогда

$$P_n(t) = \mathrm{P}\left\{\sum_{i=0}^{n-1} T_i \leqslant t < \sum_{i=0}^{n} T_i\right\}$$

и

$$S_k = \sum_{i=0}^{k-1} T_i - \text{время } k\text{-го рождения.}$$

Мы уже видели, что $P_0(t) = \exp\{-\lambda_0 t\}$. Следовательно,

$$\mathrm{P}\{T_0 \leqslant z\} = 1 - \mathrm{P}\{X(z) = 0\} = 1 - \exp\{-\lambda_0 z\},$$

т. е. T_0 имеет экспоненциальное распределение с параметром λ_0. Из постулатов (1)—(4) можно вывести, что величины T_k, $k > 0$, также имеют экспоненциальные распределения с параметрами λ_k и что они взаимно независимы (см. § 3 гл. 8, где дано формальное доказательство этого факта). Следовательно, характеристическая функция величины S_n равна

$$\varphi_n(w) = \mathrm{M}\{\exp(iwS_n)\} = \prod_{k=0}^{n-1} \mathrm{M}\{\exp(iwT_k)\} = \prod_{k=0}^{n-1}\left(\frac{\lambda_k}{\lambda_k - iw}\right). \tag{1.4}$$

В случае пуассоновского процесса, когда $\lambda_k = \lambda$ при всех k, из (1.4) видно, что величина S_n распределена в соответствии с гамма-распределением порядка n со средним n/λ.

При конкретных значениях $\lambda_k \geqslant 0$ можно последовательно проинтегрировать уравнения (1.2):

$$P_k(t) = \lambda_{k-1} \exp(-\lambda_k t) \int_0^t \exp(\lambda_k x) P_{k-1}(x)\, dx, \quad k = 1, 2, \ldots,$$

откуда ясно, что все $P_k(t) \geqslant 0$.

Но еще остается возможность того, что

$$\sum_{n=0}^{\infty} P_n(t) < 1.$$

Чтобы гарантировать регулярность процесса, т. е. дать критерий того, что будет $\sum_{n=0}^{\infty} P_n(t) = 1$ для всех t, мы должны ограничить λ_k условием

$$\sum_{n=0}^{\infty} P_n(t) = 1 \leftrightarrow \sum_{n=0}^{\infty} \frac{1}{\lambda_n} = \infty. \tag{1.5}$$

Доказательство этого дано в книге Феллера[1]) и здесь не приводится. Интуитивные соображения в пользу этого результата следующие. Время T_k между последовательными рождениями, как показано ниже, распределено экспоненциально с параметром λ_k. Следовательно, величина $\sum\limits_n \frac{1}{\lambda_n}$ равна среднему времени до того момента, когда популяция станет бесконечной. Но $1 - \sum\limits_{n=0}^{\infty} P_n(t)$ есть вероятность того, что $X(t) = \infty$.

Если $\sum \lambda_n^{-1} < \infty$, то среднее время, за которое популяция становится бесконечной, конечно. Поэтому правдоподобно, что при этом для всех $t > 0$ вероятность того, что $X(t) = \infty$, положительна.

Г. Процесс Юла

Процесс Юла является примером процесса чистого рождения, который возникает в физике и биологии. Предположим, что каждый член популяции в интервале времени длины h с вероятностью $\beta h + o(h)$ порождает нового члена $(\beta > 0)$. Кроме того предположим, что в момент 0 в популяции имеется N членов. Предполагая независимость и отсутствие взаимодействия между членами популяции, получим в силу биномиального закона

$$\begin{aligned} \mathrm{P}\{X(t+h) - X(t) = 1 \mid X(t) = n\} &= \\ &= \binom{n+N}{1} [\beta h + o(h)] [1 - \beta h + o(h)]^{n+N-1} = \\ &= (n+N)\beta h + o_n(h), \end{aligned}$$

т. е. в этом примере $\lambda_n = (n+N)\beta$. Система уравнений (1.2) в случае, когда $N = 1$, принимает вид

$$P_n'(t) = -\beta[(n+1)P_n(t) - nP_{n-1}(t)], \quad n = 0, 1, \ldots$$

с начальными условиями

$$P_0(0) = 1, \quad P_n(0) = 0, \quad n = 1, 2, \ldots.$$

Ее решение равно

$$P_n(t) = e^{-\beta t}(1 - e^{-\beta t})^n, \quad n \geqslant 0,$$

что можно проверить непосредственно.

---

[1]) Феллер В., Введение в теорию вероятностей и ее приложения, 2-е изд., т. 1, «Мир», М., 1964.

§ 2. ДОПОЛНИТЕЛЬНЫЕ СВЕДЕНИЯ О ПУАССОНОВСКИХ ПРОЦЕССАХ

В предыдущем параграфе мы определили пуассоновский процесс системой допущений (постулатов), которые довольно хорошо описывают многие реальные ситуации. Этот процесс часто называют абсолютно случайным процессом[1]), так как он «распределяет» точки «случайным образом» на прямой $[0, \infty)$ совершенно аналогично тому, как распределяются точки при равномерном распределении на конечном интервале. В частности, вероятность наступления события в некотором интервале является функцией лишь его длины, а количества событий, происходящих в двух непересекающихся интервалах, являются независимыми случайными величинами.

Рассмотрим теперь пуассоновский процесс более подробно.

А. Характеристическая функция и длительности пребывания [2])

Характеристическую функцию пуассоновского процесса $X(t)$ можно записать в виде

$$\varphi_t(w) = \mathrm{M}\{e^{iwX(t)}\} = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} e^{iwn} = \exp[\lambda t(e^{iw} - 1)].$$

Таким образом,

$$\mathrm{M}(X(t)) = \lambda t, \quad \sigma^2(X(t)) = \lambda t.$$

При обсуждении процесса чистого рождения мы показали, что

$$\mathrm{P}\{T_0 \leqslant z\} = 1 - \exp(-\lambda_0 z),$$

и упомянули, что T_k имеет экспоненциальное распределение с параметром λ_k и все T_k независимы. Для пуассоновского процесса $\lambda_k = \lambda$ при всех k, так что справедлива

Теорема 2.1. *Длительности пребывания T_k — независимые и одинаково распределенные случайные величины, имеющие экспоненциальное распределение с параметром λ.*

Строгое доказательство этой теоремы будет следовать из более общих рассмотрений § 4 гл. 8.

Из определения процесса следует большее, нежели утверждение этой теоремы. Так, время до ближайшего изменения $X(t)$ имеет точно такое же распределение, если его отсчитывать от любого момента, а не только от момента предыдущего скачка. Иначе говоря,

$$\mathrm{P}\{X(t_0 + \tau) - X(t_0) > 0\} = 1 - e^{-\lambda\tau},$$

---

[1]) Принят также термин «чисто случайный процесс». — *Прим. ред.*

[2]) В оригинале «waiting time» (длительность ожидания). — *Прим. перев.*

что было получено в § 1. Это свойство можно получить и непосредственно. Пусть $F(x) = \mathrm{P}\{X(t_0 + x) - X(t_0) > 0\}$, где t_0 — некоторый момент времени, зависящий, возможно, от истории процесса до этого момента, значение которого не изменяет данной вероятности[1]. Тогда

$$\begin{aligned} F(x+y) &= \mathrm{P}\{X(t_0+x+y) - X(t_0) > 0\} = \\ &= \mathrm{P}\{X(t_0+y) - X(t_0) > 0\} + \mathrm{P}\{X(t_0+y) - X(t_0) = 0\} \times \\ &\qquad \times \mathrm{P}\{X(t_0+x+y) - X(t_0+y) > 0 \mid X(t_0+y) - X(t_0) = 0\} \end{aligned}$$

Из определения $F(x)$, независимости приращений пуассоновского процесса и того факта (который является исходным предположением при определении пуассоновского процесса), что

$$\mathrm{P}\{X(t_0+x) - X(t_0) > 0\}$$

не зависит от t_0, получаем функциональное уравнение

$$F(x+y) = F(y) + [1 - F(y)]\, F(x).$$

То, что это уравнение определяет экспоненциальное распределение, доказывается в следующей теореме.

Теорема 2.2. *Если $F(x)$ — такое распределение, что $F(0) = 0$ и $F(x) < 1$ при некотором $x > 0$, то $F(x)$ является экспоненциальным тогда и только тогда, когда*

$$F(x+y) - F(y) = F(x)\,[1 - F(y)] \tag{*}$$

ри всех $x, y \geqslant 0$.

Доказательство. То, что экспоненциальное распределение удовлетворяет условиям теоремы, проверяется непосредственно. Чтобы доказать обратное утверждение, положим $G(x) = 1 - F(x)$. Тогда условие (*) примет вид

$$G(x+y) = G(x)\, G(y). \tag{2.1}$$

Очевидно, $G(0) = 1$, $G(x)$ — невозрастающая функция и $G(x) > 0$ при некотором $x > 0$. Предположим, что $G(x_0) = 0$ при некотором $x_0 > 0$. Из (2.1) следует, что $G(x_0) = \left[G\left(\frac{x_0}{n}\right)\right]^n$ для всякого целого $n > 0$. Следовательно, $G\left(\frac{x_0}{n}\right) = 0$. Но тогда из (2.1) следует, что $G(x) = 0$ при $x > \frac{x_0}{n}$. Так как n произвольно, то $G(x) = 0$ при всех $x > 0$, что противоречит сделанному предположению. Таким

---

[1]) Смысл этой кажущейся неясной фразы будет пояснен при обсуждении понятия «марковское время». См. § 4, гл. 8.

образом, $G(x)>0$ при всех $x>0$. Далее, для любых целых $m, n>0$ из (2.1) легко вывести, что $G\left(\frac{m}{n}\right)=[G(1)]^{\frac{m}{n}}$. Так как $G(x)$ и $[G(1)]^x$ — невозрастающие функции, совпадающие при всех рациональных x, и $[G(1)]^x$ непрерывна, то отсюда следует, что

$$G(x)=[G(1)]^x=\exp(x\ln G(1))$$

при всех $x>0$. Но $F(x)$ — функция распределения, поэтому

$$\lim_{x\to\infty} G(x)=1-\lim_{x\to\infty} F(x)=0,$$

откуда $G(1)<1$. Следовательно, $G(x)=e^{-\lambda x}$, где

$$\lambda=-\ln G(1)>0. \blacksquare$$

Другое доказательство в предположении, что G дифференцируема, следующее. Заметим, что из (2.1) получается

$$G'(x+y)=\frac{\partial}{\partial x}G(x+y)=G'(x)\,G(y),$$

$$G(x)\,G'(y)=\frac{\partial}{\partial y}G(x+y)=G'(x+y)$$

и, следовательно,

$$G'(x)=aG(x), \tag{2.2}$$

где $a=G'(y_0)[G(y_0)]^{-1}$ при некотором y_0, таком, что $G(y_0)\neq 0$. Решением уравнения (2.2) является $G(x)=Ae^{ax}$, и $A=1$, поскольку $G(0)=1-F(0)=1$. Параметр a отрицателен, так как $G(x)<1$ при некотором $x>0$.

Б. Равномерное распределение

Класс распределений, связанных с пуассоновским процессом, не ограничивается пуассоновским и экспоненциальным распределениями. Мы покажем сейчас, что возникают также равномерное и биномиальное распределения.

Рассмотрим моменты времени $\{S_i\}$, когда происходят изменения $X(t)$, т. е.

$$S_i=\sum_{k=0}^{i-1} T_k.$$

Имеем следующий результат.

Теорема 2.3. *Для любых чисел* s_i, $0\leqslant s_1\leqslant s_2\leqslant\ldots\leqslant s_n\leqslant t$,

$$\mathrm{P}\{S_i\leqslant s_i,\ i=1,\ \ldots,\ n\,|\,X(t)=n\}=\frac{n!}{t^n}\int_0^{s_1}\ldots\int_{x_{n-2}}^{s_{n-1}}\int_{x_{n-1}}^{s_n} dx_n\ldots dx_1,$$

что является распределением порядковых статистик из выборки объема n, взятой из равномерного распределения на $[0, t]$[1]).

Доказательство. Доказательство является простым следствием теоремы 2.1. Действительно,

$$\mathrm{P}\{S_1 \leqslant s_1,\ S_2 \leqslant s_2,\ \ldots,\ S_n \leqslant s_n,\ X(t)=n\} =$$
$$= \mathrm{P}\{T_0 \leqslant s_1, T_0+T_1 \leqslant s_2, \ldots, T_0+\ldots+T_{n-1} \leqslant s_n, T_0+\ldots+T_n > t\} =$$
$$= \int\limits_0^{s_1} \int\limits_0^{s_2-t_1} \int\limits_0^{s_3-(t_1+t_2)} \ldots \int\limits_0^{s_n-(t_1+\ldots+t_{n-1})} \int\limits_{t-(t_1+\ldots+t_n)}^{\infty} \lambda^{n+1} e^{-\lambda(t_1+\ldots+t_{n+1})}\, dt_{n+1}\ldots dt_1 =$$
$$= \lambda^{n+1} \int\limits_0^{s_1} \int\limits_0^{s_2-t_1} \int\limits_0^{s_3-(t_1+t_2)} \ldots \int\limits_0^{s_n-(t_1+\ldots+t_{n-1})} e^{-\lambda(t_1+\ldots+t_n)} \times$$
$$\times \left[-\frac{1}{\lambda}\exp(-\lambda t_{n+1})\right]_{t-(t_1+\ldots+t_n)}^{\infty} dt_n \ldots dt_1 =$$
$$= \lambda^n e^{-\lambda t} \int\limits_0^{s_1} \int\limits_0^{s_2-t_1} \int\limits_0^{s_3-(t_1+t_2)} \ldots \int\limits_0^{s_n-(t_1+\ldots+t_{n-1})} dt_n \ldots dt_1.$$

Если ввести новые переменные

$$u_n = t_1 + \ldots + t_n,$$
$$u_{n-1} = t_1 + \ldots + t_{n-1},$$
$$\vdots$$
$$u_1 = t_1,$$

то последнее выражение примет вид

$$\lambda^n e^{-\lambda t} \int\limits_0^{s_1} \int\limits_{u_1}^{s_2} \int\limits_{u_2}^{s_3} \ldots \int\limits_{u_{n-1}}^{s_n} du_n \ldots du_1.$$

Но

$$\mathrm{P}\{X(t)=n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

[1]) Это означает следующее. Возьмем n независимых наблюдений (реализаций) случайной величины, которая равномерно распределена на отрезке $[0, t]$. Пусть $Y_1 \leqslant Y_2 \leqslant \ldots \leqslant Y_n$ — наблюдения, упорядоченные по величине. Тогда совместное распределение величин $Y_1, \ldots, Y_n$ в точности совпадает с выражением, приведенным в теореме. Доказательство этого факта весьма простое, но более подробное обсуждение дано в гл. 9.

Следовательно,

$$P\{S_1 \leqslant s_1,\ S_2 \leqslant s_2,\ \ldots,\ S_n \leqslant s_n \mid X(t) = n\} =$$

$$= \frac{P\{S_1 \leqslant s_1, \ldots, S_n \leqslant s_n,\ X(t) = n\}}{P\{X(t) = n\}} = \frac{n!}{t^n} \int\limits_0^{s_1} \int\limits_{u_1}^{s_2} \ldots \int\limits_{u_{n-1}}^{s_n} du_n \ldots du_1.$$

B. Биномиальное распределение

Из свойств пуассоновского процесса следует, что при $u < t$ и $k < n$

$$P\{X(u) = k \mid X(t) = n\} =$$

$$= P\{X(u) = k,\ X(t) - X(u) = n - k\}\,[P\{X(t) = n\}]^{-1} = \tag{2.3}$$

$$= \frac{e^{-\lambda u} \frac{u^k}{k!} e^{-\lambda(t-u)} \frac{(t-u)^{n-k}}{(n-k)!}}{e^{-\lambda t} \frac{t^n}{n!}} = \binom{n}{k} \frac{u^k (t-u)^{n-k}}{t^n}.$$

Второй пример, в котором играет роль биномиальное распределение, можно получить, рассматривая два независимых пуассоновских процесса $X_1(t)$ и $X_2(t)$ с параметрами λ_1 и λ_2. Именно:

$$P\{X_1(t) = k \mid X_1(t) + X_2(t) = n\} = \frac{P\{X_1(t) = k,\ X_2(t) = n - k\}}{P\{X_1(t) + X_2(t) = n\}} =$$

$$= \frac{\exp(-\lambda_1 t) \frac{(\lambda_1 t)^k}{k!} \exp(-\lambda_2 t) \frac{(\lambda_2 t)^{n-k}}{(n-k)!}}{\exp[-(\lambda_1 + \lambda_2) t] \frac{(\lambda_1 + \lambda_2)^n t^n}{n!}} = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k}.$$

§ 3. МОДЕЛЬ СЧЕТЧИКА

Интересным применением пуассоновского процесса является следующая задача. Электрические импульсы случайной амплитуды X_i поступают в случайные моменты времени t_i (образуя пуассоновский поток) на детектор, реакция которого на каждый отдельный импульс в момент t выражается функцией

$$X_i \exp[-\alpha(t - t_i)]_+ = \begin{cases} 0 & \text{при} \quad t < t_i, \\ X_i \exp[-\alpha(t - t_i)] & \text{при} \quad t \geqslant t_i. \end{cases}$$

То есть в момент подачи импульса значение выходного сигнала равно амплитуде подаваемого импульса, а затем оно убывает по экспоненциальному закону. Детектор является линейным (т. е.

аддитивным). Так, если за интервал $(0, t)$ поступит N_t импульсов, то значение выхода в момент t равно

$$\eta(t) = \sum_{i=1}^{N_t} X_i \exp[-\alpha(t - t_i)]_+.$$

Типичная реализация этого процесса имеет вид, показанный на рис. 1. Мы хотели бы знать функцию распределения $\eta(t)$ для каждого t или, что эквивалентно, ее характеристическую функцию $\varphi_t(w)$.

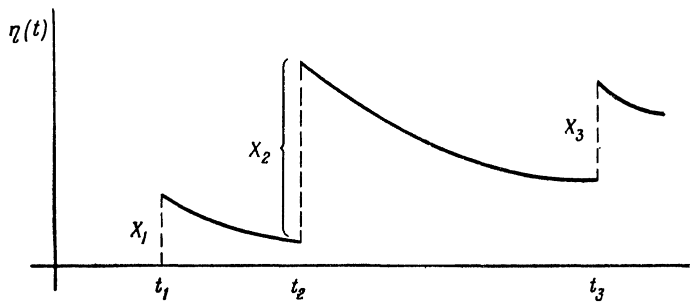


Рис. 1.

Предположим, что X_i — независимые и одинаково распределенные случайные величины с плотностью $h(x)$ и характеристической функцией

$$\psi(s) = \int_0^\infty e^{isx} h(x)\, dx.$$

Положим

$$R(v;\, t) = \mathrm{P}\{\eta(t) \leqslant v\} = \sum_{n=0}^{\infty} \mathrm{P}\{\eta(t) \leqslant v \mid N_t = n\}\, \mathrm{P}\{N_t = n\}. \qquad (3.1)$$

Конечно, $\mathrm{P}\{N_t = n\} = e^{-\lambda t}\dfrac{(\lambda t)^n}{n!}$, где λ — интенсивность пуассоновского потока, задающего моменты поступления импульсов. Из теоремы 2.3 известно, что t_i распределены, как упорядоченные наблюдения из равномерного распределения на $(0, t)$, при условии, что $N_t = n$ (т. е. на интервале $(0, t)$ поступило n импульсов). Пусть τ_i $(i = 1, 2, \ldots, N_t)$ — независимые одинаково распределенные на $(0, t)$ случайные величины, упорядочение которых по принимаемым значениям приводит к величинам t_j.

Далее, пусть $Z_1, \ldots, Z_n$ — независимые случайные величины, распределения которых совпадают с распределением X_i и которые также не зависят от $\{\tau_i\}$. Рассмотрим сумму

$$\sum_{i=1}^{n} Z_i \exp[-\alpha(t-\tau_i)]_+.$$

Определим новые случайные величины $Z_1', \ldots, Z_n'$ следующими равенствами:

$$Z_1' = Z_j, \quad \text{если} \quad \tau_j = \min(\tau_1, \ldots, \tau_n) = t_1,$$

$$Z_2' = Z_j, \quad \text{если} \quad \tau_j = t_2 - \text{второй по величине член из } \{\tau_i\},$$

$$\vdots$$

$$Z_n' = Z_j, \quad \text{если} \quad \tau_j = \max(\tau_1, \ldots, \tau_n) = t_n.$$

Неопределенность, которая возникает в случае равенства некоторых из величин τ_i, не причиняет беспокойства, поскольку вероятность этого события равна нулю. Тогда

$$\sum_{i=1}^{n} Z_i \exp[-\alpha(t-\tau_i)]_+ = \sum_{i=1}^{n} Z_i' \exp[-\alpha(t-t_i)]_+,$$

поскольку эти две суммы отличаются лишь порядком слагаемых. Далее, так как случайные величины Z_i являются независимыми, одинаково распределенными и также не зависящими от τ_i, то легко проверить, что и Z_i' — независимые случайные величины, их функция распределения совпадает с общей функцией распределения величин Z_i, и они также не зависят от τ_i. Будучи независимыми от τ_i, семейства $\{Z_i\}$ и $\{Z_i'\}$, очевидно, не зависят от t_i.

Поскольку Z_i' обладают всеми свойствами величин X_i, можно записать

$$\eta(t) = \sum_{i=1}^{n} Z_i' \exp[-\alpha(t-t_i)]_+ = \sum_{i=1}^{n} Z_i \exp[-\alpha(t-\tau_i)]_+.$$

Положим

$$Y_t(i) = Z_i \exp[-\alpha(t-\tau_i)]_+.$$

Очевидно, при фиксированном t с. в. $Y_t(i)$, $i = 1, \ldots, n$, независимы и одинаково распределены. Определим теперь характеристическую функцию величины $Y_t(k)$:

$$\theta_t(s) = \int_0^\infty e^{isy} g_t(y; k)\, dy,$$

где $g_t(y;\ k)$ — плотность величины $Y_t(k)$. Поскольку τ_k равномерно распределена на $(0,\ t)$ и τ_k и Z_k независимы, то

$$\int_0^y g_t(u,\ k)\,du = \mathrm{P}\{Y_t(k) \leqslant y\} = \mathrm{P}\left\{Z_k \exp[-\alpha(t-\tau_k)]_+ \leqslant y\right\} =$$

$$= \int_0^t \mathrm{P}\left\{Z_k \exp[-\alpha(t-\tau_k)]_+ \leqslant y \mid \tau_k = u\right\} \frac{du}{t} =$$

$$= \int_0^t \mathrm{P}\{Z_k \leqslant y e^{\alpha(t-u)}\} \frac{du}{t} = \frac{1}{t}\int_0^t H(ye^{\alpha(t-u)})\,du, \tag{3.2}$$

где H — функция распределения, соответствующая плотности h. Дифференцируя (3.2), получаем

$$g_t(y;\ k) = \frac{1}{t}\int_0^t h(ye^{\alpha(t-u)})\,e^{\alpha(t-u)}\,du.$$

Следовательно,

$$\theta_t(s) = \int_0^\infty e^{isy} g_t(y;\ k)\,dy = \frac{1}{t}\int_0^t e^{\alpha(t-u)}\left(\int_0^\infty e^{isy} h(ye^{\alpha(t-u)})\,dy\right)du =$$

(при замене переменных $ye^{\alpha(t-u)} = z$)

$$= \frac{1}{t}\int_0^t du \int_0^\infty \exp[is(e^{-\alpha(t-u)}z)]\,h(z)\,dz = \qquad \text{(в силу определения } \psi)$$

$$= \frac{1}{t}\int_0^t \psi(se^{-\alpha t}e^{\alpha u})\,du = (\text{положим } v = t-u) = \frac{1}{t}\int_0^t \psi(se^{-\alpha v})\,dv.$$

Отсюда если $r(x;\ t)$ — плотность распределения $R(x,\ t)$, то

$$\varphi_t(w) = \int_0^\infty e^{iwx} r(x;\ t)\,dx = \qquad \text{(в силу (3.1))}$$

$$= \sum_{n=0}^\infty \left(\int_0^\infty e^{iwx}\frac{d}{dx}\mathrm{P}\{\eta(t) \leqslant x \mid N_t = n\}\,dx\right) e^{-\lambda t}\frac{(\lambda t)^n}{n!} =$$

(в силу независимости величин $Y_t(k)$ при фиксированном N_t)

$$= \sum_{n=0}^\infty e^{-\lambda t}\frac{1}{n!}(\lambda t)^n [\theta_t(w)]^n = \sum_{n=0}^\infty e^{-\lambda t}\frac{1}{n!}\left(\lambda\int_0^t \psi(we^{-\alpha v})\,dv\right)^n =$$

$$= \exp\left\{-\lambda\int_0^t [1-\psi(we^{-\alpha v})]\,dv\right\}.$$

Дифференцируя по w, можно вычислить моменты $\eta(t)$. Например,

$$M[\eta(t)] = (-i)\frac{d}{dw}\varphi_t(w)\Big|_{w=0} = \lambda(-i)\psi'(0)\int_0^t e^{-\alpha v}\,dv = \\ = \lambda M(X_k)\frac{1-e^{-\alpha t}}{\alpha}.$$

§ 4. ПРОЦЕССЫ РОЖДЕНИЯ И ГИБЕЛИ

Одно из очевидных обобщений процессов чистого рождения, обсужденных в § 1, состоит в том, чтобы позволить процессу $X(t)$ как возрастать, так и убывать, например, из-за гибели членов популяции. Таким образом, если в момент t процесс находится в состоянии n, он может через некоторый случайный отрезок времени перейти в любое из соседних состояний $n+1$ или $n-1$. Возникающие при этом «процессы рождения и гибели» могут рассматриваться как процессы с непрерывным временем, служащие аналогами случайных блужданий (см. пример Б, § 2 гл. 2).

А. Постулаты

Мы предположим, что, как и в случае процессов чистого рождения, $X(t)$ является марковским процессом с состояниями 0, 1, 2, ... и что его вероятности перехода $P_{ij}(t)$ стационарны, т. е.

$$P_{ij}(t) = \mathrm{P}\{X(t+s) = j \mid X(s) = i\}.$$

Кроме того, предположим, что $P_{ij}(t)$ удовлетворяет постулатам:

1. $P_{i,\,i+1}(h) = \lambda_i h + o(h)$ при $h \downarrow 0,\ i \geqslant 0$.
2. $P_{i,\,i-1}(h) = \mu_i h + o(h)$ при $h \downarrow 0,\ i \geqslant 1$.
3. $P_{ii}(h) = 1 - (\lambda_i + \mu_i)h + o(h)$ при $h \downarrow 0,\ i \geqslant 0$.
4. $P_{ij}(o) = \delta_{ij}$.
5. $\mu_0 = 0,\ \lambda_0 > 0,\quad \mu_i,\ \lambda_i > 0,\quad i = 1, 2, \ldots$.

В каждом из этих случаев член $o(h)$ может зависеть от i. Матрица

$$\mathbf{A} = \left\| \begin{matrix} -\lambda_0 & \lambda_0 & 0 & 0 & \ldots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \ldots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \ldots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \ldots \\ \vdots & \vdots & \vdots & \vdots & . \end{matrix} \right\| \qquad (4.1)$$

называется *инфинитезимальной матрицей* процесса[1]). Параметры λ_i и μ_i называются инфинитезимальными интенсивностями рождения и гибели соответственно. В постулатах 1 и 2 предполагается, что если процесс находится в состоянии i, то за малый интервал времени вероятность того, что размер популяции увеличится или уменьшится на 1, пропорциональна длине этого интервала. Иногда рассматривают случай, когда нулевое состояние является поглощающим, т. е. $\lambda_0 = 0$ (см. § 7).

Поскольку $P_{ij}(t)$ — вероятности, то $P_{ij}(t) \geqslant 0$ и

$$\sum_{j=0}^{\infty} P_{ij}(t) = 1. \tag{4.2}$$

Используя марковское свойство процесса, можно получить уравнение Колмогорова — Чэпмена

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s). \tag{4.3}$$

Смысл этого уравнения следующий: для того чтобы перейти из состояния i в состояние j за время $t+s$, процесс $X(t)$ в момент t должен принять некоторое значение k, а затем перейти из этого состояния в j за оставшееся время s. Это аналог формулы (3.2) гл. 2 для непрерывного времени.

До сих пор рассматривались лишь вероятности перехода $P_{ij}(t)$. Для того чтобы найти вероятность события $\{X(t) = n\}$, нужно знать, в каком состоянии процесс находился в начальный момент времени, или, в более общем случае, — распределение начального состояния. Тогда

$$\mathrm{P}\{X(t) = n\} = \sum_{i=0}^{\infty} q_i P_{in}(t),$$

где

$$q_i = \mathrm{P}\{X(0) = i\}.$$

Б. Длительности пребывания

Используя принятые допущения, можно найти распределение случайной величины T_i, которая является длительностью пребывания процесса $X(t)$ в состоянии i. То есть мы найдем распределение времени T_i до первого выхода процесса из состояния i при условии, что вначале процесс находился в состоянии i. Обозначая

$$\mathrm{P}\{T_i \geqslant t\} = G_i(t),$$

[1]) В оригинале «infinitesimal generator of the process». — *Прим. перев.*

получим в силу марковского свойства, что при $h \downarrow 0$

$$G_i(t+h) = G_i(t)\,G_i(h) = G_i(t)\,(P_{ii}(h) + o(h)) = \\ = G_i(t)\,[1 - (\lambda_i + \mu_i)\,h] + o(h),$$

или

$$\frac{G_i(t+h) - G_i(t)}{h} = -(\lambda_i + \mu_i)\,G_i(t) + o(1),$$

и поэтому

$$G_i'(t) = -(\lambda_i + \mu_i)\,G_i(t). \tag{4.4}$$

Если использовать условие $G_i(0) = 1$, то найдем

$$G_i(t) = \exp[-(\lambda_i + \mu_i)\,t],$$

т. е. T_i имеет экспоненциальное распределение со средним $(\lambda_i + \mu_i)^{-1}$. Приведенное доказательство не вполне корректно, поскольку было использовано соотношение

$$G_i(h) = P_{ii}(h) + o(h)$$

без формального доказательства этого факта. Строгое доказательство справедливости (4.4) будет дано в § 4 гл. 8.

В соответствии с постулатами 1 и 2 в течение временного интервала длины h переход из состояния i в $i+1$ происходит с вероятностью $\lambda_i h + o(h)$, а в состояние $i-1$ — с вероятностью $\mu_i h + o(h)$. Интуитивно ясно, что при условии осуществления перехода в момент t вероятность того, что процесс перейдет при этом в состояние $i+1$, равна $\lambda_i(\mu_i + \lambda_i)^{-1}$, а вероятность перехода в состояние $i-1$ равна $\mu_i(\mu_i + \lambda_i)^{-1}$. Строгое доказательство этого утверждения выходит за рамки данной книги, однако обсуждение его и связанных с ним тонкостей будет дано ниже (см. гл. 8).

Развитие процесса $X(t)$ можно описать следующим образом. Процесс проводит в данном состоянии i случайное время, функция распределения которого является экспоненциальной с параметром $\lambda_i + \mu_i$. Из состояния i процесс переходит в состояния $i+1$ либо $i-1$ с вероятностями $\lambda_i(\mu_i + \lambda_i)^{-1}$ и $\mu_i(\mu_i + \lambda_i)^{-1}$ соответственно. Динамика аналогична случайному блужданию, за исключением того, что переходы осуществляются в случайные моменты времени, а не в фиксированные [1]).

Традиционная процедура построения реализаций процессов рождения и гибели состоит в задании параметров рождения и гибели $\{\lambda_i, \mu_i\}_{i=0}^{\infty}$ и построении выборочных траекторий следующим образом. Предположим, что $X(0) = i$; частица проводит случайное время (экспоненциально распределенное с параметром $\lambda_i + \mu_i$)

[1]) Весьма подробно эта аналогия рассмотрена в книге Дынкина Е. Б. и Юшкевича А. А. «Теоремы и задачи о процессах Маркова», «Наука», 1967. — *Прим. перев.*

в состоянии i, а затем переходит с вероятностью $\lambda_i(\mu_i+\lambda_i)^{-1}$ в состояние $i+1$, а с вероятностью $\mu_i(\mu_i+\lambda_i)^{-1}$ — в состояние $i-1$. Затем частица проводит случайное время в новом состоянии и вновь переходит в одно из соседних состояний и т. д. Более точно, мы выбираем значение t_1 из экспоненциального распределения с параметром $(\mu_i+\lambda_i)$, которое фиксирует время пребывания в состоянии i. Затем бросается монета, у которой вероятность выпадения герба равна $p_i=\lambda_i(\mu_i+\lambda_i)^{-1}$. Если выпадает герб (решетка), мы переводим частицу в положение $i+1$ $(i-1)$. В состоянии $i+1$ выбирается значение t_2 из экспоненциального распределения с параметром $(\mu_{i+1}+\lambda_{i+1})$, которое фиксирует время пребывания во втором по счету состоянии. Если частица при первом переходе попала в состояние $i-1$, то следующее время пребывания t_2' выбирается из экспоненциального распределения с параметром $(\mu_{i-1}+\lambda_{i-1})$. После этого вновь проводится биномиальное испытание, выбирающее следующее состояние, и т. д.

Исход этой процедуры выборок определяет реализацию процесса. Она может иметь вид

$$X(t)=\begin{cases} i, & 0<t<t_1,\\ i+1, & t_1<t<t_1+t_2,\\ i, & t_1+t_2<t<t_1+t_2+t_3,\\ \cdot & \\ \cdot & \\ \cdot & \end{cases}$$

Таким образом, делая выборки из экспоненциального и биномиального распределений соответственно, мы строим реализации процесса. Далее, возможно ввести на этом множестве реализаций (траекторий процесса) вероятностную меру таким образом, чтобы определяемые ей вероятности $P_{ij}(t)$ удовлетворяли соотношениям (4.2) и (4.3) и инфинитезимальным соотношениям (см. стр. 212). Этот результат довольно глубокий, и его строгое обсуждение лежит за пределами настоящей книги. Процесс, полученный таким образом, называется *минимальным процессом*, связанным с матрицей $\mathbf{A}$.

Приведенная выше конструкция минимального процесса является фундаментальной, поскольку инфинитезимальные параметры не определяют, вообще говоря, единственный вероятностный процесс, удовлетворяющий (4.2), (4.3) и постулатам, приведенным на стр. 212. На самом деле может быть несколько марковских процессов, обладающих одной и той же матрицей инфинитезимальных переходных вероятностей. Вопрос этот довольно сложен, и читатель может найти его в книге Чжун Кайлая[1]. В частном случае процессов рождения и гибели достаточ-

---

[1]) Чжун Кай-лай, Однородные цепи Маркова, «Мир», М., 1964.

ным условием, при котором существует единственный марковский процесс с переходными вероятностями $P_{ij}(t)$, удовлетворяющими инфинитезимальным соотношениям (4.2) и (4.3), является

$$\sum_{n=0}^{\infty} \pi_n \sum_{k=0}^{n} \frac{1}{\lambda_k \pi_k} = \infty, \tag{4.5}$$

где

$$\pi_0 = 1, \quad \pi_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n}, \quad n = 1, 2, \ldots .$$

В большинстве реальных примеров процессов рождения и гибели выполняется условие (4.5), и процесс, удовлетворяющий этому условию, определен единственным образом.

§ 5. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ДЛЯ ПРОЦЕССОВ РОЖДЕНИЯ И ГИБЕЛИ

Так же как в случае процесса чистого рождения и пуассоновского процесса, в данном случае переходные вероятности $P_{ij}(t)$ удовлетворяют системе дифференциальных уравнений, известных как *обратные дифференциальные уравнения Колмогорова*. Они имеют вид

$$\begin{aligned} &P'_{0j}(t) = -\lambda_0 P_{0j}(t) + \lambda_0 P_{1j}(t), \\ &P'_{ij}(t) = \mu_i P_{i-1,\,j}(t) - (\lambda_i + \mu_i) P_{ij}(t) + \lambda_i P_{i+1,\,j}(t), \quad i \geqslant 1. \end{aligned} \tag{5.1}$$

Начальное условие: $P_{ij}(0) = \delta_{ij}$.

Чтобы вывести их, запишем (4.3) в виде

$$\begin{aligned} P_{ij}(t+h) &= \sum_{k=0}^{\infty} P_{ik}(h) P_{kj}(t) = \\ &= P_{i,\,i-1}(h) P_{i-1,\,j}(t) + P_{ii}(h) P_{ij}(t) + P_{i,\,i+1}(h) P_{i+1,\,j}(t) + \\ &\qquad + \sum_k{}' P_{ik}(h) P_{kj}(t), \end{aligned} \tag{5.2}$$

где последняя сумма берется по всем $k \neq i-1,\ i,\ i+1$. Используя постулаты 1, 2 и 3 (§ 4), получаем

$$\begin{aligned} \sum_k{}' P_{ik}(h) P_{kj}(t) &\leqslant \sum_k{}' P_{ik}(h) = 1 - [P_{ii}(h) + P_{i,\,i-1}(h) + P_{i,\,i+1}(h)] = \\ &= 1 - [1 - (\lambda_i + \mu_i) h + o(h) + \mu_i h + o(h) + \lambda_i h + o(h)] = o(h), \end{aligned}$$

так что

$$P_{ij}(t+h) = \mu_i h P_{i-1,\,j}(t) + (1 - (\lambda_i + \mu_i) h) P_{ij}(t) + \lambda_i h P_{i+1,\,j}(t) + o(h).$$

Перенося член $P_{ij}(t)$ в левую часть, деля полученное равенство на h и переходя к пределу при $h \downarrow 0$, получим

$$P'_{ij}(t) = \mu_i P_{i-1,\,j}(t) - (\lambda_i + \mu_i) P_{ij}(t) + \lambda_i P_{i+1,\,j}(t).$$

Проведенные выкладки являются частным случаем вывода обратных дифференциальных уравнений, который дается в гл. 8, § 2.

Обратные уравнения выводятся с помощью разбиения временного интервала $(0,\ t+h)$, где h положительно и мало́, на два: $(0,\ h)$ и $(h,\ t+h)$ и отдельного рассмотрения переходов процесса на каждом из них.

В уравнениях (5.1) начальное состояние является переменным, а конечное — фиксированным параметром.

Другая ситуация возникает при разделении интервала $(0,\ t+h)$ на $(0,\ t)$ и $(t,\ t+h)$ и применении предыдущего анализа. В этом случае при более жестких условиях можно вывести следующую систему дифференциальных уравнений:

$$\begin{aligned} P'_{i0}(t) &= -\lambda_0 P_{i0}(t) + \mu_1 P_{i1}(t), \\ P'_{ij}(t) &= \lambda_{j-1} P_{i,\,j-1}(t) - (\lambda_j + \mu_j) P_{ij}(t) + \mu_{j+1} P_{i,\,j+1}(t), \quad j \geqslant 1, \end{aligned} \tag{5.3}$$

с тем же самым начальным условием $P_{ij}(0) = \delta_{ij}$. Они известны как *прямые дифференциальные уравнения Колмогорова*. Чтобы проделать это, поменяем местами t и h в (5.2) и при дополнительных предположениях (кроме постулатов 1, 2, 3) можно показать, что последний член снова равен $o(h)$. Остальное дословно повторяет предыдущий вывод. Полезность этих дифференциальных уравнений станет очевидной из примеров, которые будут разобраны ниже.

Достаточным условием для справедливости (5.3) является равенство $\frac{1}{h} P_{kj}(h) = o(1)$ при $k \neq j,\ j-1,\ j+1$, где член $o(1)$ (кроме того, что он стремится к 0) равномерно ограничен по k при фиксированном j при $h \to 0$. В этом случае нетрудно показать, что

$$\sum_k{}' P_{ik}(t) P_{kj}(h) = o(h).$$

Прежде чем переходить к примерам, обсудим кратко поведение $P_{ij}(t)$ при больших t. Можно доказать, что пределы

$$\lim_{t \to \infty} P_{ij}(t) = p_j \tag{5.4}$$

существуют, не зависят от начального состояния i и удовлетворяют уравнениям

$$\begin{aligned} -\lambda_0 p_0 + \mu_1 p_1 &= 0, \\ \lambda_{j-1} p_{j-1} - (\lambda_j + \mu_j) p_j + \mu_{j+1} p_{j+1} &= 0, \quad j \geqslant 1. \end{aligned} \tag{5.5}$$

Эти уравнения получаются из (5.3), если приравнять нулю левую часть. Сходимость $\sum_j p_j$ следует из того, что $\sum_j P_{ij}(t) = 1$. Если $\sum_j p_j = 1$, то последовательность $\{p_j\}$ называется *стационарным распределением*. Причиной для этого служит то, что p_j удовлетворяют уравнению

$$p_j = \sum_{i=0}^{\infty} p_i P_{ij}(t), \tag{5.6}$$

которое говорит, что если процесс начинается из состояния i с вероятностью p_i, то и в любой последующий фиксированный момент времени он будет находиться в состоянии i с вероятностью p_i. Доказательство (5.6) следует из (4.3) и (5.4), если устремить $t \uparrow \infty$ и использовать то, что $\sum_{i=0}^{\infty} p_i < \infty$. Решение уравнений (5.5) получается по индукции. Полагая

$$\pi_0 = 1, \quad \pi_j = \frac{\lambda_0 \lambda_1 \ldots \lambda_{j-1}}{\mu_1 \mu_2 \ldots \mu_j}, \qquad j \geqslant 1,$$

получим $p_1 = \lambda_0 \mu_1^{-1} p_0 = \pi_1 p_0$. Предполагая, что $p_k = \pi_k p_0$ при $k = 1, \ldots, j$, получаем

$$\mu_{j+1} p_{j+1} = (\lambda_j + \mu_j) \pi_j p_0 - \lambda_{j-1} \pi_{j-1} p_0 =$$
$$= \lambda_j \pi_j p_0 + (\mu_j \pi_j - \lambda_{j-1} \pi_{j-1}) p_0 = \lambda_j \pi_j p_0$$

и, наконец,

$$p_{j+1} = \pi_{j+1} p_0.$$

Для того чтобы последовательность $\{p_j\}$ являлась распределением, нужно, чтобы $\Sigma p_j = 1$. Если $\Sigma \pi_k < \infty$, то в этом случае

$$p_j = \frac{\pi_j}{\sum \pi_k}, \qquad j = 0, 1, 2, \ldots .$$

Если $\Sigma \pi_k = \infty$, то с необходимостью $p_0 = 0$ и все p_j равны нулю. Следовательно, не существует предельного стационарного распределения.

§ 6. ПРИМЕРЫ ПРОЦЕССОВ РОЖДЕНИЯ И ГИБЕЛИ

Пример 1. *Линейный рост с иммиграцией.* Процесс рождения и гибели называется *процессом с линейным ростом*, если $\lambda_n = \lambda n + a$, $\mu_n = \mu n$, где $\lambda > 0$, $\mu > 0$, $a > 0$. Такие процессы возникают при изучении биологического воспроизведения и роста популяций. Если состояние n описывает текущий размер популяции, то средняя мгновенная интенсивность роста равна $\lambda n + a$.

Аналогично вероятность того, что состояние процесса уменьшится на 1 за малый промежуток времени, равна $\mu nt + o(t)$.

Коэффициент λn соответствует естественному приросту популяции размера n, в то время как второй коэффициент a можно интерпретировать как инфинитезимальную интенсивность роста популяции за счет внешнего источника, такого, как иммиграция. Коэффициент μn, который равен средней инфинитезимальной интенсивности гибели в популяции размера n, имеет очевидную интерпретацию.

Если подставить эти значения λ_n и μ_n в (5.3), то получим

$$P'_{i0}(t) = -aP_{i0}(t) + \mu P_{i1}(t),$$
$$P'_{ij}(t) = (\lambda(j-1)+a) P_{i,j-1}(t) - ((\lambda+\mu)j + a) P_{ij}(t) + \\ + \mu(j+1) P_{i,j+1}(t), \qquad j \geqslant 1.$$

Если теперь умножить j-е уравнение на j и просуммировать уравнения, получим, что среднее

$$\mathrm{M}[X(t)] = M(t) = \sum_{j=1}^{\infty} jP_{ij}(t)$$

удовлетворяет дифференциальному уравнению

$$M'(t) = a + (\lambda - \mu) M(t)$$

с начальным условием $M(0) = i$, если $X(0) = i$. Решение этого уравнения имеет вид

$$M(t) = at + i, \qquad \text{если} \quad \lambda = \mu,$$
$$M(t) = \frac{a}{\lambda - \mu}\{e^{(\lambda-\mu)t} - 1\} - ie^{(\lambda-\mu)t}, \qquad \text{если} \quad \lambda \neq \mu.$$

Второй момент или дисперсия могут быть найдены аналогично. Интересно заметить, что $M(t) \to \infty$ при $t \to \infty$, если $\lambda \geqslant \mu$, в то время как при $\lambda < \mu$ средний размер популяции при больших t примерно равен

$$\frac{a}{\mu - \lambda}.$$

Пример 2. *Образование очереди.* Процесс образования очереди является процессом, в котором клиенты прибывают в некоторое определенное место (обслуживающий прибор), где им оказываются услуги какого-либо рода, например окно кассира в банке или место около кассира в магазине самообслуживания. Предполагается, что интервалы между прибытиями клиентов (поступлениями требований) и время, которое проведено данным клиентом на обслуживании, управляются вероятностными законами.

Длину очереди в данный момент времени t обозначим через $X(t)$ [1]).

Если в описании общего процесса рождения и гибели положить $\lambda_i = \lambda$ для всех i, то получим простой частный случай процесса обслуживания с непрерывным временем. Состояние системы при этом интерпретируется как длина очереди, в которую поступают клиенты через независимые друг от друга интервалы, имеющие экспоненциальное распределение с параметром λ, и для которой продолжительность времени обслуживания очередного клиента является случайной величиной, имеющей экспоненциальное распределение с параметром μ_n, который может зависеть от длины очереди. После завершения каждого акта обслуживания длина очереди убывает на 1, а при каждом поступлении очередь увеличивается на 1. Классический случай одноканальной системы (системы с одним обслуживающим прибором) соответствует $\mu_i = \mu$, $i \geqslant 1$, т. е. каждое время обслуживания имеет одно и то же экспоненциальное распределение с параметром μ, не зависящим от длины очереди.

Классическая модель телефонного узла может быть сформулирована как система обслуживания, описываемая процессом рождения и гибели, с бесконечным числом обслуживающих приборов, каждый из которых имеет экспоненциально распределенное время обслуживания с одним и тем же параметром μ, так что $\mu_i = i\mu$, $i \geqslant 1$. Это обосновывается следующим образом. Предположим, что очередь состоит из i клиентов, тогда, поскольку число обслуживающих приборов неограничено, все клиенты одновременно обслуживаются. Далее, времена обслуживания каждого из них независимы и экспоненциально распределены с параметром μ. Отсюда следует, что вероятностное распределение времени до того момента, пока по крайней мере один из клиентов закончит обслуживаться (т. е. времени до момента уменьшения очереди на 1), также является экспоненциальным, но с параметром $i\mu$ (читатель должен доказать это).

Кроме двух рассмотренных частных случаев, можно рассмотреть другие многочисленные модели обслуживания, выбирая соответствующим образом параметры μ_k. Например, система с n обслуживающими приборами, каждый из которых имеет экспоненциально распределенное время обслуживания с одним и тем же параметром μ, соответствовала бы случаю $\mu_k = k\mu$ при $1 \leqslant k \leqslant n$, $\mu_i = n\mu$ при $i \geqslant n$.

Для случая одного обслуживающего прибора при $\lambda < \mu$ стационарное распределение легко находится. Действительно, в этом

---

[1]) Под «очередью» здесь и ниже понимается число клиентов, находящихся в системе обслуживания, т. е. суммарное число обслуживаемых и ожидающих обслуживания клиентов. — *Прим. перев.*

случае

$$\pi_n = \frac{\lambda_0 \lambda_1 \ldots \lambda_{n-1}}{\mu_1 \mu_2 \ldots \mu_n} = \left(\frac{\lambda}{\mu}\right)^n,$$

откуда

$$p_n = \frac{\mu - \lambda}{\mu} \left(\frac{\lambda}{\mu}\right)^n, \qquad n \geqslant 0,$$

т. е. получили геометрическое распределение со средним $\lambda(\mu - \lambda)^{-1}$.

Для модели телефонного узла $\lambda_n = \lambda$, $\mu_n = n\mu$ и легко получить, что

$$p_n = e^{-\frac{\lambda}{\mu}} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!}.$$

Получили известное пуассоновское распределение со средним $\frac{\lambda}{\mu}$. Как и в примере 1, легко показать, что

$$M(t) = \sum_{j=0}^{\infty} j P_{ij}(t)$$

удовлетворяет уравнению

$$M'(t) = \lambda - \mu M(t),$$

решение которого имеет вид

$$M(t) = \frac{\lambda}{\mu}(1 - e^{-\mu t}) + i e^{-\mu t}.$$

Если положить $t \to \infty$, то $M(t) \to \frac{\lambda}{\mu}$, которое равно среднему значению стационарного распределения, приведенному выше.

Пример 3. *Некоторые генетические модели.* Рассмотрим популяцию, состоящую из N индивидуумов, которые имеют либо гены a, либо гены A. Под состоянием процесса $X(t)$ будем понимать число индивидуумов с генами a в момент t. Предположим, что вероятность того, что состояние изменится за интервал времени $(t, t+h)$, равна $\lambda h + o(h)$ и не зависит от $X(t)$, а вероятность двух или более изменений за время h равна $o(h)$.

Изменения в составе популяции происходят следующим образом. Индивидуум, который должен быть заменен другим, выбирается из популяции случайным образом. То есть если $X(t) = j$, то индивидуум с геном a отбирается для замены с вероятностью j/N, а индивидуум с геном A — с вероятностью $1 - j/N$. Мы будем называть данную замену «гибелью». Далее, «рождение» происходит по следующему правилу. Из популяции делается еще один случайный выбор, с тем чтобы определить тип нового индивидуума, который займет место погибшего. В модели вводится мутационное давление, которое создает возможность того, что тип нового индивидуума может измениться после рождения. Именно: пусть γ_1 — вероятность того, что ген a мутирует в ген A, и γ_2 — вероятность того, что ген A мутирует в ген a.

Вероятность того, что новый индивидуум, добавленный к популяции, имеет ген a, равна

$$\frac{j}{N}(1-\gamma_1)+\left(1-\frac{j}{N}\right)\gamma_2. \tag{6.1}$$

Эта формула получена следующим образом. Вероятность того, что будет выбран ген a и не произойдет мутации, равна $\frac{j}{N}(1-\gamma_1)$. Кроме того, новый ген индивидуума может иметь ген a, если будет выбран ген A, который затем мутирует в ген a. Вероятность этого события равна $\left(1-\frac{j}{N}\right)\gamma_2$. Комбинация этих двух возможностей дает формулу (6.1).

Мы утверждаем, что вероятность события $X(t+)-X(t)=1$ при условии, что в момент t имело место изменение состояния, равна

$$\left(1-\frac{j}{N}\right)\left[\frac{j}{N}(1-\gamma_1)+\left(1-\frac{j}{N}\right)\gamma_2\right], \quad \text{где} \quad X(t)=j. \tag{6.2}$$

В самом деле, число индивидуумов с геном a увеличивается только в случае гибели (замены) индивидуума с геном A. Вероятность этого равна $\left(1-\frac{j}{N}\right)$. Второй сомножитель есть вероятность того, что новый индивидуум имеет ген a (см. (6.1)).

Аналогично вероятность того, что $X(t+)-X(t)=-1$ при условии, что в момент t произошло изменение состояния, равна

$$\frac{j}{N}\left[\left(1-\frac{j}{N}\right)(1-\gamma_2)+\frac{j}{N}\gamma_1\right], \quad \text{где} \quad X(t)=j.$$

Описанный вероятностный процесс является, таким образом, процессом рождения и гибели с конечным числом состояний[1]), инфинитезимальные интенсивности рождения и гибели которого равны

$$\lambda_j=\lambda\left(1-\frac{j}{N}\right)\left[\frac{j}{N}(1-\gamma_1)+\left(1-\frac{j}{N}\right)\gamma_2\right],$$
$$\mu_j=\lambda\frac{j}{N}\left[\frac{j}{N}\gamma_1+\left(1-\frac{j}{N}\right)(1-\gamma_2)\right]$$

ссответственно при числе индивидуумов с геном a, равном j, $0\leqslant j\leqslant N$.

---

[1]) Определение процессов рождения и гибели было дано для случая бесконечного числа состояний. Случай конечного числа состояний более прост, и необходимые изменения в определениях и анализе предлагается сделать читателю.

Хотя эти параметры кажутся довольно сложными, интересно посмотреть, что́ произойдет со стационарным распределением $\{\pi_k\}_{k=0}^N$ при $N \to \infty$ и вероятностях мутации для отдельных индивидуумов γ_1 и γ_2, стремящихся к нулю так, что $\gamma_1 N \to \varkappa_1$, $\gamma_2 N \to \varkappa_2$, где $0 < \varkappa_1, \varkappa_2 < \infty$. Одновременно мы будем считать состояние процесса изменяющимся на отрезке $[0, 1]$, принимая в качестве него $\frac{j}{N}$, т. е. долю индивидуумов с геном a в популяции. Чтобы найти плотность состояния x, $0 < x < 1$, оценим π_k при $k \to \infty$, где $k = [xN]$, а $[xN]$ — наибольшее целое число, не превышающее xN. Имея это в виду, запишем

$$\lambda_j = \frac{\lambda(N-j)}{N^2}(1-\gamma_1-\gamma_2)\,j\left(1+\frac{a}{j}\right), \qquad \text{где} \quad a = \frac{N\gamma_2}{1-\gamma_1-\gamma_2},$$

$$\mu_j = \frac{\lambda(N-j)}{N^2}(1-\gamma_1-\gamma_2)\,j\left(1+\frac{b}{N-j}\right), \quad \text{где} \quad b = \frac{N\gamma_1}{1-\gamma_1-\gamma_2}.$$

Тогда

$$\ln \pi_k = \sum_{j=0}^{k-1} \ln \lambda_j - \sum_{j=1}^{k} \ln \mu_j = \sum_{j=1}^{k-1} \ln\left(1+\frac{a}{j}\right) - \\ - \sum_{j=1}^{k-1} \ln\left(1+\frac{b}{N-j}\right) + \ln Na - \ln(N-k)\,k\left(1+\frac{b}{N-k}\right).$$

Используя разложение

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots, \quad |x| < 1,$$

можно записать

$$\sum_{j=1}^{k-1} \ln\left(1+\frac{a}{j}\right) = a\sum_{j=1}^{k-1}\frac{1}{j} + c_k,$$

где c_k имеет конечный предел при $k \to \infty$. Следовательно, используя соотношение

$$\sum_{j=1}^{k-1}\frac{1}{j} \sim \ln k \quad \text{при} \quad k \to \infty,$$

имеем

$$\sum_{j=1}^{k-1} \ln\left(1+\frac{a}{j}\right) \sim \ln k^a + c_k \quad \text{при} \quad k \to \infty.$$

Аналогично можно получить

$$\sum_{j=1}^{k-1} \ln\left(1+\frac{b}{N-j}\right) \sim \ln\frac{N^b}{(N-k)^b} + d_k \quad \text{при} \quad k\to\infty,$$

где d_k имеет конечный предел при $k\to\infty$. Используя выписанные соотношения, имеем

$$\ln\pi_k \sim \ln\left(C_k \frac{k^a (N-k)^b Na}{N^b (N-k) k}\right) \quad \text{при} \quad k\to\infty, \tag{6.3}$$

где $\ln C_k = c_k + d_k$ и $C_k \to C < \infty$ при $k\to\infty$. Заметим, что $a\to\varkappa_2$ и $b - \varkappa_1$, при $N\to\infty$. Поскольку $k=[Nx]$, то при $N\to\infty$ имеем

$$\pi_k \sim C\varkappa_2 N^{\varkappa_2-1} x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}.$$

Далее, из (6.3) имеем

$$\pi_k \sim aC_k k^{a-1}\left(1-\frac{k}{N}\right)^{b-1}.$$

Следовательно,

$$\frac{1}{N^a}\sum_{k=0}^{N-1}\pi_k \sim \frac{a}{N}\sum_{k=0}^{N-1} C_k\left(\frac{k}{N}\right)^{a-1}\left(1-\frac{k}{N}\right)^{b-1}.$$

Поскольку $C_k\to C$ при $k\to\infty$, то в правой части стоит аппроксимация суммы Римана для интеграла

$$\varkappa_2 C\int_0^1 x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}\,dx.$$

Таким образом,

$$\sum_{i=0}^{N}\pi_i \sim N^{\varkappa_2}\varkappa_2 C\int_0^1 x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}\,dx,$$

так что плотность в (0,1) равна

$$\frac{\pi_k}{\sum \pi_i} \sim \frac{1}{N}\,\frac{x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}}{\int_0^1 x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}\,dx} = \frac{x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}\,dx}{\int_0^1 x^{\varkappa_2-1}(1-x)^{\varkappa_1-1}\,dx},$$

поскольку $dx \sim \frac{1}{N}$. Получили бета-распределение с параметрами $\varkappa_1$ и $\varkappa_2$.

Пример 4. *Логистический процесс.* Предположим, что мы рассматриваем популяцию, размер которой $X(t)$ изменяется в пределах от N_1 до N_2 (N_1, N_2 — целые, $N_1 < N_2$) при всех t. Пусть интенсивности рождения и гибели для каждого индивидуума в момент t равны

$$\lambda = \alpha (N_2 - X(t)), \quad \mu = \beta (X(t) - N_1),$$

а отдельные индивидуумы в популяции развиваются независимо друг от друга. Результирующие интенсивности рождения и гибели для всей популяции равны

$$\lambda_n = \alpha n (N_2 - n), \quad \mu_n = \beta n (n - N_1).$$

Чтобы показать это, заметим, что если размер популяции $X(t)$ равен n, то каждый из n индивидуумов имеет инфинитезимальную интенсивность рождения λ, так что $\lambda_n = \alpha n(N_2 - n)$. Такое же обоснование можно предложить и для интерпретации μ_n.

В этих предположениях естественно ожидать, что процесс изменяется между двумя уровнями N_1 и N_2, поскольку если, скажем, $X(t)$ близок к N_2, то интенсивность гибели высока, а интенсивность рождения низка и, следовательно, $X(t)$ стремится к N_1. В результате процесс должен флуктуировать стационарным образом между двумя уровнями N_1 и N_2.

Стационарное распределение в этом случае равно

$$p_{N_1+m} = \frac{c}{N_1+m} \binom{N_2 - N_1}{m} \left(\frac{\alpha}{\beta}\right)^m, \qquad m = 0, 1, 2, \ldots, N_2 - N_1,$$

где c — константа, определяемая из условия $\sum\limits_m p_{N_1+m} = 1$. Чтобы показать это, заметим, что

$$\pi_{N_1+m} = \frac{\lambda_0 \lambda_1 \ldots \lambda_{m-1}}{\mu_1 \mu_2 \ldots \mu_m} =$$
$$= \frac{\alpha^m N_1 (N_1+1) \ldots (N_1+m-1)(N_2-N_1) \ldots (N_2-N_1-m+1)}{\beta^m (N_1+1) \ldots (N_1+m)\, m!} =$$
$$= \frac{N_1}{N_1+m} \binom{N_2-N_1}{m} \left(\frac{\alpha}{\beta}\right)^m.$$

§ 7. ПРОЦЕССЫ РОЖДЕНИЯ И ГИБЕЛИ С ПОГЛОЩАЮЩИМИ СОСТОЯНИЯМИ

Представляет интерес рассмотреть такие процессы рождения и гибели, у которых $\lambda_0 = 0$. Это предположение делает нулевое состояние поглощающим. Если переход осуществляется из состояния 1, частица перемещается в состояние 2 с вероятностью $\lambda_1(\mu_1 + \lambda_1)^{-1}$ или попадает в состояние 0 (и остается там в дальнейшем) с вероятностью $\mu_1(\mu_1 + \lambda_1)^{-1}$. Важным примером процесса рождения и гибели, где состояние 0 поглощающее, является процесс линейного роста без иммиграции (ср. с примером 1 § 6).

8 Зак. 939

В этом случае $\lambda_n = n\lambda$ и $\mu_n = n\mu$. Поскольку рост популяции обязан исключительно существованию популяции, ясно, что когда размер популяции становится равным 0, он остается равным нулю и после этого, т. е. 0 является поглощающим состоянием.

А. Вероятность поглощения в состоянии 0

Представляет интерес нахождение вероятности поглощения в состоянии 0 при начальном состоянии i ($i \geqslant 1$). Такое поглощение не является априори достоверным событием, поскольку возможно, что частица (т. е. переменная, характеризующая состояние) будет все время блуждать по состояниям 1, 2, ... или, быть может, уходить в бесконечность.

Пусть u_i ($i = 1, 2, \ldots$) — вероятность поглощения в состоянии 0 из начального состояния i. Можно написать рекуррентное соотношение для u_i, рассматривая возможные состояния после первого перехода. Мы знаем, что при первом скачке осуществляются следующие переходы:

$$\begin{array}{lll} i \to i+1 & \text{с вероятностью} & \lambda_i (\mu_i + \lambda_i)^{-1}, \\ i \to i-1 & \text{с вероятностью} & \mu_i (\mu_i + \lambda_i)^{-1}. \end{array}$$

Непосредственно получаем

$$u_i = \frac{\lambda_i}{\mu_i + \lambda_i} u_{i+1} + \frac{\mu_i}{\mu_i + \lambda_i} u_{i-1}, \qquad i \geqslant 1, \tag{7.1}$$

где $u_0 = 1$. Другой метод получения (7.1) состоит в рассмотрении «вложенного случайного блуждания», связанного с процессом рождения и гибели. Именно рассмотрим процесс рождения и гибели только в моменты переходов. Цепь Маркова с дискретным временем, получающуюся при этом, обозначим через $\{Y_n\}_{n=0}^{\infty}$, где $Y_0 = X_0$ — начальное состояние, а Y_n ($n \geqslant 1$) — состояние после n-го перехода. Очевидно, матрица переходных вероятностей имеет вид

$$\mathbf{P} = \left\| \begin{array}{lllll} 1 & 0 & 0 & 0 & \ldots \\ q_1 & 0 & p_1 & 0 & \ldots \\ 0 & q_2 & 0 & p_2 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{array} \right\|,$$

где

$$p_i = \frac{\lambda_i}{\mu_i + \lambda_i} = 1 - q_i \qquad (i \geqslant 1).$$

Вероятность поглощения в состоянии 0 для вложенного случайного блуждания совпадает с аналогичной вероятностью для процесса рождения и гибели, поскольку оба процесса совершают одни и те же переходы.

Обратимся к задаче решения (7.1) при условиях, что $u_0 = 1$ и $0 \leqslant u_i \leqslant 1$ $(i \geqslant 1)$. Перепишем (7.1):

$$(u_{i+1} - u_i) = \frac{\mu_i}{\lambda_i}(u_i - u_{i-1}), \qquad i \geqslant 1.$$

Введем переменную $v_i = u_{i+1} - u_i$. Получим

$$v_i = \frac{\mu_i}{\lambda_i} v_{i-1}, \qquad i \geqslant 1.$$

Итерируя последнее соотношение, найдем

$$u_{i+1} - u_i = v_i = \left(\prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}\right) v_0, \qquad i \geqslant 1.$$

Суммируя эти уравнения от $i = 1$ до $i = m$, получим

$$u_{m+1} - u_1 = (u_1 - 1)\sum_{j=1}^{m}\left(\prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}\right), \qquad m \geqslant 1. \tag{7.2}$$

Поскольку u_m в силу определения ограничена числом 1, мы видим, что если

$$\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}\right) = \infty, \tag{7.3}$$

то с необходимостью $u_1 = 1$ и $u_m = 1$ при всех $m \geqslant 2$. Другими словами, если выполнено равенство (7.3), то из любого начального состояния происходит поглощение в 0 с вероятностью 1.

Пусть $0 < u_1 < 1$. Тогда, конечно,

$$\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i} \frac{\mu_j}{\lambda_j}\right) < \infty.$$

Очевидно, u_m не возрастает с ростом m, поскольку при переходе из состояния m в состояние 0 процесс попадает во все промежуточные состояния [1]). Более того, мы утверждаем, что $u_m \to 0$ при $m \to \infty$.

Если предположить противное, т. е. $u_m \geqslant \alpha > 0$ $(m \geqslant 1)$, то простые вероятностные рассуждения показывают, что $u_m \equiv 1$

[1]) Иначе говоря, событие, заключающееся в том, что за время t произошел переход $m \to 0$, содержится в событии «за время t произошел переход $k \to 0$», где $k < m$, а t — любое положительное число. — *Прим. перев.*

($m \geqslant 1$). (Читатель должен уметь доказать это формально.) Далее, устремляя $m \to \infty$ в (7.2), найдем

$$u_1 = \frac{\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{\mu_j}{\lambda_j}\right)}{1+\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{\mu_j}{\lambda_j}\right)}$$

и, кроме того имеем

$$u_{m+1} = \frac{\sum_{i=m+1}^{\infty}\left(\prod_{j=1}^{i}\frac{\mu_j}{\lambda_j}\right)}{1+\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i}\frac{\mu_j}{\lambda_j}\right)}.$$

В частном случае процесса рождения и гибели с линейным ростом, где $\mu_n = n\mu$ и $\lambda_n = n\lambda$, непосредственное вычисление показывает, что

$$u_m = \left(\frac{\mu}{\lambda}\right)^m, \quad \text{если} \quad \mu < \lambda,$$

$$u_m = 1, \quad \text{если} \quad \mu \geqslant \lambda.$$

Б. Среднее время до поглощения

Рассмотрим задачу нахождения среднего времени до поглощения, начиная из состояния m.

Предположим, что условие (7.3) выполнено, так что поглощение происходит с вероятностью 1. Заметим, что в данном случае задачу нельзя свести к изучению вложенного случайного блуждания, поскольку действительное время, проводимое процессом в каждом состоянии, существенно для нахождения среднего времени до поглощения.

Пусть ω_i — среднее время (быть может, бесконечное) до поглощения при начальном состоянии i. Рассматривая возможные состояния, следующие за первым переходом, и вспоминая тот факт, что среднее время пребывания в состоянии i равно $(\mu_i + \lambda_i)^{-1}$ (это время имеет экспоненциальное распределение с параметром $\mu_i + \lambda_i$), получим рекуррентное соотношение

$$\omega_i = \frac{1}{\mu_i+\lambda_i} + \frac{\lambda_i}{\mu_i+\lambda_i}\omega_{i+1} + \frac{\mu_i}{\mu_i+\lambda_i}\omega_{i-1}, \qquad i \geqslant 1, \tag{7.4}$$

где по условию $\omega_0 = 0$. Полагая $z_i = \omega_i - \omega_{i+1}$ и комбинируя члены в (7.4), получим

$$z_i = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i} z_{i-1}, \qquad i \geqslant 1. \tag{7.5}$$

Итерируя это соотношение, найдем

$$z_m = \frac{1}{\lambda_m} + \frac{\mu_m}{\lambda_m}\frac{1}{\lambda_{m-1}} + \frac{\mu_m\mu_{m-1}}{\lambda_m\lambda_{m-1}} z_{m-2}$$

и, наконец,

$$z_m = \sum_{i=1}^{m} \frac{1}{\lambda_i} \prod_{j=i+1}^{m} \frac{\mu_j}{\lambda_j} + \left(\prod_{j=1}^{m} \frac{\mu_j}{\lambda_j}\right) z_0.$$

$\left(\text{Произведение } \prod_{m+1}^{m} \frac{\mu_j}{\lambda_j} \text{ полагается равным } 1.\right)$ Иначе

$$\omega_m - \omega_{m+1} = \sum_{i=1}^{m} \frac{1}{\lambda_i} \prod_{j=i+1}^{m} \frac{\mu_j}{\lambda_j} - \omega_1 \prod_{j=1}^{m} \frac{\mu_j}{\lambda_j}, \qquad m \geqslant 1. \tag{7.6}$$

Более удобно записать

$$\sum_{i=1}^{m} \frac{1}{\lambda_i} \prod_{j=i+1}^{m} \frac{\mu_j}{\lambda_j} = \prod_{j=1}^{m} \frac{\mu_j}{\lambda_j} \sum_{i=1}^{m} \rho_i, \tag{7.7}$$

где

$$\rho_i = \frac{\lambda_1\lambda_2 \ldots \lambda_{i-1}}{\mu_1\mu_2 \ldots \mu_i}.$$

Затем, используя (7.7), соотношение (7.6) перепишем в виде

$$\left(\prod_{j=1}^{m} \frac{\lambda_j}{\mu_j}\right)(\omega_m - \omega_{m+1}) = \sum_{i=1}^{m} \rho_i - \omega_1. \tag{7.8}$$

Заметим, что если $\sum_{i=1}^{\infty} \rho_i = \infty$, то из (7.8) следует, что с необходимостью $\omega_1 = \infty$ [1]). В самом деле, из вероятностных соображений очевидно, что $\omega_m < \omega_{m+1}$ для всех m, а это свойство было бы нарушено для больших m, если бы мы предположили, что $\omega_1 < \infty$.

Предположим теперь, что $\sum_{i=1}^{\infty} \rho_i < \infty$. Тогда, полагая $m \to \infty$ в (7.8), получим

$$\omega_1 = \sum_{i=1}^{\infty} \rho_i - \lim_{m\to\infty} \left(\prod_{j=1}^{m} \frac{\lambda_j}{\mu_j}\right)(\omega_m - \omega_{m+1}).$$

[1]) И, следовательно, $\omega_m = \infty$, $m \geqslant 1$. — *Прим. перев.*

Можно показать, хотя и достаточно сложно, что

$$\lim_{m\to\infty}\left(\prod_{j=1}^{m}\frac{\lambda_j}{\mu_j}\right)(\omega_m-\omega_{m+1})=0.$$

Тогда, очевидно,

$$\omega_1=\sum_{i=1}^{\infty}\rho_i.$$

Подведем итоги этого параграфа в виде следующей теоремы.

Теорема 7.1. *Рассмотрим процесс рождения и гибели с параметрами* λ_n *и* μ_n, $n\geqslant 1$, $\lambda_0=0$, *так что* 0 *является поглощающим состоянием.*

Вероятность поглощения в состоянии 0 *при начальном состоянии* m *равна*

$$\begin{cases}\dfrac{\sum\limits_{i=m}^{\infty}\left(\prod\limits_{j=1}^{i}\dfrac{\mu_j}{\lambda_j}\right)}{1+\sum\limits_{i=1}^{\infty}\left(\prod\limits_{j=1}^{i}\dfrac{\mu_j}{\lambda_j}\right)}, & \text{если}\quad \sum\limits_{i=1}^{\infty}\left(\prod\limits_{j=1}^{i}\dfrac{\mu_j}{\lambda_j}\right)<\infty,\\[2ex] 1, & \text{если}\quad \sum\limits_{i=1}^{\infty}\left(\prod\limits_{j=1}^{i}\dfrac{\mu_j}{\lambda_j}\right)=\infty.\end{cases}\tag{7.9}$$

Среднее время до поглощения равно

$$\begin{cases}\infty, & \text{если}\quad \sum\limits_{i=1}^{\infty}\rho_i=\infty,\\[2ex] \sum\limits_{i=1}^{\infty}\rho_i+\sum\limits_{r=1}^{m-1}\left(\prod\limits_{k=1}^{r}\dfrac{\mu_k}{\lambda_k}\right)\sum\limits_{j=r+1}^{\infty}\rho_j, & \text{если}\quad \sum\limits_{i=1}^{\infty}\rho_i<\infty,\end{cases}\tag{7.10}$$

где $\rho_i=(\lambda_1\lambda_2\ldots\lambda_{i-1})/(\mu_1\mu_2\ldots\mu_i)$.

В случае процесса рождения и гибели с линейным ростом ($\lambda_n=n\lambda$, $\mu_n=n\mu$, $\mu>\lambda$) среднее время ω_1 до поглощения из состояния 1 равно

$$\sum_{i=1}^{\infty}\rho_i=\frac{1}{\mu}\sum_{i=1}^{\infty}\frac{1}{i}\left(\frac{\lambda}{\mu}\right)^{i-1}=\frac{1}{\lambda}\sum_{i=0}^{\infty}\int_0^{\lambda/\mu}\xi^i\,d\xi=$$

$$=\frac{1}{\lambda}\int_0^{\lambda/\mu}\frac{1}{1-\xi}\,d\xi=-\frac{1}{\lambda}\ln\left(1-\frac{\lambda}{\mu}\right).$$

§ 8. ЦЕПИ МАРКОВА С КОНЕЧНЫМ ЧИСЛОМ СОСТОЯНИЙ И НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Марковская цепь с непрерывным временем X_t $(t > 0)$ является марковским процессом с состояниями 0, 1, 2, ... Предположим, как обычно, что переходные вероятности стационарны, т. е.

$$P_{ij}(t) = \mathrm{P}\{X_{t+s} = j \mid X_s = i\}. \tag{8.1}$$

В этом параграфе мы рассмотрим лишь случай, когда пространство состояний S конечно, $S = \{0, 1, 2, \ldots, N\}$. Некоторые проблемы, касающиеся цепей Маркова с бесконечным числом состояний и непрерывным временем, рассмотрены в следующей главе.

Марковское свойство требует, чтобы $P_{ij}(t)$ удовлетворяли условиям:

(а) $P_{ij}(t) \geqslant 0$,

(б) $\sum\limits_{j=0}^{N} P_{ij}(t) = 1, \quad i, j \in S$,

(в) $P_{ik}(s+t) = \sum\limits_{j=0}^{N} P_{ij}(s) P_{jk}(t), \quad t, s \geqslant 0$ (уравнение Колмогорова — Чэпмена),

и мы требуем, кроме того, чтобы выполнялось условие

(г) $\lim\limits_{t\to 0+} P_{ij}(t) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$

Если обозначить через $\mathbf{P}(t)$ матрицу $\| P_{ij}(t) \|_{i,j=0}^{N}$, то условие (в) можно записать более компактно в матричных обозначениях

$$\mathbf{P}(t+s) = \mathbf{P}(t)\,\mathbf{P}(s), \qquad t, s \geqslant 0. \tag{8.2}$$

Условие (г) говорит о непрерывности $\mathbf{P}(t)$ при $t = 0$, поскольку из (8.2) следует, что $\mathbf{P}(0) = \mathbf{I}$ ($\mathbf{I}$ — единичная матрица). Из (8.2) вытекает, что $\mathbf{P}(t)$ непрерывна при всех $t > 0$. В самом деле, если $s = h > 0$ в (8.2), то в силу (г) имеем

$$\lim_{h\to 0+} \mathbf{P}(t+h) = \mathbf{P}(t) \lim_{h\to 0} \mathbf{P}(h) = \mathbf{P}(t)\,\mathbf{I} = \mathbf{P}(t). \tag{8.3}$$

С другой стороны, при $t > 0$ и $0 < h < t$ запишем (8.2) в виде

$$\mathbf{P}(t) = \mathbf{P}(t-h)\,\mathbf{P}(h). \tag{8.4}$$

Но $\mathbf{P}(h)$ при достаточно малых h близка к единичной матрице. Поэтому $(\mathbf{P}(h))^{-1}$ (обратная матрица к $\mathbf{P}(h)$) существует и также близка к $\mathbf{I}$. Следовательно,

$$\mathbf{P}(t) = \mathbf{P}(t) \lim_{h\to 0+} (\mathbf{P}(h))^{-1} = \lim_{h\to 0+} \mathbf{P}(t-h). \tag{8.5}$$

Предельные соотношения (8.3) и (8.5) показывают, что $\mathbf{P}(t)$ непрерывна.

В теоремах 1.1 и 1.2 гл. 8 доказано, что в общем случае для цепей Маркова с бесконечным (счетным) числом состояний и непрерывным временем существуют пределы

$$\begin{aligned} &\lim_{h\to 0+} \frac{1-P_{ii}(h)}{h} = q_i, \\ &\lim_{h\to 0+} \frac{P_{ij}(h)}{h} = q_{ij}, \qquad i \neq j, \end{aligned} \tag{8.6}$$

где $0 \leqslant q_{ij} < \infty$ $(i \neq j)$ и $0 \leqslant q_i \leqslant \infty$, т. е. q_{ij} $(i \neq j)$ всегда конечны, а q_i определены, но могут принимать бесконечные значения. Возможность $q_i = \infty$ не может осуществиться в случае марковской цепи с конечным числом состояний и непрерывным временем. Действительно, запишем соотношение

$$1 - P_{ii}(h) = \sum_{j=0,\, j\neq i}^{N} P_{ij}(h),$$

разделим его на h и устремим $h\downarrow 0$, получим

$$q_i = \sum_{j=0,\, j\neq i}^{N} q_{ij},$$

откуда следует конечность q_i.

Предполагая справедливым (8.6), найдем точное выражение для $P_{ij}(t)$ через инфинитезимальную матрицу

$$\mathbf{A} = \begin{Vmatrix} -q_0 & q_{01} & \cdots & q_{0N} \\ q_{10} & -q_1 & \cdots & q_{1N} \\ \cdot & \cdot & \cdot & \cdot \\ q_{N0} & q_{N1} & \cdots & -q_N \end{Vmatrix}.$$

Предельные соотношения (8.6) запишем в матричном виде

$$\lim_{h\to 0+} \frac{\mathbf{P}(h) - \mathbf{I}}{h} = \mathbf{A}. \tag{8.7}$$

С помощью этой формулы и равенства (8.2) находим

$$\frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} = \frac{\mathbf{P}(t)\,[\mathbf{P}(h) - \mathbf{I}]}{h} = \frac{\mathbf{P}(h) - \mathbf{I}}{h}\,\mathbf{P}(t). \tag{8.8}$$

Предел правой части существует, и, таким образом, получаются матричные дифференциальные уравнения

$$\mathbf{P}'(t) = \mathbf{P}(t)\,\mathbf{A} = \mathbf{A}\mathbf{P}(t), \tag{8.9}$$

где $\mathbf{P}'(t)$ — матрица с элементами $P'_{ij}(t)$.

Существование $P'_{ij}(t)$ является очевидным следствием (8.7) и (8.8).

Уравнения (8.9) можно решать при начальном условии $\mathbf{P}(0)=\mathbf{I}$ стандартными методами теории обыкновенных дифференциальных уравнений [1]). Получим

$$\mathbf{P}(t)=e^{\mathbf{A}t}=\mathbf{I}+\sum_{n=1}^{\infty}\frac{\mathbf{A}^h t^h}{h!}. \tag{8.10}$$

Практически мы находим собственные значения $\lambda_0, \lambda_1, \ldots, \lambda_N$ матрицы $\mathbf{A}$ и полную систему соответствующих им правых собственных векторов $\mathbf{u}^{(0)}, \ldots, \mathbf{u}^{(N)}$, если это возможно (см. приложение в конце данной книги). Затем мы пользуемся представлением

$$\mathbf{P}(t)=\mathbf{U}\Lambda(t)\,\mathbf{U}^{-1}, \tag{8.11}$$

где $\mathbf{U}$ — матрица, столбцами которой являются соответственно векторы $\mathbf{u}^{(0)}, \mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N)}$, а $\Lambda(t)$ — диагональная матрица

$$\Lambda(t)=\begin{Vmatrix} \exp(\lambda_0 t) & 0 & \ldots & 0 \\ 0 & \exp(\lambda_1 t) & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \exp(\lambda_N t) \end{Vmatrix}.$$

Строки матрицы $\mathbf{U}^{-1}$ можно рассматривать как полную систему левых собственных векторов, биортогональных к $\{\mathbf{u}^{(l)}\}_{l=0}^{N}$.

Применения формул (8.10) или (8.11) рассматриваются в задачах 18, 20 и 21 к данной главе.

ЗАДАЧИ

1. Рассмотрим пуассоновский процесс с параметром λ. При условии что за время t произошло n событий, найти плотность вероятности времени осуществления r-го события $(r<n)$.

Ответ:

$$p(x)=\begin{cases}\dfrac{n!}{(r-1)!\,(n-r)!}\dfrac{x^{r-1}}{t^r}\left(1-\dfrac{x}{t}\right)^{n-r}, & 0<x<t,\\ 0 & \text{в противном случае.}\end{cases}$$

2. Предположим, что прибор отказывает после воздействия k возмущений Пусть возмущения случаются в моменты, образующие пуассоновский поток с параметром λ. Найти плотность времени жизни T прибора.

Ответ:

$$f(t)=\begin{cases}\dfrac{\lambda^k t^{k-1}e^{-\lambda t}}{\Gamma(k)}, & t>0,\\ 0, & t\leqslant 0.\end{cases}$$

---

[1]) Коддингтон Э. А., Левинсон Н., Теория обыкновенных дифференциальных уравнений, гл. 3, ИЛ, 1958 (см. также Гантмахер Ф. Р., Теория матриц, гл. V, «Наука», 1967. — *Перев.*).

3. Пусть $(X(t), Y(t))$ — вероятностный процесс в двумерном пространстве, где $X(t)$ — пуассоновский процесс с параметром λ_1, а $Y(t)$ — пуассоновский процесс с параметром λ_2, не зависящий от $X(t)$. При условии что процесс находится в состоянии (x_0, y_0) в момент $t = 0$, $x_0 + y_0 < z$, найти вероятность пересечения процессом прямой $x + y = z$ в точке (x, y).

Ответ:

$$\begin{cases} \binom{z - x_0 - y_0}{x - x_0} \left(\dfrac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{x - x_0} \left(\dfrac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{y - y_0} & \text{при} \quad x \geqslant x_0, \ y \geqslant y_0, \\ 0 & \text{в противном случае.} \end{cases}$$

4. Рассмотрим пуассоновский процесс с параметром λ. Пусть T — время до наступления первого события, а $N(T/k)$ — число событий в следующие T/k единиц времени. Найти первые два момента величины $N\left(\frac{T}{k}\right) T$.

Ответ:

$$\mathrm{M}\left\{N\left(\frac{T}{k}\right) T\right\} = \frac{2}{\lambda k}; \quad \mathrm{M}\left\{\left(N\left(\frac{T}{k}\right) T\right)^2\right\} = \frac{6}{\lambda^2 k} + \frac{24}{\lambda^2 k^2}.$$

5. Пусть $\{X(t), t \geqslant 0\}$ — пуассоновский процесс с параметром λ. Предположим, что каждое событие «регистрируется» с вероятностью p независимо от остальных событий. Пусть $\{Y(t), t \geqslant 0\}$ — процесс, скачки которого происходят лишь в моменты наступления «зарегистрированных» событий. Доказать, что $Y(t)$ — пуассоновский процесс с параметром λp.

6. Рассмотрим n независимо действующих объектов (таких, как электрические лампочки), время безотказной работы (т. е. время жизни) которых является случайной величиной, экспоненциально распределенной с плотностью

$$f(x, \theta) = \begin{cases} \theta^{-1} \exp\left(-\dfrac{x}{\theta}\right), & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

θ — положительный параметр. Действительные длительности реализации безотказной работы становятся известными в порядке отказов объектов. Пусть

$$X_{1n} \leqslant X_{2n} \leqslant \ldots \leqslant X_{rn}$$

— длительности времен жизни первых r отказавших объектов. Найти совместную плотность величин X_{in}, $i = 1, 2, \ldots, r$.

Ответ:

$$f(x_1, x_2, \ldots, x_r) = \frac{n!}{(n-r)!} \frac{1}{\theta^r} \exp\left(-\frac{x_1 + x_2 + \ldots + x_{r-1} + (n - r + 1) x_r}{\theta}\right).$$

7. В предыдущей задаче введем

$$Y_{1n} = X_{1n},$$
$$Y_{in} = X_{in} - X_{i-1, n}, \qquad 2 \leqslant i \leqslant r.$$

Доказать, что Y_{in} независимы в совокупности, и найти функцию распределения каждой из них.

Ответ:

$$\mathrm{P}\{Y_{in} \leqslant y\} = 1 - \exp\left(-\frac{n - i + 1}{\theta} y\right).$$

8. Пусть $\{X(t), t \geqslant 0\}$ и $\{Y(t), t \geqslant 0\}$ — два независимых пуассоновских процесса с параметрами λ_1 и λ_2 соответственно. Введем процесс

$$Z(t) = X(t) - Y(t), \qquad t \geqslant 0.$$

Пространство состояний этого процесса состоит из всех целых чисел (положительных, отрицательных и нуля). Пусть

$$P_n(t) = \mathrm{P}\{Z(t) = n\}, \qquad n = 0, \pm 1, \pm 2, \ldots .$$

Доказать формулу

$$\sum_{n=-\infty}^{\infty} P_n(t)\, z^n = \exp(-(\lambda_1 + \lambda_2)\, t) \exp\left(\lambda_1 z t + \frac{\lambda_2}{z}\, t\right), \qquad |z| \neq 0.$$

Найти $\mathrm{M}(Z(t))$ и $\mathrm{M}(Z^2(t))$.

Ответ:

$$\mathrm{M}(Z^2(t)) = (\lambda_1 + \lambda_2)\, t + (\lambda_1 - \lambda_2)^2 t^2.$$

9. Рассмотрим два независимых пуассоновских процесса $X(t)$ и $Y(t)$, где $\mathrm{M}(X(t)) = \lambda t$, а $\mathrm{M}(Y(t)) = \mu t$. Пусть два последовательных события процесса $X(t)$ поисходят в моменты T и $T' > T$, так что $X(t) = X(T)$ при $T - t < T'$ и $X(T') = X(T) + 1$. Пусть $N = Y(T') - Y(T)$ — число событий процесса $Y(t)$, происшедших на интервале (T, T'). Найти распределение N.

Ответ:

$$\mathrm{P}\{N = m\} = \frac{\lambda}{\lambda + \mu}\left(\frac{\mu}{\lambda + \mu}\right)^m, \qquad m = 0, 1, 2, \ldots .$$

10. Пусть $X(t)$ — марковский процесс чистого рождения с непрерывным временем. Предположим, что

$$\mathrm{P}\{\text{событие наступит в } (t, t+h) \mid X(t) \text{ нечетно}\} = \lambda_1 h + o(h),$$

$$\mathrm{P}\{\text{событие наступит в } (t, t+h) \mid X(t) \text{ четно}\} = \lambda_2 h + o(h),$$

где $\dfrac{o(h)}{h} \to 0$ при $h \downarrow 0$. Примем $X(0) = 0$. Найти следующие вероятности:

$$P_1(t) = \mathrm{P}\{X(t) \text{ нечетно}\}, \qquad P_2(t) = \mathrm{P}\{X(t) \text{ четно}\}.$$

Указание: Вывести дифференциальные уравнения

$$P_1'(t) = -\lambda_1 P_1(t) + \lambda_2 P_2(t),$$

$$P_2'(t) = \lambda_1 P_1(t) - \lambda_2 P_2(t)$$

и решить их.

Ответ:

$$P_1(t) = \frac{\lambda_2}{\lambda_1 + \lambda_2}\left(1 - \exp\{-(\lambda_1 + \lambda_2)\, t\}\right);$$

$$P_2(t) = \frac{\lambda_1}{\lambda_1 + \lambda_2} + \frac{\lambda_2}{\lambda_1 + \lambda_2} \exp\{-(\lambda_1 + \lambda_2)\, t\}.$$

11. В условиях задачи 10 найти $\mathrm{M}(X(t))$.

Ответ:

$$\mathrm{M}(X(t)) = \frac{2\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}\, t + \frac{(\lambda_1 - \lambda_2)\, \lambda_2}{(\lambda_1 + \lambda_2)^2}\left[\exp\{-(\lambda_1 + \lambda_2)\, t\} - 1\right].$$

12. Пусть $g(t)$ — условная интенсивность исчезновения частицы в момент t при условии, что она не исчезла до t, т. е. $\mathrm{P}\{$исчезновение произошло в $(t, t+h) \mid$ исчезновения не было до момента $t\} = g(t)h + o(h)$ при $h \downarrow 0$. Предположим, что $g(t)$ положительна и непрерывна на $(0, \infty)$. Найти выражение для $F(t) = \mathrm{P}\{$исчезновение произошло в некоторый момент $\tau < t\}$ через $g(\cdot)$.

Указание: Вывести дифференциальное уравнение для $F(t)$.

Ответ: $F(t) = 1 - \exp\left[-\int\limits_0^t g(t)\,d\tau\right]$.

13. Рассмотрим пуассоновский процесс с переменной интенсивностью, т. е. вероятность осуществления события E в интервале $(t, t+h)$ не зависит от предыстории процесса и равна $\lambda(t)h + o(h)$ $(h \to 0)$. (Заметим, что λ может теперь зависеть от t.)

(а) Доказать, что вероятность неосуществления события E на отрезке $[0, s]$ равна

$$\exp\left(-\int\limits_0^s \lambda(\xi)\,d\xi\right).$$

(б) Доказать, что вероятность наступления k событий на отрезке $[0, s]$ равна

$$\frac{1}{k!}\left(\int\limits_0^s \lambda(\xi)\,d\xi\right)^k \exp\left(-\int\limits_0^s \lambda(\xi)\,d\xi\right).$$

14. Рассмотрим процесс чистой гибели, в котором $\mu_n = n\mu$, $n = 1, 2, \ldots$, т. е. $\mathrm{P}\{X(t+h) = j \mid X(t) = k\} = 0$ при $j > k$, $t > 0$, $k > 0$. Предположим, что начальный размер популяции равен i. Найти $P_n(t) = \mathrm{P}\{X(t) = n\}$, $M(X(t))$ и $\sigma^2(X(t))$.
Ответ:

$$P_n(t) = \binom{i}{n} e^{-(i-n)\mu t}\left(1 - e^{-\mu t}\right)^n,$$
$$M(X(t)) = ie^{-\mu t},$$
$$\sigma^2(X(t)) = ie^{-\mu t}\left(1 - e^{-\mu t}\right).$$

15. Рассмотрим процесс Юла с параметром β и начальным состоянием $N = 1$. Предположим, что первый индивидуум может погибнуть, причем вероятность гибели в интервале $(t, t+h)$ при условии, что индивидуум не погиб до момента t, равна $\mu h + o(h)$. Найти распределение общего числа потомков всех поколений единственного индивидуума в момент гибели «родоначальника».

Ответ: Вероятность того, что общее число потомков всех поколений в момент гибели «родоначальника» равна n, выражается формулой

$$\int\limits_0^\infty e^{-\beta t}\left(1 - e^{-\beta t}\right)^n \mu e^{-\mu t}\,dt = \frac{\mu}{\beta}\,\frac{\Gamma\left(\frac{\mu}{\beta} + 1\right)\Gamma(n+1)}{\Gamma\left(n + \frac{\mu}{\beta} + 2\right)}.$$

16. Телефонный узел имеет m каналов. Моменты поступления вызовов образуют пуассоновский поток с параметром λ. Они принимаются (обслуживаются), если имеется свободный канал. В противном случае они теряются. Продолжительность каждого разговора — д. с. в. с экспоненциальным распределением, имеющим параметр μ. Длительности отдельных разговоров — независимые случайные величины. Найти стационарное распределение числа занятых каналов.
Ответ:

$$p_n = \frac{\left(\frac{\lambda}{\mu}\right)^n \frac{1}{n!}}{\sum\limits_{k=0}^m \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}}, \qquad n = 0, 1, 2, \ldots, m.$$

17. Найти стационарное распределение для процесса рождения и гибели с линейным ростом при $\lambda < \mu$ (см. пример 1 § 6).

Ответ:

$$p_n = \left(\frac{\lambda}{\mu}\right)^n \frac{\left(\frac{a}{\lambda}\right)\left(\frac{a}{\lambda}+1\right)\ldots\left(\frac{a}{\lambda}+n-1\right)}{n!} \frac{1}{\left(1-\frac{\lambda}{\mu}\right)^{\frac{a}{\lambda}}}.$$

18. Цепь Маркова с непрерывным временем имеет два состояния: 0 и 1. Время пребывания в состоянии 0 распределено экспоненциально с параметром $\lambda > 0$. Время пребывания в состоянии 1 распределено экспоненциально с параметром $\mu > 0$. Найти вероятность $P_{00}(t)$ нахождения процесса в состоянии 0 в момент t, если в момент 0 он находится в состоянии 0.

Ответ:

$$P_{00}(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}.$$

19. Пусть в задаче 18 $\lambda = \mu$, а $N(t)$ — число моментов изменения состояний процесса за время $t \geqslant 0$. Найти вероятностное распределение $N(t)$.

Ответ:

$$\mathrm{P}\{N(t) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

20. Имеется два трансатлантических кабеля, каждый из которых может передавать одновременно только одно телеграфное сообщение. Время до отказа каждого из них имеет одно и то же экспоненциальное распределение с параметром λ. Время ремонта каждого кабеля имеет одно и то же экспоненциальное распределение с параметром μ. При условии, что в момент 0 оба кабеля находятся в рабочем состоянии, найти вероятность того, что если в момент t одновременно поступают два сообщения, то они найдут оба кабеля исправными.

Указание: Рассматриваемый процесс является цепью Маркова с тремя состояниями и непрерывным временем.

Ответ:

$$\frac{\mu^2}{(\lambda+\mu)^2} + \frac{\lambda^2 e^{-2(\lambda+\mu)t}}{(\lambda+\mu)^2} + \frac{2\lambda\mu}{(\lambda+\mu)^2} e^{-(\lambda+\mu)t}.$$

21. Предположим, что некоторый механизм подвержен отказам двух типов. Пусть вероятность отказа первого типа в интервале $(t, t+h)$ равна $\lambda_1 h + o(h)$, а вероятность отказа второго типа в том же интервале равна $\lambda_2 h + o(h)$. В состоянии отказа производится ремонт, длительность которого экспоненциально распределена с параметром, зависящим от типа отказа. Пусть μ_1 и μ_2 — значения этих параметров. Найти вероятность того, что механизм работает в момент t.

Ответ:

$$\mathbf{P}(t) = e^{\mathbf{Q}t}, \quad \text{где} \quad \mathbf{Q} = \begin{Vmatrix} -(\lambda_1+\lambda_2) & \lambda_1 & \lambda_2 \\ \mu_1 & -\mu_1 & 0 \\ \mu_2 & 0 & -\mu_2 \end{Vmatrix}.$$

22. Система состоит из N идентичных компонент, каждая из которых работает независимо от других случайное время до отказа. Предположим, что время до отказа имеет экспоненциальное распределение с параметром λ. Когда какая-либо компонента отказывает, она ремонтируется. Время ремонта случайно с экспоненциальной функцией распределения, имеющей параметр μ. Говорят, что система находится в состоянии n в момент t, если в этот момент ровно n компонент находится в ремонте. Этот процесс является процессом рождения и гибели. Найти его инфинитезимальные параметры.

23. В задаче 22 предположим, что вначале все N компонент работают. Найти распределение $F(t)$ до первого момента, когда имеются две неработающие компоненты.

Ответ: Преобразование Лапласа $\varphi(s)$ функции $F(t)$ равно

$$\varphi(s)=\frac{N(N-1)\lambda^2}{s^2+s[(2N-1)\lambda+\mu]+N(N-1)\lambda^2}.$$

В случае $\lambda=\mu$

$$1-F(t)=\frac{\sqrt{N}(N-1)}{2}\{\exp[(-N+\sqrt{N})\lambda t]-\exp[(-N-\sqrt{N})\lambda t]\}.$$

24. Рассмотрим следующий непрерывный аналог модели Эренфестов (см. гл. 2, стр. 43). Имеется $2N$ шаров, занумерованных от 1 до $2N$. В момент 0 каждый шар с равной вероятностью может быть помещен в одну из двух урн. В дальнейшем шары независимым образом перемещаются в случайные моменты времени из одной урны в другую в соответствии со следующими правилами. Шар с вероятностью $\frac{1}{2}h+o(h)$ попадает из одной урны в другую в течение интервала $(t, t+h)$ и с вероятностью $1-\frac{h}{2}+o(h)$ остается в той же урне в этом интервале. Перемещения на непересекающихся интервалах времени независимы. Пусть $X(t)$ — число шаров в урне 1 в момент t. Обозначим

$$P_{jk}(t)=\mathrm{P}\{X(t)=k\mid X(0)=j\},\qquad j,\ k=0,\ 1,\ \ldots,\ 2N.$$

Доказать формулу

$$g(t,s)=\sum_{k=0}^{2N}P_{jk}(t)s^k=2^{-2N}\left[1-e^{-t}+\left(1+e^{-t}\right)s\right]^j\left[1+e^{-t}+\left(1-e^{-t}\right)s\right]^{2N-j}.$$

Указание: Ввести случайные величины

$$X_i(t)=\begin{cases}1, & \text{если } i\text{-й шар находится в урне 1 в момент } t,\\ 0 & \text{в противном случае.}\end{cases}$$

Тогда $X(t)=\sum\limits_{i=1}^{2N}X_i(t)$. Показать, что

$$\mathrm{P}\{X_i(t)=X_i(0)\}=\frac{1+e^{-t}}{2},\qquad i=1,\ 2,\ \ldots,\ 2N.$$

25. Рассмотрим процесс $X(t)$ рождения и гибели с линейным ростом с параметрами λ, μ и $a=0$. Предположим, что $X(0)=1$. Найти распределение числа живущих индивидуумов в момент первой гибели.

Ответ:

$$\mathrm{P}\{\text{число рождений до первой гибели равно } k\}=\frac{\lambda}{\mu+\lambda}\left(\frac{\lambda}{\mu+\lambda}\right)^k.$$

26. Для процесса $X(t)$ рождения и гибели с линейным ростом при $\lambda=\mu$ (см. пример 1 § 6) доказать, что

$$u(t)=\mathrm{P}\{X(t)=0\mid X(0)=1\}$$

удовлетворяет интегральному уравнению

$$u(t)=\frac{1}{2}\int_0^t 2\lambda e^{-2\lambda\tau}\,d\tau+\frac{1}{2}\int_0^t 2\lambda e^{-2\lambda\tau}\,[u(t-\tau)]^2\,d\tau.$$

Указание: Заметить, что время до наступления первого события (рождения или гибели) распределено экспоненциально с параметром 2λ.

27 (продолжение). Показать, что $u(t)$ удовлетворяет дифференциальному уравнению Рикатти:

$$u'(t)+2\lambda u(t)=\lambda+\lambda u^2(t), \qquad u(0)=0.$$

28 (продолжение). Найти $u(t)$.
Ответ:

$$u(t)=\frac{\lambda t}{1+\lambda t}.$$

29 (продолжение). Найти $\mathrm{P}\{X(t)=0\mid X(0)=1, X(T)=0\}$ при $0<t<T$.

30. Рассмотрим процесс рождения и гибели с инфинитезимальными параметрами λ_n, μ_n. Показать, что среднее время достижения состояния $r+1$ из состояния 0 равно

$$\sum_{n=0}^{r}\pi_n\sum_{k=0}^{n}\frac{1}{\lambda_k\pi_k}$$

где π_n определяется в (4.5).

Указание: Пусть T_n^* — время первого достижения состояния $n+1$ из состояния n. Вывести рекуррентное соотношение для $\mathrm{M}\left(T_n^*\right)$.

31. Начнем наблюдать радиоактивный атом в момент 0. Он распадается и перестает быть радиоактивным в момент t, $t>0$, определяемый распределением

$$F(\tau)=\begin{Bmatrix} 0, & \tau<0, \\ 1-e^{-\lambda\tau}, & \tau\geqslant 0 \end{Bmatrix}=\mathrm{P}\{t\leqslant\tau\}.$$

Рассмотрим состояние атома в момент t как случайную величину

$$x_t=\begin{cases} 0, & \text{если атом радиоактивен в момент } t, \\ 1, & \text{если атом не радиоактивен в момент } t. \end{cases}$$

Переменная $\{x_t\}$ образует случайный процесс.

Предположим, что в момент 0 мы начинаем наблюдать N независимых радиоактивных атомов, которые описываются, как показано выше, величинами x_t^i, $i=1, 2, \ldots, N$. Пусть $X_t=\sum_{i=1}^{N} x_t^i$. Тогда $\{X_t\}$ — также случайный процесс. Показать, что при $t\ll\frac{1}{\lambda}$ $\left(t\text{ мало по сравнению с }\frac{1}{\lambda}\right)$ и достаточно большом N процесс $\{X_t\}$ хорошо аппроксимируется пуассоновским процессом $Y(t)$ с параметром λNt.

32. Предположим, что в задаче 31 нельзя считать $t \ll \lambda^{-1}$.

(1) Является ли процесс процессом с независимыми приращениями?
(2) Является ли он стационарным?
(3) Имеет ли он стационарные переходные вероятности?
(4) Является ли он марковским?

Ответ: (1) да, (2) нет, (3) да, (4) да.

33. Пусть $\mathfrak{M}$ — процесс рождения и гибели с непрерывным временем, где $\lambda_n = \lambda > 0$, $n \geqslant 0$, $\mu_0 = 0$, $\mu_n > 0$, $n \geqslant 1$. Пусть $\pi = \sum\limits_n \pi_n < \infty$, где $\pi_n = \lambda^n (\mu_1, \mu_2 \ldots \mu_n)^{-1}$, так что $\dfrac{\pi_i}{\pi}$ — стационарное распределение процесса. Предположим, что начальное состояние — д. с. в, распределение которой совпадает со стационарным распределением процесса. Доказать, что число случаев гибели на отрезке $[0, t]$ имеет распределение Пуассона с параметром λt.

Указание: Пусть $a_k(t)$ — вероятность того, что число случаев гибели за время t равно k. Получить дифференциальное уравнение

$$a_k'(t) = -\lambda a_k(t) + \lambda a_{k-1}(t), \quad k \geqslant 1.$$

34. Следующее построение является одним из способов задания многомерного пуассоновского процесса для случая двух измерений. Пусть $(X(t), Y(t))$ — пара, где $X(t) = \alpha(t) + \gamma(t)$, $Y(t) = \beta(t) + \gamma(t)$, $\alpha(t)$, $\beta(t)$, $\gamma(t)$ — три независимых пуассоновских процесса с параметрами λ_1, λ_2 и λ_3 соответственно. Найти производящую функцию распределения $(X(t), Y(t))$.

Ответ:

$$\sum_{i,j} \mathrm{P}\{X(t) = i,\ Y(t) = j\} x^i y^j = \exp\{t(\lambda_1 x + \lambda_2 y + \lambda_3 xy - \lambda_1 - \lambda_2 - \lambda_3)\}.$$

35. Рассмотрим процесс Юла $\{N_t,\ t \geqslant 0\}$ с интенсивностью рождения λ и начальным размером популяции 1. Найти функцию распределения величины $N_t(x)$, равной числу членов популяции возраста не большего x в момент t.

Ответ:

$$\mathrm{P}\{N_t(x) = k\} = \frac{e^{-\lambda t}\left(1 - e^{-\lambda x}\right)^k}{\left[1 - e^{-\lambda x} + e^{-\lambda t}\right]^{k+1}}.$$

36. Пусть $\{X_i(t),\ t \geqslant 0\}$, $i = 1, 2$, — два независимых процесса Юла с одним и тем же параметром λ. Пусть $X_i(0) = n_i$, $i = 1, 2$. Найти распределение величины $X_1(t)$ при условии, что $X_1(t) + X_2(t) = N$ $(N \geqslant n_1 + n_2)$.

Ответ:

$$\mathrm{P}\{X_1(t) = k \mid X_1(t) + X_2(t) = N\} = \frac{\dbinom{k-1}{n_1-1}\dbinom{N-k-1}{n_2-1}}{\dbinom{N-1}{n_1+n_2-1}}.$$

37 (продолжение). Доказать предельное соотношение

$$\lim_{t\to\infty} \mathrm{P}\left\{\frac{X_1(t)}{X_1(t) + X_2(t)} \leqslant x\right\} = \frac{(n_1+n_2-1)!}{(n_1-1)!\,(n_2-1)!} \int_0^x y^{n_1-1}(1-y)^{n_2-1}\,dy.$$

Указание: Пусть $N \to \infty$ и $k \to \infty$, так что $\frac{k}{N} \to y$ $(0 < y < 1)$. Тогда с помощью формулы Стирлинга установить предельное соотношение

$$\lim_{\substack{\frac{k}{N} \to y,\ N \to \infty \\ k \to \infty}} \frac{\binom{k-1}{n_1-1}\binom{N-k-1}{n_2-1}}{\binom{N-1}{n_1+n_2-1}} = \frac{(n_1+n_2-1)!}{(n_1-1)!\,(n_2-1)!}\, y^{n_1-1}\,(1-y)^{n_2-1}.$$

Использовать его для доказательства равенства

$$\lim_{t\to\infty} \mathrm{P}\left\{ y \leqslant \frac{X_1(t)}{X_1(t)+X_2(t)} \leqslant y+h \right\} = \frac{(n_2+n_1-1)!}{(n_1-1)!\,(n_2-1)!}\, h y^{n_1-1}\,(1-y)^{n_2-1} + o(h).$$

НЕКОТОРЫЕ ЭЛЕМЕНТАРНЫЕ ЗАДАЧИ

1. Пусть $\{X(t),\ t \geqslant 0\}$ и $\{Y(t),\ t \geqslant 0\}$ — независимые пуассоновские процессы с параметрами λ_1 и λ_2 соответственно. Пусть $Z_1(t) = X(t) + Y(t)$, $Z_2(t) = X(t) - Y(t)$, $Z_3(t) = X(t) + k$, где k — положительное целое число. Определить, который из перечисленных выше процессов является пуассоновским, и найти его параметр.

2. Телеграммы поступают на телеграф в соответствии с пуассоновским процессом, в среднем 3 телеграммы за час.

(а) Какова вероятность того, что за утро (от 8 до 12) не поступит ни одной телеграммы?

(б) Каково распределение момента поступления первой дневной телеграммы?

3. Пусть $X(t)$ — однородный пуассоновский процесс с параметром λ. Найти ковариацию между $X(t)$ и $X(t+\tau)$, $t, \tau > 0$, т. е. определить $\mathrm{M}[(X(t) - \mathrm{M}(X(t)))(X(t+\tau) - \mathrm{M}(X(t+\tau)))]$.

4. В молекулярной биологии возникает следующая задача. Предполагается, что поверхность бактерии содержит несколько точек, где могут закрепляться молекулы, пришедшие извне, если они имеют правильное строение. Молекулы, имеющие такое строение, будем называть допусгимыми. Рассмотрим фиксированную точку и примем, что молекулы прибывают в эту точку в соответствии с пуассоновским процессом, имеющим параметр μ. Долю β этих молекул составляют допустимые. Недопустимые молекулы остаются в точке экспоненциально (с параметром λ) распределенное время. Пока они находятся в рассматриваемой точке, они препятствуют другим молекулам. Допустимые молекулы занимают это место навечно, также препятствуя в этом другим молекулам. Какова вероятность того, что рассматриваемая точка свободна в момент t?

Указание: Свести задачу к изучению цепи Маркова с непрерывным временем и 3 состояниями.

Ответ:

$$\frac{\beta\mu}{s_2 - s_1}\left[\left(1 + \frac{\lambda}{s_1}\right)e^{s_1 t} - \left(1 + \frac{\lambda}{s_2}\right)e^{s_2 t}\right],$$

где s_1, s_2 — корни уравнения $s^2 + s(\lambda + \mu) + \mu\beta\lambda = 0$.

ЗАМЕЧАНИЯ

Пуассоновские процессы и процессы рождения и гибели играют фундаментальную роль в теории и приложениях, охватывающих модели создания запасов и массового обслуживания, рост популяций, технические системы и т. д. Элементарные обсуждения пуассоновских и родственных им процессов можно найти в любом учебнике по случайным процессам.

ЛИТЕРАТУРА

1. Феллер В., Введение в теорию вероятностей и ее приложения, т. 1, «Мир», М., 1967.
2. Bailey N. T., The Elements of Stochastic Processes with Applications to the Natural Sciences, Wiley, New York, 1964.
3. Blanc-Lapierre A., Fortet R., Théorie des Fonctions Alétoire, Masson, Paris, 1953.

Глава 8

ЦЕПИ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Цель этой главы состоит в том, чтобы познакомить читателя с некоторыми из наиболее изученных вопросов и понятий, возникающих при исследовании цепей Маркова с непрерывным параметром (временем). Как и прежде, мы будем рассматривать лишь однородный случай.

§ 1. СВОЙСТВА ДИФФЕРЕНЦИРУЕМОСТИ ПЕРЕХОДНЫХ ВЕРОЯТНОСТЕЙ

Пусть X_t — марковский процесс с дискретным множеством состояний и непрерывным временем, матрица переходных вероятностей которого $\| P_{ij}(t) \|_{i,\, j=0}^{\infty}$. Таким образом,

$$P\{X(t+s)=j \mid X(s)=i\}=P_{ij}(t).$$

Кроме обычных ограничений, накладываемых на переходную матрицу $\| P_{ij}(t) \|$, т. е.

$$\text{(а)}\quad P_{ij}(t) \geqslant 0, \qquad t>0,$$

$$\text{(б)}\quad \sum_{j} P_{ij}(t)=1, \qquad t>0,$$

$$\text{(в)}\quad \sum_{k} P_{ik}(t)\, P_{kj}(h)=P_{ij}(t+h), \qquad t,\ h>0,$$

предположим, что $P_{ij}(t)$ непрерывны при $t>0$ и что

$$\text{(г)}\quad \lim_{t \to 0} P_{ij}(t)=\begin{cases} 1, & \text{если} \quad i=j, \\ 0, & \text{если} \quad i \neq j \end{cases}$$

(см. также задачу 3). Такую переходную матрицу часто называют «стандартной». Оказывается, что из условий (а)—(г) можно получить гораздо больше следствий, чем можно было бы ожидать. Одним из таких результатов является дифференцируемость $P_{ij}(t)$ при всех $t \geqslant 0$. Мы докажем лишь гораздо более простое утверждение, что функции $P_{ij}(t)$ дифференцируемы (т. е. имеют правосторонние производные) при $t=0$. Рассмотрим сначала $P_{ii}(t)$.

Теорема 1.1. *При каждом i предел*

$$-P'_{ii}(0)=\lim_{t\to 0}\frac{1-P_{ii}(t)}{t}$$

существует, но может быть бесконечным.

Доказательство. Покажем сначала, что $P_{ii}(t)>0$ при всех $t>0$. Действительно, в силу (г)[1]) для любого i существует число $\varepsilon>0$, такое, что $P_{ii}(t)>0$ при $0\leqslant t\leqslant\varepsilon$. Далее, путем последовательного применения (в) можно получить

$$P_{ij}(t_1+\ldots t_n)=\sum_{k_1,\ldots,k_n}P_{ik_1}(t_1)P_{k_1k_2}(t_2)\ldots P_{k_{n-1},j}(t_n). \quad (1.1)$$

Полагая $t_1=\ldots=t_n=\frac{t}{n}$, $i=j$ и беря в правой части лишь члены, соответствующие $k_1=k_2=\ldots=k_{n-1}=i$, получим

$$P_{ii}(t)\geqslant\left[P_{ii}\left(\frac{t}{n}\right)\right]^n. \quad (1.2)$$

При достаточно больших n, очевидно, $\frac{t}{n}\leqslant\varepsilon$. Следовательно, $P_{ii}\left(\frac{t}{n}\right)>0$, таким образом, $P_{ii}(t)>0$. Пусть $-\ln P_{ii}(t)=\varphi(t)$. Это определение корректно, поскольку $P_{ii}(t)>0$. Так же как и (1.2), можно доказать справедливость неравенства

$$P_{ii}(t+s)\geqslant P_{ii}(t)P_{ii}(s).$$

Беря логарифм от обеих частей, получаем неравенство полуаддитивности для $\varphi(t)$:

$$\varphi(t+s)\leqslant\varphi(t)+\varphi(s).$$

Поскольку $0<P_{ii}(t)\leqslant 1$, то $\varphi(t)\geqslant 0$. Положим

$$q_i=\sup_{t>0}\frac{\varphi(t)}{t}.$$

Тогда $0\leqslant q_i\leqslant\infty$, так как $\varphi(t)\geqslant 0$ при $t>0$. Если $q_i<\infty$, то существует $t_0>0$, для которого $\frac{\varphi(t_0)}{t_0}\geqslant q_i-\varepsilon$. Для любого t величину t_0 можно представить в виде $t_0=nt+\delta$, где $n\geqslant 0$, $0\leqslant\delta<t$. Тогда

$$\varphi(t_0)\leqslant\varphi(nt)+\varphi(\delta)\leqslant\varphi((n-1)t)+\varphi(t)+\varphi(\delta)\leqslant\ldots$$
$$\ldots\leqslant n\varphi(t)+\varphi(\delta)$$

и, таким образом,

$$q_i-\varepsilon\leqslant\frac{\varphi(t_0)}{t_0}\leqslant\frac{n\varphi(t)+\varphi(\delta)}{t_0}=\frac{nt}{t_0}\,\frac{\varphi(t)}{t}+\frac{\varphi(\delta)}{t_0}.$$

---

[1]) И непрерывности $P_{ii}(t)$, $t>0$. — *Прим. перев.*

Следовательно,

$$q_i - \varepsilon \leqslant \varliminf_{t\to 0}\left[\frac{nt}{t_0}\frac{\varphi(t)}{t} + \frac{\varphi(\delta)}{t_0}\right].$$

Но при $t \to 0$ имеем $\frac{nt}{t_0} \to 1$ и $\varphi(\delta) \to 0$ (поскольку $P_{ii}(\delta) \to 1$ при $\delta \to 0$). Отсюда

$$\varliminf_{t\to 0}\left[\frac{nt}{t_0}\frac{\varphi(t)}{t} + \frac{\varphi(\delta)}{t_0}\right] = \varliminf_{t\to 0}\frac{\varphi(t)}{t}.$$

Далее, из определения q_i

$$\varlimsup_{t\to 0}\frac{\varphi(t)}{t} \leqslant q_i.$$

Комбинируя последние три неравенства, получаем

$$q_i - \varepsilon \leqslant \varliminf_{t\to 0}\frac{\varphi(t)}{t} \leqslant \varlimsup_{t\to 0}\frac{\varphi(t)}{t} \leqslant q_i.$$

Поскольку ε произвольно, имеем

$$\varliminf_{t\to 0}\frac{\varphi(t)}{t} = \varlimsup_{t\to 0}\frac{\varphi(t)}{t} = q_i.$$

Если $q_i = \infty$, то можно вместо $q_i - \varepsilon$ написать сколь угодно большое число M и затем получить, что $M \leqslant \varliminf_{t\to 0}\frac{\varphi(t)}{t}$. Таким образом, $\infty = \varliminf_{t\to 0}\frac{\varphi(t)}{t}$. В любом случае

$$\lim_{t\to 0}\frac{\varphi(t)}{t} = q_i.$$

Далее,

$$\lim_{t\to 0}\frac{1 - P_{ii}(t)}{t} = \lim_{t\to 0}\frac{1 - e^{-\varphi(t)}}{\varphi(t)}\frac{\varphi(t)}{t} = q_i. \blacksquare$$

Теорема 1.2. *Для любых i и j, $i \neq j$, предел*

$$P'_{ij}(0) = \lim_{t\to 0}\frac{P_{ij}(t)}{t}$$

существует и конечен.

Доказательство. Для любого фиксированного $h > 0$ $\|P_{ij}(h)\|$ является матрицей переходных вероятностей цепи Маркова $\{X_{nh}\}$. Очевидно, $P^n_{ij}(h) = P_{ij}(nh)$. Введем вероятности ${}_jP^0_{ii}(h) = 1$ и

$${}_jP^n_{ii}(h) = \mathrm{P}\{X_{nh} = i;\ X_{vh} \neq j;\ 1 \leqslant v < n \mid X_0 = i\},$$
$$f^n_{ij}(h) = \mathrm{P}\{X_{nh} = j;\ X_{vh} \neq j,\ 1 \leqslant v < n \mid X_0 = i\}.$$

Тогда

$$P_{ij}(nh) \geqslant \sum_{v=0}^{n-1} {}_jP_{ii}^v(h)\, P_{ij}(h)\, P_{jj}((n-v-1)\,h), \tag{1.3}$$

поскольку каждое слагаемое в правой части соответствует некоторому возможному пути, ведущему из состояния i в состояние j за n шагов (относительно шага длины h); эти пути несовместны, но, вообще говоря, не исчерпывают всех путей. Член ${}_jP_{ii}^v(h)\,P_{ij}(h)$ является вероятностью события, заключающегося в том, что последнее попадание в i перед попаданием в j происходит на v-м шаге. (Соотношение (1.3) также появлялось при обсуждении теорем для отношений в гл. 5.)

Далее, аналогичным образом получаем

$$P_{ii}(vh) = {}_jP_{ii}^v(h) + \sum_{m=1}^{v-1} f_{ij}^m(h)\, P_{ji}((v-m)\,h).$$

Первое слагаемое есть вероятность достижения состояния i на v-м шаге без попадания до этого в состояние j. Члены, стоящие под знаком суммы, учитывают достижение состояния j в некоторый промежуточный момент. Поскольку

$$\sum_{m=1}^{v-1} f_{ij}^m(h) \leqslant 1,$$

то

$${}_jP_{ii}^v(h) \geqslant P_{ii}(vh) - \max_{1 \leqslant m < v} P_{ji}((v-m)\,h). \tag{1.4}$$

Далее, из условия (г) [1]) следует, что для любого $\varepsilon > 0$ и любых фиксированных i, j $(i \neq j)$ существует число t_0, такое, что

$$\max_{0 \leqslant t \leqslant t_0} P_{ji}(t) < \varepsilon, \qquad \min_{0 \leqslant t \leqslant t_0} P_{ii}(t) > 1-\varepsilon, \qquad \min_{0 \leqslant t \leqslant t_0} P_{jj}(t) > 1-\varepsilon.$$

Следовательно, если $nh < t_0$ и $v \leqslant n$, то из (1.4) получаем

$${}_jP_{ii}^v(h) > 1 - 2\varepsilon.$$

Подставляя эту оценку в (1.3), находим

$$P_{ij}(nh) \geqslant (1-2\varepsilon) \sum_{v=0}^{n-1} P_{ij}(h)(1-\varepsilon) \geqslant (1-3\varepsilon)\, nP_{ij}(h),$$

или

$$\frac{P_{ij}(nh)}{nh} \geqslant (1-3\varepsilon)\, \frac{P_{ij}(h)}{h}, \quad \text{если} \quad nh < t_0. \tag{1.5}$$

---

[1]) И непрерывности $P_{ij}(t)$. — *Прим. перев.*

Положим

$$q_{ij} = \varlimsup_{t \to 0} \frac{P_{ij}(t)}{t}.$$

Тогда из (1.5) следует, что $q_{ij} < \infty$. Действительно, если бы $q_{ij} = \infty$, то можно было бы найти сколько угодно малое h, для которого $\frac{P_{ij}(h)}{h}$ сколь угодно велико. Выбирая n' так, чтобы $\frac{t_0}{2} < n'h < t_0$, из (1.5) мы бы получили, что $\frac{P_{ij}(n'h)}{n'h}$ можно сделать сколь угодно большим, но в то же самое время

$$\frac{P_{ij}(n'h)}{n'h} < \frac{\varepsilon}{n'h} < \frac{2\varepsilon}{t_0}.$$

Полученное противоречие доказывает, что $q_{ij} < \infty$. Оставшаяся часть доказательства носит чисто аналитический характер и является следствием из (1.5). В силу определения q_{ij} существует $t_1 < t_0$, такое, что

$$\frac{P_{ij}(t_1)}{t_1} < q_{ij} + \varepsilon.$$

Поскольку $P_{ij}(t)$ непрерывна, можно найти настолько малое h_0, что

$$\frac{P_{ij}(t)}{t} < q_{ij} + \varepsilon \quad \text{при} \quad t \in I = [t_1 - h_0,\ t_1 + h_0]. \tag{1.6}$$

Далее, для любого $h < h_0$ определим целое число n_h, такое, что $n_h h \in I$. Тогда, используя (1.5) и (1.6), найдем

$$(1 - 3\varepsilon)\frac{P_{ij}(h)}{h} \leqslant \frac{P_{ij}(n_h h)}{n_h h} < q_{ij} + \varepsilon, \qquad h < h_0,$$

откуда заключаем, что

$$(1 - 3\varepsilon)\varlimsup_{h \to 0} \frac{P_{ij}(h)}{h} \leqslant q_{ij} + \varepsilon.$$

В силу произвольности ε получаем

$$\varlimsup_{h \to 0} \frac{P_{ij}(h)}{h} \leqslant q_{ij}.$$

Утверждение теоремы следует теперь из определения q_{ij}. ■

Если взять в качестве примера процессы рождения и гибели, то

$$q_i = \lambda_i + \mu_i, \quad q_{ij} = \begin{cases} \lambda_i, & j = i + 1, \\ 0, & j \neq i - 1,\ i + 1,\ i, \qquad i = 0,\ 1,\ \ldots, \\ \mu_i, & j = i - 1. \end{cases}$$

В общем случае

$$\sum_{j \neq i} q_{ij} \leqslant q_i \quad \text{при всех } i.$$

Действительно, поскольку

$$\sum_{j \neq i} P_{ij}(h) = 1 - P_{ii}(h),$$

то для любого конечного N

$$\sum_{j=1,\ j \neq i}^{N} P_{ij}(h) \leqslant 1 - P_{ii}(h).$$

Деля на h и устремляя $h \to 0$, получим неравенство

$$\sum_{j=1,\ j \neq i}^{N} q_{ij} \leqslant q_i.$$

Поскольку N произвольно, а все слагаемые неотрицательны, получаем требуемое утверждение.

§ 2. КОНСЕРВАТИВНЫЕ ПРОЦЕССЫ. ПРЯМЫЕ И ОБРАТНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Говорят, что цепь Маркова с непрерывным временем «консервативна», если

$$\sum_{j \neq i} q_{ij} = q_i < \infty \quad \text{при всех } i.$$

Заметим, что процесс рождения и гибели консервативен. Докажем теперь, что для консервативной цепи Маркова все $P_{ij}(t)$ не только дифференцируемы, если $q_i < \infty$ $(i \geqslant 0)$, но и удовлетворяют системе дифференциальных уравнений, известных как обратные уравнения Колмогорова. (Частный случай процесса рождения и гибели рассмотрен в гл. 7, § 5.) Напомним, однако, читателю, что дифференцируемость $P_{ij}(t)$ следует непосредственно из условий (а)—(г). Предположение о консервативности делает доказательство чрезвычайно простым. В самом деле,

$$P_{ij}(s+t) - P_{ij}(t) = \sum_k P_{ik}(s) P_{kj}(t) - P_{ij}(t) =$$
$$= \sum_{k \neq i} P_{ik}(s) P_{kj}(t) + [P_{ii}(s) - 1] P_{ij}(t).$$

Деля на s и устремляя $s \to 0+$, формально получаем *обратные уравнения*

$$P'_{ij}(t) = \sum_{k \neq i} q_{ik} P_{kj}(t) - q_i P_{ij}(t) \quad \text{для всех } i, \tag{2.1}$$

Чтобы строго вывести эти уравнения, следует показать, что

$$\lim_{s\to 0+} \frac{1}{s} \sum_{k\neq i} P_{ik}(s) P_{kj}(t) = \sum_{k\neq i} q_{ik} P_{kj}(t).$$

Далее,

$$\varliminf_{s\to 0} \frac{1}{s} \sum_{k\neq i} P_{ik}(s) P_{kj}(t) \geqslant \varliminf_{s\to 0} \frac{1}{s} \sum_{k=1,\ k\neq i}^{N} P_{ik}(s) P_{kj}(t) = \sum_{k=1,\ k\neq i}^{N} q_{ik} P_{kj}(t)$$

для любого N, и поэтому

$$\varliminf_{s\to 0} \frac{1}{s} \sum_{k\neq i} P_{ik}(s) P_{kj}(t) \geqslant \sum_{k\neq i} q_{ik} P_{kj}(t). \tag{2.2}$$

С другой стороны, при $N > i$

$$\sum_{k\neq i} P_{ik}(s) P_{kj}(t) \leqslant \sum_{k=1,\ k\neq i}^{N} P_{ik}(s) P_{kj}(t) + \sum_{k=N+1}^{\infty} P_{ik}(s) =$$

$$= \sum_{k=1,\ k\neq i}^{N} P_{ik}(s) P_{kj}(t) + 1 - P_{ii}(s) - \sum_{k=1,\ k\neq i}^{N} P_{ik}(s).$$

Деля на s и беря $\varlimsup\limits_{s\to 0+}$ от обеих частей, получаем

$$\varlimsup_{s\to 0} \frac{1}{s} \sum_{k\neq i} P_{ik}(s) P_{kj}(t) \leqslant \sum_{k=1,\ k\neq i}^{N} q_{ik} P_{kj}(t) + q_i - \sum_{k=1,\ k\neq i}^{N} q_{ik}.$$

Устремляя $N \to \infty$ и используя консервативность процесса, мы видим, что

$$\varlimsup_{s\to 0+} \frac{1}{s} \sum_{k\neq i} P_{ik}(s) P_{kj}(t) \leqslant \sum_{k\neq i} q_{ik} P_{kj}(t).$$

Сравнивая это неравенство с (2.2), заключаем, что предел

$$\lim_{s\to 0} \frac{1}{s} \sum_{k\neq i} P_{ik}(s) P_{kj}(t)$$

существует и равен

$$\sum_{k\neq i} q_{ik} P_{kj}(t).$$

Аналогичным образом можно формально вывести систему так называемых *прямых уравнений*. Запишем

$$P_{ij}(s+t) - P_{ij}(s) = \sum_{k} P_{ik}(s) P_{kj}(t) - P_{ij}(s) = \sum_{k} P_{ik}(s) [P_{kj}(t) - \delta_{kj}].$$

Деля на t и переходя к пределу при $t \to 0$, формально получаем прямые уравнения

$$P'_{ij}(s) = \sum_{k \neq j} P_{ik}(s) q_{kj} - P_{ij}(s) q_j \quad \text{для всех } i, j. \tag{2.3}$$

Вопрос о справедливости этих уравнений существенно более сложен, чем для обратных уравнений, и мы его затрагивать не будем. Обе системы уравнений в матричных обозначениях принимают весьма простой вид. В самом деле, рассмотрим бесконечную матрицу $\mathbf{A} = \| a_{ij} \|$, элементы которой равны

$$a_{ij} = \begin{cases} q_{ij}, & i \neq j, \\ -q_i, & i = j, \end{cases}$$

и которая называется *инфинитезимальной матрицей* процесса. Обратные уравнения могут быть компактно записаны в виде следующего матричного дифференциального уравнения:

$$\mathbf{P}'(t) = \mathbf{A}\mathbf{P}(t),$$

а прямые уравнения —

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{A},$$

где

$$\mathbf{P}(t) = \| P_{ij}(t) \|.$$

§ 3. ПОСТРОЕНИЕ ЦЕПИ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ С ПОМОЩЬЮ ЕЕ ИНФИНИТЕЗИМАЛЬНЫХ ПАРАМЕТРОВ

Интересным и важным вопросом в теории цепей Маркова с непрерывным параметром является следующий. Предположим, что дано множество неотрицательных чисел $\{q_{ij}\}$, обладающих свойством

$$\sum_{j \neq i} q_{ij} \leqslant q_{ii} \quad \text{для всех } i.$$

(В целях единства обозначений мы иногда пишем q_{ii} вместо q_i, как это сделано выше.) Существует ли цепь Маркова, т. е. стандартная матрица переходных вероятностей $\| P_{ij}(t) \|$, для которой

$$P'_{ij}(0) = q_{ij}, \quad i \neq j,$$
$$P'_{ii}(0) = -q_{ii} \quad ?$$

Задача становится более определенной, если предположить, что

$$\sum_{j \neq i} q_{ij} = q_{ii} < \infty \quad \text{для всех } i,$$

так как в этом случае известно, что любая цепь Маркова, связанная с $\{q_{ij}\}$, должна удовлетворять по крайней мере обратным уравнениям.

Практическая важность этих вопросов покоится на том факте (как мы уже видели, в частности, для процессов рождения и гибели, см. гл. 7), что довольно часто цепь Маркова с непрерывным временем определена таким образом, что мы вынуждены выводить обратные уравнения. Затем следует попытка решить эти уравнения для того, чтобы вычислить полную функцию перехода процесса.

В настоящее время для общего случая не получены определенные результаты. Известно, что при условии

$$\sum_{j \neq i} q_{ij} = q_{ii} \quad \text{для всех } i$$

существует по крайней мере одна отвечающая этим параметрам переходная матрица $\mathbf{P}(t)$, и если их существует более одной, то существует и бесконечно много. Конечно, больше известно о частных видах матрицы $\mathbf{A} = \|q_{ij}\|$, например в случае процесса рождения и гибели. В этом частном случае известна полная классификация всех процессов, отвечающих заданной инфинитезимальной матрице. Эти процессы в основном отличаются поведением на границе, т. е. видом выборочных функций в ∞. Напомним читателю, что в частном случае процесса рождения и гибели матрица $\mathbf{A}$ должна удовлетворять условию (4.5) гл. 7 для того, чтобы процесс определялся единственным образом.

Для общего случая цепей Маркова с непрерывным временем задача классификации инфинитезимальной матрицы $\mathbf{A}$ и отвечающих ей процессов достаточно сложна, и мы лишь отошлем интересующегося читателя к соответствующей литературе (см. ссылки в конце данной главы).

Пусть для состояния i выполняется условие $0 < q_i < \infty$. Дадим теперь некоторую интерпретацию элементам матрицы $\mathbf{A}$, аналогичную интерпретации, данной в случае процессов рождения и гибели величинам $\lambda_i + \mu_i$ и $p_i = \lambda_i(\mu_i + \lambda_i)^{-1}$. Напомним, что в этом случае мы формально доказали, что $(\lambda_i + \mu_i)^{-1}$ — среднее время пребывания в состоянии i, а $\lambda_i(\mu_i + \lambda_i)^{-1}$ — вероятность перехода в состояние $i+1$ из состояния i при условии совершения какого-либо перехода. Результаты для случая общей цепи Маркова с непрерывным временем аналогичные.

Пусть $t > 0$ фиксировано, а $n > 0$ — произвольное положительное целое число. Предположим, что процесс начинается из состояния i. Тогда рассмотрим

$$\mathrm{P}\left\{X(\tau) = i \quad \text{при} \quad \tau = 0, \ \frac{t}{n}, \ \frac{2t}{n}, \ \frac{3t}{n}, \ \ldots, \ t\right\} = \left[P_{ii}\left(\frac{t}{n}\right)\right]^n.$$

Поскольку

$$\frac{1-P_{ii}(t)}{t}=q_i+o(1) \qquad (o(1)\to 0 \text{ при } t\to 0+),$$

то

$$\left[P_{ii}\left(\frac{t}{n}\right)\right]^n=\left(1-\frac{t}{n}q_i+o\left(\frac{t}{n}\right)\right)^n=\exp\left(n\ln\left[1-\frac{tq_i}{n}+o\left(\frac{t}{n}\right)\right]\right).$$

Используя разложение логарифма вида $\ln(1-x)=-x+\theta(x)x^2$, верное при $|x|\leqslant\frac{1}{2}$ и $|\theta|\leqslant 1$, где $x=\frac{tq_i}{n}+o\left(\frac{t}{n}\right)$, и затем переходя к пределу при $n\to\infty$, получим

$$\lim_{n\to\infty}\left[P_{ii}\left(\frac{t}{n}\right)\right]^n=\exp(-q_it).$$

Но (см. гл. 1, стр. 18) вероятность

$$\mathrm{P}\{X(\tau)=i,\ 0\leqslant\tau\leqslant t\}$$

совпадает с

$$\lim_{n\to\infty}\mathrm{P}\left\{X(\tau)=i,\ \tau=0,\ \frac{t}{n},\ \frac{2t}{n},\ \ldots,\ \frac{n-1}{n}t,\ t\right\}$$

(при этом неявно предполагается, что процесс является сепарабельным), откуда видно, что вероятность пребывания в состоянии i в течение отрезка времени, не меньшего t, равна $\exp(-q_it)$.

Другими словами, время пребывания процесса в состоянии i имеет экспоненциальное распределение с параметром q_i. Это является строгим доказательством (в общем случае цепей Маркова с непрерывным временем) того, что было эвристически показано для частного случая процессов рождения и гибели (см. стр. 214).

Состояние i, для которого $0\leqslant q_i<\infty$, называется *устойчивым*. Оно называется поглощающим, если $q_i=0$; очевидно, если процесс перешел в такое состояние i, то он останется там навсегда. Конечно, в этом случае

$$\mathrm{P}\{X(\tau)=i,\ 0\leqslant\tau<t\mid X(0)=i\}=\exp(-q_it)=1$$

для всех t. С другой стороны, если $q_i>0$, то время пребывания в состоянии i является случайной величиной, функция распределения которой есть «настоящая» экспонента, и, следовательно, выход из состояния i осуществляется за конечное время.

Состояние i, для которого $q_i=\infty$, называется *мгновенным*. Среднее время пребывания в таком состоянии равно нулю. Наименование «мгновенное» обязано тому факту, что время пребывания процесса в состоянии i равно нулю, т. е., попадая в это состояние, процесс мгновенно его покидает [1]).

Теория цепей Маркова с непрерывным временем, имеющих мгновенные состояния, крайне сложна, особенно при рассмотрении выборочных траекторий процесса. Дело усложняется еще тем, что марковские цепи могут состоять только из мгновенных состояний. Имеет смысл найти технические задачи, отвечающие таким примерам. Но в то же самое время утешительно сознавать, что почти все цепи Маркова с непрерывным временем, возникающие в приложениях, имеют только устойчивые состояния. Действительно, в большинстве случаев, представляющих интерес, изучаемый процесс обычно определяется инфинитезимальными характеристиками как известными данными. Чтобы завершить теорию, необходимо установить существование процесса (т. е. определить выборочные траектории и их вероятностные законы), обладающего заданной инфинитезимальной матрицей.

Наиболее элементарные учебники и обсуждения цепей Маркова с непрерывным временем избегают этого аспекта задачи (так же поступим и мы), делая упор преимущественно на нахождение функций распределения процесса и вычисление различных вероятностных характеристик, представляющих интерес. Вычисление переходной функции для всех t традиционно сопровождается выводом обратных дифференциальных уравнений и, в лучшем случае, их решением. Такой подход был основой нашего рассмотрения процессов рождения и гибели (см. гл. 7).

В оставшейся части главы мы будем рассматривать цепи Маркова с непрерывным временем, имеющие лишь устойчивые состояния. Наша следующая задача — придать физический смысл параметрам q_{ij}. Действительно, если процесс консервативен, то $\dfrac{q_{ij}}{q_i}$ можно интерпретировать как условные вероятности того, что произойдет переход из состояния i в состояние j. Чтобы показать это, рассмотрим

$$R_{ij}(h) = \mathrm{P}\{X(h) = j \mid X(0) = i,\ X(h) \neq i\}, \quad j \neq i,$$

---

[1]) На первый взгляд кажется, что если i — мгновенное состояние, то $P_{ii}(t)$ не является непрерывной функцией в нуле. Однако это не так. Если $X(0) = i$, то, хотя процесс и покинет сразу это состояние, он будет «достаточно часто» в него возвращаться. Именно если обозначить через $\theta_i(t)$ общее время пребывания процесса в состоянии i на отрезке $[0, t]$, то оказывается, что

$$\mathrm{P}\left\{\lim_{t\to 0} \frac{1}{t}\theta_i(t) = 1 \mid X(0) = i\right\} = 1.$$

Это обеспечивает непрерывность функции $P_{ii}(t)$, когда i — мгновенное состояние (см., например, Прохоров Ю. В., Розанов Ю. А., Теория вероятностей, «Наука», 1967, гл. V, § 1). — *Прим. перев.*

и найдем $\lim_{h \downarrow 0} R_{ij}(h)$. Он равен вероятности перехода из состояния i в состояние j при условии, что переход осуществляется. Далее, в силу определения $R_{ij}(h)$ имеем

$$R_{ij}(h) = \frac{P_{ij}(h)}{1 - P_{ii}(h)}.$$

Деля числитель и знаменатель на h, устремляя $h \downarrow 0$ и используя результаты теорем 1.1 и 1.2, получаем искомую формулу

$$\lim_{h \downarrow 0} R_{ij}(h) = \frac{q_{ij}}{q_i}, \quad j \neq i.$$

Если взять сумму от правой части (по j), то она будет равна 1, так как процесс консервативный.

Мы отметили выше, что для любых инфинитезимальных параметров $q_{ij} \geqslant 0$ $(i \neq j)$ и q_i $(0 < q_i < \infty)$ $(i \geqslant 0)$, удовлетворяющих $q_i \geqslant \sum_{j \neq i} q_{ij}$, может существовать одна или бесконечно много цепей Маркова с непрерывным временем, имеющих одну и ту же инфинитезимальную матрицу $\mathbf{A}$. В случае консервативной инфинитезимальной матрицы (т. е. $q_i = \Sigma q_{ij}$ при всех i) существует один специальный процесс (*минимальный процесс*), для которого можно просто описать выборочные траектории. Построение минимального процесса для случая процессов рождения и гибели было показано в гл. 7, § 4. В общем случае метод остается тем же. Опишем кратко основные идеи этого построения для общего случая. Типичная реализация, начинающаяся из некоторого состояния, скажем i, имеет следующий вид. Возьмем выборку из экспоненциального распределения с параметром q_i. Она определяет время пребывания в состоянии i. В конце этого интервала частица перемещается в состояние j с вероятностью $\frac{q_{ij}}{q_i}$ $(j \neq i)$. В новом состоянии, скажем j', она пребывает случайное время (экспоненциально распределенное с параметром $q_{j'}$). По окончании времени пребывания в состоянии j' она переходит в новое состояние j с вероятностью $\frac{q_{j'j}}{q_{j'}}$ $(j \neq j')$; там частица проводит случайное время с соответствующим экспоненциальным законом распределения, затем снова совершает переход и т. д. С помощью этой последовательной процедуры мы строим все возможные реализации процесса. Используя довольно глубокие методы теории меры, можно найти переходную функцию $\|\tilde{P}_{ij}(t)\|$, имеющую заданную инфинитезимальную матрицу.

Другой путь описания минимального процесса состоит в следующем. Переходная матрица $\|\tilde{P}_{ij}(t)\|$ определяется через матрицы вероятностей различных переходов, совершаемых только за конечное число скачков. Более определенно: мы вводим в рассмотрение вероятность $P_{ij}(t; N)$ перехода из состояния i в состояние j за время t и за число переходов, не превосходящее N. В частности, в соответствии со смыслом инфинитезимальных параметров

$$P_{ij}(t;\ 0) = \begin{cases} \exp(-q_i t), & i = j, \\ 0, & i \neq j, \end{cases}$$

и можно выписать рекуррентное соотношение, связывающее $P_{ij}(t; N)$ с $P_{ij}(t; N-1)$ следующим образом.

Рассмотрим сначала $P_{ii}(t; N)$. В соответствии с тем, произошел переход до момента t или нет, возникают две возможности. Время пребывания в состоянии i распределено экспоненциально с параметром q_i, так что с вероятностью $\exp(-q_i t)$ перехода не происходит. Предположим, что первый переход произошел в интервале от τ до $\tau + d\tau$ (вероятность этого события равна $q_i \exp(-q_i\tau)\,d\tau$) и что при этом состояние приняло значение $j \neq i$. (Вероятность этого последнего события равна $\frac{q_{ij}}{q_i}$.) Вероятность возвратиться в состояние i в оставшееся время $t-\tau$ не более чем за $N-1$ переходов равна $P_{ji}(t-\tau; N-1)$. В силу формулы полной вероятности имеем

$$P_{ii}(t;\ N) = \exp(-q_i t) + \sum_{j \neq i} \frac{q_{ij}}{q_i} \int_0^t P_{ji}(t-\tau;\ N-1)\, q_i \exp(-q_i\tau)\, d\tau.$$

С помощью аналогичного перечисления различных возможностей получим равенство

$$P_{ij}(t;\ N) = \sum_{k \neq i} q_{ik} \int_0^t P_{kj}(t-\tau;\ N-1) \exp(-q_i\tau)\, d\tau, \qquad i \neq j.$$

Можно показать, что

$$\lim_{N\to\infty} P_{ij}(t;\ N) = \tilde{P}_{ij}(t)$$

(см. ссылки в конце главы).

Здесь мы вынуждены принести извинения перед читателем за то, что ввели множество понятий, важных в теории марковских процессов, и почти не изучали их. Эти понятия содержат в себе много тонкостей и патологий, изучение которых далеко выходит за рамки этого учебника. Можно лишь надеяться, что читатель продолжит изучение этих вопросов по другим отличным книгам, рассматривающим данный предмет.

§ 4. СТРОГО МАРКОВСКОЕ СВОЙСТВО

Строгое рассмотрение длительностей пребывания для марковского процесса, или даже для цепей Маркова с непрерывным временем, тесно связано с так называемым строго марковским свойством. Хотя полное обсуждение этого вопроса требует привлечения языка теории меры, основную идею можно объяснить гораздо более просто.

Для того чтобы развить основную идею, необходимо ввести случайную величину специального типа, связанную со случайным процессом, известную под различными названиями: *случайная величина, не зависящая от будущего*, или *марковский момент*, или *время остановки* (мы будем использовать термин «марковский момент»). Пусть σ является неотрицательной случайной величиной, связанной с данным процессом $\{X_t\}$ с непрерывным параметром. Другими словами, свяжем с каждой выборочной функцией X_t неотрицательное (возможно, бесконечное) число, которое обозначим через $\sigma(X_t)$. Говорят, что случайная величина σ является марковским моментом относительно $\{X_t\}$, если она обладает следующим свойством: если X_t и Y_t — две выборочные функции процесса, такие, что $X_\tau = Y_\tau$ при $0 \leqslant \tau \leqslant s$ и $\sigma(X_t) < s$, то $\sigma(X_t) = \sigma(Y_t)$.

На описательном языке, который полезен для понимания существенных свойств таких случайных величин, можно сказать, что случайная величина σ «следит» за выборочной функцией, начиная с $t = 0$ до некоторого момента σ_0, который зависит только от значений рассматриваемой выборочной функции на отрезке $[0, \sigma_0]$, и момент σ_0, в который σ прекращает «слежение», является значением $\sigma(X_t)$ [1].

Для того чтобы привести пример марковского момента, предположим, что задана цепь Маркова, начальное состояние которой , т. е. $X(0) = j_0$. Если положить

$$\sigma(X_t) = \inf\{\tau \mid X_\tau = i\},$$

где i — фиксированное состояние, и принять $\inf = \infty$, если не существует τ, для которого $X_\tau = i$, то σ является марковским моментом. Формальное доказательство этого простое. Действительно, пусть $\sigma(X_t) < s$ для фиксированной выборочной функции X_t. Это означает, что существует момент $t' < s$, для которого $X_{t'} = i$. Далее,

---

[1] Строго говоря, это описание несколько более ограничительное, чем данное ранее определение. Это случилось из-за того, что ранее было условие $\sigma(X_t) < s$, а не $\sigma(X_t) \leqslant s$. Этого ограничения можно избежать, сказав, что для каждого $\varepsilon > 0$ значение $\sigma(X_t)$, если σ прекращает «слежение» в момент σ_0, лежит в интервале $(\sigma_0 - \varepsilon, \sigma_0)$. Однако здесь это является совершенно ненужным усложнением.

если $X_\tau = Y_\tau$ при $0 \leqslant \tau \leqslant s$, то $Y_{t'} = i$ и заведомо $\sigma(Y_t) \leqslant \sigma(X_t)$. Следовательно, в силу симметрии $\sigma(X_t) = \sigma(Y_t)$. Такая случайная величина σ называется временем первого достижения состояния i. Аналогично марковским моментом является время первого достижения любого конечного множества состояний C, не содержащего $X(0) = j_0$ и определяемого следующим образом:

$$\sigma(X_t) = \inf\{\tau \mid X_\tau \in C\}.$$

Доказательство этого повторяет предыдущие рассуждения. Тривиальным марковским моментом является $\sigma \equiv \text{const}$. Читатель без труда должен построить другие примеры марковских моментов.

Если ограничиться интуитивным уровнем, то марковское свойство однородных марковских процессов утверждает следующее. Если известны значения $X(s_i)$ при $0 \leqslant s_1 < s_2 < \ldots < s_n = t_0$ ($t_0 > 0$ фиксировано), то вероятностное распределение величин

$$X(t_0 + t_1),\ X(t_0 + t_2),\ \ldots,\ X(t_0 + t_k) \qquad (0 < t_1 < t_2 < \ \ldots \ < t_k) \tag{4.1}$$

зависит только от значения $X(t_0)$. Более точно, вероятностное распределение величин (4.1) при условии, что известны значения $X(s)$ в моменты $0 \leqslant s_1 < s_2 < \ldots < s_n = t_0$ [или даже, в более общем случае, вся история процесса $X(s)$ вплоть до момента t_0 $(0 \leqslant s \leqslant t_0)$], совпадает с вероятностным распределением величин

$$X(t_1),\ X(t_2),\ \ldots,\ X(t_k) \quad \text{при известном} \quad X(0).$$

Другими словами, можно найти вероятностный закон для величин (4.1), перемещая шкалу времени так, чтобы $t_0 = 0$, и беря в качестве начального состояния значение $X(t_0)$.

Интуитивно кажется правдоподобным, что то же самое соотношение должно выполняться, если заменить фиксированное значение t_0 на «марковский момент» σ. Более точно, предположим, что мы хотим найти вероятностное распределение величины

$$X(\sigma + \tau) \quad \text{при известном} \quad X(\sigma) = x, \tag{4.2}$$

т. е. вероятностное распределение $X(\tau)$, где $\tau > 0$ — время, прошедшее с момента σ, при известном значении X в момент σ. В случайной величине σ «учитывается» поведение процесса только до момента σ включительно, но не далее, хотя само значение σ не обязательно фиксировано и может изменяться от одной реализации к другой. Другими словами, это случайный момент.

Кажется естественной возможность считать, что марковское свойство справедливо для случайного момента σ. Отсюда тогда следовало бы, что вероятностное распределение величины (4.2) совпадает с законом распределения величины

$$X(t) \quad \text{при условии, что} \quad X(0) = x. \tag{4.3}$$

Этот факт не является прямым следствием марковского свойства, поскольку в первоначальной формулировке говорилось о фиксированных моментах. Утверждение, что (4.2) и (4.3) «управляются» одним и тем же вероятностным законом, и выражает строго марковское свойство. Более точно, если для любого марковского момента σ вероятностное распределение величин

$$X(t_1+\sigma),\ X(t_2+\sigma),\ \ldots,\ X(t_k+\sigma), \qquad t_1<t_2<\ \ldots\ <t_k, \tag{4.4}$$

при известных $X(s)$, $s\leqslant\sigma$, и $X(\sigma)=x$ совпадает с вероятностным распределением величин

$$X(t_1),\ X(t_2),\ \ldots,\ X(t_k) \qquad (t_1<\ \ldots\ <t_k)$$

при известном $X(0)=x$, то говорят, что процесс Маркова обладает *строго марковским свойством.* Существуют примеры марковских процессов, не являющихся строго марковскими.

Результат такого рода, который следует из (4.4), крайне важен для вычисления различных представляющих интерес вероятностных характеристик. Действительно, один из основных приемов анализа случайных процессов и вычисления вероятностных характеристик состоит в получении рекуррентных соотношений, использующих обычно первые или последние моменты осуществления определенных событий. Рассмотрим пример. Предположим, что мы хотим найти $P_{ij}(t)$, представляя рассматриваемое событие через момент первого достижения состояния j. Пусть σ_{ij} — момент первого достижения состояния j, начиная из состояния i. Выше было указано, что моменты первого достижения конечного множества состояний являются марковскими, в частности σ_{ij} — марковский момент. Пусть

$$F_{ij}(s)=\mathrm{P}\{\sigma_{ij}\leqslant s\}.$$

Разумеется, начиная с момента σ_{ij}, когда частица попадает в состояние j, вероятностное распределение ее будущей истории такое же, как если бы $\sigma_{ij}=0$, а начальное состояние цепи было бы j и цепь Маркова «управлялась» бы матрицей $\|P_{kl}(t)\|$ обычным образом. Это положение справедливо лишь при условии выполнения строго марковского свойства. Тогда верно соотношение

$$P_{ij}(t)=\int\limits_0^t P_{jj}(t-s)\,dF_{ij}(s), \tag{4.5}$$

которое является непрерывным аналогом формулы (5.9) гл. 2. Читатель может интерпретировать $dF_{ij}(s)$ как функцию плотности $f_{ij}(s)ds$ величины σ_{ij}, когда эта плотность существует. Формула (4.5) следует из формулы полной вероятности, где $dF_{ij}(s)$ — вероятность того, что $s\leqslant\sigma_{ij}\leqslant s+ds$, а $P_{jj}(t-s)$ — переходная вероятность, т. е. вероятность того, что если момент σ_{ij} начальный (частица тогда с необходимостью находится в состоянии j), то $t-s$ единиц времени спустя процесс будет снова в состоянии j.

Можно выписать множество других соотношений типа восстановления, подобных (4.5) (родственных формулам (2.1), (2.2) и (2.3) гл. 5). Мы еще раз подчеркнем, что такие соотношения справедливы в общем случае лишь при выполнении строго марковского свойства.

Тот факт, что длительности последовательных пребываний в данном состоянии (или в двух фиксированных состояниях) являются независимыми случайными величинами, есть непосредственное следствие строго марковского свойства. (Читатель должен уметь обосновать это формально.)

Поскольку строго марковское свойство играет фундаментальную роль при анализе марковских процессов, мы в заключение приведем наиболее приятный результат. *Любая консервативная цепь Маркова с непрерывным временем, имеющая лишь устойчивые состояния, — строго марковский процесс.* Или, в более общей формулировке, марковское свойство выполняется для любого марковского момента σ, такого, что $X(\sigma) \neq \infty$ с вероятностью 1.

С практической точки зрения это означает, что почти все соотношения типа восстановления для таких процессов вполне корректны и могут использоваться безбоязненно. Доказательство приведенного выше результата, более подробное обсуждение строго марковского свойства для цепей Маркова с непрерывным временем и другие фундаментальные вопросы можно найти в книге Чжун Кай-лая [1].

ЗАДАЧИ

1. Пусть $\mathbf{P}(t) = \| P_{ij}(t) \|_{i,\,j=0}^{N}$ — матрица переходных вероятностей цепи Маркова с конечным множеством состояний и непрерывным временем. Показать, что $\det[\mathbf{P}(t)] > 0$ при всех $t > 0$.

2. Пусть $\mathbf{P}$ — стохастическая матрица размера 2×2, т. е. $\mathbf{P} = \begin{Vmatrix} \alpha & \beta \\ \gamma & \delta \end{Vmatrix}$, где $\alpha, \beta, \gamma, \delta \geqslant 0$, $\alpha+\beta=1$, $\gamma+\delta=1$. Доказать, что тогда и только тогда существует непрерывное семейство $\mathbf{P}(t)$, $t > 0$, стохастических матриц 2×2, таких, что $\mathbf{P}(1) = \mathbf{P}$, когда

$$\det \mathbf{P} = \alpha\delta - \gamma\beta > 0 \quad \text{и} \quad \alpha + \delta > 1.$$

Указание: Использовать тот факт, что по предположению $\mathbf{P}$ можно представить в виде $e^{\mathbf{A}}$, где $\mathbf{A}$ имеет вид $\mathbf{A} = \begin{Vmatrix} -a & a \\ b & -b \end{Vmatrix}$ $(a,\ b > 0)$.

3. Показать, что если $\mathbf{P}(t) = \| P_{ij}(t) \|_{i,\,j=0}^{\infty}$ удовлетворяет постулатам (а) — (г) § 1, то $\mathbf{P}(t)$ непрерывна при всех $t > 0$.
Указание: См. гл. 7, § 8.

4. Рассмотрим неприводимую цепь Маркова с непрерывным временем и конечным числом состояний $1, 2, \ldots, N$. Пусть q_{ij}, $i, j = 1, \ldots, N$, — инфинитезимальные параметры процесса. Предположим, что $q_{ij} = q_{ji}$ для всех $i, j = 1,$

2, ..., N. Пусть $P_i(t)$ — вероятность того, что процесс находится в момент t в состоянии i. Введем функцию

$$E(t) = -\sum_{i=1}^{N} P_i(t) \ln P_i(t),$$

где $x \ln x = 0$ при $x = 0$.

Доказать, что $E(t)$ — неубывающая функция $t \geqslant 0$.

Указание: Доказать сначала, что

$$P_i'(t) = \sum_{j=1}^{N} q_{ij}\left[P_j(t) - P_i(t)\right].$$

Используя эту формулу, доказать затем, что

$$\frac{dE(t)}{dt} = \frac{1}{2} \sum_{i,\, j \neq 1}^{N} q_{ij}\left[P_j(t) - P_i(t)\right]\left[\ln P_j(t) - \ln P_i(t)\right].$$

5. Пусть X_t $(t > 0)$ — консервативная цепь Маркова с непрерывным временем, у которой все состояния устойчивы, т. е. $\sum_{j \neq i} q_{ij} = q_i$. Предположим, что $q_{ij} > 0$, $i \neq j$, и $0 < q_i < \infty$. Показать, что процесс X_t возвратен, если существует последовательность $z = (z_0, z_1, z_2, \ldots)$, такая, что (1) $z_n \to \infty$ при $n \to \infty$, (2) $-q_i z_i + \sum_{j \neq i} q_{ij} z_j \leqslant 0$, $i \geqslant 1$. (Возвратность означает, что в каждое состояние процесс возвращается бесконечное число раз, или, что то же, суммарное время пребывания в каждом состоянии бесконечно.)

Указание: Использовать вложенную цепь Маркова $\{Y_n\}$, $n = 0, 1, 2, \ldots$, со стационарной матрицей переходных вероятностей $\mathbf{P} = \| P_{ij} \|$, где

$$P_{ij} = \frac{q_{ij}}{q_i}, \qquad i \neq j,$$
$$P_{ii} = 0.$$

Вложенная цепь Маркова $\{Y_n\}$ попросту фиксирует последовательность состояний, через которые проходит процесс X_t безотносительно к длительностям пребывания в различных состояниях. Использовать теорему 4.2 гл. 3.

6. В предположениях задачи 5 показать, что X_t невозвратен, если существует ограниченная непостоянная последовательность $z = (z_0, z_1, z_2, \ldots)$, такая, что

$$-z_i q_i + \sum_{j \neq i} q_{ij} z_j = 0, \qquad i \geqslant 1.$$

Указание: См. гл. 3, § 4.

7. Показать, что сумма двух независимых пуассоновских процессов — снова пуассоновский процесс.

8. *Процесс восстановления* является целочисленным случайным процессом, который равен числу точек (событий) в интервале $(0, t)$, если интервалы между наступлениями событий — независимые одинаково распределенные случайные величины с общей функцией распределения $F(x)$, $x \geqslant 0$ $(F(x) = 0,\ x < 0)$; F непрерывна при $x = 0$. *Общим процессом восстановления* назовем такой процесс, у которого общая функция распределения $F(x)$ имеет скачок величины q в нуле. Показать, что общий процесс восстановления эквивалентен обычному, у кото-

рого число одновременно наступаемых событий в момент 0 и в другие «вызывающие» моменты — независимые одинаково распределенные случайные величины R_0, R_1, R_2, ... с распределением

$$\mathrm{P}\{R_i = n\} = pq^n, \qquad n = 0, 1, 2, \ldots$$

для всех $i = 0, 1, 2, \ldots$, где $p = 1 - q$.

9. Доказать, что если сумма двух независимых процессов восстановления есть пуассоновский процесс, то оба процесса восстановления — пуассоновские.

10. Показать, что для неприводимой непериодической возвратной цепи Маркова

$$\lim_{n \to \infty} \left(P_{00}^n\right)^{1/n} = 1.$$

Указание: Использовать полуаддитивные функции, введенные в процессе доказательства теоремы 1.1.

11. Рассмотрим цепь Маркова с непрерывным временем и конечным (скажем, равным N) числом состояний. Пусть $\mathbf{A}$ — инфинитезимальная матрица. Показать, что $\mathbf{A}$ имеет ранг $N - 1$.

12. Рассмотрим консервативную цепь Маркова с непрерывным временем и двумя состояниями. Пусть

$$\mathbf{A} = \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}$$

инфинитезимальная матрица процесса. Показать, что в этом случае обратные уравнения имеют вид

$$P''_{ij}(t) - (\mathrm{Sp}\,\mathbf{A})\, P'_{ij}(t) + (\det \mathbf{A})\, P_{ij}(t) = 0, \qquad i, j = 1, 2\ [1].$$

Решить эти уравнения для $\mathbf{A} = \begin{Vmatrix} -1 & 1 \\ 1 & -1 \end{Vmatrix}$.

ЗАМЕЧАНИЯ

Эта глава содержит лишь малую часть материала обстоятельной книги Чжун Кай-лая [1]. Основные идеи этой главы, развитые для случая общих марковских процессов, можно найти в книгах Дынкина [2] и Дуба [3].

ЛИТЕРАТУРА

1. Чжун Кай-лай, Однородные цепи Маркова, гл. 3 и 4, «Мир», 1964.
2. Дынкин Е. Б., Марковские процессы, Физматгиз, 1963.
3. Дуб Дж. Л., Вероятностные процессы, ИЛ, 1956.

[1]) Если $\mathbf{A} = \| a_{ij} \|$, то $\mathrm{Sp}\,\mathbf{A} = \sum_i a_{ii}$ — след матрицы $\mathbf{A}$. — *Прим. перев.*

Глава 9

ПОРЯДКОВЫЕ СТАТИСТИКИ, ПУАССОНОВСКИЕ ПРОЦЕССЫ И ИХ ПРИЛОЖЕНИЯ

В этой главе приводятся различные специальные приложения некоторых методов, связанных с пуассоновскими процессами и суммами независимых случайных величин.

§ 1. ПОРЯДКОВЫЕ СТАТИСТИКИ И ИХ СВЯЗЬ С ПУАССОНОВСКИМИ ПРОЦЕССАМИ

Пусть $Y_1, Y_2, \ldots, Y_n$ суть n независимых одинаково распределенных случайных величин, имеющих непрерывную строго возрастающую функцию распределения $F(y)$. Определим случайные величины $Y_1^*, \ldots, Y_n^*$ следующим образом:

Y_i^* есть i-я в порядке возрастания величина среди $Y_1, Y_2, \ldots, Y_n$ [1]). В частности,

$$Y_1^* = \min\{Y_1, Y_2, \ldots, Y_n\} \quad \text{и} \quad Y_n^* = \max\{Y_1, Y_2, \ldots, Y_n\}.$$

Очевидно,

$$Y_1^* \leqslant Y_2^* \leqslant \ldots \leqslant Y_n^*.$$

Y_i^* называется *i-й порядковой статистикой* выборки $(Y_1, \ldots, Y_n)$, а набор $(Y_1^*, \ldots, Y_n^*)$ — *множеством порядковых статистик объема n*, соответствующих выборке $(Y_1, \ldots, Y_n)$.

В этой главе рассматриваются распределения порядковых статистик выборок, их связь с пуассоновскими процессами и другие приложения. Сначала, однако, сделаем существенное упрощение без потери общности.

Положим

$$X_i = F(Y_i), \qquad i = 1, \ldots, n,$$

и найдем распределение величины X_i:

$$P\{X_i < x\} = P\{F(Y_i) < x\} = P\{Y_i < F^{-1}(x)\} =$$
$$= F\left(F^{-1}(x)\right) = x, \qquad 0 \leqslant x \leqslant 1, \qquad i = 1, \ldots, n, \qquad (1.1)$$

---

[1]) В случае равенства делается произвольный выбор среди равных по величине Y_i. В действительности событие, состоящее в том, что по крайней мере два значения Y_i будут равны, имеет вероятность 0, и его можно не принимать во внимание.

где F^{-1} — функция, обратная к F, — определена единственным образом в силу сделанных предположений относительно F. Далее, поскольку $0 \leqslant F(y) \leqslant 1$, то

$$P\{X_i < x\} = \begin{cases} 0, & \text{если } x < 0, \\ 1, & \text{если } x > 1, \end{cases} \quad i = 1, \ldots, n. \tag{1.2}$$

Таким образом, в силу (1.1) и (1.2) X_i распределена равномерно на $[0, 1]$ при всех $i = 1, \ldots, n$ независимо от вида непрерывной строго возрастающей функции F. Заметим, что отношение порядка среди $\{Y_i^*\}$ сохраняется при преобразовании $X_i = F(Y_i)$. Это означает, что вместо исследования порядковых статистик $\{Y_i^*\}$, соответствующих выборке общего вида $(Y_1, \ldots, Y_n)$, можно изучать порядковые статистики $\{X_i^*\}$ для равномерного на отрезке $[0, 1]$ распределения.

Поэтому в дальнейшем мы ограничимся изучением порядковых статистик

$$X_1^* \leqslant X_2^* \leqslant \ldots \leqslant X_n^* \tag{1.3}$$

для выборки независимых равномерно распределенных на отрезке $[0, 1]$ случайных величин

$$X_1, X_2, \ldots, X_n.$$

Тот факт, что для случайных величин $X_1^*, X_2^*, \ldots, X_n^*$ выполняется (1.3), ясно указывает, то они не являются независимыми. Мы сначала найдем совместное распределение порядковых статистик $X_1^*, \ldots, X_n^*$ или, вернее, его функцию плотности, которую обозначим через $f^*(x_1, \ldots, x_n)$. Ее существование будет очевидно из доказательства. Выбирая $0 < x_1 < x_2 < \ldots < x_n < 1$ и достаточно малые приращения $h_1, h_2, \ldots, h_n$ так, чтобы интервалы $(x_1, x_1 + h_1)$, $(x_2, x_2 + h_2)$, $\ldots$, $(x_n, x_n + h_n)$ не пересекались, получаем

$$\int_{x_n}^{x_n+h_n} \ldots \int_{x_1}^{x_1+h_1} f^*(x_1, \ldots, x_n)\, dx_1 \ldots dx_n =$$

$$= P\{x_i \leqslant X_i^* < x_i + h_i, \quad i = 1, 2, \ldots, n\} =$$

$$= \sum_{\substack{\text{по всем перестановкам} \\ \sigma \text{ из } (1, 2, \ldots, n)}} P\{x_i \leqslant X_{i\sigma} < x_i + h_i, \ i = 1, 2, \ldots, n\} =$$

$$= \sum_{\sigma} \prod_{i=1}^{n} P\{x_i \leqslant X_{i\sigma} < x_i + h_i\} = \quad (\text{в силу независимости величин } X_i)$$

$$= \sum_{\sigma} \prod_{i=1}^{n} h_i = n!\, h_1 h_2 \ldots h_n; \tag{1.4}$$

здесь использовались независимость, равномерное распределение величин X_i $(i = 1, \ldots, n)$ и тот факт, что число всевозможных перестановок индексов $1, 2, \ldots, n$ равно $n!$.

Из (1.4) следует, что совместная плотность порядковых статистик $X_1^*, \ldots, X_n^*$ равна

$$f^*(x_1, \ldots, x_n) = \begin{cases} n!, & 0 \leqslant x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n \leqslant 1, \\ 0 & \text{в противном случае.} \end{cases} \tag{1.5}$$

Такое же доказательство показывает, что если бы величины $X_1, \ldots, X_n$ были взяты из равномерного распределения на отрезке $[a, b]$, то соответствующая совместная плотность порядковых статистик равнялась бы

$$f^*(x_1, \ldots, x_n) = \begin{cases} \dfrac{n!}{(b-a)^n}, & a \leqslant x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n \leqslant b, \\ 0 & \text{в противном случае.} \end{cases} \tag{1.6}$$

Мы уже сталкивались с плотностью (1.5) при обсуждении пуассоновских процессов (см. стр. 206). Именно: пусть $\{Y(t), 0 \leqslant t \leqslant 1\}$ — пуассоновский процесс, в частности, при каждом $t \in [0, 1]$ пусть $Y(t)$ будет дискретной случайной величиной с распределением

$$p_k(t) = \begin{cases} e^{-\lambda t} \dfrac{(\lambda t)^k}{k!}, & k = 0, 1, 2, \ldots, \\ 0 & \text{в противном случае,} \end{cases}$$

где λ — фиксированный действительный параметр.

Предположим, что $Y(0) = 0$. Тогда при условии, что $Y(1) = n$ (целое число), на отрезке $[0, 1]$ будет ровно n временны́х точек, в которых $Y(t)$ совершает «скачок». Точное положение этих точек зависит от случая и определяется случайными величинами

$$T_1, T_2, \ldots, T_n \qquad (T_1 < T_2 < \ldots < T_n)$$

со значениями из отрезка $[0, 1]$.

Утверждается следующее. При условии

$$Y(1) = n$$

случайные величины $T_1, T_2, \ldots, T_n$ распределены как множество порядковых статистик объема n, взятых из равномерного распределения на отрезке $[0, 1]$.

Доказательство этого положения непосредственно вытекает из уже полученных результатов. Действительно, вывод условной плотности распределения величин $T_1, T_2, \ldots, T_n$ при условии, что $Y(1) = n$, был дан в теореме 2.3 гл. 7. Сравнение с (1.5) показывает, что эти формулы совпадают, и утверждение, таким образом, доказано.

Соответствие между порядковыми статистиками и условными моментами наступления событий пуассоновских процессов упрощает вывод других свойств порядковых статистик.

Например, пусть, как и ранее, $X_1^*, X_2^*, \ldots, X_n^*$ — порядковые статистики выборки объема n из равномерного распределения. Мы утверждаем, что совместное распределение величин $X_1^*, X_2^*, \ldots$ $\ldots, X_{k-1}^*$ при условии, что $X_k^* = c_k, \ldots, X_n^* = c_n$, совпадает с распределением порядковых статистик для выборки $X_1, \ldots, X_{k-1}$, где каждая д. с. в. равномерно распределена на $[0, c_k]$. Чтобы проверить этот факт, переформулируем задачу в терминах событий пуассоновского процесса.

Пусть $0 \leqslant T_1 \leqslant T_2 \leqslant \ldots \leqslant T_n \leqslant 1$ — моменты наступления событий пуассоновского процесса $Y(t)$ при условии, что $Y(1) = n$. Предположим, что налагаются дополнительные условия $T_k = c_k$, $T_{k+1} = c_{k+1}, \ldots, T_n = c_n$ и требуется определить совместное распределение величин $T_1, T_2, \ldots, T_{k-1}$. Так как $Y(t)$ — процесс с независимыми приращениями, то очевидно, что если сделано предположение $T_k = c_k$, или, что эквивалентно, $Y(c_k - \varepsilon) = k - 1$ (где $\varepsilon > 0$ и достаточно мало), то вся информация, относящаяся к величинам $T_1, T_2, \ldots, T_{k-1}$, содержится в этом предположении. Но при этом условии $T_1, \ldots, T_{k-1}$ распределены как порядковые статистики выборки объема $k - 1$ из равномерного распределения на отрезке $[0, c_k]$. Таким образом, условное совместное распределение величин $X_1^*, X_2^*, \ldots, X_{k-1}^*$ при условии $X_k^* = c_k, \ldots, X_n^* = c_n$ совпадает с совместным распределением порядковых статистик объема $k - 1$, взятых из равномерного распределения на отрезке $[0, c_k]$. Соответствующая условная плотность равна

$$f^*(x_1, \ldots, x_{k-1} \mid c_k, \ldots, c_n) = \begin{cases} \dfrac{(k-1)!}{c_k^{k-1}}, & 0 \leqslant x_1 \leqslant \ldots \leqslant x_{k-1} \leqslant c_k, \\ 0 & \text{в противном случае.} \end{cases} \quad (1.7)$$

В точности таким же образом можно доказать, что условная плотность величин $X_{k+1}^*, \ldots, X_n^*$ при условии, что значения первых k порядковых статистик равны

$$X_1^* = c_1, \ldots, X_k^* = c_k,$$

совпадает с совместной плотностью $n - k$ порядковых статистик выборки из $n - k$ независимых случайных величин, распределенных на отрезке $[c_k, 1]$:

$$f^*(x_{k+1}, \ldots, x_n \mid c_1, \ldots, c_k) = \begin{cases} \dfrac{(n-k)!}{(1-c_k)^{n-k}}, & c_k \leqslant x_{k+1} \leqslant \ldots \leqslant x_n \leqslant 1, \\ 0 & \text{в противном случае.} \end{cases}$$

$$(1.8)$$

Формулы (1.7) и (1.8) показывают, что эти совместные условные плотности зависят лишь от c_k и не зависят от c_i, $i = k+1, \ldots, n$ и $i = 1, 2, \ldots, k-1$ соответственно. Это означает, что совместные плотности величин $X_1^*, \ldots, X_{k-1}^*$ и $X_{k+1}^*, \ldots, X_n^*$ при одном и том же единственном условии $X_k^* = c_k$ имеют соответственно вид

$$f^*(x_1, \ldots, x_{k-1} \mid c_k) = \begin{cases} \dfrac{(k-1)!}{c_k^{k-1}}, & 0 \leqslant x_1 \leqslant \ldots \leqslant x_{k-1} \leqslant c_k, \\ 0 & \text{в противном случае} \end{cases} \tag{1.9}$$

и

$$f^*(x_{k+1}, \ldots, x_n \mid c_k) = \begin{cases} \dfrac{(n-k)!}{(1-c_k)^{n-k}}, & c_k \leqslant x_{k+1} \leqslant \ldots \leqslant x_n \leqslant 1, \\ 0 & \text{в противном случае.} \end{cases} \tag{1.10}$$

Формулы (1.7) и (1.8) совместно с (1.9) и (1.10) также показывают, что множества величин $X_1^*, \ldots, X_{k-1}^*$ и $X_{k+1}^*, \ldots, X_n^*$ (условно) независимы при условии, что $X_k^* = c_k$.

Более того, два множества случайных величин $X_1^*, \ldots, X_i^*$ и $X_{k+1}^*, \ldots, X_n^*$ $(i < k)$ условно независимы при известных значениях остальных величин $X_{i+1}^*, \ldots, X_k^*$.

Пользуясь этим, можно получить совместную плотность для любого числа последовательных порядковых статистик. Так, совместная плотность величин $X_1^*, \ldots, X_i^*, X_{k+1}^*, \ldots, X_n^*$ $(i < k)$ при условии, что $X_{i+1}^* = x_{i+1}, \ldots, X_k^* = x_k$, равна

$$f^*(x_1, \ldots, x_i, x_{k+1}, \ldots, x_n \mid x_{i+1}, \ldots, x_k) = \frac{f^*(x_1, \ldots, x_n)}{f^*(x_{i+1}, \ldots, x_k)}. \tag{1.11}$$

С другой стороны, в силу вышеприведенного утверждения о независимости левая часть равна

$$f^*(x_1, \ldots, x_i \mid x_{i+1}, \ldots, x_k) f^*(x_{k+1}, \ldots, x_n \mid x_{i+1}, \ldots, x_k) =$$
$$= f^*(x_1, \ldots, x_i \mid x_{i+1}) f^*(x_{k+1}, \ldots, x_n \mid x_k) = \frac{i!}{x_{i+1}^i} \frac{(n-k)!}{(1-x_k)^{n-k}}.$$

Отсюда, а также из (1.11) и (1.5) получаем для $0 \leqslant i < k \leqslant n$

$$f^*(x_{i+1}, \ldots, x_k) = \begin{cases} \dfrac{n!}{i!\,(n-k)!} x_{i+1}^i (1-x_k)^{n-k}, & 0 \leqslant x_{i+1} \leqslant \ldots \leqslant x_k \leqslant 1, \\ 0 & \text{в противном случае.} \end{cases} \tag{1.12}$$

В частности, (1.12) при $i+1=k$ дает маргинальную (частную) плотность величины X_k^*

$$f^*(x_k)=\begin{cases}\dfrac{n!}{(k-1)!\,(n-k)!}\,x_k^{k-1}\left(1-x_k\right)^{n-k}, & 0\leqslant x_k\leqslant 1,\\ 0 & \text{в противном случае,}\end{cases}\tag{1.13}$$

которая является плотностью бета-распределения.

Порядковые статистики $X_1^*, X_2^*, \ldots, X_n^*$ делят интервал $(0, 1)$ на $n+1$ непересекающихся интервалов, имеющих длины

$$U_1=X_1^*,\ U_2=X_2^*-X_1^*,\ \ldots,\ U_n=X_n^*-X_{n-1}^*,\ U_{n+1}=1-X_n^*.$$

Очевидно, $U_1, U_2, \ldots, U_n, U_{n+1}$ не являются независимыми случайными величинами, поскольку $\sum_{i=1}^{n+1} U_i=1$. Записывая преобразование переменных $\left(x_1^*, \ldots, x_n^*\right)$ в $\left(u_1, \ldots, u_n\right)$

$$\begin{aligned} u_1&=x_1^*,\\ u_2&=-x_1^*+x_2^*,\\ &\ \vdots\\ u_n&=\qquad\qquad -x_{n-1}^*+x_n \end{aligned}\tag{1.14}$$

и вычисляя якобиан этого преобразования, который в данном случае равен тождественно 1, можно найти совместную плотность $g(u_1, \ldots, u_n)$ случайных величин $\{U_1, \ldots, U_n\}$. Именно

$$g(u_1, \ldots, u_n)=\begin{cases} n!, & u_i\geqslant 0\ (i=1, \ldots, n),\ \sum_{i=1}^{n} u_i\leqslant 1,\\ 0 & \text{в противном случае.}\end{cases}\tag{1.15}$$

Таким образом, можно сказать, что случайные величины $\{U_1, \ldots, U_n\}$ распределены равномерно в области

$$u_i\geqslant 0,\quad i=1, \ldots, n,\quad \sum_{i=1}^{n} u_i\leqslant 1.$$

Это также определяет распределение случайных величин $\{U_1, \ldots, U_n, U_{n+1}\}$ в области

$$u_i \geqslant 0, \quad i=1, \ldots, n, \; n+1, \quad \sum_{i=1}^{n+1} u_i = 1.$$

Покажем теперь, что совместное распределение величин $\{U_1, \ldots, U_{n+1}\}$ совпадает с распределением величин

$$Y_1/S, \ldots, Y_n/S, \; Y_{n+1}/S,$$

где $S = Y_1 + \ldots + Y_n + Y_{n+1}$, а Y_i, $i = 1, \ldots, n, n+1$, — независимые экспоненциально распределенные случайные величины (с параметром λ). Этот результат может быть доказан с помощью введения пуассоновского процесса и анализа задачи в новых терминах. Для разнообразия дадим прямое доказательство. Для этого запишем совместную плотность величин $(Y_1, \ldots, Y_{n+1})$

$$f(y_1, \ldots, y_{n+1}) = \begin{cases} \lambda^{n+1} \exp\left(-\lambda \sum_{i=1}^{n+1} y_i\right), & y_i \geqslant 0, \; i=1, \ldots, n+1, \\ 0 & \text{в противном случае} \end{cases}$$

и сделаем преобразование

$$v_1 = \frac{y_1}{y_1 + \ldots + y_{n+1}},$$

$$v_2 = \frac{y_2}{y_1 + \ldots + y_{n+1}},$$

$$\vdots$$

$$v_n = \frac{y_n}{y_1 + \ldots + y_{n+1}},$$

$$v_{n+1} = y_1 + \ldots y_{n+1}.$$

Обратное преобразование имеет вид

$$y_1 = v_1 v_{n+1},$$

$$y_2 = v_2 v_{n+1},$$

$$\vdots$$

$$y_n = v_n v_{n+1},$$

$$y_{n+1} = v_{n+1}[1 - (v_1 + \ldots + v_n)].$$

Отсюда можно найти якобиан

$$J=\begin{vmatrix} v_{n+1} & 0 & 0 & \dots & 0 & v_1 \\ 0 & v_{n+1} & 0 & \dots & 0 & v_2 \\ 0 & 0 & v_{n+1} & \dots & 0 & v_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & v_{n+1} & v_n \\ -v_{n+1} & -v_{n+1} & -v_{n+1} & \dots & -v_{n+1} & 1-v_1-v_2-\dots-v_n \end{vmatrix}=$$

$$=\begin{vmatrix} v_{n+1} & 0 & 0 & \dots & 0 & v_1 \\ 0 & v_{n+1} & 0 & \dots & 0 & v_2 \\ 0 & 0 & v_{n+1} & \dots & 0 & v_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ & & & & v_{n+1} & v_n \\ 0 & 0 & 0 & \dots & 0 & 1 \end{vmatrix}=v_{n+1}^n.$$

Следовательно, совместная плотность величин

$$Y_1/S,\ Y_2/S,\ \dots,\ Y_n/S,\ S$$

равна

$$f(v_1,\ \dots,\ v_n,\ v_{n+1})=$$

$$=\begin{cases} \lambda^{n+1}\exp(-\lambda v_{n+1})\,v_{n+1}^n, & v_i\geqslant 0,\ i=1,\ \dots,\ n+1,\ \sum\limits_{i=1}^{n+1} v_i=1, \\ 0 & \text{в противном случае.} \end{cases}$$

Отсюда следует, что S и случайный вектор $(Y_1/S,\ \dots,\ Y_n/S)$ независимы и имеют следующие маргинальные плотности:

$$f(v_{n+1})=\begin{cases} \dfrac{\lambda^{n+1}}{n!}\exp\left(-\lambda v_{n+1}\right)v_{n+1}^n, & v_{n+1}\geqslant 0, \\ 0 & \text{в противном случае} \end{cases}$$

и

$$f(v_1,\ \dots,\ v_n)=\begin{cases} n!, & v_i\geqslant 0,\ i=1,\ \dots,\ n,\ \sum\limits_{i=1}^{n} v_i\leqslant 1, \\ 0 & \text{в противном случае.} \end{cases} \tag{1.16}$$

Поскольку (1.16) согласуется с (1.15) и поскольку

$$\frac{Y_1}{S}+\frac{Y_2}{S}+\dots+\frac{Y_n}{S}+\frac{Y_{n+1}}{S}=1,$$

утверждение о равенстве распределений величин

$$(U_1, \ldots, U_{n+1}) \quad \text{и} \quad (Y_1/S, \ldots, Y_{n+1}/S)$$

доказано.

§ 2. ЗАДАЧА О БАЛЛОТИРОВКЕ

Теперь мы хотим показать некоторые применения пуассоновских процессов и соответствующих порядковых статистик к анализу различных случайных величин, связанных с эмпирическими функциями распределения. Для этого мы сначала получим некоторые результаты, известные под названием *задачи о баллотировке*, которые представляют значительный интерес и ценность сами по себе.

Задача о баллотировке может быть сформулирована следующим образом.

При баллотировке (выборах) и общем количестве избирателей c кандидаты A и B получают a и b голосов соответственно, $a + b = c$. В течение подсчета голосов лидерство кандидатов может все время меняться. Задача о баллотировке (в ее простейшем варианте) состоит в следующем: предполагая, что $a \geqslant b$, найти вероятность того, что кандидат A при подсчете голосов будет всегда впереди (хотя бы на один голос).

Непосредственное решение задачи о баллотировке следующее. Рассмотрим фиксированное размещение a символов A и b символов B на окружности. Для данного размещения определим число начальных позиций, при отсчете от которых, скажем по часовой стрелке, символы A всегда будут лидировать в счете.

Чтобы найти эти позиции, исключим последовательно все соседние пары AB, проходя для этого окружность, возможно, несколько раз. В результате останется $a - b$ символов A. Легко понять, что оставшиеся места и являются теми исходными позициями, при отсчете от которых символы A всегда лидируют. Отсюда следует, что вероятность того, что A всегда лидирует (возможное наблюдение является описанным размещением или одной из его циклических перестановок), равна $(a - b)/(a + b)$. Эта вероятность не зависит от выбора исходной последовательности. Отсюда вероятность того, что кандидат A всегда лидирует при подсчете голосов, равна $(a - b)/(a + b)$.

Проведенный анализ весьма элегантен. Однако, поскольку мы имеем в виду другие обобщения, сформулируем теперь задачу о баллотировке в терминах более общей «урновой схемы» и проанализируем ее структуру.

В урне a карт, на которых написан нуль, и b карт, на которых написана двойка, $a + b = c$, $a \geqslant b$. Карты вынимаются из урны одна за другой случайным образом без возвращения до тех пор, пока не будет вынута последняя карта. Пусть ν_i — случайная величина, равная числу, написанному на i-й карте, $i = 1, \ldots, c$. Тогда задача о баллотировке будет решена, если найти

$$\mathrm{P}\{\nu_1 + \nu_2 + \ldots + \nu_r < r,\ r = 1, 2, \ldots, c\}. \tag{2.1}$$

Это утверждение следует из того факта, что если среди первых r чисел имеется α нулей и β двоек $(\alpha + \beta = r)$, то условие $\nu_1 + \ldots + \nu_r < r$ означает, что $\alpha \cdot 0 + \beta \cdot 2 < \alpha + \beta$, или $\beta < \alpha$. Очевидно, (2.1) является вероятностью того, что это неравенство $(\beta < \alpha)$ выполняется для $r = 1, 2, \ldots, c$.

Чтобы найти вероятность (2.1), заметим сначала, что $(\nu_1, \nu_2, \ldots, \nu_c)$ является набором из a нулей и b двоек и все из $c!/(a!b!)$ возможных перестановок равновероятны. Это означает, что для любого r $(r = 1, \ldots, c)$ и любого множества $i_1, \ldots, i_r$ различных чисел из набора $\{1, \ldots, c\}$ совместное распределение $(\nu_{i_1}, \nu_{i_2}, \ldots, \nu_{i_r})$ совпадает с распределением $(\nu_1, \nu_2, \ldots, \nu_r)$. Говорят, что случайные величины $\nu_1, \ldots, \nu_r$, обладающие таким свойством, *перестановочны*. (Независимые одинаково распределенные случайные величины являются перестановочными.) Тогда, поскольку

$$\nu_1 + \nu_2 + \ldots + \nu_c = a \cdot 0 + b \cdot 2 = 2b, \tag{2.2}$$

имеем

$$\sum_{i=1}^{c} \mathrm{M}\{\nu_i\} = \mathrm{M}\{\nu_1 + \nu_2 + \ldots + \nu_c\} = 2b.$$

Поскольку д. с. в. $\nu_1, \nu_2, \ldots, \nu_n$ перестановочны, все ν_i имеют одно и то же маргинальное распределение и, следовательно,

$$c\mathrm{M}\{\nu_i\} = 2b, \text{ или } \mathrm{M}\{\nu_i\} = 2b/(a+b), \qquad i = 1, 2, \ldots, c. \tag{2.3}$$

Используя этот факт, докажем теперь по индукции относительно c, что

$$\mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, c \mid \nu_1 + \nu_2 + \ldots + \nu_c = 2b\} = \\ = 1 - 2b/c = (a-b)/(a+b). \tag{2.4}$$

Покажем сначала, что (2.4) справедливо для $c = 1$. Но из того, что $c = 1$, следует, что $a = 1$, $b = 0$, поскольку неравенство $a > b \geqslant 0$ предполагалось в постановке задачи. Тогда, очевидно,

$$\mathrm{P}\{\nu_1 < 1\} = \mathrm{P}\{\nu_1 = 0\} = 1.$$

Утверждение (2.4) тривиально для $2b = c$, поскольку в этом случае $\nu_1 + \ldots + \nu_c = 2b = c$.

Предположим теперь, что (2.4) выполнено для всех $c \leqslant n - 1$ и $0 \leqslant 2b < c$. Мы хотим доказать, что (2.4) справедливо также для $c = n$ и $0 \leqslant 2b < n$.

Пусть b' — целое число, такое, что $0 \leqslant b' \leqslant b$. Тогда

$$\begin{aligned} &\mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, c \mid \nu_1 + \ldots + \nu_{2b} = 2b'\} = \\ &\quad = \mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, 2b \mid \nu_1 + \ldots + \nu_{2b} = 2b'\}, \end{aligned} \tag{2.5}$$

так как неравенство $\nu_1 + \ldots + \nu_r < r$ всегда выполнено при $r = 2b + 1, \ldots, c$ в силу условия (2.2). Но правая часть (2.5) является выражением того же типа, что и левая часть (2.4) при условии вида (2.2). Действительно, в (2.4) заменены лишь c на $2b$ и $2b$ на $2b'$. Используя гипотезу индукции, где c и b заменены соответственно на $2b$ и b', получаем

$$\begin{aligned} \mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, 2b \mid \nu_1 + \ldots + \nu_{2b} = 2b'\} &= \\ &= 1 - 2b'/2b. \end{aligned} \tag{2.6}$$

Чтобы завершить доказательство, запишем

$$\begin{aligned} &\mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, n\} = \\ &\quad = \sum_{b'=0}^{b} \mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, n \mid \nu_1 + \ldots + \nu_{2b} = 2b'\} \times \\ &\times \mathrm{P}\{\nu_1 + \ldots + \nu_{2b} = 2b'\} = \sum_{b'=0}^{b} (1 - 2b'/2b)\, \mathrm{P}\{\nu_1 + \ldots + \nu_{2b} = 2b'\} \end{aligned} \tag{2.7}$$

в силу (2.6). Но

$$\sum_{b'=0}^{b} 2b' \mathrm{P}\{\nu_1 + \ldots + \nu_{2b} = 2b'\} = \mathrm{M}\{\nu_1 + \ldots + \nu_{2b}\} = 2b \cdot 2b/n$$

в силу (2.1). Следовательно, из (2.7) вытекает, что

$$\begin{aligned} &\mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, n\} = \\ &= \sum_{b'=0}^{b} \mathrm{P}\{\nu_1 + \ldots + \nu_{2b} = 2b'\} - \frac{1}{2b} \sum_{b'=0}^{b} 2b' \mathrm{P}\{\nu_1 + \ldots + \nu_{2b} = 2b'\} = \\ &\qquad = 1 - \frac{1}{2b}\, 2b\, \frac{2b}{n} = 1 - \frac{2b}{n}. \end{aligned}$$

Таким образом, (2.4) выполняется для $c = n$ и $0 \leqslant 2b < n$. Это завершает доказательство.

Обобщение задачи о баллотировке состоит в предположении о наличии урны с c картами, на которых написаны числа $k_1, \ldots, k_c$, такие, что

$$k_i \text{ — неотрицательные целые числа, } i = 1, \ldots, c,$$
$$k_1 + \ldots + k_c = k, \qquad 0 \leqslant k \leqslant c.$$

Если ν_i снова означает число на i-й по счету карте, вынутой из урны без возвращения, то (2.4) для общего случая примет вид

$$\mathrm{P}\{\nu_1 + \ldots + \nu_r < r,\ r = 1, \ldots, c\} = 1 - k/c. \tag{2.8}$$

Доказательство этого более общего утверждения можно провести, почти дословно повторяя доказательство (2.4). Мы опускаем детали. (Читателю предлагается провести доказательство самостоятельно.)

Дальнейшее обобщение ведет к следующей задаче. Пусть $\nu_1, \ldots, \nu_n$ — неотрицательные перестановочные случайные величины и

$$\nu_1 + \ldots + \nu_n = y \quad \text{(фиксированное число).} \tag{2.9}$$

Пусть $\tau_1, \ldots, \tau_n$ — порядковые статистики для n независимых наблюдений случайной величины, равномерно распределенной на

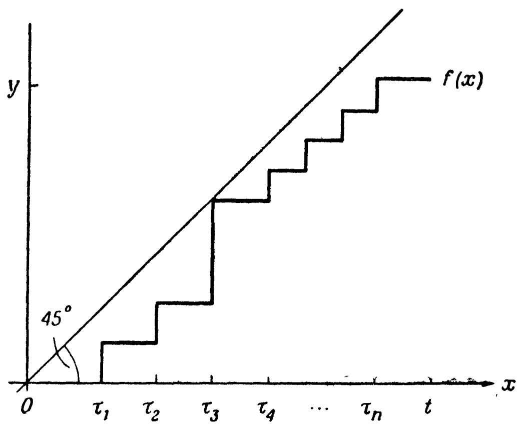


Рис. 1.

отрезке $[0, t]$. Предположим далее, что случайные величины $\nu_i (i = 1, \ldots, n)$ не зависят от величин $\tau_j (j = 1, \ldots, n)$.

Введем в рассмотрение ступенчатую функцию (рис. 1)

$$f(x) = \begin{cases} 0, & 0 \leqslant x < \tau_1, \\ \nu_1 + \ldots + \nu_r, & \tau_r \leqslant x < \tau_{r+1}, \ r = 1, \ldots, n-1, \\ \nu_1 + \ldots + \nu_n, & \tau_n \leqslant x < t, \end{cases}$$

и поставим вопрос, чему равна вероятность того, что график $f(x)$ не пересечет прямой $y = x$ (рис. 1). Аналитическая формулировка и решение этой задачи таковы:

$$\mathrm{P}\{\nu_1 + \ldots + \nu_r \leqslant \tau_r, \ r = 1, \ldots n\} = \begin{cases} 1 - y/t, & 0 \leqslant y \leqslant t, \\ 0, & y > t. \end{cases} \tag{2.10}$$

Докажем (2.10) по индукции. При $n = 1$

$$P\{\nu_1 \leqslant \tau_1\} = \begin{cases} 1 - y/t, & 0 \leqslant y \leqslant t, \\ 0, & y > t, \end{cases}$$

поскольку $\nu_1 = y$, а τ_1 распределена равномерно на $[0, t]$.

Предположим теперь, что (2.10) выполняется для $n - 1$. Накладывая условия

$$\nu_1 + \ldots + \nu_{n-1} = z \qquad (0 \leqslant z \leqslant y)$$

и

$$\tau_n = u \qquad (0 \leqslant u \leqslant t), \tag{2.11}$$

найдем условную вероятность

$$P\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n \mid \nu_1 + \ldots + \nu_{n-1} = z,\ \tau_n = u\}. \tag{2.12}$$

Предположим сначала, что $y \leqslant u \leqslant t$. Тогда в силу (2.9) неравенство

$$\nu_1 + \ldots + \nu_n \leqslant \tau_n$$

будет заведомо выполнено при условии (2.11). Следовательно, вероятность (2.12) равна

$$P\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n-1 \mid \nu_1 + \ldots + \nu_{n-1} = z,\ \tau_n = u\} = \\ = \begin{cases} 1 - z/u, & 0 \leqslant z \leqslant u, \\ 0, & z > u, \end{cases}$$

в силу гипотезы индукции, поскольку $(\tau_1, \ldots, \tau_{n-1})$ при условии (2.11) являются порядковыми статистиками для $n - 1$ независимых случайных величин, равномерно распределенных на $[0, u]$.

Далее, помня, что ν_i не зависит от τ_n при всех $i = 1, \ldots, n$, рассмотрим плотность

$$\varphi(z) = \frac{d}{dz} P\{\nu_1 + \ldots + \nu_{n-1} \leqslant z\}.$$

Тогда для $y \leqslant u \leqslant t$

$$P\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n \mid \tau_n = u\} = \\ = \int_0^y P\{\nu_1 + \ldots + \nu_r \leqslant \tau_r, r = 1, \ldots, n \mid \nu_1 + \ldots + \nu_{n-1} = z,\ \tau_n = u\} \times \\ \times \varphi(z)\,dz = \int_0^y \left(1 - \frac{z}{u}\right) \varphi(z)\,dz = 1 - \frac{1}{u} \int_0^y z\varphi(z)\,dz = \\ = 1 - \frac{1}{u} M(\nu_1 + \ldots + \nu_{n-1}) = 1 - \frac{1}{u}\,\frac{n-1}{n}\,y, \tag{2.13}$$

так как из (2.9) и перестановочности $\nu_1, \ldots, \nu_n$ следует, что

$$\mathrm{M}\{\nu_i\} = y/n, \qquad i = 1, 2, \ldots, n.$$

При $u < y$, однако,

$$\mathrm{P}\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n \mid \nu_n = u\} = 0, \tag{2.14}$$

поскольку при условии (2.11) неравенство

$$\nu_1 + \ldots + \nu_n \leqslant \tau_n$$

выполнено быть не может. Далее, в силу (1.13) плотность величины τ_n равна

$$\psi(u) = \begin{cases} n(u/t)^{n-1}(1/t), & 0 \leqslant u \leqslant t, \\ 0 & \text{в противном случае.} \end{cases}$$

Тогда в силу (2.13) и (2.14) при $0 \leqslant y \leqslant t$

$$\mathrm{P}\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n\} =$$

$$= \int_0^t \mathrm{P}\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n \mid \tau_n = u\}\, \psi(u)\, du =$$

$$= \int_y^t \left(1 - \frac{1}{u}\, \frac{n-1}{n}\, y\right) n \left(\frac{u}{t}\right)^{n-1} \frac{du}{t} = 1 - \frac{y}{t}.$$

При $y > t$, очевидно,

$$\mathrm{P}\{\nu_1 + \ldots + \nu_r \leqslant \tau_r,\ r = 1, \ldots, n\} = 0.$$

Это доказывает равенство (2.10).

Мы теперь в состоянии дать некоторые приложения упомянутых идей к порядковым статистикам.

§ 3. ЭМПИРИЧЕСКИЕ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И ПОРЯДКОВЫЕ СТАТИСТИКИ

Важный класс задач, связанных с порядковыми статистиками, касается эмпирических функций распределения случайных величин. Если X — случайная величина с функцией распределения $F(x)$ и

$$(X_1^*, X_2^*, \ldots, X_n^*)$$

— множество порядковых статистик, соответствующих выборке размера n из $F(x)$, то *эмпирическая функция распределения*

$F_n(x)$ величины X является случайной величиной, определяемой следующим образом:

$$F_n(x) = \begin{cases} 0, & \text{если } x < X_1^*, \\ k/n, & \text{если } X_k^* \leqslant x < X_{k+1}^*, \quad k = 1, \ldots, n-1, \\ 1, & \text{если } x \geqslant X_n^*. \end{cases} \tag{3.1}$$

Пусть X — случайная величина с непрерывной строго возрастающей функцией распределения $F(x)$ и эмпирической функцией распределения $F_n(x)$. Мы хотим найти вероятность

$$\mathrm{P}\{F_n(x) < \gamma F(x), \quad -\infty < x < \infty\}. \tag{3.2}$$

Как мы уже видели в начале этой главы, без потери общности можно считать, что X равномерно распределена на отрезке $[0, 1]$.

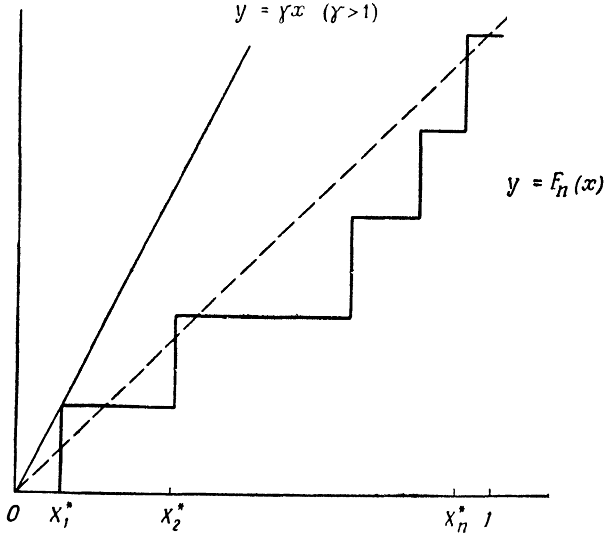


Рис. 2.

В самом деле, заменим X на величину $F(X)$, соответствующие наблюдения которой равны $F(X_1)$, $F(X_2)$, ..., $F(X_n)$. Тогда (3.2) сведется к

$$\mathrm{P}\{F_n(x) < \gamma x, \quad 0 \leqslant x \leqslant 1\}, \tag{3.3}$$

где $F_n(x)$ — теперь эмпирическая функция распределения, связанная с равномерным распределением.

На рис. 2 показана типичная реализация $F_n(x)$ для равномерного распределения.

Докажем, что

$$\mathrm{P}\{F_n(x) < \gamma x,\ 0 < x \leqslant 1\} = \begin{cases} 0, & \text{если } \gamma \leqslant 1, \\ 1 - 1/\gamma, & \text{если } \gamma > 1. \end{cases} \tag{3.4}$$

Если $\gamma \leqslant 1$, то равенство (3.4) очевидно, поскольку не выполнено условие $F_n(1) < \gamma$. Результат для случая $\gamma > 1$ является следствием равенства (2.10). Необходимо лишь сделать подстановку $t = 1$, $y = 1/\gamma$. В самом деле, ясно (рис. 2), что событие

$$F_n(x) < \gamma x, \qquad 0 < x \leqslant 1,$$

произойдет тогда и только тогда, когда

$$X_k^* > k/(n\gamma), \qquad k = 1, 2, \ldots, n, \tag{3.5}$$

где X_k^* $(k = 1, 2, \ldots, n)$ — порядковые статистики объема n, соответствующие равномерному распределению на отрезке $[0, 1]$. Случайные величины $\nu_1, \nu_2, \ldots, \nu_n$ (см. (2.10)) в данном случае являются фиксированными числами $\nu_i = 1/(n\gamma)$ $(i = 1, \ldots, n)$, которые, несомненно, перестановочны. Тогда (3.5) можно записать в виде

$$\nu_1 + \nu_2 + \ldots + \nu_k < X_k^*, \qquad k = 1, 2, \ldots, n,$$

где $\nu_1 + \nu_2 + \ldots + \nu_n = 1/\gamma = y$.

Равенство (3.4) теперь следует из (2.10). Заметим, что правая часть (3.4) не зависит от n.

Следующая задача, относящаяся к бросанию монеты, также включает в себя эмпирическую функцию распределения. Пусть X — случайная величина с непрерывной строго возрастающей функцией распределения $F(x)$. Пусть $(X_1, \ldots, X_n)$ и $(Y_1, \ldots, Y_n)$ — две независимые случайные выборки объема n из распределения величины X. Пусть $(X_1^*, \ldots, X_n^*)$ и $(Y_1^*, \ldots, Y_n^*)$ — соответствующие порядковые статистики. Образуем эмпирические функции распределения:

$$F_n(x) = \begin{cases} 0, & \text{если } x < X_1^*, \\ k/n, & \text{если } X_k^* \leqslant x < X_{k+1}^*, \qquad k = 1, \ldots, n-1, \\ 1, & \text{если } x \geqslant X_n^*, \end{cases}$$
$$G_n(y) = \begin{cases} 0, & \text{если } y < Y_1^*, \\ l/n, & \text{если } Y_l^* \leqslant y < Y_{l+1}^*, \qquad l = 1, \ldots, n-1, \\ 1, & \text{если } y \geqslant Y_n^*. \end{cases} \tag{3.6}$$

Можно снова предположить, что $F(x)$ — равномерное распределение на отрезке $[0, 1]$. Как показывает рис. 3, $F_n(x) = G_n(x) = 0$ при $x < \min(X_1^*, Y_1^*)$ и $F_n(x) = G_n(x) = 1$ при $x \geqslant \max(X_n^*, Y_n^*)$.

В интервале $I=(\min(X_1^*, Y_1^*), \max(X_n^*, Y_n^*))$ один из двух графиков может лежать полностью ниже другого или, как на рис. 3, они могут пересекаться. Перейдем к определению вероятности первого из двух упомянутых событий, т. е. найдем

$$P\{F_n(x) \neq G_n(x) \text{ в любом подинтервале в } I\}. \qquad (3.7)$$

Эта вероятность может быть интерпретирована следующим образом. Будем рассматривать $(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$ как од-

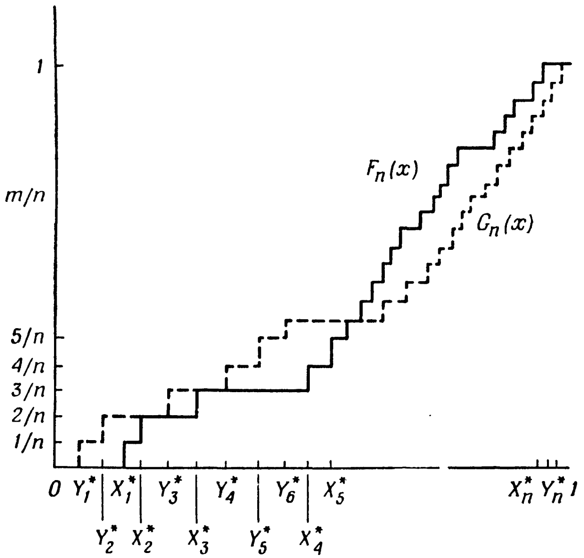


Рис. 3.

ну выборку и образуем соответствующие порядковые статистики. Возможный вид их может быть следующим (рис. 3):

$$(Y_1^*, Y_2^*, X_1^*, X_2^*, Y_3^*, X_3^*, Y_4^*, Y_5^*, Y_6^*, X_4^*, X_5^*, \ldots, X_n^*, Y_n^*).$$

Так как обе выборки взяты из одного распределения, все такие перестановки равновероятны. Обозначим порядковые статистики вектором

$$(Z_1^*, Z_2^*, \ldots, Z_{2n}^*), \qquad (3.8)$$

где половина величин Z_k^* является статистиками X^*, а другая половина — статистиками Y^*. Для каждого k можно найти отношение

$$\rho_k = \frac{\text{число величин } X_i^* \leqslant Z_k^*}{\text{число величин } Y_j^* \leqslant Z_k^*}.$$

Очевидно, $\rho_{2n} = 1$. Поскольку график $F_n(x)$ будет лежать ниже графика $G_n(x)$ при всех $x \in I$ тогда и только тогда, когда $\rho_k < 1$ при всех $k = 2, 3, \ldots, 2n-1$, выражение (3.7) эквивалентно следующему:

$$P\{\rho_k < 1,\ k = 2,\ 3,\ \ldots,\ 2n-1, \text{ или } \rho_k > 1,\ k = 2,\ 3,\ \ldots,\ 2n-1\}. \tag{3.9}$$

Другое графическое представление события, записанного в фигурных скобках в формуле (3.9), можно дать следующим образом. Рассмотрим ступенчатую функцию (рис. 4), которая совершает

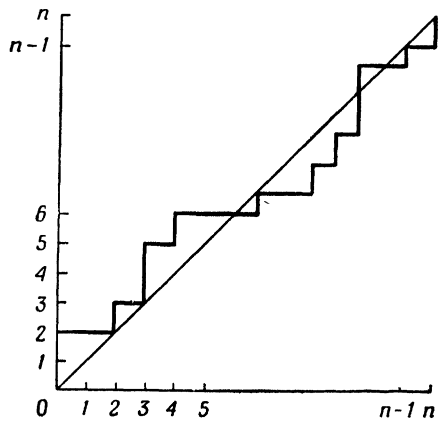


Рис. 4.

горизонтальный единичный скачок всякий раз, когда в перестановке порядковых статистик встречается X_i^*, и вертикальный единичный скачок всякий раз, когда встречается Y_j^*. График этой ступенчатой функции лежит строго выше прямой с наклоном 45° (кроме граничных точек, где он будет совпадать с ней) тогда и только тогда, когда $\rho_k < 1$ при всех $k = 2, 3, \ldots, 2n-1$. Следовательно, выражение (3.9) (а также (3.7)) есть вероятность того, что график ступенчатой функции на рис. 4 лежит полностью по одну сторону от прямой с наклоном 45° (кроме граничных точек).

Эта задача может быть интерпретирована как последовательное бросание монеты. Предположим, что мы проводим серию из $2n$ бросаний симметричной монеты. Предполагая, что в конце серии общее число выпадений герба равно общему числу выпадений решетки (и оба они равны n), можно спросить, чему равна вероятность того, что число выпадений герба всегда больше числа выпадений решетки, наблюдаемых по ходу игры, или наоборот (число выпадений решетки всегда больше числа выпадений герба при каждом бросании)?

Если сопоставить горизонтальным скачкам события, состоящие в выпадении герба, а вертикальным скачкам — в выпадении решетки, то таким образом каждой серии ставится в соответствие ступенчатая функция типа показанной на рис. 4. Следовательно, вероятность того, что число выпадений герба превышает число выпадений решетки (или наоборот) в течение всей серии из $2n$ бросаний при условии, что они равны в конце, равна вероятности (3.9).

Перейдем теперь к вычислению этой вероятности, для которой было дано несколько интерпретаций. Для этой цели мы ради удобства будем ссылаться на типичные случайные ступенчатые функции вида функции, представленной на рис. 4. Графики этих ступенчатых функций идут из точки $(0, 0)$ в точку (n, n) за $2n$ скачков, n из которых должны быть вертикальными. Очевидно, общее число таких функций равно $\binom{2n}{n}$. Чтобы сосчитать те из них, графики которых не имеют общих точек (кроме граничных) с прямой, имеющей наклон 45°, достаточно в силу симметрии сосчитать те ступенчатые функции, графики которых проходят ниже указанной прямой, и удвоить это число. Однако график любой ступенчатой функции, лежащий ниже прямой с наклоном 45°, проходит на первом шаге через точку $(1, 0)$ и затем идет к точке (n, n). Очевидно, существует $\binom{2n-1}{n}$ ступенчатых функций, графики которых ведут из точки $(1, 0)$ в точку (n, n).

Мы хотим сосчитать лишь те из них, которые лежат ниже прямой с наклоном 45°. Чтобы найти это число, мы сосчитаем число ступенчатых функций, графики которых ведут из точки $(1, 0)$ в точку (n, n) и имеют по крайней мере одну общую точку с прямой с наклоном 45°, а затем вычтем это число из $\binom{2n-1}{n}$. Покажем сначала, что любая ступенчатая функция, график которой ведет из $(1, 0)$ в (n, n), имеет хотя бы одну общую точку с прямой с наклоном 45° и заканчивается вертикальным скачком (рис. 5), соответствует ступенчатой функции, график которой ведет из $(1, 0)$ в (n, n), пересекает прямую с наклоном 45° и заканчивается горизонтальным скачком. Чтобы показать это, возьмем ступенчатую функцию, график которой соприкасается с прямой с наклоном 45° и заканчивается вертикальным скачком; пусть (k, k) — точка их последнего соприкосновения перед достижением (n, n). Отразим часть графика между точками (k, k) и (n, n) симметрично относительно указанной прямой (см. пунктирную линию на рис. 5). Очевидно, такой процесс устанавливает взаимно однозначное соответствие между двумя видами графиков ступенчатых функций — касающихся прямой с наклоном 45° и заканчивающихся вертикальным скачком и оканчивающихся горизонтальным скачком после соприкосновения с прямой с наклоном 45°. Таким образом, нужно лишь подсчитать ступенчатые функции, графики которых ведут из (1, 0) в (n, n), пересекают прямую с наклоном 45° и оканчиваются горизонтальным скачком. Но, очевидно, график *любой* ступенчатой функции, ведущий из (1, 0) в (n, n) и оканчивающийся горизонтальным скачком, должен проходить через точку ($n - 1$, n), а таких графиков существует ровно $\binom{2n-2}{n}$.

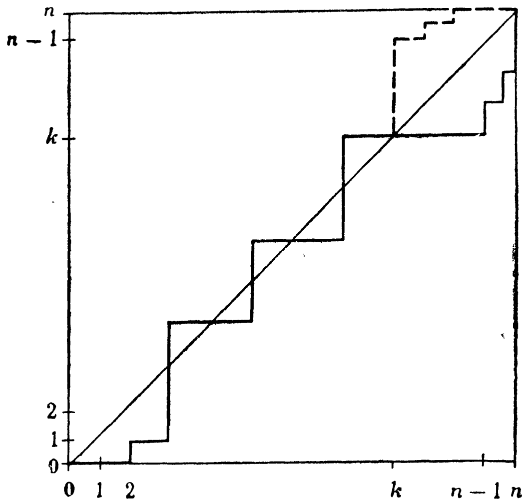


Рис. 5.

В силу принципа отражения мы знаем, что число ступенчатых функций, графики которых ведут из (1, 0) в (n, n), пересекают прямую с наклоном 45° или касаются ее и оканчиваются вертикальным скачком, также равно $\binom{2n-2}{n}$. Тогда число ступенчатых функций, графики которых ведут из (1, 0) в (n, n) и всегда остаются ниже прямой с наклоном 45° (кроме точки (n, n)), равно

$$\binom{2n-1}{n} - 2\binom{2n-2}{n} = \frac{1}{n-1}\binom{2n-2}{n}.$$

Это также равно числу ступенчатых функций, графики которых ведут из (0, 0) в (n, n) и остаются ниже прямой с наклоном 45° (кроме граничных точек).

Следовательно, вероятность того, что график ступенчатой функции (рис. 4), выбранной случайным образом, не будет иметь общих точек, кроме граничных, с прямой с наклоном 45°, равна

$$2\frac{1}{n-1}\binom{2n-2}{n}\bigg/\binom{2n}{n} = \frac{1}{2n-1}. \qquad (3.10)$$

Это также вероятность, задаваемая формулами (3.7) и (3.9). Результат (3.10) также может быть выведен с помощью непосредственного приложения теоремы о баллотировке.

Наконец, результат (3.10) можно получить из результатов, касающихся бросания монеты и полученных в гл. 4. Соответствующим примером была первая из цепей Маркова, рассматривавшаяся в гл. 4, § 4. В нашей формулировке искомая вероятность является в точности вероятностью того, что первый переход из состояния 1 в состояние 0 произойдет при $(2n-1)$-м испытании при условии, что общее число выпадений герба равно общему числу выпадений решетки за $2n$ испытаний.

Безусловная вероятность того, что первый переход из состояния 1 в состояние 0 произойдет при $(2n-1)$-м испытании, вычислена в гл. 4, § 6. Было показано, что она равна $\frac{1}{2n-1}2^{-2n}\binom{2n}{n}$. Следовательно, условная вероятность такого первого перехода равна в точности $\frac{1}{\mu_{2n}}f_{00}^{2n}=\frac{1}{2n-1}$, как и должно быть.

§ 4. НЕКОТОРЫЕ ПРЕДЕЛЬНЫЕ РАСПРЕДЕЛЕНИЯ ДЛЯ ЭМПИРИЧЕСКИХ ФУНКЦИЙ РАСПРЕДЕЛЕНИЯ

Следующая лемма представляет интерес сама по себе и будет использована при нахождении предельного распределения для эмпирических функций распределения.

Лемма 4.1. *Пусть* $\mathbf{x}=(x_1, \ldots, x_n)$ — *вектор в n-мерном евклидовом пространстве. Предположим, что*

$$x_1+x_2+\ldots+x_n=0$$

и что $\sum_{k=i+1}^{j} x_k \neq 0$ *для* $1 \leqslant i \leqslant j \leqslant n$.

Пусть $x_{k+n}=x_k$, *и пусть* $\mathbf{x}(k)=(x_k, x_{k+1}, \ldots, x_n, \ldots, x_{n+k-1})$ — *циклическая перестановка компонент вектора* $\mathbf{x}$, $k=1, 2, \ldots, n$. *Тогда для любого* $r=0, 1, \ldots, n-1$ *существует ровно один из векторов* $\mathbf{x}(k)$, $k=1, \ldots, n$, *такой, что среди частичных сумм его компонент (взятых по порядку, начиная с первой) в точности r сумм положительны.*

Доказательство. Пусть $s_k=x_1+x_2+\ldots+x_k$, $k=1, \ldots, n$, и $s_0=0$, $s_{k+n}=s_k$, $k=0, 1, \ldots, n$. Тогда все s_k, $k=1, 2, \ldots, n$, различны, поскольку если

$$s_i=s_j \text{ для } 1 \leqslant i < j \leqslant n,$$

то $x_1+\ldots+x_i=x_1+\ldots+x_i+x_{i+1}+\ldots+x_j$, т. е. $x_{i+1}+\ldots+x_j=0$, что противоречит сделанным предположениям. Очевидно, частичные суммы взятых по порядку (начиная с первой) компонент вектора $\mathbf{x}(k)$ равны

$$s_k - s_{k-1},\ s_{k+1} - s_{k-1},\ \ldots,\ s_n - s_{k-1},\ s_{n+1} - s_{k-1},\ \ldots,\ s_{k+n-1} - s_{k-1}. \tag{4.1}$$

Эта последовательность статистически эквивалентна следующей:

$$s_k - s_{k-1},\ s_{k+1} - s_{k-1},\ \ldots,\ s_n - s_{k-1},\ s_1 - s_{k-1},\ \ldots,\ s_{k-1} - s_{k-1}. \tag{4.2}$$

Пусть теперь

$$s_1^*,\ s_2^*,\ \ldots,\ s_n^*$$

— единственная перестановка $s_1, s_2, \ldots, s_n$, для которой

$$s_1^* > s_2^* > \ldots > s_n^*.$$

Такая перестановка действительно существует (и единственна), поскольку все s_k $(k = 1, \ldots, n)$ различны.

Число положительных членов в последовательности (4.2) (или, что то же, в последовательности (4.1)) совпадает с числом положительных членов в последовательности

$$s_1^* - s_{k-1},\ s_2^* - s_{k-1},\ \ldots,\ s_n^* - s_{k-1}, \tag{4.3}$$

поскольку это попросту перестановка (4.2). Далее, для любого $r = 0, 1, \ldots, n-1$ будет существовать ровно одно число k $(1 \leqslant k \leqslant n)$, такое, что в точности r членов в (4.3) (а следовательно, и в (4.1)) будут положительны. Чтобы показать это, достаточно выбрать k так, чтобы

$$s_{k-1} = s_{r+1}^*.$$

Тогда $s_1^* - s_{k-1} > 0$, $s_2^* - s_{k-1} > 0$, $\ldots$, $s_r^* - s_{k-1} > 0$, а $s_{r+1}^* - s_{k-1} = 0$, $s_{r+2}^* - s_{k-1} < 0$, $\ldots$, $s_n^* - s_{k-1} < 0$. ■

Эта лемма имеет геометрическую интерпретацию. Отвлечемся от предположения $s_n = 0$. Отметим точки $(0, 0)$, $(1, s_1)$, $\ldots$, (n, s_n) на плоскости в декартовой системе координат. Соединим соседние точки отрезками прямых, как показано на рис. 6. Получившаяся ломаная называется суммарным полигоном вектора $\mathbf{x} = (x_1, \ldots, x_n)$. Точно таким же образом можно получить суммарные полигоны векторов $\mathbf{x}(k)$, компоненты которых являются циклическими перестановками компонент вектора $\mathbf{x}$. Отрезок прямой, соединяющий точки $(0, 0)$ и (n, s_n), называется хордой суммарного полигона. Рассмотрим точку пересечения P этой хорды с вертикальной прямой, проходящей через точку (k, s_k). Из элементарных геометрических соображений ясно, что ордината P равна $(k/n)s_n$. Следовательно, расстояние по вертикали от точки (k, s_k) до хорды полигона равно $s_k - (k/n)s_n$.

Далее, вектор с компонентами

$$(x_1 - s_n/n,\ x_2 - s_n/n,\ \ldots,\ x_n - s_n/n),$$

очевидно, удовлетворяет условиям леммы, включая и требование того, чтобы n-я частичная сумма была равна нулю. Лемма утверждает, что среди суммарных полигонов, соответствующих n циклическим перестановкам вектора $\mathbf{x}$, для любого $r = 0, 1, \ldots, n-1$ существует ровно один, у которого в точности r вершин находятся

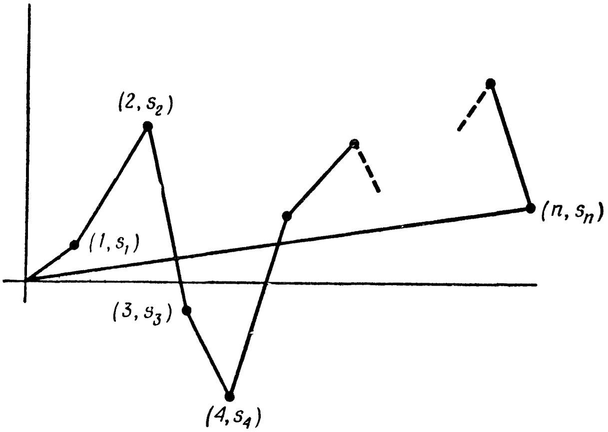


Рис. 6.

выше хорды. В частности, для случая $r = 0$ нужно взять циклическую перестановку, начинающуюся с номера $1 + k_0$, где k_0 — индекс, при котором достигается $\max\limits_k (s_k - (k/n) s_n)$. Случай $r = n - 1$ соответствует номеру $1 + k_1$, где k_1 — индекс, при котором достигается $\min\limits_k (s_k - (k/n) s_n)$, и т. д.

В заключение главы укажем на применение леммы 4.1 к анализу некоторых д. с. в., связанных с эмпирическими функциями распределения. Пусть, как обычно, $F_n(x)$ — эмпирическая функция распределения, построенная по выборке объема n из равномерного распределения на $[0, 1]$. Введем две случайные величины U_n и V_n. Положим

$$U_n \doteq \{\text{общая длина всех интервалов на оси } x, \text{ на которых } F_n(x) > x\} \tag{4.4}$$

(рис. 7). Для реализации, показанной на рисунке, U_n равна сумме длин всех отрезков, отмеченных жирной линией.

Определим

$$V_n = \inf\{x \mid F_n(x) - x = \max_{0 \leqslant x \leqslant 1} [F_n(x) - x]\}. \tag{4.5}$$

Величина $\max\limits_{0 \leqslant x \leqslant 1} [F_n(x) - x] = D_n^+$, обычно называемая односторонней статистикой Колмогорова — Смирнова, является важной характеристикой наблюдения, используемой в статистических тестах при решении, принадлежит ли исследуемая выборка заданному равномерному распределению.

Наша цель в данном параграфе — определить закон распределения величин U_n и V_n. Замечательным результатом является то,

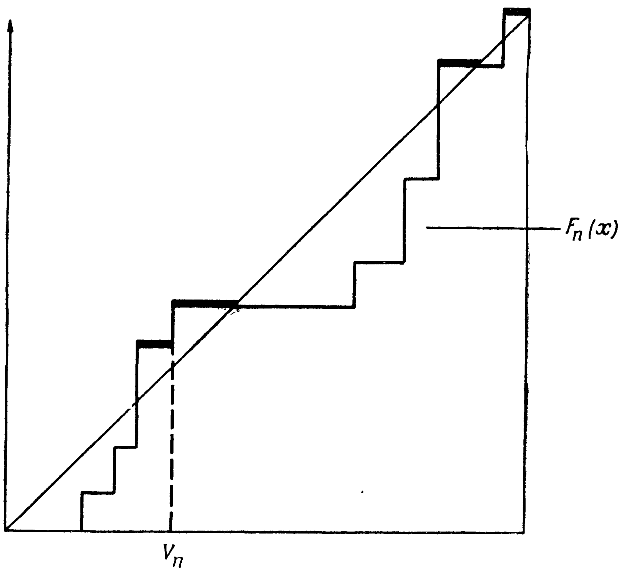


Рис. 7.

что при любом n любая из этих величин распределена равномерно на отрезке [0, 1].

Чтобы доказать это утверждение, рассмотрим пуассоновский процесс $X(t)$, $0 \leqslant t \leqslant 1$, с параметром 1. Далее, разделим интервал (0, 1) на $r+1$ частей:

$$\left(0, \frac{1}{r+1}\right), \left(\frac{1}{r+1}, \frac{2}{r+1}\right), \ldots, \left(\frac{r}{r+1}, 1\right),$$

где $r+1$ — простое число, большее n. (Причина этому будет ясна из дальнейшего.) Приращения

$$X\left(\frac{1}{r+1}\right), \; X\left(\frac{2}{r+1}\right) - X\left(\frac{1}{r+1}\right), \; \ldots, \; X(1) - X\left(\frac{r}{r+1}\right)$$

— независимые и одинаково распределенные по закону Пуассона случайные величины. Обозначим эти приращения через W_1, W_2, ...

..., W_{r+1} соответственно и определим случайные величины

$$Y_i = W_i - \frac{n}{r+1}, \qquad i = 1, \ldots, r+1,$$

которые, очевидно, независимы и одинаково распределены. Образуем последовательность частичных сумм

$$S_k = Y_1 + \ldots + Y_k, \qquad k = 1, 2, \ldots, r+1,$$

и заметим, что

$$\mathrm{P}\{S_i = 0\} = 0, \qquad i = 1, 2, \ldots, r. \tag{4.6}$$

В самом деле, из равенства $S_i = 0$ следует, что $(r+1)\, X\left(\frac{i}{r+1}\right) = ni$. Но этого не может быть в силу предположения о том, что $r+1$ больше n, i и не делится ни на n, ни на i, поскольку $r+1$ — простое число[1]). Следовательно, равенство (4.6) доказано.

Аналогично можно показать, что $\mathrm{P}\{S_j - S_i = 0\} = 0$ для любых $j > i$. Это означает, что с вероятностью 1 никакая из частичных сумм $S_j - S_i$ $(0 \leqslant i < j < r+1,\ S_0 = 0)$ не равна нулю.

Пусть

N_r — число положительных членов в последовательности S_1, S_2, ..., S_r,

L_r — наименьший индекс j, для которого $S_j = \max(0, S_1, \ldots, S_r)$.

Для последовательности $x_i = S_i - S_{i-1}$, $i = 1, \ldots, r$, выполнены гипотезы леммы 4.1, если $S_j - S_i \neq 0$ $(j > i)$ и $S_{r+1} \neq 0$. Но указанные события имеют вероятность 1. В соответствии с леммой

$$\mathrm{P}\{L_r = m \mid S_{r+1} = 0\} = \mathrm{P}\{N_r = m \mid S_{r+1} = 0\} = \frac{1}{r+1} \quad (m = 0, 1, \ldots, r). \tag{4.7}$$

Если $X(1) = n$, то $X(t)/n$ распределена по закону $F_n(t)$, $0 \leqslant t \leqslant 1$. (Это следует из результатов § 1.) Следовательно, можно определить величины U_n, V_n для $\frac{X(t)}{n}$, $0 \leqslant t \leqslant 1$, при условии, что $X(1) = n$. Утверждается, что

$$\left| U_n - \frac{N_r}{r+1} \right| < \frac{A}{r+1}, \quad \left| V_n - \frac{L_r}{r+1} \right| < \frac{B}{r+1}, \tag{4.8}$$

где постоянные A и B зависят от n, но не от r.

Первое из неравенств (4.8) доказывается следующим образом. Если

$$S_{k+1} = X\left(\frac{k+1}{r+1}\right) - n\,\frac{k+1}{r+1} > 0$$

[1]) Следует напомнить, что $i \leqslant r < r+1$, а случайная величина X может принимать лишь целочисленные неотрицательные значения. — *Прим. перев.*

и если $X(t)$ не имеет скачка в интервале $\left(\frac{k}{r+1}, \frac{k+1}{r+1}\right)$, то при любом t, $\frac{k}{r+1} \leqslant t \leqslant \frac{k+1}{r+1}$,

$$X(t) - nt > 0. \tag{4.9}$$

Поскольку $X(t)$ имеет ровно n скачков (при условии $X(1) = n$), то существует самое большее n интервалов длиной не более $1/(r+1)$ каждый, для которых из условия $S_i > 0$ не следует справедливость неравенства

$$X(t) - nt > 0, \quad i/(r+1) \leqslant t \leqslant i + 1(r+1).$$

Но величина $N_r/(r+1)$ равна числу положительных S_i, умноженному на длину $1/(r+1)$. Учитывая сказанное выше, можно заключить, что $N_r/(r+1)$ может отличаться от U_n не более чем на $n/(r+1)$. Таким образом, $A \leqslant n$. Аналогично можно получить второе соотношение (4.8). Таким образом, при условии $X(1) = n$ обе абсолютные величины в (4.8) стремятся по вероятности к нулю при $r \to \infty$ [1]). Так как $N_r/(r+1)$ и $L_r/(r+1)$ распределены асимптотически равномерно на отрезке $[0, 1]$ при $r \to \infty$ ($r+1$ стремится к ∞, принимая лишь простые значения), это завершает доказательство следующей теоремы.

Теорема 4.1. *Пусть $F_n(x)$ — эмпирическая функция распределения, построенная по выборке объема n из равномерного распределения на отрезке* $[0, 1]$. *Рассмотрим случайные величины U_n и V_n, определенные формулами* (4.4) *и* (4.5). *Тогда*

$$\mathrm{P}\{U_n \leqslant x\} = \mathrm{P}\{V_n \leqslant x\} = x, \qquad 0 \leqslant x \leqslant 1.$$

ЗАДАЧИ

1. Пусть $X_1, X_2, \ldots, X_n$ — независимые одинаково распределенные случайные величины с непрерывной функцией распределения $F(x)$. Обозначим k-е по величине из чисел

$$X_1, X_2, \ldots, X_n$$

через X_{nk}; таким образом,

$$X_{n1} \leqslant X_{n2} \leqslant \ldots \leqslant X_{nn}.$$

Найти функцию распределения $F_{nk}(x)$ величины X_{nk}.

Ответ:

$$F_{nk}(x) = \sum_{l=k}^{n} \binom{n}{l} [F(x)]^l [1 - F(x)]^{n-l}.$$

---

[1]) Из (4.8) следует, что указанные величины стремятся к 0 с вероятностью 1. — *Прим. перев.*

2. При обозначениях задачи 1 показать, что

$$F_{nk}(x) = k\binom{n}{k}\int\limits_{1-F(x)}^{1} t^{n-k}(1-t)^{k-1}\,dt.$$

Указание: Взять интеграл по частям.

3. При обозначениях задачи 1 найти

$$\mathrm{P}\{X_{nk} > y,\ X_{n+1,k} \leqslant x\} \quad \text{для } x \leqslant y.$$

Ответ:

$$\binom{n}{k-1}[F(x)]^k[1-F(y)]^{n-k+1}.$$

4. Пусть X_i^*, $i = 1, 2, \ldots, n$, — порядковые статистики для равномерного на отрезке $[0, 1]$ распределения. Показать, что $\ln X_k^*$ имеет такое же распределение, как и $-\sum\limits_{j=k}^{n}\theta_j/j$, где θ_j — независимые случайные величины с экспоненциальным распределением с параметром 1.

Указание: Использовать соответствие между пуассоновскими процессами и порядковыми статистиками для равномерного распределения.

5. Доказать, что $\left(X_i^*/X_{i+1}^*\right)^i$, $i = 1, 2, \ldots, n$, где по определению $X_{n+1}^* = 1$, — независимые равномерно распределенные на $[0, 1]$ случайные величины. Здесь $X_1^*, \ldots, X_n^*$ — порядковые статистики, соответствующие выборке объема n из равномерного распределения.

6. Пусть $X_1, X_2, \ldots, X_n$ — независимые выборки из равномерного на отрезке $[0, 1]$ распределения. Найти распределение случайной величины $P = \prod\limits_{i=1}^{n} X_i$.

Указание: Либо вычислить непосредственно, либо ввести новые переменные

$$X_i = \exp(-Y_i), \qquad i = 1, 2, \ldots, n.$$

Ответ:

$$\mathbf{P}\{P \leqslant p\} = \int\limits_0^p \int\limits_t^1 \int\limits_{\zeta_2}^1 \cdots \int\limits_{\zeta_{n-1}}^1 \frac{d\zeta_2\, d\zeta_3 \ldots d\zeta_n}{\zeta_2\zeta_3 \ldots \zeta_n}\,dt = \frac{(-1)^n}{n!}\int\limits_0^p t(\ln t)^n\,dt.$$

7. Пусть $X_1, X_2, \ldots, X_k$ — независимые одинаково распределенные случайные величины с функцией распределения $F(x)$ и плотностью $f(x)$. Пусть $X_1^* \leqslant X_2^* \leqslant \ldots \leqslant X_k^*$ — соответствующие порядковые статистики. Пусть $Y_1, Y_2, \ldots$ — независимые одинаково распределенные случайные величины с распределением $F(x)$. Определим целочисленную случайную величину N соотношением

$$Y_i \leqslant X_k^*, \quad i = 1, 2, \ldots, N-1,$$

но $Y_N > X_k^*$. Аналогично определим M как случайную величину, для которой

$$Y_i \leqslant X_{k-1}^*, \qquad i = 1, 2, \ldots, M-1,$$

но

$$Y_M > X^*_{k-1}.$$

Найти распределения случайных величин N и M.

Ответ:

$$P\{N=n\} = \frac{k}{(n+k)(n+k-1)},$$

$$P\{M=n\} = \frac{2k(k-1)}{(n+k)(n+k-1)(n+k-2)}.$$

Задачи 8—15 *имеют дело со следующими объектами.*

Пусть $X_1, \ldots, X_{n-1}$ — независимые случайные величины, каждая из которых равномерно распределена на отрезке [0, 1]. Пусть

$$X^*_1 \leqslant X^*_2 \leqslant \ldots \leqslant X^*_{n-1}$$

— упорядоченные значения величин X_i. Эти порядковые статистики разбивают единичный отрезок на n непересекающихся интервалов, имеющих длины

$$U_1 = X^*_1,\ U_2 = X^*_2 - X^*_1,\ \ldots,\ U_n = 1 - X^*_{n-1}.$$

Пусть

$$U^*_1 \leqslant U^*_2 \leqslant \ldots \leqslant U^*_n$$

— упорядоченные значения величин U_i.

В тексте этой главы установлены следующие результаты.

(А) Случайный вектор

$$(Y_1/S_n,\ Y_2/S_n,\ \ldots,\ Y_n/S_n)$$

(где Y_i — независимые экспоненциально распределенные д. с. в. с параметром 1) и д. с. в. $S_n = Y_1 + \ldots + Y_n$ независимы.

(Б) Случайные векторы $(U_1, \ldots, U_n)$ и

$$(Y_1/S_n,\ Y_2/S_n,\ \ldots,\ Y_n/S_n)$$

одинаково распределены.

(В) Если S — случайная величина, распределенная, как S_n, и независимая от вектора $(U_1, U_2, \ldots, U_n)$, то случайные векторы $(SU_1, SU_2, \ldots, SU_n)$ и $(Y_1, \ldots, Y_n)$ распределены одинаково.

8. Найти $M\left(U_1^{i_1} U_2^{i_2} \ldots U_n^{i_n}\right)$, где $i_1, i_2, \ldots, i_n$ — неотрицательные целые числа.

Ответ:

$$\frac{n!\, i_1! \ldots i_n!}{(n+i_1+\ldots+i_n)!}.$$

9. Доказать, что

$$X^*_1/X^*_2,\ X^*_2/X^*_3,\ \ldots,\ X^*_n/1$$

— независимые случайные величины.

Указание: См. задачу 5.

10. Пусть $S_k = Y_1 + \ldots + Y_k$. Показать, что

$$S_1/S_2,\ S_2/S_3,\ \ldots,\ S_{n-1}/S_n$$

— независимые случайные величины.

11. Если i, j, k и l — различные индексы, то показать, что

$$U_i/U_j,\ U_k/U_l$$

— независимые случайные величины.

12. Найти распределение величины

$$\frac{U_1 + \ldots + U_r}{U_{r+1} + \ldots + U_{r+s}}.$$

Ответ: Отношение двух независимых д. с. в., каждая из которых имеет гамма-распределение.

13. Показать, что

$$F(t_1, \ldots, t_n) = \mathrm{P}\{U_1 > t_1,\ U_2 > t_2,\ \ldots,\ U_n > t_n\} = \begin{cases} 1, & \text{если } t < 0, \\ (1-t)^n, & \text{если } 0 \leqslant t \leqslant 1, \\ 0, & \text{если } t > 1, \end{cases}$$

где $t = t_1 + \ldots + t_n$.

14. Пусть $Y_1^* \leqslant Y_2^* \leqslant \ldots \leqslant Y_n^*$ — упорядоченные значения величин $Y_1, \ldots, Y_n$. Найти совместное распределение величин

$$Z_1 = nY_1^*,\ Z_2 = (n-1)\left(Y_2^* - Y_1^*\right),\ \ldots,\ Z_n = \left(Y_n^* - Y_{n-1}^*\right).$$

Указание: См. задачу 5.

Ответ: Независимые экспоненциально (с параметром 1) распределенные случайные величины.

15. Показать, что векторы $(U_1, U_2, \ldots, U_n)$ и $\left(nU_1^*,\ (n-1)\left(U_2^* - U_1^*\right), \ldots, U_n^* - U_{n-1}^*\right)$ имеют одно и то же совместное распределение.

16. Пусть n точек выбраны на отрезке прямой длины L «случайным образом» (т. е. в соответствии с равномерным распределением). Показать, что если $0 < d < L/(n-1)$, то вероятность того, что никакие две точки не будут расположены ближе друг к другу, чем на расстояние d, равна $([L-(n-1)d]/L)^n$.

Указание: Установить и интерпретировать соотношение

$$\frac{n!}{L^n} \int\limits_{(n-1)d}^{L} dx_n \int\limits_{(n-2)d}^{x_n - d} dx_{n-1} \ldots \int\limits_{2d}^{x_4 - d} dx_3 \int\limits_{d}^{x_3 - d} dx_2 \int\limits_{0}^{x_2 - d} dx_1 = \frac{(L-(n-1)d)^n}{L^n}.$$

*17. Предположим, что n точек выбраны независимо друг от друга в соответствии с равномерным распределением на окружности. Показать, что вероятность того, что длины фиксированных j дуг (из образованных с помощью указанных точек n дуг) будут больше α, равна

$$\pi_j = \begin{cases} (1 - j\alpha/t)^{n-1}, & \text{если } 0 \leqslant j\alpha < t, \\ 0 & \text{в противном случае,} \end{cases}$$

где t — длина окружности.

Указание: В силу круговой симметрии выбрать любую из n точек за начало отсчета и предположить, что остальные $n-1$ точек были выбраны случайным образом на отрезке $[0, 1]$ независимо друг от друга. Пусть

$$0 \leqslant X_1^* \leqslant X_2^* \leqslant \ldots \leqslant X_{n-1}^* \leqslant t$$

— порядковые статистики, отмечающие расстояния от начала отсчета до первой, второй, ..., $(n-1)$-й точки соответственно. Использовать результаты задачи 16.

*18. В задаче 17 показать, что вероятность того, что в точности k из n дуг между соседними точками будут иметь длины, превышающие α, равна

$$P_k = \binom{n}{k} \sum_{j=k}^{[t/\alpha]} (-1)^{j-k} \binom{n-k}{n-j} \left(1 - \frac{j\alpha}{t}\right)^{n-1}.$$

Указание: Пусть V_k — вероятность того, что из фиксированных k дуг каждая длиннее α, а оставшиеся $n-k$ дуг короче α. Заметить, что

$$P_k = \binom{n}{k} V_k, \qquad k = 0, 1, \ldots, n.$$

Затем установить формулу

$$\pi_j = \sum_{k=j}^{n} \binom{n-j}{n-k} V_k,$$

где вероятность π_j определена в задаче 17. Показать, что, решив эти уравнения, получим

$$V_k = \sum_{j=k}^{n} (-1)^{j-k} \binom{n-k}{n-j} \pi_j.$$

$\left(\text{Напомним, что } \binom{n}{r} = 0 \text{ при } r > n.\right)$

19. Пусть n импульсов поступают на счетчик в моменты $t_1, \ldots, t_n$, где $t_1, t_2, \ldots, t_n$ распределены как порядковые статистики для равномерного распределения на отрезке $[0, 1]$. Всякий раз после того, как счетчик зарегистрирует импульс, у него имеется «мертвый период» длительности τ, в который он не регистрирует импульсы, даже если они на него и поступают. Интервал $(0, \tau)$ также считается «мертвым периодом». Найти вероятность того, что счетчик зарегистрирует первые k импульсов, которые на него поступают (т. е. $t_i - t_{i-1} \geqslant \tau$, $i = 1, 2, \ldots, k$, $t_0 = 0$).

Указание: Использовать метод решения задачи 16.

Ответ:

$$\begin{cases} (1-k\tau)^n, & \text{если } k\tau \leqslant 1, \\ 0, & \text{если } k\tau > 1. \end{cases}$$

20 (продолжение). В задаче 19 найти плотность $f(y)$ д. с. в. Y, которая равна моменту поступления n-го импульса, при условии, что счетчик регистрирует первые k поступивших импульсов $(n > k)$.

Ответ:

$$f(y) = \frac{n(y-k\tau)^{n-1}}{(1-k\tau)^n}, \qquad k\tau \leqslant y \leqslant 1.$$

21. Используя обозначения задачи 1, скажем, что ранг величины X_j в множестве $X_1, X_2, \ldots, X_n$ равен r, если $X_j = X_{nr}$ (см. задачу 1). Далее, пусть R_j — ранг величины X_j в множестве $X_1, X_2, \ldots, X_j$. Показать, что случайные величины $R_1, R_2, \ldots, R_n$ независимы и

$$\mathrm{P}\{R_n = r\} = 1/n, \qquad r = 1, 2, \ldots, n.$$

Указание: Доказать, что

$$\mathrm{P}\{R_1 = r_1,\ R_2 = r_2,\ \ldots,\ R_n = r_n\} = 1/n!,$$
$$r_1 = 1;\ r_2 = 1,\ 2;\quad r_3 = 1,\ 2,\ 3;\ \ldots;\ r_n = 1,\ 2,\ \ldots,\ n.$$

22. Пусть $X_1, X_2, \ldots$ — независимые одинаково распределенные положительные случайные величины с непрерывной функцией распределения $F(x)$. Доказать, что

$$\mathrm{P}\{X_k > \max(X_1, X_2, \ldots, X_{k-1})\} = 1/k.$$

23. Пусть $X_1, X_2, \ldots$ — независимые одинаково распределенные д. с. в.; определим

$$M_n = \max(X_1, X_2, \ldots, X_n).$$

Пусть $N_n(\lambda, \mu)$ $(0 < \lambda < \mu)$ — число членов в последовательности

$$X_{[n\lambda]},\ X_{[n\lambda]+1},\ X_{[n\lambda]+2},\ \ldots,\ X_{[n\mu]},$$

таких, что $X_i = M_i$. Здесь, как обычно, символ $[\cdot]$ означает взятие целой части. Показать, что вероятностная производящая функция величины $N_n(\lambda, \mu)$ равна

$$\mathrm{M}\left(z^{N_n(\lambda,\,\mu)}\right) = \exp\left\{\sum_{i=[\lambda n]}^{[\mu n]} \ln\left(1 - \frac{1-z}{i}\right)\right\} = \prod_{i=[\lambda n]}^{[\mu n]} \left(1 - \frac{1-z}{i}\right).$$

Указание. Ввести д. с. в.

$$C(i) = \begin{cases} 1, & \text{если}\quad X_i = M_i, \\ 0 & \text{в противном случае}. \end{cases}$$

Используя задачу 21, показать, что $C(i)$ — независимые случайные величины, и затем найти

$$\prod_{i=[n\lambda]}^{[n\mu]} \mathrm{M}\left(z^{C(i)}\right).$$

24 (продолжение). Показать, что предельное распределение величины $N_n(\lambda, \mu)$ при $n \to \infty$ является пуассоновским с параметром $\ln(\mu/\lambda)$.

25. Пусть $W_{1,\,n\mu}$ — индекс последнего максимального члена последовательности $X_1, X_2, \ldots, X_{[n\mu]}$. Доказать, что величина $\dfrac{1}{n\mu} W_{1,\,n\mu}$ имеет предельное при $n \to \infty$ распределение, являющееся равномерным на отрезке $[0, 1]$.

26. Пусть $W_1, W_2, \ldots$ — независимые и равномерно распределенные на $[0, 1]$ д. с. в. Пусть N — число индексов i, удовлетворяющих соотношению $t < \prod_{j=1}^{i} W_j < 1$, где t — фиксированное число, $0 < t < 1$. Найти функцию распределения величины N.

Ответ: N имеет пуассоновское распределение с параметром $-\ln t$.

27. Пусть $X_1, X_2, \ldots, X_n$ суть n наблюдений случайной величины с непрерывной функцией распределения $F(x)$. Пусть N_n — число индексов k, для которых $X_k > \max(X_1, X_2, \ldots, X_{k-1})$, где $k = 1, 2, \ldots, n$. Показать, что производящая функция для N_n равна

$$\sum_{r=1}^{n} \mathrm{P}\{N_n = r\}\, s^r = \binom{s+n-1}{n}.$$

28. В предыдущей задаче установить соотношение

$$\lim_{n\to\infty} \frac{M(N_n)}{\ln n} = 1.$$

29. Пусть $Z_i = (X_i, Y_i)$, $i = 1, \ldots, n$, суть n пар действительных случайных величин, причем все из них независимы и одинаково распределены с непрерывной функцией распределения $F(x)$. Говорят, что вектор $Z_i = (X_i, Y_i)$ — *допустимый*, если не существует другого $Z_j = (X_j, Y_j)$, у которого $X_j \geqslant X_i$, $Y_j \geqslant Y_i$. Пусть I_n — число допустимых векторов в последовательности $Z_1, Z_2, \ldots, Z_n$.

Без потери общности предположим, что Z_i перенумерованы в таком порядке, что

$$X_1 \leqslant X_2 \leqslant \ldots \leqslant X_n.$$

Доказать, что

$$I_n = \sum_{i=1}^{n} U_i,$$

где U_i — *независимые* случайные величины, определяемые равенством

$$U_i = \begin{cases} 1, & \text{если ранг } Y_i \text{ в множестве } Y_i, Y_{i+1}, \ldots, Y_n \text{ равен } n-i+1, \\ 0 & \text{в противном случае.} \end{cases}$$

Указание. Использовать результат задачи 21.

30. При условиях задачи 29 найти вероятностную производящую функции $g(t)$ величины I_n (см. задачу 27).

Ответ:

$$g(t) = \frac{1}{n!}\,\frac{\Gamma(t+n)}{\Gamma(t)}.$$

31. Пусть $X_1, X_2, \ldots$ — последовательность независимых одинаково распределенных случайных величин с функцией распределения $F(x)$ и плотностью $f(x)$. Пусть n_1 — целочисленная случайная величина, равная наименьшему индексу, такому, что $X_{n_1} > X_1$. Далее, пусть n_2 — наименьший индекс ($>n_1$), такой, что $X_{n_2} > X_{n_1}$, и вообще пусть n_r — наименьший индекс ($>n_{r-1}$), для которого $X_{n_r} > X_{n_{r-1}}$. Найти функцию распределения величины X_{n_r}.

Указание: Накладывая условие на значения величин $X_1, X_{n_1}, X_{n_2}, \ldots, X_{n_{r-1}}$, доказать формулу

$$P\left\{X_{n_r} \leqslant z\right\} = \\ = \int_{-\infty}^{z} f(x_r)\,dx_r \int_{-\infty}^{x_r} \frac{f(x_{r-1})\,dx_{r-1}}{1-F(x_{r-1})} \int_{-\infty}^{x_{r-1}} \frac{f(x_{r-2})\,dx_{r-2}}{1-F(x_{r-2})} \ldots \int_{-\infty}^{x_1} \frac{f(x_0)\,dx_0}{1-F(x_0)}.$$

Затем упростить ее. Другой путь решения — использование индукции по r.

Ответ: Плотность величины X_{n_r} равна

$$f_{X_{n_r}}(z) = \frac{[-\ln(1-F(z))]^r}{r!} f(z), \qquad z > 0.$$

32. Пусть $X_1, X_2, \ldots$ — последовательность независимых одинаково распределенных случайных величин с функцией распределения $F(x)$ и плотностью $f(x)$. Пусть n_i — д. с. в., определенные в задаче 31. Найти распределение случайной величины $n_r - n_{r-1}$.

Ответ:

$$P\{n_r - n_{r-1} = i\} = \sum_{k=0}^{i-1} (-1)^k \binom{i-1}{k} \frac{1}{(k+2)^r}.$$

Задачи 33—37 являются вариантами задачи о баллотировке.

33. В качестве модификации задачи о баллотировке, приведенной в тексте, найти вероятность того, что A никогда не будет отставать при подсчете голосов.

Указание: Рассмотрим более общую урновую схему, когда в урне имеется c карт, на которых написаны числа $k_1, k_2, \ldots, k_c$. Пусть $n_0, n_1, n_2, \ldots$ — количество нулей, единиц, двоек, ... среди $k_1, k_2, \ldots, k_c$. Предположим, что $k_1 + \ldots + k_c = k$ $(0 \leqslant k \leqslant c)$. Пусть ν_i $(i = 1, 2, \ldots, c)$ — д. с. в., значение которой равно числу на i-й вынутой карте. За один раз вынимается одна карта; обратно в урну она не возвращается. Обозначим

$$P_j(c, k; n_0, n_1, \ldots) = P\{\nu_1 + \ldots + \nu_r < r + j, \ r = 1, 2, \ldots, c\}.$$

Поместив в урну дополнительную карту с числом 0, получим рекуррентную формулу

$$P_{j-1}(c+1, k; n_0+1, n_1, n_2, \ldots) = \frac{n_0+1}{c+1} P_j(c, k; n_0, n_1, \ldots) +$$
$$+ \sum_{i=1}^{j-1} \frac{n_i}{c+1} P_{j-i}(c, k-i; n_0+1, n_1, \ldots, n_{i-1}, n_i - 1, n_{i+1}, \ldots).$$

С помощью метода, использованного в тексте (где $n_0 = a$, $n_2 = b$, $a + b = c$), показать, что вероятность того, что A никогда не будет отставать при подсчете голосов, равна

$$P\{\nu_1 + \ldots + \nu_r < r + 1, \ r = 1, 2, \ldots, c\}.$$

Ответ: $(a - b + 1)/(a + 1)$.

34. Рассмотрим урну, содержащую a белых и b черных шаров $(a > b)$. Последовательно вынимаются все шары случайным образом без возвращения. Одновременно производится случайное блуждание по неотрицательным целым числам, которое начинается с 1. Направление шага (вправо или влево на 1) зависит соответственно от того, белый или черный очередной шар вынут из урны. Чему равна вероятность того, что за $a + b$ шагов хотя бы один раз будет достигнута точка 0?

Ответ:

$$b/(a+1).$$

35. Пусть a и b — положительные целые числа, $a > b$. Доказать следующие равенства:

$$(1)\ \frac{b}{a+b} +$$
$$+ \sum_{k=1}^{b-1} (2k+1)^{-1} \binom{2k+1}{k} \frac{a(a-1)\ldots(a-k+1)\, b(b-1)\ldots(b-k)}{(a+b)(a+b-1)\ldots(a+b-2k)} = \frac{b}{a+1};$$

$$(2)\ \frac{a-b}{a+b} +$$
$$+ \sum_{r=1}^{b} \frac{a(a-1)\ldots(a-r+1)\, b(b-1)\ldots(b-r+1)}{(a+b)(a+b-1)\ldots(a+b-2r+1)} \binom{2r}{r} \frac{a-b}{a+b-2r} = \frac{b}{a+1}.$$

Указание: (1) Рассмотреть ситуацию задачи 34. Выразить событие, заключающееся в достижении начала координат, через момент первого достижения начала координат.

(2) Это равенство получается при рассмотрении последнего момента, в который количества вынутых белых и черных шаров равны.

Должны быть использованы результаты задачи о баллотировке, задачи 34 и вывод формулы (6.12) в гл. 4.

36. При условиях задачи 34 доказать, что вероятность достижения начала координат равна

$$p_n(a,\ b)=\frac{b\,(b-1)\ldots(b-n+1)}{(a+1)\,(a+2)\ldots(a+n)},$$

если случайное блуждание начинается в точке n. Здесь n — положительное целое число, $n \leqslant b$.

Указание: Найти рекуррентную формулу для $p_n(a, b)$. Затем использовать индукцию по n совместно с результатом задачи 34, т. е. равенством $p_1(a, b) = b/(a+1)$.

37. Пусть $F_n(x)$ — эмпирическая функция распределения, соответствующая n наблюдениям из равномерного распределения на $[0, 1]$. Найти вероятность

$$P_n(a,\ \gamma)=\mathrm{P}\,\{F_n(x)<a+\gamma x \quad \text{при всех} \quad 0\leqslant x\leqslant 1\},$$

где $(n-1)/n \leqslant a \leqslant 1$, $\gamma > 0$, $\gamma + a > 1$.

Ответ: $1-[(1-a)/\gamma]^n$.

38. Пусть $a < (n-1)/n$, а все остальные условия — те же, что и в задаче 37. Доказать, что

$$P_n(a,\ \gamma)=\int\limits_{(1-a)/\gamma}^{1} P_{n-1}\left(\frac{n}{n-1}\,a,\ \frac{n}{n-1}\,\gamma t\right) n t^{n-1}\,dt.$$

Указание: Наложить условие на максимальное значение реализации случайной величины и использовать формулу полной вероятности.

39. Пусть

$$C_n(a,\ \gamma,\ i)=\binom{n}{i}\left(\frac{n-i}{n\gamma}-\frac{a}{\gamma}\right)^{n-i}\left(1-\left(\frac{n-i}{n\gamma}-\frac{a}{\gamma}\right)\right)^{i-1}\left(\frac{\gamma+a-1}{\gamma}\right).$$

Показать, что

$$P_n(a,\ \gamma)=1-\sum_{i=0}^{k} C_n(a,\ \gamma,\ i),$$

где целое число k определяется соотношением

$$k/n\leqslant 1-a<(k+1)/n.$$

Указание: Рассмотрим вспомогательное событие, заключающееся в том, что $F_n(x) \geqslant a+\gamma x$ при некотором x. Оно происходит тогда и только тогда, когда

$$F_n\left(\frac{n-r}{n\gamma}-\frac{a}{\gamma}\right)\geqslant\frac{n-r}{n} \quad \text{при некотором } r=1,\ 2,\ \ldots,\ k.$$

Введем событие A_i, заключающееся в том, что минимальный индекс r, для которого выполняется неравенство

$$F_n\left(\frac{n-r}{n\gamma}-\frac{a}{\gamma}\right)\geqslant\frac{n-r}{n},$$

равен i $(i=0,\ 1,\ \ldots,\ k)$. Найти $\sum_{i=0}^{k} \mathrm{P}\{A_i\}$.

Другим методом решения является использование результата задачи 37 и проведение индукции по n.

ЗАМЕЧАНИЯ

В этой главе мы придерживаемся точки зрения Реньи [1] на порядковые статистики и пуассоновские процессы.

Комбинаторные методы, примененные во второй половине главы, в основном базируются на работе Такача, которая пока еще не вышла в виде книги[1]).

Подробное обсуждение порядковых статистик с точки зрения классической статистики можно найти в книге Уилкса [2]. См. также приводимые там ссылки.

ЛИТЕРАТУРА

1. Rényi A., Wahrscheinlichkeitsrechnung, mit einem Anhang über Informationstheorie, Deutscher Verlag der Wissenschaften, Berlin, 1962.
2. Уилкс С., Математическая статистика, «Наука», М., 1967.

[1]) Монография Л. Такача «Комбинаторные методы в теории случайных процессов» вышла в оригинале практически одновременно с книгой Карлина; русский ее перевод вышел в 1971 г. (изд-во «Мир»). — *Прим. ред.*

Глава 10

БРОУНОВСКОЕ ДВИЖЕНИЕ

§ 1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

Математический анализ некоторых специальных классов случайных процессов развит весьма детально. Наиболее известен процесс броуновского движения, который исторически был первым среди досконально исследованных процессов. Мы лишь слегка коснемся некоторых из наиболее интересных его особенностей и надеемся вызвать у читателя интерес к элегантной и совершенной теории этого процесса.

Броуновское движение как физическое явление было открыто английским ботаником Броуном в 1827 году. Математическое описание этого явления было выведено из законов физики Эйнштейном в 1905 году. С тех пор в этой области отмечен значительный прогресс. Физическая теория была далее усовершенствована Смолуховским, Фоккером, Планком, Бюргером, Ферсом, Орнштейном, Уленбеком, Чандрасекаром, Крамерсом и другими. Математическая теория развивалась медленнее, потому что точное математическое описание модели связано с рядом трудностей, тогда как некоторые из вопросов, на которые физики получили ответы из данной модели, были весьма простыми и интуитивно ясными. Многие из ответов были получены эвристическим путем Башелье в его книге (1912 г.), тогда как первое математически четкое построение теории было дано Винером в его диссертации (1918 г.) и более поздних работах (см. ссылки в конце данной главы).

В терминах нашей общей классификации случайных процессов процесс броуновского движения является примером марковского процесса с непрерывным временем и непрерывным пространством состояний. Мы обсудим только одномерный случай.

Пусть $X(t)$ — положение частицы (как функция времени) в броуновском движении (см. стр. 22). Пусть x_0 — положение частицы в момент t_0, т. е. $X(t_0) = x_0$, а $p(x, t \mid x_0)$ — условная плотность вероятности величины $X(t + t_0)$ при условии, что $X(t_0) = x_0$. Мы постулируем, что вероятностный закон, «управляющий» переходами, стационарен во времени (поскольку $p(x, t \mid x_0)$ не зависит от начального момента t_0).

Поскольку $p(x, t \mid x_0)$ — плотность вероятности, то

$$p(x, t \mid x_0) \geqslant 0, \qquad \int_{-\infty}^{\infty} p(x, t \mid x_0)\, dx = 1. \tag{1.1}$$

Далее, предположим, что $X(t+t_0)$ при малых t с большой вероятностью находится вблизи точки $X(t_0)=x_0$. Это достигается условием

$$\lim_{t\to 0} p(x,\ t\,|\,x_0)=0 \quad \text{при} \quad x\neq x_0. \tag{1.2}$$

Исходя из физических законов, Эйнштейн показал, что функция $p(x,t|x_0)$ должна удовлетворять уравнению в частных производных

$$\frac{\partial p}{\partial t}=D\,\frac{\partial^2 p}{\partial x^2}, \tag{1.3}$$

которое называется уравнением диффузии; D является коэффициентом диффузии. Малые частицы совершают броуновское движение благодаря столкновениям с молекулами окружающих их газа или жидкости. Величина D находится из формулы $D=2RT/Nf$, где R — газовая постоянная, T — температура, N — число Авогадро, f — коэффициент трения. Выбирая соответствующий масштаб, можно получить $D=1/2$. Тогда непосредственной проверкой устанавливаем, что функция

$$p(x,\ t\,|\,x_0)=\frac{1}{\sqrt{2\pi t}}\exp\left[-(x-x_0)^2/2t\right] \tag{1.4}$$

является решением уравнения (1.3), точнее единственным его решением, удовлетворяющим условиям (1.1) и (1.2). (Вопрос о единственности решения уравнения (1.3) должен быть сформулирован строго, а его анализ требует большой аккуратности и выходит за рамки данной книги. Прилежный читатель может обратиться к книге Ито и Маккина, упомянутой в ссылках в конце главы.)

Другим подходом к (1.3) является аппроксимация с помощью дискретного случайного блуждания[1]). Рассмотрим симметричное случайное блуждание на целочисленной решетке (см. пример Б, гл. 2, § 2). Пусть $p_k(n)$ — вероятность того, что частица при случайном блуждании в момент n оказывается на расстоянии k справа от исходной точки $(-\infty<k<\infty)$. Уравнение Колмогорова — Чэпмена (формула (3.2) гл. 2) принимает в этом случае вид

$$p_k(n+1)=\frac{1}{2}\,p_{k+1}(n)+\frac{1}{2}\,p_{k-1}(n),$$

или

$$p_k(n+1)-p_k(n)=\frac{1}{2}\left[p_{k+1}(n)-2p_k(n)+p_{k-1}(n)\right]. \tag{1.5}$$

Слева стоит дискретный аналог производной по времени, а справа — с коэффициентом 1/2 — дискретный аналог второй производ-

[1]) Подробности можно найти в книге Е. Б. Дынкина и А. А. Юшкевича «Теоремы и задачи о процессах Маркова», «Наука», 1967. — *Прим. перев.*

ной по пространственной переменной. Переходя соответствующим образом к пределу при одновременном стремлении времени между переходами и величины шага к нулю, можно получить (1.3) из (1.5).

В частности, пусть интервал времени между переходами равен Δ, а величина каждого шага равна η. Тогда аналог соотношения (1.5) имеет вид

$$\frac{p_{k\eta}((n+1)\Delta)-p_{k\eta}(n\Delta)}{\Delta}=\frac{(1/2)\left[p_{(k+1)\eta}(n\Delta)-2p_{k\eta}(n\Delta)+p_{(k-1)\eta}(n\Delta)\right]}{\Delta}. \quad (1.6)$$

Пусть теперь Δ и η стремятся к 0 так, что сохраняется соотношение $\Delta=\eta^2$, и в то же самое время пусть n и k стремятся к ∞ так, чтобы $k\eta\to x$, а $n\Delta\to t$. Тогда $p_{k\eta}(n\Delta)\to p(x, t\,|\,0)$ и соотношение (1.6) формально переходит в (1.3).

Мы не будем пытаться проводить эту процедуру более строго. Это просто по существу, но требует довольно тонкого анализа.

При другом способе перехода к пределу используется центральная предельная теорема. Имеем

$$p_k(n)=\mathrm{P}\{X_1+X_2+\ldots+X_n=k\},$$

где $\{X_i\}$ — последовательность исходов бросания симметричной монеты (т. е. $X_i=1$, если выпал герб, и $X_i=-1$, если выпала решетка, причем вероятность каждого исхода равна 1/2). По центральной предельной теореме (см. гл. 1, § 1)

$$\lim_{n\to\infty}\sum_{k=-\infty}^{\sqrt{n}\,x}p_k(n)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-u^2/2)\,du. \quad (1.7)$$

Предельное соотношение (1.6) и равенство (1.7) в сущности эквивалентны и связаны друг с другом «принципом инвариантности для случайных процессов». Приведенные эвристические выводы могут быть сделаны строгими, но это выходит за пределы данной книги.

§ 2. СОВМЕСТНЫЕ ВЕРОЯТНОСТИ ДЛЯ БРОУНОВСКОГО ДВИЖЕНИЯ

Переходная плотность (1.4) дает вероятностное распределение только для величины $X(t)-X(0)$. Полное описание процесса броуновского движения дается следующим определением.

Определение 2.1. *Броуновское движение является случайным процессом* $\{X(t);\ t\geqslant 0\}$ *со следующими свойствами:*

(а) *любое приращение* $X(t+s)-X(s)$ *нормально распределено со средним значением* 0 *и дисперсией* ct; $c>0$ — *фиксированный постоянный параметр*;

(б) *для любых двух непересекающихся интервалов* (t_1, t_2), (t_3, t_4), $t_1 < t_2 \leqslant t_3 < t_4$, *приращения* $X(t_4) - X(t_3)$ *и* $X(t_2) - X(t_1)$ — *независимые случайные величины с распределениями, указанными в пункте* (а); *аналогичное свойство имеет место и для* n *непересекающихся интервалов, где* n — *произвольное положительное целое число.*

Таким образом, мы постулируем, что смещение $X(t+s) - X(s)$ не зависит от прошлого, или, другими словами, если известно $X(s) = x_0$, то никакая дополнительная информация о поведении $X(\tau)$ при $\tau < s$ не влияет на наше знание вероятностного закона, «управляющего» приращением $X(t+s) - X(s)$. Формально это означает, что при $t > t_0 > t_1 > \ldots > t_n$

$$P\{X(t) \leqslant x \mid X(t_0) = x_0,\ X(t_1) = x_1,\ \ldots,\ X(t_n) = x_n\} = \\ = P\{X(t) \leqslant x \mid X(t_0) = x_0\}. \quad (2.1)$$

Это говорит о марковском характере процесса. Подчеркнем, однако, что предположение (б) о независимости приращений на самом деле более ограничительно, чем марковское свойство. В силу пункта (а) определения при $c = 1$ имеем

$$P\{X(t) \leqslant x \mid X(t_0) = x_0\} = P\{X(t) - X(t_0) \leqslant x - x_0\} = \\ = \frac{1}{\sqrt{2\pi(t-t_0)}} \int_{-\infty}^{x-x_0} \exp\left[-\frac{\alpha^2}{2(t-t_0)}\right] d\alpha. \quad (2.2)$$

Согласованность пункта (б) определения с пунктом (а) следует из известных свойств нормального распределения, например если $t_1 \leqslant t_2 \leqslant t_3$, то

$$X(t_3) - X(t_1) = [X(t_3) - X(t_2)] + [X(t_2) - X(t_3)].$$

Слагаемые справа являются независимыми случайными величинами со средними 0 и дисперсиями $t_3 - t_2$ и $t_2 - t_1$ соответственно. Следовательно, их сумма распределена нормально со средним 0 и дисперсией $t_3 - t_1$, как и должно быть.

Используя (2.1) и (2.2), нетрудно найти совместную плотность величин $X(t_1), X(t_2), \ldots, X(t_n)$ $(0 < t_1 < t_2 < \ldots < t_n)$ при условии, что $X(0) = 0$. В самом деле, для этого нам необходимо лишь знать плотность вероятности того, что $X_1 = X(t_1) = x_1$, $X_2 - X_1 = x_2 - x_1$ и т. д. и, наконец, $X_n - X_{n-1} = x_n - x_{n-1}$. В силу пункта (б) определения сразу получаем следующее выражение для функции плотности:

$$f(x_1, \ldots, x_n) = p(x_1, t_1)\, p(x_2 - x_1, t_2 - t_1) \ldots p(x_n - x_{n-1}, t_n - t_{n-1}), \quad (2.3)$$

где

$$p(x, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right).$$

Имея формулу (2.3), можно в принципе найти любые интересующие нас условные вероятности.

Из марковского свойства следует, что если $t_1 < t_2 < t_3$, то условная плотность величины $X(t_3)$ при известных $X(t_1)$ и $X(t_2)$ совпадает с условной плотностью величины $X(t_3)$ при известной величине $X(t_2)$.

Однако плотность величины $X(t_2)$ при известных $X(t_1)$ и $X(t_3)$ также представляет интерес. Предположим для определенности, что $X(t_1) = X(t_3) = 0$ и $t_1 = 0$, $t_3 = 1$, а $t_2 = t$ $(0 < t < 1)$.

В силу (2.3) совместная плотность $X(t)$ и $X(1)$ равна

$$f(x,\ y) = \frac{1}{2\pi\sqrt{t(1-t)}} \exp\left\{-\frac{1}{2}\left[\frac{x^2}{t} + \frac{(y-x)^2}{1-t}\right]\right\}.$$

Отсюда следует, что условная плотность величины $X(t)$ при условии $X(0) = X(1) = 0$, которую мы обозначим через $f_t(x \mid X(0) = X(1) = 0)$, равна

$$\frac{1}{\sqrt{2\pi t(1-t)}} \exp\left[-\frac{1}{2}\frac{x^2}{t(1-t)}\right], \qquad -\infty < x < \infty.$$

В частности, $\mathrm{M_c}(X(t)) = 0$, а $\mathrm{M_c}(X^2(t)) = t(1-t)$, где $\mathrm{M_c}$ означает математическое ожидание, взятое при условии $X(0) = X(1) = 0$. Такими же методами можно получить более общий интерполяционный результат.

Теорема 2.1. *Условная плотность величины $X(t)$, $t_1 < t < t_2$, при условии, что $X(t_1) = A$, а $X(t_2) = B$, является нормальной со средним*

$$A + \frac{B-A}{t_2 - t_1}(t - t_1)$$

и дисперсией

$$\frac{(t_2 - t)(t - t_1)}{t_2 - t_1}.$$

Это можно свести к разобранному случаю следующим образом. Рассматриваемая условная случайная величина $X(t)$, т. е. д. с. в. $X(t)$ при условии, что $X(t_1) = A$ и $X(t_2) = B$, имеет такую же плотность, как и случайная величина $A + X(t - t_1)$ при условии, что $X(0) = 0$, $X(t_2 - t_1) = B - A$, которая в свою очередь имеет такую же плотность, как и случайная величина

$$A + X(t - t_1) + \frac{t - t_1}{t_2 - t_1}(B - A)$$

при условии, что $X(0) = 0$ и $X(t_2 - t_1) = 0$.

§ 3. НЕПРЕРЫВНОСТЬ ТРАЕКТОРИЙ И ИХ МАКСИМАЛЬНЫЕ ЗНАЧЕНИЯ

Броуновское движение — физический процесс, и это наводит на мысль, что возможные реализации $X(t)$ как графики изменения положения частицы (т. е. выборочные траектории), движения которой определяются непрерывными столкновениями с молекулами окружающей среды, являются непрерывными функциями времени. Этот факт является верным, но его строгое доказательство требует довольно тонкого анализа. Выборочные траектории $X(t)$, хотя и непрерывны, но весьма причудливы, и у них ни в одной точке не существует производной. Этот факт является довольно глубоким. Исчерпывающее описание структуры выборочных траекторий процесса броуновского движения можно найти у П. Леви, а также у Ито и Маккина (см. ссылки в конце этой главы).

Используя свойство непрерывности траекторий, покажем, как вычислить различные интересные вероятностные характеристики броуновского движения. Первый пример иллюстрирует использование так называемого *принципа отражения.*

Имея в виду непрерывность $X(t)$, рассмотрим множество траекторий $X(t)$, $0 \leqslant t \leqslant T$, $X(0)=0$, обладающих тем свойством, что $X(T)>a$ $(a>0)$. Поскольку $X(t)$ непрерывна и $X(0)=0$, существует момент τ (который сам является случайной величиной, зависящей от рассматриваемой выборочной траектории), в который $X(t)$ впервые достигает значения a.

При $t>\tau$ построим отражение $X(t)$ относительно прямой $x=a$ и обозначим

$$\tilde{X}(t)=\begin{cases} X(t) & \text{для} \quad t<\tau, \\ a-[X(t)-a] & \text{для} \quad t>\tau \end{cases}$$

(рис. 1). Заметим, что $\tilde{X}(T)<a$, поскольку $X(T)>a$. Так как вероятностный закон развития процесса при $t>\tau$ при условии, что $X(\tau)=a$, симметричен по отношению к значениям $x>a$ и $x<a$ и не зависит от предыстории процесса до момента τ, операция отражения ставит в соответствие каждой выборочной траектории с $X(T)>a$ две равновероятные выборочные траектории $X(t)$ и $\tilde{X}(t)$, такие, что

$$\max_{0 \leqslant u \leqslant T} X(u) \geqslant a \quad \text{и} \quad \max_{0 \leqslant u \leqslant T} \tilde{X}(u) \geqslant a.$$

Обратно, в силу самой природы этого соответствия каждая траектория $X(t)$, для которой $\max\limits_{0 \leqslant u \leqslant T} X(u) \geqslant a$, является одной из двух равновероятных выборочных траекторий, одна из которых такова, что $X(T)>a$, если $X(T) \neq a$. Но $\mathrm{P}\{X(T)=a\}=0$. Таким образом, можно заключить, что если $X(0)=0$, то

$$P\{\max_{0\leqslant u\leqslant T} X(u)\geqslant a\}=2P\{X(T)>a\}=\frac{2}{\sqrt{2\pi T}}\int_a^\infty \exp\left(-\frac{x^2}{2T}\right)dx. \qquad (3.1)$$

Приведенное выше рассуждение нельзя считать исчерпывающим, хотя этот метод и типичен для многих задач анализа случайных процессов с непрерывными траекториями. (Такие процессы называются диффузионными.) Строгое обоснование включало

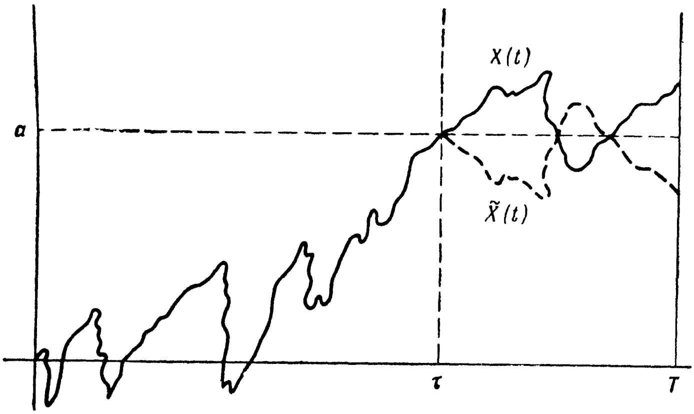


Рис. 1.

бы использование строго марковского свойства для марковского времени (см. гл. 8, § 4), соответствующего первому прохождению уровня a.

С помощью этого же «принципа отражения» можно теперь решить следующую задачу: найти

$$A(x, y)=P\{X(t)>y, \min_{0\leqslant u\leqslant t} X(u)>0 \mid X(0)=x\} \qquad (3.2)$$

для $x>0$ и $y>0$. Чтобы найти (3.2), запишем очевидное соотношение

$$P\{X(t)>y \mid X(0)=x\}=$$
$$=A(x, y)+P\{X(t)>y, \min_{0\leqslant u\leqslant t} X(u)\leqslant 0 \mid X(0)=x\} \qquad (3.3)$$

и применим принцип отражения к последнему члену.

Рис. 2 поясняет рассуждение. Можно заключить, что

$$P\{X(t)>y, \min_{0\leqslant u\leqslant t} X(u)\leqslant 0 \mid X(0)=x\}=$$
$$=P\{X(t)<-y, \min_{0\leqslant u\leqslant t} X(u)\leqslant 0 \mid X(0)=x\}=$$
$$=P\{X(t)<-y \mid X(0)=x\}. \qquad (3.4)$$

Обосновать (3.4) можно следующим образом. Рассмотрим траекторию, начинающуюся в точке $x > 0$, удовлетворяющую условию $X(t) > y$ и достигающую значения 0 в некоторый промежуточный момент τ. Отражая такую траекторию относительно нуля после момента τ, получаем траекторию, начинающуюся из x и достигающую значения, меньшего, чем $-y$, в момент t. Отсюда получается первое равенство в соотношении (3.4). Второе равенство

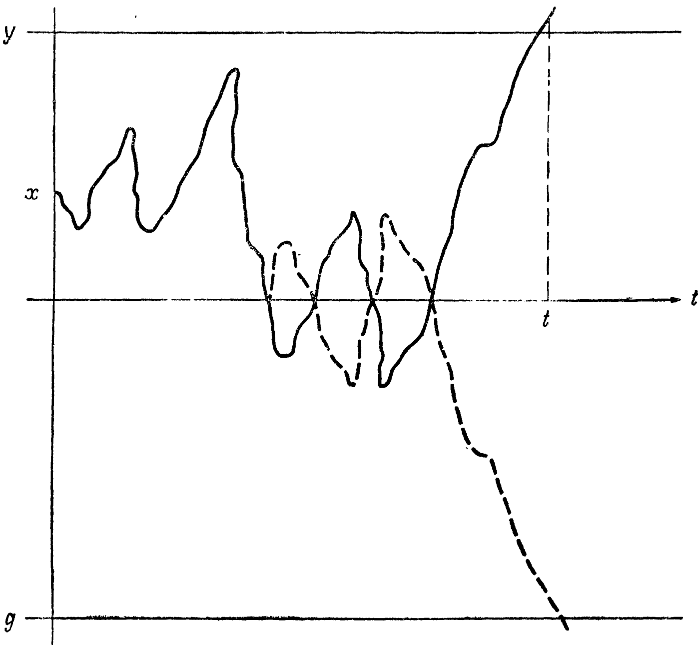


Рис. 2.

очевидно, поскольку условие $\min\limits_{0 \leqslant u \leqslant t} X(u) \leqslant 0$ излишне при выполнении неравенства $X(t) < -y$ $(y > 0)$ [1]. Подставляя (3.4) в (3.3), получаем

$$A(x, y) = \frac{1}{\sqrt{2\pi t}} \int\limits_{y}^{\infty} \exp\left[-\frac{(x-u)^2}{2t}\right] du - \frac{1}{\sqrt{2\pi t}} \int\limits_{-\infty}^{-y} \exp\left[-\frac{(x-u)^2}{2t}\right] du.$$

С помощью соотношения (3.4) можно определить распределение момента первого достижения уровня $a > 0$ при условии, что

[1]) Т. е. $\{X(t) < -y \ (y > 0)\} \subset \{\min\limits_{0 \leqslant u \leqslant t} X(u) \leqslant 0\}$ и, следовательно, $\mathrm{P}\{X(t) < -y \ (y > 0), \min\limits_{0 \leqslant u \leqslant t} X(u) \leqslant 0\} = \mathrm{P}\{X(t) < -y \ (y > 0)\}$. — *Прим. перев.*

$X(0)=0$. Пусть T_a — момент первого достижения траекторией $X(t)$ значения a, где $X(0)=0$. Тогда, очевидно,

$$\mathrm{P}\{T_a \leqslant t\} = \mathrm{P}\{\max_{0\leqslant u\leqslant t} X(u) \geqslant a \mid X(0)=0\}. \tag{3.5}$$

Но в соответствии с (3.1)

$$\mathrm{P}\{\max_{0\leqslant u\leqslant t} X(u) \geqslant a \mid X(0)=0\} = 2\mathrm{P}\{X(t)>a\} =$$

$$= \frac{2}{\sqrt{2\pi t}} \int_c^\infty \exp\left[-\frac{1}{2}\frac{x^2}{t}\right] dx$$

и поэтому

$$\mathrm{P}\{T_a \leqslant t\} = \sqrt{\frac{2}{\pi t}} \int_a^\infty \exp\left[-\frac{1}{2}\frac{x^2}{t}\right] dx.$$

Делая замену переменной $x = y\sqrt{t}$, получаем

$$\mathrm{P}\{T_a \leqslant t\} = \sqrt{\frac{2}{\pi}} \int_{a/\sqrt{t}}^\infty \exp\left[-\frac{y^2}{2}\right] dy. \tag{3.6}$$

Функция плотности случайной величины T_a получается дифференцированием (3.6) по t. Таким образом,

$$f_{T_a}(t \mid X(0)=0)\,dt = \frac{a}{\sqrt{2\pi}} t^{-3/2} \exp\left[-\frac{a^2}{2t}\right] dt. \tag{3.7}$$

В силу симметрии и пространственной однородности процесса броуновского движения мы заключаем, что

$$\mathrm{P}\{\min_{0\leqslant u\leqslant t} X(u) \leqslant 0 \mid X(0)=a\} = \qquad \text{(в силу симметрии)}$$

$$= \mathrm{P}\{\max_{0\leqslant u\leqslant t} X(u) \geqslant 0 \mid X(0)=-a\} = \qquad \text{(в силу однородности)}$$

$$= \mathrm{P}\{\max_{0\leqslant u\leqslant t} X(u) \geqslant a \mid X(0)=0\} = \mathrm{P}\{T_a \leqslant t\} = \qquad \text{(в силу (3.7))}$$

$$= \frac{a}{\sqrt{2\pi}} \int_0^t u^{-3/2} \exp\left[-\frac{a^2}{2u}\right] du, \qquad a>0. \tag{3.8}$$

Другой способ получения результата (3.8) состоит в следующем. Если $X(t_0)=a$, то вероятность $P(a)$ того, что $X(t)$ имеет хотя бы один нуль между t_0 и t_1, равна

$$P(a) = \frac{|a|}{\sqrt{2\pi}} \int_0^{t_1-t_0} \exp\left[-\frac{a^2}{2u}\right] u^{-3/2}\, du. \tag{3.9}$$

С помощью этой формулы можно найти вероятность α того, что если $X(0)=0$, то $X(t)$ обращается в нуль хотя бы один раз на интервале (t_0, t_1).

В действительности мы налагаем условие на возможные значения $X(t_0)$. Так, если $X(t_0)=a$, то вероятность того, что $X(t)$ обращается в нуль на интервале (t_0, t_1) хотя бы раз, равна $P(a)$. В силу формулы полной вероятности[1])

$$\alpha = \int_0^\infty P(a)\,\mathrm{P}\{|X(t_0)|=a \mid X(0)=0\}\,da =$$

$$= \sqrt{\frac{2}{\pi t_0}} \int_0^\infty P(a) \exp\left[-\frac{a^2}{2t_0}\right] da. \tag{3.10}$$

Подставляя в (3.10) выражение (3.9) и меняя затем порядок интегрирования, получим

$$\alpha = \sqrt{\frac{2}{\pi t_0}} \int_0^\infty \exp\left[-\frac{a^2}{2t_0}\right] \frac{a}{\sqrt{2\pi}} \left(\int_0^{t_1-t_0} \exp\left[-\frac{a^2}{2u}\right] u^{-3/2}\,du \right) da =$$

$$= \frac{1}{\pi\sqrt{t_0}} \int_0^{t_1-t_0} u^{-3/2} \left(\int_0^\infty a \exp\left[-\frac{a^2}{2}\left(\frac{1}{u}+\frac{1}{t_0}\right)\right] da \right) du. \tag{3.11}$$

Внутренний интеграл может быть взят, и после упрощения получаем

$$\alpha = \frac{\sqrt{t_0}}{\pi} \int_0^{t_1-t_0} \frac{du}{(t_0+u)\sqrt{u}}.$$

После замены переменных $u = t_0 v^2$ получим выражение

$$\alpha = \frac{2}{\pi} \int_0^{\sqrt{(t_1-t_0)/t_0}} \frac{dv}{1+v^2} = \frac{2}{\pi} \operatorname{arctg} \sqrt{\frac{t_1-t_0}{t_0}},$$

которое можно записать с помощью стандартных тригонометрических соотношений в виде

$$\sqrt{\frac{t_0}{t_1}} = \cos\frac{\pi\alpha}{2}, \quad \text{или} \quad \alpha = \frac{2}{\pi} \arccos \sqrt{\frac{t_0}{t_1}}.$$

В результате получаем следующую теорему.

---

[1]) В первом интеграле сомножитель $\mathrm{P}\{|X(t_0)|=a \mid X(0)=0\}$ является условной плотностью вероятности величины $X(t_0)$, а не условной вероятностью, т. е. здесь $\mathrm{P}\{|X(t_0)|=a \mid X(0)=0\} = p(a, t_0 \mid 0) + p(-a, t_0 \mid 0) = 2p(a, t_0 \mid 0)$, $a > 0$. — *Прим. перев.*

Теорема 3.1. *Вероятность того, что $X(t)$ имеет по крайней мере один нуль в интервале (t_0, t_1) при условии, что $X(0)=0$, равна*

$$\alpha = \frac{2}{\pi} \arccos \sqrt{\frac{t_0}{t_1}}.$$

Проведенные вычисления и обсуждения являются лишь введением в сложную структуру процесса броуновского движения. Существует множество фундаментальных связей между процессами броуновского движения и различными типами случайных процессов, возникающих в физике, биологии и экономике. Мы отсылаем читателя к специальным книгам, посвященным этой теме.

ЗАДАЧИ

Следующие задачи имеют дело с броуновским движением, определенным на стр. 299, $c=1$ (т. е. $\sigma^2(X(t))=t$ и $X(0)=0$). Они дополняют предыдущий текст и раскрывают другие интересные свойства обычного процесса броуновского движения.

1. Доказать, что

$$P\{\max_{0\leqslant u\leqslant t} X(u) < \xi,\ x \leqslant X(t) \leqslant x+dx\} \sim$$
$$\sim \frac{1}{\sqrt{2\pi t}} \left\{ \exp\left(-\frac{x^2}{2t}\right) - \exp\left(-\frac{(2\xi - x)^2}{2t}\right) \right\} dx$$

для $\xi > x$, $\xi \geqslant 0$.

Указание: Использовать принцип отражения или продифференцировать соответствующим образом переписанную формулу для $A(x, y)$ (см. формулу (3.3)).

2. Определить совместную плотность для $M(t) = \max\limits_{0\leqslant u\leqslant t} X(u)$ и $X(t)$.

Ответ:

$$\begin{cases} \dfrac{2(2\xi - x)}{t\sqrt{2\pi t}} \exp\left(-\dfrac{(2\xi-x)^2}{2t}\right), & 0\leqslant\xi,\ x\leqslant\xi; \\ 0, & \xi < x. \end{cases}$$

3. Доказать, что $P\{M(t) > \xi \mid X(t) = M(t)\} = \exp(-\xi^2/2t)$.

Указание: Пусть $Y(t) = M(t) - X(t)$. Найти условное распределение для $M(t)$ при условии, что $Y(t) = 0$.

4. Пусть T_0 — максимальный нуль функции $X(\tau)$, не превышающий t. Доказать формулу

$$P\{T_0 < t_0\} = \frac{2}{\pi} \arcsin \sqrt{t_0/t}.$$

Указание: Использовать теорему 3.1.

5. Пусть T_1 — минимальный нуль функции $X(\tau)$, превышающий t. Показать, что

(а) $P\{T_1 < t_1\} = \dfrac{2}{\pi} \arccos \sqrt{t/t_1}$,

(б) $P\{T_0 < t_0,\ T_1 > t_1\} = \dfrac{2}{\pi} \arcsin \sqrt{t_0/t_1}$.

6. Доказать, что $|X(t)|$ — марковский процесс с непрерывным временем. (Использовать симметрию $X(t)$ относительно начала координат.)

7. Доказать, что $Y(t) = M(t) - X(t)$ — марковский процесс с непрерывным временем.

Указание: Учесть, что для $t' < t$

$$Y(t) = \max\{\max_{t' \leqslant u \leqslant t} (X(u) - X(t')), Y(t')\} - (X(t) - X(t')).$$

8. Показать, что случайные процессы $Y(t) = M(t) - X(t)$ и $|X(t)|$ эквивалентны. (Говорят, что два процесса эквивалентны, если у них совпадают все конечномерные распределения.)

Указание: Поскольку $|X(t)|$ и $Y(t)$ — марковские процессы, достаточно доказать, что плотности функций распределения

$$\mathrm{P}\{Y(t) < y \mid Y(t_0) = y_0,\ t_0 < t\} \quad \text{и} \quad \mathrm{P}\{|X(t)| < y \mid |X(t_0)| = y_0,\ t_0 < t\}$$

совпадают.

Для вычисления первого члена использовать задачу 2 и представление $Y(t)$, данное в задаче 7.

9. Доказать, что вероятность того, что $Y(t)$ обратится хотя бы один раз в 0 на интервале (t_0, t_1), равна $(2/\pi)\arccos\sqrt{t_0/t_1}$.

10. Пусть T_1^* (T_0^*) — минимальный (максимальный) нуль функции $Y(\tau) = M(\tau) - X(\tau)$, превышающий (не превышающий) t. Показать, что T_0^* и T_1^* обладают теми же распределениями, что и величины T_0 и T_1 соответственно, определенные в задачах 4 и 5.

11. Доказать, что

$$\mathrm{P}\{X(\tau) \neq 0,\ 0 < t < \tau < u < 1 \mid X(0) = X(1) = 0\} = \frac{2}{\pi}\arccos\sqrt{\frac{u-t}{u(1-t)}}.$$

Указание: Вычислить

$$2\int_{a=0}^{\infty}\int_{\tau=u}^{1} \mathrm{P}\{X(t) = a,\ T(a) = \tau - t,\ X(1-\tau) = 0 \mid X(0) = 0\} \times$$
$$\times [\mathrm{P}\{X(1) = 0 \mid X(0) = 0\}]^{-1}\, da\, d\tau,$$

где $T(a)$ — время первого достижения броуновским процессом уровня 0, если $X(0) = a$ (см. формулу (3.7)). Тогда равенство

$$\frac{d}{du}\left[\frac{2}{\pi}\arccos\sqrt{\frac{u-t}{u(1-t)}}\right] =$$
$$= -\frac{2}{\pi}\frac{1}{\sqrt{1-(u-t)/u(1-t)}}\cdot\frac{1}{\sqrt{1-t}}\cdot\frac{t}{2\sqrt{1-\left(\frac{t}{u}\right)}\,u^2} =$$
$$= -\frac{\sqrt{t}}{\pi}\frac{1}{u\sqrt{(1-u)(u-t)}}$$

доказывает искомый результат.

12. Пусть $T(\lambda)$ — время первого прохождения уровня $\lambda > 0$, если $X(0) = 0$. Доказать, что распределение величины $T(\lambda_1 + \lambda_2)$ совпадает с распределением суммы $T(\lambda_1)$ и $T(\lambda_2)$, где $T(\lambda_1)$ и $T(\lambda_2)$ рассматриваются как независимые случайные величины, $\lambda_1, \lambda_2 > 0$.

13. Найти условную вероятность того, что $X(t)$ не обращается в нуль на интервале (t_0, t_2) при условии, что $X(t)$ не обращается в нуль на интервале (t_0, t_1), $0 < t_0 \leqslant t_1 \leqslant t_2$.

Ответ:

$$\frac{\arcsin\sqrt{t_0/t_2}}{\arcsin\sqrt{t_0/t_1}}.$$

14. Показать, что вероятность того, что $X(t)$ не обращается в 0 на интервале $(0, t_2)$ при условии, что $X(t)$ не обращается в 0 на интервале $(0, t_1)$, $0 < t_1 < t_2$, равна $\sqrt{t_1/t_2}$.

Указание: Найти

$$P\{X(t) \neq 0,\ 0 < t_0 \leqslant t \leqslant t_2 \mid X(t) \neq 0,\ 0 < t_0 \leqslant t \leqslant t_1\},$$

а затем положить $t_0 \to 0$.

15. Показать, что вероятность события $|X(t_1) - X(t_0)| > \xi$ при условии, что $X(t)$ принимает какое-то экстремальное значение ($X(t)$ имеет два экстремальных значения) на интервале (t_0, t_1) либо в точке t_0, либо в t_1, равна $\exp(-\xi^2/2(t_1 - t_0))$, $t_0 > 0$.

Указание: Доказать следующие утверждения.

(1) Событие, упомянутое в задаче, может произойти лишь при осуществлении одной из четырех возможностей: (A) $X(t_0)$ является минимумом, (B) $X(t_0)$ является максимумом, (C) $X(t_1)$ является минимумом, (D) $X(t_1)$ является максимумом.

(2) Условная вероятность осуществления какого-либо одного из событий (A), (B), (C), (D) при условии, что имеет место любое другое из этих событий, равна нулю.

(3) $P\{|X(t_1) - X(t_0)| > \xi \mid$ (A) или (B), или (C), или (D) имеют место$\}$ $= \sum\limits_{\alpha = (A), (B), (C), (D)} P\{|X(t_1) - X(t_0)| > \xi \mid \alpha\}\, P\{\alpha \mid$ (A) или (B), или (C), или (D) имеют место$\} = \exp\{-\xi^2/2(t_1 - t_0)\}$ (использовать задачу 3 и принцип отражения).

16. Доказать, что

$$P\{X(\tau) \neq 0,\ \tau \in (0, t) \mid X(0) = a,\ X(t) = b,\ ab > 0\} = 1 - e^{-2ab/t}.$$

Указание: Использовать функцию $A(x, y)$.

17. Проверить, что $M(X(t)X(s) \mid X(0) = 0) = \min(t, s)$.

18. Доказать, что следующие пары случайных величин имеют одни и те же распределения при всех $t > 0$:

(а) $X(c^2t)/c$, $c > 0$ и $X(t)$,

(б) $tX(1/t)$ и $X(t)$.

19. Доказать равенство

$$P\{M(1) \leqslant x \mid X(1) \leqslant 0\} = \frac{1}{\pi}\int_0^1 P\{M(u) \leqslant x \mid X(u) = 0\} \frac{1}{\sqrt{u(1-u)}}\, du.$$

Указание: Рассмотреть последний момент времени (меньший 1), когда $X(t)$ обращается в 0, при условии, что $X(1) \leqslant 0$. Использовать также тот факт, что условие $X(1) \leqslant 0$ эквивалентно условию $-\infty < X(1) < \infty$ в силу принципа отражения.

20. Доказать, что $\mathrm{P}\{X(1) \leqslant x \mid X(u) \geqslant 0,\ 0 \leqslant u \leqslant 1\} = 1 - \exp(-x^2/2)$.

Указание: Использовать пространственную однородность броуновского движения. Тогда

$$\mathrm{P}\{X(1) \leqslant x \mid X(u) \geqslant 0,\ 0 \leqslant u \leqslant 1\} =$$
$$= \mathrm{P}\{X(1) \leqslant x \mid X(u) \leqslant X(1),\ 0 \leqslant u \leqslant 1\} = 1 - \exp(-x^2/2)$$

(см. задачу 3).

21. Доказать, что для $\alpha,\ \beta > 0$

$$\mathrm{P}\{X(u) < \alpha u + \beta,\ 0 \leqslant u \leqslant 1 \mid X(0) = X(1) = 0\} = 1 - e^{-2\beta(\beta+\alpha)}.$$

Указание: Использовать теорему 2.1 для доказательства равенства

$$\mathrm{P}\{X(u) < \alpha u + \beta,\ 0 \leqslant u \leqslant 1 \mid X(0) = X(1) = 0\} =$$
$$= \mathrm{P}\{X(u) < 0,\ 0 \leqslant u \leqslant 1 \mid X(0) = -\beta,\ X(1) = -\beta - \alpha\},$$

а затем использовать задачу 16.

22. Рассмотрим семейство случайных величин $W(t) = e^{X(t)-t/2}$. Показать, что $\{W(t),\ t \geqslant 0\}$ является мартингалом, т. е.

$$\mathrm{M}(W(t) \mid W(t_1),\ W(t_2),\ \ldots,\ W(t_k)) = W(t_1),$$

где $t > t_1 > t_2 > \ldots > t_k > 0$.

23. Пусть $0 = t_0^n < t_1^n < t_2^n < \ldots < t_n^n = 1$ — последовательность точек разбиения единичного отрезка, таких, что

$$\max_{0 \leqslant i \leqslant n} \left[t_i^n - t_{i-1}^n\right] \to 0 \quad \text{при} \quad n \to \infty.$$

Пусть

$$U_n = \sum_{i=0}^{n-1} \left[X\left(t_{i-1}^n\right) - X\left(t_i^n\right)\right]^2, \quad n = 1,\ 2,\ \ldots.$$

(а) Доказать, что $\mathrm{M}(U_n) = 1$ для всех n.

(б) Доказать, что $\lim\limits_{n\to\infty} \mathrm{M}\left(U_n^2\right) = 1$.

24. Доказать, что плотность

$$p(t,\ x,\ y) = \frac{1}{\sqrt{2\pi t}} \exp\left[-(x-y)^2/2t\right]$$

удовлетворяет уравнению теплопроводности

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}.$$

25. Какова вероятность того, что частица, находившаяся сначала в 0 и совершающая броуновское движение, достигает уровня b прежде, чем уровня $-c$ $(b > 0,\ c > 0)$?

Указание: Пусть $u(x)$ $(-c < x < b)$ — вероятность того, что броуновская частица достигает из точки x уровня b прежде, чем $-c$. Доказать, что $u(x)$ удовлетворяет уравнению

$$u(x) = \frac{1}{\sqrt{2\pi t}} \int_{-c}^{b} \exp\left[-(x-y)^2/2t\right] u(y)\, dy + o(t^r), \quad t \to 0,$$

при любом $r > 0$. Затем доказать, полагая $t \to 0$, что $u(x)$ удовлетворяет дифференциальному уравнению $u''(x) = 0$ и граничному условию

$$u(-c) = 0, \quad u(b) = 1.$$

Ответ:

$$u(x) = (x + c)/(b + c).$$

Двумерным процессом броуновского движения $\mathbf{X}(t) = (X_1(t), X_2(t))$ *является процесс с независимыми приращениями* (см. гл. 1), *такой, что* $X_1(0) = X_2(0) = 0$ *и совместная плотность* $\mathbf{X}(t)$ *равна*

$$f(x_1, x_2) = \frac{1}{2\pi t} \exp\left[-\frac{x_1^2 + x_2^2}{2t}\right] = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x_1^2}{2t}\right) \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{x_2^2}{2t}\right).$$

26. Пусть $\mathbf{X}(t)$ — двумерный процесс броуновского движения. Пусть

$$R(t) = \sqrt{X_1^2(t) + X_2^2(t)}.$$

Доказать, что $R(t)$ — марковский процесс с непрерывным временем, пространство состояний которого есть $(0, \infty)$.

ЗАМЕЧАНИЯ

Содержание этой главы полностью основывается на классической работе Леви [1], в особенности на ее гл. 6.

Мы рекомендуем также блестящую монографию Каца [2], где можно найти приложения броуновского движения к статистической механике и математическому анализу.

Выдающейся работой по диффузионным процессам, которая завершает и глубоко обобщает работу Леви, является книга Ито и Маккина [3].

ЛИТЕРАТУРА

1. Lévy P., Processus stochastiques et mouvement Brownian, Gauthier-Villars, Paris, 1948.
2. Кац М., Вероятность и смежные вопросы в физике, М., «Мир», 1965.
3. Ито К., Маккин Г., Диффузионные процессы и их траектории, М., «Мир», 1968.

Глава 11

ВЕТВЯЩИЕСЯ ПРОЦЕССЫ

§ 1. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ С ДИСКРЕТНЫМ ВРЕМЕНЕМ

Ветвящиеся процессы были введены в качестве примеров цепей Маркова в гл. 2, § 2. Существуют многочисленные примеры марковских ветвящихся процессов, которые возникают естественным образом в различных научных дисциплинах. Перечислим некоторые из наиболее интересных процессов.

(а) *Электронные умножители.* Электронный умножитель является устройством, усиливающим слабый поток электронов. На пути электронов, испускаемых источником, устанавливается ряд пластин. Каждый электрон, ударяясь о первую пластину, порождает случайное число новых электронов, которые в свою очередь ударяются о следующую пластину и в свою очередь порождают электроны и т. д. Пусть X_0 — число исходных электронов, X_1 — число электронов, испущенных первой пластиной благодаря соударению с X_0 исходными электронами. Вообще пусть X_n — число электронов, эмиттированных n-й пластиной благодаря столкновению с электронами, испущенными $(n-1)$-й пластиной. Последовательность случайных величин $X_0, X_1, X_2, \ldots, X_n$ образует ветвящийся процесс.

(б) *Нейтронная цепная реакция.* При случайном столкновении с нейтроном ядро расщепляется. В результате деления испускается случайное число новых нейтронов. Каждый из этих вторичных нейтронов может бомбардировать другие ядра, производя случайное число новых нейтронов, и т. д. В этом случае первоначальное число нейтронов равно $X_0 = 1$. Первое поколение нейтронов включает все произведенные при делении, вызванном исходным нейтроном. Размер первого поколения является случайной величиной X_1. В общем случае размер n-го поколения X_n складывается из случайного количества нейтронов, произведенных при бомбардировках ядер X_{n-1} нейтронами $(n-1)$-го поколения.

(в) *Выживание фамилий.* Фамилию наследуют только сыновья. Предположим, что каждый индивидуум с вероятностью p_k имеет k потомков мужского пола. Далее, каждый индивидуум порождает 1-е, 2-е, ..., n-е, ... поколения потомков. Можно исследовать распределение такой случайной величины, как число потомков в n-м поколении, или определить вероятность того, что фамилия исчезнет. Такие вопросы будут изучаться в данной главе при общем анализе ветвящихся процессов.

(г) *Выживание мутантных генов.* Каждый отдельный ген имеет возможность породить k «потомков», $k = 1, 2, \ldots,$ которые являются генами того же типа. Однако любой отдельный ген также может трансформироваться в другой тип, называемый мутантным геном, который может стать первым в последовательности поколений мутантных генов. Представляет интерес вероятность выживания мутантных генов в популяции исходных генов.

Все приведенные выше примеры имеют следующую структуру. Пусть X_0 — размер исходной популяции. Каждый индивидуум *независимо от других* порождает k новых индивидуумов с вероятностью p_k, где

$$p_k \geqslant 0, \quad k = 0, 1, 2, \ldots, \quad \sum_{k=0}^{\infty} p_k = 1. \tag{1.1}$$

Общее количество всех прямых потомков индивидуумов исходной популяции образует первое поколение, размер которого мы обозначим через X_1. Каждый индивидуум первого поколения вновь независимым образом порождает потомство в соответствии с распределением (1.1). Общее количество всех потомков образует второе поколение размера X_2. Вообще n-е поколение складывается из потомков $(n-1)$-го поколения, каждый из членов которого независимо порождает k потомков с вероятностью p_k, $k = 0, 1, 2, \ldots$ Размер популяции n-го поколения обозначается через X_n. Величины X_n образуют последовательность целочисленных случайных величин, связанных в цепь Маркова.

§ 2. СООТНОШЕНИЯ ДЛЯ ПРОИЗВОДЯЩЕЙ ФУНКЦИИ, ОПИСЫВАЮЩЕЙ ВЕТВЯЩИЙСЯ ПРОЦЕСС

Получим некоторые соотношения для производящих функций распределений величин X_n. Предположим, что исходная популяция состоит из одного индивидуума, т. е. $X_0 = 1$. Очевидно, для любого $n = 0, 1, 2, \ldots$ можно записать

$$X_{n+1} = \sum_{r=1}^{X_n} \xi_r,$$

где ξ_r $(r \geqslant 1)$ — независимые одинаково распределенные случайные величины с распределением

$$\mathrm{P}\{\xi_r = k\} = p_k, \quad k = 0, 1, 2, \ldots, \quad \sum p_k = 1.$$

Введем производящие функции

$$\varphi(s)=\sum_{k=0}^{\infty} p_k s^k$$

и

$$\varphi_n(s)=\sum_{k=0}^{\infty} \mathrm{P}\{X_n=k\}\, s^k, \quad n=0,\ 1,\ 2,\ \ldots .$$

Ясно, что

$$\varphi_0(s)=s, \quad \text{а} \quad \varphi_i(s)=\varphi(s).$$

Далее,

$$\begin{aligned}
\varphi_{n+1}(s) &= \sum_{k=0}^{\infty} \mathrm{P}\{X_{n+1}=k\}\, s^k = \\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \mathrm{P}\{X_{n+1}=k \mid X_n=j\}\,\mathrm{P}\{X_n=j\}\, s^k = \\
&= \sum_{k=0}^{\infty} s^k \sum_{j=0}^{\infty} \mathrm{P}\{X_n=j\}\,\mathrm{P}\{\xi_1+\ \ldots\ +\xi_j=k\} = \\
&= \sum_{j=0}^{\infty} \mathrm{P}\{X_n=j\} \sum_{k=0}^{\infty} \mathrm{P}\{\xi_1+\ \ldots\ +\xi_j=k\}\, s^k. \qquad (2.1)
\end{aligned}$$

Поскольку ξ_r $(r=1,\ 2,\ \ldots,\ j)$ — независимые одинаково распределенные случайные величины с общей производящей функцией $\varphi(s)$, сумма $\xi_1+\ldots+\xi_j$ имеет производящую функцию $[\varphi(s)]^j$. Таким образом,

$$\varphi_{n+1}(s)=\sum_{j=0}^{\infty} \mathrm{P}\{X_n=j\}\,[\varphi(s)]^j.$$

Но правая часть этого равенства равна производящей функции $\varphi_n(\cdot)$ с аргументом $\varphi(s)$. Таким образом,

$$\varphi_{n+1}(s)=\varphi_n(\varphi(s)). \qquad (2.2)$$

Итерируя это соотношение, получаем

$$\begin{aligned}
\varphi_{n+1}(s) &= \varphi_n(\varphi(s))=\varphi_{n-1}(\varphi(\varphi(s)))=\varphi_{n-1}(\varphi_2(s))= \\
&= \varphi_{n-2}(\varphi_2(\varphi(s)))=\varphi_{n-2}(\varphi_3(s)).
\end{aligned}$$

По индукции получаем, что для любого $k=0,\ 1,\ \ldots,\ n$

$$\varphi_{n+1}(s)=\varphi_{n-k}(\varphi_{k+1}(s));$$

в частности, при $k=n-1$

$$\varphi_{n+1}(s)=\varphi(\varphi_n(s)). \qquad (2.3)$$

Если вместо условия $X_0 = 1$ рассмотреть более общее условие $X_0 = i_0$, где i_0 — любое постоянное число, то

$$\varphi_0(s) = s^{i_0}, \quad \text{а} \quad \varphi_1(s) = [\varphi(s)]^{i_0},$$

потому что

$$X_1 = \sum_{j=1}^{i_0} \xi_j.$$

Равенство

$$\varphi_{n+1}(s) = \varphi_n(\varphi(s))$$

сохраняется, но соотношение (2.3) не выполняется[1]).

С помощью (2.2) можно теперь вычислить математическое ожидание и дисперсию величины X_n. Далее всюду будет предполагаться (если не оговорено противное), что $X_0 = 1$. Предположим, что

$$m = \mathrm{M}(X_1) \quad \text{и} \quad \sigma^2 = \sigma^2(X_1) = \mathrm{M}(X_1^2) - (\mathrm{M}(X_1))^2$$

существуют и конечны.

Очевидно, $\mathrm{M}(X_n) = \varphi_n'(1)$. Дифференцируя (2.2) и полагая $s = 1$, получаем (поскольку $\varphi(1) = 1$), что $\varphi_{n+1}'(1) = \varphi_n'(1)\varphi'(1)$, или по индукции

$$\varphi_{n+1}'(1) = \varphi'(1)\varphi_n'(1) = [\varphi'(1)]^2 \varphi_{n-1}'(1) = [\varphi'(1)]^3 \varphi_{n-2}'(1).$$

По индукции получаем

$$\varphi_{n+1}'(1) = [\varphi'(1)]^n \varphi_1'(1) = [\varphi'(1)]^{n+1}.$$

Но $\varphi'(1) = \varphi_1'(1) = \mathrm{M}(X_1) = m$. Таким образом,

$$\mathrm{M}(X_{n+1}) = m^{n+1}. \tag{2.4}$$

Чтобы найти $\sigma^2(X_{n+1})$, заметим сначала, что

$$\varphi_n''(1) = \sum_{k=2}^{\infty} k(k-1)\,\mathrm{P}\{X_n = k\} = \mathrm{M}(X_n^2) - \mathrm{M}(X_n) = \mathrm{M}(X_n^2) - \varphi_n'(1)$$

и отсюда

$$\sigma^2(X_n) = \varphi_n''(1) + \varphi_n'(1) - [\varphi_n'(1)]^2.$$

---

[1]) Вместо (2.3) выполняется равенство

$$\varphi_{n+1}(s) = \varphi_1(\varphi_n(s)) = [\varphi(\varphi_n(s))]^{i_0}.$$

В общем случае если X_0 — случайная величина с производящей функцией $\psi(s) = \sum_{k=0}^{\infty} \mathrm{P}\{X_0 = k\} s^k$, то эквивалентом соотношения (2.3) будет

$$\varphi_{n+1}(s) = \psi[\varphi(\varphi_n(s))].$$

Соотношение (2.2) сохраняется и в этом общем случае. — *Прим. перев.*

Дифференцируя (2.3) дважды и полагая $s = 1$, получаем

$$\varphi''_{n+1}(1) = \varphi''(1)\left[\varphi'_n(1)\right]^2 + \varphi'(1)\,\varphi''_n(1).$$

Поскольку $\varphi'(1) = m$, а $\varphi''(1) = \mathrm{M}(X_1^2) - \mathrm{M}(X_1) = \sigma^2 + m^2 - m$, имеем

$$\varphi''_{n+1}(1) = Lm^{2n} + m\varphi''_n(1),$$

где $L = \sigma^2 + m^2 - m$. По индукции

$$\varphi''_{n+1}(1) = L\{m^{2n} + m^{2n-1}\} + m^2\varphi''_{n-1}(1) = \ldots$$
$$\ldots = L\{m^{2n} + m^{2n-1} + \ldots + m^n\}.$$

Таким образом,

$$\sigma^2(X_{n+1}) = (\sigma^2 + m^2 - m)\{m^{2n} + m^{2n-1} + \ldots + m^n\} + m^{n+1} - m^{2n+2} =$$
$$= \sigma^2\{m^{2n} + m^{2n-1} + \ldots + m^n\} = \sigma^2 m^n \frac{m^{n+1} - 1}{m - 1}, \quad \text{если} \quad m \neq 1.$$

Если $m = 1$, то

$$\sigma^2(X_{n+1}) = (n + 1)\,\sigma^2.$$

Мы получили формулы $\mathrm{M}(X_n) = m^n$ и

$$\sigma^2(X_n) = \begin{cases} \sigma^2 m^{n-1}\dfrac{m^n - 1}{m - 1}, & \text{если} \quad m \neq 1, \\ n\sigma^2, & \text{если} \quad m = 1. \end{cases}$$

Таким образом, дисперсия увеличивается (уменьшается) со скоростью прогрессии, если $m > 1$ $(m < 1)$, и изменяется линейно, если $m = 1$. Такое поведение лежит в основе многих результатов для ветвящихся процессов.

§ 3. ВЕРОЯТНОСТИ ВЫРОЖДЕНИЯ

Мы хотим найти вероятность того, что популяция выродится, т. е. $\mathrm{P}\{X_n = 0$ для некоторого $n\}$. Очевидно, если $X_n = 0$, то $X_k = 0$ при всех $k > n$.

Заметим сначала, что вырождения никогда не произойдет, если вероятность того, что индивидуум не порождает ни одного потомка, равна нулю, т. е. если $p_0 = 0$. Таким образом, при исследовании вероятности вырождения предположим, что $0 < p_0 < 1$. Пусть

$$q_n = \mathrm{P}\{X_n = 0\} = \varphi_n(0).$$

В силу формулы (2.3)

$$q_{n+1} = \varphi_{n+1}(0) = \varphi(\varphi_n(0)) = \varphi(q_n). \tag{3.1}$$

Так как $\varphi(s)$ — строго возрастающая функция (степенной ряд с неотрицательными коэффициентами и $p_0 < 1$) и $q_1 = \varphi_1(0) = p_0 > 0$, то $q_2 = \varphi(q_1) > \varphi(0) = q_1$. Предположим, что $q_n > q_{n-1}$.

Тогда $q_{n+1} = \varphi(q_n) > \varphi(q_{n-1}) = q_n$. Этим доказано, что $q_1, q_2, \ldots$ $\ldots, q_n, \ldots$ — монотонно возрастающая последовательность, ограниченная единицей. Следовательно, существует

$$\pi = \lim_{n\to\infty} q_n$$

и $0 < \pi \leqslant 1$. Поскольку $\varphi(s)$ непрерывна при $0 \leqslant s < 1$ (непрерывность в точке $s = 1$ следует из леммы Абеля, см. гл. 2, лемма 5.1), полагая $n \to \infty$ в (3.1), получаем

$$\pi = \varphi(\pi). \tag{3.2}$$

Поскольку q_n — вероятность вырождения популяции не позже, чем за n поколений, то π — вероятность вырождения популяции, и из (3.2) следует, что π — корень уравнения

$$\varphi(s) = s. \tag{3.3}$$

Покажем теперь, что π — наименьший положительный корень уравнения (3.3). Пусть s_0 — положительный корень уравнения (3.3). Тогда $q_1 = \varphi(0) < \varphi(s_0) = s_0$. Предположим, что $q_n < s_0$. Тогда в силу (3.1) $q_{n+1} = \varphi(q_n) < \varphi(s_0) = s_0$. Таким образом, по индукции показано, что $q_n < s_0$ для всех n. Отсюда следует, что $\pi = \lim\limits_{n\to\infty} q_n \leqslant s_0$, т. е. π — наименьший положительный корень уравнения (3.3).

Теперь предположим, что $p_0 + p_1 < 1$. Тогда $\varphi(s)$ — выпуклая функция при $0 < s \leqslant 1$, поскольку $\varphi''(s) = \sum\limits_{k=2}^{\infty} k(k-1)p_k s^{k-2} > 0$. Следовательно, график функции $\varphi(s)$ может пересекать прямую с наклоном 45°, идущую из начала координат, самое большее в двух точках. Мы знаем, что $\varphi(1) = 1$, и поэтому пересечение определенно имеет место в точке $(1, 1)$. Очевидно, может иметь место один из двух случаев, представленных на рис. 1 и 2. Если $m = \varphi'(1) > 1$, то тангенс угла наклона касательной к графику $\varphi(s)$ в точке $s = 1$ больше 1 и имеет место случай, представленный на рис. 1. В этом случае $0 < \pi < 1$. Если $m = \varphi'(1) \leqslant 1$, то тангенс угла наклона касательной в точке $s = 1$ меньше или равен 1 и имеет место ситуация, представленная на рис. 2. Тогда с необходимостью $\pi = 1$. Таким образом, мы доказали, что вероятность вырождения равна 1, если $m \leqslant 1$, и меньше 1, если $m > 1$. Другими словами, вырождение определенно имеет место тогда и только тогда, когда среднее число потомков от одного индивидуума не превышает 1.

Далее, заметим, что $\varphi(s) \leqslant \varphi(\pi)$ при $0 \leqslant s \leqslant \pi$ (рис. 2). По индукции имеем $\varphi_n(s) \leqslant \pi$ $(0 \leqslant s \leqslant \pi)$ для всех n.

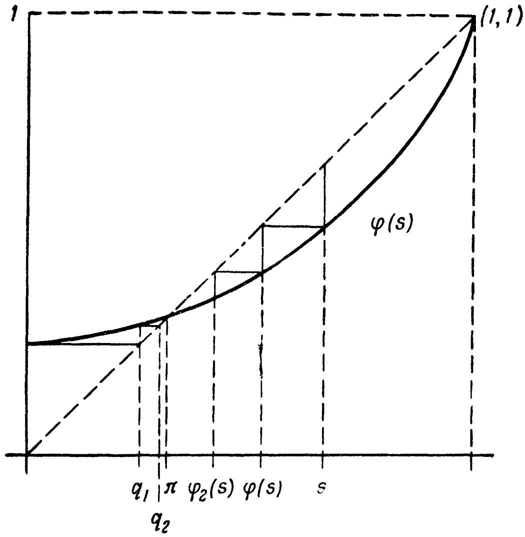


Рис. 1.

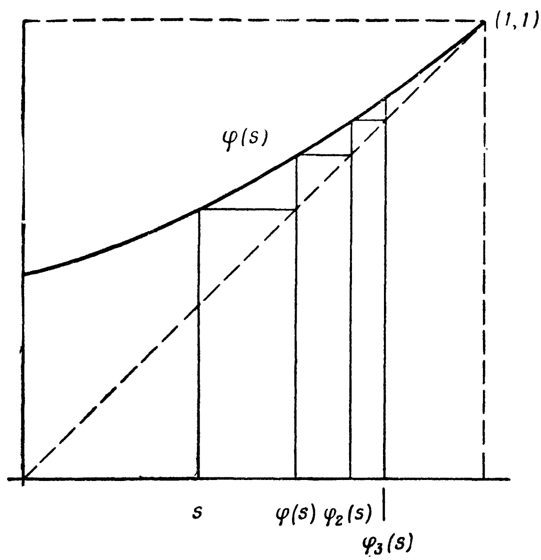


Рис. 2.

Но $\varphi_n(s) \geqslant \varphi_n(0) = q_n$, и, таким образом, $q_n \leqslant \varphi_n(s) \leqslant \pi$. Пусть $n \to \infty$. Тогда

$$\lim_{n\to\infty} \varphi_n(s) = \pi \quad \text{при} \quad 0 \leqslant s \leqslant \pi.$$

В случае $m > 1$, когда $\pi < s < 1$, имеем $\pi < \varphi(s) < s < 1$ (рис. 1). По индукции

$$\pi < \varphi_n(s) < \varphi_{n-1}(s) < \ldots \quad (\pi < s < 1).$$

Отсюда

$$\lim_{n\to\infty} \varphi_n(s) \geqslant \pi. \tag{3.4}$$

Предел в (3.4) должен равняться π, так как если бы $\lim\limits_{n\to\infty} \varphi_n(s) = \alpha > \pi$, то $\varphi(\alpha) < \alpha$, и указанная в (3.4) сходимость была бы невозможна в силу соотношения $\lim\limits_{n\to\infty} \varphi_{n+1}(s) = \lim\limits_{n\to\infty} \varphi(\varphi_n(s))$. Таким образом,

$$\lim_{n\to\infty} \varphi_n(s) = \pi \quad \text{при} \quad 0 \leqslant s < 1.$$

Из того факта, что $\varphi_n(s)$ сходится к постоянной π при $0 \leqslant s < 1$, следует, что в разложении

$$\varphi_n(s) = \sum_{k=0}^{\infty} \mathrm{P}\{X_n = k\} s^k$$

первый коэффициент $\mathrm{P}\{X_n = 0\}$ сходится к π при $n \to \infty$, а все остальные коэффициенты $\mathrm{P}\{X_n = k\}$ сходятся к 0 при $n \to \infty$, $k = 1, 2, \ldots$.

Следовательно, при любом значении $m = \mathrm{M}(X_1) > 1$ вероятность того, что n-е поколение будет состоять из любого положительного конечного числа индивидуумов, стремится к 0 при $n \to \infty$, в то время как вероятность вырождения стремится к π. В этом случае мы скажем, что $X_n \to \infty$ при $n \to \infty$ с вероятностью $1 - \pi$.

Этот результат является также следствием общей теории цепей Маркова, поскольку марковская цепь, определяемая последовательностью $X_0, X_1, X_2, \ldots$, имеет единственное поглощающее состояние $\{0\}$, и поэтому $\lim\limits_{n\to\infty} P_{ij}^n = 0$, $1 \leqslant i, j < \infty$, так как i и j являются автоматически переходными (невозвратными) состояниями.

В заключение параграфа отметим интересное свойство, состоящее в том, что условное математическое ожидание величины X_{n+r} (r — положительное целое число) при условии, что известно X_n, равно $m^r X_n$, т. е. $\mathrm{M}(X_{n+r} | X_n) = m^r X_n$. Чтобы доказать это, рассмотрим сначала случай $r = 1$:

$$\mathrm{M}\{X_{n+1} \mid X_n\} = \mathrm{M}\left\{ \sum_{j=1}^{X_n} \xi_j \mid X_n \right\} = X_n \mathrm{M}(\xi_j) = m X_n.$$

Предположим теперь, что соотношение установлено для числа r, и докажем его для $r+1$. Имеем

$$M\{X_{n+r+1} \mid X_n\} = M\{M[X_{n+r+1} \mid X_{n+r}, X_{n+r-1}, \ldots, X_n] \mid X_n\} = \\ = M\{M[X_{n+r+1} \mid X_{n+r}] \mid X_n\},$$

где использована марковская природа последовательности $\{X_n\}$. Но $M[X_{n+r+1} \mid X_{n+r}] = X_{n+r}$, и в силу индукции $M[mX_{n+r} \mid X_n] = m^{r+1}X_n$. Таким образом,

$$M\{X_{n+r} \mid X_n\} = X_n m^r \quad \text{для} \quad r = 0, 1, 2, \ldots, \quad n = 0, 1, 2, \ldots. \tag{3.5}$$

Рассмотрим теперь случайные величины

$$W_n = \frac{X_n}{m^n}, \quad n = 0, 1, 2, \ldots.$$

Тогда в силу (3.5) имеем

$$M\{W_{n+r} \mid W_n\} = \frac{1}{m^{n+r}} M\{X_{n+r} \mid X_n\} = \frac{1}{m^{n+r}} X_n m^r = W_n.$$

Можно записать, что для $r, n = 0, 1, 2, \ldots$

$$M\{W_{n+r} \mid W_n, W_{n-1}, \ldots, W_1, W_0\} = W_n, \tag{3.6}$$

откуда следует, что последовательность $\{W_n\}_{n=0}^{\infty}$ является мартингалом.

§ 4. ПРИМЕРЫ

(1) Пусть $\varphi(s) = p_0 + p_1 s$, $0 < p_0 < 1$. Соответствующий ветвящийся процесс является процессом чистой гибели. В любой период времени каждый отдельный индивидуум погибает с вероятностью p_0 и остается жить с вероятностью $p_1 = 1 - p_0$.

(2) Пусть $\varphi(s) = p_0 + p_2 s^2$ $(0 < p_0 < 1,\ p_0 + p_2 = 1)$. Такая производящая функция соответствует ветвящемуся процессу, в котором в каждом поколении отдельный индивидуум либо погибает, либо порождает двух потомков.

(3) Рассмотрим пример, где каждый индивидуум порождает N или 0 прямых потомков с вероятностями p и q соответственно. Таким образом, $p_0 = q$, $p_N = p$ и $p_k = 0$ для $k \neq 0, N$. Тогда

$$\varphi(s) = q + ps^N. \tag{4.1}$$

(4) Каждый индивидуум может иметь k потомков, где k — случайная величина с биномиальным распределением вероятностей с параметрами N и p. Тогда

$$\varphi(s) = (q + ps)^N. \tag{4.2}$$

(5) В связи с примером (г), описанным в начале этой главы, часто предполагается, что случайное число k ($k = 0, 1, 2, \ldots$) прямых «потомков» мутантного гена имеет пуассоновское распределение со средним $\lambda = 1$. Тогда $\varphi(s) = e^{s-1}$ и $\pi = 1$.

Основания для выбора именно этого распределения следующие. Во многих популяциях производится большое число зигот (оплодотворенных яиц), и лишь малая доля их доживает до зрелости. События, состоящие в оплодотворении и выживании до зрелости, подчиняются закону независимых биномиальных испытаний. Число испытаний (т. е. число зигот) настолько велико, что действительное число зрелых потомков следует пуассоновскому распределению. Именно следствие закона редких событий оправдывает применение пуассоновского приближения. Оно кажется вполне приемлемым в модели роста популяции редких мутантных генов. Если мутантный ген обладает некоторым благоприятствующим (или неблагоприятствующим) биологическим признаком, то берется пуассоновское распределение со средним $\lambda > 1$ (или $\lambda < 1$). В этом случае

$$\varphi(s) = e^{\lambda(s-1)} \tag{4.3}$$

и $0 < \pi < 1$ тогда и только тогда, когда $\lambda > 1$.

В случае гетерогенной (неоднородной) популяции мутантных генов можно предположить, что вероятностное распределение числа потомков является пуассоновским, но со случайным средним.

Например, можно представить большую географическую область, в каждой подобласти которой ветвящийся процесс характеризуется производящей функцией пуассоновского распределения с параметром λ. Предположим далее, что значение λ изменяется в зависимости от подобласти и распределение значения λ во всей области является гамма-распределением. Формально постулируется, что вероятность того, что мутантный ген имеет в точности k прямых «потомков», равна

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots,$$

где λ — случайная величина, имеющая гамма-распределение с плотностью

$$f(\lambda) = \begin{cases} \dfrac{1}{\Gamma(\alpha)}\left(\dfrac{q}{p}\right)^{\alpha}\lambda^{\alpha-1}\exp\left(-\dfrac{q}{p}\lambda\right), & \lambda \geqslant 0, \\ 0 & \text{в противном случае;} \end{cases}$$

здесь q, p, α — положительные постоянные, $q + p = 1$. Если усреднить по параметру λ, то получим вероятность того, что индивидуум имеет k потомков:

$$P\{\xi = k\} = \int_0^{\infty} P\{\xi = k \mid \lambda\} f(\lambda)\, d\lambda.$$

Производящая функция равна

$$\varphi(s)=\sum_{k=0}^{\infty}\mathrm{P}\{\xi=k\}s^k=$$

$$=\sum_{k=0}^{\infty}\int_0^{\infty}\exp(-\lambda)\frac{1}{k!}\lambda^k\frac{1}{\Gamma(\alpha)}\left(\frac{q}{p}\right)^{\alpha}\lambda^{\alpha-1}\exp\left(-\frac{q}{p}\lambda\right)^k s^k\,d\lambda=$$

$$=\int_0^{\infty}\exp(-\lambda)\frac{1}{\Gamma(\alpha)}\left(\frac{q}{p}\right)^{\alpha}\lambda^{\alpha-1}\exp\left(-\frac{q}{p}\lambda\right)\left(\sum_{k=0}^{\infty}\frac{1}{k!}(\lambda s)^k\right)d\lambda=$$

$$=\int_0^{\infty}\exp\left\{-\left(\frac{q}{p}+1-s\right)\lambda\right\}\frac{1}{\Gamma(\alpha)}\left(\frac{q}{p}\right)^{\alpha}\lambda^{\alpha-1}\,d\lambda=$$

$$=\left(\frac{q/p}{(q/p)+1-s}\right)^{\alpha}=\left(\frac{q}{1-ps}\right)^{\alpha}.$$

Мы получили производящую функцию отрицательного биномиального распределения.

(6) В примерах (2) — (4) неизвестно выражение для производящей функции $\varphi_n(s)$ в замкнутой форме. Пример, рассматриваемый ниже, допускает достаточно полный анализ. Найдем, в частности, производящую функцию для n-го поколения. Пусть

$$p_k=bc^{k-1},\quad k=1,\ 2,\ \ldots,$$

и

$$p_0=1-\sum_{k=1}^{\infty}p_k,$$

где $b,\ c>0$ и $b+c<1$. Тогда

$$p_0=1-b\sum_{k=1}^{\infty}c^{k-1}=1-\frac{b}{1-c}=\frac{1-b-c}{1-c},$$

а соответствующая производящая функция равна

$$\varphi(s)=1-\frac{b}{1-c}+bs\sum_{k=1}^{\infty}(cs)^{k-1}=\frac{1-(b+c)}{1-c}+\frac{bs}{1-cs}. \tag{4.4}$$

Заметим, что $\varphi(s)$ имеет вид функции дробно-линейного преобразования

$$f(s)=\frac{\alpha+\beta s}{\gamma+\delta s},\quad \alpha\delta-\beta\gamma\neq 0. \tag{4.5}$$

Выпишем несколько элементарных свойств дробно-линейных преобразований, необходимых для дальнейшего.

(i) Итерации дробно-линейных преобразований вновь являются дробно-линейными преобразованиями, так как если $f(s)$ задается соотношением (4.5), то из элементарных выкладок следует, что

$$f(f(s)) = \frac{\alpha(\gamma+\beta) + (\alpha\delta + \beta^2)s}{\alpha\delta + \gamma^2 + \delta(\gamma+\beta)s}.$$

(ii) Всегда существует два конечных (возможно, совпадающих) решения уравнения $f(s) = s$. Эти решения называются неподвижными точками $f(\cdot)$. Если $f(s)$ — производящая функция, то одна из неподвижных точек $s_1 = 1$, и мы увидим, что другая неподвижная точка s_0 меньше 1, равна 1 или больше 1 в соответствии с тем, будет ли значение $f'(1)$ больше, равно или меньше 1.

Для производящей функции (4.4) можно непосредственно проверить, что вторая неподвижная точка при $c > 0$, $b + c < 1$ равна

$$s_0 = \frac{1-b-c}{c(1-c)}.$$

(iii) Для любых двух точек s_i, $i = 0, 1$, легко показать, что

$$\frac{f(s) - f(s_i)}{s - s_i} = \frac{\gamma\beta - \alpha\delta}{(\gamma + \delta s)(\gamma + \delta s_i)};$$

следовательно,

$$\frac{f(s) - f(s_0)}{f(s) - f(s_1)} = \left(\frac{\gamma + \delta s_1}{\gamma + \delta s_0}\right)\left(\frac{s - s_0}{s - s_1}\right). \tag{4.6}$$

Если теперь принять, что s_0 и s_1 — две (несовпадающие) неподвижные точки $f(\cdot)$, и обозначить $w = f(s)$, то равенство (4.6) примет вид

$$\frac{w - s_0}{w - s_1} = \varkappa\,\frac{s - s_0}{s - s_1}, \tag{4.7}$$

где $\varkappa$ можно найти из (4.6) или, что проще, из (4.5), полагая $s = 0$.

Используя (4.7), нетрудно получить итерации $f_n(s) = w_n$ функции $f(s)$:

$$\frac{w_2 - s_0}{w_2 - s_1} = \varkappa\,\frac{w_1 - s_0}{w_1 - s_1} = \varkappa\left(\varkappa\,\frac{s - s_0}{s - s_1}\right)$$

и вообще

$$\frac{w_n - s_0}{w_n - s_1} = \varkappa^n\,\frac{s - s_0}{s - s_1}. \tag{4.8}$$

Для производящей функции геометрического распределения, задаваемой формулой (4.4), замечая, что неподвижные точки равны $s_0 = (1 - b - c)/c(1 - c)$ и $s_1 = 1$, получаем

$$\varkappa = \frac{(1-c)^2}{b} = \frac{1}{m},$$

где m — среднее геометрического распределения. Для $m \neq 1$ две неподвижные точки s_0 и 1 различны. Следовательно, решая (4.8) относительно w_n, находим

$$w_n = \frac{s_0 - (1/m^n)\,[(s-s_0)/(s-1)]}{1-(1/m^n)\,[(s-s_0)/(s-1)]}, \quad m \neq 1, \tag{4.9}$$

или

$$\varphi_n(s) = 1 - m^n\left(\frac{1-s_0}{m^n - s_0}\right) + \frac{m^n\,[(1-s_0)/(m^n-s_0)]^2\,s}{1-[(m^n-1)/(m^n-s_0)]\,s}. \tag{4.10}$$

Вероятность вырождения за n поколений равна

$$\mathrm{P}\{X_n = 0\} = \varphi_n(0) = 1 - m^n\left(\frac{1-s_0}{m^n-s_0}\right).$$

Заметим, что это выражение сходится к s_0 при $n \to \infty$, если $m > 1$, и к 1, если $m \leqslant 1$. Вероятностное распределение размера популяции в n-м поколении $\mathrm{P}\{X_n = k\}$, $k = 1, 2, \ldots$, можно найти разложением (4.10) в степенной ряд по s. Если определить время до вырождения T как наименьший индекс n, для которого $X_n = 0$, т. е. время первого попадания в состояние 0, то

$$\mathrm{P}\{T \leqslant n\} = \mathrm{P}\{X_n = 0\} = \varphi_n(0),$$

$$\mathrm{P}\{T = n\} = \mathrm{P}\{T \leqslant n\} - \mathrm{P}\{T \leqslant n-1\} = \varphi_n(0) - \varphi_{n-1}(0).$$

В случае $m > 1$ имеем

$$\mathrm{P}\{T = n\} = 1 - m^n\left(\frac{1-s_0}{m^n-s_0}\right) - 1 + m^{n-1}\left(\frac{1-s_0}{m^{n-1}-s_0}\right) =$$

$$= m^{n-1}s_0\frac{(m-1)(1-s_0)}{(m^n-s_0)(m^{n-1}-s_0)}, \quad n = 1, 2, \ldots.$$

Если $m = 1$, то $b = (1-c)^2$; в этом случае уравнение $\varphi(s) = s$ имеет двойной корень $s = 1$ и не имеет других корней. Таким образом,

$$\varphi(s) = c + \frac{(1-c)^2\,s}{1-cs} = \frac{c-(2c-1)\,s}{1-cs}.$$

Тогда

$$\varphi_2(s) = \varphi(\varphi(s)) = \frac{c-(2c-1)\,[(c-(2c-1)\,s)/(1-cs)]}{1-c\,[(c-(2c-1)\,s)/(1-cs)]} = \frac{2c-(3c-1)\,s}{1+c-2cs}$$

и по индукции

$$\varphi_n(s) = \frac{nc-[(n+1)\,c-1]\,s}{1+(n-1)\,c-ncs}. \tag{4.11}$$

Вероятности вырождения для случая $m = 1$ равны

$$\mathrm{P}\{X_n = 0\} = \varphi_n(0) = \frac{nc}{1+(n-1)\,c}, \quad n = 1, 2, \ldots.$$

Время до вырождения T имеет распределение

$$\mathrm{P}\{T = n\} = \frac{nc}{1+(n-1)\,c} - \frac{(n-1)\,c}{1+(n-2)\,c} = \frac{c\,(1-c)}{[1+(n-1)\,c]\,[1+(n-2)\,c]}.$$

§ 5. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ С ДВУМЯ ТИПАМИ ЧАСТИЦ

Обобщим развитую выше теорию на случай двух измерений. Рассмотрим популяцию организмов или объектов, в которой различаются два типа этих объектов. Индивидуумы каждого типа могут в общем случае порождать потомков обоих типов независимо от других индивидуумов. Пусть U_n и V_n — число индивидуумов типов I и II соответственно в n-м поколении. Можно записать

$$U_{n+1} = \sum_{j=1}^{U_n} \xi_j^{(1)} + \sum_{j=1}^{V_n} \xi_j^{(2)},$$

$$V_{n+1} = \sum_{j=1}^{U_n} \zeta_j^{(1)} + \sum_{j=1}^{V_n} \zeta_j^{(2)},$$

где $\left(\xi_j^{(i)}, \zeta_j^{(i)}\right)$ — независимые одинаково распределенные случайные векторы с распределением

$$P\left\{\xi_j^{(i)} = k,\ \zeta_j^{(i)} = l\right\} = p_i(k, l), \quad k, l = 0,\ 1,\ 2,\ \ldots,$$
$$j = 1,\ 2,\ \ldots,\ i = 1,\ 2.$$

Здесь $p_i(k, l) \geqslant 0$, $\sum\limits_{k,\, l=0}^{\infty} p_i(k, l) = 1$, $i = 1,\ 2$.

Другими словами, $p_1(k, l)$ и $p_2(k, l)$ — вероятности того, что индивидуум типа I или II соответственно порождает $k + l$ прямых потомков, из которых k имеют тип I и l — тип II.

Предположим, что процесс начинается, когда имеется один индивидуум, т. е. предположим, что

$$U_0 = 1, \quad V_0 = 0 \tag{5.1}$$

или

$$U_0 = 0, \quad V_0 = 1. \tag{5.2}$$

Введем две двумерные производящие функции

$$\varphi^{(i)}(s,\ t) = \sum_{k,\, l=0}^{\infty} p_i(k, l)\, s^k t^l, \quad i = 1, 2,$$

т. е.

$$\varphi_n^{(1)}(s,\ t) = \sum_{k,\, l=0}^{\infty} P\{U_n = k,\ V_n = l \mid U_0 = 1,\ V_0 = 0\}\, s^k t^l,$$

$$\varphi_n^{(2)}(s,\ t) = \sum_{k,\, l=0}^{\infty} P\{U_n = k,\ V_n = l \mid U_0 = 0,\ V_0 = 1\}\, s^k t^l.$$

Производящая функция распределения (5.1) имеет вид

$$\varphi_0^{(1)}(s,\ t) \equiv s,$$

а распределения (5.2) —

$$\varphi_0^{(2)}(s,\ t) \equiv t,$$

Кроме того,

$$\varphi_1^{(i)}(s, t) = \varphi^{(i)}(s, t), \quad i = 1, 2.$$

Обобщая метод, примененный для одномерного случая, можно показать, что

$$\varphi_{n+m}^{(i)}(s, t) = \varphi_m^{(i)}(\varphi_n^{(1)}(s, t), \varphi_n^{(2)}(s, t)), \quad i = 1, 2, \quad n, m = 0, 1, 2, \ldots. \tag{5.3}$$

Это двумерный аналог формулы (2.3).

Чтобы обобщить формулу (3.5), введем следующие обозначения. Пусть $\mathbf{X}_n = (U_n, V_n)$ — двумерный вектор с компонентами U_n и V_n. Положим

$$\begin{aligned} m_{11} &= \mathrm{M}\{U_1 \mid U_0 = 1, \ V_0 = 0\} = \mathrm{M}(\xi^{(1)}), \\ m_{12} &= \mathrm{M}\{V_1 \mid U_0 = 1, \ V_0 = 0\} = \mathrm{M}(\zeta^{(1)}), \\ m_{21} &= \mathrm{M}\{U_1 \mid U_0 = 0, \ V_0 = 1\} = \mathrm{M}(\xi^{(2)}), \\ m_{22} &= \mathrm{M}\{V_1 \mid U_0 = 0, \ V_0 = 1\} = \mathrm{M}(\zeta^{(2)}) \end{aligned}$$

и введем матрицу математических ожиданий

$$\mathbf{M} = \begin{Vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{Vmatrix}.$$

Таким образом, m_{11} и m_{12} — средние количества потомков типов I и II соответственно, порожденных единственным родителем типа I. Тогда обобщением (3.5) является матричное равенство

$$\mathrm{M}[\mathbf{X}_{n+r} \mid \mathbf{X}_n] = \mathbf{X}_n \mathbf{M}^r, \quad r, n = 0, 1, 2, \ldots. \tag{5.4}$$

Доказательство его для $r = 1$ проводится непосредственно:

$$\begin{aligned} &\mathrm{M}[\mathbf{X}_{n+r} \mid \mathbf{X}_n] = \\ &= \left(\mathrm{M}\left[\sum_{j=1}^{U_n} \xi_j^{(1)} + \sum_{j=1}^{V_n} \xi_j^{(2)} \,\Big|\, (U_n, V_n)\right], \ \mathrm{M}\left[\sum_{j=1}^{U_n} \zeta_j^{(1)} + \sum_{j=1}^{V_n} \zeta_j^{(2)} \,\Big|\, (U_n, V_n)\right]\right) = \\ &= (m_{11}U_n + m_{21}V_n, \ m_{12}U_n + m_{22}V_n) = (U_n, V_n) \begin{Vmatrix} m_{11} & m_{22} \\ m_{21} & m_{22} \end{Vmatrix} = \mathbf{X}_n \mathbf{M}. \end{aligned}$$

Предположим теперь, что соотношение (5.4) выполняется для r, и докажем его для $r + 1$. В силу того что последовательность $\{\mathbf{X}_n\}$ образует цепь Маркова, имеем

$$\begin{aligned} \mathrm{M}[\mathbf{X}_{n+r+1} \mid \mathbf{X}_n] &= \mathrm{M}\{\mathrm{M}[\mathbf{X}_{n+r+1} \mid \mathbf{X}_{n+r}, \ \ldots, \ \mathbf{X}_n] \mid \mathbf{X}_n\} = \\ &= \mathrm{M}\{\mathrm{M}[\mathbf{X}_{n+r+1} \mid \mathbf{X}_{n+r}] \mid X_n\} = \mathrm{M}\{\mathbf{X}_{n+r}\mathbf{M} \mid \mathbf{X}_n\} = \\ &= \mathrm{M}\{\mathbf{X}_{n+r} \mid \mathbf{X}_n\} \cdot \mathbf{M} = (\text{по индукции}) = \mathbf{X}_n \mathbf{M}^{r+1}. \end{aligned}$$

Тем самым (5.4) доказано.

Введем следующие вероятности поглощения для двумерного ветвящегося процесса:

$$\pi^{(1)} = \mathrm{P}\{U_n = V_n = 0 \text{ при некотором } n \mid U_0 = 1,\ V_0 = 0\},$$
$$\pi^{(2)} = \mathrm{P}\{U_n = V_n = 0 \text{ при некотором } n \mid U_0 = 0,\ V_0 = 1\}.$$

«Одномерная» теория распространяется на этот случай с той лишь разницей, что роль математического ожидания m здесь играет наибольшее собственное значение ρ матрицы $\mathbf{M}$.

Мы отсылаем читателя к приложению и, в частности, к теореме Фробениуса относительно матриц с неотрицательными элементами. Там доказано, что если $\mathbf{M}$ — матрица с положительными элементами (что условимся обозначать $\mathbf{M} \gg \mathbf{0}$), то собственное значение, максимальное по модулю, является положительным и, следовательно, действительным. Это собственное значение будет обозначаться через $\rho(\mathbf{M}) = \rho$.

Удобно ввести следующие векторные обозначения:

$$\mathbf{u} = (s, t),$$
$$\boldsymbol{\varphi}(\mathbf{u}) = (\varphi^{(1)}(s, t),\ \varphi^{(2)}(s, t)),$$
$$\boldsymbol{\varphi}_n(\mathbf{u}) = (\varphi_n^{(1)}(s, t),\ \varphi_n^{(2)}(s, t)),$$
$$\boldsymbol{\pi} = (\pi^{(1)}, \pi^{(2)}),$$
$$\mathbf{1} = (1, 1).$$

Теперь можно доказать такую теорему.

Теорема 5.1. *Предположим, что компоненты вектора* $\boldsymbol{\varphi}(\mathbf{u})$ *не являются линейными функциями* s *и* t *и что* $\mathbf{M} \gg \mathbf{0}$ *(все элементы матрицы* $\mathbf{M}$ *положительны). Тогда* $\boldsymbol{\pi} = \mathbf{1}$, *если максимальное собственное значение* ρ *матрицы* $\mathbf{M}$ *не превышает* 1, *и* $\boldsymbol{\pi} \ll \mathbf{1}$, *если* $\rho > 1$. *(Обозначение* $\mathbf{u} \ll \mathbf{v}$ $(\mathbf{u} \leqslant \mathbf{v})$ *означает, что вектор* $\mathbf{v} - \mathbf{u}$ *имеет положительные (неотрицательные) компоненты.) В случае* $\rho > 1$ $\boldsymbol{\pi}$ *— наименьшее неотрицательное решение уравнения*

$$\mathbf{u} = \boldsymbol{\varphi}(\mathbf{u}), \quad \mathbf{u} \ll \mathbf{1}. \tag{5.5}$$

Доказательство. Рассмотрим случай $\rho \leqslant 1$. В соответствии с общей теорией цепей Маркова мы знаем, что если цепь имеет единственное поглощающее состояние, то все состояния, из которых оно может быть достигнуто, — переходные (невозвратные). Двумерный процесс $\mathbf{X}_n = (U_n, V_n)$ является как раз таким процессом: начало координат является единственным поглощающим состоянием, и оно может быть достигнуто из любого другого. Это является следствием того факта, что $\boldsymbol{\varphi}(\mathbf{u})$ не имеет линейных компонент и $\rho \leqslant 1$. Таким образом, все состояния, за исключением начала координат, невозвратны. Следовательно, есл $|\mathbf{X}_n| = U_n + V_n$, то

$$P\{0 < |\mathbf{X}_n| < N \text{ для бесконечно многих } n\} = 0$$

для любого положительного N (см. теорему 7.1 гл. 2). Это означает, что

$$P\{|\mathbf{X}_n| \to 0\} + P\{|\mathbf{X}_n| \to \infty\} = 1.$$

Из формулы (5.4) имеем $M[\mathbf{X}_n|\mathbf{X}_0] = \mathbf{X}_0\mathbf{M}^n$. Но из теоремы 2.3 приложения следует, что $\frac{1}{\rho^n}\mathbf{M}^n$ сходится покомпонентно при $n \to \infty$. Следовательно, в случае $\rho \leqslant 1$ компоненты векторов $M[\mathbf{X}_n|\mathbf{X}_0]$ остаются ограниченными при $n \to \infty$. Отсюда следует, что событие $|\mathbf{X}_n| \to \infty$ происходит с вероятностью 0. Следовательно, $P\{|\mathbf{X}_n| \to 0\} = 1$, или, что то же, $U_n \to 0$ и $V_n \to 0$ при $n \to \infty$ с вероятностью 1. Таким образом, если $\rho \leqslant 1$, то

$$\pi^{(1)} = \pi^{(2)} = 1.$$

Рассмотрим теперь случай $\rho > 1$. Из формулы (5.3) при $s = t = 0$ имеем

$$\varphi_{n+1}^{(i)}(0,\ 0) = \varphi^{(i)}\left(\varphi_n^{(1)}(0,\ 0),\ \varphi_n^{(2)}(0,\ 0)\right), \qquad i = 1,\ 2. \tag{5.6}$$

Пусть

$$q_n^{(1)} = \varphi_n^{(1)}(0,\ 0) = P\left\{U_n = V_n = 0 \mid U_0 = 1,\ V_0 = 0\right\},$$
$$q_n^{(2)} = \varphi_n^{(2)}(0,\ 0) = P\left\{U_n = V_n = 0 \mid U_0 = 0,\ V_0 = 1\right\}.$$

Тогда (5.6) можно переписать в виде

$$q_{n+1}^{(i)} = \varphi^{(i)}\left(q_n^{(1)},\ q_n^{(2)}\right), \qquad i = 1,\ 2. \tag{5.7}$$

Поскольку $\varphi^{(i)}(s, t)$ является возрастающей функцией s и t (причем если увеличиваются обе эти переменные одновременно, то — строго возрастающей) и поскольку $q_1^{(i)} = \varphi_1^{(i)}(0,\ 0) > 0$, $i = 1,\ 2$, то имеем, очевидно,

$$q_2^{(i)} = \varphi^{(i)}\left(q_1^{(1)},\ q_1^{(2)}\right) > \varphi^{(i)}(0,\ 0) = q_1^{(i)}, \qquad i = 1,\ 2.$$

Тогда по индукции

$$q_{n+1}^{(i)} = \varphi^{(i)}\left(q_n^{(1)},\ q_n^{(2)}\right) > \varphi^{(i)}\left(q_{n-1}^{(1)},\ q_{n-1}^{(2)}\right) = q_n^{(i)}, \qquad i = 1,\ 2.$$

Следовательно, $q_n^{(i)}$, $n = 1,\ 2,\ \ldots$, при любом $i = 1,\ 2$ — монотонно возрастающая последовательность, ограниченная сверху числом 1, и

$$\lim_{n \to \infty} q_n^{(i)} = \pi^{(i)} \leqslant 1, \qquad i = 1,\ 2.$$

Пусть в соотношении (5.7) $n \to \infty$. Тогда

$$\pi^{(i)} = \varphi^{(i)}(\pi^{(1)}, \pi^{(2)}), \qquad i = 1, 2,$$

или, в векторных обозначениях, $\boldsymbol{\pi} = \boldsymbol{\varphi}(\boldsymbol{\pi})$. Докажем теперь, что $\boldsymbol{\pi} \ll \mathbf{1}$ и что это — единственное решение уравнения (5.5) при указанных условиях. Раскладывая $\varphi_n^{(i)}(\cdot, \cdot)$ по формуле Тейлора в окрестности точки $(1, 1)$, получаем выражение

$$\varphi_n^{(i)}(1-s, 1-t) = \varphi_n^{(i)}(1, 1) + \left(\frac{\partial \varphi_n^{(i)}(s, t)}{\partial s}\bigg|_{s=t=1}\right) s + \\ + \left(\frac{\partial \varphi_n^{(i)}(s, t)}{\partial t}\bigg|_{s=t=1}\right) t + o(|s| + |t|), \qquad (5.8)$$

которое справедливо при $|1-s| \leqslant 1$, $|1-t| \leqslant 1$ и достаточно малых s и t. Символ $o(\cdot)$ означает, что $[o(|s| + |t|)]/(|s| + |t|) \to 0$ при $|s| + |t| \to 0$. Очевидно,

$$\frac{\partial \varphi_n^{(1)}(s, t)}{\partial s}\bigg|_{s=t=1} = \mathrm{M}[U_n \mid U_0 = 1, V_0 = 0] = m_{11}^{(n)},$$

$$\frac{\partial \varphi_n^{(1)}(s, t)}{\partial t}\bigg|_{s=t=1} = \mathrm{M}[V_n \mid U_0 = 1, V_0 = 0] = m_{12}^{(n)},$$

$$\frac{\partial \varphi_n^{(2)}(s, t)}{\partial s}\bigg|_{s=t=1} = \mathrm{M}[U_n \mid U_0 = 0, V_0 = 1] = m_{21}^{(n)},$$

$$\frac{\partial \varphi_n^{(2)}(s, t)}{\partial t}\bigg|_{s=t=1} = \mathrm{M}[V_n \mid U_0 = 0, V_0 = 1] = m_{22}^{(n)}.$$

Перепишем (5.8) в векторной форме

$$\boldsymbol{\varphi}_n(\mathbf{1} - \mathbf{u}) = \mathbf{1} - \mathbf{M}^{(n)}\mathbf{u} + \mathbf{o}(|s| + |t|), \qquad (5.9)$$

где

$$\mathbf{M}^{(n)} = \begin{Vmatrix} m_{11}^{(n)} & m_{12}^{(n)} \\ m_{21}^{(n)} & m_{22}^{(n)} \end{Vmatrix}$$

и $|\mathbf{u}| < \varepsilon$. Из соотношения $\mathrm{M}[\mathbf{X}_n \mid \mathbf{X}_0] = \mathbf{X}_0\mathbf{M}^n$ следует, что $\mathbf{M}^{(n)} = \mathbf{M}^n$. Пусть норма вектора $\mathbf{v} = (v_1, v_2)$ определяется как сумма абсолютных значений его координат: $|\mathbf{v}| = |v_1| + |v_2|$. Докажем теперь, что для достаточно больших n

$$|\mathbf{M}^n\mathbf{u}| > 2|\mathbf{u}|, \qquad \mathbf{u} = (s, t), \qquad (5.10)$$

при условии, что $\mathbf{u} \geqslant \mathbf{0}$. Действительно, в соответствии с теоремой 2.3 приложения

$$\mathbf{M}^n\mathbf{u} = \rho^n \begin{Vmatrix} x_1^0 y_1^0 & x_1^0 y_2^0 \\ x_2^0 y_1^0 & x_2^0 y_2^0 \end{Vmatrix} \mathbf{u} + \mathbf{o}(\rho^n)\,\mathbf{u},$$

где ρ — максимальное собственное значение $\mathbf{M}$, а $\mathbf{x}^{(0)} = (x_1^0, x_2^0)$, $\mathbf{y}^0 = (y_1^0, y_2^0)$ — отвечающие ему единственные (с точностью до постоянного множителя) соответственно левый и правый собственные векторы, нормированные таким образом, что $x_1^0 y_1^0 + x_2^0 y_2^0 = 1$. Обозначение $\mathbf{o}(\rho^n)$ является обобщением общепринятого. Именно $\mathbf{o}(\rho^n)/\rho^n$ — вектор, каждый элемент которого стремится к 0 при $n \to \infty$. Множитель $\mathbf{o}(\rho^n)$ не зависит от $\mathbf{u}$. Он равен разности между $\frac{1}{\rho^n}\mathbf{M}^n$ и его предельным (при $n \to \infty$) значением. Перепишем вышеприведенное выражение в виде

$$\mathbf{M}^n\mathbf{u} = \rho^n (y_1^0 s + y_2^0 t)\, \mathbf{x}^0 + o(\rho^n)\, \mathbf{u}, \qquad \mathbf{u} = (s, t).$$

Если $\mathbf{u} \geqslant \mathbf{0}$, то получим очевидную оценку

$$|\, \mathbf{M}^n\mathbf{u} \,| \geqslant \rho^n [x_1^0 + x_2^0] \min(y_1^0, y_2^0) |\, \mathbf{u} \,| + o(\rho^n) |\, \mathbf{u} \,|.$$

Поскольку $\rho > 1$, то при достаточно большом n получим неравенство (5.10). Комбинируя (5.9) и (5.10), получаем

$$|\, \mathbf{1} - \varphi_n(\mathbf{1} - \mathbf{u}) \,| > 2 \,|\, \mathbf{u} \,|$$

при условии, что $\mathbf{1} \geqslant \mathbf{u} \geqslant \mathbf{0}$, $|\mathbf{u}|$ достаточно мал, а n достаточно велико, скажем $n \geqslant n_0$. Пусть $\mathbf{v} = \mathbf{1} - \mathbf{u}$. Тогда

$$|\, \mathbf{1} - \varphi_n(\mathbf{v}) \,| > 2 \,|\, \mathbf{1} - \mathbf{v} \,| \tag{5.11}$$

при всех $\mathbf{0} \leqslant \mathbf{v} \leqslant \mathbf{1}$, удовлетворяющих условию $|\mathbf{1} - \mathbf{v}| < \varepsilon$, и $n \geqslant n_0$. Используем (5.11) для того, чтобы доказать, что $\boldsymbol{\pi} \ll \mathbf{1}$. Предположим, что $\boldsymbol{\pi} = \mathbf{1}$, т. е. $\pi^{(i)} = 1$ для $i = 1, 2$. Тогда величина $q_n^{(i)} = \varphi_n^{(i)}(0) \geqslant 0$ стремится к 1 при $n \to \infty$. Из (5.3) следует, что

$$\varphi_{n+N}(\mathbf{0}) = \varphi_n(\varphi_N(\mathbf{0})).$$

Используя (5.11), где $\mathbf{v} = \varphi_N(\mathbf{0})$, имеем

$$|\, \mathbf{1} - \varphi_{n+N}(\mathbf{0}) \,| = |\, \mathbf{1} - \varphi_n(\varphi_N(\mathbf{0})) \,| > 2 |\, \mathbf{1} - \varphi_N(\mathbf{0}) \,|, \tag{5.12}$$

если $|\mathbf{1} - \varphi_N(\mathbf{0})| < \varepsilon$, а этого можно добиться, выбрав N достаточно большим. Однако соотношение (5.12) противоречит предположению о том, что $\varphi_n^{(i)}(0) \to 1$ при $n \to \infty$. Таким образом, равенство $\pi^{(1)} = \pi^{(2)} = 1$ невозможно. Предположим теперь, что $\pi^{(1)} < 1$ и $\pi^{(2)} = 1$. Тогда

$$\pi^{(1)} = \varphi^{(1)}(\pi^{(1)}, 1)$$

и

$$1 = \pi^{(2)} = \varphi^{(2)}(\pi^{(1)}, 1).$$

Таким образом, имеем

$$\varphi^{(2)}(1, 1) = 1, \quad \varphi^{(2)}(\pi^{(1)}, 1) = 1,$$

где $\pi^{(1)} < 1$. Поскольку $\varphi^{(2)}(s, t)$ монотонна по s, $\varphi^{(2)}(s, 1)$ должна быть постоянной на отрезке $\pi^{(1)} \leqslant s \leqslant 1$:

$$\frac{\partial \varphi^{(2)}(s, 1)}{\partial s} = 0, \qquad \pi^{(1)} \leqslant s \leqslant 1,$$

и, следовательно,

$$m_{21}^{(2)} = \left.\frac{\partial \varphi^{(2)}(s, t)}{\partial s}\right|_{s=t=1} = 0,$$

что противоречит сделанному предположению о том, что $\mathbf{M} \gg \mathbf{0}$. Аналогичным образом можно доказать, что случай $\pi^{(1)} = 1$, $\pi^{(2)} < 1$ невозможен. Таким образом, установлено, что $\boldsymbol{\pi} \ll \mathbf{1}$. Проверка того, что $\boldsymbol{\pi}$ меньше любой другой положительной неподвижной точки, производится следующим образом. Пусть вектор $\boldsymbol{\pi}^* > \mathbf{0}$ удовлетворяет уравнению $\boldsymbol{\varphi}(\boldsymbol{\pi}^*) = \boldsymbol{\pi}^*$. В силу монотонности имеем $\boldsymbol{\pi}^* = \boldsymbol{\varphi}(\boldsymbol{\pi}^*) > \boldsymbol{\varphi}_1(0, 0)$. Итерируя, получаем $\boldsymbol{\pi}^* \geqslant \boldsymbol{\varphi}_n(0, 0)$ и, переходя к пределу, находим, что $\boldsymbol{\pi}^* \geqslant \boldsymbol{\pi}$. ∎

Мы можем усилить результат теоремы 5.1 следующим образом.

Теорема 5.2. *В предположениях теоремы* 5.1 *для любого вектора* $\mathbf{q}$ *из единичного квадрата, отличного от* $\mathbf{1}$,

$$\lim_{n \to \infty} \boldsymbol{\varphi}_n(\mathbf{q}) = \boldsymbol{\pi}.$$

Доказательство. Предположим сначала, что $0 \leqslant q^{(i)} < 1$ $(i = 1, 2)$. Если N — положительное целое число, то разложение Тейлора функции $\varphi_n^{(1)}(\mathbf{q})$ имеет вид

$$\varphi_n^{(1)}(\mathbf{q}) = \mathrm{P}\{|\mathbf{X}_n| = 0 \mid U_0 = 1,\ V_0 = 0\} + \\ + \sum_{0 < |\mathbf{x}| \leqslant N} \mathrm{P}\{\mathbf{X}_n = \mathbf{x} \mid U_0 = 1,\ V_0 = 0\} (q^{(1)})^{x_1} (q^{(2)})^{x_2} + \\ + \sum_{|\mathbf{x}| > N} \mathrm{P}\{\mathbf{X}_n = \mathbf{x} \mid U_0 = 1,\ V_0 = 0\} (q^{(1)})^{x_1} (q^{(2)})^{x_2}.$$

Последняя сумма ограничена величиной

$$(\max(q^{(1)}, q^{(2)}))^N \mathrm{P}\{|\mathbf{X}_n| > N\} \leqslant (\max(q^{(1)}, q^{(2)}))^N,$$

которая при $N \to \infty$ стремится к 0, поскольку $\max(q^{(1)}, q^{(2)}) < 1$.

Каждый коэффициент первой суммы стремится к нулю при $n \to \infty$, поскольку $|\mathbf{X}_n|$ стремится либо к 0, либо к ∞. Этот факт базируется на том, что все конечные состояния, не совпадающие с $(0, 0)$, невозвратны. Отсюда следует, что при фиксированном N и $n \to \infty$ первая сумма стремится к 0. Следовательно,

$$\lim_{n \to \infty} \varphi_n^{(1)}(\mathbf{q}) = \lim_{n \to \infty} \mathrm{P}\{|\mathbf{X}_n| = 0 \mid U_0 = 1,\ V_0 = 0\} = \lim_{n \to \infty} \varphi_n^{(1)}(\mathbf{0}) = \pi^{(1)},$$

как и утверждается в теореме. Аналогично

$$\lim_{n \to \infty} \varphi_n^{(2)}(\mathbf{q}) = \pi^{(2)}.$$

Если одна из величин $q^{(i)}$ $(i = 1, 2)$ равна 1, но не обе одновременно, то $\varphi_1(\mathbf{q}) = (\varphi^{(1)}(\mathbf{q}), \varphi^{(2)}(\mathbf{q}))$ является положительным вектором, каждая компонента которого строго меньше 1. Применяя предыдущие рассуждения к $\varphi_1(\mathbf{q})$, получаем

$$\lim_{n\to\infty} \varphi_n(\varphi_1(\mathbf{q})) = \boldsymbol{\pi} = \lim_{n\to\infty} \varphi_{n+1}(\mathbf{q}). \blacksquare$$

Следствие 5.1. *Единственными решениями уравнения* (5.5) *являются* $\mathbf{1}$ *и* $\boldsymbol{\pi}$.

§ 6. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ С НЕСКОЛЬКИМИ ТИПАМИ ЧАСТИЦ

Обобщение теории, развитой в предыдущем параграфе, на случай большего количества типов повторяет дословно все, сделанное для случая двух типов. Доказательства не содержат новых идей или приемов. Мы приведем лишь сами результаты. Прилежный читатель должен провести полностью все доказательства.

Мы рассмотрим ветвящийся процесс для p типов частиц. Различные типы могут отвечать реальным различным мутантным формам некоторого организма или могут соответствовать одному организму, где отдельный тип отвечает возрасту или другому подобному свойству. То, что рассматривается лишь конечное число типов, можно интерпретировать, например, как установление конечного набора возрастных классификаций.

В случае излучения фотонов, возникающего в электронном ливне космических лучей, тип может соответствовать уровню энергии фотона.

Типу i соответствует производящая функция

$$f^{(i)}(s_1, \ldots, s_p) = \sum_{r_1, \ldots, r_p=0}^{\infty} p^{(i)}(r_1, \ldots, r_p) s_1^{r_1} \ldots s_p^{r_p},$$
$$|s_1| \leqslant 1, \ldots, |s_p| \leqslant 1, \qquad i = 1, \ldots, p,$$

где $p^{(i)}(r_1, \ldots, r_p)$ — вероятность того, что один объект типа i породит r_1 объектов типа 1, r_2 объектов типа 2, ..., r_p объектов типа p. Введем вектор $\mathbf{s} = (s_1, \ldots, s_p)$.

Пусть $f_n^{(i)}(\mathbf{s})$ — производящая функция числа индивидуумов в n-м поколении при условии, что первоначально был лишь один индивидуум типа i. Аналогично (5.3) имеем

$$f_{n+1}^{(i)}(\mathbf{s}) = f^{(i)}\left(f_n^{(1)}(\mathbf{s}), f_n^{(2)}(\mathbf{s}), \ldots, f_n^{(p)}(\mathbf{s})\right), \quad f_0^{(i)}(\mathbf{s}) = s_i, \quad n = 0, 1, \ldots, \quad i = 1, \ldots, p.$$

Пусть $\mathbf{Z}_n = \left(Z_n^{(1)}, \ldots, Z_n^{(p)}\right)$ — вектор, представляющий размер популяции, состоящей из p типов в n-м поколении. Аналог соотношения (5.4) имеет вид

$$M(\mathbf{Z}_{n+m} \mid \mathbf{Z}_n) = \mathbf{Z}_n \mathbf{M}^m,$$

где $\mathbf{M} = \|m_{ij}\|_{i,\, j=1}^{p}$ — матрица первых моментов:

$$m_{ij} = M\left(\mathbf{Z}_1^{(j)} \mid \mathbf{Z}_0 = \mathbf{e}_i\right) = \frac{\partial f^{(i)}(1, 1, \ldots, 1)}{\partial s_j}, \qquad i, j = 1, \ldots, p,$$

а $\mathbf{e}_i$ — вектор, у которого i-я компонента равна 1, а остальные — 0.

Приведем теперь аналог теоремы 5.1 для p типов. Предположим, что $m_{ij} > 0$ для всех i, j. (Достаточно потребовать, чтобы $m_{ij}^{(n)} > 0$ для некоторого n

и всех i, j.) Пусть $\pi^{(i)}$ — вероятность вырождения, если первоначально имеется лишь один объект типа i ($i = 1, \ldots, p$), т. е.

$$\pi^{(i)} = \mathrm{P}\{\mathbf{Z}_n = \mathbf{0} \quad \text{для некоторого} \quad n \mid \mathbf{Z}_0 = \mathbf{e}_i\}.$$

Вектор $(\pi^{(1)}, \ldots, \pi^{(p)})$ обозначим через $\boldsymbol{\pi}$ и введем вектор $\mathbf{1} = (1, 1, \ldots, 1)$.

Теорема 6.1. *Пусть* $m_{ij} > 0$, $i, j = 1, \ldots, p$, *и пусть* ρ — *максимальное собственное значение матрицы* $\mathbf{M}$. *Если* $\rho \leqslant 1$, *то* $\boldsymbol{\pi} = \mathbf{1}$. *Если* $\rho > 1$, *то* $\mathbf{0} \leqslant \boldsymbol{\pi} \ll \mathbf{1}$ *и* $\boldsymbol{\pi}$ *удовлетворяет уравнению*

$$\pi^{(i)} = f^{(i)}(\boldsymbol{\pi}), \qquad i = 1, \ldots, p.$$

§ 7. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Для ветвящихся процессов, изучавшихся в § 1—6, время существования одного поколения было фиксировано. Хотя некоторые явления, в особенности экспериментальные испытания, удовлетворяют этому условию, более естественным представляется процесс с непрерывным временем. Поэтому описание ветвящихся процессов с непрерывным временем представляет несомненный интерес.

В данном параграфе мы исследуем структуру *однородных по времени марковских ветвящихся процессов*; в § 11 предположение о том, что процесс является марковским, будет снято. Определим марковский ветвящийся процесс с непрерывным временем, имеющий в качестве состояния $X(t)$ число частиц в момент t при условии, что $X(0) = 1$, через инфинитезимальные параметры процесса. Пусть

$$\delta_{1k} + a_k h + o(h), \qquad k = 0,\ 1,\ 2,\ \ldots \tag{7.1}$$

(см. § 4 гл. 7 и § 3 гл. 8) — вероятность того, что одна частица породит k частиц (или объектов) за малый интервал $(t, t+h)$ длительности h. В (7.1) δ_{1k} обозначает, как обычно, символ Кронекера. Предположим, что $a_1 \leqslant 0$, $a_k \geqslant 0$ для $k = 0, 2, 3, \ldots$ и

$$\sum_{k=0}^{\infty} a_k = 0. \tag{7.2}$$

Далее, постулируем, что отдельные частицы действуют независимо друг от друга и всегда «управляются» инфинитезимальными параметрами (7.1). Заметим, что при этом предполагается однородность по времени, поскольку a_k не зависят от моментов превращения (или расщепления) частиц.

Другим способом выражения инфинитезимальных переходов является разбиение времени на участки, где не происходит превращений, и рассмотрение различных вариантов превращений. Так, каждый объект живет случайное экспоненциально распределенное время со средним значением λ^{-1}, $\lambda = a_0 + a_2 + a_3 + \ldots$, По завершении времени жизни он порождает случайное число D потомков, представляющих из себя такие же объекты. Вероятностное распределение D имеет вид

$$P\{D = k\} = \frac{a_k}{a_0 + a_2 + a_3 + \dots}, \qquad k = 0,\ 2,\ 3,\ \dots.$$

Время жизни и число потомков для отдельных индивидуумов — независимые случайные величины, и их распределения не меняются от индивидуума к индивидууму. Принимая во внимание предположения о независимости, в особенности то свойство, что отдельные индивидуумы действуют независимо друг от друга, можно (7.1) представить как инфинитезимальные переходные вероятности:

$$P\{X(t+h) = n+k-1 \mid X(t) = n\} = na_k h + o(h), \tag{7.3}$$

$$P\{X(t+h) = n \mid X(t) = n\} = 1 - na_1 h + o(h), \tag{7.4}$$

где $o(h)/h \to 0$ при $h \to 0+$.

Мы уже имели дело с примером ветвящегося процесса с непрерывным временем, когда рассматривали процесс рождения и гибели. Действительно, если положить $a_2 = \lambda$, $a_0 = \mu$, $a_1 = -(\lambda + \mu)$, $a_k = 0$, $k \neq 0, 1, 2$, то величину $(\lambda + \mu)^{-1}$ можно интерпретировать как интенсивность наступления событий рождения и гибели; $\lambda/(\lambda + \mu)$ $(\mu/(\lambda + \mu))$ — вероятность рождения (гибели) при условии, что какое-то из этих событий произошло. Так определенный случайный процесс, состоянием которого является размер популяции, есть не что иное, как процесс рождения и гибели с линейным ростом (см. гл. 7, § 6).

Как указано в гл. 8, задача построения марковского процесса, соответствующего заданной инфинитезимальной матрице, не является легкой. Еще более трудная задача — проверить, что построенный процесс имеет реализации, соответствующие предположениям для ветвящегося процесса, т. е. отдельные частицы порождают независимые семейства, потомки действуют независимо друг от друга и т. д. Мы не будем касаться анализа этого построения, так как он выходит за рамки данной книги. Более подготовленный читатель может эти вопросы найти в книге Харриса (см. литературу в конце данной главы); мы же сошлемся на § 3 гл. 8, где обсуждаются связи между марковскими процессами и инфинитезимальными матрицами.

Пусть $P_{ij}(t)$ (предположим, что всюду далее она определена корректно) — вероятность того, что размер популяции в момент t равен j при условии, что в момент 0 он был равен i, или $P_{ij}(t) = P\{X(t+s) = j \mid X(s) = i\}$. Как показывает обозначение, эта вероятность зависит лишь от прошедшего времени, т. е. процесс имеет стационарные переходные вероятности. Введем производящую функцию

$$\phi(t;\ s) = \sum_{j=0}^{\infty} P_{1j}(t)\, s^j. \tag{7.5}$$

Поскольку индивидуумы развиваются независимо, имеем основное соотношение (ср. со стр. 315)

$$\sum_{j=0}^{\infty} P_{ij}(t)\, s^j = [\phi(t;\ s)]^i. \tag{7.6}$$

Формула (7.6) характеризует ветвящиеся процессы и выделяет их среди других цепей Маркова с непрерывным временем. Она выражает то свойство, что различные индивидуумы (частицы) порождают независимые реализации процесса, где потомство одних индивидуумов не влияет на потомство других. Другими словами, популяция $X(t;\ i)$, порожденная i объектами, статистически совпадает в момент t с суммой i независимых популяций, каждая из которых порождена одним объектом.

В силу однородности по времени уравнения Колмогорова — Чэпмена имеют вид

$$P_{ij}(t+\tau) = \sum_{k=0}^{\infty} P_{ik}(t)\, P_{kj}(\tau). \tag{7.7}$$

С помощью (7.5), (7.6) и (7.7) получим

$$[\phi(t+\tau;\ s)]^i = \sum_{j=0}^{\infty} P_{ij}(t+\tau)\, s^j = \sum_{j=0,}^{\infty}\sum_{k=0}^{\infty} P_{ik}(t)\, P_{kj}(\tau)\, s^j =$$
$$= \sum_{k=0}^{\infty} P_{ik}(t) \sum_{j=0}^{\infty} P_{kj}(\tau)\, s^j = \sum_{k=0}^{\infty} P_{ik}(t)\, [\phi(\tau;\ s)]^k = [\phi(t;\ \phi(\tau;\ s))]^i$$

и, в частности,

$$\phi(t+\tau;\ s) = \phi(t;\ \phi(\tau;\ s)). \tag{7.8}$$

Соотношение (7.8) является непрерывным аналогом итерационной формулы, введенной в § 2, которая является основной для ветвящихся процессов с дискретным временем. Введем производящую функцию инфинитезимальных параметров (7.1). Именно пусть

$$u(s) = \sum_{k=0}^{\infty} a_k s^k.$$

Проводимый ниже анализ является формальным. Рассмотрим

$$\phi(h;\ s)=\sum_{j=0}^{\infty}P_{1j}(h)\,s^j=\sum_{j=0}^{\infty}(\delta_{1j}+a_jh+o(h))\,s^j=$$
$$=s+h\sum_{j=0}^{\infty}a_js^j+o(h)=s+hu(s)+o(h). \tag{7.9}$$

Из (7.8) при $\tau=h$ имеем

$$\phi(t+h;\ s)=\phi(t;\ \phi(h;\ s))=\phi(t;\ s+hu(s)+o(h))$$

и, разлагая правую часть по второй переменной по формуле Тейлора, получаем

$$\phi(t+h;\ s)=\phi(t;\ s)+\frac{\partial\phi(t;\ s)}{\partial s}hu(s)+o(h).$$

Тогда

$$\frac{\phi(t+h;\ s)-\phi(t;\ s)}{h}=\frac{\partial\phi(t;\ s)}{\partial s}u(s)+\frac{o(h)}{h}.$$

Устремив h к 0, получим

$$\frac{\partial\phi(t;\ s)}{\partial t}=\frac{\partial\phi(t;\ s)}{\partial s}u(s). \tag{7.10}$$

Это уравнение в частных производных относительно функции двух переменных $\phi(t;\ s)$, удовлетворяющей начальному условию

$$\phi(0;\ s)\equiv\sum_{j=0}^{\infty}P_{1j}(0)\,s^j\equiv s. \tag{7.11}$$

Если известна $u(s)$, то уравнение (7.10) с начальным условием (7.11) можно решить относительно $\phi(t;\ s)$.

Дифференциальное уравнение (7.10) является формой прямых дифференциальных уравнений Колмогорова, которые были свернуты в уравнение относительно производящей функции исходных переходных вероятностей.

Можно вывести второе дифференциальное уравнение относительно ϕ, которое соответствует обратным уравнениям Колмогорова. Для этого положим в (7.8) $t=h$. Получим

$$\phi(h+\tau;\ s)=\phi(h;\ \phi(\tau;\ s)).$$

Вновь используем (7.9) и разложение Тейлора. Тогда

$$\phi(h+\tau;\ s)=\phi(\tau;\ s)+hu(\phi(\tau;\ s))+o(h).$$

Это выражение можно переписать в виде

$$\frac{\phi(\tau+h;\ s)-\phi(\tau;\ s)}{h}=u(\phi(\tau;\ s))+\frac{o(h)}{h}. \tag{7.12}$$

Устремляя $h \to 0+$ и заменяя τ на t, получаем

$$\frac{\partial \phi(t;\ s)}{\partial t} = u(\phi(t;\ s)). \tag{7.13}$$

Это обыкновенное дифференциальное уравнение. Начальное условие для него вновь (7.11). Позже будет показано, как можно эффективно решить уравнение (7.13).

§ 8. ВЕРОЯТНОСТИ ВЫРОЖДЕНИЯ ДЛЯ ВЕТВЯЩИХСЯ ПРОЦЕССОВ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Решим сначала более легкую задачу нахождения среднего значения $X(t)$. Для этого продифференцируем (7.10) по s и изменим порядок дифференцирования в левой части. Получим

$$\frac{\partial}{\partial t}\frac{\partial \phi(t;\ s)}{\partial s} = \frac{\partial^2 \phi(t;\ s)}{\partial s^2}u(s) + \frac{\partial \phi(t;\ s)}{\partial s}u'(s). \tag{8.1}$$

Положим $s = 1$. Тогда, поскольку $u(1) = 0$ (условие (7.2)), имеем

$$\frac{\partial m(t)}{\partial t} = u'(1)\, m(t), \tag{8.2}$$

где

$$m(t) = \mathrm{M}(X(t)) = \left.\frac{\partial \phi(t;\ s)}{\partial s}\right|_{s=1}.$$

Решение уравнения (8.2) имеет вид

$$m(t) = \exp[u'(1)\, t], \tag{8.3}$$

поскольку $m(0) = 1$, если предположить, что $X(0) = 1$.

Займемся теперь задачей о вырождении. Для этого всюду далее в этом параграфе будем предполагать, что $a_0 > 0$, так как в противном случае вырождение невозможно. Достаточно рассмотреть случай, когда в момент 0 имеется один индивидуум. Действительно, из (7.6) известно, что

$$\sum_{j=0}^{\infty} P_{ij}(t)\, s^j = \left[\sum_{j=0}^{\infty} P_{1j}(t)\, s^j\right]^i,$$

следовательно,

$$P_{i0}(t) = [P_{10}(t)]^i.$$

Но $P_{i0}(t)$ — вероятность того, что популяция, имеющая в начальный момент размер i, выродится к моменту t. Интуитивно очевидно, что $P_{i0}(t)$ — неубывающая функция t. Докажем это формально, используя (7.8). В самом деле,

$$P_{i0}(t+\tau) = [\phi(t+\tau;\ 0)]^i = [\phi(t;\ \phi(\tau,\ 0))]^i \geqslant [\phi(t,\ 0)]^i = P_{i0}(t),$$

где использован тот факт, что $\phi(t, s)$ является степенным рядом по s с неотрицательными коэффициентами и, следовательно, возрастающей функцией s.

Вероятность вырождения можно определить как вероятность того, что популяция, порожденная одним индивидуумом, исчезнет, т. е.

$$q = \lim_{t\to\infty} P_{10}(t).$$

Используя теорию ветвящихся процессов с дискретным временем (§ 3), легко найти вероятность вырождения в случае непрерывного времени. Пусть t_0 — любое фиксированное положительное число; рассмотрим процесс с дискретным временем

$$X(0),\ X(t_0),\ X(2t_0),\ \ldots,\ X(nt_0),\ \ldots,$$

где $X(t)$ — размер популяции в момент t, соответствующий исходному ветвящемуся процессу с непрерывным временем, $X(0) = 1$. Поскольку по предположению $X(t)$ — марковский процесс, процесс с дискретным временем, то $Y_n = X(nt_0)$ образует, очевидно, цепь Маркова. Более того, он является ветвящимся процессом с дискретным временем. В самом деле, в силу гипотезы об однородности процесса $X(t)$ и в силу (7.6) получим

$$\sum_{k=0}^{\infty} \mathrm{P}\{Y_{n+1} = k \mid Y_n = i\}\, s^k = \mathrm{M}\left[s^{Y_{n+1}} \mid Y_n = i\right] =$$
$$= \mathrm{M}\left[s^{X((n+1)t_0)} \mid X(nt_0) = i\right] = \mathrm{M}\left[s^{X(t_0)} \mid X(0) = i\right] =$$
$$= [\phi(t_0;\ s)]^i = \{\mathrm{M}\left[s^{X(t_0)} \mid X(0) = 1\right]\}^i = \{\mathrm{M}\left[s^{Y_1} \mid Y_0 = 1\right]\}^i.$$

Отсюда следует, что $\{Y_n\}$ является ветвящимся процессом. Производящая функция числа потомков одного индивидуума для этого процесса равна $\phi(t_0, s)$. Следовательно, вероятность вырождения процесса $\{Y_n\}$ является наименьшим неотрицательным корнем уравнения

$$\phi(t_0;\ s) = s, \tag{8.4}$$

как это показано в § 3. Но

$$\mathrm{P}\{Y_n = 0 \text{ при некотором } n\} = \lim_{n\to\infty} \mathrm{P}\{Y_n = 0\} =$$
$$= \lim_{n\to\infty} \mathrm{P}\{X(nt_0) = 0\} = \lim_{t\to\infty} \mathrm{P}\{X(t) = 0\} = q.$$

Следовательно, *вероятность вырождения q ветвящегося процесса $X(t)$ с непрерывным временем является наименьшим неотрицательным корнем уравнения* (8.4), *где t_0 — любое положительное число.*

Поскольку q — корень уравнения (8.4) при любом t_0, следует ожидать, что вероятность q можно найти из уравнения, не содержащего этого параметра. Это в действительности так и утверждается следующей теоремой.

Теорема 8.1. *Вероятность вырождения q является наименьшим неотрицательным корнем уравнения*

$$u(s) = 0. \tag{8.5}$$

Следовательно, $q = 1$ тогда и только тогда, когда $u'(1) \leqslant 0$. (Вспомним, что $u(s) = \sum_{k=0}^{\infty} a_k s^k = a_1 s + [a_0 + a_2 s^2 + \ldots] = a_1 s + g(s)$.)

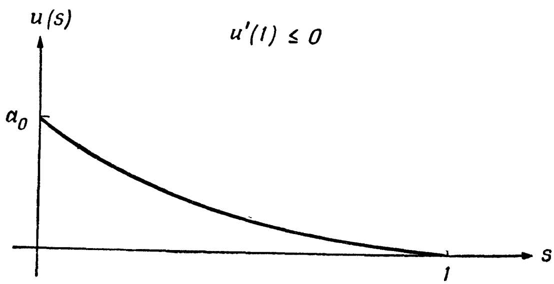


Рис. 3

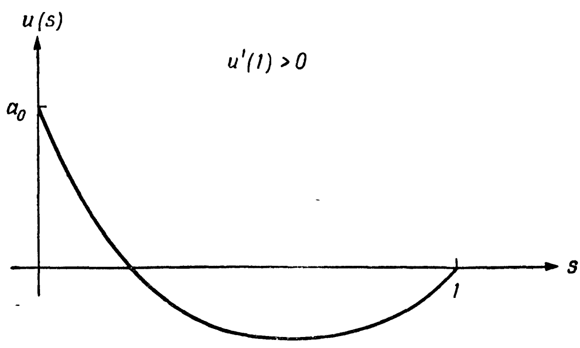


Рис. 4.

Доказательство. Поскольку q удовлетворяет уравнению (8.4) при любом t_0, из (7.12) следует, что

$$0 = u(q) + o(h)/h \quad \text{при любом} \quad h > 0.$$

Устремляя h к $0+$, получаем $u(q) = 0$.

Поскольку $u''(s) = \sum_{k=2}^{\infty} a_k k(k-1) s^{k-2} \geqslant 0$, функция $u(s)$ выпукла на отрезке $[0, 1]$. Так как $u(1) = 0$, а $u(0) = a_0 > 0$, то $u(s)$

может иметь самое большее один нуль в интервале $(0, 1)$. Случаи $u'(1) \leqslant 0$ и $u'(1) > 0$ представлены на рис. 3 и 4 соответственно. Заметим, что $M(X(t_0)) = M(Y_1) > 1$ тогда и только тогда, когда $u'(1) > 0$ [1]). Это означает, что для ветвящегося процесса с дискретным временем $X(nt_0), n = 0, 1, \ldots$ ($t_0 > 0$ и фиксировано), вырождение в этом случае происходит с вероятностью <1 и, следовательно, это же справедливо и для процесса $X(t)$. Вероятность вырождения q в этом случае с необходимостью является наименьшим нулем функции $u(s)$ на $[0, 1]$. Аналогичным образом показывается, что если $u'(1) \leqslant 0$, то $q = 1$. В любом из этих случаев q — наименьший неотрицательный корень уравнения (8.5).

§ 9. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ ВЕТВЯЩИХСЯ ПРОЦЕССОВ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Вернемся к задаче решения обыкновенного дифференциального уравнения (7.13), анализу и интерпретации его свойств при $t \to \infty$. Поскольку $\exp[u'(1)t]$ — среднее число частиц в момент t, мы будем различать тип поведения в зависимости от того, какое из условий $u'(1) < 0$, $u'(1) = 0$ или $u'(1) > 0$ выполняется. Мы обсудим лишь случай $u'(1) < 0$ при дополнительном предположении $u''(1) < \infty$. Сначала докажем, что функция

$$B(s) = \frac{1}{u(s)} - \frac{1}{u'(1)(s-1)}$$

ограничена и, следовательно, интегрируема при $0 \leqslant s < 1$. В самом деле, разлагая $u(s)$ в окрестности точки $s = 1$, получаем формулу

$$u(s) = u(1) + u'(1)(s-1) + R(s)(s-1)^2, \qquad s \leqslant 1,$$

где

$$\lim_{s \to 1-} R(s) = \frac{u''(1)}{2!} < \infty. \tag{9.1}$$

Вспоминая, что $u(q) = u(1) = 0$, получаем

$$\frac{1}{u(s)} = \frac{1}{u'(1)(s-1) + R(s)(s-1)^2} = \frac{1}{u'(1)(s-1)} \frac{1}{1 + [R(s)(s-1)/u'(1)]} =$$

$$= \frac{1}{u'(1)(s-1)} \left\{ 1 - \frac{R(s)(s-1)/u'(1)}{1 + [R(s)(s-1)/u'(1)]} \right\}.$$

Следовательно,

$$B(s) = -\frac{R(s)/[u'(1)]^2}{1 + [R(s)(s-1)/u'(1)]}, \tag{9.2}$$

и как следствие (9.1) получаем, что $B(s)$ ограничена в окрестности точки $s = 1-$. Функция $B(s)$ заведомо ограничена при всех s, не

[1]) См. формулу (8.3). — *Прим. перев.*

принадлежащих окрестности точки $s=1$, т. е. $0 \leqslant s \leqslant 1-\delta$; это следует из определения, поскольку в исследуемом случае ($u'(1)<0$) $u(s)=0$ только при $s=1$. Таким образом, $B(s)$ ограничена при $0 \leqslant s < 1$ при условии, что $u''(1) < \infty$ и $u'(1) < 0$. В силу этого можно ввести функцию

$$K(s)=\int\limits_1^s \left[\frac{1}{u(x)}-\frac{1}{u'(1)(x-1)}\right]dx+\frac{\ln(1-s)}{u'(1)}, \tag{9.3}$$

поскольку интеграл существует и конечен. Заметим, что

$$K'(s)=\frac{1}{u(s)}>0 \quad \text{при} \quad 0 \leqslant s < 1$$

снова в силу предположения $u'(1) < 0$. Это означает, что $K(s)$ — строго возрастающая непрерывная функция. Следовательно, отображение

$$w=K(s) \tag{9.4}$$

имеет обратное, которое является непрерывной строго возрастающей функцией

$$s=K^{-1}(w)=L(w), \quad L(K(s))=s, \tag{9.5}$$

обладающей тем свойством, что при изменении s на интервале $[0,1)$ функция w изменяется на интервале $[K(0), \infty)$, $K(0) < 0$. Теперь у нас есть все необходимое, чтобы найти искомое решение (7.13) при начальном условии (7.11). Разделяя в (7.13) переменные и интегрируя, получаем точную формулу относительно $\phi(t; s)$:

$$\int\limits_s^{\phi(t;\, s)} \frac{dx}{u(x)}=t.$$

Совершая очевидные преобразования и используя определение $K(\cdot)$, находим

$$t=\int\limits_s^{\phi(t;\, s)} \frac{dx}{u(x)}=\int\limits_s^{\phi(t;\, s)} \left[\frac{1}{u(x)}-\frac{1}{u'(1)(x-1)}\right]dx+\frac{1}{u'(1)}\ln(1-x)\Big|_s^{\phi(t;\, s)}=$$

$$=\int\limits_1^{\phi(t;\, s)} \left[\frac{1}{u(x)}-\frac{1}{u'(1)(x-1)}\right]dx+\frac{1}{u'(1)}\ln[1-\phi(t;\, s)]-$$

$$-\int\limits_1^s \left[\frac{1}{u(x)}-\frac{1}{u'(1)(x-1)}\right]dx-\frac{\ln(1-s)}{u'(1)}=K(\phi(t;\, s))-K(s).$$

Отсюда

$$K(\phi(t;\, s))=t+K(s).$$

Поскольку существует обратная функция, то

$$\phi(t;\ s) = K^{-1}(t + K(s)), \qquad 0 \leqslant s < 1, \qquad t \geqslant 0. \tag{9.6}$$

В предположениях $u'(1) < 0$ и $u''(1) < \infty$ можно также получить некоторые асимптотические результаты относительно вероятности вырождения к моменту t $(t \to \infty)$. Так как $B(s)$ ограничена при $0 \leqslant s < 1$ и $\lim\limits_{s \to 1-} B(s)$ существует, можно записать (см. (9.3))

$$K(s) = \frac{\ln(1-s)}{u'(1)} + C \cdot (1-s) + o(1-s). \tag{9.7}$$

Здесь C — отрицательная постоянная. В самом деле,

$$\lim_{x \to 1-} \left[\frac{1}{u(x)} - \frac{1}{(x-1)\,u'(1)}\right] = C = -\frac{u''(1)}{2\,[u'(1)]^2}.$$

Перепишем равенство (9.7) в виде

$$\ln(1-s) = u'(1)\,K(s) - Cu'(1)(1-s) + o(1-s),$$

откуда

$$1 - s = \exp[u'(1)\,K(s)]\exp[-Cu'(1)(1-s)]\exp[o(1-s)].$$

Но

$$\exp[o(1-s)] = 1 + o(1-s),$$

а

$$\exp[-Cu'(1)(1-s)] = 1 - Cu'(1)(1-s) + o(1-s).$$

Отсюда

$$1 - s = \exp[u'(1)\,K(s)]\,[1 - Cu'(1)(1-s) + o(1-s)]. \tag{9.8}$$

Следовательно,

$$\lim_{s \to 1-} \frac{1-s}{\exp[u'(1)\,K(s)]} = 1.$$

С помощью этого предельного соотношения можно записать (9.8) в виде

$$1 - s = \\ = \{\exp[u'(1)\,K(s)]\}\{1 - Cu'(1)\exp[u'(1)\,K(s)] + o(\exp[u'(1)\,K(s)])\}.$$

Заменяя s на $K^{-1}(w)$ (см. (9.5)), получаем

$$1 - K^{-1}(w) = \exp[u'(1)\,w]\{1 - Cu'(1)\exp[u'(1)\,w] + o(\exp[u'(1)\,w])\}; \tag{9.9}$$

при этом соотношение $s \to 1-$ эквивалентно $w \to \infty$. Теперь с помощью равенств (9.6) и (9.9) можно найти вероятность того, что вырождение не произойдет к моменту t:

$$\begin{aligned} 1 - P_{10}(t) &= 1 - \phi(t;\ 0) = 1 - K^{-1}(t + K(0)) = \\ &= \{\exp[u'(1)(K(0)+t)]\}\{1 - Cu'(1)\exp[u'(1)(K(0)+t)] + \\ &\qquad + o(\exp[u'(1)(K(0)+t)])\} = \\ &= \exp\{u'(1)\,K(0)\}\exp\{u'(1)\,t\} + O(\exp\{2u'(1)\,t\}) + o(\exp[u'(1)\,t]), \end{aligned}$$

или

$$1 - P_{10}(t) = \exp[u'(1)K(0)]\,m(t) + o(\exp[u'(1)t\,]). \qquad (9.10)$$

Другой асимптотический результат (при $t \to \infty$) можно получить следующим образом. Условная производящая функция величины $X(t)$ при условии, что $X(t) \neq 0$, определена равенством

$$g(z,\ t) = \sum_{k=0}^{\infty} \mathrm{P}\{X(t) = k \mid X(t) \neq 0\}\, z^k, \qquad 0 \leqslant z < 1.$$

Но

$$\mathrm{P}\{X(t) = k \mid X(t) \neq 0\} = \frac{\mathrm{P}\{X(t) = k,\ X(t) \neq 0\}}{\mathrm{P}\{X(t) \neq 0\}} =$$

$$= \begin{cases} 0, & \text{если} \quad k = 0, \\ \dfrac{\mathrm{P}\{X(t) = k\}}{1 - \mathrm{P}\{X(t) = 0\}}, & \text{если} \quad k \neq 0. \end{cases}$$

Таким образом,

$$g(z,\ t) = \sum_{k=1}^{\infty} \frac{\mathrm{P}\{X(t) = k\}}{1 - \mathrm{P}\{X(t) = 0\}} z^k = \frac{\phi(t;\ z) - \phi(t;\ 0)}{1 - \phi(t;\ 0)} =$$

$$= \frac{K^{-1}(t + K(z)) - K^{-1}(t + K(0))}{1 - K^{-1}(t + K(0))} =$$

$$= \frac{[1 - K^{-1}(t + K(0))] - [1 - K^{-1}(t + K(z))]}{1 - K^{-1}(t + K(0))},$$

где использована формула (9.6) для $\phi(t;\ s)$.

Подставляя выражение для $K^{-1}(w)$ (формула (9.9)), получаем

$$g(z;\ t) =$$

$$= \frac{e^{u'(1)(t+K(0))}[1 + O(e^{u'(1)(t+K(0))})] - e^{u'(1)(t+K(z))}[1 + O(e^{u'(1)(t+K(z))})]}{e^{u'(1)(t+K(0))}[1 + o(e^{u'(1)(t+K(0))})]} =$$

$$= 1 - e^{u'(1)[K(z) - K(0)]}\, \frac{1 + O(e^{u'(1)(t+K(z))})}{1 + o(e^{u'(1)(t+K(0))})}.$$

Пусть теперь $t \to \infty$. Тогда отношение, стоящее в правой части, стремится к 1, если $u'(1) < 0$. Следовательно,

$$\lim_{t \to \infty} g(z;\ t) = g(z) = 1 - \exp\{u'(1)[K(z) - K(0)]\}.$$

В силу (9.3), однако,

$$K(z) - K(0) = \int_0^z \left[\frac{1}{u(x)} - \frac{1}{u'(1)(x-1)}\right] dx + \frac{\ln(1-z)}{u'(1)} =$$

$$= \int_0^z \frac{dx}{u(x)} - \frac{\ln(1-x)}{u'(1)}\bigg|_0^z + \frac{\ln(1-z)}{u'(1)} = \int_0^z \frac{dx}{u(x)}.$$

Следовательно, при $t \to \infty$ предельная производящая функция равна

$$g(z) = 1 - \exp\left[u'(1) \int_0^z \frac{dx}{u(x)}\right] = \sum_{k=1}^{\infty} \lim_{t\to\infty} \mathrm{P}\{X(t) = k \mid X(t) \neq 0\} z^k.$$

Подведем итоги предыдущему обсуждению в виде следующей теоремы.

Теорема 9.1. *Рассмотрим ветвящийся процесс $X(t)$ с непрерывным временем, определяемый инфинитезимальной производящей функцией*

$$u(s) = \sum_{k=0}^{\infty} a_k s^k, \tag{9.11}$$

где интерпретация последовательности $\{a_k\}$ дана в (7.1) *и выполняется условие* (7.2). *Предположим, что $u''(1) < \infty$. Предположим, далее, что $u'(1) < 0$, так что вероятность вырождения q равна единице (см. теорему* 8.1). *Тогда*

$$\begin{aligned}\phi(t;\ s) &= \sum_{k=0}^{\infty} \mathrm{P}\{X(t) = k \mid X(0) = 1\} s^k = \\ &= K^{-1}(t + K(s)), \quad t \geqslant 0, \quad |s| < 1,\end{aligned} \tag{9.12}$$

где функция $K(s)$ определена соотношением (9.3). *Вероятность того, что до момента t не произойдет вырождения, стремится к* 0 *как экспонента в соответствии с соотношением*

$$\lim_{t\to\infty} \frac{1 - \mathrm{P}_{10}(t)}{\exp[u'(1)K(0)]\exp[u'(1)t]} = 1.$$

Кроме того, случайная величина $X(t)$ (при условии, что $X(t) > 0$) имеет предельное распределение с производящей функцией

$$\begin{aligned}g(z,\ t) &= \frac{\sum_{k=1}^{\infty} \mathrm{P}\{X(t) = k \mid X(0) = 1\} z^k}{1 - \mathrm{P}\{X(t) = 0\}} \to \\ &\to 1 - \exp\left[u'(1)\int_0^z \frac{dx}{u(x)}\right] \quad \text{при} \quad t \to \infty.\end{aligned} \tag{9.13}$$

Мы приведем без доказательства следующую предельную теорему для случаев $u'(1) = 0$ и $u'(1) > 0$. Ее доказательство более сложно, хотя и аналогично предыдущему по существу.

Теорема 9.2. (1) *Предположим, что* $u'(1)=0$ *и* $u'''(1)<\infty$. *Тогда*

$$\mathrm{P}\{X(t)>0 \mid X(0)=1\} \sim \frac{2}{u''(t)}\frac{1}{t}, \quad t\to\infty,$$

и

$$\lim_{t\to\infty}\mathrm{P}\left\{\frac{2X(t)}{u''(1)\,t}>\lambda \mid X(t)>0\right\}=e^{-\lambda}, \quad \lambda>0.$$

(2) *Если* $u'(1)>0$ *и* $u''(1)<\infty$, *то случайная величина*

$$Z(t)=\frac{X(t)}{\exp[u'(1)\,t]}$$

имеет предельное распределение при $t\to\infty$ [1]).

§ 10. ВЕТВЯЩИЙСЯ ПРОЦЕСС С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ И ДВУМЯ ТИПАМИ ЧАСТИЦ

Рассмотрим два различных типа частиц, которые мы будем называть частицами типов 1 и 2 соответственно. Ветвящийся марковский процесс с непрерывным временем будет определяться соответсгвующими инфинитезимальными параметрами. Именно мы постулируем, что каждая частица типа i ($i=1, 2$) может в течение интервала $(t, t+h)$ независимо от прошлого и независимо от истории или текущего состояния любой другой частицы любого типа превратиться в k_1 частиц типа 1 и k_2 частиц типа 2 с вероятностями

$$\delta_{1k_1}\delta_{0k_2}+a^{(1)}_{k_1,\,k_2}h+o(h) \quad \text{для} \quad i=1,$$

$$\delta_{0k_1}\delta_{1k_2}+a^{(2)}_{k_1,\,k_2}h+o(h) \quad \text{для} \quad i=2,$$

где k_1, $k_2=0, 1, 2, \ldots$, а δ_{ij} — символ Кронекера. Заметим, что мы снова постулируем однородность по времени для процесса, поскольку постоянные $a^{(i)}_{k_1,\,k_2}$ не зависят от времени. Введенные параметры подчиняются следующим ограничениям:

$a^{(1)}_{1,\,0}\leqslant 0,\quad a^{(2)}_{0,\,1}\leqslant 0,$

$a^{(1)}_{k_1,\,k_2}\geqslant 0$ для всех k_1, $k_2=0, 1, 2, \ldots$, кроме $k_1=1$, $k_2=0$;

$a^{(2)}_{k_1,\,k_2}\geqslant 0$ для всех k_1, $k_2=0, 1, 2, \ldots$, кроме $k_1=0$, $k_2=1$;

$$\sum_{k_1,\,k_2=0}^{\infty} a^{(i)}_{k_1,\,k_2}=0, \quad i=1,2.$$

[1]) Более конкретно, случайная величина $Z(t)$ сходится при $t\to\infty$ к случайной величине W с вероятностью 1. Для величины W можно найти вид производящей функции (см. книгу Т. Харриса, упоминаемую в списке литературы в конце главы). — *Прим. перев.*

Введем две производящие функции

$$u^{(i)}(s_1, s_2) = \sum_{k_1, k_2=0}^{\infty} a^{(i)}_{k_1, k_2} s_1^{k_1} s_2^{k_2}, \quad i = 1, 2 \quad (|s_1| \leqslant 1, \ |s_2| \leqslant 1).$$

Пусть $P_{k_1, k_2;\, j_1, j_2}(t)$ — вероятность того, что в момент t популяция состоит из j_1 объектов типа 1 и j_2 объектов типа 2 при условии, что в момент 0 было k_1 объектов типа 1 и k_2 объектов типа 2. Поскольку инфинитезимальные параметры $a^{(i)}_{k_1, k_2}$ не зависят от времени, переходные вероятности с необходимостью стационарны. Определим производящие функции

$$\phi^{(1)}(t;\ s_1, s_2) = \sum_{j_1, j_2=0}^{\infty} P_{1, 0;\, j_1, j_2}(t)\, s_1^{j_1} s_2^{j_2},$$

$$\phi^{(2)}(t;\ s_1, s_2) = \sum_{j_1, j_2=0}^{\infty} P_{0, 1;\, j_1, j_2}(t)\, s_1^{j_1} s_2^{j_2}.$$

Тогда аналогично одномерному случаю получим

$$\sum_{j_1, j_2=0}^{\infty} P_{k_1, k_2;\, j_1, j_2}(t)\, s_1^{j_1} s_2^{j_2} = \\ = [\phi^{(1)}(t;\ s_1, s_2)]^{k_1}\, [\phi^{(2)}(t;\ s_1, s_2)]^{k_2}, \qquad k_1, k_2 = 1, 2, \ldots . \tag{10.1}$$

Фактически соотношение (10.1) можно считать определяющим для ветвящихся процессов с двумя типами частиц и непрерывным временем. Другими словами, можно сказать, что любая переходная функция, удовлетворяющая (10.1), порождает марковский ветвящийся процесс с двумя типами частиц и непрерывным временем. Марковский характер процесса выражается уравнениями Колмогорова — Чэпмена

$$P_{k_1, k_2;\, j_1, j_2}(t + \tau) = \sum_{l_1, l_2=0}^{\infty} P_{k_1, k_2;\, l_1, l_2}(t)\, P_{l_1, l_2;\, j_1, j_2}(\tau). \tag{10.2}$$

Из (10.1) и (10.2) следует

$$\phi^{(1)}(t + \tau;\ s_1, s_2) = \sum_{j_1, j_2=0}^{\infty} P_{1, 0;\, j_1, j_2}(t + \tau)\, s_1^{j_1} s_2^{j_2} =$$

$$= \sum_{j_1, j_2=0}^{\infty} \sum_{l_1, l_2=0}^{\infty} P_{1, 0;\, l_1, l_2}(t)\, P_{l_1, l_2;\, j_1, j_2}(\tau)\, s_1^{j_1} s_2^{j_2} =$$

$$= \sum_{l_1, l_2=0}^{\infty} P_{1, 0;\, l_1, l_2}(t) \sum_{j_1, j_2=0}^{\infty} P_{l_1, l_2;\, j_1, j_2}(\tau)\, s_1^{j_1} s_2^{j_2} =$$

$$= \sum_{l_1, l_2=0}^{\infty} P_{1, 0;\, l_1, l_2}(t)\, [\phi^{(1)}(\tau;\ s_1, s_2)]^{l_1}\, [\phi^{(2)}(\tau;\ s_1, s_2)]^{l_2} =$$

$$= \phi^{(1)}(t;\ \phi^{(1)}(\tau;\ s_1, s_2),\ \phi^{(2)}(\tau;\ s_1, s_2)).$$

Аналогичное соотношение выполняется и для функции $\phi^{(2)}(t;\ s_1,\ s_2)$. Таким образом,

$$\phi^{(i)}(t+\tau;\ s_1,\ s_2) = \phi^{(i)}(t;\ \phi^{(1)}(\tau;\ s_1,\ s_2),\ \phi^{(2)}(\tau;\ s_1,\ s_2)),\quad i=1,2. \tag{10.3}$$

Кроме того,

$$\phi^{(1)}(h;\ s_1,\ s_2) = \sum_{j_1,\ j_2=0}^{\infty} P_{1,\,0;\ j_1,\ j_2}(h)\, s_1^{j_1} s_2^{j_2} =$$

$$= \sum_{j_1,\ j_2=0}^{\infty} \left[\delta_{1j_1}\delta_{0j_2} + a^{(1)}_{j_1,\ j_2}h + o(h)\right] s_1^{j_1} s_2^{j_2} = s_1 + hu^{(1)}(s_1,\ s_2) + o(h);$$

аналогичное равенство имеет место для $\phi^{(2)}(h;\ s_1,\ s_2)$. Таким образом,

$$\phi^{(i)}(h;\ s_1,\ s_2) = s_i + hu^{(i)}(s_1,\ s_2) + o(h),\qquad i=1,\ 2. \tag{10.4}$$

Получим теперь уравнения в частных производных, которым удовлетворяют функции $\phi^{(i)}(t;\ s_1,\ s_2)$ $(i=1,\ 2)$, аналогичные уравнениям (7.10) и (7.13). Для этого положим сначала $\tau = h$ и подставим (10.4) в (10.3). Используя формулу Тейлора, получаем

$$\phi^{(i)}(t+h;\ s_1,\ s_2) =$$

$$= \phi^{(i)}(t;\ s_1 + hu^{(1)}(s_1,\ s_2) + o(h),\ s_2 + hu^{(2)}(s_1,\ s_2) + o(h)) =$$

$$= \phi^{(i)}(t;\ s_1,\ s_2) = \frac{\partial\phi^{(i)}(t;\ s_1,\ s_2)}{\partial s_1} hu^{(1)}(s_1,\ s_2) +$$

$$+ \frac{\partial\phi^{(i)}(t;\ s_1,\ s_2)}{\partial s_2} hu^{(2)}(s_1,\ s_2) + o(h).$$

Разделив обе части на h и положив $h \to 0$, формально получим дифференциальные уравнения

$$\frac{\partial\phi^{(i)}(t;\ s_1,\ s_2)}{\partial t} = \frac{\partial\phi^{(i)}(t;\ s_1,\ s_2)}{\partial s_1} u^{(1)}(s_1,\ s_2) +$$

$$+ \frac{\partial\phi^{(i)}(t;\ s_1,\ s_2)}{\partial s_2} u^{(2)}(s_1,\ s_2),\qquad i=1,\ 2. \tag{10.5}$$

Обратимся вновь к равенству (10.3) и положим на этот раз $t = h$. Используя (10.4), получаем формулу.

$$\phi^{(i)}(h+\tau;\ s_1,\ s_2) = \phi^{(i)}(h;\ \phi^{(1)}(\tau;\ s_1,\ s_2),\ \phi^{(2)}(\tau;\ s_1,\ s_2)) =$$

$$= \phi^{(i)}(\tau;\ s_1,\ s_2) + hu^{(i)}(\phi^{(1)}(\tau;\ s_1,\ s_2),\ \phi^{(2)}(\tau;\ s_1,\ s_2)) + o(h).$$

Разделив на h, положив $h \to 0$ и заменив τ на t, получим вторую систему дифференциальных уравнений

$$\frac{\partial\phi^{(i)}(t;\ s_1,\ s_2)}{\partial t} = u^{(i)}(\phi^{(1)}(t;\ s_1,\ s_2),\ \phi^{(2)}(t;\ s_1,\ s_2)),\qquad i=1,\ 2. \tag{10.6}$$

Начальные условия для уравнений (10.5) и (10.6) таковы:

$$\phi^{(i)}(0;\ s_1,\ s_2) = s_i,\qquad i=1,\ 2.$$

С помощью (10.5) и (10.6) можно получить системы обыкновенных дифференциальных уравнений, которым удовлетворяют моменты искомых случайных величин. Здесь мы не будем входить в детали этих вычислений.

Дадим теперь несколько применений и примеров ветвящихся процессов с двумя типами частиц и непрерывным временем.

Пример 1. В первом примере рассматривается ветвящийся процесс с иммиграцией. Рассмотрим одномерный ветвящийся процесс с непрерывным временем и обобщим его, допустив, кроме ветвления, миграцию частиц в систему. Напомним, что

$$\delta_{1k}+a_k h+o(h), \qquad k=0,\ 1,\ 2,\ \ldots,$$

— вероятность того, что какая-либо частица превратится в k частиц за малый интервал $(t,\ t+h)$ независимо от предыстории процесса. Включим процесс иммиграции в популяцию следующим образом. Пусть

$$\delta_{0k}+b_k h+o(h), \qquad k=0,\ 1,\ 2,\ \ldots,$$

— вероятность того, что (независимо от предыстории процесса) k частиц того же типа добавятся к популяции за временной интервал $(t,\ t+h)$. Заметим, что параметры a_k так же, как и параметры b_k, по предположению не зависят от времени; иначе говоря, соответствующие переходные вероятности стационарны. Величины a_k и b_k подчиняются условиям

$$\begin{gathered}
a_1\leqslant 0, \quad b_0\leqslant 0,\\
a_k\geqslant 0 \quad \text{для} \quad k=0,\ 2,\ 3,\ \ldots,\\
b_k\geqslant 0 \quad \text{для} \quad k=1,\ 2,\ 3,\ \ldots,\\
\sum_{k=0}^{\infty} a_k=\sum_{k=0}^{\infty} b_k=0.
\end{gathered}$$

Пусть

$$\begin{aligned}
P_k(t)=\mathrm{P}\{&\text{размер популяции в момент } t \text{ равен } k \mid \text{в момент } t=0\\
&\text{размер популяции } =0\}=\\
&=\mathrm{P}\{X(t)=k\mid X(0)=0\}, \qquad k=0,\ 1,\ 2,\ \ldots, \qquad (10.7)
\end{aligned}$$

и пусть

$$\phi(t;\ s)=\sum_{k=0}^{\infty} P_k(t)\, s^k. \qquad (10.8)$$

Наша цель — найти вероятности $P_k(t)$ или — если это невозможно — получить некоторые их свойства.

Введем производящие функции

$$u(s)=\sum_{k=0}^{\infty} a_k s^k, \quad v(s)=\sum_{k=0}^{\infty} b_k s^k.$$

Можно представить одномерный ветвящийся процесс с непрерывным временем и иммиграцией в виде ветвящегося процесса с двумя типами частиц.

Идея, лежащая в основе упомянутого представления, заключается в следующем. Имеются два типа частиц: частицы типа 1 — реальные, в то время как частицы типа 2 — фиктивные. Реальная частица по окончании своего времени жизни (которое является экспоненциально распределенной случайной величиной с параметром $\lambda = a_0 + a_2 + a_3 + \ldots$) порождает k новых реальных частиц с вероятностью $\lambda^{-1} a_k$ $(k = 0, 2, 3, \ldots)$. Фиктивная частица также живет случайный отрезок времени (экспоненциально распределенный с параметром $\bar{\lambda} = b_1 + b_2 + \ldots$) и в конце его порождает l реальных частиц и одну фиктивную с вероятностью $\bar{\lambda}^{-1} b_l$ $(l = 1, 2, 3, \ldots)$. Заметим, что $\sum_{l=1}^{\infty} \bar{\lambda}^{-1} b_l = 1$. Потомство фиктивных частиц соответствует иммиграции. Таким образом,

$$a^{(1)}_{k_1, k_2} = \begin{cases} a_{k_1}, & \text{если } k_2 = 0, \\ 0, & \text{если } k_2 \neq 0, \end{cases}$$

$$a^{(2)}_{k_1, k_2} = \begin{cases} b_{k_1}, & \text{если } k_2 = 1, \\ 0, & \text{если } k_2 \neq 1. \end{cases}$$

Тогда в соответствии с обозначениями, введенными в начале этого параграфа, имеем

$$u^{(1)}(s_1, s_2) = \sum_{k_1, k_2 = 0}^{\infty} a^{(1)}_{k_1, k_2} s_1^{k_1} s_2^{k_2} = \sum_{k_1 = 0}^{\infty} a_{k_1} s_1^{k_1},$$

$$u^{(2)}(s_1, s_2) = \sum_{k_1, k_2 = 0}^{\infty} a^{(2)}_{k_1, k_2} s_1^{k_1} s_2^{k_2} = s_2 \sum_{k_1 = 0}^{\infty} b_{k_1} s_1^{k_1}.$$

Таким образом,

$$u^{(1)}(s_1, s_2) = u(s_1), \quad u^{(2)}(s_1, s_2) = s_2 v(s_1).$$

В рассматриваемом частном случае дифференциальное уравнение (10.5) принимает вид

$$\frac{\partial \phi^{(i)}(t; s_1, s_2)}{\partial t} = \frac{\partial \phi^{(i)}(t; s_1, s_2)}{\partial s_1} u(s_1) + \frac{\partial \phi^{(i)}(t; s_1, s_2)}{\partial s_2} s_2 v(s_1), \tag{10.9}$$

$$i = 1, 2,$$

а дифференциальное уравнение (10.6) —

$$\frac{\partial \phi^{(1)}(t;\ s_1,\ s_2)}{\partial t} = u\left(\phi^{(1)}(t;\ s_1,\ s_2)\right), \tag{10.10}$$

$$\frac{\partial \phi^{(2)}(t;\ s_1,\ s_2)}{\partial t} = [\phi^{(2)}(t;\ s_1,\ s_2)]\, v\left(\phi^{(1)}(t;\ s_1,\ s_2)\right). \tag{10.11}$$

Установим теперь соответствие между вероятностями $P_{0,1;\,j_1,j_2}(t)$, введенными для процесса с двумя типами частиц, и вероятностями, определяемыми соотношением (10.7). В соответствии с введенными обозначениями начальное состояние (0, 1) для процесса с двумя типами частиц означает, что в момент 0 реальные частицы отсутствуют, а имеется лишь фиктивная частица — «потенциальные иммигранты». В силу определения очевидно, что

$$P_{0,\,1;\,j_1,\,j_2}(t) = \begin{cases} P_{j_1}(t), & \text{если } j_2 = 1, \\ 0, & \text{если } j_2 \neq 1, \end{cases}$$

и, следовательно,

$$\phi^{(2)}(t;\ s_1,\ s_2) = s_2 \phi(t;\ s_1). \tag{10.12}$$

Тогда из (10.9) получаем

$$\frac{\partial \phi(t;\ s)}{\partial t} = \frac{\partial \phi(t;\ s)}{\partial s} u(s) + \phi(t;\ s)\, v(s), \tag{10.13}$$

где вместо s_1 мы написали s. Начальное условие имеет вид

$$\phi(0;\ s) = 1. \tag{10.14}$$

Вместо того чтобы решать это дифференциальное уравнение, решим систему уравнений (10.10) и (10.11), что легче. Уравнение (10.10) можно рассмотреть методами, примененными для анализа уравнения (7.13). Решение уравнения (10.10) можно представить в виде, аналогичном (9.6). Обозначим его через $f(t;\ s)$; здесь вместо s_1 и s_2 записано сокращенно s. В силу (10.12) уравнение (10.11) принимает вид

$$\frac{\partial \phi(t;\ s)}{\partial t} = \phi(t;\ s)\, v\left(f(t;\ s)\right) \tag{10.15}$$

с начальным условием (10.14). Решение уравнения (10.15) равно

$$\phi(t;\ s) = \exp\left[\int_0^t v\left(f(\tau;\ s)\right) d\tau\right].$$

Пример 2. В заключение параграфа опишем простой *немарковский* одномерный ветвящийся процесс (бинарного деления) с непрерывным временем, который можно свести к марковскому ветвящемуся процессу с двумя типами частиц. Предположим, что частица имеет распределение времени жизни с плотностью

$$\frac{\lambda^2}{2} te^{-\lambda t} \tag{10.16}$$

(гамма-распределение порядка 2). По истечении времени жизни частица заменяется на две частицы такого же типа, каждая из которых независима от другой и исходной частицы и имеет плотность (10.16) времени жизни.

Марковские процессы в общем случае характеризуются тем свойством, что время пребывания в любом состоянии распределено экспоненциально. В данной главе время пребывания в данном состоянии определяется временем жизни частицы. Если оно распределено экспоненциально, то процесс роста популяции этих частиц является марковским. В рассматриваемом случае время жизни распределено не по экспоненте, а по свертке двух экспонент.

Пусть $X(t)$ — число частиц в момент t; предположим, что $X(0) = 1$. Поскольку (10.16) — плотность суммы двух независимых экспоненциально (с параметром λ) распределенных д. с. в., можно считать, что каждая частица как бы проходит две фазы развития, каждая из которых имеет экспоненциально распределенную с параметром λ длительность. Такой процесс нетрудно свести к марковскому ветвящемуся процессу с двумя типами частиц. Вместо того чтобы говорить о двух фазах жизни одной и той же частицы, будем говорить о двух типах частиц. Частица типа 1 имеет экспоненциально распределенное с параметром λ время жизни и по окончании его превращается в частицу типа 2. Частица типа 2 имеет экспоненциально распределенное с параметром λ время жизни и по окончании его превращается в две частицы типа 1. Таким образом, в обозначениях данного параграфа имеем

$$a^{(1)}_{k_1, k_2} = \begin{cases} -1, & \text{если} \quad k_1 = 1, \ k_2 = 0, \\ +1, & \text{если} \quad k_1 = 0, \ k_2 = 1, \\ 0 & \text{в противном случае,} \end{cases}$$

$$a^{(2)}_{k_1, k_2} = \begin{cases} +1, & \text{если} \quad k_1 = 2, \ k_2 = 0, \\ -1, & \text{если} \quad k_1 = 0, \ k_2 = 1, \\ 0 & \text{в противном случае,} \end{cases}$$

где для простоты предполагается, что $\lambda = 1$. Тогда

$$u^{(1)}(s_1, s_2) = -s_1 + s_2$$

и

$$u^{(2)}(s_1, s_2) = s_1^2 - s_2.$$

Соотношения (10.5) и (10.6) примут соответствующий частный вид. Производящая функция величины $X(t)$ при условии $X(0) = 1$ может быть получена из $\phi^{(1)}(t;\ s_1,\ s_2)$ при $s_1 = s_2 = s$.

§ 11. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ, ЗАВИСЯЩИЕ ОТ ВОЗРАСТА

В этом параграфе мы рассмотрим модель ветвящегося процесса, где каждый объект (или частица, или индивидуум) имеет случайное время жизни с произвольным распределением, а по окончании его порождает своих потомков. Этот процесс следует сравнить с ветвящимися процессами, у которых время жизни одного объекта фиксировано или экспоненциально распределено. Предположим, что отдельный объект имеет время жизни случайной длительности T с плотностью $f(t)$, т. е. вероятность того, что время жизни объекта лежит в интервале $(t,\ t + dt)$, равна $f(t)dt$. Предположим далее, что по окончании времени жизни объект порождает два новых объекта такого же типа, времена жизни которых будут независимыми случайными величинами с той же плотностью $f(t)$. По окончании своего времени жизни каждый объект снова породит два новых объекта того же типа, и этот процесс продолжается бесконечно долго. Пусть $N(t)$ — число объектов в момент t; обозначим его вероятностное распределение через

$$p_k(t) = \mathrm{P}\{N(t) = k\}, \qquad k = 0,\ 1,\ 2,\ \ldots .$$

Очевидно, $p_0(t) = 0$ при всех $t \geqslant 0$, поскольку в наличии всегда будет по крайней мере один объект. Действительно, до первого «раздвоения» будет в точности один объект, а после него — по крайней мере 2 объекта. Таким образом,

$$p_1(t) = \mathrm{P}\{N(t) = 1\} = \mathrm{P}\{T > t\} = 1 - F(t),$$

где

$$F(t) = \mathrm{P}\{T \leqslant t\} = \int_0^t f(\tau)\,d\tau$$

— функция распределения величины T.

Пусть $G(s,\ t)$ — производящая функция величины $N(t)$, т. е.

$$G(s,\ t) = \sum_{k=0}^{\infty} p_k(t)\,s^k = \sum_{k=1}^{\infty} p_k(t)\,s^k.$$

Получим интегральное уравнение относительно $G(s,\ t)$. Вероятность $p_k(t)$ того, что в момент t имеется ровно k объектов, можно выразить следующим образом. Предположим, что первое ветвление происходит на интервале $(\tau, \tau + d\tau)$, $0 \leqslant \tau \leqslant t$, вероятность чего равна $f(\tau)d\tau$, и за оставшееся время $t - \tau$ оба новых независимо развивающихся объекта породят в сумме k потомков. Естественно, что момент первого ветвления τ может принимать любые значения из отрезка $[0, t]$. Таким образом, из формулы полной вероятности следует, что

$$p_k(t) = \int_0^t f(\tau)\,d\tau \sum_{l=1}^{k} p_l(t-\tau)\,p_{k-l}(t-\tau), \qquad k=2,\ 3,\ \ldots,$$

$$p_1(t) = 1 - F(t).$$

В силу определения $G(s, t)$ имеем

$$G(s,\ t) = [1 - F(t)]\,s + \sum_{k=2}^{\infty} s^k \int_0^t d\tau f(\tau) \sum_{l=1}^{\infty} p_l(t-\tau)\,p_{k-l}(t-\tau) =$$

$$= [1 - F(t)]\,s + \sum_{k=0}^{\infty} s^k \int_0^t d\tau f(\tau) \sum_{l=0}^{k} p_l(t-\tau)\,p_{k-l}(t-\tau).$$

Поскольку все входящие под знак суммирования величины неотрицательны, знаки суммирования и интегрирования можно поменять местами. Тогда

$$G(s,\ t) = \int_0^t d\tau f(\tau) \sum_{k=0}^{\infty} s^k \sum_{l=0}^{k} p_l(t-\tau)\,p_{k-l}(t-\tau) + [1 - F(t)]\,s.$$

Двойная сумма в правой части является производящей функцией двукратной свертки последовательности $\{p_k(t-\tau)\}$. Таким образом,

$$G(s,\ t) = \int_0^t [G(s,\ t-\tau)]^2 f(\tau)\,d\tau + [1 - F(t)]\,s. \tag{11.1}$$

К сожалению, в общем случае данное интегральное уравнение не решается. Решим его для частного случая, когда T имеет экспоненциальное распределение с плотностью

$$f(t) = \lambda e^{-\lambda t}, \qquad t \geqslant 0. \tag{11.2}$$

Процесс, соответствующий данному частному случаю, является процессом чистого рождения Юла. Действительно, если вначале имеется n объектов, то время до первого ветвления является случайной величиной $Z = \min(X_1, X_2, \ldots, X_n)$, где X_i независимы и имеют плотность (11.2). Распределение величины Z является экспоненциальным с параметром $n\lambda$. Следовательно, вероятность того, что ветвление произойдет на интервале длины h, равна $n\lambda h + o(h)$. Когда это случается, популяция увеличивается до размера $n+1$

и временной интервал до следующего ветвления будет распределен экспоненциально с параметром $(n+1)\lambda$ и т. д. Анализ этого примера с точки зрения процессов чистого рождения был дан в гл. 7, § 1. Другой метод, излагаемый ниже, имеет самостоятельный интерес.

Если $1 - F(t) = e^{-\lambda t}$, то уравнение (11.1) примет вид

$$G(s,\ t)\, e^{\lambda t} = \lambda \int_0^t [G(s,\ t-\tau)]^2 e^{\lambda(t-\tau)}\, d\tau + s.$$

Произведем замену переменных $u = t - \tau$ и получим

$$G(s,\ t)\, e^{\lambda t} = \lambda \int_0^t [G(s,\ u)]^2 e^{\lambda u}\, du + s.$$

Продифференцируем полученное равенство по t:

$$e^{\lambda t} G'(s,\ t) + \lambda e^{\lambda t} G(s,\ t) = \lambda [G(s,\ t)]^2 e^{\lambda t},$$

где

$$G'(s,\ t) = \frac{\partial}{\partial t} G(s,\ t).$$

Разделив обе части на $e^{\lambda t}$, получим дифференциальное уравнение Бернулли

$$G'(s,\ t) = \lambda [G(s,\ t)]^2 - \lambda G(s,\ t). \tag{11.3}$$

Чтобы решить это дифференциальное уравнение, можно разделить переменные:

$$\frac{dG(s,\ t)}{G(s,\ t)\,[G(s,\ t) - 1]} = \lambda\, dt.$$

Тогда решение имеет вид

$$\frac{G(s,\ t) - 1}{G(s,\ t)} = C(s)\, e^{\lambda t},$$

или

$$G(s,\ t) = \frac{1}{1 - C(s)\, e^{\lambda t}}, \tag{11.4}$$

где $C(s)$ не зависит от t, но, быть может, зависит от s. Чтобы найти $C(s)$, положим в (11.4) $t = 0$. Поскольку

$$p_k(0) = \begin{cases} 0, & \text{если} \quad k \neq 1, \\ 1, & \text{если} \quad k = 1, \end{cases} \qquad s \equiv G(s,\ 0) = \frac{1}{1 - C(s)},$$

мы имеем

$$C(s) = (s-1)/s.$$

Таким образом, решение уравнения (11.3), а также уравнения (11.1) в случае экспоненциального времени жизни равно

$$G(s,\ t) = \frac{se^{-\lambda t}}{1-\left(1-e^{-\lambda t}\right)s}. \tag{11.5}$$

Чтобы найти точный вид $p_k(t)$, разложим (11.5) в ряд по степеням s:

$$G(s,\ t) = e^{-\lambda t} s \sum_{k=0}^{\infty} (1-e^{-\lambda t})^k s^k.$$

Очевидно, что

$$p_k(t) = e^{-\lambda t}(1-e^{-\lambda t})^{k-1}, \qquad k = 1,\ 2,\ \ldots.$$

Хотя мы не можем в общем случае решить интегральное уравнение (11.1), можно получить из него уравнение относительно среднего $m(t) = \mathrm{M}(N(t))$. Вспомним, что

$$\left.\frac{\partial G(s,\ t)}{\partial s}\right|_{s=1} = \sum_{k=1}^{\infty} k p_k(t) = m(t).$$

Дифференцируя (11.1) по s, получаем

$$\frac{\partial G(s,\ t)}{\partial s} = 2\int_0^t G(s,\ t-\tau)\frac{\partial G(s,\ t-\tau)}{\partial s} f(\tau)\,d\tau + 1 - F(t).$$

Положим теперь $s = 1$ и учтем, что

$$G(1,\ t-\tau) = \sum_{k=1}^{\infty} p_k(t-\tau) = 1;$$

тогда

$$m(t) = 2\int_0^t m(t-\tau) f(\tau)\,d\tau + 1 - F(t).$$

Это интегральное уравнение является примером так называемого уравнения восстановления. Его характерной особенностью является наличие неизвестной функции под знаком интеграла в виде свертки. Существует множество методов, которые описывают асимптотические свойства решения $m(t)$ уравнения восстановления.

Очевидным обобщением описанной модели является допущение о том, что объект по истечении его времени жизни порождает ровно r новых объектов того же типа, где r — фиксированное целое число, $r \geqslant 2$. Легко видеть, что интегральное уравнение (11.1) заменяется тогда следующим:

$$G(s,\ t) = \int_0^t [G(s,\ t-\tau)]^r f(\tau)\,d\tau + [1 - F(t)]\,s.$$

Дальнейшее обобщение этой модели состоит в допущении того, что любой объект по истечении своего времени жизни может породить случайное число новых объектов того же типа, например можно предположить, что объект порождает l новых объектов с вероятностью q_l, $l = 0, 1, 2, \ldots$. Пусть

$$h(s) = \sum_{l=0}^{\infty} q_l s^l$$

— соответствующая производящая функция. При этом интегральное уравнение можно получить следующим образом. Предположим, что первое ветвление произошло в интервале $(\tau, \tau + d\tau)$, $0 \leqslant \tau \leqslant t$, и при этом образовалось l новых объектов. Это событие имеет вероятность $q_l f(\tau)\,d\tau$. Тогда в течение оставшегося времени $t - \tau$ каждый из l объектов может породить любое число новых объектов, но так, чтобы их суммарное число в момент t равнялось k. В силу формулы полной вероятности имеем

$$p_k(t) = \int_0^t d\tau f(\tau) \sum_{l=0}^{\infty} q_l \sum_{k_1+k_2+\ldots+k_l=k} p_{k_1}(t-\tau)\,p_{k_2}(t-\tau)\,\ldots\,p_{k_l}(t-\tau),$$
$$k = 2, 3, \ldots,$$

$$p_1(t) = [1 - F(t)] + \int_0^t d\tau f(\tau) \sum_{l=0}^{\infty} q_l \sum_{k_1+k_2+\ldots+k_l=1} p_{k_1}(t-\tau)\,p_{k_2}(t-\tau)\,\ldots\,p_{k_l}(t-\tau)\ [1]).$$

Тогда производящая функция равна

$$G(s, t) = [1 - F(t)]\,s + \sum_{k=0}^{\infty} s^k \int_0^t d\tau f(\tau) \sum_{l=0}^{\infty} q_l \sum_{k_1+\ldots+k_l=k} p_{k_1}(t-\tau)\,\ldots\,p_{k_l}(t-\tau) =$$
$$= [1 - F(t)]\,s + \int_0^t d\tau f(\tau) \sum_{l=0}^{\infty} q_l \sum_{k=0}^{\infty} s^k \sum_{k_1+\ldots+k_l=k} p_{k_1}(t-\tau)\,\ldots\,p_{k_l}(t-\tau).$$

Но сумма

$$\sum_{k=0}^{\infty} s^k \sum_{k_1+\ldots+k_l=k} p_{k_1}(t-\tau)\,\ldots\,p_{k_l}(t-\tau)$$

[1]) Отличие от первоначально рассмотренного случая бинарного деления в выражении для $p_1(t)$ проистекает от того, что если в момент t размер популяции равен 1, то это означает, что либо до момента t не произошло ни одного ветвления (вероятность чего равна $1 - F(t)$), либо ветвление произошло до момента t и к моменту t размер популяции вновь стал равным 1 (вероятность этого события выражается вторым слагаемым). — *Прим. перев.*

является производящей функцией l-кратной свертки последовательности $\{p_k(t-\tau)\}$ (читатель должен это проверить). Таким образом, эта сумма равна l-й степени производящей функции

$$\sum_{k=0}^{\infty} s^k p_k(t-\tau) = G(s,\ t-\tau),$$

т. е. величине $[G(s, t-\tau)]^l$, и поэтому

$$G(s,\ t) = [1-F(t)]\,s + \int_0^t d\tau f(\tau) \sum_{l=0}^{\infty} q_l [G(s,\ t-\tau)]^l.$$

Но сумма под интегралом равна производящей функции $h(\cdot)$ в точке $G(s,\ t-\tau)$, и окончательно интегральное уравнение принимает вид

$$G(s,\ t) = \int_0^t h(G(s,\ t-\tau))\,f(\tau)\,d\tau + [1-F(t)]\,s.$$

ЗАДАЧИ

1. Пусть X_n — ветвящийся процесс, $X_0 = 1$. Для произвольного, но фиксированного положительного целого k определим последовательность

$$Y_r = X_{rk}, \qquad r = 0,\ 1,\ 2,\ \ldots .$$

Показать, что последовательность $\{Y_r,\ r = 0,\ 1,\ 2,\ \ldots\}$ образует ветвящийся процесс. Кроме того, доказать, что если $\varphi(s)$ — производящая функция числа прямых потомков одного индивидуума для процесса $\{X_n\}$, а $\varphi_n(x)$ — ее n-я итерация, то $\varphi_k(s)$ — производящая функция числа прямых потомков одного индивидуума для процесса $\{Y_r\}$.

2. Пусть $f(s) = 1 - p(1-s)^{\beta}$, где p и β — постоянные, $0 < p < 1$, $0 < \beta < 1$. Доказать, что $f(s)$ — производящая функция и ее итерации имеют вид

$$f_n(s) = 1 - p^{1+\beta+\ldots+\beta^{n-1}}(1-s)^{\beta^n}, \qquad n = 1,\ 2,\ \ldots .$$

3. Предположим, что $f(s)$ — производящая функция, а $h(s)$ — функция, такая, что

$$g(s) = h^{-1}[f(h(s))]$$

— производящая функция. Проверить, что

$$g_n(s) = h^{-1}[f_n(h(s))]$$

— производящая функция, где f_n и g_n — функциональные итерации функций f и g соответственно.

4. В качестве иллюстрации к задаче 3 возьмем

$$f(s) = \frac{s}{m-(m-1)s}, \qquad m > 1,$$

и $h(s) = s^k$, $k > 0$ — целое число.

Доказать, что $g(s) = h^{-1}f(h(s))$ — производящая функция и что n-я итерация функции g равна

$$g_n(s) = \frac{s}{(m^n - (m^n - 1)\, s^k)^{1/k}}.$$

5. Показать, что $M\left(\sum_{n=1}^{\infty} X_n\right) = m/(1-m)$, если $m = M(X_1) < 1$, $X_0 = 1$, а $\{X_n\}$ — ветвящийся процесс.

6. В момент 0 в культуре клеток крови имеется один эритроцит. Через минуту эритроцит погибает и заменяется одной из следующих комбинаций с соответствующей вероятностью (показана справа):

2 эритроцита	$^1/_4$,
1 эритроцит, 1 лейкоцит	$^2/_3$,
2 лейкоцита	$^1/_{12}$.

Каждый эритроцит живет одну минуту и «порождает потомство» указанным образом. Каждый лейкоцит живет одну минуту и погибает, не оставляя никакого потомства. Предположим, что клетки развиваются независимо.

(а) Чему равна вероятность того, что к моменту $n + 1/2$ не появилось ни одного лейкоцита?

(б) Чему равна вероятность того, что культура клеток крови погибает?

Ответ: (а) $(^1/_4)^{2^{n+1}-1}$; (б) $^1/_3$.

7. Пусть $f(s) = as^2 + bs + c$, где a, b, $c > 0$ и $f(1) = 1$. Предположим, что вероятность вырождения равна d $(0 < d < 1)$. Доказать, что $d = c/a$.

8. Предположим, что в ветвящемся процессе число потомков исходной частицы имеет распределение с производящей функцией $f(s)$. Каждый член первого поколения порождает случайное число потомков, функция распределения которого имеет производящую функцию $g(s)$. Следующее поколение вновь имеет производящую функцию f, затем g и т. д.

Исходя из общих принципов (т. е. не используя какие-либо общие результаты, полученные в § 2 для процессов с несколькими типами частиц), найти вероятность вырождения процесса и среднее число частиц в n-м поколении (например, для четных n). Изменятся ли эти величины, если поменять f и g местами?

9. Рассмотрим ветвящийся процесс с дискретным временем $\{X_n\}$, где $X_0 = 1$. Доказать простое неравенство $P\{X_n > L$ при некотором $0 \leqslant n \leqslant m \mid X_m = 0\} \leqslant [P\{X_m = 0\}]^L$.

10. Для исследования некоторого урологического процесса была введена следующая модель. Предположим, что бактерии растут в соответствии с процессом Юла с параметром λ (см. § 1 гл. 7). В каждую единицу времени каждая бактерия уничтожается с вероятностью p. Чему равна производящая функция числа бактерий, существующих в момент n?

Указание: Эта производящая функция является n-й итерацией производящей функции, характеризующей некоторый ветвящийся процесс.

Ответ: $f_n(s)$ является n-й итерацией функции

$$f(s) = \frac{e^{-\lambda}(p + qs)}{1 - \left(1 - e^{-\lambda}\right)(p + qs)}.$$

11. (а) «Зрелые» индивидуумы порождают потомство в соответствии с производящей функцией $f(s)$. Предположим, что имеется популяция, состоящая из k «незрелых» индивидуумов, каждый из которых с вероятностью p достигает зрелости и затем размножается независимо от других индивидуумов. Найти производящую функцию числа (незрелых) индивидуумов в следующем поколении.

(б) Найти производящую функцию числа «зрелых» индивидуумов в следующем поколении при условии, что в родительском поколении имеется k «зрелых» индивидуумов.

Ответ: (а) $(1-p+pf(s))^k$; (б) $(f(1-p+ps))^k$.

12. Показать, что распределения (а) и (б) в задаче 11 имеют одно и то же среднее, но в общем случае разные дисперсии.

13. Рассмотрим ветвящийся процесс $\{X_n\}$ с дискретным временем и производящей функцией

$$\varphi(s)=\frac{1-(b+c)}{1-c}+\frac{bs}{1-cs}, \qquad 0<c<b+c<1,$$

где $(1-b-c)/c(1-c)>1$. Предположим, что $X_0=1$. Найти условное предельное распределение

$$\lim_{n\to\infty} \mathrm{P}\{X_n=k \mid X_n>0\}.$$

Ответ:

$$\left(1-\frac{1}{s_0}\right)\left(\frac{1}{s_0}\right)^{k-1}, \quad s_0=\frac{1-b-c}{c(1-c)}.$$

14. Предположим, что в предыдущей задаче $1-b-c=c(1-c)$. Определить $\mathrm{P}\{X_n>0\}$.

Ответ: $(1-c)/[1+(n-1)c]$.

15. При условиях задачи 14 доказать, что $\mathrm{P}\{X_n\leqslant nx \mid X_n>0\}$ сходится при $n\to\infty$ к экспоненциальному распределению.

Указание: Найти преобразование Лапласа распределения величины X_n/n при условии $X_n>0$ и найти его предел при $n\to\infty$.

Ответ: Экспоненциальное распределение с параметром $(1-c)/c$.

16. Рассмотрим ветвящийся процесс с начальным размером популяции N и производящей функцией

$$\varphi(s)=q+ps, \quad q,\ p>0, \quad q+p=1.$$

Найти распределение времени T вырождения популяции.

Ответ: $\mathrm{P}\{T=n\}=(1-p^{n+1})^N-(1-p^n)^N$.

17. Пусть $\{X_n, n\geqslant 0\}$ — ветвящийся процесс с соответствующей производящей функцией $\varphi(s)$. Пусть Y_n — суммарное число индивидуумов в первых n поколениях, т. е.

$$Y_n=X_0+X_1+\ldots+X_n, \qquad n=0,\ 1,\ 2,\ \ldots, \quad X_0=1.$$

Пусть $F_n(x)$ — производящая функция величины Y_n. Доказать функциональное соотношение

$$F_{n+1}(s)=s\varphi(F_n(s)), \qquad n=0,\ 1,\ 2,\ \ldots.$$

18. Пусть $\varphi(s)$ — производящая функция числа потомков одного индивидуума в ветвящемся процессе, в котором в момент 0 имеется один индивидуум. Пусть $\varphi_n(s)$ — его n-я итерация. Предположим, кроме того, что имеется иммиграция в популяцию. Размер иммиграции за время одного поколения описывается производящей функцией $h(s)$. Рассмотрим ветвящийся процесс с иммиграцией, переходная матрица которого задается соотношением

$$\sum_{j=0}^{\infty} P_{ij}s^j=[\varphi(s)]^i\,h(s).$$

Доказать, что матрица переходных вероятностей за n шагов определяется соотношением

$$\sum_{j=0}^{\infty} P_{ij}^n s^j = [\varphi_n(s)]^i \, h(\varphi_{n-1}(s)) \, h(\varphi_{n-2}(s)) \ldots h(\varphi(s)) \, h(s).$$

19. Предположим, что в ветвящемся процессе с иммиграцией (см. задачу 18) $\varphi'(1) = m < 1$. Доказать, что соответствующая цепь Маркова имеет стационарное распределение с производящей функцией $\pi(s) = \sum_{r=0}^{\infty} \pi_r s^r$, удовлетворяющей функциональному уравнению

$$\pi(\varphi(s)) \, h(s) = \pi(s).$$

20. В предположениях задачи 19 для частного случая $\varphi(s) = q + ps$ $(0 < p < 1, q + p = 1)$ и $h(s) = e^{s-1}$ найти стационарное распределение.

21. Найти $\sigma^2(X(t))$, где $X(t)$ — ветвящийся процесс с непрерывным временем и $X(0) = 1$.

Ответ:

$$\sigma^2(X(t)) = \begin{cases} \left[\dfrac{u''(1) - u'(1)}{u'(1)}\right] e^{u'(1)\,t}\left(e^{u'(1)\,t} - 1\right), & \text{если } u'(1) \neq 0, \\ u''(1)\,t, & \text{если } u'(1) = 0. \end{cases}$$

22. Найти производящую функцию $\varphi(t; s)$ для ветвящегося процесса с непрерывным временем и производящей функцией

$$u(s) = s^k - s \quad (k \geqslant 2 \text{ — целое число}).$$

Указание: Решить уравнение

$$\frac{\partial \varphi(t; s)}{\partial t} = u(\varphi(t; s)), \quad \varphi(0; s) = s.$$

Ответ:

$$\varphi(t; s) = s\left[e^{(k-1)\,t} - \left(e^{(k-1)\,t} - 1\right) s^{k-1}\right]^{-1/(k-1)}.$$

23. Найти производящую функцию $\varphi(t; s)$ для ветвящегося процесса с непрерывным временем и производящей функцией инфинитезимальных параметров

$$u(s) = 1 - s - \sqrt{1 - s}.$$

Ответ:

$$\varphi(t; s) = 1 - \left[1 - e^{-t/2} + e^{-t/2}\sqrt{1 - s}\right]^2.$$

24. Рассмотрим ветвящийся процесс с непрерывным временем и начальным размером популяции N. Обозначим размер популяции в момент t через $X_N(t)$. Пусть $\lambda_k h + o(h)$, $k = 0, 2, 3,$ — вероятность того, что индивидуум породит k индивидуумов в интервале $(t, t + h)$. Предположим, что $o(h) \to 0$ при $h \to 0$ равномерно по k и t. Пусть $1 - \lambda h + o(h)$ — вероятность того, что в интервале $(t, t + h)$ ветвления нет $(\lambda = \lambda_0 + \lambda_2 + \lambda_3 + \ldots)$. Предположим, что $\sum k^2 \lambda_k < \infty$. Предположим также, что индивидуумы развиваются независимо. Пусть $Y(t)$ — число моментов ветвления на отрезке $[0, t]$.

Показать, что

$$\mathrm{P}\{Y(t + T) - Y(t) = 0 \mid X_N(t) = n\} = e^{-n\lambda T}.$$

25. Рассмотрим процесс Юла, в котором каждый член популяции с вероятностью $\beta h + o(h)$ порождает k новых членов, а с вероятностью $1 - \beta h + o(h)$ не происходит рождений на интервале длины h ($\beta > 0$, k — положительное целое число). Предположим, что в момент 0 имеется N членов популяции.

(а) Пусть $X(t)$ — число моментов ветвления до момента t. Найти характер роста $\mathrm{M}(X(t))$.

(б) Пусть τ_n — момент n-го ветвления. Найти плотность величины τ_n.

Указание: Учесть, что

$$\mathrm{P}\{\tau_n \leqslant t \mid \tau_{n-1} = \xi\} = \begin{cases} 1 - \exp\{-[(n-1)k + N]\beta(t-\xi)\}, & \xi \leqslant t, \\ 0, & \xi > t, \end{cases}$$

и получить рекуррентную формулу для плотности величины τ_n через плотность величины τ_{n-1}.

Ответ: (а) $e^{-k\beta t}\,\mathrm{M}(X(t)) \to N$ при $t \to \infty$;

(б) плотность величины τ_n равна

$$f_n(t) = \frac{N(N+k)\dots[N+(n-1)k]}{(n-1)!\,k^{n-1}}\,\beta e^{-N\beta t}\left(1 - e^{-k\beta t}\right)^{n-1}.$$

***26.** Рассмотрим простой процесс рождения и гибели (линейный рост без иммиграции), т. е. $\lambda_n = \lambda n$ и $\mu_n = \mu n$, где $\lambda > 0$, $\mu > 0$ и $\mu > \lambda$. Пусть $Z(t)$ — размер популяции в момент t. С помощью подходящей интерпретации показать, что период занятости для процесса обслуживания с бесконечным числом обслуживающих устройств, где интервалы между поступающими требованиями распределены по закону $1 - e^{-\lambda t}$, а функция распределения времени обслуживания равна $1 - e^{-\mu t}$, имеет такое же распределение, что и величина $\int\limits_0^\infty Z(t)\,dt$, где $Z(0) = 1$.

27. Пусть $\{X_n\}$ — ветвящийся процесс с соответствующей производящей функцией $\varphi(s)$, и пусть $\varphi_n(s) = \sum\limits_{k=0}^{\infty} \mathrm{P}\{X_n = k\}\,s^k$. Предположим, что $\varphi'(1) > 1$. Пусть $\widetilde{X}_n$ — число всех частиц в n-м поколении, имеющих бесконечное число поколений потомков.

Показать, что производящая функция величины $\widetilde{X}_n$ равна

$$\sum_{k=0}^{\infty} \mathrm{P}\{\widetilde{X}_n = k \mid \widetilde{X}_0 = X_0 = 1\}\,s^k = \frac{\varphi_n(s(1-q)+q) - q}{1-q},$$

где q — вероятность вырождения.

Указание: Заметить, что при $k \geqslant 1$

$$\mathrm{P}\{\widetilde{X}_n = k \mid \widetilde{X}_0 = 1,\ X_0 = 1\} = \frac{\sum\limits_{l=k}^{\infty} \mathrm{P}\{\widetilde{X}_n = k,\ X_n = l \mid X_0 = 1\}}{\mathrm{P}\{\widetilde{X}_0 = 1 \mid X_0 = 1\}}.$$

28. Популяция состоит из двух типов индивидуумов — мужских и женских особей. Предположим, что все женские особи могут производить потомство в соответствии с производящей функцией $f(x)$ при условии, что в популяции имеется по крайней мере одна мужская особь. Если в потомстве с вероятностью α появляются женские особи, то чему равна производящая функция числа родившихся женских особей при условии рождения хотя бы одной мужской особи.

Ответ: $\dfrac{f(\alpha s + (1-\alpha)) - f(\alpha s)}{1 - f(\alpha)}$.

ЗАМЕЧАНИЯ

Материал этой главы следует книге Т. Харриса [1] по ветвящимся процессам, в которой также содержится обширная библиография по данному предмету и его приложениям.

ЛИТЕРАТУРА

1. Харрис Т., Теория ветвящихся случайных процессов, изд-во «Мир», 1966.

Глава 12

СОСТАВНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

В этой главе будет рассмотрен ряд отдельных вероятностных моделей, имеющих отношение к приложениям в астрономии, биологии, технике и физике. Эти процессы имеют своими составными частями различные классические процессы, включая пуассоновские, ветвящиеся и процессы роста диффузионного типа. Во всех случаях подчиненный процесс определяется через основной, и его исследуют, изучая распределения состояний основного процесса. В § 1 будут рассмотрены многомерные пуассоновские процессы, а в следующем параграфе будет дано их приложение к астрономии. Понятие многомерного пуассоновского процесса будет играть важную роль при определении каскадных, или составных, случайных процессов. Некоторые из них будут изучены в последующих параграфах этой главы (см., например, § 2).

В § 3 будет исследована вероятностная модель роста и иммиграции. В § 4 будет определен случайный процесс роста двух типов индивидуумов — нормального и мутантного. Популяция «диких» (т. е. нормальных, немутантных) типов растет детерминированным образом, в то время как популяция мутантных типов растет в соответствии с законами марковского ветвящегося процесса. Кроме того, каждый индивидуум нормального типа в момент гибели (время его жизни распределено экспоненциально) превращается в мутантный тип, а затем развивается подобно другим индивидуумам мутантного типа.

В § 5 и 6 рассматриваются вероятностные модели роста, учитывающие фактор географического распределения и распространения популяции, а также и процесс естественного роста.

Исследуемые случайные процессы типичны для широкого класса общих каскадных процессов. Цель настоящей главы — познакомить читателя с задачами, связанными с комбинациями случайных процессов, имеющими множество приложений и требующими для своего анализа большого искусства.

В § 7 дан обзор некоторых детерминированных моделей роста популяций, учитывающих их возрастную структуру.

§ 1. МНОГОМЕРНЫЕ ОДНОРОДНЫЕ ПУАССОНОВСКИЕ ПРОЦЕССЫ

Пуассоновские процессы были введены в гл. 7; при этом параметр процесса являлся действительным положительным числом $t \geqslant 0$ и обычно считался временем. Введем теперь вариант пуассоновского процесса, у которого значение параметра определяется мерой множества на плоскости, в трехмерном пространстве или в пространствах более общего вида. Цель данного параграфа — определить некоторые варианты многомерных пуассоновских процессов и описать некоторые примеры этих процессов и их приложения.

В гл. 7 пуассоновский процесс $X(t)$, $t \geqslant 0$, был введен аксиоматически. Было доказано, что $X(t)$ имеет вероятностное распределение

$$\mathrm{P}\{X(t) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad t \geqslant 0, \qquad k = 0,\ 1,\ 2,\ \ldots,$$

где λ — положительная постоянная, интерпретируемая как средняя интенсивность наступления событий. В данном параграфе мы введем постулаты, характеризующие однородный пространственный пуассоновский процесс $X(S)$, где параметр S является произвольной областью плоскости или пространства, имеющей конечную меру, а $X(S)$ обладает вероятностным распределением

$$\mathrm{P}\{X(S) = k\} = e^{-\lambda A(S)} \frac{[\lambda A(S)]^k}{k!}, \qquad k = 0,\ 1,\ 2,\ \ldots. \tag{1.1}$$

Здесь λ — положительная постоянная, называемая интенсивностью (или параметром) процесса, а $A(S)$ — площадь или объем области S в зависимости от того, является ли область S частью плоскости или пространства. Введем следующие постулаты:

(1) $X(S)$ принимает лишь неотрицательные целочисленные значения и $0 < \mathrm{P}\{X(S) = 0\} < 1$, если $A(S) > 0$.

(2) Вероятностное распределение величины $X(S)$ зависит только от $A(S)$, и, кроме того, если $A(S) \to 0$, то $P\{X(S) \geqslant 1\} \to 0$.

(3) Если $S_1, S_2, \ldots, S_n$ $(n \geqslant 1)$ — непересекающиеся области, то $X(S_1), \ldots, X(S_n)$ — независимые в совокупности случайные величины и

$$X(S_1 \cup \ldots \cup S_n) = X(S_1) + \ldots + X(S_n).$$

(4) Выполняется требование

$$\lim_{A(S) \to 0} \frac{\mathrm{P}\{X(S) \geqslant 1\}}{\mathrm{P}\{X(S) = 1\}} = 1.$$

Прежде чем переходить к описательному обсуждению этих аксиом, полезно привести некоторые примеры.

(а) В трехмерном пространстве $X(S)$ может представлять собой число звезд, расположенных в области S.

(б) На плоскости $X(S)$ может представлять собой число бактерий определенного вида, содержащихся в области S.

Объяснение и интерпретация введенных аксиом вполне очевидны. Постулат (2) утверждает, что $X(S)$ зависит не от вида области S, а только от ее площади или объема. Такое предположение представляется разумным (см. примеры (а) и (б)). В соответствии с постулатом (3), если использовать терминологию примера (а), количества звезд, содержащихся в непересекающихся областях, являются независимыми случайными величинами, а значение $X(S)$ для суммарной области является суммой значений $X(\cdot)$ для составляющих областей. По-видимому, предположение о независимости является разумным приближением к реальной ситуации распределения звезд. Постулат (4) достаточно понятен интуитивно и не требует объяснений.

Основной целью этого параграфа является доказательство следующей теоремы:

Теорема 1.1. *Если случайный процесс* $X(S)$, *определенный относительно областей* S *евклидова пространства размерности* n, *удовлетворяет постулатам* (1) — (4), *то* $X(S)$ *имеет распределение* (1.1).

Доказательство. Рассмотрим произвольную область S, такую, что $0 < A(S) < \infty$. Разобьем S на непересекающиеся области $S_1, S_2, \ldots, S_n$ равной площади (объема), т. е.

$$S_1 \cup S_2 \cup \ldots \cup S_n = S,$$

$$S_i \cap S_j = \varnothing, \ i \neq j \qquad (\varnothing \text{ — пустое множество}),$$

$$A(S_i) = \frac{1}{n} A(S) \qquad \text{для всех } i = 1, 2, \ldots, n.$$

Тогда в силу постулата (3)

$$\mathrm{P}\{X(S) = 0\} = \mathrm{P}\{X(S_1 \cup \ldots \cup S_n) = 0\} = \\ = \mathrm{P}\{X(S_1) + \ldots + X(S_n) = 0\}.$$

Но из постулата (1) следует, что событие $\sum_{i=1}^{n} X(S_i) = 0$ может произойти тогда и только тогда, когда $X(S_i) = 0$ при всех $i = 1, 2, \ldots, n$; тогда, используя независимость величин $X(S_i)$, $i = 1, 2, \ldots, n$ (постулат (3)), получим

$$\mathrm{P}\{X(S) = 0\} = \mathrm{P}\{X(S_i) = 0, \ i = 1, 2, \ldots, n\} = \prod_{i=1}^{n} \mathrm{P}\{X(S_i) = 0\}.$$

Из постулата (2) следует, что $\mathrm{P}\{X(S_i) = 0\}$ зависит только от $A(S_i) = \frac{1}{n} A(S)$. Следовательно,

$$\mathrm{P}\{X(S_1) = 0\} = \mathrm{P}\{X(S_2) = 0\} = \ldots = \mathrm{P}\{X(S_n) = 0\}.$$

Таким образом, получаем

$$P\{X(S)=0\}=[P\{X(S_n)=0\}]^n. \tag{1.2}$$

Далее,

$$P\{X(S_n)=0\}=1-P\{X(S_n)\geqslant 1\}.$$

Взяв логарифм от обеих частей (1.2), что допустимо в силу постулата (1), получаем

$$-\ln P\{X(S)=0\}=-n\ln[1-P\{X(S_n)\geqslant 1\}]=$$
$$=n\left[P\{X(S_n)\geqslant 1\}+\frac{1}{2}(P\{X(S_n)\geqslant 1\})^2+\ \ldots\right], \tag{1.3}$$

где использовано разложение

$$-\ln(1-x)=x+\frac{1}{2}x^2+\frac{1}{3}x^3+\ \ldots,$$

справедливое при $-1\leqslant x<1$. Очевидно, что $0\leqslant P\{X(S_n)\geqslant 1\}<1$, поскольку в противном случае было бы $P\{X(S_n)=0\}=0$, откуда $P\{X(S)=0\}=0$. Это, однако, невозможно в силу постулата (1), поскольку по предположению $A(S)>0$. Формулу (1.3) можно переписать в виде

$$-\ln P\{X(S)=0\}=$$
$$=nP\{X(S_n)=1\}\left[\frac{P\{X(S_n)\geqslant 1\}}{P\{X(S_n)=1\}}(1+O(P\{X(S_n)\geqslant 1\}))\right]. \tag{1.4}$$

Символ $O(P\{X(S_n)\geqslant 1\})$ имеет обычный смысл, т. е. величина

$$\frac{O(P\{X(S_n)\geqslant 1\})}{P\{X(S_n)\geqslant 1\}}$$

ограничена при $n\to\infty$. Заметим, что в силу постулата (2) имеем

$$P\{X(S_n)\geqslant 1\}\to 0 \text{ при } n\to\infty,$$

поскольку при этом $A(S_n)=\frac{1}{n}A(S)\to 0$. Далее, из постулата (4) следует, что

$$\lim_{n\to\infty}\frac{P\{X(S_n)\geqslant 1\}}{P\{X(S_n)=1\}}=\lim_{A(S_n)\to 0}\frac{P\{X(S_n)\geqslant 1\}}{P\{X(S_n)=1\}}=1.$$

Следовательно, из (1.4) при $n\to\infty$ получаем

$$-\ln P\{X(S)=0\}=\lim_{n\to\infty}nP\{X(S_n)=1\}. \tag{1.5}$$

В силу постулата (1) левая часть равенства с необходимостью должна быть положительной и конечной.

Рассмотрим производящие функции величин $X(S)$ и $X(S_n)$:

$$g(s) = \mathrm{M}\left(s^{X(S)}\right) = \sum_{k=0}^{\infty} \mathrm{P}\{X(S) = k\}\, s^k,$$

$$g_n(s) = \mathrm{M}\left(s^{X(S_n)}\right) = \sum_{k=0}^{\infty} \mathrm{P}\{X(S_n) = k\}\, s^k.$$

В силу постулатов (2) и (3)

$$g(s) = \mathrm{M}\left(s^{X(S)}\right) = \mathrm{M}\left(s^{X(S_1) + \dots + X(S_n)}\right) = \prod_{i=1}^{n} \mathrm{M}\left(s^{X(S_i)}\right) = \left[\mathrm{M}\left(s^{X(S_n)}\right)\right]^n,$$

т. е.

$$g(s) = [g_n(s)]^n. \tag{1.6}$$

Можно записать $g_n(s)$ в виде

$$g_n(s) = \mathrm{P}\{X(S_n) = 0\} + \mathrm{P}\{X(S_n) = 1\}\, s + \mathrm{P}\{X(S_n) > 1\}\, \theta(s),$$

где $|\theta(s)| \leqslant 1$, но

$$\mathrm{P}\{X(S_n) = 0\} = 1 - \mathrm{P}\{X(S_n) = 1\} - \mathrm{P}\{X(S_n) > 1\}.$$

Следовательно, подставляя вместо $\mathrm{P}\{X(S_n) = 0\}$ это выражение, получим

$$g_n(s) = 1 + (s-1)\, \mathrm{P}\{X(S_n) = 1\} + (\theta(s) - 1)\, \mathrm{P}\{X(S_n) > 1\}. \tag{1.7}$$

Используем теперь постулат (4), который утверждает, что

$$\frac{\mathrm{P}\{X(S_n) > 1\}}{\mathrm{P}\{X(S_n) = 1\}} \to 0 \text{ при } n \to \infty. \tag{1.8}$$

Кроме того, выполняется условие $\mathrm{P}\{X(S_n) = 1\} \to 0$ при $n \to \infty$. В самом деле, мы выше установили (см. (1.6)), что

$$g(s) = \left[\mathrm{M}\left(s^{X(S_n)}\right)\right]^n,$$

или, что то же самое,

$$[g(s)]^{\frac{1}{n}} = \sum_{k=0}^{\infty} \mathrm{P}\{X(S_n) = k\}\, s^k, \qquad 0 \leqslant s \leqslant 1.$$

В силу гипотезы (1) имеем $g(0) = \mathrm{P}\{X(S) = 0\} > 0$ и, следовательно,

$$\mathrm{P}\{X(S_n) = 1\} = \frac{d}{ds}[g(s)]^{\frac{1}{n}}\Big|_{s=0} = \frac{1}{n}\frac{g'(s)}{g(s)}[g(s)]^{\frac{1}{n}}\Big|_{s=0} =$$

$$= \frac{1}{n}\frac{g'(0)}{g(0)}[g(0)]^{\frac{1}{n}} \to 0 \text{ при } n \to \infty.$$

Далее, пользуясь разложением

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \ldots = z + o(z), \qquad |z| \to 0,$$

из формул (1.6), (1.7) и (1.8) получим

$$\ln g(s) = n \ln g_n(s) = n[(s-1)\,\mathrm{P}\{X(S_n)=1\} + \\ + (\theta(s)-1)\,\mathrm{P}\{X(S_n)>1\} + o(\mathrm{P}\{X(S_n)=1\})].$$

Взяв предел при $n \to \infty$ от обеих частей равенства и вновь использовав постулат (4) (в виде соотношения (1.8)), получим

$$\ln g(s) = (s-1) \lim_{n\to\infty} n\mathrm{P}\{X(S_n)=1\}. \tag{1.9}$$

Из (1.5) и (1.9) следует, что

$$\ln g(s) = -(s-1)\ln \mathrm{P}\{X(S)=0\},$$

или

$$g(s) = \exp[(s-1)(-\ln \mathrm{P}\{X(S)=0\})]. \tag{1.10}$$

Это выражение является производящей функцией пуассоновского распределения с математическим ожиданием

$$\mathrm{M}(X(S)) = g'(1) = -\ln \mathrm{P}\{X(S)=0\}.$$

Но математическое ожидание является неотрицательной аддитивной функцией, зависящей лишь от $A(S)$, откуда получаем

$$-\ln \mathrm{P}\{X(S)=0\} = \lambda A(S). \tag{1.11}$$

Формально последнее утверждение доказывается следующим образом. Пусть f — функция, удовлетворяющая равенству

$$\mathrm{M}(X(S)) = f(A(S)),$$

которое следует из постулата (2). Докажем теперь, что f — линейная функция. Пусть S_1 и S_2 — два непересекающихся множества, таких, что $A(S_1)$, $A(S_2) < \infty$; тогда

$$\mathrm{M}(X(S_1 \cup S_2)) = f(A(S_1) + A(S_2))$$

в силу аддитивности $A(S)$. С другой стороны, в силу постулата (3) имеем

$$\mathrm{M}(X(S_1 \cup S_2)) = \mathrm{M}(X(S_1)) + \mathrm{M}(X(S_2)) = f(A(S_1)) + f(A(S_2)).$$

Поскольку $A(S)$ может изменяться от 0 до ∞, то

$$f(x+y) = f(x) + f(y) \quad \text{для любых} \quad x,\ y \geqslant 0. \tag{1.12}$$

Кроме того, в силу определения $f(x) \geqslant 0$ и, очевидно, $f(0) = 0$. Единственным решением (1.12), обладающим указанными свойствами, является линейная функция $f(x) = \lambda x$ с некоторой постоянной λ (см. стр. 205) [1]). Равенство (1.11) доказано.

Из замечания после формулы (1.5) следует, что λ — действительный параметр. Подставляя в (1.10) равенство (1.11), получаем

$$g(s) = e^{\lambda A(S)(s-1)},$$

или, что то же самое,

$$g(s) = e^{-\lambda A(S)} \sum_{k=0}^{\infty} \frac{[\lambda A(S)]^k}{k!} s^k.$$

Очевидно, это доказывает равенство (1.1), утверждающее, что вероятностное распределение величины $X(S)$ является пуассоновским. Доказательство теоремы 1.1 завершено. ■

Исследуем дальнейшие свойства распределения случайного процесса, характеризуемого постулатами (1) — (4). Удобно говорить, что событие $\{X(S) = k\}$ состоит в том, что «в области S существуют в точности k точек».

Покажем теперь, что если процесс $X(S)$ удовлетворяет постулатам (1)—(4), т. е. является пуассоновским процессом на плоскости или в пространстве, то при условии, что в области S положительной площади существует в точности одна точка (т. е. $X(S) = 1$, $A(S) > 0$), местоположение этой точки является случайным с равномерным распределением в S. В самом деле, пусть $S = S_1 \cup S_2$, где S_1 и S_2 не пересекаются. Тогда в силу постулата (3)

$$\mathrm{P}\{X(S_1) = 1 \mid X(S) = 1\} = \frac{\mathrm{P}\{X(S_1) = 1,\ X(S) = 1\}}{\mathrm{P}\{X(S) = 1\}} =$$
$$= \frac{\mathrm{P}\{X(S_1) = 1,\ X(S_2) = 0\}}{\mathrm{P}\{X(S) = 1\}} = \frac{\mathrm{P}\{X(S_1) = 1\}\,\mathrm{P}\{X(S_2) = 0\}}{\mathrm{P}\{X(S) = 1\}} =$$
$$= \frac{\exp[-\lambda A(S_1)]\,\lambda A(S_1)\exp[-\lambda A(S_2)]}{\exp[-\lambda A(S)]\,\lambda A(S)}.$$

Поскольку S_1 и S_2 не пересекаются и $S_1 \cup S_2 = S$, то $A(S) = A(S_1) + A(S_2)$ и $\exp[-\lambda A(S_1)]\exp[-\lambda A(S_2)] = \exp[-\lambda A(S)]$. Таким образом, имеем

$$\mathrm{P}\{X(S_1) = 1 \mid X(S) = 1\} = \frac{A(S_1)}{A(S)},$$

[1]) Цитированное здесь доказательство опирается на факт монотонности $f(x)$. Указанное свойство, очевидно, выполняется: если $x, y > 0$, то можно выбрать S_1, S_2 так, что $x = A(S_1)$, $y = A(S_2)$; отсюда

$$X(S_1) \leqslant X(S_1 \cup S_2);\ f(x) = \mathrm{M}(X(S_1)) \leqslant \mathrm{M}(X(S_1 \cup S_2)) = f(x+y),$$

что и требовалось доказать. — *Прим. ред.*

а это и выражает тот факт, что местоположение точки в S имеет равномерное распределение.

Этот результат можно обобщить следующим образом.

Теорема 1.2. *Если $X(S)$ удовлетворяет постулатам* (1) — (4), *то при условии $X(S) = k$, где $A(S) > 0$, местоположения этих k точек являются независимыми случайными величинами, равномерно распределенными в S.*

Замечание. Утверждение о том, что k точек в S независимы и равномерно распределены, будет означать, что для любых n непересекающихся областей $S_1, S_2, \ldots, S_n$, $\bigcup_{i=1}^{n} S_i = S$, и любых целых чисел $k_1, k_2, \ldots, k_n$, $\sum_{i=1}^{n} k_i = k$, выполняется условие $\mathrm{P}\{k_1$ точек лежит в S_1; k_2 точек лежит в S_2; $\ldots$; k_n точек лежит в $S_n \mid X(S)=k\} = \dfrac{k!}{k_1!\,k_2!\ldots k_n!}\left[\dfrac{A(S_1)}{A(S)}\right]^{k_1}\left[\dfrac{A(S_2)}{A(S)}\right]^{k_2}\cdots\left[\dfrac{A(S_n)}{A(S)}\right]^{k_n}$.

Доказательство. Пусть

$$S = S_1 \cup S_2 \cup \ldots \cup S_n,$$

где $S_1, S_2, \ldots, S_n$ — непересекающиеся области; тогда для любых неотрицательных целых чисел $k_1, k_2, \ldots, k_n$, $k_1 + k_2 + \ldots + k_n = k$,

$$\mathrm{P}\{X(S_1) = k_1;\ X(S_2) = k_2;\ \ldots;\ X(S_n) = k_n \mid X(S) = k\} =$$

$$= \frac{\mathrm{P}\{X(S_1) = k_1;\ X(S_2) = k_2;\ \ldots;\ X(S_n) = k_n\}}{\mathrm{P}\{X(S) = k\}} =$$

$$= \frac{\mathrm{P}\{X(S_1) = k_1\}\,\mathrm{P}\{X(S_2) = k_2\}\ldots\mathrm{P}\{X(S_n) = k_n\}}{\mathrm{P}\{X(S) = k\}} =$$

$$= \frac{e^{-\lambda A(S_1)}\frac{1}{k_1!}[\lambda A(S_1)]^{k_1}e^{-\lambda A(S_2)}\frac{1}{k_2!}[\lambda A(S_2)]^{k_2}\ldots e^{-\lambda A(S_n)}\frac{1}{k_n!}[\lambda A(S_n)]^{k_n}}{e^{-\lambda A(S)}\frac{1}{k!}[\lambda A(S)]^{k}} =$$

$$= \frac{k!}{k_1!\,k_2!\ldots k_n!}\left[\frac{A(S_1)}{A(S)}\right]^{k_1}\left[\frac{A(S_2)}{A(S)}\right]^{k_2}\cdots\left[\frac{A(S_n)}{A(S)}\right]^{k_n},$$

поскольку $A(S_1) + A(S_2) + \ldots + A(S_n) = A(S)$. ■

§ 2. ПРИМЕНЕНИЕ МНОГОМЕРНЫХ ПУАССОНОВСКИХ ПРОЦЕССОВ В АСТРОНОМИИ

Рассмотрим звезды, распределенные в пространстве в соответствии с трехмерным пуассоновским процессом $X(S)$, описанным в § 1. Пусть $\mathbf{x}$ и $\mathbf{y}$ — трехмерные векторы. Предположим, что интенсивность света, создаваемая в точке $\mathbf{x}$ звездой, находящейся в точке $\mathbf{y}$, равна $f(\mathbf{x}, \mathbf{y}, \alpha)$. Здесь α — действительный случайный параметр, зависящий от яркости звезды, находящейся в точке $\mathbf{y}$. Предположим, что параметры α, соответствующие различным звездам, являются независимыми одинаково распределенными случайными величинами с общей плотностью распределения $k(\cdot)$. Предположим также, что общая интенсивность света, создаваемая в точке $\mathbf{x}$ световыми сигналами от различных звезд, является суммой составляющих интенсивностей. Пусть $Y(\mathbf{x}, S)$ — общая интенсивность света, создаваемая в точке $\mathbf{x}$ всеми звездами, локализованными в области S, т. е.

$$Y(\mathbf{x}, S) = \sum_{\mathbf{y}_r \in S} f(\mathbf{x}, \mathbf{y}_r, \alpha_r).$$

Заметим, что данная сумма содержит случайное (но конечное с вероятностью 1) число членов. Мы желаем найти распределение величины $Y(\mathbf{x}, S)$. Задача будет решена, если будет найдено преобразование Лапласа $g(z; \mathbf{x}, S)$ этого распределения, т. е.

$$g(z; \mathbf{x}, S) = \mathrm{M}(e^{-zY(\mathbf{x}, S)}) = \int_0^\infty e^{-z\xi} h(\xi; \mathbf{x}, S)\, d\xi, \qquad \operatorname{Re} z > 0,$$

где $h(\cdot; \mathbf{x}, S)$ — плотность распределения величины $Y(\mathbf{x}, S)$.

Конечно, в принципе, зная преобразование Лапласа, можно найти стандартным образом моменты величины $Y(\mathbf{x}, S)$, и вообще по формуле обращения можно определить функцию h через g. Вычисления по этой формуле в рассматриваемом случае довольно громоздкие, и поэтому мы не будем здесь их проводить.

По формуле полной вероятности имеем

$$g(z; \mathbf{x}, S) = \mathrm{M}(e^{-zY(\mathbf{x}, S)}) = \sum_{k=0}^\infty \mathrm{M}[e^{-zY(\mathbf{x}, S)} \mid X(S) = k]\, \mathrm{P}\{X(S) = k\}.$$

Но из теоремы 1.2 известно, что при условии $X(S) = k$ эти k точек распределены равномерно как k независимых случайных величин в области S. Следовательно,

$$\mathrm{M}[e^{-zY(\mathbf{x}, S)} \mid X(S) = k] = (\mathrm{M}[e^{-zY(\mathbf{x}, S)} \mid X(S) = 1])^k.$$

Чтобы найти $M[e^{-zY(\mathbf{x}, S)} \mid X(S) = 1]$, заметим, что $Y(\mathbf{x}, S) = f(\mathbf{x}, \mathbf{y}, \alpha)$ при условии $X(S) = 1$, где $\mathbf{y}$ — местоположение единственной звезды в S, а α — соответствующий случайный параметр, отражающий ее яркость. Далее, поскольку положение этой звезды распределено в S равномерно, имеем

$$\mathrm{M}[e^{-zY(\mathbf{x}, S)} \mid X(S) = 1] = \frac{1}{A(S)} \int_S \left[\int_{-\infty}^{+\infty} e^{-zf(\mathbf{x}, \mathbf{y}, \alpha)} k(\alpha)\, d\alpha \right] d\mathbf{y},$$

где интеграл по переменной $\mathbf{y}$ понимается как тройной интеграл по области S. С помощью выведенных выше соотношений очевидным образом получаем

$$
\begin{aligned}
g(z;\ \mathbf{x},\ S) &= \\
&= \sum_{k=0}^{\infty} \left\{ \frac{1}{A(S)} \int\limits_S \left[\int\limits_{-\infty}^{\infty} e^{-zf(\mathbf{x},\ \mathbf{y},\ \alpha)} k(\alpha)\, d\alpha \right] d\mathbf{y} \right\}^k \exp[-\lambda A(S)] \frac{[\lambda A(S)]^k}{k!} = \\
&= e^{-\lambda A(S)} \exp\left\{ \lambda \int\limits_S \left[\int\limits_{-\infty}^{\infty} e^{-zf(\mathbf{x},\ \mathbf{y},\ \alpha)} k(\alpha)\, d\alpha \right] d\mathbf{y} \right\} = \\
&= \exp\left\{ \lambda \int\limits_S \left[\int\limits_{-\infty}^{\infty} e^{-zf(\mathbf{x},\ \mathbf{y},\ \alpha)} k(\alpha)\, d\alpha - 1 \right] d\mathbf{y} \right\},
\end{aligned}
$$

так как $\int\limits_S d\mathbf{y} = A(S)$. Мы определили $g(z; \mathbf{x}, S)$ через $f(\mathbf{x}, \mathbf{y}, \alpha)$, $k(\alpha)$ и S, которые можно считать известными или получаемыми на основе других данных.

§ 3. ИММИГРАЦИЯ И РОСТ ПОПУЛЯЦИЙ

Модель, изучаемая в данном параграфе, описывает однотипную популяцию, развивающуюся из исходной популяции, а процесс роста этой популяции соответствует марковскому ветвящемуся процессу. Кроме того, в дополнение к самопроизвольному росту популяции имеется приток иммигрантов того же типа, которые в дальнейшем развиваются, как и остальные члены популяции. Процесс поступления иммигрантов в общем случае является случайным. Для определенности опишем процесс, являющийся моделью роста популяций бактерий.

Рассмотрим колонию бактерий, состоящую из n_0 индивидуумов. Предположим, что каждая бактерия независимо от остальных порождает поколение потомков, которые в свою очередь производят следующее поколение потомков и т. д. Рост популяции, развивающейся из одной бактерии, описывается марковским ветвящимся процессом с непрерывным временем. Пусть $F(s, t)$ — вероятностная производящая функция размера в момент t популяции, развившейся из одной бактерии. Очевидно, размер популяции, развившейся из колонии, имевшей в момент 0 размер n_0, является случайной величиной с производящей функцией $[F(s,\ t)]^{n_0}$. Предположим далее, что иммиграция новых бактерий происходит в моменты t_j, $j = 1, 2, \ldots, N$. Каждый иммигрант порождает потомство таким же образом, как и исходные n_0 бактерий, независимо от них и от других бактерий. Размер в момент t популяции, порожденной иммигрантом, прибывшим в момент t_j, имеет производящую функцию $F(s, t-t_j)$. Общий размер популяции в момент t имеет производящую функцию

$$[F(s, t)]^{n_0} \prod_{j=1}^{N} F(s, t-t_j),$$

поскольку каждая бактерия развивается независимо от других. Предположим теперь, что иммиграция происходит не в фиксированные моменты t_j, а в моменты, образующие пуассоновский поток с параметром r. Наша задача — выразить производящую функцию общего размера популяции через $F(s, t)$ и r.

Моменты иммиграции t_j являются случайными величинами, и их число $N(t)$ за отрезок времени $[0, t]$ — также случайная величина с пуассоновским распределением, имеющим параметр rt. Пусть $Y_j(t, t_j)$ — размер популяции к моменту t, развившейся из одного иммигранта, поступившего в колонию в момент t_j, $j = 1, 2, \ldots, N(t)$; тогда

$$Y(t) = \sum_{j=1}^{N(t)} Y_j(t, t_j)$$

есть число всех бактерий в момент t, «предками» которых были иммигранты, поступившие за отрезок времени $[0, t]$.

Производящая функция величины $Y(t)$ может быть получена стандартным образом с помощью наложения условия на $N(t)$ и формулы полной вероятности:

$$\mathrm{M}[s^{Y(t)}] = \sum_{k=0}^{\infty} \mathrm{M}[s^{Y(t)} \mid N(t) = k]\, \mathrm{P}\{N(t) = k\}. \tag{3.1}$$

Из результатов § 2 гл. 7 известно, что совместное распределение на $[0, t]$ моментов поступления t_j, $j = 1, 2, \ldots, N(t)$, при известном их числе $N(t) = k$ совпадает с распределением порядковых статистик k независимых равномерно распределенных на $[0, t]$ случайных величин. Поэтому

$$\mathrm{M}[s^{Y(t)} \mid N(t) = k] = \frac{k!}{t^k} \int_0^t dt_1 \int_{t_1}^t dt_2 \ldots \int_{t_{k-1}}^t dt_k \mathrm{M}\left[s^{\sum_{j=1}^{k} Y_j(t, t_j)}\right] =$$

$$= \frac{1}{t^k} \int_0^t dt_1 \int_0^t dt_2 \ldots \int_0^t dt_k \mathrm{M}\left[s^{\sum_{j=1}^{k} Y_j(t, t_j)}\right],$$

поскольку интегрируемая функция симметрична относительно $t_1, t_2, \ldots, t_k$. Далее, поскольку различные иммигранты создают независимо развивающиеся популяции, имеем

$$\mathrm{M}\left[s^{\sum_{j=1}^{k} Y_j(t,\, t_j)}\right] = \prod_{j=1}^{k} \mathrm{M}\left[s^{Y_j(t,\, t_j)}\right] = \prod_{j=1}^{k} F(s,\ t - t_j).$$

Следовательно,

$$\mathrm{M}\left[s^{Y(t)} \mid N(t) = k\right] = \frac{1}{t^k} \int_0^t dt_1 \int_0^t dt_2 \ldots \int_0^t dt_k \prod_{j=1}^{k} F(s,\ t - t_j) =$$

$$= \frac{1}{t^k} \prod_{j=1}^{k} \int_0^t F(s,\ t - t_j)\, dt_j = \left(\frac{1}{t} \int_0^t F(s,\ t - \tau)\, d\tau\right)^k.$$

Подставляя эту формулу в (3.1) и учитывая, что $N(t)$ является пуассоновским процессом, получим

$$\mathrm{M}\left[s^{Y(t)}\right] = \sum_{k=0}^{\infty} \left\{\frac{1}{t} \int_0^t F(s,\ t - \tau)\, d\tau\right\}^k e^{-rt} \frac{(rt)^k}{k!} =$$

$$= \exp\left\{r \int_0^t [F(s,\ t - \tau) - 1]\, d\tau\right\}. \qquad (3.2)$$

Следовательно, производящая функция общего размера популяции в момент t равна

$$G(s,\ t) = [F(s,\ t)]^{n_0} \exp\left\{r \int_0^t [F(s,\ t - \tau) - 1]\, d\tau\right\}. \qquad (3.3)$$

П р и м е р. В качестве примера предположим, что каждая отдельная бактерия развивается в соответствии с процессом Юла $X(t)$ с параметром $\beta > 0$ (см. § 2 гл. 7); и размер популяции в момент t, развившейся из одной бактерии, имевшейся в момент $t = 0$, имеет распределение

$$P_k(t) = \mathrm{P}\{X(t) = k \mid X(0) = 1\},$$

где

$$P_k(t) = e^{-\beta t}(1 - e^{-\beta t})^{k-1}, \qquad k = 1,\ 2,\ \ldots$$

Производящая функция для процесса Юла равна

$$F(s,\ t) = e^{-\beta t} \sum_{k=1}^{\infty} (1 - e^{-\beta t})^{k-1} s^k = \frac{s e^{-\beta t}}{1 - (1 - e^{-\beta t})\, s}.$$

В соответствии с (3.2) производящая функция размера популяции, развившейся из иммигрантов, имеет вид

$$M[s^{Y(t)}] = \exp\left\{ r \int_0^t \frac{se^{-\beta(t-\tau)}}{1 - s + se^{-\beta(t-\tau)}}\, d\tau - rt \right\} =$$

$$= \exp\left\{ -\frac{r}{\beta} \ln\left[1 - s + se^{-\beta t}\right] - rt \right\} =$$

$$= e^{-rt}\left(1 - s + se^{-\beta t}\right)^{-\frac{r}{\beta}} = e^{-rt}\left[1 - (1 - e^{-\beta t})\, s\right]^{-\frac{r}{\beta}}.$$

Если учесть исходную популяцию бактерий, то в соответствии с (3.3) производящая функция общего размера популяции в момент t имеет вид

$$G(s,\ t) = \exp\left[-(r + \beta n_0)\, t\right]\left[1 - (1 - e^{-\beta t})\, s\right]^{-\left(n_0 + \frac{r}{\beta}\right)}.$$

Средний размер популяции в момент t равен $\frac{\partial}{\partial s} G(s,\ t)\,|_{s=1}$, или

$$\left(n_0 + \frac{r}{\beta}\right)(e^{\beta t} - 1).$$

Моменты более высоких порядков можно найти, последовательно дифференцируя производящую функцию в точке $s = 1$.

§ 4. ВЕРОЯТНОСТНЫЕ МОДЕЛИ МУТАЦИИ И РОСТА

Часто в микробиологических популяциях, первоначально однородных, один или более индивидуумов изменяются в мутантную форму и затем эта форма размножается. Мутация может соответствовать, например, иммунитету от вируса, который сохраняется в потомстве, или в общем случае некоторому свойству, отличающему мутантную форму от исходной (от индивидуумов «дикого типа»). Исследуем теперь модель, описывающую случайные флуктуации мутантного роста. Предположим, что материнская, или родительская колония состоит в момент $t = 0$ из N индивидуумов и растет детерминированным образом так, что ее размер в момент t равен Ne^t. Далее, предположим, что каждый индивидуум дикого типа с вероятностью $\rho h + o(h)$ изменяется в мутантную форму за отрезок времени $[t,\ t + h]$. Поскольку родительская популяция в момент t имеет размер Ne^t, а индивидуумы развиваются независимо, вероятность образования некоторого мутанта на отрезке $[t, t + h]$ равна

$$\rho N e^t h + o(h).$$

Кроме того, мы постулируем, что вероятность двух или более мутаций на отрезке $[t, t + h]$ есть $o(h)$. Из приведенной формулировки следует, что число индивидуумов мутантного вида, как функция времени, описывается неоднородным пуассоновским процессом с функцией интенсивности $r(t) = \rho N e^t$ (см. задачу 12 гл. 7).

Используя решение задач 12 и 13 гл. 7, можно показать, что вероятностная производящая функция числа событий, происшедших за отрезок времени $[0, t]$, для неоднородного пуассоновского процесса с параметром $r(t)$ равна

$$\varphi(t, s) = e^{-m(t)(1-s)},$$

где

$$m(t) = \int_0^t r(\tau)\, d\tau = \rho N (e^t - 1).$$

Таким образом, в нашем частном случае выполняется равенство

$$\varphi(t, s) = \exp\{\rho N (e^t - 1)(s - 1)\}. \tag{4.1}$$

Предположим теперь, что каждый мутант порождает свой собственный процесс роста, и пусть $F(s, t)$ — вероятностная производящая функция числа потомков одного мутанта спустя время t после его возникновения. В данной модели мы предположим, что в мутантной популяции не происходит гибели, т. е. $F(0, t) \equiv 0$.

Пусть $H(s, t; N)$ — вероятностная производящая функция числа мутантов в момент t при условии, что родительская колония состояла из N индивидуумов в момент $t = 0$ и в этот момент в ней не было ни одного мутантного индивидуума. Наша цель — найти $H(s, t; N)$ через $F(s, t)$ и параметры ρ и N, которые в рассматриваемой задаче считаются известными. Для этого введем вероятности

$P_k(t) = \mathrm{P}\{$в момент t существует ровно k потомков одного мутантного индивидуума, возникшего в момент $t = 0\}$,

$h_k(t; N) = \mathrm{P}\{$в момент t существует ровно k мутантов при условии, что родительская популяция в момент $t = 0$ имела размер N и не содержала ни одного мутанта$\}$.

В таком случае

$$F(s, t) = \sum_{k=1}^{\infty} P_k(t)\, s^k, \tag{4.2}$$

поскольку по предположению $P_0(t) \equiv 0$ и

$$H(s, t; N) = \sum_{k=0}^{\infty} h_k(t; N)\, s^k. \tag{4.3}$$

Из (4.1) следует, что

$$\begin{aligned} h_0(t; N) &= \mathrm{P}\{\text{первая мутация происходит позже } t\} = \\ &= \varphi(t, 0) = \exp[-\rho N (e^t - 1)] = 1 - K(t; N), \end{aligned} \tag{4.4}$$

где $K(t; N)$ — функция распределения момента рождения первого мутанта при исходной популяции, состоящей из N индивидуумов. Ее плотность равна

$$-\frac{dh_0(t,\ N)}{dt} = \rho N e^t \exp[-\rho N(e^t - 1)].$$

Событие, состоящее в том, что в момент t имеется ровно $k+1$ мутантных индивидуумов ($k = 0, 1, 2, \ldots$), произойдет, если первая мутация случится в момент τ ($0 \leqslant \tau \leqslant t$) и появившийся мутант и родительская популяция (в данном случае размера Ne^τ) вместе породят k мутантов за оставшееся время $t - \tau$. Вероятность того, что до момента τ не произойдет ни одной мутации, равна $\exp\{-\rho N(e^\tau - 1)\}$. В интервале $(\tau,\ \tau + d\tau)$ мутация произойдет с вероятностью $\rho N e^\tau d\tau + o(d\tau)$. Наконец, вероятность того, что мутантный организм и родительская колония размера Ne^τ породят в точности k мутантов за время $t - \tau$, равна

$$\sum_{l=0}^{k} P_{k-l}(t-\tau)\, h_l(t-\tau;\ Ne^\tau)$$

(напомним, что $P_0(t) \equiv 0$). Но момент τ может быть любым между 0 и t, поэтому в силу формулы полной вероятности получаем

$$h_k(t;\ N) = \int_0^t \exp[-\rho N(e^\tau - 1)]\, \rho N e^\tau\, d\tau \times \\ \times \left(\sum_{l=0}^{k} P_{k-l}(t-\tau)\, h_l(t-\tau;\ Ne^\tau) \right), \qquad k = 1,\ 2,\ \ldots.$$

Теперь можно записать соответствующую производящую функцию. Получим формулу

$$H(s,\ t;\ N) = \exp[-\rho N(e^t - 1)] + \sum_{k=1}^{\infty} s^k \int_0^t \exp[-\rho N(e^\tau - 1)] \times \\ \times \rho N e^\tau \left(\sum_{l=0}^{k} P_{k-l}(t-\tau)\, h_l(t-\tau;\ Ne^\tau) \right) d\tau = \\ = \exp[-\rho N(e^t - 1)] + \int_0^t \exp[-\rho N(e^\tau - 1)] \times \\ \times \rho N e^\tau \left(\sum_{l=0}^{\infty} h_l(t-\tau;\ Ne^\tau)\, s^l \sum_{k=l}^{\infty} P_{k-l}(t-\tau)\, s^{k-l} \right) d\tau,$$

где мы использовали гипотезу $P_0(t) \equiv 0$. В силу (4.2) и (4.3) это соотношение можно переписать в более простом виде

$$H(s, t; N) = \exp[-\rho N(e^t - 1)] +$$
$$+ \rho N \int_0^t e^y \exp[-\rho N(e^y - 1)] F(s, t-y) H(s, t-y; Ne^y)\, dy. \qquad (4.5)$$

Это интегральное уравнение относительно $H(s, t; N)$, довольно сложное. Однако его можно решить, используя следующий прием. Пусть $\xi(t; N)$ — число мутантов в момент t при условии, что в момент $t = 0$ родительская популяция состоит из N индивидуумов. Поскольку мутации происходят в соответствии с неоднородным пуассоновским процессом и индивидуумы развиваются независимо, отсюда следует, что величина $\xi(t; N)$ удовлетворяет функциональному уравнению

$$\xi(t; N_1) + \xi(t; N_2) = \xi(t; N_1 + N_2). \qquad (4.6)$$

В силу независимости $\xi(t; N_1)$ и $\xi(t; N_2)$ и определения

$$H(s, t; N) = \mathrm{M}(s^{\xi(t; N)})$$

заключаем, что

$$H(s, t; N_1) H(s, t; N_2) = H(s, t; N_1 + N_2)$$

для всех неотрицательных целых чисел $N_1, N_2, \ldots$. Отсюда, очевидно, следует, что

$$H(s, t; N) = [H(s, t; 1)]^N,$$

т. е.

$$H(s, t; N) = e^{NL(s, t)}, \qquad (4.7)$$

где $L(s, t) = \ln H(s, t; 1)$. Нам осталось найти функцию $L(s, t)$. Для этого подставим формулу (4.7) в (4.5) и разделим обе части на N:

$$\frac{1}{N}\{\exp[NL(s, t)] - \exp[-\rho N(e^t - 1)]\} =$$
$$= \rho \int_0^t e^y \exp[-\rho N(e^y - 1)] F(s, t-y) \exp[Ne^y L(s, t-y)]\, dy. \qquad (4.8)$$

Равенство (4.8) справедливо лишь при неотрицательных целых N. Однако мы будем оперировать с ним, как если бы оно было справедливым при всех $N > 0$. (Это можно сделать, соответствующим образом изменяя ρ. Мы не будем входить в детали данного вопроса, который является достаточно тонким.)

Положим теперь $N \to 0$ в (4.8), тогда правая часть устремится к пределу

$$\rho \int_0^t e^y F(s,\ t-y)\,dy.$$

В левой части имеем

$$\lim_{N\to 0} \frac{1}{N}\{\exp[NL(s,\ t)] - \exp[-\rho N(e^t-1)]\} =$$
$$= \lim_{u\to 0}\left\{\frac{1}{u}[\exp(uL(s,\ t))-1] + \frac{1}{u}[1-\exp(-\rho u(e^t-1))]\right\} =$$
$$= \frac{d}{du}\exp[uL(s,\ t)]\Big|_{u=0} - \frac{d}{du}\exp[-\rho u(e^t-1)]\Big|_{u=0} =$$
$$= L(s,\ t) + \rho(e^t-1).$$

Следовательно, формально получается

$$L(s,\ t) = -\rho(e^t-1) + \rho\int_0^t e^\tau F(s,\ t-\tau)\,d\tau \tag{4.9}$$

и функция $H(s,\ t;\ N)$ определяется равенством (4.7).

Чтобы найти среднее число мутантов в момент t, положим

$$\nu(t) = \frac{\partial F(s,\ t)}{\partial s}\Big|_{s=1}; \tag{4.10}$$

$\nu(t)$ равно среднему числу потомков одного мутанта спустя время t после его рождения. Из (4.7) и (4.10) имеем

$$\mathrm{M}(\xi(t;\ N)) = \frac{\partial H(s,\ t;\ N)}{\partial s}\Big|_{s=1} = e^{NL(1,\ t)}N\,\frac{\partial L(s,\ t)}{\partial s}\Big|_{s=1},$$

но

$$L(1,\ t) = -\rho(e^t-1) + \rho\int_0^t e^\tau\,d\tau = 0,$$

поскольку

$$F(1,\ t-\tau) = \sum_{k=0}^{\infty} P_k(t-\tau) = 1.$$

Из (4.9) и (4.10)

$$\frac{\partial L(s,\ t)}{\partial s}\Big|_{s=1} = \rho\int_0^t e^\tau \nu(t-\tau)\,d\tau.$$

Следовательно,

$$M(\xi(t;\ N)) = \rho N \int_0^t e^{\tau} \nu(t-\tau)\, d\tau. \tag{4.11}$$

Если для не слишком больших t можно приближенно записать

$$\nu(t) \sim n_0 e^t \qquad (n_0 = \text{const}),$$

то из (4.11) следует, что $M(\xi(t; N)) \sim \rho N n_0 t e^t$.

§ 5. ЭКСПОНЕНЦИАЛЬНЫЙ РОСТ ОДНОМЕРНОЙ ПОПУЛЯЦИИ

Другой пример экспоненциального роста популяций дает описываемая ниже модель. Ядерные частицы расположены на бесконечной прямой. При их расщеплении «потомки» рассеиваются в соответствии с некоторым вероятностным законом. Более определенно, предположим, что «потомок» частицы, расположенной в точке x, будет находиться в точке $x+y$ с плотностью вероятности $f(y)$,

$$f(y) \geqslant 0, \quad -\infty < y < \infty, \quad \int_{-\infty}^{\infty} f(y)\, dy = 1.$$

Заметим, что $f(y)$ зависит только от расстояния y между «родительской» частицей и «потомком» и не зависит от местонахождения «родительской» частицы. Для простоты предположим сначала, что каждая частица делится в точности на две новые. Если в момент 0 имеется одна частица в точке $x=0$, то ее «потомство» назовем первым поколением; «потомство» первого поколения образует второе поколение и т. д.

Введем случайную величину $Z_n(x;\ 0)$, равную числу частиц n-го поколения, расположенных на полуинтервале $(-\infty,\ x]$, если нулевое поколение состояло из одной частицы, расположенной в точке $x=0$. Положим

$$p_k^{(n)}(x) = P\{Z_n(x;\ 0) = k\}.$$

Предположим, что мы поместили исходную частицу в точку u. Пусть $Z_n(x; u)$ — число частиц n-го поколения, расположенных в $(-\infty, x]$, при условии, что нулевое поколение состояло из одной частицы, расположенной в точке u. В силу пространственной однородности закона распределения разброса потомков интуитивно ясно, что

$$P\{Z_n(x;\ u) = k\} = P\{Z_n(x-u;\ 0) = k\} = p_k^{(n)}(x-u). \tag{5.1}$$

Формальное доказательство этого читателю предлагается провести самостоятельно.

Введем производящую функцию

$$g_n(s;\ x) = \sum_{k=0}^{\infty} p_k^{(n)}(x)\, s^k \tag{5.2}$$

и среднее

$$e_n(x) = \mathrm{M}(Z_n(x;\ 0)) = \sum_{k=1}^{\infty} k p_k^{(n)}(x) = g_n'(1;\ x), \tag{5.3}$$

где штрихом обозначена производная по s.

Событие, заключающееся в том, что в $(n+1)$-м поколении на полуинтервале $(-\infty, x]$ будет в точности k частиц, произойдет, если два «потомка» исходной частицы, расположенной в точке $x = 0$, будут локализованы в интервалах $(u, u + du)$ и $(v, v + dv)$ соответственно, где $-\infty < u,\ v < \infty$, и каждая из новых частиц будет иметь через n поколений в $(-\infty, x]$ такое число «потомков», что в сумме они составят k. Вероятность того, что две частицы первого поколения будут расположены в интервалах $(u,\ u + du)$ и $(v, v + dv)$ соответственно, равна

$$f(u) f(v)\, du\, dv.$$

Вероятность того, что эти две частицы через n поколений дадут в сумме k «потомков», находящихся в полуинтервале $(-\infty, x]$, равна

$$\sum_{l=0}^{k} p_l^{(n)}(x-u)\, p_{k-l}^{(n)}(x-v)$$

(см. (5.1)). Далее, величины u и v могут принимать любые значения на действительной прямой независимо друг от друга. Следовательно,

$$p_k^{(n+1)}(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} du\, dv f(u) f(v) \sum_{l=0}^{k} p_l^{(n)}(x-u)\, p_{k-l}^{(n)}(x-v),$$
$$k = 0,\ 1,\ 2,\ \ldots.$$

Переходя к производящей функции, находим:

$$g_{n+1}(s;\ x) = \sum_{k=0}^{\infty} s^k \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u) f(v) \sum_{l=0}^{k} p_l^{(n)}(x-u)\, p_{k-l}^{(n)}(x-v)\, du\, dv =$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u) f(v) \sum_{l=0}^{\infty} p_l^{(n)}(x-u)\, s^l \sum_{k=l}^{\infty} p_{k-l}^{(n)}(x-v)\, s^{k-l}\, du\, dv =$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_n(s;\ x-u)\, g_n(s;\ x-v) f(u) f(v)\, du\, dv =$$
$$= \int_{-\infty}^{\infty} g_n(s;\ x-u) f(u)\, du \int_{-\infty}^{\infty} g_n(s;\ x-v) f(v)\, dv,$$

т. е.

$$g_{n+1}(s;\ x)=\left\{\int_{-\infty}^{\infty} g_n(s;\ x-u)f(u)\,du\right\}^2. \tag{5.4}$$

Если в качестве обобщения предположить, что каждая частица порождает r новых частиц, где $r>0$ — целое число, то вместо (5.4) получим формулу

$$g_{n+1}(s;\ x)=\left\{\int_{-\infty}^{\infty} g_n(s;\ x-u)f(u)\,du\right\}^r.$$

Если в качестве дальнейшего обобщения предположить, что каждая частица порождает r новых частиц с вероятностью a_r, то точно таким же образом получим формулу

$$g_{n+1}(s;\ x)=A\left(\int_{-\infty}^{\infty} g_n(s;\ x-u)f(u)\,du\right),$$

где

$$A(z)=\sum_{r=0}^{\infty} a_r z^r$$

есть производящая функция числа новых частиц, производимых при каждом расщеплении.

Математическое ожидание $\mathrm{M}_{n+1}(x)$ числа частиц в $(n+1)$-м поколении можно найти обычным путем:

$$\mathrm{M}_{n+1}(x)=g'_{n+1}(1;\ x)=$$
$$=A'\left(\int_{-\infty}^{\infty} g_n(1;\ x-u)f(u)\,du\right)\int_{-\infty}^{\infty} g'_n(1;\ x-u)f(u)\,du. \tag{5.5}$$

Из (5.1) следует, что

$$g_n(1;\ x-u)=\sum_{k=0}^{\infty} p_k^{(n)}(x)=1,$$

и

$$A'(1)=\sum_{r=0}^{\infty} ra_r=m$$

есть среднее число новых частиц, производимых при каждом расщеплении. Формула (5.5) приобретает вид

$$\mathrm{M}_{n+1}(x)=m\int_{-\infty}^{\infty}\mathrm{M}_n(x-u)f(u)\,du. \tag{5.6}$$

Это рекуррентное соотношение можно легко решить. Будем рассматривать исходную частицу, находящуюся в точке $x = 0$, как нулевое поколение; тогда, очевидно,

$$M_0(x) = \begin{cases} 0, & \text{если } x < 0, \\ 1, & \text{если } x \geqslant 0. \end{cases}$$

Из (5.6) получаем

$$M_1(x) = m \int_{-\infty}^{x} f(u)\,du = mF(x),$$

где $F(x)$ есть функция распределения, соответствующая плотности $f(x)$. Далее,

$$M_2(x) = m^2 \int_{-\infty}^{\infty} F(x-u)\,f(u)\,du = m^2 F^{(2)}(x),$$

где $F^{(2)}(x)$ — двукратная свертка функции $F(x)$. По индукции с очевидностью получаем

$$M_n(x) = m^n F^{(n)}(x), \tag{5.7}$$

где $F^{(n)}(x)$ есть n-кратная свертка функции $F(x)$.

Если плотность $f(y)$ имеет дисперсию σ^2 и среднее μ, то в силу центральной предельной теоремы

$$F^{(n)}(n\mu + \xi\sigma\sqrt{n}) \to \Phi(\xi), \qquad n \to \infty,$$

где $\Phi(\xi)$ — стандартная нормальная функция распределения, т. е.

$$\Phi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} \exp\left(-\frac{\eta^2}{2}\right) d\eta.$$

Отсюда получаем асимптотическую формулу для $M_n(x)$, а именно

$$\frac{M_n(n\mu + \xi\sigma\sqrt{n})}{m^n} \to \Phi(\xi) \quad \text{при } n \to \infty,$$

которая представляет некоторый самостоятельный интерес.

§ 6. ВЕРОЯТНОСТНАЯ МОДЕЛЬ РОСТА ПОПУЛЯЦИИ В ПРОСТРАНСТВЕ И ВРЕМЕНИ

Предположим, что некоторые растения распределены в пространстве в соответствии с двумерным распределением Пуассона, имеющим интенсивность λ. (Мы рассматриваем модель распределения растений в двумерном пространстве, но все рассуждения без изменений могут быть перенесены на трехмерный случай.) Предположим, что каждое родительское растение, чье местоположение описывается двумерным вектором $\mathbf{r}_0$, порождает независимо от других растений случайное число потомков с производящей функцией $H(s)$, т. е.

$$H(s) = \sum_{k=0}^{\infty} h_k s^k,$$

где h_k — вероятность того, что родительское растение породит k потомков. Предположим также, что потомство одного родителя, расположенного в точке $\mathbf{r}_0$, распределяется независимым образом вокруг точки $\mathbf{r}_0$ в соответствии с двумерной плотностью $f(\mathbf{r}-\mathbf{r}_0)$, зависящей лишь от вектора $\mathbf{r}-\mathbf{r}_0$, например можно взять двумерную нормальную плотность

$$f(\mathbf{r}-\mathbf{r}_0) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 \exp\left\{-\frac{1}{2\sigma^2}\left[(x-x_0)^2+(y-y_0)^2\right]\right\}.$$

Таким образом, вероятность того, что данный потомок родителя, находящегося в точке $\mathbf{r}_0$, будет находиться в области R, равна

$$p = \int_R f(\mathbf{r}-\mathbf{r}_0)\,d\mathbf{r}. \tag{6.1}$$

Если родительское растение, находящееся в точке $\mathbf{r}_0$, имеет в точности n потомков, то число потомков этого родителя, находящихся в области R, имеет биномиальное распределение с параметрами p и n, где p задается равенством (6.1). Общее число потомков одного родителя является случайной величиной с вероятностной производящей функцией $H(s)$. С помощью обычного метода использования формулы полной вероятности можно показать, что производящая функция числа потомков одного родительского растения, расположенного в точке $\mathbf{r}_0$, которые находятся в области R, равна $H[1+p(s-1)]$, где p задается равенством (6.1). Наша цель — найти производящую функцию числа потомков в области R, порожденных всеми родителями из области S. Для этой цели введем следующие обозначения.

Пусть $X(S)$ — число родительских растений в области S;
$Y(\mathbf{r}_0, R)$ — число потомков одного родительского растения, находящегося в точке $\mathbf{r}_0$, которые расположены в области R;
$Y(S, R)$ — суммарное число потомков в области R, которые порождены всеми родителями из области S.

Тогда в этих обозначениях имеем

$$Y(S, R) = \sum_{\mathbf{r}_0 \in S} Y(\mathbf{r}_0, R).$$

Если S имеет конечную площадь, то сумма справа с вероятностью 1 содержит конечное число членов, поскольку число родителей в S

имеет пуассоновское распределение с параметром $\lambda A(S)$ ($A(S)$ — площадь S). Далее, величина $X(S)$ описывается двумерным пуассоновским процессом с интенсивностью λ. Производящую функцию величины $Y(S, R)$ можно найти с помощью формулы полной вероятности, налагая условие на значения $X(S)$. Таким образом, получаем

$$g(s) = \mathrm{M}(s^{Y(S,R)}) = \sum_{k=0}^{\infty} \mathrm{M}[s^{Y(S,R)} \mid X(S) = k]\, \mathrm{P}\{X(S) = k\}. \qquad (6.2)$$

Поскольку родительские растения развиваются независимо, то

$$\mathrm{M}[s^{Y(S,R)} \mid X(S) = k] = \{\mathrm{M}[s^{Y(S,R)} \mid X(S) = 1]\}^k. \qquad (6.3)$$

Кроме того, из теории пространственных пуассоновских процессов известно, что при условии $X(S) = 1$ выполняется равенство

$$Y(S, R) = Y(\mathbf{r}_0, R),$$

где вектор $\mathbf{r}_0$ распределен равномерно в S. В таком случае имеем

$$\mathrm{M}[s^{Y(S,R)} \mid X(S) = 1] = \mathrm{M}[s^{Y(\mathbf{r}_0,R)} \mid \mathbf{r}_0 \text{ равномерно распределен в } S] = \frac{1}{A(S)} \int_S H[1 + p(s-1)]\, d\mathbf{r}_0, \qquad (6.4)$$

где p определяется равенством (6.1). Последнее равенство выполняется, так как было доказано, что $H[1 + p(s - 1)]$ является производящей функцией числа потомков в R, порожденных одним родителем, расположенным в точке $\mathbf{r}_0$. Здесь, однако, вектор $\mathbf{r}_0$ равномерно распределен в S, и поэтому мы получили равенство (6.4). Из (6.2), (6.3) и (6.4) следует

$$g(s) = \sum_{k=0}^{\infty} \left\{ \frac{1}{A(S)} \int_S H[1 + p(s-1)]\, d\mathbf{r}_0 \right\}^k e^{-\lambda A(S)} \frac{[\lambda A(S)]^k}{k!},$$

поскольку $X(S)$ имеет пуассоновское распределение с параметром $\lambda A(S)$. Последнюю формулу можно переписать в более простом виде:

$$g(s) = \exp\left\{ \lambda \int_S (H[1 + p(s-1)] - 1)\, d\mathbf{r}_0 \right\}, \qquad (6.5)$$

где

$$p = \int_R f(\mathbf{r} - \mathbf{r}_0)\, d\mathbf{r}.$$

В выражении (6.5) часто в качестве S берут все двумерное пространство.

Формула (6.5) остается в силе, если растения распределены в пространстве в соответствии с трехмерным пуассоновским распределением, а $\mathbf{r}$, $\mathbf{r}_0$, R, S означают трехмерные векторы и области

13 Зак. 939

соответственно. В любом (дву- или трехмерном) случае, если область R достаточно мала, вероятность p приближенно равна

$$f(\mathbf{r}-\mathbf{r}_0)\,A(R).$$

В качестве $\mathbf{r}$ можно взять любой вектор из области R, $\mathbf{r}\in R$. Тогда равенство (6.5) перепишется в виде

$$g(s)\approx\exp\left\{\lambda\int\limits_S (H[1+f(\mathbf{r}-\mathbf{r}_0)\,A(R)(s-1)]-1)\,d\mathbf{r}_0\right\}.$$

Если S — все пространство (дву- или трехмерное), то можно записать

$$g(s)\approx\exp\left\{\lambda\int (H[1+f(\mathbf{u})\,A(R)(s-1)]-1)\,d\mathbf{u}\right\},\tag{6.6}$$

где интеграл (двойной или тройной) берется по всему пространству.

Для примера рассмотрим в качестве f нормальное распределение на плоскости:

$$f(\mathbf{u})=\frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2\sigma^2}(x^2+y^2)\right],\quad\text{где }\mathbf{u}=(x,\ y).$$

Пусть вероятностное распределение числа потомков одного родителя равно

$$h_k=\mu^k(1-\mu),\qquad k=0,\ 1,\ 2,\ \ldots,$$

где μ — постоянная, $0<\mu<1$. Тогда $H(s)=\sum\limits_{k=0}^{\infty}\mu^k s^k(1-\mu)=$ $=(1-\mu)/(1-\mu s)$. Подставляя в (6.6) и упрощая, получим при малых $A(R)$ приближенное равенство

$$g(s)\approx\exp\left\{\lambda\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\frac{\mu A(R)(s-1)\exp\left[-(2\sigma^2)^{-1}(x^2+y^2)\right]dx\,dy}{2\pi\sigma^2(1-\mu)-\mu A(R)(s-1)\exp\left[-(2\sigma^2)^{-1}(x^2+y^2)\right]}\right\}.$$

После перехода к полярным координатам r и θ выражение для $g(s)$ примет вид

$$g(s)\approx\exp\left\{2\pi\lambda\int\limits_0^{\infty}\frac{\mu A(R)(s-1)\,r\exp\left(-(2\sigma^2)^{-1}r^2\right)dr}{2\pi\sigma^2(1-\mu)-\mu A(R)(s-1)\exp\left[-(2\sigma^2)^{-1}r^2\right]}\right\}=$$

$$=\exp\left\{2\pi\lambda\int\limits_0^{1}\frac{\mu A(R)(s-1)\,\sigma^2\,dz}{2\pi\sigma^2(1-\mu)-\mu A(R)(s-1)\,z}\right\}\quad\left(\text{где }z=\exp\left(-\frac{r^2}{2\sigma^2}\right)\right)$$

$$=\exp\left\{-2\pi\lambda\sigma^2\ln\left[1-\frac{\mu A(R)(s-1)}{2\pi\sigma^2(1-\mu)}\right]\right\}=$$

$$=[1-\beta(s-1)]^{-k},$$

где

$$\beta = \frac{\mu A(R)}{2\pi\sigma^2 (1-\mu)}, \qquad k = 2\pi\lambda\sigma^2.$$

Таким образом $g(s)$ есть производящая функция отрицательного биномиального распределения.

§ 7. ДЕТЕРМИНИРОВАННЫЙ РОСТ ПОПУЛЯЦИИ С РАСПРЕДЕЛЕНИЕМ ПО ВОЗРАСТАМ

В этом параграфе обсуждаются некоторые простые детерминированные модели роста популяций, учитывающие возрастную структуру популяции. Вероятностный вариант этих процессов роста весьма сложен и выходит за рамки данной книги.

А. Простая модель роста

Рассмотрим сначала однородную популяцию. Пусть

$N(t)$ — размер популяции в момент t,

$\nu(t)\,dt$ — число потомков, порождаемых каждым индивидуумом за «малый» интервал $(t,\ t+dt)$. Более точно, число потомков, порожденных индивидуумом за интервал $(t,\ t+h)$, равно $\nu(t)h + o(h)$.

В этих обозначениях имеем

$$N(t+h) = N(t) + N(t)\,\nu(t)\,h + o(h),$$

$$\frac{N(t+h) - N(t)}{h} = \nu(t)\,N(t) + \frac{o(h)}{h}.$$

Переходя к пределу при $h \to 0$ в обеих частях равенства получим

$$\frac{dN(t)}{dt} = \nu(t)\,N(t). \tag{7.1}$$

Решение этого уравнения равно

$$N(t) = N(0)\exp\left(\int\limits_0^t \nu(\tau)\,d\tau\right), \tag{7.2}$$

где $N(0)$ — начальный размер популяции. Если интеграл $\int\limits_0^t \nu(\tau)\,d\tau$ расходится при $t \to \infty$, то размер популяции увеличивается до бесконечности[1]). Если $\nu(\tau) = \nu$ постоянна, то $N(t) = N(0)e^{\nu t}$ и популяция увеличивается экспоненциальным образом с интенсивностью ν.

---

[1]) Особо следует оговорить случай, когда $\int\limits_0^t \nu(\tau)\,d\tau \to \infty$ при $t \to t_0 < \infty$. В этом случае решение (7.2) определено лишь при $t < t_0$. — *Прим. перев.*

Б. Модель, в которой размер популяции влияет на рост

В рассмотренной выше модели увеличение размера популяции не влияло на ее рост. Учтем теперь это влияние, допустив зависимость $\nu(t)$ от $N(t)$. В частности, предположим, что

$$\nu(t)=\begin{cases} \beta\left(1-\dfrac{N(t)}{\alpha}\right) & \text{при } N(t)\leqslant\alpha, \\ 0 & \text{в противном случае,} \end{cases}$$

где α и β — положительные числа. Заметим, что размер популяции не может быть больше α. В этом случае уравнение (7.1) примет вид

$$\frac{dN(t)}{dt}=N(t)\beta\left(1-\frac{N(t)}{\alpha}\right)=\beta N(t)-\frac{\beta}{\alpha}N^2(t). \tag{7.3}$$

Решая уравнение методом разделения переменных, получим

$$N(t)=\frac{\alpha N(0)\,e^{\beta t}}{\alpha+N(0)\left[e^{\beta t}-1\right]}=\frac{\alpha N(0)}{\alpha e^{-\beta t}+N(0)-N(0)\,e^{-\beta t}}. \tag{7.4}$$

Анализ решения показывает, что $N(t)\to\alpha$ при $t\to\infty$.

В. Влияние возрастной структуры

Рассмотрим теперь влияние возрастной структуры на процесс роста популяции.

Введем следующие обозначения:

$\rho(u,\ t)$ — функция частоты индивидуумов возраста u в популяции в момент t, т. е. функция $\rho(u,\ t)$ обладает тем свойством, что $\int\limits_{u_1}^{u_2}\rho(u,\ t)\,du$ — доля индивидуумов в популяции в момент t, возраст которых лежит в интервале $(u_1,\ u_2)$. Реальное число индивидуумов такого возраста равно, естественно, $N(t)\int\limits_{u_1}^{u_2}\rho(u,\ t)\,du$; (7.5)

$b(t)$ — интенсивность рождения новых индивидуумов в популяции в момент t. Более точно, $\int\limits_{t_1}^{t_2}b(t)\,dt$ — число новых индивидуумов, рожденных в интервале времени $(t_1,\ t_2)$; (7.6)

$\lambda(u)\,dt$ — среднее число потомков одного индивидуума возраста u в dt единиц времени;

(7.7)

$l(u)$ — вероятность того, что время жизни индивидуума превышает u;

(7.8)

$c(u)$ — инфинитезимальная интенсивность гибели, т. е. вероятность того, что индивидуум возраста u погибнет в следующие h единиц времени, равна $c(u)h + o(h)$.

(7.9)

Соотношение между $l(\cdot)$ и $c(\cdot)$ может быть получено следующим образом. При известных u, $h \geqslant 0$ время жизни индивидуума превышает $u + h$ тогда и только тогда, когда он является живым в момент u (отсчитываемый от рождения) и не погибает в следующие h единиц времени. Таким образом, получаем

$$l(u+h) = l(u)[1 - c(u)h] + o(h)$$

и

$$\frac{l(u+h) - l(u)}{h} = -\frac{l(u)c(u)h + o(h)}{h}.$$

Переходя к пределу при $h \to 0$, получаем

$$\frac{dl(u)}{du} = -l(u)c(u).$$

Решая уравнение, находим

$$l(u) = l(0)\exp\left[-\int_0^u c(\xi)\,d\xi\right] = \exp\left[-\int_0^u c(\xi)\,d\xi\right], \tag{7.10}$$

поскольку $l(0) = 1$.

При рассмотрении влияния возрастной структуры на растущую популяцию мы будем интересоваться функцией $b(t)$, т. е. функции $\lambda(u)$, $l(u)$ и $c(u)$ будут считаться известными, а задача будет состоять в нахождении $b(u)$. Интенсивность рождения новых индивидуумов в популяции в момент t имеет две составляющие. Одна составляющая, скажем $b_0(t)$, является интенсивностью воспроизводства для тех индивидуумов в популяции в момент t, которые уже существовали в момент 0. Плотность индивидуумов возраста u в популяции в момент 0 равна $\rho(u, 0)$. Вероятность того, что индивидуум, имеющий в момент 0 возраст u, будет жить в момент t (в это время его возраст будет равен $u + t$), равна $\frac{l(t+u)}{l(u)}$. Следовательно, доля тех индивидуумов, которые имеют в момент 0 возраст u и которые доживут до момента t, равна $[l(t+u)/l(u)]\rho(u, 0)$. Интенсивность воспроизводства для индивидуумов возраста $t+u$ равна $\lambda(t+u)$. Усредняя по всевозможным возрастам, получаем

$$b_0(t)=N(0)\int_0^\infty \lambda(t+u)\frac{l(t+u)}{l(u)}\rho(u, 0)\,du. \tag{7.11}$$

Другой составляющей $b(t)$ является интенсивность воспроизводства новых индивидуумов в момент t для тех индивидуумов в популяции, которые родились после момента 0. Интенсивность рождения новых индивидуумов в популяции в момент τ равна $b(\tau)$. При $0<\tau\leqslant t$ вероятность того, что индивидуум, рожденный в момент τ, будет жить в момент t (в это время его возраст будет равен $t-\tau$), равна $l(t-\tau)$. Интенсивность воспроизводства для индивидуумов возраста $t-\tau$ равна $\lambda(t-\tau)$. Отсюда получаем

$$b(t)=b_0(t)+\int_0^t \lambda(t-\tau)\,l(t-\tau)\,b(\tau)\,d\tau, \tag{7.12}$$

где функция $b_0(t)$ дается формулой (7.11).

Соотношение (7.12) является непрерывной формой уравнения восстановления (см. § 1 гл. 3). Его решение можно найти с помощью метода последовательных приближений.

Пример. Предположим, что как интенсивность рождения, так и инфинитезимальная интенсивность гибели являются постоянными, не зависящими от возраста, т. е. $\lambda(u)=\lambda$, $c(u)=c$. Тогда в силу (7.10) вероятность того, что индивидуум доживет до возраста u, равна

$$l(u)=e^{-cu}. \tag{7.13}$$

Предположим, что в момент 0 рождается первый индивидуум, тогда

$$N(0)\int_0^\infty \rho(u, 0)\,du=1. \tag{7.14}$$

В действительности возрастная плотность $\rho(u, 0)$ является дельта-функцией, соответствующей вырожденному распределению, сконцентрированному в точке $u=0$. Из (7.11), (7.13) и (7.14) заключаем, что

$$b_0(t)=\lambda e^{-ct}. \tag{7.15}$$

Следовательно, в силу (7.13) и (7.15) уравнение (7.12) приобретает вид

$$b(t)=\lambda e^{-ct}+\lambda\int_0^t e^{-c(t-\tau)}b(\tau)\,d\tau. \tag{7.16}$$

Решим уравнение (7.16) относительно функции $b(\cdot)$. Умножим обе части уравнения на e^{ct}, получим

$$e^{ct}b(t) = \lambda + \lambda \int_0^t e^{c\tau} b(\tau)\, d\tau$$

и обозначим

$$f(\tau) = e^{c\tau} b(\tau). \tag{7.17}$$

Уравнение относительно $f(\cdot)$ имеет вид

$$f(t) = \lambda + \lambda \int_0^t f(\tau)\, d\tau.$$

Очевидно, $f(0) = \lambda$. При дифференцировании обеих частей по t получаем $f'(t) = \lambda f(t)$. Таким образом, $f(t) = \lambda e^{\lambda t}$ и, подставляя это выражение в (7.17), получаем

$$b(t) = \lambda e^{(\lambda - c)t}. \tag{7.18}$$

Определив в этом примере $b(t)$, используем этот результат для нахождения возрастной структуры, которая задается величиной $N(t)\rho(u, t)$. Поскольку мы предположили при выводе (7.18), что в момент 0 рождается первый индивидуум, нам нужно рассмотреть лишь случай $u \leqslant t$. Индивидуум в момент t имеет возраст u в том и только том случае, если он родился в момент $t - u$. Интенсивность рождения новых индивидуумов в популяции в момент $t - u$ равна $b(t-u) = \lambda e^{(\lambda - c)(t-u)}$. Вероятность того, что индивидуум доживет до возраста u, равна e^{-cu} в силу (7.13). Отсюда

$$N(t)\rho(u, t) = e^{-cu} b(t-u), \qquad u \leqslant t. \tag{7.19}$$

Подставляя (7.18) в (7.19), получаем

$$N(t)\rho(u, t) = e^{-cu}\lambda e^{(\lambda - c)(t-u)} = \lambda e^{-\lambda u} e^{(\lambda - c)t}, \qquad u \leqslant t. \tag{7.20}$$

Обратимся теперь к исходной формулировке, сохраняя лишь предположение о том, что в момент 0 рождается первый индивидуум; тогда вывод равенства (7.19) сохранится, и в общем случае при $u \leqslant t$ имеем

$$N(t)\rho(u, t) = b(t-u)\, l(u). \tag{7.21}$$

Здесь $b(t)$ определяется как решение уравнения восстановления (7.12) при условии $b_0(t) = 0$, поскольку в момент $t = 0$ не было живущих индивидуумов. Таким образом,

$$b(t) = \int_0^t \lambda(t-\tau)\, l(t-\tau)\, b(\tau)\, d\tau = \int_0^t b(t-u)\, \lambda(u)\, l(u)\, du.$$

Пусть $\phi(u)=\lambda(u)l(u)$, тогда уравнение примет вид

$$b(t)=\int_0^t b(t-u)\phi(u)\,du. \qquad (7.22)$$

В качестве пробного решения уравнения (7.22) выберем

$$b(t)=e^{\gamma t} \qquad (7.23)$$

($\gamma=\mathrm{const}$), где постоянную γ следует выбрать так, чтобы уравнение (7.22) выполнялось при больших t. Подставляя (7.23) в (7.22), получим условие

$$e^{\gamma t}=\int_0^t e^{\gamma(t-u)}\phi(u)\,du=e^{\gamma t}\int_0^t e^{-\gamma u}\phi(u)\,du,$$

или

$$\int_0^t e^{-\gamma u}\phi(u)\,du=1. \qquad (7.24)$$

Мы интересуемся предельным поведением (при $t\to\infty$) возрастной структуры популяции. Устремим в (7.24) t к ∞, получим

$$R(\gamma)=\int_0^\infty e^{-\gamma u}\phi(u)\,du=1. \qquad (7.25)$$

В силу определения $R(\gamma)$ является строго убывающей функцией по γ. Следовательно, уравнение (7.25) имеет максимум один положительный корень. Пусть

$$R=\int_0^\infty \phi(u)\,du=\int_0^\infty \lambda(u)\,l(u)\,du.$$

Величина R называется репродуктивным числом индивидуума и равна среднему числу потомков, порожденных индивидуумом за время жизни. Иногда R называют мальтусовской интенсивностью.

Если $R>1$, то $R(\gamma)$ имеет вид, показанный на рис. 1, и решение $\gamma_0>0$ уравнения (7.25) существует.

В этом случае $b(t)$ асимптотически пропорциональна $\exp(\gamma_0 t)$ и рост популяции — экспоненциальный. Если $R<1$, то $R(\gamma)$ имеет вид, показанный на рис. 2, и уравнение (7.25) имеет решение $\gamma_0<0$. В этом случае $b(t)$ асимптотически пропорциональна $\exp(\gamma_0 t)$, а популяция вымирает с экспоненциальной скоростью.

Если $R=1$, то задача должна исследоваться вероятностными методами.

Полученные выше результаты найдены эвристическим путем. Предположения и анализ, необходимые для того, чтобы придать строгий смысл решению, выходят за рамки данной книги. Основной вывод, который получен, — тот, что при соответствующих усло-

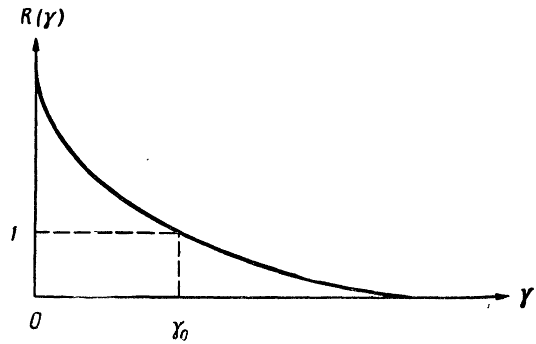


Рис. 1.

виях популяция растет экспоненциальным образом. Еще одно подтверждение этого явления мы получим, рассматривая ниже дискретную модель (§ 8).

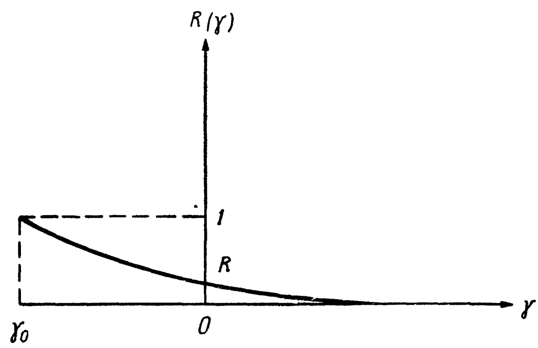


Рис. 2.

Продолжим теперь эвристические рассуждения. В случае $R > 1$ найдем асимптотическую возрастную структуру, задаваемую плотностью $\rho(u, t)$. В силу (7.21)

$$\rho(u, t) = \frac{b(t-u)\, l(u)}{N(t)}, \qquad u \leqslant t.$$

Далее, $b(t-u)$ асимптотически пропорциональна $\exp[\gamma_0(t-u)]$. Следовательно, $N(t)$ асимптотически пропорциональна $\exp(\gamma_0 t)$, поэтому $\rho(u, t)$ асимптотически пропорциональна

$$\frac{\exp[\gamma_0(t-u)]\, l(u)}{\exp(\gamma_0 t)} = \exp(-\gamma_0 u)\, l(u).$$

Коэффициент пропорциональности можно найти, используя тот факт, что плотность является вероятностным распределением. Следовательно, возрастная структура популяции при больших t задается асимптотической функцией плотности

$$\rho(u, t) \sim \frac{\exp(-\gamma_0 u)\, l(u)}{\int\limits_0^\infty \exp(-\gamma_0 x)\, l(x)\, dx}.$$

§ 8. ДИСКРЕТНАЯ ВОЗРАСТНАЯ МОДЕЛЬ

Рассмотрим теперь строгий подход к задаче, изучавшейся в предыдущем параграфе, в случае дискретного времени. Пусть $t = 0, 1, 2, \ldots$, и пусть

$n_x^{(t)}$ — число индивидуумов возраста x в момент t;

P_x — доля индивидуумов возраста x, доживающих до возраста $x+1$;

F_x — число индивидуумов, рождающихся от каждого родителя возраста x в следующий момент времени.

Предполагается, что F_x и P_x не зависят от t и $P_x > 0$, $x = 0, \ldots, m-1$, $F_x > 0$, $x = 0, 1, 2, \ldots, m$. Предположим, что возраст любого индивидуума ограничен числом m, т. е. положим $P_m = 0$; тогда переходные соотношения для возрастной структуры между моментами $t = 0$ и $t = 1$ имеют вид

$$\begin{aligned} n_0^{(1)} &= \sum_{x=0}^{m} n_x^{(0)} F_x, \\ n_1^{(1)} &= P_0 n_0^{(0)}, \\ n_2^{(1)} &= P_1 n_1^{(0)}, \\ &\vdots \\ n_m^{(1)} &= P_{m-1} n_{m-1}^{(0)}. \end{aligned}$$

Запишем эти соотношения в матричном виде

$$\mathbf{n}^{(1)} = \mathbf{M}\mathbf{n}^{(0)}, \tag{8.1}$$

где $\mathbf{n}^{(0)}=\left(n_0^{(0)}, n_1^{(0)}, \ldots, n_m^{(0)}\right)$, $\mathbf{n}^{(1)}=\left(n_0^{(1)}, n_1^{(1)}, \ldots, n_m^{(1)}\right)$ и матрица $\mathbf{M}$ записывается в виде

$$\mathbf{M}=\begin{Vmatrix} F_0 & F_1 & F_2 & \ldots & \ldots & F_m \\ P_0 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & P_1 & \ldots & \ldots & \ldots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \ldots & P_{m-1} & 0 \end{Vmatrix}. \tag{8.2}$$

$\mathbf{M}$ является матрицей с неотрицательными элементами. Свойства таких матриц даны в § 2 приложения. Поскольку P_x и F_x не зависят от времени, точно такие же переходные соотношения справедливы для любых двух последовательных моментов времени. Таким образом, применяя последовательно формулу (8.1), получим

$$\mathbf{n}^{(t)}=\mathbf{M}^t\mathbf{n}, \quad \mathbf{n}^{(t)}=\left(n_0^{(t)}, n_1^{(t)}, \ldots, n_m^{(t)}\right), \qquad t=1, 2, 3, \ldots, \tag{8.3}$$

где мы обозначили $n_x=n_x^{(0)}$, $x=1, 2, \ldots, m$.

При достаточно больших t все элементы матрицы $\mathbf{M}^t$ строго положительны. Кроме того, существует собственное значение $\lambda_0>0$, которое по абсолютной величине строго превосходит другие собственные значения (теорема 2.2 приложения). Для любого вектора $\mathbf{n}$ вектор $\mathbf{M}^t\mathbf{n}$ асимптотически равен $\lambda_0^t\mathbf{z}$, где $\mathbf{z}$ пропорционален правому собственному вектору матрицы $\mathbf{M}$, отвечающему собственному значению λ_0. Асимптотически при $t\to\infty$ соотношение (8.3) принимает вид

$$\mathbf{n}^{(t)}=\mathbf{M}^t\mathbf{n}\sim\lambda_0^t\mathbf{z}=[\exp(t\ln\lambda_0)]\,\mathbf{z}$$

и величина $\ln\lambda_0$ соответствует критическому значению γ_0, введенному при эвристическом рассмотрении варианта с непрерывным временем. Популяция растет экспоненциально, если $\lambda_0>1$, и вымирает экспоненциально быстро, если $\lambda_0<1$.

ЗАДАЧИ

1. Предположим, что некоторый объект имеет целочисленные размеры и что за время h объект с вероятностью $h+o(h)$ вырастает на одну единицу независимо от предыстории роста и от текущего размера. При условии, что объект первоначально имеет размер 1, найти вероятности того, что за время t он вырастет до размера n.

Указание: Рост объекта образует процесс Юла; см. § 1 гл. 7.

Ответ: $e^{-t}\left(1-e^{-t}\right)^{n-1}$.

2. N бактерий распределены равномерно и независимо на слайде микроскопа площади A. Для наблюдения произвольным образом выбирается область площади a. Найти вероятность того, что на этой площади имеются k бактерий.

Ответ:

$$p(k) = \binom{N}{k}\left(\frac{a}{A}\right)^k\left(1-\frac{a}{A}\right)^{N-k}.$$

3. Показать, что при $N \to \infty$ и $a \to 0$ так, что $\left(\frac{a}{A}\right)N \to c$ $(0 < c < \infty)$,

$$p(k) \to e^{-c}\frac{c^k}{k!}.$$

*4. Предположим, что между бактериями происходит реакция, если две или более бактерий отстоят друг от друга меньше, чем на расстояние r. Найти функцию распределения числа реакций в области A при $r \to 0$, $N \to \infty$ и $\frac{1}{A}\pi r^2 N^2 \to$ $\to \lambda$, $0 < \lambda < \infty$.

Ответ: $p(l) = e^{-\lambda}\dfrac{\lambda^l}{l!}$.

5. Рассмотрим двумерное пуассоновское распределение (с интенсивностью ν) частиц на плоскости. Найти функцию распределения $F_D(x)$ расстояния между произвольной частицей и ее ближайшим соседом. Найти среднее расстояние.

Ответ:

$$F_D(x) = 1 - \exp(-\nu\pi x^2); \quad \mathrm{M}(D) = \frac{1}{2\sqrt{\nu}}.$$

6. Решить задачу 5 в трехмерном случае.

Ответ:

$$F_D(x) = 1 - \exp\left(-\nu\frac{4}{3}\pi x^3\right); \quad \mathrm{M}(D) = \frac{\Gamma\left(\frac{1}{3}\right)}{(36\nu\pi)^{1/3}}.$$

7. Предположим, что прибор подвержен влиянию одного из k возможных воздействий $E_1, E_2, \ldots, E_k$, которые могут иметь место с соответствующими вероятностями $c_1, c_2, \ldots, c_k$ $\left(\sum_{j=1}^{k} c_j = 1\right)$. При каждом воздействии опасные пики нагрузки происходят в соответствии с пуассоновским потоком с параметром λ_j, $j = 1, 2, \ldots, k$. Условная вероятность того, что прибор выйдет из строя при пиковой нагрузке воздействия E_j, равна p_j. Найти вероятность того, что прибор выйдет из строя за время t.

Ответ:

$$\mathrm{P}\{T < t\} = 1 - \sum_{j=1}^{k} c_j \exp(-\lambda_j p_j t).$$

8. Группа из n инженеров работает над проектом. Время безошибочной работы инженера имеет функцию распределения $F(t)$. Если ошибка сделана, то с вероятностью p это ошибка типа I, а с вероятностью $1 - p$ — типа II. Ошибка типа I настолько серьезна, что если кто-либо допустит ее в любой момент времени, то весь проект наверняка обречен на неудачу. Ошибка же типа II настолько незначительна, что она может испортить весь проект, только если все инженеры независимо друг от друга допустят ее. Найти вероятность того, что в момент t проект еще не обречен на неудачу.

Указание: Найти вероятность того, что в точности k инженеров сделали ошибки типа II, а остальные не допустили ошибок к моменту t. Показать, что эта вероятность равна $\binom{n}{k}[(1-p)F(t)]^k[1-F(t)]^{n-k}$.

Ответ: $[1-pF(t)]^n-[(1-p)F(t)]^n$.

9. Рассмотрим схему, состоящую из m параллельных подсистем. Каждая подсистема состоит из n одинаковых элементов, включенных последовательно. Предположим, что время безотказной работы каждого элемента имеет функцию распределения $F(t)$. Показать, что вероятность того, что схема будет работать в момент t, равна $1-[1-(1-F(t))^n]^m$.

10. Рассмотрим электрические элементы $i=1,2,\ldots,m$ сложной системы S. Пусть $F_i(t)$ — функция распределения времени безотказной работы i-го элемента. Пусть $1-p_i$ — условная вероятность того, что если i-й элемент выходит из строя, то вся система перестает работать.

(i) Говорят, что система элементов является полупараллельной, если она отказывает либо (с вероятностью 1) при отказе всех элементов, либо (с вероятностью $1-p_i$) при отказе одного элемента (скажем, i-го).

(ii) Говорят, что система элементов является последовательной, если она отказывает при отказе любого элемента.

Надежностью системы в момент t называется вероятность того, что система S работает.

(1) Предположим, что $F_i(t)=F(t)$, $p_i=p$ $(i=1,2,\ldots,m)$ и рассмотрим полупараллельную систему. Доказать, что надежность системы в момент t равна $[1-F(t)+pF(t)]^m-[pF(t)]^m$.

(2) Пусть выполнены условия пункта (1). Предположим, что $F(t)=1-e^{-\lambda t}$ $(\lambda>0)$. Доказать, что среднее время безотказной работы системы равно

$$\frac{1}{\lambda}\sum_{k=1}^{m}\frac{1}{k}p^{m-k}.$$

Указание для пункта (2): Пусть u_m — среднее время безотказной работы полупараллельной системы, состоящий из m компонент. Вывести рекуррентную формулу

$$u_m=\frac{1}{m\lambda}+pu_{m-1}.$$

***11.** Предположим, что новые мутантные виды возникают в соответствии с пуассоновским потоком с параметром ν. Популяция, порожденная каждым новым мутантом, развивается в соответствии с процессом рождения и гибели, имеющим параметры $\lambda_n=n\lambda$, $\mu_n=n\mu$, где $\mu>\lambda$. Различные мутанты порождают независимо развивающиеся популяции. (Напомним, что поскольку $\mu>\lambda$, каждая такая популяция вырождается за конечное время с вероятностью 1, см. § 7 гл. 7.) Показать, что число мутантных популяций L_t, существующих в момент t, имеет пуассоновское распределение.

Ответ: Пусть $\Omega(\xi)$ — функция распределения времени вырождения линейного процесса рождения и гибели $(\lambda_n=n\lambda,\ \mu_n=n\mu)$ с единичным начальным условием. Параметр пуассоновского распределения равен

$$\nu\int_0^t[1-\Omega(\xi)]\,d\xi.$$

12 (продолжение). Найти предельное распределение величины L_t при $t \to \infty$. *Ответ:* Пуассоновское распределение с параметром

$$\nu \int_0^\infty [1 - \Omega(\xi)]\, d\xi = \nu \frac{\mu}{\lambda}\left(-\ln\left(1 - \frac{\lambda}{\mu}\right)\right).$$

13. Рассмотрим множество кругов на плоскости, центры которых распределены в соответствии с пространственным пуассоновским процессом с параметром $\lambda|A|$, где $|A|$ — площадь множества A. (Таким образом, число центров $\xi(A)$, расположенных в множестве A, имеет распределение $\mathrm{P}\{\xi(A) = k\} = e^{-\lambda|A|}\frac{1}{k!}(\lambda|A|)^k$.) Радиус каждого круга по предположению является случайной величиной, независимой от местоположения центра круга, с плотностью $f(r)$ и конечным вторым моментом. Показать, что семейство случайных величин $C(r)$ — число кругов, покрывающих начало координат, центр которых отстоит от начала на расстояние, меньшее r, — является пуассоновским процессом с непрерывным параметром r (см. задачу 13 гл. 7).

Указание: Доказать, что

(1) вероятность того, что существует круг с центром, лежащим в кольце с внутренним и внешним радиусами r и $r + dr$ соответственно, покрывающий начало координат, равна

$$\lambda 2\pi r\, dr \int_r^\infty f(\rho)\, d\rho + o(dr);$$

(2) события, определяемые непересекающимися кольцами, независимы. Показать, что $C(r)$ является неоднородным пуассоновским процессом с параметром

$$\lambda(r) = 2\pi\lambda r \int_r^\infty f(\rho)\, d\rho.$$

14. Показать, что число кругов, покрывающих начало координат (см. задачу 13), является пуассоновской случайной величиной с параметром

$$\lambda \int_0^\infty \pi r^2 f(r)\, dr.$$

15. Рассмотрим сферы в трехмерном пространстве, центры которых распределены по пуассоновскому закону с параметром $\lambda|A|$, где $|A|$ — объем множества A. Радиусы всех сфер распределены по закону $F(r)$ с плотностью $f(r)$, имеющему конечный третий момент. Показать, что число сфер, содержащих внутри себя точку t, является пуассоновской случайной величиной с параметром

$$\frac{4}{3}\lambda\pi \int_0^\infty r^3 f(r)\, dr.$$

16. Предположим, что частицы прибывают в соответствии с пуассоновским потоком, имеющим параметр λ. По прибытии частица принимает одно из r $(r \geqslant 1)$ состояний с вероятностями $p_1, p_2, \ldots, p_r$ соответственно. После этого состояние каждой частицы изменяется в соответствии с законами однородного по времени марковского процесса с переходными вероятностями $P_{ij}(t) = \mathrm{P}\{$состояние в момент t равно $j \mid$ состояние в момент 0 было $i\}$, $i, j = 1, \ldots, r$. Пусть $\mathbf{Y}(t) = \{X_1(t), \ldots, X_r(t);\ t \geqslant 0\}$ — векторный случайный процесс, где $X_i(t)$ — число частиц, находящихся в состоянии i в момент t. Доказать, что $\{\mathbf{Y}(t);\ t \geqslant 0\}$ — однородный по времени марковский процесс. (Состояния можно интерпретировать как различные стадии некоторой болезни.)

17 (продолжение). Рассмотрим одну частицу типа i в момент 0. Пусть

$$\delta_{ij}(t)=\begin{cases}1, & \text{если в момент } t \text{ эта частица находится в состоянии } j;\\ 0 & \text{в противном случае.}\end{cases}$$

Доказать соотношение

$$\mathrm{M}\left\{z_1^{\delta_{i1}(t)} z_2^{\delta_{i2}(t)} \ldots z_r^{\delta_{ir}(t)}\right\}=\sum_{j=1}^{r} z_j P_{ij}(t).$$

18 (продолжение). Доказать, что производящая функция вектора $\{X_1(t),\ldots$ $\ldots, X_r(t)\}$ равна

$$\varphi(z_1, z_2, \ldots, z_r;\ t)=\mathrm{M}\left\{z_1^{X_1(t)} z_2^{X_2(t)} \ldots z_r^{X_r(t)}\right\}=$$

$$=\exp\left\{\lambda\sum_{i=1}^{r} p_i \sum_{k=1}^{r}(z_k-1)\int_0^t P_{ik}(\tau)\,d\tau\right\}.$$

Указание (см. анализ § 3).

1. Наложить условие на число частиц, прибывших до момента t, и воспользоваться формулой полной вероятности.

2. Использовать тот факт, что в пуассоновском процессе при известном числе событий, происшедших до момента t, моменты наступления событий распределены независимо и равномерно в $(0, t)$.

3. Использовать тот факт, что отдельные частицы развиваются независимо друг от друга.

19. Рассмотрим следующую модель роста популяции, состоящей из двух типов частиц — нормального и мутантного. Продолжительность жизни частиц нормального типа имеет экспоненциальное распределение со средним значением λ^{-1}. По окончании времени жизни нормальная частица с вероятностью p порождает две нормальные частицы, а с вероятностью $q = 1 - p$ — одну нормальную и одну мутантную. Все потомки нормального типа развиваются так же, как родители. Частица мутантного типа живет случайное время, имеющее экспоненциальное распределение со средним значением μ^{-1}, и затем порождает две мутантные частицы, развивающиеся таким же образом, как их родитель. Предположим, что все частицы развиваются независимо. В начале имеется одна частица нормального типа. Пусть $\{X(t),\ Y(t)\}$ — число частиц нормального и мутантного типов в момент t соответственно. Найти производящие функции $\psi_1(z, t) = \mathrm{M}(z^{X(t)})$ и $\psi_2(z, t) = \mathrm{M}(z^{Y(t)})$.

Указание: (а) Показать, что $\{X(t)\}$ — процесс Юла с параметром λp.

(б) Вывести следующее интегральное уравнение относительно $\psi = \psi_2$, налагая условие на время первого деления и пользуясь формулой полной вероятности:

$$\psi(z, t)=e^{-\lambda t}+\int_0^t \lambda e^{-\lambda\tau}\left[p\psi^2(z, t-\tau)+q\psi(z, t-\tau)\,g(z, t-\tau)\right]d\tau,$$

где

$$g(z, t)=ze^{-\mu t}\left[1-z\left(1-e^{-\mu t}\right)\right]^{-1}.$$

Произвести под интегралом замену переменных $t - \tau = u$ и, продифференцировав затем уравнение по t, получить

$$\frac{\partial\psi}{\partial t}=-\lambda\psi+\lambda\psi\,(p\psi+qg).$$

Пусть $\varphi = \frac{1}{\psi}$; тогда

$$\frac{\partial \varphi}{\partial t} = \lambda\varphi - \lambda p - \lambda q g \varphi.$$

Решив это уравнение, найти, что

$$\psi = F(z, t)\left[1 - \lambda_p \int_0^t F(u, t)\, du\right]^{-1},$$

где $F(z, t) = e^{-\lambda t}\left[1 - z\left(1 - e^{-\mu t}\right)\right]^{-\frac{\lambda q}{\mu}}$. В частном случае $\lambda = \mu$ показать, что

$$\frac{1}{\psi} = 1 + \frac{1-z}{z} e^{\mu t}\left\{1 - \left(1 - z + z e^{-\mu t}\right)^q\right\}.$$

20. Пусть $\{X_i(t), t \geqslant 0\}_{i=1}^n$ — независимые пуассоновские процессы с параметром λ. Найти функцию распределения времени до того момента, пока во всех процессах не произойдет хотя бы одно событие.

Ответ: $\mathrm{P}\{T \leqslant t\} = \left(1 - e^{-\lambda t}\right)^n$.

21. Некоторый эксперимент имеет N возможных исходов, i-й из которых происходит с вероятностью p_i, $\sum\limits_{i=1}^N p_i = 1$. Пусть T — число испытаний, необходимое для того, чтобы осуществились все различные исходы. Показать, что

$$\mathrm{M}(T) = \int_0^\infty t d\left[\prod_{i=1}^n \left(1 - e^{-tp_i}\right)\right].$$

Указание: Отождествить проведение эксперимента с моментом наступления события пуассоновского процесса с параметром $\lambda = 1$. Использовать результат задачи 41 (стр. 527).

ЗАМЕЧАНИЯ

Источником материала для данной главы послужили книги Бартлетта [1] и Харриса [2].

Дальнейшее обсуждение составных случайных процессов с акцентом на их приложениях содержится в книге Бартлетта [3].

Обширная библиография до 1960 г. собрана в книге Баруча-Рида [4].

ЛИТЕРАТУРА

1. Bartlett M. S., Stochastic Population Models in Ecology and Epidemiology, Wiley, New York, 1960.
2. Харрис Т. Е., Теория ветвящихся случайных процессов, «Мир», М., 1966.
3. Бартлетт М. С., Введение в теорию случайных процессов, ИЛ, М., 1958.
4. Баруча-Рид А. Т., Элементы теории марковских процессов и их приложения, «Наука», М., 1969.

Глава 13

ДЕТЕРМИНИРОВАННЫЕ И СЛУЧАЙНЫЕ ГЕНЕТИЧЕСКИЕ И ЭКОЛОГИЧЕСКИЕ ПРОЦЕССЫ

Одной из главных областей применения вероятностных идей являются системы, изучаемые в генетике. В § 1 мы приведем качественное описание основных понятий, лежащих в основе генетического механизма, а в следующих двух параграфах сконцентрируем наше внимание на некоторых классических детерминированных моделях. В § 4 будет рассмотрено несколько вероятностных моделей генетических систем, а также будет обсуждена их взаимосвязь с ветвящимися процессами. Представленный ниже материал следует расценивать как введение в богатый и чрезвычайно интересный раздел генетики популяции, связанный с ее внутренней вероятностной структурой.

§ 1. ГЕНЕТИЧЕСКИЕ МОДЕЛИ. ОПИСАНИЕ ГЕНЕТИЧЕСКОГО МЕХАНИЗМА

В ядре клетки (рис. 1) находятся *хромосомы*, которые «ответственны» за наследственные характеристики организма. Их число различно для различных организмов [1]).

Единица, контролирующая полностью или частично наследование некоторого признака, называется *геном*. Место, которое ген занимает в хромосоме, называется *локусом*.

Набор хромосом может содержать по одной хромосоме каждого типа (*гаплоиды*), как у некоторых грибов, по паре хромосом (*диплоиды*), как у млекопитающих, или группы большего размера (триплоиды, тетраплоиды и в общем случае полиплоиды), как у многих растений.

В каждом локусе гены могут принадлежать к различным контрастирующим типам, которые называются *аллелями* [2]). Различные возможные комбинации аллелей в соответствующих локусах разных хромосом называются *генотипами*. При диплоидии, если имеются аллели A и a, возможны следующие генотипы: AA, Aa,

---

[1]) У большинства видов организмов число хромосом в клетке колеблется от 10 до 50. В клетках человека и многих животных содержится по 46 хромосом. — *Прим. перев.*

[2]) Так, некоторый ген, «отвечающий», например, за карий цвет глаз, является аллелем гена, «отвечающего» за голубой цвет глаз. — *Прим. перев.*

aa. Генотипы *AA* и *aa* называются *гомозиготными*, а генотип *Aa* — *гетерозиготным.*

Фенотипами называются различные внешние проявления аллельных комбинаций. Например, если *A* доминирует над *a*, то раз-

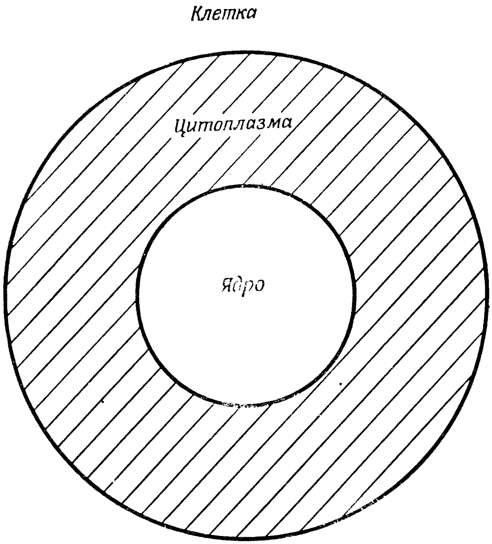


Рис. 1.

ные генотипы *AA* и *Aa* будут определять один и тот же фенотип, а комбинация *aa* определяет другой фенотип.

А. Размножение и закон Харди — Вейнберга

Существует два способа клеточного деления. Более распространенным из них является *митоз* (рис. 2). При этом способе каждая хромосома делится на две и, когда клетка расщепляется на две дочерние клетки, каждая из вновь образованных клеток получает по одной хромосоме из каждой пары. Таким образом, каждая новая клетка имеет полный набор хромосом.

Процесс второго типа — *мейоз* (рис. 3) — происходит при половом размножении диплоидных клеток. Каждая хромосома при этом удваивается. Затем образуются четыре новые клетки, каждая из которых содержит одну хромосому. Таким образом, каждая из четырех новых клеток содержит половину нормального набора хромосом. Такая клетка называется *гаметой.* Когда две гаметы соединяются, образуется оплодотворенная клетка — *зигота,* которая имеет полный набор хромосом.

Рассмотрим теперь какой-либо локус в паре хромосом у диплоидного организма для случая, когда возможны два аллеля A и a. Повсюду в этом параграфе будет рассматриваться бесконечная популяция, различные поколения которой не смешиваются.

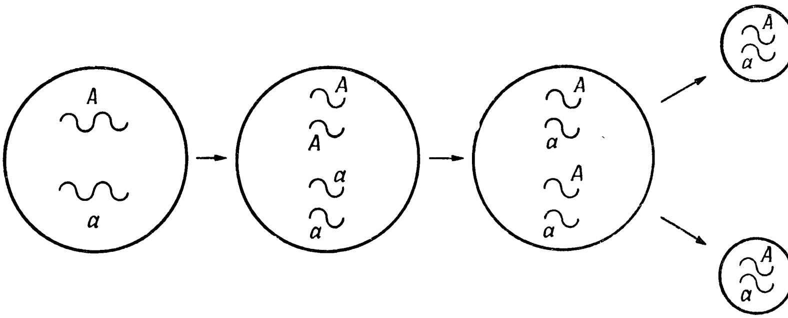


Рис. 2.

Предположим, что скрещивается генотип AA с генотипом Aa. Первый из них производит только гаметы A, а второй — гаметы A и a в равных пропорциях. При рассматриваемом спаривании (скрещивании) гаметы, образованные генотипом AA, по-видимому, одинаково часто будут соединяться с различными гаметами, производимыми генотипом Aa. Таким образом, для случая спаривания

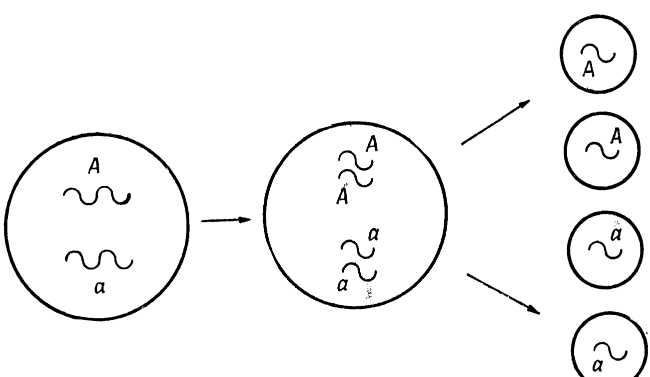


Рис. 3.

$AA \times Aa$ частота потомков типа AA равна 1/2 и такая же частота для потомков типа Aa. Аналогично при спаривании $AA \times aa$ образуются только потомки типа Aa. Спаривание $Aa \times Aa$ дает следующее соотношение потомков: $\frac{1}{4} AA$, $\frac{1}{2} Aa$, $\frac{1}{4} aa$ и т. д. Эти пропорции исходов различных возможных спариваний известны под названием *отношения при расщеплении.*

Рассмотрим некоторую раздельнополую популяцию с тремя генотипами AA, Aa, aa, в которой осуществляется случайное спаривание. Пусть исходное соотношение генотипов следующее: доля AA равна u, доля Aa равна v и доля aa равна w, $u+v+w=1$. В табл. I показаны результаты случайного спаривания.

Таблица I

Скрещивание	Типы потомков	Частота данного типа скрещивания
$AA \times AA$	AA	u^2
$AA \times Aa$	$\frac{1}{2}AA + \frac{1}{2}Aa$	$2uv$
$AA \times aa$	Aa	$2uw$
$Aa \times Aa$	$\frac{1}{4}AA + \frac{1}{2}Aa + \frac{1}{4}aa$	v^2
$Aa \times aa$	$\frac{1}{2}Aa + \frac{1}{2}aa$	$2vw$
$aa \times aa$	aa	w^2

Следовательно, в следующем поколении эти три генотипа будут представлены в отношении

$$\begin{array}{ccc} AA & Aa & aa \\ u^2+uv+\frac{1}{4}v^2 & 2uw+vw+uv+\frac{1}{2}v^2 & w^2+wv+\frac{1}{4}v^2 \end{array}$$

Положим $p=u+\frac{v}{2}$, $q=w+\frac{v}{2}$, тогда эти пропорции примут вид

$$\begin{array}{ccc} AA & Aa & aa \\ p^2 & 2qp & q^2 \end{array}$$

и $p+q=1$. Точно таким же образом мы найдем, что после следующего спаривания генотипы распределяются в отношении

$$\begin{array}{ccc} AA & Aa & aa \\ (p^2+pq)^2=p^2 & 2(p^2+pq)(q^2+pq)=2pq & (q^2+pq)^2=q^2 \end{array}$$

Таким образом, начиная со второго поколения, соотношение между генотипами не меняется, т. е. устойчивое состояние достигается за одно поколение. Этот результат известен, как *закон Харди — Вейнберга.*

Можно получить этот же результат следующим образом. Рассмотрим популяцию гамет: если начальное соотношение между генотипами то же, что и выше, то соответствующая доля гамет A

в популяции равна $u+\frac{v}{2}=p$, а гамет a равна $w+\frac{v}{2}=q$, так что в следующем поколении генотипы представлены в отношении

$$\begin{array}{ccc} AA & Aa & aa \\ p\cdot p=p^2 & 2pq & qq=q^2 \end{array}$$

Рассматривая совокупность гамет в новой популяции, найдем:

$$\text{доля } (A)=p^2+pq=p(p+q)=p,$$
$$\text{доля } (a)=q^2+pq=q(p+q)=q,$$

что указывает на достигнутую устойчивость. Факт совпадения результатов, полученных обоими методами, известен в виде следующего предложения: «случайное слияние гамет» эквивалентно случайному спариванию. (Вероятностная интерпретация этой теоремы содержится в задаче 17.)

Теперь мы рассмотрим, как три фактора — мутация, миграция и отбор — влияют на популяцию генотипов в предположении о случайном спаривании.

Б. Мутация

1. Рассмотрим гаплоидную популяцию генов A и a. Предположим, что в промежутке между любыми двумя последующими поколениями доля μ генов a мутирует (изменяется) в форму A. Если p_n — доля генов a в n-м поколении, то $1-p_n$ — доля генов A в n-м поколении. Тогда $p_{n+1}=p_n(1-\mu)=p_{n-1}(1-\mu)^2$ и окончательно $p_n=p_0(1-\mu)^n$, т. е. если $0<\mu<1$, то $p_n\to 0$ по геометрической прогрессии со знаменателем $1-\mu$.

2. Пусть теперь между двумя последующими поколениями наряду с рассмотренной мутацией доля $\nu>0$ генов A мутирует в форму a. Тогда

$$p_{n+1}=p_n(1-\mu)+\nu(1-p_n)=p_n(1-\mu-\nu)+\nu,$$

что эквивалентно соотношению

$$p_{n+1}-\frac{\nu}{\mu+\nu}=\left(p_n-\frac{\nu}{\mu+\nu}\right)(1-\mu-\nu).$$

По индукции получаем

$$p_{n+1}-\frac{\nu}{\mu+\nu}=\left(p_{n-1}-\frac{\nu}{\mu+\nu}\right)(1-\mu-\nu)^2=\ldots=\left(p_0-\frac{\nu}{\mu+\nu}\right)(1-\mu-\nu)^n.$$

Снова сходимость имеет геометрический характер. Заметим, что если $\mu+\nu>1$, $1-\mu-\nu<0$, то p_n будет сходиться к пределу немонотонно, если же $\mu+\nu\leqslant 1$, то сходимость монотонная[1]).

[1]) Заметим, что поскольку $\mu,\nu<1$, то в любом случае $|1-\mu-\nu|<1$ и, следовательно, сходимость действительно имеет место: $p_n\xrightarrow[n\to\infty]{}\frac{\nu}{\mu+\nu}$. — *Прим. перев.*

В. Миграция

Предположим, что мы имеем 2 популяции, составленные из генов A и a, — внутреннюю популяцию, в которой частоты этих генов равны p_n и q_n соответственно (и их мы хотим найти), и внешнюю «большую» популяцию, в которой гены A и a имеют постоянные частоты P и Q. «Миграция» означает, что доля I всей внутренней популяции за время одного поколения замещается таким же количеством индивидуумов, взятых из внешней популяции.

Тогда $p_{n+1} = p_n(1 - I) + PI$. Поскольку величина P постоянна, полученное рекуррентное соотношение совпадает, очевидно, с описывающим процесс мутации, если взять $\nu = PI$, $\mu = (1 - P)I$. Отсюда следует, что p_n сходятся к величине $\frac{\nu}{\mu + \nu} = P$.

Г. Отбор

Существует 2 вида отбора — *гаметический*, который отражает относительные преимущества некоторых *гамет* над другими, и *зиготный*, который отражает преимущества некоторых *генотипов* над другими. Такие преимущества могут соответствовать различным смертностям, относительной плодовитости и т. д.

(а) *Гаметический отбор*

Пусть 1 и $1 - s$ $(0 < s < 1)$ — меры относительного селективного преимущества гаметы A по сравнению с гаметой a. Пусть p_n и q_n означают их частоты в популяции в n-м поколении. Эти данные сведены в таблицу II.

Таблица II

	Гамета A	Гамета a	
Частота	p_n	q_n	$p_n + q_n = 1$
Относительное преимущество	1	$1 - s$	

Сформируем следующее поколение гамет A и a, приписывая имеющимся частотам гамет веса 1 и $1 - s$ соответственно. Получим

$$p_{n+1} = p_n [p_n + (1 - s) q_n]^{-1}; \quad q_{n+1} = (1 - s) q_n [p_n + (1 - s) q_n]^{-1}.$$

Сомножитель $[p_n + (1 - s) q_n]^{-1}$ является нормирующим. Имеем

$$p_{n+1} = \frac{p_n}{1 - q_n s}; \quad q_{n+1} = \frac{(1 - s) q_n}{1 - q_n s}.$$

Отсюда

$$\frac{q_{n+1}}{p_{n+1}} = (1-s)\frac{q_n}{p_n} = \ldots = (1-s)^{n+1}\frac{q_0}{p_0} \to 0 \quad \text{при} \quad n \to \infty.$$

Следовательно, $q_n \to 0$ и $p_n \to 1$, т. е. популяция генов a вырождается.

(б) *Зиготный отбор. Общий случай*

Рассмотрим диплоидную популяцию, в которой каждый генотип характеризуется индексом σ, отражающим его преимущество или недостаток по сравнению с другими типами. Предположим, что относительное преимущество не зависит от времени и частот гамет в n-м поколении.

Таблица III

	AA	Aa	aa
Частота в n-м поколении	p_n^2	$2p_nq_n$	q_n^2
Относительное преимущество	σ_1	σ_2	σ_3

Пусть p_n и q_n — частоты A и a соответственно в n-м поколении (см. таблицу III).

Зиготные частоты AA, Aa и aa в $(n+1)$-м поколении равны соответственно

$$\frac{\sigma_1 p_n^2}{w}, \quad \frac{2\sigma_2 p_n q_n}{w}, \quad \frac{\sigma_3 q_n^2}{w},$$

где величина $w = \sigma_1 p_n^2 + 2\sigma_2 p_n q_n + \sigma_3 q_n^2$ называется *функцией приспособленности* и w^{-1} играет роль нормирующего множителя. Следовательно, частоты гамет A и a равны соответственно

$$p_{n+1} = \frac{\sigma_1 p_n^2 + \sigma_2 p_n q_n}{w}, \quad q_{n+1} = \frac{\sigma_3 q_n^2 + \sigma_2 p_n q_n}{w}. \tag{1.1}$$

Таким образом,

$$\Delta p_n = p_{n+1} - p_n = p_n q_n \frac{(\sigma_1 - \sigma_2)\, p_n + (\sigma_2 - \sigma_3)\, q_n}{w}. \tag{1.2}$$

Равновесной частотой является любая величина $\tilde{p}$ $(0 \leqslant \tilde{p} \leqslant 1)$, такая, что из равенства $p_n = \tilde{p}$ следует $p_{n+1} = \tilde{p}$. Она называется устойчивой, если для значений p_n, близких к $\tilde{p}$, переменная p_{n+k} при $k \to \infty$ стремится к $\tilde{p}$. Равновесная частота, не являющаяся устойчивой, называется неустойчивой. Определим *полиморфизм* как такое устойчивое равновесное состояние, в котором представлен каждый генотип. Такая ситуация имеет место, когда гетерозиготный индивидуум имеет селективное преимущество перед каждым гомозиготным индивидуумом, как, например, при наследовании признака серповидноклеточной анемии.

Пример полиморфизма: $\sigma_2 > \sigma_1$, $\sigma_2 > \sigma_3$. Такая ситуация обычно называется гетерозиготным преимуществом, или доминированием. Рассмотрим случаи, когда Δp_n равна нулю.

(а) $p_n = 1$. Это равновесное состояние. Однако если p_n близко к 1, то $\Delta p_n < 0$ и переменная p_n удаляется от 1. Это неустойчивое равновесие.

(б) $p_n = 0$. Это состояние также неустойчивое, поскольку $\Delta p_n > 0$ при малых p_n.

(в) $\frac{p_n}{q_n} = \frac{(\sigma_2 - \sigma_3)}{(\sigma_2 - \sigma_1)}$, или, что то же, $p_n = p^*$, где

$$p^* = \frac{\sigma_2 - \sigma_3}{2\sigma_2 - \sigma_1 - \sigma_3}, \qquad 0 < p^* < 1.$$

Пусть $p = p^* + \delta$. Тогда $\Delta p \simeq [(p^*q^*)/(w(p^*))](\sigma_1 + \sigma_3 - 2\sigma_2)\delta$ имеет знак, обратный знаку δ, так как $\sigma_1 + \sigma_3 - 2\sigma_2 < 0$. Отсюда следует, что положение равновесия устойчивое. Чуть более подробный анализ показывает, что процесс (1.1) сходится к равновесному состоянию p^* при любом начальном состоянии p, таком, что $p(1-p) \neq 0$.

Д. Модели двуполых популяций

Рассмотрим теперь популяцию, состоящую из мужских и женских особей, спаривание при этом происходит только между представителями различных полов. Если ограничиться рассмотрением лишь генов, локализованных на аутосомных (неполовых) хромосомах (как это было в предыдущем случае), то процесс случайного спаривания можно описать точно таким же образом, как и выше. Предположим, что частоты генотипов внутри мужской и женской популяций равны:

Мужская популяция			Женская популяция		
AA	Aa	aa	AA	Aa	aa
r_1	$2s_1$	t_1	u_1	$2v_1$	w_1

Гаметы имеют следующие частоты:

Мужская популяция		Женская популяция	
A	a	A	a
$p_1 = r_1 + s_1$	$q_1 = s_1 + t_1$	$\bar{p}_1 = u_1 + v_1$	$\bar{q}_1 = v_1 + w_1$

Для случайного спаривания, которое эквивалентно случайному выбору из мужских и женских совокупностей гамет, мы найдем, что частоты генотипов в *обеих* мужской и женской популяциях во втором поколении равны $p_1\bar{p}_1$, $p_1\bar{q}_1 + \bar{p}_1 q_1$, $q_1\bar{q}_1$ для AA, Aa и aa соответственно. Как и в случае самооплодотворения, устойчивое распределение достигается за одно поколение.

Однако, если рассмотреть половые хромосомы, ситуация резко изменится. Причина этого состоит в том, что в некоторых диплоидных клетках могут находиться два различных типа хромосом X и Y. У млекопитающих женские особи характеризуются наличием двух гомологичных [1]) X-хромосом, в то время как у мужских особей спарены хромосомы X и Y. Более того, многие, если не все, гены из X-хромосомы не имеют комплементов на Y-хромосоме. В рассматриваемом случае генотип мужской особи полностью определяется аллелями, локализованными на X-хромосоме. Такие гены и контролируемые ими признаки называются *сцепленными с полом.*

Рассмотрим распределение сцепленного с полом гена в последовательных поколениях. Предположим, что начальные частоты генотипов у мужских и женских особей равны

Женские особи (XX)			Мужские особи (XY)	
AA	Aa	aa	A	a
r_0	$2s_0$	t_0	p_0	q_0

Заметим, что генотип у мужской особи определяется только одним символом гена, поскольку Y-хромосома не содержит ни одного гена, содержащегося в X-хромосоме.

Частоты гамет у женских особей равны

$$\begin{array}{cc} A & a \\ r_0 + s_0 & s_0 + t_0 \end{array}$$

В предположении о случайном спаривании, или, что то же, о случайном выборе из совокупности гамет, частоты генотипов в следующем поколении равны у женских особей:

$$\begin{array}{ccc} AA & Aa & aa \\ r_1 = (r_0 + s_0)\, p_0 & 2s_1 = (r_0 + s_0)\, q_0 + (s_0 + t_0)\, p_0 & t_1 = (s_0 + t_0)\, q_0 \end{array}$$

и у мужских особей

$$\begin{array}{cc} A & a \\ p_1 = r_0 + s_0 & q_1 = s_0 + t_0 \end{array}$$

Каждая вновь образованная мужская зигота получает Y-хромосому из мужской популяции и X-хромосому из женской. Следовательно, доля генов A в мужской популяции в точности равна доле генов A в предыдущем поколении женской популяции.

[1]) Гомологичными называются хромосомы, образующие пару (в диплоидном случае). Все неполовые гомологичные хромосомы тождественны. — *Прим. перев.*

В общем случае если p_n — частота аллеля A в совокупности женских гамет в n-м поколении, то p_{n-1} — частота аллеля A в мужской популяции в n-м поколении. Следовательно, $(n+1)$-е поколение женских генотипов имеет состав

$$\begin{array}{ccc} AA & Aa & aa \\ p_n p_{n-1} & p_n q_{n-1} + q_n p_{n-1} & q_n q_{n-1} \end{array}$$

Отсюда можно вывести рекуррентное соотношение

$$p_{n+1} = p_n p_{n-1} + \frac{1}{2}(p_n q_{n-1} + q_n p_{n-1}),$$

которое сводится к следующему:

$$2[p_{n+1} - p_n] = -(p_n - p_{n-1}),$$

или $2\Delta_n = -\Delta_{n-1}$, где $\Delta_n = p_{n+1} - p_n$. По индукции получаем

$$\Delta_n = \left(-\frac{1}{2}\right)^n \Delta_0,$$

или

$$p_{n+1} = p_n + \left(-\frac{1}{2}\right)^n \Delta_0.$$

Отсюда

$$p_{n+1} = \left(-\frac{1}{2}\right)^n \Delta_0 + \left(-\frac{1}{2}\right)^{n-1} \Delta_0 + \ldots + \Delta_0 + p_0 = p_0 + \Delta_0 \left(\frac{1 - \left(-\frac{1}{2}\right)^{n+1}}{3/2}\right).$$

В пределе

$$\alpha = \lim p_n = p_0 + \frac{2}{3}(p_1 - p_0) = \frac{2}{3} p_1 + \frac{1}{3} p_0;$$

получили выражение для предельной частоты генов A в мужской популяции. В женской популяции предельные частоты генотипов имеют вид

$$\begin{array}{ccc} AA & Aa & aa \\ \alpha^2 & 2\alpha(1-\alpha) & (1-\alpha)^2 \end{array}$$

Это, конечно, в точности соответствует частотам Харди—Вейнберга.

§ 2. ИНБРИДИНГ[1]

Существует два типа неслучайного спаривания. Первый из них, называемый *ассортативным спариванием*, подразумевает обязательность спаривания подобных (или неподобных) генотипов (или фе-

---

[1]) Инбридингом в биологии называется спаривание двух близкородственных индивидуумов. — *Прим. перев.*

нотипов). Очевидным примером является спаривание генотипов XX и XY при половом размножении. Вторым типом неслучайного спаривания является спаривание близкородственных индивидуумов.

А. Самооплодотворение

Первая модель инбридинга, которую мы рассмотрим, относится к самооплодотворению. Каждый индивидуум при этом обладает обеими гаметами. Допускаются только спаривания типов $AA \times \times AA$, $aa \times aa$, $Aa \times Aa$, т. е. спариваться могут только индивидуумы, имеющие один и тот же генотип. Например, у цветов пыльник выделяет пыльцу (мужские гаметы), а в пестике развиваются семяпочки (женские гаметы) и при этом часто происходит своего рода инбридинг.

Если начальные частоты зигот AA, Aa и aa равны соответственно u_0, v_0 и w_0, то аналогичные частоты для следующего поколения равны

$$u_1 = u_0 + \frac{1}{4} v_0, \quad v_1 = \frac{1}{2} v_0, \quad w_1 = w_0 + \frac{1}{4} v_0,$$

поскольку при спаривании $Aa \times Aa$ производятся потомки в отношении $\frac{1}{4} AA$, $\frac{1}{2} Aa$, $\frac{1}{4} aa$. С каждым последующим поколением частота гетерозигот Aa уменьшается в 2 раза. В n-м поколении частоты равны

$$u_n = u_0 + \frac{1}{4} v_0 \left[1 + \frac{1}{2} + \ldots + \left(\frac{1}{2}\right)^{n-1}\right] = u_0 + \frac{1}{2} v_0 \left[1 - \left(\frac{1}{2}\right)^n\right],$$

$$v_n = \left(\frac{1}{2}\right)^n v_0,$$

$$w_n = w_0 + \frac{1}{4} v_0 \left[1 + \frac{1}{2} + \ldots + \left(\frac{1}{2}\right)^{n-1}\right] = w_0 + \frac{1}{2} v_0 \left[1 - \left(\frac{1}{2}\right)^n\right].$$

Предельные частоты при $n \to \infty$ равны: $u = u_0 + \frac{v_0}{2}$, $v = 0$, $w = w_0 + \frac{v_0}{2}$.

Б. Спаривание сибсов

В следующей модели рассматривается спаривание потомков одних и тех же родителей[1]), или спаривание сибсов. Опять целью является определение предельных частот различных генотипов, участвующих в спаривании. Однако вместо непосредственного рассмотрения частот генотипов мы проследим за флуктуациями частот спаривающихся особей в последовательных поколениях.

Если ограничиться рассмотрением двух аллелей A и a, то существуют три возможных генотипа для каждого родителя и шесть

[1]) В оригинале «brother — sister mating». — *Прим. перев.*

возможных вариантов сочетания, или скрещивания генотипов, т. е. существует шесть различных вариантов спаривания: $AA \times AA$, $aa \times aa$, $AA \times Aa$, $aa \times Aa$, $Aa \times Aa$, $AA \times aa$. Предположим, что частоты каждого из этих вариантов известны. Мы хотим вычислить эти частоты для следующего поколения. Для того чтобы сделать это, мы сейчас найдем распределение генотипов потомков, а затем, предполагая, что спариваться могут только сибсы, определим частоту получаемых при этом различных видов спаривания потомков.

Например, все дети, получаемые при спаривании $AA \times AA$, имеют тип AA, и поэтому среди сибсов допустимым является лишь спаривание вида $AA \times AA$. При спаривании $AA \times Aa$ производятся потомки в отношении $\frac{1}{2} AA$, $\frac{1}{2} Aa$. Частоты различных вариантов спаривания среди сибсов, очевидно, равны $\frac{1}{4} AA \times AA$, $\frac{1}{2} AA \times Aa$, $\frac{1}{4} Aa \times Aa$. Аналогичным образом можно рассмотреть и другие возможности. Так, при спаривании $Aa \times Aa$ потомки производятся в отношении $\frac{1}{4} AA$, $\frac{1}{2} Aa$, $\frac{1}{4} aa$. Частоты различных вариантов спаривания среди сибсов можно легко вычислить. Они равны

$$\frac{1}{16} AA \times AA, \quad \frac{1}{16} aa \times aa, \quad \frac{1}{4} Aa \times Aa,$$
$$\frac{1}{8} AA \times aa, \quad \frac{1}{4} AA \times Aa, \quad \frac{1}{4} aa \times Aa.$$

Другие сочетания исследуются аналогичным образом.

Можно проанализировать изменения частот спаривающихся типов от поколения к поколению с помощью цепей Маркова. Для этого рассмотрим цепь Маркова с шестью состояниями, соответствующими шести вариантам спаривания:

состояние 1 — $AA \times AA$, состояние 2 — $AA \times Aa$, состояние 3 — $AA \times aa$, состояние 4 — $Aa \times Aa$, состояние 5 — $Aa \times aa$, состояние 6 — $aa \times aa$.

Эта цепь Маркова имеет следующую матрицу перехода:

$$\mathbf{A} = \begin{Vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{16} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{16} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}.$$

В строках матрицы записаны частоты различных вариантов спаривания среди сибсов, родившихся в результате соответствующего номеру строки варианта спаривания. Имеются два поглощающих состояния: 1 и 6.

Из теории марковских цепей известно, что в конце концов произойдет поглощение в одном из этих двух состояний. Вероятность поглощения из переходных состояний (состояний 2—5) представляет интерес, поскольку она, очевидно, соответствует той интенсивности, с которой исчезают гетерозиготные генотипы. Эта интенсивность равна, конечно, наибольшему собственному значению матрицы

$$\mathbf{B} = \left\| \begin{matrix} \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{matrix} \right\|,$$

которая получится при исключении первого и шестого состояний. Поскольку последствия вариантов спаривания $AA \times Aa$ и $aa \times Aa$ симметричны, можно дóлжным образом объединить первую и четвертую строки матрицы $\mathbf{B}$. При этом получим матрицу

$$\mathbf{C} = \left\| \begin{matrix} \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{8} & \frac{1}{4} \end{matrix} \right\|.$$

Собственные значения $\mathbf{C}$ являются корнями уравнения

$$\left(\frac{1}{2} - \lambda\right)\left[\lambda^2 - \lambda\frac{1}{4} - \frac{1}{8}\right] + \frac{1}{8}\lambda = 0,$$

а именно

$$\lambda_1 = \frac{1}{4}, \quad \lambda_2 = \frac{1 - \sqrt{5}}{4}, \quad \lambda_3 = \frac{1 + \sqrt{5}}{4}.$$

Следовательно, скорость приближения к гомозиготному состоянию (фиксации) равна

$$\lambda_3 = \frac{1}{4}(1 + \sqrt{5}),$$

т. е. с каждым поколением частота вариантов спаривания различных генотипов уменьшается в λ_3 раз.

Теперь не представляет труда вычислить при любых начальных частотах $(u_1, u_2, \ldots, u_6)$ вариантов спаривания финальную частоту поглощения, или фиксации в состояниях 1 ($AA \times AA$) или 6 ($aa \times \times aa$). Точно так же мы определим частоту фиксации двух гомозиготных генотипов.

Вероятность x_i фиксации в состоянии 1 (спаривание $AA \times AA$) при начальном состоянии i, $i = 1, 2, \ldots, 6$, является решением уравнения

$$\sum_{j=1}^{6} a_{ij} x_j = x_i,$$

где полагается $x_1 = 1$, $x_6 = 0$ и $A = \|a_{ij}\|$. Нетрудно найти, что

$$x_2 = \frac{3}{4}, \quad x_3 = \frac{1}{2}, \quad x_4 = \frac{1}{2}, \quad x_5 = \frac{1}{4}.$$

Если начальные частоты различных состояний равны $(u_1, u_2, \ldots, u_6)$, то вероятность фиксации в состоянии 1, очевидно, равна

$$\sum_{i=1}^{6} u_i x_i = u_1 + \frac{3}{4} u_2 + \frac{1}{2} u_3 + \frac{1}{2} u_4 + \frac{1}{4} u_5.$$

В. Гены, идентичные по происхождению

Еще одно определение степени инбридинга в популяции дается через понятие коэффициента инбридинга. *Коэффициент инбридинга* индивидуума X определяется как вероятность того, что два гена в одном локусе у индивидуума X *идентичны по происхождению.* Под этим мы понимаем, что два гена являются «копиями» одного и того же гена некоторого индивидуума из предшествующего поколения.

Рассмотрим теперь конечную популяцию, состоящую из N индивидуумов, воспроизводящихся бесполым путем, т. е. каждый индивидуум «поставляет» бесконечное число копий своих двух генов в совокупность гамет, в которой затем происходит случайное спаривание с целью образования новых N индивидуумов[1]). Пусть F_n — коэффициент инбридинга в n-м поколении популяции, т. е. F_n — вероятность того, что два гена, выбранные случайным образом, идентичны по происхождению.

В совокупности гамет происходит их спаривание с целью образования N новых индивидуумов. Предположим, что выбрана некоторая гамета и нужно (случайным образом) выбрать вторую гамету с тем, чтобы спарить с первой. Вероятность того, что ген, содержащийся во второй гамете, будет копией того же самого гена, что и в первой, равна $\frac{1}{2N}$, так как каждый из $2N$ типов гамет, произведенных N индивидуумами, равновероятен. При этом два спа-

[1]) Точнее, каждый индивидуум поставляет *одно и то же число* генов k и рассматривается предельный случай $k \to \infty$. При этом понятие частоты не теряет смысла. — *Прим. перев.*

ренных гена являются, конечно, идентичными по происхождению. Теперь, если два спаренных гена не являются копиями одного и того же гена (вероятность чего равна $1-\frac{1}{2N}$), вероятность того, что они идентичны по происхождению, равна F_n. Следовательно,

$$F_{n+1}=\frac{1}{2N}+\frac{2N-1}{2N}F_n,$$

$$H_n=1-F_n=\left(1-\frac{1}{2N}\right)H_{n-1}=\left(1-\frac{1}{2N}\right)^2 H_{n-2}=\ \dots\ =\left(1-\frac{1}{2N}\right)^n H_0.$$

Таким образом, вероятность того, что некоторый индивидуум в n-м поколении — гетерозигот, стремится к 0 при $n\to\infty$ со скоростью $\left(1-\frac{1}{2N}\right)^n$. Поскольку было сделано предположение о большом размере популяции, эта скорость весьма низка. Значение проведенного анализа состоит в получении качественного результата о том, что конечность популяции влечет за собой (в какой-то мере) инбридинг.

Г. Бесконечная диплоидная популяция и коэффициент инбридинга

Предположим, что имеется бесконечная диплоидная популяция, у которой в одном локусе ген A может находиться с частотой p, а ген a — с частотой $q=1-p$. Предположим, что популяция имеет постоянный коэффициент инбридинга F, который имеет тот же смысл, что и в предыдущем параграфе. Доля F популяции состоит из гомозигот, у которых гены идентичны по происхождению. Очевидно, доли гомозигот AA и aa с генами, идентичными по происхождению, равны соответственно Fp и Fq. Доля всех зигот, в которых гены не идентичны по происхождению, равна $1-F$. Следовательно, доли типов AA, Aa и aa соответственно равны $(1-F)p^2$, $(1-F)2pq$ и $(1-F)q^2$. Поэтому если u, v и w — частоты генотипов AA, Aa и aa соответственно, то

$$\begin{aligned}u&=(1-F)p^2+Fp=p^2+Fpq,\\ v&=(1-F)\cdot 2pq,\\ w&=(1-F)q^2+Fq=q^2+Fpq.\end{aligned}\tag{2.1}$$

Очевидно, $u+v+w=1$, как и следовало ожидать. Эти частоты являются модификациями частот Харди — Вейнберга. Заметим, что частота гена A равна $u+\frac{1}{2}v=p^2+pq=p$ и остается постоянной, если $F=\mathrm{const}$. Значение коэффициента F становится ясным, если заметить, что при $F=0$ u, v, w являются частотами Харди — Вейнберга, описывающими популяцию без инбридинга, в то время как при $F=1$ частота гетерозиготных индивидуумов равна нулю, т. е. популяция достигает состояния максимального инбридинга.

Д. Популяция с самооплодотворением и случайным спариванием

Предположим, что доля s популяции воспроизводится самооплодотворением, в то время как в другой ее части имеет место случайное спаривание. Соответствующие схемы размножения показаны в табл. IV, где $p = u + \frac{1}{2}v$, $q = 1 - p = w + \frac{1}{2}v$.

Таблица IV

	AA	Aa	aa
Частоты	u	v	w
Потомство от самооплодотворения	$su\,(AA)$	$sv\left(\frac{1}{4}AA + \frac{1}{2}Aa + \frac{1}{4}aa\right)$	$sw\,(aa)$
Потомство от случайного спаривания	$(1-s)\,p^2\,(AA)$	$(1-s)\,2pq\,(Aa)$	$(1-s)\,q^2\,(aa)$

Частоты генотипов во втором поколении равны

$$u' = (1-s)\,p^2 + su + \frac{1}{4}sv,$$

$$v' = 2(1-s)\,pq + \frac{1}{2}sv,$$

$$w' = (1-s)\,q^2 + sw + \frac{1}{4}sv.$$

Кроме того,

$$p' = u' + \frac{1}{2}v' = (1-s)\,p + s\left(u + \frac{1}{2}v\right) = (1-s)\,p + sp = p.$$

Таким образом, p остается постоянным от поколения к поколению, как и следовало, конечно, ожидать, поскольку мы не рассматриваем какие-либо факторы, могущие влиять на частоту гена.

Следовательно, можно записать

$$v_{n+1} = 2(1-s)\,pq + \frac{1}{2}sv_n =$$

$$= 2(1-s)\,pq\left(1 + \frac{1}{2}s + \left(\frac{1}{2}s\right)^2 + \ldots + \left(\frac{1}{2}s\right)^n\right) + \left(\frac{1}{2}s\right)^{n+1} v_0 =$$

$$= \frac{4(1-s)}{2-s}\,pq\left[1 - \left(\frac{1}{2}s\right)^{n+1}\right] + \left(\frac{1}{2}s\right)^{n+1} v_0,$$

$$u_n = (1-s)\,p^2 + s\left(u_{n-1} + \frac{1}{4}\,v_{n-1}\right) = \left(\text{поскольку} \quad u_{n-1} + \frac{1}{2}\,v_{n-1} = p\right)$$
$$= (1-s)\,p^2 + s\left(p - \frac{1}{4}\,v_{n-1}\right) = p^2 + pqs + \frac{1}{4}\,sv_{n-1} =$$
$$= p^2 + pqs\,\frac{1}{4}\,s\,\frac{4\,(1-s)}{2-s}\,pq\left[1 - \left(\frac{1}{2}\,s\right)^{n-1}\right] - \frac{1}{4}\,s\left(\frac{1}{2}\,s\right)^{n-1} v_0 =$$
$$= p^2 + pqs\left[1 - \frac{1-s}{2-s}\left(1 - \left(\frac{1}{2}\,s\right)^{n-1}\right)\right] - \frac{1}{4}\,s\left(\frac{1}{2}\,s\right)^{n-1} v_0.$$

Следовательно, предельное значение u при $n \to \infty$ равно

$$p^2 + pqs\,\frac{1}{2-s} = p^2 + pq\,\frac{s}{2-s}\,.$$

Аналогично предельные значения v и w равны соответственно $2pq\left(1 - \frac{s}{2-s}\right)$ и $q^2 + pq\left(\frac{s}{1-s}\right)$. Таким образом, обращаясь к (2.1), мы видим, что в равновесном состоянии F имеет значение $\frac{s}{2-s}$. Чем больше s, тем медленнее достигается равновесное состояние, и наоборот.

В частности, при $s = 0$, т. е. при полностью случайном спаривании, коэффициент F равен нулю и инбридинга нет. Если $s = 1$ или происходит только самооплодотворение, то $F = 1$ и в равновесном состоянии популяция полностью гомозиготна.

§ 3. ПОЛИПЛОИДЫ

Многие организмы, в особенности растения, являются полиплоидами, т. е. гомологичные хромосомы встречаются группами размера более чем 2. Рассмотрим $2n$-плоидную популяцию (т. е. наборы гомологичных хромосом имеют размер $2n$), у которой в каждом рассматриваемом локусе имеется лишь два аллеля. Будем обозначать комбинации зигот (генотипы) через $A^s a^{2n-s}$ $(s = 0, 1, \ldots, 2n)$, где s — число хромосом, несущих ген A, а остальные $2n - s$ хромосом несут ген a. Возможные гаметические комбинации имеют вид

$$A^r a^{n-r}, \qquad r = 0,\ 1,\ \ldots,\ n.$$

Расщепление гамет в полиплоиде более сложное, нежели в случае диплоидных организмов. Это объясняется тем фактом, что в течение мейотического цикла происходит процесс удвоения, когда каждая хромосома расщепляется вдоль своей длины на две одинаковые хроматиды, которые, однако, остаются соединенными в пары с помощью центрального стерженька — центромера. Между хроматидами гомологичных хромосом имеет место некоторый обмен генетическим материалом. (Это называют явлением *кроссинговера*, или

14 Зак. 939

перекреста.) Природа расщепления гамет зависит от расстояния рассматриваемого локуса от центромера. Два наиболее распространенных способа расщепления — *хромосомное и хроматидное.*

В случае хромосомного расщепления родитель $A^s a^{2n-s}$ будет производить гаметы $A^r a^{n-r}$ с частотой

$$\frac{1}{\binom{2n}{n}}\binom{s}{r}\binom{2n-s}{n-r}, \qquad r=0,\ 1,\ \ldots,\ s,$$

а в случае хроматидного расщепления — с частотой

$$\frac{1}{\binom{4n}{n}}\binom{2s}{r}\binom{4n-2s}{n-r}, \qquad r=0,\ 1,\ \ldots,\ s.$$

В случае $n=1$ эти выражения совпадают.

Получим аналог закона Харди — Вейнберга для локуса при хромосомном расщеплении. Если в популяции из $2n+1$ гамет, в которой частота генов A равна p, а частота генов a равна q, происходит случайное слияние гамет, то частота особей $A^s a^{2n-s}$ в следующем поколении равна $\binom{2n}{s} p^s q^{2n-s}$. Соответствующие гаметы $A^r a^{n-r}$ имеют частоту

$$\sum_{s=0}^{2n}\binom{s}{r}\binom{2n-s}{n-r}\binom{2n}{n}^{-1}\binom{2n}{s}p^s q^{2n-s}=$$

$$=\sum_s \frac{s!}{r!\,(s-r)!}\,\frac{(2n-s)!}{(n-r)!\,(n+r-s)!}\,\frac{(n!)^2}{2n!}\,\frac{2n!}{(2n-s)!\,s!}\,p^s q^{2n-s}=$$

$$=\sum_s \frac{n!}{r!\,(n-r)!}\,p^r q^{n-r}\,\frac{n!}{(s-r)!\,(n-s+r)!}\,p^{s-r}q^{n-s+r}=$$

$$=\binom{n}{r}p^r q^{n-r}\sum_{s=r}^{n+r}\binom{n}{s-r}p^{s-r}q^{n-s+r}=$$

$$=\binom{n}{r}p^r q^{n-r}(p+q)^n=\binom{n}{r}p^r q^{n-r}.$$

Таким образом, ген A, например, имеет частоту

$$\sum_r \frac{r}{n}\binom{n}{r}p^r q^{n-r}=p.$$

Можно сформулировать следующее утверждение о сходимости: пусть $p_k(A^r a^{n-r})$, $r=0,\ 1,\ \ldots,\ n,$ — частота гаметы $A^r a^{n-r}$ в k-м поколении. Рассмотрим случайное спаривание и хромосомное расщепление. Тогда

$$\lim_{k\to\infty} p_k\left(A^r a^{n-r}\right) = \binom{n}{r} p^r (1-p)^{n-r},$$

где p — начальная частота аллеля A в популяции. Сходимость имеет скорость геометрической прогрессии со знаменателем $\frac{n-1}{2n-1}$, т. е.

$$\left| p_k\left(A^r a^{n-r}\right) - \binom{n}{r} p^r (1-p)^{n-r} \right| \leqslant C\lambda^k, \qquad r = 1,\ 2,\ \ldots,\ n,$$

для всех k, где $\lambda = (n-1)/(2n-1)$.

Мы не будем доказывать эту теорему. Ее можно пояснить с помощью следующих эвристических рассуждений:

Пусть $B_r = \binom{n}{r} p^r (1-p)^{n-r}$. Для образования гаметы выберем n хромосом из совокупности размера $2n$. Если n хромосом взять от различных предков, то имеет место хромосомное расщепление нулевого поколения популяции, и рассматриваемые частоты являются биномиальными, т. е. $p_k\left(A^r a^{n-r}\right) = B_r$. Мы определенно ожидаем, что при больших n это соответствует действительности.

Но предположим теперь, что $n-2$ хромосомы получены от различных предков, а две — от одного и того же предка (другие возможности можно не учитывать, поскольку они гораздо менее вероятны).

Пусть $1-\rho^{(k)}$ — вероятность того, что все n генов в гамете k-го поколения получены от различных предков нулевого поколения.

Тогда $p_k\left(A^r a^{n-r}\right) = \left(1-\rho^{(k)}\right) B_r + \rho^{(k)} C_r$, где C_r не зависит от n, а зависит лишь от числа «родственных» хромосом в гамете. Следовательно,

$$p_k\left(A^r a^{n-r}\right) = B_r + \rho^{(k)}\left(C_r - B_r\right).$$

Для того чтобы получить рекуррентное соотношение относительно $\rho^{(k+1)}$, заметим, что $\rho^{(k+1)}$ зависит от вероятности спаривания между гаметой, имеющей хромосомы, пришедшие от различных предков, и гаметой, у которой имеется хотя бы две «родственные» хромосомы. Но мы можем распорядиться произвольным образом лишь $(n-2)$ хромосомами — две «родственные» должны остаться в паре. Мы пренебрегаем случаем, когда имеется большее число «родственных» хромосом. Тогда

$$\rho^{(k+1)} = 2\rho^{(k)}\left(1-\rho^{(k)}\right) \frac{\binom{2}{2}\binom{2n-2}{n-2}}{\binom{2n}{n}}.$$

Пренебрегая в правой части членом, содержащим $[\rho^{(k)}]^2$, получим [1])

$$\rho^{(k+1)} = 2\rho^{(k)} \frac{\binom{2}{2}\binom{2n-2}{n-2}}{\binom{2n}{n}} = 2\rho^{(k)} \frac{1}{2} \frac{n-1}{2n-1} = \rho^{(k)} \frac{n-1}{2n-1}.$$

Таким образом, $p_k(A^r a^{n-r}) = B_r + \rho^{(k)} [C_r - B_r]$ сходится к B_r, поскольку $\rho^{(k)} \to 0$ геометрически с параметром $(n-1)/(2n-1)$.

§ 4. МАРКОВСКИЕ ПРОЦЕССЫ, ПОРОЖДАЕМЫЕ ПРЯМЫМ ПРОИЗВЕДЕНИЕМ ВЕТВЯЩИХСЯ ПРОЦЕССОВ

Существует тесная связь между теорией ветвящихся процессов и некоторыми вероятностными моделями частот, используемыми в генетике и экологии популяций. Мы введем в рассмотрение класс конечных цепей Маркова специального вида, который включает в себя как частные случаи цепи, рассмотренные в § 2 гл. 2 в качестве примеров приложения к генетике.

Пусть $f(s) = \sum_{n=0}^{\infty} a_n s^n$, $\sum_{n=0}^{\infty} a_n = 1$, $a_n \geqslant 0$. Тогда марковский ветвящийся процесс определяется как цепь Маркова, для которой элемент матрицы переходных вероятностей P_{ij} определяется, как коэффициент при s^j в разложении $f^i(s)$ (см. гл. 11). Вероятностное распределение, определяемое последовательностью $\{a_i\}$, называется распределением числа потомков для одного индивидуума, а $f^i(s)$ — вероятностной производящей функцией для числа независимо порожденных потомков i индивидуумов. Аналогично процесс $\{X_n\}$ является ветвящимся, если $X_{n+1} = \sum_{k=1}^{i} \xi_k = j$, где $X_n = i$ — число индивидуумов в текущем поколении, ξ_k — число потомков k-го индивидуума в текущем поколении, а j — число индивидуумов в последующем поколении. Случайные величины $\xi_1, \xi_2, \ldots, \xi_i$ являются независимо распределенными с общей вероятностной производящей функцией $f(s)$.

---

[1]) Заметим, что поскольку

$$\rho^{(k+1)} \leqslant 2\rho^{(k)} \frac{\binom{2}{2}\binom{2n-2}{n-2}}{\binom{2n}{n}},$$

то действительные значения $\rho^{(k)}$ не превышают решения «линеаризованной» системы. В частности, если решение линеаризованной системы сходится к 0 при $k \to \infty$, то это же справедливо и для исходной системы. — *Прим. перев.*

Рассмотрим теперь две популяции типов A и a, каждая из которых независимым образом размножается, как ветвящийся процесс. Это — двумерный процесс, развертывающийся как последовательность пар случайных величин $Z_n = (X_n; Y_n)$, где X_n и Y_n — соответственно числа типов A и a в n-м поколении. Состояние процесса описывается парой (i, j), $i \geqslant 0$, $j \geqslant 0$, где i и j — текущие значения числа типов A и a соответственно. В такой формулировке компоненты X_n и Y_n образуют независимые ветвящиеся процессы. Поэтому можно назвать Z_n *прямым произведением ветвящихся процессов* $\{X_n\}$ и $\{Y_n\}$.

Пусть $f(z)$ — производящая функция распределения числа потомков индивидуумов типа A, а $g(w)$ — производящая функция распределения числа потомков индивидуумов типа a. Если индивидуумы развиваются независимо друг от друга, то компоненты матрицы переходных вероятностей вычисляются по формуле

$$\mathrm{P}\{X(1)=j_1,\ Y(1)=j_2 \mid X(0)=i_1,\ Y(0)=i_2\} = P_{(i_1,\, i_2),\,(j_1,\, j_2)} =$$
$$= \text{коэффициент при } z^{j_1} w^{j_2} \text{ в разложении } f^{i_1}(z)\, g^{i_2}(w).$$

Можно сформулировать марковскую модель частот, связанную с ветвящимся процессом для двух типов индивидуумов, в которой размер популяции постоянен во времени. Имея это в виду, найдем сначала вероятность того, что в поколении потомков имеется j индивидуумов типа A при условии, что общее число потомков равно M и текущее поколение состоит из i индивидуумов типа A и $N-i$ индивидуумов типа a:

$$\mathrm{P}\{X(1)=j \mid X(0)=i,\ Y(0)=N-i,\ X(1)+Y(1)=M\} =$$
$$= \frac{\mathrm{P}\{X(1)=j,\ Y(1)=M-j \mid X(0)=i,\ Y(0)=N-i\}}{\mathrm{P}\{X(1)+Y(1)=M \mid X(0)=i,\ Y(0)=N-i\}} =$$
$$= \frac{\text{коэффициент при } z^j w^{M-j} \text{ в разложении } f^i(z)\, g^{N-i}(w)}{\text{коэффициент при } t^M \text{ в разложении } f^i(t)\, g^{N-i}(t)}. \tag{4.1}$$

Здесь использован тот факт, что $f^i(t)\, g^{N-i}(t)$ — производящая функция общего числа потомков, не различаемых по типам. Если $N = M$, то можно считать выражение (4.1) определяющим матрицу переходных вероятностей $P_{ij}(N)$ конечной цепи Маркова с пространством состояний $\{0, 1, 2, \ldots, N\}$.

Отношение номера состояния к числу N можно интерпретировать как частоту типа A во всей популяции. Построенная цепь Маркова, следовательно, является дискретной вероятностной моделью, которая описывает изменения частот типов в популяции постоянного размера, состоящей из двух типов индивидуумов. Изменения частоты в каждом из поколений определяются видом матрицы переходных вероятностей цепи Маркова.

Размер популяции N является параметром, который часто для упрощения обозначений будет опускаться. Таким образом,

$$P_{ij} = \mathrm{P}\{X(1) = j \mid X(1) + Y(1) = N,\ X(0) = i,\ Y(0) = N - i\} =$$
$$= \frac{\text{коэффициент при } z^j w^{N-j} \text{ в разложении } f^i(z)\, g^{N-i}(w)}{\text{коэффициент при } t^N \text{ в разложении } f^i(t)\, g^{N-i}(t)}. \tag{4.2}$$

Марковскую цепь, характеризуемую матрицей переходных вероятностей (4.2), иногда называют цепью Маркова, порожденной прямым произведением марковских ветвящихся процессов.

Мы проиллюстрируем возможности данного класса цепей Маркова на трех простейших примерах и укажем их место в прикладных задачах.

(1) В качестве первого примера рассмотрим прямое произведение пуассоновских ветвящихся процессов, когда

$$f(z) = e^{\lambda(z-1)}, \quad g(w) = e^{\mu(w-1)}, \quad \lambda,\ \mu > 0.$$

При этом

$$P_{ij} = \frac{\text{коэффициент при } z^j \text{ в разложении } e^{i\lambda(z-1)}}{\text{коэффициент при } t^N \text{ в разложении } e^{(i\lambda + (N-i)\mu)(t-1)}} \times$$
$$\times \left[\text{коэффициент при } w^{N-j} \text{ в разложении } e^{(N-i)\mu(w-1)}\right] =$$
$$= \frac{e^{-i\lambda}\left[\dfrac{(i\lambda)^j}{j!}\right]\left\{e^{-(N-i)\mu}\dfrac{1}{(N-j)!}[(N-i)\mu]^{N-j}\right\}}{e^{-(i\lambda+(N-i)\mu)}\left\{[i\lambda + (N-i)\mu]^N \dfrac{1}{N!}\right\}} =$$
$$= \binom{N}{j}\left[\frac{i\lambda}{i\lambda + (N-i)\mu}\right]^j \left[\frac{(N-i)\mu}{i\lambda + (N-i)\mu}\right]^{N-j}, \qquad i,\ j = 0,\ 1,\ 2,\ \ldots,\ N. \tag{4.3}$$

Если $\lambda = \mu$, то получаем выражение

$$P_{ij} = \binom{N}{j}\left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}, \qquad i,\ j = 0,\ 1,\ \ldots,\ N, \tag{*4.4}$$

не зависящее от λ. Это в точности совпадает с вероятностной моделью биномиального выбора (пример Ж § 2 гл. 2), предложенной С. Райтом для изучения флуктуаций частоты гена в гаплоидной популяции при воздействии случайного дрейфа. Если $\lambda \neq \mu$, то (4.3) сводится к модели отбора Райта, в которой $\dfrac{\lambda}{\mu} = 1 + \sigma$, а σ — коэффициент отбора, выражающий преимущество индивидуумов типа A над индивидуумами типа a.

В общем случае, соответствующем формуле (4.1), мы скажем, что отбор имеет место, если $f(z) \not\equiv g(z)$, т. е. если вероятностные производящие функции числа потомков двух типов различны.

(2) Пусть $f(z) = g(z) = (q + pz)^2$, $q = 1 - p$, $0 < p < 1$. Порожден-

ная цепь Маркова (4.2) в данном случае принимает вид

$$P_{ij}=\left[\binom{2N}{N}\right]^{-1}\binom{2i}{j}\binom{2(N-i)}{N-j},\qquad i,\ j=0,\ 1,\ 2,\ \ldots,\ N. \tag{4.5}$$

Интерпретация этой цепи Маркова как модели полисомного наследования дана в примере 3 § 2 гл. 2.

(3) Пусть

$$f(z)=\frac{q^{\alpha}}{(1-pz)^{\alpha}},\quad g(z)=\frac{q^{\beta}}{(1-pz)^{\beta}},\quad q=1-p,\quad 0<p<1,\quad \alpha,\ \beta>0.$$

Тогда, используя разложение

$$\frac{1}{(1-pz)^{\gamma}}=\sum_{n=0}^{\infty}\binom{\gamma+n-1}{n}p^n z^n,$$

мы видим, что порожденная цепь Маркова имеет следующие переходные вероятности

$$P_{ij}=\left[\binom{i\alpha+(N-i)\beta+N-1}{N}\right]^{-1}\binom{i\alpha+j-1}{j}\binom{(N-i)\beta+N-i-1}{N-j}, \tag{4.6}$$
$$i,\ j=0,\ 1,\ \ldots,\ N.$$

Эта цепь Маркова может интерпретироваться как модель роста гетерозиготной популяции (см. стр. 321).

Явления миграции и мутации могут быть учтены в частотной модели, представляемой порожденной марковской цепью, следующим образом. Сначала рассмотрим явление иммиграции. Предположим, что кроме размножения происходит процесс иммиграции в систему независимо от текущего размера популяции. Пусть $h(s)$ и $k(s)$ — вероятностные производящие функции числа иммигрантов типов A и a, поступающих в систему за время одного поколения соответственно. Совместная производящая функция $U(z,\ w)=\sum_{n,\,m=0}^{\infty}u_{n,\,m}z^n w^m$ для вероятностей числа индивидуумов типов A и a в популяции в следующем поколении равна $U(z,w)=f^i(z)\,h(z)\,g^{N-i}(w)\,k(w)$. Здесь

$$u_{n,\,m}=\mathrm{P}\{X(1)=n,\ Y(1)=m\mid X(0)=i,\ Y(0)=N-i\}.$$

Для того чтобы построить порожденную цепь Маркова, введем условие $X(1)+Y(1)=N$. Порожденная марковская цепь имеет переходные вероятности

$$P_{ij}=\mathrm{P}\{X(1)=j\mid X(0)=i,\ Y(0)=N-i,\ X(1)+Y(1)=N\}=$$
$$=\frac{\text{коэффициент при } z^j w^{N-j} \text{ в разложении } f^i(z)\,h(z)\,g^{N-i}(w)\,k(w)}{\text{коэффициент при } t^N \text{ в разложении } f^i(t)\,h(t)\,g^{N-i}(t)\,k(t)}, \tag{4.7}$$
$$i,\ j=0,\ 1,\ \ldots,\ N.$$

Явление мутации можно учесть в порожденной цепи Маркова двумя способами. Предположим, что каждый индивидуум типа A может мутировать в индивидуум типа a с вероятностью α_1 ($0 \leqslant \alpha_1 \leqslant 1$) и каждый индивидуум типа a может мутировать в индивидуум типа A с вероятностью α_2 ($0 \leqslant \alpha_2 \leqslant 1$). По отношению к мутации индивидуумы являются независимыми. Для того чтобы образовать следующее поколение, мы можем постулировать, что процесс мутации следует за процессом роста (ветвящегося размножения) или наоборот. Эти два постулата приводят к различным цепям Маркова.

Наиболее просто описать функционирование модели мутации непосредственно в терминах соответствующих вероятностных производящих функций. Рассмотрим сначала случай, когда мутация происходит после воспроизведения. Пусть $f(z)$ и $g(w)$, как и прежде, — производящие функции числа потомков одного родителя типов A и a соответственно. Предположим, что при следующем процессе воспроизведения каждый потомок индивидуума типа A (типа a) может производить индивидуумы обоих видов с производящей функцией $A(z, w)$ ($B(z, w)$). Другими словами, мы постулируем две фазы процесса воспроизведения: первая соответствует обычному процессу размножения, в котором потомки являются копиями родителей, в то время как вторая фаза соответствует своего рода процессу трансформации и роста, когда один индивидуум одного типа может превращаться в индивидуумов другого типа, т. е. коэффициент $a_{m,n}$ в разложении $A(z, w)$ является вероятностью того, что на второй фазе потомок типа A произведет m индивидуумов типа A и n индивидуумов типа a. Производящая функция, описывающая вид популяции потомков одного индивидуума типа A (типа a) после обеих фаз воспроизводства, равна $f(A(z, w))$ $(g(B(z, w)))$.

Механизм мутации получается при следующем ограничении:

$$A(z, w) = (1-\alpha_1)z + \alpha_1 w, \quad B(z, w) = \alpha_2 z + (1-\alpha_2)w. \tag{4.8}$$

При этом вторая фаза воспроизводства сводится к процессу трансформации С вероятностью $1-\alpha_1$ потомок типа A сохраняет свой тип, а с вероятностью α_1 мутирует в тип a. Аналогичную интерпретацию имеет и производящая функция $B(z, w) = \alpha_2 z + (1-\alpha_2)w$.

В результате производящая функция популяции потомков, полученных при размножении одного индивидуума типа A и подверженных затем мутационному давлению, равна

$$f(z, w) = f[(1-\alpha_1)z + \alpha_1 w]. \tag{4.9}$$

Аналогичным образом можно описать популяцию потомков одного родителя типа a, учитывая мутационное давление, с помощью производящей функции

$$g(z, w) = g[\alpha_2 z + (1-\alpha_2)w]. \tag{4.10}$$

С другой стороны, если мутационные превращения происходят до воспроизводства, производящие функции принимают вид

$$f^*(z, w) = (1-\alpha_1) f(z) + \alpha_1 g(w),$$
$$g^*(z, w) = \alpha_2 f(z) + (1-\alpha_2) g(w).$$

Проведенное выше обсуждение приводит к следующей более общей конструкции.

Мы постулируем, что каждый индивидуум типа A может производить потомков обоих типов. Обозначим вероятностную производящую функцию размера популяции потомков через $f(z, w)$. Аналогично предположим, что индивидуум типа a может производить индивидуумов обоих типов, и пусть $g(z, w)$ — производящая функция популяции его потомков. Пусть $h(z, w)$ — производящая функция числа индивидуумов типов A и a, иммигрирующих в систему в течение каждого периода. Пусть $(X(n), Y(n)) = Z(n)$, $n = 0, 1, \ldots,$ — результирующий двумерный ветвящийся процесс. Вероятностная производящая функция величины $Z(1)$ при начальных условиях $X(0) = i$, $Y(0) = k$ равна

$$[f(z, w)]^i [g(z, w)]^k h(z, w). \tag{4.11}$$

Матрица переходных вероятностей порожденной цепи Маркова при условии, что размер популяции фиксирован числом N, вычисляется обычным образом. Получим

$$P\{X(1) = j \mid X(0) = i,\ Y(0) = N-i,\ X(1)+Y(1) = N\} = P_{ij} =$$
$$= \frac{\text{коэффициент при } z^j w^{N-j} \text{ в разложении } f^i(z, w)\, g^{N-i}(z, w)\, h(z, w)}{\text{коэффициент при } t^N \text{ в разложении } f^i(t, t)\, g^{N-i}(t, t)\, h(t, t)}. \tag{4.12}$$

В частности, если

$$f(z, w) = e^{-\lambda} \exp\{\lambda[(1-\alpha_1) z + \alpha_1 w]\},$$
$$g(z, w) = e^{-\mu} \exp\{\mu[\alpha_2 z + (1-\alpha_2) w]\},$$
$$h(z, w) = \exp\{[a(z-1) + b(w-1)]\},$$

то переходные вероятности (4.12) порожденной цепи Маркова равны

$$P_{ij} = \binom{N}{j} \frac{[(1-\alpha_1)\lambda i + \mu(N-i)\alpha_2 + a]^j\, [\alpha_1 \lambda i + \mu(N-i)(1-\alpha_2) + b]^{N-j}}{[\lambda i + \mu(N-i) + a + b]^N}, \tag{4.13}$$
$$i, j = 0, 1, \ldots, N.$$

Эта конечная цепь Маркова была предложена С. Райтом для изучения флуктуаций частоты гена при учете отбора, мутации, миграции и генетического отбора.

§ 5. МОДЕЛИ РОСТА ПОПУЛЯЦИЙ С НЕСКОЛЬКИМИ ТИПАМИ ИНДИВИДУУМОВ

Теория, развитая выше, непосредственно обобщается на случай произвольного числа типов. Рассмотрим многомерный ветвящийся процесс для p типов индивидуумов, обозначенных через $A_1, A_2, \ldots$ $\ldots, A_p$. Предположим, что индивидуум типа A_k за одно поколение порождает потомство всех типов в соответствии с производящей функцией

$$f_k(s_1, s_2, \ldots, s_p) = \sum_{\nu_i \geqslant 0} a^{(k)}_{\nu_1, \nu_2, \ldots, \nu_p} s_1^{\nu_1} s_2^{\nu_2} \ldots s_p^{\nu_p}, \quad k = 1, 2, \ldots, p, \tag{5.1}$$

где $a^{(k)}_{\nu_1, \nu_2, \ldots, \nu_p}$ — вероятность того, что индивидуум типа A_k произведет потомство, состоящее из ν_1 индивидуумов типа A_1, ν_2 индивидуумов типа A_2 и т. д. Предполагается, что индивидуумы действуют независимо.

Пусть $(X_1(n), X_2(n), \ldots, X_p(n))$ означает соответствующий ветвящийся процесс, где $X_k(n)$ является числом индивидуумов типа A_k в начале n-го поколения. Вероятностная производящая функция числа потомков за одно поколение равна

$$f_1^{i_1}(s_1, \ldots, s_p) f_2^{i_2}(s_1, \ldots, s_p) \ldots f_p^{i_p}(s_1, \ldots, s_p) h(s_1, \ldots, s_p), \tag{5.2}$$

где $X_k(0) = i_k$ $(k = 1, 2, \ldots, p)$, а $h(s_1, \ldots, s_p)$ — производящая функция числа иммигрантов различных типов.

Как и в случае двух типов, цепь Маркова, порождаемая путем фиксации размера популяции, может интерпретироваться как частотная модель. Мы опишем ее точную структуру. Пространство состояний будет содержать все совокупности $\mathbf{k} = (k_1, k_2, \ldots, k_p)$ неотрицательных целых чисел, удовлетворяющих ограничению $\sum_{\nu=1}^{p} k_\nu = N$. Матрица переходных вероятностей строится с помощью ветвящегося процесса следующим образом. Пусть $\mathbf{k} = (k_1, k_2, \ldots$ $\ldots, k_p)$, $\mathbf{l} = (l_1, l_2, \ldots, l_p)$. Тогда

$$P_{\mathbf{kl}} =$$

$$= \mathrm{P}\left\{ X_1(1) = l_1,\ X_2(1) = l_2, \ldots, X_p(1) = l_p \,\middle|\, \begin{array}{l} X_\nu(0) = k_\nu, \quad \nu = 1, \ldots, p, \\ \sum_{\nu=1}^{p} X_\nu(0) = N,\ \sum_{\nu=1}^{p} X_\nu(1) = N \end{array} \right\} =$$

$$= \frac{\text{коэффициент при } s_1^{l_1} s_2^{l_2} \ldots s_p^{l_p} \text{ в разложении } f_1^{k_1}(\mathbf{s}) f_2^{k_2}(\mathbf{s}) \ldots f_p^{k_p}(\mathbf{s})\, h(\mathbf{s})}{\text{коэффициент при } t^N \text{ в разложении } f_1^{k_1}(\mathbf{t}) f_2^{k_2}(\mathbf{t}) \ldots f_p^{k_p}(\mathbf{t})}, \tag{5.3}$$

где $\mathbf{s} = (s_1, \ldots, s_p)$, $\mathbf{t} = (t_1, \ldots, t_p)$.

В частном случае

$$f_i(\mathbf{s}) = \exp\left\{\lambda_i\left(\sum_{\nu=1}^{p} \alpha_{i\nu}s_\nu - 1\right)\right\}, \qquad i = 1, 2, \ldots, p, \tag{5.4}$$

$$h(\mathbf{s}) = \exp\left\{\sum_{\nu=1}^{p} c_\nu (s_\nu - 1)\right\}, \tag{5.5}$$

где

$$\alpha_{i\nu} \geqslant 0, \quad \sum_{\nu=1}^{p} \alpha_{i\nu} = 1, \ \lambda_i > 0, \qquad \nu, \ i = 1, \ldots, p,$$
$$c_\nu > 0, \qquad \nu = 1, \ldots, p;$$

переходные вероятности (5.3) равны

$$P_{\mathbf{kl}} = \frac{\binom{N}{l_1, l_2, \ldots, l_p} \prod_{\nu=1}^{p} \left[\sum_{i=1}^{p} k_i\lambda_i\alpha_{i\nu} + c_\nu\right]^{l_\nu}}{\left[\sum_{i,\nu} k_i\lambda_i\alpha_{i\nu} + \sum_{\nu} c_\nu\right]^N}, \tag{5.6}$$

где

$$\binom{N}{l_1, l_2, \ldots, l_p} = \frac{N!}{l_1!\, l_2! \ldots l_p!}.$$

Параметры, встречающиеся в выражении (5.6), должны быть интерпретированы следующим образом: $\alpha_{i\nu}$ является вероятностью того, что индивидуум типа A_i после рождения будет мутировать в тип A_ν, λ_i представляет собой относительный коэффициент отбора (т. е. приспособленности) типа A_i, а c_i является средней интенсивностью, с которой индивидуумы типа A_i иммигрируют в популяцию. Переходные вероятности (5.6) являются точным многомерным аналогом марковской модели Райта частоты гена при учете мутации, миграции и отбора.

Вероятностная модель самооплодотворения при диплоидии

Вероятностная модель самооплодотворения может быть построена аналогично (5.3) следующим образом. Рассмотрим случай трех генотипов AA, Aa и aa. Каждый индивидуум спаривается с индивидуумом такого же типа. Пусть $f_1(u)$ — вероятностная производящая функция числа потомков, получающихся от одного спаривания вида $AA \times AA$. Все потомки, конечно, имеют вновь тип AA. Пусть $f_3(w)$ — вероятностная производящая функция числа потомков, производимых от одного спаривания $aa \times aa$. Наконец, пусть $f_2(v)$ — вероятностная производящая функция числа потомков, порождаемых при спаривании $Aa \times Aa$. В последнем случае каждый потомок имеет генотип AA, Aa или aa с вероятностями $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ соответственно.

Пусть число индивидуумов типов AA, Aa и aa в начальном поколении равнялось i_1, i_2 и i_3 соответственно. Флуктуации популяции из-за самооплодотворения и роста описываются марковским ветвящимся процессом, у которого совместная производящая функция общего числа потомков в следующем поколении равна

$$f_1^{i_1}(u)\, f_2^{i_2}\left(\frac{1}{4}u+\frac{1}{2}v+\frac{1}{4}w\right) f_3^{i_3}(w) \tag{5.7}$$

при начальном состоянии (i_1, i_2, i_3). Коэффициенты при степенях u, v и w в разложении (5.7) являются соответственно вероятностями того или иного числа потомков AA, Aa и aa.

Сформулируем теперь частотную модель, порождаемую фиксированием размера популяции ветвящегося процесса (5.7). При этом подобно (5.3) можно найти, что

$$P_{\mathbf{i},\mathbf{j}} = \frac{\text{коэффициент при } u^{j_1}v^{j_2}w^{j_3} \text{ в разложении } f_1^{i_1}(u)\, f_2^{i_2}\left(\frac{1}{4}u+\frac{1}{2}v+\frac{1}{4}w\right) f_3^{i_3}(w)}{\text{коэффициент при } t^N \text{ в разложении } f_1^{i_1}(t)\, f_2^{i_2}(t)\, f_3^{i_3}(t)}, \tag{5.8}$$

где $\mathbf{i} = (i_1, i_2, i_3)$, $\mathbf{j} = (j_1, j_2, j_3)$,

$$j_1+j_2+j_3 = i_1+i_2+i_3 = N, \quad i_\nu \geqslant 0, \quad j_\nu \geqslant 0.$$

Переходные вероятности (5.8) являются частным случаем (5.3).

Эту конечную цепь Маркова можно рассматривать как вероятностный аналог детерминированного процесса самооплодотворения, сформулированного в § 1. Если взять в качестве рассматриваемых производящих функций пуассоновские, то цепь Маркова с переходными вероятностями (5.8) будет описывать процесс самооплодотворения с биномиальным выбором.

§ 6. СОБСТВЕННЫЕ ЗНАЧЕНИЯ ЦЕПЕЙ МАРКОВА, ПОРОЖДЕННЫХ ПРЯМЫМ ПРОИЗВЕДЕНИЕМ ВЕТВЯЩИХСЯ ПРОЦЕССОВ

В предыдущих двух параграфах мы рассмотрели несколько цепей Маркова, связанных с ветвящимися процессами. В настоящем параграфе мы изучим некоторые вопросы, относящиеся к структуре этих цепей. Более точно, будет найдено все множество собственных значений и описаны некоторые свойства соответствующих векторов. В § 8 мы укажем на различные применения и интерпретации этих результатов.

Пусть задана производящая функция

$$f(s) = \sum_{n=0}^{\infty} a_n s^n, \quad a_n \geqslant 0, \quad \sum_{n=0}^{\infty} a_n = 1. \tag{6.1}$$

Рассмотрим порожденную цепь Маркова с матрицей переходных вероятностей $\mathbf{P} = \|P_{ij}\|$, где

$$P_{ij} = \frac{\text{коэффициент при } s^j t^{N-j} \text{ в разложении } f^i(s) f^{N-i}(t)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)} =$$

$$= \frac{\text{коэффициент при } s^j \omega^N \text{ в разложении } f^i(s\omega) f^{N-i}(\omega)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)}, \tag{6.2}$$

$$i, j = 0, 1, \ldots, N.$$

Для того чтобы найти собственные значения матрицы $\mathbf{P}$, проанализируем производящую функцию вероятностей $\{P_{ij}\}$:

$$G(s) = \sum_{j=0}^{N} P_{ij} s^j = \frac{\text{коэффициент при } \omega^N \text{ в разложении } f^i(s\omega) f^{N-i}(\omega)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)}. \tag{6.3}$$

Дифференцируя $G(s)$ по s, получим

$$G'(s) = \sum_{j=0}^{N} P_{ij} j s^{j-1} =$$

$$= \frac{\text{коэффициент при } \omega^N \text{ в разложении } i f^{i-1}(s\omega) f'(s\omega) \omega f^{N-i}(\omega)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)} =$$

$$= \frac{\text{коэффициент при } \omega^{N-1} \text{ в разложении } i f^{i-1}(s\omega) f'(s\omega) f^{N-i}(\omega)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)}. \tag{6.4}$$

Полагая $s = 1$, имеем

$$G'(1) = \sum_{j=0}^{N} P_{ij} j = i \left\{ \frac{\text{коэффициент при } \omega^{N-1} \text{ в разложении } f^{N-1}(\omega) f'(\omega)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)} \right\} = \lambda_1 i, \tag{6.5}$$

где

$$\lambda_1 = \frac{\text{коэффициент при } \omega^{N-1} \text{ в разложении } f^{N-1}(\omega) f'(\omega)}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)}$$

Таким образом, если мы будем трактовать $\mathbf{P}$ как матрицу, задающую линейное преобразование в линейном пространстве размерности $N+1$, то из соотношения (6.5) следует, что $\mathbf{Pv} = \lambda_1 \mathbf{v}$, где $\mathbf{v} = (0, 1, 2, \ldots, N)$. Таким образом, $(0, 1, 2, \ldots, N)$ является собственным вектором матрицы $\mathbf{P}$, соответствующим собственному значению λ_1. Очевидно, $\mathbf{Pe} = 1 \cdot \mathbf{e}$, где $\mathbf{e} = (1, 1, \ldots, 1)$, поскольку $\sum_j P_{ij} = 1$ при всех i. Эти два результата являются частным проявлением формулы

$$\sum_j P_{ij} Q_r(j) = \lambda_r Q_r(i), \qquad i = 0, 1, \ldots, N, \tag{6.6}$$

где $Q_r(\cdot)$ является полиномом степени r. При $r=0, 1$ имеем $Q_0(x)\equiv 1$ и $Q_1(x)=x$.

Наша цель теперь — доказать формулу (6.6) для всех r. Для этого продифференцируем (6.4) еще раз по s:

$$G''(s)=\sum_{j=0}^{N} P_{ij}[j(j-1)]\,s^{j-2}=$$

$$=\frac{\text{коэффициент при } \omega^{N-1} \text{ в разложении } \{i(i-1)f^{i-2}(s\omega)(f'(s\omega))^2\omega f^{N-i}(\omega)+if^{i-1}(s\omega)f''(s\omega)\omega f^{N-i}(\omega)\}}{\text{коэффициент при } \omega^{N} \text{ в разложении } f^{N}(\omega)}.$$

Положим $s=1$, тогда

$$G''(1)=\sum_{j=0}^{N} P_{ij}[j(j-1)]=$$

$$=\frac{\text{коэффициент при } \omega^{N-2} \text{ в разложении } [i(i-1)f^{N-2}(\omega)(f'(\omega))^2+if^{N-1}(\omega)f''(\omega)]}{\text{коэффициент при } \omega^{N} \text{ в разложении } f^{N}(\omega)}.$$

Пусть

$$\lambda_2=\frac{\text{коэффициент при } \omega^{N-2} \text{ в разложении } f^{N-2}(\omega)(f'(\omega))^2}{\text{коэффициент при } \omega^{N} \text{ в разложении } f^{N}(\omega)}.$$

Тогда

$$\sum_{j=0}^{N} P_{ij}[j(j-1)]=\lambda_2[i(i-1)]+i\varepsilon_2, \qquad i=0,\ 1,\ \ldots,\ N,$$

где

$$\varepsilon_2=\frac{\text{коэффициент при } \omega^{N-2} \text{ в разложении } f^{N-1}(\omega)f''(\omega)}{\text{коэффициент при } \omega^{N} \text{ в разложении } f^{N}(\omega)}.$$

Таким образом,

$$\sum_{j=0}^{N} P_{ij}j^2=\lambda_2 i^2+(\text{полином степени} \leqslant 1 \text{ по } i).$$

Следовательно, нами установлена формула

$$\sum_{j=0}^{N} P_{ij}j^r=\lambda_2 i^r+u_{r-1}(i), \qquad i=0,\ 1,\ \ldots,\ N,$$

где $u_{r-1}(\cdot)$ — полином степени $\leqslant r-1$, по крайней мере для $r=0,\ 1,\ 2$. В действительности же справедлив следующий более общий результат.

Теорема 6.1. *Пусть* $\lambda_0 = 1$,

$$\lambda_r = \frac{\text{коэффициент при } \omega^{N-r} \text{ в разложении } f^{N-r}(\omega)\,[f'(\omega)]^r}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)}, \tag{6.7}$$

$$r = 1,\ 2,\ \ldots,\ N.$$

Тогда

$$\sum_{j=0}^{N} P_{ij} j^r = \lambda_r i^r + u_{r-1}(i), \tag{6.8}$$

где $u_{r-1}(i)$ — *полином степени* $\leqslant r-1$ *по* i, $r = 0, 1, \ldots, N$. (*Условимся, что полином степени* -1 *есть тождественный нуль.*)

Доказательство. Если продифференцировать (6.3) k раз, левая часть примет вид

$$G^{(k)}(s) = \sum_{j=0}^{N} P_{ij}\,[j(j-1)\ldots(j-k+1)]\,s^{j-k} = \\ = \sum_{j=0}^{N} P_{ij}\,[j^k + v_{k-1}(j)]\,s^{j-k}, \tag{6.9}$$

где $v_{k-1}(\cdot)$ — полином степени $\leqslant k-1$. Если продифференцировать $f^i(s\omega) f^{N-i}(\omega)$ в правой части k раз по s, то получим

$$\frac{d^k}{ds^k}[f^i(s\omega) f^{N-i}(\omega)] = f^{N-i}(\omega)\frac{d^k}{ds^k}[f^i(s\omega)] = \\ = [i(i-1)\ldots(i-k+1)]\,f^{i-k}(s\omega)\,(f'(s\omega))^k f^{N-i}(\omega)\,\omega^k + \\ + \text{остальные члены}. \tag{6.10}$$

Первый член появляется при последовательном дифференцировании $f^i(s\omega)$ k раз. Оставшиеся члены возникают при дифференцировании других сомножителей, таких, как $f'(s\omega)$, $f''(s\omega)$ и т. д., которые появляются в процессе дифференцирования исходного выражения. Степени этих возникающих сомножителей не зависят от i. Это означает, что результат указанных операций дифференцирования есть линейная комбинация различных степеней f, коэффициенты которой содержат в качестве сомножителей максимум $k-1$ из чисел i, $i-1$, $i-2$, $\ldots$, $i-k+1$, т. е. коэффициенты являются полиномами степени не выше $k-1$ по отношению к переменной i. Каждое слагаемое второго члена выражения (6.10) зависит от f через выражение вида

$$f^{i-l}(s\omega)\,f^{N-i}(\omega)\,(f'(s\omega))^{\nu_1}\,[f''(s\omega)]^{\nu_2}\ldots$$

при некотором $0 \leqslant l \leqslant k$ и т. д.

Группируя члены в равенстве (6.10), получим

$$\frac{d^k}{ds^k}\left[f^i(s\omega)f^{N-i}(\omega)\right] = i^k f^{i-k}(s\omega)\left[f'(s\omega)\right]^k f^{N-i}(\omega)\,\omega^k + u_{k-1}(i;\,f), \qquad (6.11)$$

где $u_{k-1}(i;\,f)$ — полином от переменной i степени не выше $k-1$ при $s=1$. Фактически при $s=1$ степени члена $f^{i-l}(s\omega)f^{N-i}(\omega)$ складываются, и зависимость от i сохраняется лишь в коэффициентах при $f^{N-l}(\omega)$.

Предположим, что равенство (6.8) доказано для всех степеней $\leqslant r-1$. Тогда в силу гипотезы индукции

$$G^{(r)}(1) = \sum_{j=0}^{N} P_{ij}j^r + \sum_{j=0}^{N} P_{ij}v_{r-1}(j) = \sum_{j=0}^{N} P_{ij}j^r + \tilde{v}_{r-1}(i), \qquad (6.12)$$

где $v_{r-1}(\cdot)$ и $\tilde{v}_{r-1}(\cdot)$ — полиномы степеней не выше $r-1$. В силу (6.11)

$$G^{(r)}(1) = \frac{\text{коэффициент при } \omega^N \text{ в разложении } \left[\dfrac{d^r}{ds^r} f^i(s\omega)\, f^{N-i}(\omega)\right]_{s=1}}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)} = i^r\lambda_r + u_{r-1}(i). \qquad (6.13)$$

Сравнивая (6.12) и (6.13), видим, что утверждение индукции выполняется и для r. Это завершает доказательство формулы (6.8). ∎

Докажем теперь основную теорему данного параграфа, которая раскрывает природу собственных значений матрицы **P**.

Теорема 6.2. (1) *Собственными значениями матрицы* **P** *являются числа* $\lambda_0, \lambda_1, \ldots, \lambda_N$, *определяемые равенством* (6.7). *Более того,*

(2) *числа* λ_i *удовлетворяют соотношению* $1=\lambda_0=\lambda_1>\lambda_2>\ldots$ $\ldots>\lambda_N>0$, *если* $a_0a_1a_2>0$, *где* $f(s)=\sum_{n=0}^{\infty} a_n s^n$.

Доказательство. (1) Введем базис из $N+1$ векторов размерности $N+1:\pi_r = (\pi_r(0), \pi_r(1), \ldots, \pi_r(N))$, $r=0, 1, \ldots, N$, где $\pi_r(x)=x^r$. Эти векторы, очевидно, линейно независимы, поскольку в противном случае мы могли бы построить полином N-й степени, имеющий $N+1$ корней, что невозможно.

Найдем теперь вид матрицы **P** в специально выбранном базисе $\{\pi_0, \pi_1, \ldots, \pi_N\}$. Понятно, что i-й столбец матрицы **P** состоит из коэффициентов вектора $\mathbf{P}\pi_i$, который представлен в виде линейной комбинации векторов $\{\pi_j\}_{j=0}^{N}$.

Из соотношения (6.8) следует, что матрица **P** в новом базисе имеет вид

$$\left\| \begin{matrix} 1 & 0 & * & * & & & \\ 0 & \lambda_1 & * & * & & & \\ . & 0 & \lambda_2 & * & & & \\ . & 0 & 0 & \lambda_3 & & & \\ . & . & 0 & 0 & \cdot & & \\ \vdots & \vdots & \vdots & \vdots & & \cdot & \\ \vdots & \vdots & \vdots & \vdots & & & \cdot \\ 0 & 0 & 0 & 0 & & & \end{matrix} \right\|,$$

где элементы, обозначенные $*$, соответствуют полиному u_{r-1} в (6.8). Таким образом, матрица **P**, представленная в новом базисе, является треугольной и ее собственные значения, очевидно, равны $1 = \lambda_0, \lambda_1, \ldots, \lambda_N$. Поскольку собственные значения инвариантны относительно выбора базиса, доказательство пункта (1) закончено.

(2) Заметим, что для любого степенного ряда $\pi(\omega) = \sum_{m=0}^{\infty} c_m \omega^m$ коэффициент при ω^M, очевидно, равен μ^{-1} [коэффициент при ω^{M-1} в разложении $\pi'(\omega)$]. Полагая $M = N - r$ и $\pi(\omega) = f^{N-r}(\omega)\,(f'(\omega))^r$, получим

коэффициент при ω^{N-r} в разложении $\{f^{N-r}(\omega)(f'(\omega))^r\} =$

$$= \frac{1}{N-r}\Big[\text{коэффициент при } \omega^{N-r-1} \text{ в разложении}$$
$$\left\{(N-r) f^{N-r-1}(\omega)(f'(\omega))^{r+1} + r f^{N-r}(\omega)(f'(\omega))^{r-1} f''(\omega)\right\}\Big] =$$
$$= \text{коэффициент при } \omega^{N-r-1} \text{ в разложении}$$
$$\left\{ f^{N-r-1}(\omega)(f'(\omega))^{r+1} + \frac{r}{N-r} f^{N-r}(\omega)(f'(\omega))^{r-1} f''(\omega) \right\}.$$

Если разделить обе части на коэффициент при ω^N в разложении $f^N(\omega)$, то, вспоминая определение λ_r, получим соотношение

$$\lambda_r = \lambda_{r+1} + \frac{\text{коэффициент при } \omega^{N-r-1} \text{ в разложении } \left\{\frac{r}{N-r} f^{N-r}(\omega)\,(f'(\omega))^{r-1} f''(\omega)\right\}}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)},$$
$$r \leqslant N-1. \tag{6.14}$$

При $r = 0$ второе слагаемое в правой части, очевидно, равно 0, таким образом, $1 = \lambda_0 = \lambda_1$.

Поскольку $f(\omega)$ — вероятностная производящая функция, то ее коэффициенты неотрицательны. Следовательно, коэффициент при ω^{N-r-1} в разложении $f''(\omega) f^{N-r}(\omega)\,[f'(\omega)]^{r-1}$ неотрицателен, так что всегда $\lambda_r \geqslant \lambda_{r+1}$. Из предположения о том, что $a_0 a_1 a_2 > 0$, непосредственно вытекает, что второе слагаемое в (6.14) положительно при $r > 0$. (На самом деле достаточны гораздо более слабые предположения.) Поэтому $\lambda_r > \lambda_{r+1}$ $(r \geqslant 1)$, что и требовалось доказать. ■

Из теоремы 6.2 известны все собственные значения матрицы (6.2). В предположениях пункта (2), т. е. $a_0 a_1 a_2 > 0$, где

$$f(s) = \sum_{n=0}^{\infty} a_n s^n,$$

собственные значения $\lambda_r < 1$, $r = 2, 3, \ldots, N$, все различны. Обратимся теперь к задаче описания собственных векторов, соответствующих этим собственным значениям.

Мы уже показали, что собственному значению $1 = \lambda_0 = \lambda_1$ соответствует пара правых собственных векторов $\mathbf{e} = (1, 1, 1, \ldots, 1)$ и $\mathbf{u} = (0, 1, 2, \ldots, N)$. Интуитивно более привлекательна пара векторов, являющихся линейной комбинацией указанных, а именно $\left(1, \frac{N-1}{N}, \ldots, \frac{1}{N}, 0\right) = \mathbf{e} - \frac{\mathbf{u}}{N}$ и $\left(0, \frac{1}{N}, \frac{2}{N}, \ldots, 1 = \frac{N}{N}\right) = \frac{\mathbf{u}}{N}$. Для компонент векторов, представленных в такой форме, можно дать вероятностную интерпретацию (см. § 2 гл. 4).

Анализ выражения для P_{ij} показывает, что матрица $\mathbf{P}$ имеет вид

$$\mathbf{P} = \begin{Vmatrix} 1 & 0 & 0 & \ldots & 0 \\ * & * & * & \ldots & * \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ * & * & * & & * \\ 0 & 0 & 0 & \ldots & 1 \end{Vmatrix},$$

откуда с очевидностью следует, что 0 и N — поглощающие состояния. Более того, из предположения о том, что $0 < a_0 < 1$, следует, что 0 и N достижимы за один переход из каждого состояния $\{1, 2, \ldots, N-1\}$ с положительной вероятностью. Следовательно, 0 и N — единственные поглощающие состояния.

Следующая теорема характеризует правые собственные векторы, соответствующие собственным значениям λ_r $(r = 2, 3, \ldots, N)$.

Теорема 6.3. *Существуют полиномы $Q_r(x)$ степени r, для которых*

$$\mathbf{P}\mathbf{Q}_r = \lambda_r \mathbf{Q}_r.$$

(*Вектор $\mathbf{Q}_r$ равен* $(Q_r(0), Q_r(1), \ldots, Q_r(N))$.)

Доказательство. Теорема уже доказана при $r=0,\ 1$. Если $r>1$, мы знаем, что

$$\sum_{j=0}^{N} P_{ij}j^r = \lambda_r i^r + H_{r-1}(i),$$

где $H_{r-1}(\cdot)$ — полином степени $\leqslant r-1$. Выберем $Q_{r+1}(j)$ в виде

$$Q_{r+1}(j) = j^{r+1} + a_1 j^r + a_2 j^{r-1} + \ldots + a_{r+1} \qquad (j=0,\ 1,\ \ldots,\ N),$$

где постоянные $a_1,\ a_2,\ \ldots,\ a_{r+1}$ нужно подобрать. Необходимо выбрать $\{a_r\}$ так, чтобы соотношение

$$\sum_{j=0}^{N} P_{ij}Q_{r+1}(j) = \left(\lambda_{r+1}i^{r+1} + H_r(i)\right) + a_1\left(\lambda_r i^r + H_{r-1}(i)\right) + \\ + a_2\left(\lambda_{r-1}i^{r-1} + H_{r-2}(i)\right) + \ldots = \lambda_{r+1}Q_{r+1}(i) = \\ = \lambda_{r+1}\left(i^{r+1} + a_1 i^r + \ldots + a_{r+1}\right) \tag{6.15}$$

выполнялось при всех i. Приравнивая коэффициенты при i^r в обеих частях равенства, получим условие

$$a_1\lambda_{r+1} = a_1\lambda_r + \text{известная часть } H_r(i). \tag{6.16}$$

Поскольку $\lambda_{r+1} < \lambda_r$ при $r \geqslant 1$, можно найти такое число a_1, чтобы равенство (6.16) удовлетворялось. Затем, приравнивая коэффициенты при i^{r-1} в обеих частях (6.15) и используя то обстоятельство, что a_1 известно, получим соотношение вида

$$a_2\lambda_{r-1} + \text{известный член} = a_2\lambda_{r+1}. \tag{6.17}$$

Найдем a_2 из (6.17). Это возможно, поскольку $\lambda_{r+1} < \lambda_{r-1}$. Продолжая процедуру таким же образом, можно вычислить последовательно все коэффициенты $a_3,\ a_4,\ \ldots,\ a_{r+1}$ так, чтобы (6.15) выполнялось тождественно по i. ■

Вышеприведенное построение доказывает существование полного базиса из собственных векторов специального вида (т. е. полученных с помощью полиномов). Из теории матриц (см. стр. 505) следует, что существует полная система левых собственных векторов. Предположим, что мы записали правые собственные векторы в следующем порядке:

$$\begin{aligned} \varphi_0 &= \left(1,\ 1-\frac{1}{N},\ 1-\frac{2}{N},\ \ldots,\ 0\right) && \text{для} \quad \lambda_0 = 1, \\ \varphi_1 &= \left(0, \frac{1}{N},\ \frac{2}{N},\ \ldots,\ 1\right) && \text{для} \quad \lambda_1 = 1, \\ \varphi_r &= (Q_r(0),\ Q_r(1),\ \ldots,\ Q_r(N)) && \text{для} \quad \lambda_r,\ r=2,\ \ldots,\ N, \end{aligned} \tag{6.18}$$

где $Q_r(x)$ — полином, определенный в теореме 6.3.

Обозначим биортогональную систему левых собственных векторов через

$$\begin{aligned} \psi_0 &= (1,\ 0,\ 0,\ \ldots,\ 0,\ 0) && \text{для} \quad \lambda_0 = 1, \\ \psi_1 &= (0,\ 0,\ 0,\ \ldots,\ 0,\ 1) && \text{для} \quad \lambda_1 = 1, \\ \psi_r &= (\psi_r(0),\ \psi_r(1),\ \ldots,\ \psi_r(N)) && \text{для} \quad \lambda_r,\ r = 2,\ \ldots,\ N. \end{aligned} \tag{6.19}$$

По-видимому, довольно трудно выразить φ_r и ψ_r в замкнутом виде. Тем не менее φ_r может быть в принципе вычислен рекуррентно (фактически с помощью того же метода, который использовался при доказательстве теоремы 6.3).

Для вычисления ψ_r также можно предложить рекуррентный метод. Можно доказать, что компоненты вектора ψ_r имеют вид

$$\psi_r(i) = \binom{N}{i}(-1)^i R_{N-r}(i), \qquad i = 0,\ 1,\ 2,\ \ldots,\ N,$$

где $R_m(\cdot)$ — полином степени m, который можно найти рекуррентным образом.

Мы завершим параграф несколькими конкретными примерами.

Примеры. (А) $f(\omega) = e^{\lambda(\omega-1)}$. Тогда при $r > 1$

$$\lambda_r = \frac{\text{коэффициент при } \omega^{N-r} \text{ в разложении } f^{N-r}(\omega)\,(f'(\omega))^r}{\text{коэффициент при } \omega^N \text{ в разложении } f^N(\omega)} =$$

$$= \frac{\text{коэффициент при } \omega^{N-r} \text{ в разложении } e^{\lambda(N-r)(\omega-1)}\lambda^r e^{r\lambda(\omega-1)}}{\text{коэффициент при } \omega^N \text{ в разложении } e^{\lambda N(\omega-1)}} =$$

$$= \frac{\lambda^r\left[(\lambda N)^{N-r}/(N-r)!\right]}{\left[(\lambda N)^N/N!\right]} = \frac{N!}{(N-r)!}\,\frac{1}{N^r}.$$

(Б) $f(s) = (1-p+ps)^\gamma$, γ — целое число > 1,

$$\lambda_r = \frac{\text{коэффициент при } \omega^{N-r} \text{ в разложении } (1-p+p\omega)^{\gamma(N-r)}\left[\gamma p\,(1-p+p\omega)^{\gamma-1}\right]^r}{\text{коэффициент при } \omega^N \text{ в разложении } (1-p+p\omega)^{\gamma N}} =$$

$$= \gamma^r\binom{N\gamma-r}{N-r}\binom{N\gamma}{N}^{-1}.$$

(В) $f(s) = [(1-p)/(1-ps)]^\alpha$, $\alpha > 0$,

$$\lambda_r = \frac{\begin{array}{l}\text{коэффициент при } \omega^{N-r} \text{ в разложении} \\ \qquad [(1-p)/(1-p\omega)]^{\alpha(N-r)}\left[(\alpha p\,(1-p)^\alpha)/(1-p\omega)^{\alpha+1}\right]^r\end{array}}{\text{коэффициент при } \omega^N \text{ в разложении } [(1-p)/(1-p\omega)]^{\alpha N}} =$$

$$= \frac{(1-p)^{\alpha N}(\alpha p)^r\cdot\{\text{коэффициент при } \omega^{N-r} \text{ в разложении } [1/(1-p\omega)]^{\alpha N+r}\}}{(1-p)^{\alpha N}\cdot\{\text{коэффициент при } \omega^N \text{ в разложении } [1/(1-p\omega)]^{\alpha N}\}} =$$

$$= (-1)^r\alpha^r\binom{-\alpha N-r}{N-r}\binom{-\alpha N}{N}^{-1} = \alpha^r\binom{\alpha N+N-1}{N-r}\binom{\alpha N+N-1}{N}^{-1}.$$

Значение λ_2 для этих частных случаев представляет интерес, поскольку оно характеризует скорость приближения системы к поглощающему состоянию или, в генетической терминологии, скорость приближения к гомозиготному состоянию. Для примера (А)

$$\lambda_2 = \frac{N!}{(N-2)!} \quad \frac{1}{N^2} = 1 - \frac{1}{N}.$$

Для примера (Б)

$$\lambda_2 = \gamma^2 \binom{N\gamma - 2}{N-2} \binom{N\nu}{N}^{-1} = \frac{\gamma^2 N (N-1)}{N\gamma (N\gamma - 1)} = \frac{\gamma (N-1)}{N\gamma - 1} = 1 - (\gamma - 1) \frac{1}{N\gamma - 1},$$

следовательно, асимптотически по N

$$\lambda_2 \sim 1 - \frac{\gamma - 1}{\gamma} \frac{1}{N}.$$

Аналогично для случая (В)

$$\lambda_2 = \alpha^2 \binom{\alpha N + N - 1}{N-2} \binom{\alpha N + N - 1}{N}^{-1} = \frac{\alpha^2 N (N-1)}{\alpha N (\alpha N + 1)} = 1 - (1 + \alpha) \frac{1}{1 + \alpha N},$$

так что асимптотически

$$\lambda_2 \sim 1 - \frac{\alpha + 1}{\alpha} \cdot \frac{1}{N}.$$

Заметим, что в пуассоновском случае значение λ_2 (для больших N) является промежуточным между соответствующими значениями для биномиального распределения потомков и отрицательного биномиального распределения.

§ 7. СОБСТВЕННЫЕ ЗНАЧЕНИЯ ДЛЯ МОДЕЛИ МУТАЦИИ С НЕСКОЛЬКИМИ ТИПАМИ ИНДИВИДУУМОВ

В предыдущем параграфе мы нашли вид собственных значений и собственных векторов матрицы перехода порожденной цепи Маркова (6.2). В этом параграфе мы будем изучать порожденную марковскую цепь для p типов ($p > 2$) индивидуумов при общих предположениях относительно мутации. В частности, будет найден вид собственных значений и собственных векторов цепи Маркова с переходными вероятностями

$$P_{\mathbf{i}, \mathbf{k}} = \frac{\text{коэффициент при } s_1^{k_1} s_2^{k_2} \ldots s_p^{k_p} \text{ в разложении } \prod\limits_{\nu=1}^{p} f^{i_\nu}\left(\sum\limits_{\mu=1}^{p} a_{\nu\mu} s_\mu\right)}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)}, \tag{7.1}$$

где $\mathbf{i} = (i_1, i_2, \ldots, i_p)$, $\mathbf{k} = (k_1, k_2, \ldots, k_p)$, $\sum\limits_{\nu=1}^{p} i_\nu = \sum\limits_{\nu=1}^{p} k_\nu = N$, $i_\nu \geqslant 0$ и $k_\nu \geqslant 0$ при всех ν. $P_{\mathbf{i}, \mathbf{k}}$ задают условные (при условии, что общее число потомков равно N) частоты p типов индивидуумов, когда ν-й тип производит потомство с производящей функцией $f(s)$, после чего каждый из потомков мутирует в тип A_j с вероятностью $\alpha_{\nu j}$ ($j = 1, 2, \ldots, p$).

В силу определения матрица $\mathbf{\Gamma} = \| \alpha_{\nu\mu} \|_{\nu, \mu=1}^{p}$ является стохастической, т. е. $\alpha_{\nu\mu} \geqslant 0$, $\sum\limits_{\mu=1}^{p} \alpha_{\nu\mu} = 1$, $\nu = 1, \ldots, p$. Мы предположим, что матрица $\mathbf{\Gamma}$ диагонализируема (см. приложение). Это означает, что ее p собственным значениям соответствуют p линейно независимых собственных векторов. Пусть $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(p)}$ — полное семейство правых собственных векторов, причем можно взять $\mathbf{u}^{(1)} = (1, 1, 1, \ldots, 1)$, поскольку сумма элементов любой строки матрицы $\mathbf{\Gamma}$ равна 1.

Выражение (7.1) можно переписать в следующем виде:

$$P_{\mathbf{i}, \mathbf{k}} = \frac{\text{коэффициент при } s_p^N t_1^{k_1} t_2^{k_2} \ldots t_p^{k_p} \text{ в разложении } \prod\limits_{\nu=1}^{p} f^{i_\nu}\left(s \sum\limits_{\mu=1}^{p} \alpha_{\nu\mu} t_\mu\right)}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)}. \qquad (7.2)$$

Производящая функция, соответствующая вероятностям $P_{\mathbf{i}, \mathbf{k}}$, равна

$$G(t_1, t_2, \ldots, t_p) = \sum_{\mathbf{k}} P_{\mathbf{i}, \mathbf{k}} t_1^{k_1} t_2^{k_2} \ldots t_p^{k_p} =$$

$$= \frac{\text{коэффициент при } s^N \text{ в разложении } \prod\limits_{\nu=1}^{p} f^{i_\nu}\left(s \sum\limits_{\mu=1}^{p} \alpha_{\nu\mu} t_\mu\right)}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)}. \qquad (7.3)$$

Прежде чем находить все собственные значения матрицы $\mathbf{P} = \| P_{\mathbf{i}, \mathbf{k}} \|$, мы рассмотрим два частных случая, которые помогут пояснить общий метод нахождения собственных значений матрицы (7.1).

Случай 1. Дифференцируя уравнение (7.3) по t_μ и полагая затем $t_1 = t_2 = \ldots = t_p = 1$, получим

$$\frac{\partial}{\partial t_\mu} G(t_1, \ldots, t_p) \big|_{t_1 = t_2 = \ldots = t_p = 1} = \sum_{\mathbf{k}} k_\mu P_{\mathbf{i}, \mathbf{k}} =$$

$$= \text{коэффициент при } s^N \text{ в разложении}$$

$$\left\{ \frac{s i_1 \alpha_{1\mu} f^{N-1}(s) f'(s) + s i_2 \alpha_{2\mu} f^{N-1}(s) f'(s) + \ldots + s i_p \alpha_{p\mu} f^{N-1}(s) f'(s)}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)} \right\} =$$

$$= \frac{\left(\sum\limits_{\nu=1}^{p} i_\nu \alpha_{\nu\mu}\right) \left[\text{коэффициент при } s^{N-1} \text{ в разложении } f^{N-1}(s) f'(s)\right]}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)}. \qquad (7.4)$$

Равенство (7.4) выполнено при всех $\mathbf{i}$. Положим

$$\lambda_1 = \frac{\text{коэффициент при } s^{N-1} \text{ в разложении } f^{N-1}(s)\, f'(s)}{\text{коэффициент при } s^{N} \text{ в разложении } f^{N}(s)}.$$

Тогда равенство (7.4) можно записать в виде (справедливом при всех $\mathbf{i}$)

$$\sum_{\mathbf{k}} k_\mu P_{\mathbf{i},\mathbf{k}} = \lambda_1 \sum_{\nu=1}^{p} i_\nu \alpha_{\nu\mu}, \qquad \mu = 1, 2, \ldots, p. \tag{7.5}$$

Далее, умножим обе части равенства (7.5) на $u_\mu^{(q)}$ и просуммируем по μ, где $\mathbf{u}^{(q)} = \left(u_1^{(q)}, u_2^{(q)}, \ldots, u_p^{(q)}\right)$ — q-й собственный вектор матрицы $\mathbf{\Gamma} = \| \alpha_{\nu\mu} \|$, $q = 1, 2, \ldots, p$. Получим

$$\sum_{\mu} u_\mu^{(q)} \sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} k_\mu = \lambda_1 \sum_{\mu} u_\mu^{(q)} \sum_{\nu=1}^{p} i_\nu \alpha_{\nu\mu},$$

или, меняя порядок суммирования,

$$\sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} \left(\sum_{\mu=1}^{p} u_\mu^{(q)} k_\mu \right) = \lambda_1 \sum_{\nu=1}^{p} i_\nu \sum_{\mu=1}^{p} \alpha_{\nu\mu} u_\mu^{(q)} \quad \text{для всех } \mathbf{i}.$$

Так как $\mathbf{u}^{(q)}$ — собственный вектор матрицы $\|\alpha_{\nu\mu}\|$, соответствующий собственному значению γ_q, то

$$\sum_{\mu=1}^{p} \alpha_{\nu\mu} u_\mu^{(q)} = \gamma_q u_\nu^{(q)}.$$

Отсюда

$$\sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} \left(\sum_{\mu=1}^{p} u_\mu^{(q)} k_\mu \right) = \lambda_1 \gamma_q \sum_{\nu=1}^{p} i_\nu u_\nu^{(q)}. \tag{7.6}$$

Для дальнейшего удобно ввести вектор L_q, $\mathbf{k}$-я компонента которого равна

$$\sum_{\mu=1}^{p} u_\mu^{(q)} k_\mu = L_q(\mathbf{k}).$$

(Заметим, что $L_q(\mathbf{k})$ — линейная функция от $\mathbf{k}$, в частности $L_1(\mathbf{k}) = \sum_{\mu=1}^{p} k_\mu = N$.) Равенство (7.6) примет вид

$$\sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} L_q(\mathbf{k}) = \lambda_1 \gamma_q L_q(\mathbf{i}), \qquad q = 1, 2, \ldots, p \quad \text{для всех } \mathbf{i}, \tag{7.7}$$

откуда следует, что $\lambda_1 \gamma_q$ является собственным значением матрицы $\mathbf{P}$, а $L_q(\mathbf{i})$ — соответствующим ему собственным вектором.

Заметим, что $L_1(\mathbf{i}) \equiv N$ (константа, не зависящая от $\mathbf{i}$). Мы будем в какой-то мере выделять вектор $L_1(\mathbf{i})$ из остальных векторов

$L_q(\mathbf{i})$, $q = 2, \ldots, p$, которые являются (не постоянными) линейными функциями от $\mathbf{i}$. Собственные векторы $L_q(\mathbf{i})$, соответствующие собственным значениям $\lambda_1\gamma_q$ $(q = 1, 2, \ldots, p)$, линейно независимы. Это следствие линейной независимости собственных векторов $\mathbf{u}^{(q)}$ матрицы $\|\alpha_{\nu\mu}\|$. В самом деле, если бы векторы $L_q(\mathbf{i})$ были линейно зависимы, то для некоторых действительных постоянных a_n, не всех равных нулю, было бы $\sum_{n=1}^{p} a_n L_n(\mathbf{i}) = 0$. Но тогда

$$0 \equiv \sum_{n=1}^{p} a_n L_n(\mathbf{i}) \equiv \sum_{n=1}^{p} a_n \sum_{\nu=1}^{p} u_\nu^{(n)} i_\nu = \sum_{\nu=1}^{p} i_\nu \left(\sum_{n=1}^{p} a_n u_\nu^{(n)} \right).$$

Так как это выполняется при всех i, то отсюда следует

$$\sum_{n=1}^{p} a_n u_\nu^{(n)} = 0, \qquad \nu = 1, 2, \ldots, p,$$

что противоречит линейной независимости векторов $\mathbf{u}^{(q)}$.

Поскольку $L_1(\mathbf{i}), L_2(\mathbf{i}), \ldots, L_p(\mathbf{i})$ линейно независимы, то легко показать, что любая линейная функция $\mathfrak{L}(\mathbf{i})$ может быть представлена в виде линейной комбинации векторов $\{L_q(\mathbf{i})\}_{q=1}^{p}$, т. е. существуют постоянные $b_1, \ldots, b_p$, такие, что

$$\mathfrak{L}(\mathbf{i}) = \sum_{\nu=1}^{p} b_\nu L_\nu(\mathbf{i}) \quad \text{для всех } \mathbf{i}. \tag{7.8}$$

Из (7.7) и (7.8) вытекает, что матрица $\mathbf{P}$ переводит линейные функции в линейные, т. е. если $\mathfrak{L}(\mathbf{i})$ линейна, то функция

$$\sum_{\mathbf{j}} P_{\mathbf{i},\mathbf{j}} \mathfrak{L}(\mathbf{j}) = \mathfrak{L}^*(\mathbf{i}) \tag{7.9}$$

также линейна по $\mathbf{i}$. Мы используем это свойство ниже.

Случай 2. Он аналогичен случаю 1, но требует для своего доказательства несколько более сложных выкладок и в некотором смысле более общих рассуждений.

Дифференцируя равенство (7.3) по любым двум аргументам t_μ (они могут и совпадать) и полагая затем $t_1 = t_2 = \ldots = t_p = 1$, получим в левой части равенства

$$\frac{\partial^2}{\partial t_m\, \partial t_n} G(t_1, \ldots, t_p)\Big|_{t_1=t_2=\ldots=t_p=1} = \sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} \left(k_m k_n + \mathfrak{L}_{m,n}(\mathbf{k})\right), \tag{7.10}$$

$$m, n = 1, 2, \ldots, p,$$

где

$$\mathfrak{L}_{m,n}(\mathbf{k}) = \begin{cases} \mathbf{k}m, & \text{если } m = n, \\ 0 & \text{в противном случае.} \end{cases}$$

После дифференцирования правой части равенства (7.3) получим выражение

$$\frac{\text{коэффициент при } s^N \text{ в разложении } \left\{ s^2 \left(\sum_{\nu=1}^{p} i_\nu \alpha_{\nu m} \right) \left(\sum_{\nu=1}^{p} i_\nu \alpha_{\nu n} \right) f^{N-2}(s)\,[f'(s)]^2 \right\}}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)} +$$

$$+ \text{линейная функция от переменных } (i_1, i_2, \ldots, i_p). \qquad (7.11)$$

Введем в соответствии с предыдущим величину

$$\lambda_2 = \frac{\text{коэффициент при } s^{N-2} \text{ в разложении } f^{N-2}(s)\,[f'(s)]^2}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)}.$$

Далее, умножим выражение (7.10) на $u_m^{(q)} u_n^{(q')}$ $(q, q' = 1, 2, \ldots, p)$ и просуммируем по всем m и n. Получим

$$\sum_{m=1}^{p} \sum_{n=1}^{p} u_m^{(q)} u_n^{(q')} \sum_{\mathbf{k}} P_{\mathbf{i},\,\mathbf{k}} \left(k_m k_n + \mathfrak{L}_{m,\,n}(\mathbf{k}) \right),$$

или, меняя порядок суммирования,

$$\sum_{\mathbf{k}} P_{\mathbf{i},\,\mathbf{k}} \left[L_q(\mathbf{k})\, L_{q'}(\mathbf{k}) + \mathfrak{L}^*(\mathbf{k}) \right]. \qquad (7.12)$$

Выполняя аналогичные операции над выражением (7.11), получим

$$\lambda_2 \left[\sum_{\nu=1}^{p} i_\nu \sum_{m=1}^{p} \alpha_{\nu m} u_m^{(q)} \right] \left[\sum_{\nu=1}^{p} i_\nu \sum_{n=1}^{p} \alpha_{\nu n} u_n^{(q')} \right] + \sum_{m=1}^{p} \sum_{n=1}^{p} u_m^{(q)} u_n^{(q')} L_{m,\,n}(\mathbf{i}), \qquad (7.13)$$

где $L_{m,\,n}(\mathbf{i})$ при всех m и n — линейная функция от $\mathbf{i}$. Мы снова воспользуемся тем фактом, что $\sum_{m=1}^{p} \alpha_{\nu m} u_m^{(q)} = \gamma_q u_\nu^{(q)}$ по определению собственных векторов матрицы $\Gamma = \| \alpha_{\nu\mu} \|$. Выражение (7.13) тогда примет вид

$$\lambda_2 \gamma_q \gamma_{q'} \left(\sum_{\nu=1}^{p} u_\nu^{(q)} i_\nu \right) \left(\sum_{\nu=1}^{p} u_\nu^{(q')} i_\nu \right) + \sum_{n=1}^{p} \sum_{m=1}^{p} u_m^{(q)} u_n^{(q')} L_{m,\,n}(\mathbf{i}) =$$

$$= \lambda_2 \gamma_q \gamma_{q'} L_q(\mathbf{i})\, L_{q'}(\mathbf{i}) + L^*(\mathbf{i}),$$

где $L^*(\mathbf{i}) = \sum_{m=1}^{p} \sum_{n=1}^{p} u_m^{(q)} u_n^{(q')} \cdot L_{m,\,n}(\mathbf{i})$ есть линейная функция от $\mathbf{i} = (i_1, i_2, \ldots, i_p)$.

Результатом этих вычислений является равенство

$$\sum_{\mathbf{k}} P_{\mathbf{i},\,\mathbf{k}} \left[L_q(\mathbf{k})\, L_{q'}(\mathbf{k}) + \mathfrak{L}^*(\mathbf{k}) \right] = \lambda_2 \gamma_q \gamma_{q'} L_q(\mathbf{i}) + L^*(\mathbf{i}) \quad \text{при всех } \mathbf{i}.$$

Перенося $\sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} \mathfrak{L}^*(\mathbf{k})$ в правую часть, получим

$$\sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} L_q(\mathbf{k}) L_{q'}(\mathbf{k}) = \lambda_2 \gamma_q \gamma_{q'} L_q(\mathbf{i}) L_{q'}(\mathbf{i}) + \tilde{\mathfrak{L}}(\mathbf{i}), \tag{7.14}$$

где $\tilde{\mathfrak{L}}(\mathbf{i}) = L^*(\mathbf{i}) - \sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} \mathfrak{L}^*(\mathbf{k})$ — линейная функция от $\mathbf{i}$ (в силу (7.9)).

Для простоты предположим (на самом деле это не существенно), что $0 \neq \lambda_1 \gamma_q \neq \lambda_2 \gamma_q \gamma_{q'} \neq 0$ при всех q, q' $(2 \leqslant q, q' \leqslant p)$. Мы не будем рассматривать случай $q = 1$, так что $L_q(\mathbf{k}) L_{q'}(\mathbf{k})$ — квадратичные функции переменных $(k_1, k_2, \ldots, k_p)$.

Если бы в равенстве (7.14) не было линейного члена $\tilde{\mathfrak{L}}(\mathbf{i})$, то это было бы соотношение, определяющее собственное значение.

Мы утверждаем, что $\lambda_2 \gamma_q \gamma_{q'}$ — собственное значение матрицы (7.1). Покажем, как построить собственный вектор, соответствующий значению $\lambda_2 \gamma_q \gamma_{q'}$.

Предположим, что существует собственный вектор следующего вида:

$$a_{qq'}(i_1, i_2, \ldots, i_p) = L_q(\mathbf{i}) L_{q'}(\mathbf{i}) + K(\mathbf{i}),$$

где форма $K(\mathbf{i})$ линейна по $\mathbf{i}$ и пока произвольна. Вектор $a_{qq'}$ — нетривиальный, поскольку квадратичный член не может быть тождественно равным линейному члену. Представим $K(\mathbf{i})$ в виде $\sum_{k=1}^{p} b_k L_k(\mathbf{i})$, что может быть сделано, поскольку $L_k(\mathbf{i})$ $(k = 1, 2, \ldots, p)$ образуют полную систему линейно независимых линейных собственных векторов (см. также (7.8)).

Следующий метод построения собственных векторов аналогичен методу, использованному при нахождении собственных векторов матрицы перехода для модели с двумя типами индивидуумов при отсутствии мутации (см. § 6). В вышеприведенное выражение для собственных векторов входят неизвестные коэффициенты b_k. Для того чтобы $a_{qq'}(\mathbf{i})$ был собственным вектором, нужно, чтобы удовлетворялось равенство

$$\sum_{\mathbf{k}} P_{\mathbf{i},\mathbf{k}} \left\{ L_q(\mathbf{k}) L_{q'}(\mathbf{k}) + \sum_{\nu=1}^{p} b_\nu L_\nu(\mathbf{k}) \right\} =$$
$$= \lambda_2 \gamma_q \gamma_{q'} \left[L_q(\mathbf{i}) L_{q'}(\mathbf{i}) + \sum_{k=1}^{q} b_k L_k(\mathbf{i}) \right] \quad \text{для всех } \mathbf{i}. \tag{7.15}$$

Из равенств (7.7) и (7.14) следует, что левая часть (7.15) равна

$$\lambda_2 \gamma_q \gamma_{q'} L_q(\mathbf{i}) L_{q'}(\mathbf{i}) + \tilde{\mathfrak{L}}(\mathbf{i}) + \lambda_1 \sum_{k=1}^{p} b_k \gamma_k L_k(\mathbf{i}).$$

Приравняем это выражение правой части (7.15). Приведя подобные члены, получим

$$\tilde{\mathfrak{L}}(\mathbf{i})+\lambda_1\sum_{k=1}^{p} b_k\gamma_k L_k(\mathbf{i})=\lambda_2\gamma_q\gamma_{q'}\sum_{k=1}^{p} b_k L_k(\mathbf{i}) \quad \text{для всех } \mathbf{i}. \tag{7.16}$$

Так как $\tilde{\mathfrak{L}}(\mathbf{i})$ линейна по $\mathbf{i}$, можно записать

$$\tilde{\mathfrak{L}}(\mathbf{i})=\sum_{k=1}^{p} c_k L_k(\mathbf{i}),$$

где коэффициенты c_k определяются единственным образом из вида $\tilde{\mathfrak{L}}(\mathbf{i})$ и являются известными величинами. Таким образом,

$$\sum_{k=1}^{p}[\lambda_1 b_k\gamma_k+c_k-\lambda_2\gamma_q\gamma_{q'}b_k]\,L_k(\mathbf{i})=0;$$

поскольку $L_k(\mathbf{i})$, $k=1, \ldots, p$, линейно независимы, то

$$\lambda_1 b_k\gamma_k+c_k-\lambda_2\gamma_q\gamma_{q'}b_k=0, \qquad k=1,\ 2,\ \ldots,\ p,$$

откуда

$$b_k=\frac{c_k}{\lambda_2\gamma_q\gamma_{q'}-\lambda_1\gamma_k}, \qquad k=1,\ 2,\ \ldots,\ p.$$

Собственный вектор может быть теперь записан в виде

$$a_{qq'}(\mathbf{i})=L_q(\mathbf{i})\,L_{q'}(\mathbf{i})+\sum_{k=1}^{p}\frac{c_k}{\lambda_2\gamma_q\gamma_{q'}-\lambda_1\gamma_k}L_k(\mathbf{i}), \tag{7.17}$$

где $q,\ q'=2,\ 3,\ \ldots,\ p$.

Мы доказали, что каждому собственному значению $\lambda_2\gamma_q\gamma_{q'}$ соответствует собственный вектор $a_{qq'}(\mathbf{i})$. Все собственные векторы $a_{qq'}(\mathbf{i})$ линейно независимы, поскольку квадратичные части $L_q(\mathbf{i})\,L_{q'}(\mathbf{i})$, очевидно, линейно независимы.

Общий случай. Рассмотрим теперь общий случай. Дифференцируя равенство (7.3) r_i раз по t_i $(i=1,\ 2,\ \ldots,\ p)$ и полагая затем $t_1=t_2=\ldots=t_p=1$, получим

$$\sum_{\mathbf{k}} P_{\mathbf{i},\,\mathbf{k}}(k_1)_{r_1}(k_2)_{r_2}\ldots(k_p)_{r_p}=$$
$$=\lambda_R\left(\sum_{l=1}^{p} i_l\alpha_{l1}\right)^{r_1}\left(\sum_{l=1}^{p} i_l\alpha_{l2}\right)^{r_2}\ldots\left(\sum_{l=1}^{p} i_l\alpha_{lp}\right)^{r_p}+$$
$$+\text{полином от } (i_1,\ i_2,\ \ldots,\ i_p) \text{ степени } <R, \tag{7.18}$$

где $R=\sum_{i=1}^{p} r_i$,

$$\lambda_R=\frac{\text{коэффициент при } s^{N-R} \text{ в разложении } f^{N-R}(s)\,[f'(s)]^R}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)},$$

а $(k_l)_m$ определяется так:

$$(k_l)_m = k_l(k_l - 1) \ldots (k_l - m + 1).$$

Первый член в правой части равенства (7.18) является однородным полиномом от $(i_1, i_2, \ldots, i_p)$ степени R. Он получен при однократном дифференцировании правой части (7.3). Как только при дифференцировании появляется член $f''(s)$, степень коэффициента при нем относительно $\mathbf{i}$ обязана быть меньше R, поскольку для того, чтобы получить $f''(s)$, нужно продифференцировать $f'(s)$, а $f'(s)$ не возведено в степень i_ν $(\nu = 1, 2, \ldots, p)$. Используя равенство (7.18), можно найти собственные значения матрицы переходных вероятностей (7.1).

Первое собственное значение равно 1 и соответствующий ему собственный вектор имеет равные между собой компоненты. Причиной этому служит то, что сумма элементов строк матрицы переходных вероятностей равна 1. При $R = 1$ мы имеем случай 1, а при $R = 2$ — случай 2, рассмотренные выше. В случае 1 мы получили $\binom{p-2}{0} + \binom{p-1}{1}$ линейно независимых собственных векторов. Член $\binom{p-2}{0} = 1$ отвечает собственному значению 1. Член $\binom{p-1}{1} = p - 1$ отвечает собственным значениям $\lambda_1\gamma_q$, $q = 2, 3, \ldots, p$. В случае 2 было $\frac{p}{2}$ линейно независимых собственных векторов, соответствующих $\frac{1}{2}p(p-1)$ собственным значениям вида $\lambda_2\gamma_q\gamma_{q'}$, $q, q' = 2, \ldots, p$.

Процедура, аналогичная использованной в случаях 1 и 2, может быть применена для построения $\binom{r+p-2}{r}$ линейно независимых собственных векторов, которые являются полиномами степени r от переменных

$$\left(\sum_{\nu=1}^{p} u_\nu^{(1)} i_\nu\right) \equiv N, \quad L_2(\mathbf{i}) = \sum_{\nu=1}^{p} u_\nu^{(2)} i_\nu, \ \ldots,$$
$$L_q(\mathbf{i}) = \sum_{\nu=1}^{p} u_\nu^{(q)} i_\nu, \ \ldots, \quad L_p(\mathbf{i}) = \sum_{\nu=1}^{p} u_\nu^{(p)} i_\nu$$

и соответствуют собственным значениям

$$\lambda_r\gamma_{q_1}\gamma_{q_2} \ldots \gamma_{q_r}, \quad 2 \leqslant q_1 \leqslant q_2 \leqslant \ldots \leqslant q_r \leqslant p. \tag{7.19}$$

Точное выражение для собственных векторов имеет вид

$$a_{q_1, q_2, \ldots, q_r}(i_1, i_2, \ldots, i_p) = \prod_{k=1}^{r} (L_{q_k}(\mathbf{i})) +$$
$$+ \text{полином от } (i_1, \ldots, i_p) \text{ степени } < r. \tag{7.20}$$

Число собственных значений вида $\lambda_r\gamma_{q_1}\gamma_{q_2}\ldots\gamma_{q_r}$, $q_\nu=2, \ldots, p$, можно найти, рассматривая соответствующую задачу о размещении, где k шаров нужно разместить в n ячеек. В нашем случае k равно степени r соответствующего собственного вектора, а n равно числу различных чисел γ_q ($q = 2, \ldots, p$), т. е. $p-1$. При такой интерпретации число различных собственных значений равно числу сочетаний из $r+p-2$ по r. Поэтому число различных собственных векторов вида

$$\lambda_r\gamma_{q_1}\gamma_{q_2}\ldots\gamma_{q_r}, \qquad q_i=2, \ldots, p, \tag{7.21}$$

равно $\binom{r+p-2}{r}=\binom{r+p-2}{p-2}$.

Общее число линейно независимых собственных векторов равно

$$\binom{p-2}{0}+\binom{p-1}{1}+\binom{p}{2}+\ldots+\binom{r+p-2}{r}+\ldots+\binom{N+p-2}{N}=$$
$$=\binom{N+p-1}{N},$$

что в точности совпадает с числом состояний цепи Маркова (7.1). Подытожим полученные результаты в следующей теореме.

Теорема 7.1. *Пусть* $\mathbf{\Gamma}=\|\alpha_{\nu\mu}\|_{\nu,\mu=1}^{p}$ — *стохастическая матрица* $\left(\text{т. е. } \alpha_{\nu\mu}\geqslant 0,\ \sum_{\mu}\alpha_{\nu\mu}=1,\ \nu=1, 2, \ldots, p\right)$, *соответствующая неприводимой цепи Маркова. Пусть* $\gamma_1=1, \gamma_2, \ldots, \gamma_p$ — *ее собственные значения, и предположим, что* $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(p)}$ *образуют полное семейство соответствующих им собственных векторов, т. е.* $\mathbf{\Gamma}$ *диагонализируема. Более того, предположим, что* $\lambda_r\gamma_{q_1}\gamma_{q_2}\ldots\gamma_{q_r}\neq$ $\neq\lambda_{r+1}\gamma_{q'_1}\gamma_{q'_2}\ldots\gamma_{q'_{r+1}}$ *для любых* γ_{q_k} *и* $\gamma_{q'_k}$ $(2\leqslant q_k, q'_k\leqslant p)$, $r\geqslant 1$, *где*

$$\lambda_r=\frac{\text{коэффициент при } s^{N-r} \text{ в разложении } f^{N-r}(s)\,(f'(s))^r}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)}. \tag{7.22}$$

Рассмотрим цепь Маркова для p типов индивидуумов с матрицей переходных вероятностей, задаваемой равенством (7.1). *При* $r\geqslant 1$ *существует* $\binom{r+p-2}{r}$ *линейно независимых собственных векторов, которые получаются из полиномов степени r от переменных*

$$L_q(\mathbf{i})=\sum_{\nu=1}^{p}u_\nu^{(q)}i_\nu, \qquad q=2, \ldots, p.$$

Вид собственных значений дается формулой (7.21), *а правые собственные векторы имеют вид* (7.20). *На эти собственные векторы* *и постоянный вектор натянуто пространство размерности* $\binom{N+p-1}{N}$, *равной порядку матрицы* $\|P_{\mathrm{i,\,k}}\|$.

В качестве примера применим теорему 7.1 к случаю, когда $\mathbf{\Gamma}$ — единичная матрица. Другими словами, нет мутационного давления. В этом случае цепь Маркова (7.1) является непосредственным обобщением модели с двумя типами индивидуумов на случай p типов $A_1, A_2, \ldots, A_p$. Каждый из них размножается независимо в соответствии с законами развития ветвящегося процесса, характеризуемого вероятностной производящей функцией $f(s)$. Порожденная цепь Маркова имеет следующие вероятности перехода:

$$P_{\mathrm{i,\,j}} = \frac{\text{коэффициент при } s_1^{j_1} s_2^{j_2} \ldots s_p^{j_p} \text{ в разложении } \prod\limits_{\nu=1}^{p} f^{i_\nu}(s_\nu)}{\text{коэффициент при } t^N \text{ в разложении } f^N(t)}. \tag{7.23}$$

В этом случае $\gamma_1 = \gamma_2 = \ldots = \gamma_p = 1$. Пусть константа λ_r определена равенством (7.22). Теорема 7.1 утверждает, что λ_r является собственным значением кратности $\binom{r+p-2}{r}$ $(r = 0,\ 1,\ 2,\ \ldots,\ N)$, т. е. существуют $\binom{r+p-2}{r}$ линейно независимых собственных векторов, соответствующих собственному значению λ_r. Кратность возникает из-за того, что $\lambda_r \gamma_{q_1} \gamma_{q_2} \ldots \gamma_{q_r} = \lambda_r$ и не зависит от выбора $\gamma_{q_1} = \gamma_{q_2} = \ldots = \gamma_{q_r} = 1$. Правые собственные векторы, соответствующие λ_r, являются полиномами степени r от переменных

$$L_q(\mathbf{i}) = \sum_\nu u_\nu^{(q)} i_\nu = i_q \qquad (q = 2,\ 3,\ \ldots,\ p),$$

поскольку можно выбрать $u_\nu^{(q)} = \delta_{\nu q}$.

§ 8. ВЕРОЯТНОСТНЫЙ СМЫСЛ СОБСТВЕННЫХ ЗНАЧЕНИЙ

Мы можем теперь, используя теорему 7.1, дать вероятностную трактовку собственным значениям

$$\lambda_0 = 1, \quad \lambda_r = \frac{\text{коэффициент при } s^{N-r} \text{ в разложении } f^{N-r}(s)\,[f'(s)]^r}{\text{коэффициент при } s^N \text{ в разложении } f^N(s)} \tag{8.1}$$

и соответствующим собственным векторам матрицы переходных вероятностей (7.1), которая описывает процесс размножения p типов индивидуумов в отсутствие мутации. Напомним, что в этом случае собственное значение λ_r имеет кратность $\binom{r+p-2}{r}$.

Начнем обсуждение со случая $p = 2$. В этом случае все собственные значения, отличные от $\lambda_0 = \lambda_1 = 1$, имеют кратность 1.

Из теоремы 6.3 мы знаем, что матрица (7.1) диагонализируема. Используя множество попарно ортогональных собственных векторов, можно записать

$$P_{ij}^t = \sum_{r=0}^{N} \lambda_r^t \varphi_r(i) \psi_r(j), \tag{8.2}$$

где φ_r — r-й правый собственный вектор, а ψ_r — r-й левый собственный вектор, упоминаемые в теореме 6.3. (Здесь t — неотрицательное целое число, а φ_r и $\psi_{r'}$ попарно ортогональны.) Целесообразно выделить два члена в (8.2), соответствующих $\lambda_0 = \lambda_1 = 1$. Получим

$$P_{ij}^t = \varphi_0(i)\psi_0(j) + \varphi_1(i)\psi_1(j) + \sum_{r=2}^{N} \lambda_r^t \varphi_r(i)\psi_r(j). \tag{8.3}$$

Из соотношений (6.18) и (6.19) имеем

$$\begin{gathered} \psi_0(j) = \delta_{0j}, \qquad \psi_1(j) = \delta_{Nj}, \\ \varphi_0(i) = \frac{N-i}{N}, \quad \varphi_1(i) = \frac{i}{N}. \end{gathered} \tag{8.4}$$

Сумма в правой части (8.3) стремится к нулю при $t \to \infty$, как λ_2^t ($\lambda_2 < 1$). Более того, из выражений для $\psi_0(j)$ и $\psi_1(j)$ следует, что при $0 < j < N$ первые два члена в правой части (8.3) равны нулю. Отсюда

$$\lim_{t\to\infty} P_{ij}^t = \varphi_0(i)\psi_0(j) + \varphi_1(i)\psi_1(j) = \begin{cases} 0, & j \neq 0, N, \\ \dfrac{N-i}{N}, & j = 0, \\ \dfrac{i}{N}, & j = N, \end{cases} \tag{8.5}$$

и имеет место экспоненциальная сходимость (с параметром λ_2)[1]), т. е. λ_2 — «скорость приближения к гомозиготному состоянию». Аналогично вероятность того, что система не находится в гомозиготном состоянии (0 или N), имеет порядок λ_2^t при $t \to \infty$. Далее, так как $\lambda_3 < \lambda_2$, то

$$\lim_{t\to\infty} \frac{P_{ij}^t}{\lambda_2^t} = \varphi_2(i)\psi_2(j), \qquad i, j \neq 0, N. \tag{8.6}$$

Легко показать, что $\varphi_2(i) \not\equiv 0$ при $0 < i < N$. В самом деле, в противном случае было бы $\varphi_2(i) \equiv 0$ при $0 < i < N$. Но всегда $\varphi_2(0) = 0 = \varphi_2(N)$, поскольку φ_2 ортогонален ψ_0 и ψ_1. Таким образом, $\varphi_2(i) \equiv 0$ при $0 \leqslant i \leqslant N$, что невозможно, поскольку φ_2 — собственный вектор, соответствующий значению λ_2.

---

[1]) То есть $\left| P_{ij}^t - \varphi_0(i)\psi_0(j) + \varphi_1(i)\psi_1(j) \right| \sim \lambda_2^t$. — *Прим. перев.*

Нормализуя φ_2 так, чтобы $\varphi_2(i^0) > 0$ при некотором i^0, мы заключаем, используя (8.6), что $\psi_2(j) \geqslant 0$, $0 < j < N$. Мы утверждаем, что $\psi_2(j) \neq 0$, $0 < j < N$. Доказательство аналогично уже использованному при выводе утверждения, что $\varphi_2(i) \neq 0$. Мы доказали, что $\sum_{j=1}^{N-1} \psi_2(j) > 0$. Можно показать, что $\psi_2(j) > 0$, $0 < j < N$.

В самом деле, поскольку $0 < a_0 a_1 a_2$ (где $f(s) = \sum a_n s^n$), то переходные состояния $T = \{1, 2, \ldots, N-1\}$ все сообщающиеся, т. е. начиная из любого состояния из T возможно (с положительной вероятностью) достичь любого другого состояния из T. Отсюда следует, что показатель, с которым P_{ij}^t стремится к нулю (при $t \to \infty$, $i, j \in T$), не зависит от выбора i и j в T. (Читателю предлагается строго доказать это.)

Мы уже доказали, что при некоторых $i = i_0$ и $j = j_0$ $\varphi_2(i_0)\psi_2(j_0) > 0$. Но

$$P_{ij}^t \sim \lambda_2^t \varphi_2(i)\psi_2(j), \qquad i, j \in T.$$

Следовательно, $\varphi_2(i)\psi_2(j) > 0$ при всех i, $j \in T$, откуда получается, что $\varphi_2(i)$ не меняет знака при $i \in T$. Это же справедливо для $\psi_2(j)$, $j \in T$. Мы условились, что после нормализации $\psi_2(j) > 0$ при всех $j \in T$.

Заметим, что $\psi_2(0) < 0$ и $\psi_2(N) < 0$, поскольку вектор ψ_2 ортогонален φ_0 и φ_1. Из соотношения (8.6) можно теперь получить, что предельное условное распределение состояния j при условии, что $j \neq 0, N$, равно

$$\frac{\psi_2(j)}{\sum_{j=1}^{N-1} \psi_2(j)}. \tag{8.7}$$

Обратимся теперь к вероятностной интерпретации собственных значений и векторов в случае цепи Маркова для трех типов индивидуумов, матрица переходных вероятностей которой задается соотношением (7.23) при $p = 3$. Обозначим симплекс в пространстве состояний через Δ_3, а его ребра — через E_1, E_2 и E_3, т. е.

$$\Delta_3 = \{\mathbf{i} = (i_1, i_2, i_3) \mid i_1 \geqslant 0,\ i_2 \geqslant 0,\ i_3 \geqslant 0,\ i_2 + i_2 + i_3 = N,\ i_\nu - \text{целые числа}\},$$

а E_k состоит из всевозможных векторов $\mathbf{i} \in \Delta_3$, для которых $i_k = 0$ $(k = 1, 2, 3)$. Кратность собственного значения λ_r равна $\binom{r+p-2}{r} = \binom{r+1}{r} = r + 1$. Запишем соответствующие правые и левые собственные векторы в виде

$$\alpha_{rk} = \alpha_{rk}(i_1, i_2, i_3), \quad \beta_{rk} = \beta_{rk}(i_1, i_2, i_3), \qquad k = 0, 1, \ldots, r, \tag{8.8}$$

где (i_1, i_2, i_3) — компоненты векторов; здесь всюду $i_1, i_2, i_3 \geqslant 0$, $i_1 + i_2 + i_3 = N$. Векторы (8.8) в силу построения попарно ортогональны.

Заметим, что собственное значение $1 = \lambda_0 = \lambda_1$ имеет кратность 3 (λ_0 — однократный корень и λ_1 — двукратный), как и следовало ожидать, поскольку существует три поглощающих состояния $(N, 0, 0)$, $(0, N, 0)$ и $(0, 0, N)$. Левые собственные векторы, соответствующие этим собственным значениям, совпадают соответственно со стационарными распределениями каждого из возвратных классов и, следовательно, имеют вид

$$\beta_{00}(\mathbf{i}) = \begin{cases} 1, & \text{если} \quad i_1 = N, \\ 0 & \text{в противном случае}, \end{cases}$$

$$\beta_{10}(\mathbf{i}) = \begin{cases} 1, & \text{если} \quad i_2 = N, \\ 0 & \text{в противном случае}, \end{cases}$$

$$\beta_{11}(\mathbf{i}) = \begin{cases} 1, & \text{если} \quad i_3 = N, \\ 0 & \text{в противном случае}. \end{cases}$$

Соответствующие правые собственные векторы (они являются линейными функциями от $\mathbf{i}$), которые попарно ортогональны с $\{\beta_{00}, \beta_{10}, \beta_{11}\}$, равны соответственно

$$\alpha_{00}(\mathbf{i}) = \frac{i_1}{N}, \quad \alpha_{10}(\mathbf{i}) = \frac{i_2}{N}, \quad \alpha_{11}(\mathbf{i}) = \frac{i_3}{N}.$$

О левых собственных векторах можно сказать больше. Заметим, что если начальное состояние цепи Маркова принадлежит ребру E_1, то все последующие состояния также принадлежат E_1. Интуитивно это очевидно, поскольку принадлежность начального состояния ребру E_1 означает, что индивидуумы типа 1 отсутствуют и не могут быть произведены с помощью индивидуумов типа 2 и 3. Непосредственной проверкой нетрудно найти, что в матрице переходных вероятностей (7.23)

$$P_{\mathbf{i},\mathbf{j}} = 0, \quad \text{если} \quad \mathbf{i} \in E_1, \ \mathbf{j} \notin E_1. \tag{8.9}$$

Более того, если и $\mathbf{i}$ и $\mathbf{j}$ принадлежат E_1, то матрица $\| P_{\mathbf{i},\mathbf{j}} \|$ сводится к матрице цепи Маркова для индивидуумов двух типов 2 и 3.

Рассмотрим теперь вектор с компонентами

$$\beta_{20}(\mathbf{i}) = \begin{cases} \psi_2(i_2), & \mathbf{i} \in E_1, \\ 0, & \mathbf{i} \notin E_1, \end{cases} \tag{8.10}$$

где ψ_2 — второй левый собственный вектор матрицы (6.2). По определению ψ_2 имеем

$$\sum_{\mathbf{i} \in E_1} \psi_2(i_2) P_{\mathbf{i},\mathbf{j}} = \lambda_2 \psi_2(i_2), \qquad \mathbf{j} \in E_1, \tag{8.11}$$

где $\mathbf{i} = (i_1, i_2, i_3)$, $\mathbf{j} = (j_1, j_2, j_3)$. Принимая во внимание (8.9) и (8.11), получим

$$\sum_{\mathbf{i}} \beta_{20}(\mathbf{i}) P_{\mathbf{i},\,\mathbf{j}} = \lambda_2 \beta_{20}(\mathbf{j}),$$

где $\beta_{20}(\mathbf{i})$ определяется равенством (8.10).

Аналогично пусть

$$\beta_{21}(\mathbf{i}) = \begin{cases} \psi_2(i_3), & \mathbf{i} \in E_2, \\ 0, & \mathbf{i} \notin E_2, \end{cases}$$
$$\beta_{22}(\mathbf{i}) = \begin{cases} \psi_2(i_1), & \mathbf{i} \in E_3, \\ 0, & \mathbf{i} \notin E_3. \end{cases} \tag{8.12}$$

Это также левые собственные векторы, соответствующие собственному значению λ_2. Соотношения (8.10) и (8.12) показывают, что нулевые значения векторов β_{20}, β_{21} и β_{23} сосредоточены в различных ребрах E_1, E_2 и E_3 соответственно, и поэтому они, очевидно, линейно независимы.

Вообще для каждого левого собственного вектора ψ_k (см. (6.19)) цепи Маркова в случае двух типов можно построить три левых собственных вектора матрицы (7.23), соответствующих собственному значению λ_k. Аналогично построениям (8.10) и (8.12) можно взять

$$\begin{aligned} \beta_{k0} &= \psi_k \text{ на } E_1 \text{ и } 0 \text{ в противном случае,} \\ \beta_{k1} &= \psi_k \text{ на } E_2 \text{ и } 0 \text{ в противном случае,} \\ \beta_{k2} &= \psi_k \text{ на } E_3 \text{ и } 0 \text{ в противном случае.} \end{aligned} \tag{8.13}$$

Эти векторы линейно независимы, поскольку их ненулевые значения сосредоточены в различных ребрах Δ_3.

Из общей теории следует, что собственному значению λ_3 соответствует четыре собственных вектора. Векторы β_{30}, β_{31}, β_{32}, задаваемые равенством (8.13), — три из них, и все они равны нулю при $\mathbf{i} \in \overset{0}{\Delta}_3$ ($\overset{0}{\Delta}_3$ — внутренность Δ_3). Мы утверждаем, что $\beta_{33}(\mathbf{i})$ не может тождественно равняться нулю в $\overset{0}{\Delta}_3$.

Формальное доказательство этого следующее. Каждый собственный вектор (для случая двух типов) ψ_r ($r = 2, \ldots, N$) определяет собственный вектор (для случая трех типов), ненулевые компоненты которого соответствуют точкам на E_1. Аналогично мы получаем $N-1$ линейно независимых собственных векторов, ненулевые компоненты которых сосредоточены в E_2, а также $N-1$ собственных векторов, соответствующих E_3. Наконец, мы имеем векторы β_{00}, β_{10} и β_{11}, ненулевые значения которых сосредоточены в вершинах Δ_3. В целом это дает $3N$ линейно независимых векторов, ненулевые компоненты которых соответствуют границе Δ_3. Обозначим множество этих собственных векторов через V

На границе Δ_3 имеется ровно $3N$ состояний. Следовательно, на перечисленные собственные векторы натянуто линейное пространство всех векторов, имеющих ненулевые координаты только на границе Δ_3.

Далее, вектор β_{33} не равняется ψ_r (при некотором r), поскольку он соответствует собственному значению λ_3, а мы уже установили соответствие между $(\beta_{30}, \beta_{31}, \beta_{32})$ и ψ_3. Если бы $\beta_{33}(\mathbf{j}) = 0$ при $\mathbf{j} \in \Delta_3^0$, то β_{33} линейно зависел бы от перечисленных векторов. Но это невозможно, поскольку β_{33} по построению не зависит от β_{30}, β_{31} и β_{32} и, очевидно, не зависит от остальных векторов из V, так как они соответствуют собственным значениям, отличным от λ_3.

Выше доказано, что $\beta_{33}(\mathbf{j}) \not\equiv 0$ при $\mathbf{j} \in \Delta_3^0$. Аналогично $\alpha_{33}(\mathbf{i}) \not\equiv 0$ при $\mathbf{i} \in \Delta_3^0$. Действительно, предположим противное, что $\alpha_{33}(\mathbf{i}) \equiv 0$ при $\mathbf{i} \in \Delta_3^0$. Тогда, поскольку вектор α_{33} ортогонален любому вектору из V, получаем, что $\alpha_{33} = 0$. То есть $\alpha_{33}(\mathbf{i}) \in 0$ при $\mathbf{i} \in \Delta_3$, что противоречит определению α_{33}. В самом деле, отсюда следует, что значения $\alpha_{33}(\mathbf{i})$ с необходимостью равны нулю на границе Δ_3.

Матрицу перехода можно представить в виде

$$P_{\mathbf{i},\mathbf{j}}^t = \sum_{r=0}^{N} \lambda_r^t \sum_{k=0}^{r} \alpha_{rk}(\mathbf{i}) \beta_{rk}(\mathbf{j}), \qquad \mathbf{i} = (i_1, i_2, i_3), \quad \mathbf{j} = (j_1, j_2, j_3). \tag{8.14}$$

Перепишем (8.14), выделяя члены, включающие первые три собственных значения:

$$\begin{aligned} P_{\mathbf{i},\mathbf{j}}^t &= \alpha_{00}(\mathbf{i})\beta_{00}(\mathbf{j}) + \alpha_{10}(\mathbf{i})\beta_{10}(\mathbf{j}) + \alpha_{11}(\mathbf{i})\beta_{11}(\mathbf{j}) + \\ &+ \lambda_2^t\left[\alpha_{20}(\mathbf{i})\beta_{20}(\mathbf{j}) + \alpha_{21}(\mathbf{i})\beta_{21}(\mathbf{j}) + \alpha_{22}(\mathbf{i})\beta_{22}(\mathbf{j})\right] + \sum_{r=3}^{N} \lambda_r^t \left(\sum_{k=0}^{r} \alpha_{rk}(\mathbf{i})\beta_{rk}(\mathbf{j})\right). \end{aligned} \tag{8.15}$$

Можно дать теперь обещанную ранее вероятностную интерпретацию собственных значений и собственных векторов, используя при этом специальный вид векторов β_{00}, β_{10}, β_{01}, β_{20}, β_{21}, β_{22}.

(а) Скорость поглощения (или фиксации, или приближения к гомозиготному состоянию) в вершинах равна λ_2, поскольку для $\mathbf{j}$, не совпадающих с вершиной Δ_3, выражение для $P_{\mathbf{i},\mathbf{j}}^t$ в (8.15) сводится к сумме от $r = 2$ и далее.

Распределение для больших t при условии, что фиксация не происходит и что первый тип не исчезает из популяции, пропорционально $\beta_{20}(\mathbf{j})$, где $\beta_{20}(\mathbf{j})$ — собственный вектор с ненулевыми компонентами, сосредоточенными в ребре, отвечающем только второму и третьему типам. Собственные векторы $\beta_{21}(\mathbf{j})$ и $\beta_{22}(\mathbf{j})$ имеют аналогичную интерпретацию. Доказательство этого повторяет по существу доказательство (8.7).

(б) Скорость поглощения в ребрах (т. е. скорость, с которой один из типов, все равно какой, выбывает из популяции) равна λ_3. Фактически при $\mathbf{i}, \mathbf{j} \in \Delta_3^0$ (внутренность Δ_3)

$$P_{\mathbf{i}, \mathbf{j}}^t = \sum_{r=3}^{N} \lambda_r^t \left(\sum_{k=0}^{r} \alpha_{rk}(\mathbf{i}) \beta_{rk}(\mathbf{j}) \right).$$

Доминирующим членом является первый, поскольку коэффициенты при λ_3^t, в частности $\alpha_{33}(\mathbf{i})\beta_{33}(\mathbf{j})$, отличны от нуля в некоторых точках Δ_3^0, в то время как все ненулевые компоненты векторов $\beta_{30}(\mathbf{j})$, $\beta_{31}(\mathbf{j})$, $\beta_{32}(\mathbf{j})$ сосредоточены на границе Δ_3. В действительности же $\alpha_{33}(\mathbf{i})\beta_{33}(\mathbf{j}) > 0$ при всех $\mathbf{i}, \mathbf{j} \in \Delta_3^0$. В самом деле, мы знаем, что все состояния из Δ_3^0 являются сообщающимися. Более того, уже доказано, что $\alpha_{33}(\mathbf{i}_0)\beta_{33}(\mathbf{j}_0) \neq 0$ при некоторых $\mathbf{i}_0, \mathbf{j}_0 \in \Delta_3^0$, и поскольку

$$0 \leqslant P_{\mathbf{i}, \mathbf{j}}^t \sim \lambda_3^t \beta_{33}(\mathbf{j}) \alpha_{33}(\mathbf{i}) \quad \text{при} \quad t \to \infty \quad \text{для всех} \quad \mathbf{i}, \mathbf{j} \in \Delta_3^0,$$

то $\alpha_{33}(\mathbf{i}_0)\beta_{33}(\mathbf{j}_0) \geqslant 0$. Но скорость сходимости к нулю с необходимостью одна и та же для всех состояний $\mathbf{i}, \mathbf{j} \in \Delta_3^0$, поскольку они сообщаются. Следовательно,

$$\alpha_{33}(\mathbf{i})\beta_{33}(\mathbf{j}) > 0, \qquad \mathbf{i}, \mathbf{j} \in \Delta_3^0.$$

Это означает, что $\beta_{33}(\mathbf{j})$ и $\alpha_{33}(\mathbf{i})$ не меняют знака при $\mathbf{i}, \mathbf{j} \in \Delta_3^0$. Мы можем, следовательно, без потери общности выбрать $\beta_{33}(\mathbf{j}) > 0$ и $\alpha_{33}(\mathbf{i}) > 0$ при всех $\mathbf{i} \in \Delta_3^0$, $\mathbf{j} \in \Delta_3^0$.

Предполагая эти свойства выполненными для β_{33} и α_{33}, можно утверждать, что условное распределение вектора состояния $\mathbf{j}$ при больших t при условии, что все типы индивидуумов имеются в наличии, асимптотически равно

$$\frac{\beta_{33}(\mathbf{j})}{\sum\limits_{\mathbf{j} \in \Delta_3^0} \beta_{33}(\mathbf{j})}, \quad \mathbf{j} \in \Delta_3^0.$$

Вероятностный смысл правых собственных векторов очевиден. Так,

$\alpha_{00}(\mathbf{i})$ — вероятность поглощения в вершине $(N, 0, 0)$ при начальном состоянии $\mathbf{i}$,

$\alpha_{10}(\mathbf{i})$ — вероятность поглощения в вершине $(0, N, 0)$ при начальном состоянии $\mathbf{i}$,

$\alpha_{11}(\mathbf{i})$ — вероятность поглощения в вершине $(0, 0, N)$ при начальном состоянии $\mathbf{i}$.

Аналогичным образом можно дать интерпретацию и для векторов $\alpha_{20}(\mathbf{i})$, $\alpha_{21}(\mathbf{i})$ и $\alpha_{22}(\mathbf{i})$. Для этой цели рассмотрим предельную условную вероятность того, что поглощения в вершине не произойдет. Очевидно, отсюда можно получить распределение на ребрах Δ_3. При векторе $\mathbf{j}$, не являющемся вершиной, но лежащем на ребре Δ_3, очевидно,

$$P^t_{\mathbf{i},\mathbf{j}} \sim \lambda_2^t\left[\alpha_{20}(\mathbf{i})\beta_{20}(\mathbf{j}) + \beta_{21}(\mathbf{j})\alpha_{21}(\mathbf{i}) + \alpha_{22}(\mathbf{i})\beta_{22}(\mathbf{j})\right], \qquad t \to \infty. \tag{8.16}$$

Напомним, что

$$\beta_{20}(\mathbf{j}) = \begin{cases} \psi_2(\mathbf{j}), & \mathbf{j} \in E_1, \\ 0 & \text{в противном случае} \end{cases}$$

и аналогичные равенства имеют место и для других величин. Из (8.7) заключаем, что $\psi_2(\mathbf{j}) > 0$ всюду, кроме вершин. Следовательно,

$$\sum_{\mathbf{i},\,\mathbf{j} \neq \text{вершине}} \beta_{20}(\mathbf{j}) = \sum_{\mathbf{i},\,\mathbf{j} \neq \text{вершине}} \beta_{21}(\mathbf{j}) = \sum_{\mathbf{i},\,\mathbf{j} \neq \text{вершине}} \beta_{22}(\mathbf{j}) = \sum_{l} \psi_2(l) > 0.$$

Поскольку β_{20}, β_{21} и β_{22} — неотрицательные векторы, каждый из которых принимает ненулевые значения всюду, кроме вершин, принадлежащих различным ребрам, то $\alpha_{20}(\mathbf{i})$, $\alpha_{21}(\mathbf{i})$ и $\alpha_{22}(\mathbf{i})$ также неотрицательны при всех $\mathbf{i}$, не являющихся вершинами, а по крайней мере один из этих векторов положителен. Предельное условное распределение $P^t_{\mathbf{i},\mathbf{j}}$ $(t \to \infty)$ для $\mathbf{j}$, не являющихся вершинами, равно

$$\frac{\alpha_{20}(\mathbf{i})\beta_{20}(\mathbf{j}) + \alpha_{21}(\mathbf{i})\beta_{21}(\mathbf{j}) + \alpha_{22}(\mathbf{i})\beta_{22}(\mathbf{j})}{\left(\sum_{l} \psi_2(l)\right)[\alpha_{20}(\mathbf{i}) + \alpha_{21}(\mathbf{i}) + \alpha_{22}(\mathbf{i})]}. \tag{8.17}$$

Заметим, что для каждой пары $\mathbf{i}$ и $\mathbf{j}$ только одно слагаемое в числителе положительно, поскольку произведение любых двух $\beta_{20}(\mathbf{j})\beta_{21}(\mathbf{j}) = 0$ при $\mathbf{j}$, не являющемся вершиной. Вероятность того, что при начальном состоянии $\mathbf{i}$ поглощение произойдет на грани E_1, а не на E_2 или E_3 (вершины автоматически исключаются), получается суммированием выражения (8.17) по всевозможным индексам $\mathbf{j} \in E_1$, не соответствующим вершинам. Поскольку $\beta_{21}(\mathbf{j}) = \beta_{22}(\mathbf{j}) = 0$ при $\mathbf{j} \in E_1$, то отсюда следует, что

$$\mathrm{P}\{\text{поглощение произойдет в } E_1, \text{ а не в } E_2 \cup E_3 \mid X_0 = \mathbf{i}\} = \frac{\alpha_{20}(\mathbf{i})}{\alpha_{20}(\mathbf{i}) + \alpha_{21}(\mathbf{i}) + \alpha_{22}(\mathbf{i})}.$$

Аналогично,

$$\mathrm{P}\{\text{поглощение произойдет в } E_2, \text{ а не в } E_1 \cup E_3 \mid X_0 = \mathbf{i}\} = \frac{\alpha_{21}(\mathbf{i})}{\alpha_{20}(\mathbf{i}) + \alpha_{21}(\mathbf{i}) + \alpha_{22}(\mathbf{i})}.$$

И, наконец,

$$\mathrm{P}\{\text{поглощение произойдет в } E_3, \text{ а не в } E_1 \cup E_2 \mid X_0 = \mathbf{i}\} = \frac{\alpha_{22}(\mathbf{i})}{\alpha_{20}(\mathbf{i}) + \alpha_{21}(\mathbf{i}) + \alpha_{22}(\mathbf{i})}.$$

Это показывает, что $\alpha_{20}(\mathbf{i})$, $\alpha_{21}(\mathbf{i})$ и $\alpha_{22}(\mathbf{i})$ положительны при всех $\mathbf{i} \in \Delta_3^0$, $\alpha_{20}(\mathbf{i}) = 0$ при $\mathbf{i} \in E_2^0 \cup E_3^0$, $\alpha_{21}(\mathbf{i}) = 0$ при $\mathbf{i} \in E_1^0 \cup E_3^0$ и $\alpha_{22}(\mathbf{i}) = 0$ при $\mathbf{i} \in E_1^0 \cup E_2^0$ (верхний индекс 0 определяет внутренность рассматриваемого множества).

Все приведенные рассмотрения распространяются на общий случай нескольких типов. Мы приведем соответствующие результаты без доказательства. Доказательство их получается путем обобщения примененных выше методов.

Теорема 8.1. *Пусть* $\mathbf{P}$ — *матрица переходных вероятностей* (7.23) *порожденной цепи Маркова для* p *типов в отсутствие мутаций. Собственные значения* $\mathbf{P}$ *не зависят от числа типов и равны*

$$\lambda_0 = 1,$$

$$\lambda_r = \frac{\text{коэффициент при } s^{N-r} \text{ в разложении } [f(s)]^{N-r}[f'(s)]^r}{\text{коэффициент при } s^N \text{ в разложении } [f(s)]^N}, \quad r = 1, 2, \ldots, N.$$

Собственное значение λ_r *имеет кратность* $\binom{r+p-2}{r}$. *Пусть* Δ_n — *симплекс состояний цепи Маркова, т. е.* $\Delta_p = \left\{\mathbf{i} \mid \mathbf{i} = (i_1, i_2, \ldots, i_p),\right.$ $\left. i_\nu \geqslant 0 - \text{целые числа}, \sum_\nu i_\nu = N\right\}$.

Вероятностная интерпретация собственных значений следующая:

(1) *Скорость поглощения приближения к гомозиготному состоянию равна* λ_2, *т. е. если* $\mathbf{i}$ *и* $\mathbf{j}$ *не совпадают с вершинами, то* $P_{\mathbf{i},\mathbf{j}}^t \sim c_{\mathbf{i},\mathbf{j}}\lambda_2^t$ $(t \to \infty)$, *где* $c_{\mathbf{i},\mathbf{j}}$ — *постоянная, зависящая от* $\mathbf{i}$ *и* $\mathbf{j}$, *но не зависящая от* t.

(2) *Скорость, с которой в популяции вырождаются все, кроме* k *типов* $(k \leqslant p)$ *(все равно, каких именно), равна* λ_k, *т. е. вероятность того, что популяция в* t*-м поколении содержит по крайней мере* k *типов, имеет порядок* λ_k^t. *Другой способ выражения этого же свойства — это сказать, что скорость, с которой происходит поглощение в* $(k-1)$*-мерной границе множества* Δ_p, *равна* λ_k.

Более определенно, если $\mathbf{i}$ *и* $\mathbf{j}$ *принадлежат либо внутренности множества* Δ_p, *либо внутренности его границы, имеющей размерность* $\geqslant k$, *то*

$$P_{\mathbf{i},\mathbf{j}}^t \sim c'_{\mathbf{i},\mathbf{j}}\lambda_k^t, \quad t \to \infty,$$

где $c'_{\mathbf{i},\mathbf{j}}$ — *постоянная, зависящая от* $\mathbf{i}$ *и* $\mathbf{j}$ *и не зависящая от* t. *В частности, вероятность того, что популяция содержит все типы индивидуумов в* t*-м поколении, стремится к нулю со скоростью* λ_p^t, *т. е.* $P_{\mathbf{i},\mathbf{j}}^t \sim c'_{\mathbf{i},\mathbf{j}}\lambda_p^t$ *при* $\mathbf{i}, \mathbf{j} \in \Delta_p^0$ *(внутренность* Δ_p*), данное соотношение асимптотически точно при* $t \to \infty$.

Перечисленные выше результаты были детально доказаны для случая $p=3$. Интересующийся читатель может распространить доказательство на общий случай.

ЗАДАЧИ

1. Рассмотрим популяцию, изменяющуюся непрерывно во времени, с частотами генотипов $P(t)$ для AA, $Q(t)$ для Aa и $R(t)$ для aa. Предположим, что в любом интервале времени длины h (h мало) доля λh ($\lambda > 0$ и фиксировано) популяции, выбранная случайным образом, заменяется на популяцию такого же размера, которая получается путем случайного спаривания выбранной доли популяции. Показать, что при $t \to \infty$ $P(t) \to \alpha^2$, $Q(t) \to 2\alpha\beta$, $R(t) \to \beta^2$ для некоторых α, β, где $\alpha + \beta = 1$ ($\alpha > 0$, $\beta > 0$) и сходимость при этом экспоненциальная $\left(\sim e^{-\lambda t}\right)$.

2. Рассмотрим диплоидную популяцию с коэффициентами отбора 1, 1 и $1-s$ для генотипов AA, Aa и aa соответственно. Предположим также, что в каждом поколении доля μ генов A мутирует в гены a. Пусть s и μ достаточно малы так, чтобы влияния отбора и мутации были аддитивны. Показать, что стационарная устойчивая частота гена равна $q \sim \sqrt{\dfrac{\mu}{s}}$ и достигается при условии, что $\mu < s$, и при достаточно малой величине s.

3. Рассмотрим гаплоидную популяцию с коэффициентами отбора $1-s$ и 1 для генов a и A соответственно. Предположим, что доля μ генов A мутирует в гены a в каждом поколении. Предположим также, что s и μ достаточно малы, так что влияния отбора и мутации аддитивны. Показать, что $q \sim \dfrac{\mu}{s}$ — устойчивая стационарная частота гена и достигается при условии $\mu < s$ и достаточно малой величине s.

4. Для локуса, сцепленного с полом, с двумя аллелями A и a допустимые виды спаривания приведены в таблице. (Y означает мужскую половую хромосому, которая не содержит генов.)

Женские особи		Мужские особи
AA	$\times$	AY
aa	$\times$	aY
AA	$\times$	aY
aa	$\times$	AY
Aa	$\times$	aY или AY

Показать, что в случае спаривания сибсов матрица имеет вид

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{Vmatrix},$$

где типы спаривания очевидным образом объединены в три группы. Найти собственные значения, собственные векторы и частоты различных типов спаривания в n-м поколении, получившемся в результате инбридинга при произвольном начальном типе спаривания.

5. Для задачи 4 определить предельные относительные частоты типов спаривания $AA \times A\mathrm{Y}$ и $aa \times a\mathrm{Y}$, когда начальный тип спаривания был $Aa \times a\mathrm{Y}$.

6. Пусть тетраплоидная популяция при отсутствии преимуществ при отборе является объектом для хроматидного расщепления. Обозначим генотипы $\{A^4, A^3a, A^2a^2, Aa^3, a^4\}$. Пусть начальные доли гамет равны

$$\begin{array}{ccc} AA & Aa & aa \\ x & 2y & z \end{array}$$

Показать, что соответствующие доли в следующем поколении равны

$$\begin{aligned} x_1 &= x^2 + \frac{15}{28}\cdot 4xy + \frac{6}{28}(4y^2+2xz) + \frac{1}{28}(4yz), \\ 2y_1 &= f(x, y, z), \\ z_1 &= g(x, y, z), \end{aligned}$$

и найти $f(x, y, z)$ и $g(x, y, z)$.

7. В условиях задачи 6 показать, что частоты x_1, y_1, z_1 можно записать в виде

$$\begin{aligned} x_1 &= \frac{2}{7}x + \frac{4}{7}p^2 + \frac{1}{7}p, \\ 2y_1 &= \frac{2}{7}(2y) + \frac{4}{7}(2pq), \\ z_1 &= \frac{2}{7}z + \frac{4}{7}q^2 + \frac{1}{7}q, \end{aligned}$$

где $p = x + y$ и $q = y + z$.

8. Используя результат задачи 7, доказать, что частоты гамет в n-м поколении сходятся показательно с параметром $\frac{2}{7}$, и найти предельные (равновесные) доли гамет

9. Рассмотрим тетраплоидную популяцию с двумя генами A и a при хромосомном расщеплении. Предположим, что осуществляется инбридинг в форме самооплодотворения. Доказать, что сходимость к гомозиготному состоянию показательная с параметром $\frac{5}{6}$.

10. Рассмотрим тетраплоидную популяцию с двумя генами при хромосомном расщеплении с возможными зиготами (генотипами) A^4, A^3a, A^2a^2, Aa^3, a^4. Свяжем с каждой из них величины (коэффициенты отбора) α, β, γ, β, α соответственно.

(1) Доказать, что множество частот $1, 0, 0, 0, 0$ зигот $A^4, A^3a, \ldots, a^4$ соответственно устойчиво тогда и только тогда, когда $\alpha > \frac{1}{3}$ и $\beta < \alpha$.

(2) Показать, что частоты (u, v, w, v, u) являются равновесными тогда и только тогда, когда или $u = w$, или $v^2 + 2uw = 6\alpha uw$.

11. *В самостерильных* популяциях, встречающихся у некоторых диплоидных растений, пыльца не оплодотворяет растения, имеющие гены того же типа, что и содержащиеся в пыльце. Следовательно, при этом невозможно формирование гомозиготных индивидуумов. Пусть p_n — частота какого-либо аллеля в трехаллельной системе [1]). Найти рекуррентное соотношение для p_n в случае беско-

[1]) В n-м поколении. — *Прим. перев.*

нечной популяции и определить предельное значение для $p_n (n \to \infty)$. Показать, что задача сводится к исследованию двухаллельной не самостерильной системы с очень большим $(\to \infty)$ коэффициентом мутации.

12. Рассмотрим популяцию из N индивидуумов, воспроизводящихся бесполым путем. Написать рекуррентное соотношение для коэффициента инбридинга F_n (см. стр. 414), если предполагается, что с вероятностью u произвольный аллель за время одного поколения мутирует в аллель другого типа. Чему равно равновесное значение для F_n?

Ответ: $F = 2N - (2N - 1)(1 - u)^2$.

13. Рассмотрим диплоидную популяцию, в которой потомки спариваются со своими родителями и затем считаются родителями для следующего поколения. Найти матрицу переходных вероятностей для частот различных типов спариваний.

Указание: Различные типы спариваний приведены в следующей таблице.

	Потомки	Родители
1	*Aa*	*AA*
2	*Aa*	*aa*
3	*AA*	*Aa*
4	*aa*	*Aa*
5	*Aa*	*Aa*
6	*AA*	*aa*
7	*aa*	*AA*
8	*AA*	*AA*
9	*aa*	*aa*

Случаи 6 и 7 не реализуются после первого поколения, поскольку в результате могут производиться лишь гетерозиготные потомки. В силу симметрии случаи 1 и 2 могут быть сгруппированы так же, как и случаи 3 и 4.

14. В условиях задачи 13 показать, что сходимость к финальному распределению имеет скорость $(1 + \sqrt{5})/4$.

15. (Самооплодотворение при тетраплоидии) Предположим, что в тетраплоидной популяции в каждом локусе может находиться четыре различных гена. Обозначим их через a, b, c, d. Один из способов воспроизведения тетраплоидов заключается в том, что от каждого родителя случайным образом берутся по два гена и комбинируются, чтобы получить требуемую комбинацию генов для потомка. Рассмотрим популяцию, в которой все потомки получаются путем самооплодотворения, т. е. одни и те же индивидуумы выступают одновременно в роли родителей мужского и женского родов. Например, индивидуум типа *aaab* после самооплодотворения производит потомка одного из следующих типов: *aaaa*, *aaab*, *aabb* в отношении $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$. Обозначим индивидуума, содержащего лишь однотипные гены (*aaaa*, *bbbb* и т. д.), через G_4, индивидуума, у которого три гена одного типа, а четвертый — другого, — через G_{31} (*aaab*, *bbbd* и т. д.); индивидуума, у которого два гена одного типа, а другие два гена также имеют одинаковый тип, но отличный от первого, — через G_{22} (*aabb*, *aacc* и т. д.), индивидуума, у которого два гена одного типа, а другие два — различных типов, — через G_{211} (*aabc*, *ddca* и т. д.) и индивидуума типа *abcd* — через G_{1111}. Пусть $p_1^{(n)}, p_2^{(n)}, p_3^{(n)}, p_4^{(n)}, p_5^{(n)}$ — доли типов $G_4, G_{31}, G_{22}, G_{211}, G_{1111}$ соответственно в n-м поколении. Определить матрицу перехода этой цепи Маркова.

Ответ:

$$\mathbf{P} = \left\| \begin{matrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{18} & \frac{4}{9} & \frac{1}{2} & 0 & 0 \\ \frac{1}{36} & \frac{2}{9} & \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{matrix} \right\|.$$

16. В задаче 15 проанализировать **P** с тем, чтобы найти ее собственные значения, а также скорость сходимости к состоянию G_4.

Ответ: Собственные значения равны

$$1, \frac{5}{6}, \frac{1}{2}, \frac{1}{6}, \frac{1}{6}.$$

17. *Случайное спаривание генотипов.* Рассмотрим простейший случай, когда имеются лишь две возможные гаметы (*альтернирующих* генов), обозначенные через A и a. При этом возможны генотипы AA, Aa и aa.

В модели случайного спаривания N детей производится от N родителей в соответствии со следующим вероятностным механизмом. Два родителя выбираются случайным образом (с замещением) из родительской популяции, содержащей i генотипов AA, j генотипов Aa и k генотипов aa $(i+j+k=N)$. От каждого родителя гамета берется случайным образом. Эта процедура повторяется N раз.

Случайное слияние гамет. В этой модели рассматриваются сами гаметы, а не генотипы. Родительская популяция состоит из $2N$ гамет; i из них имеют тип A, а j — тип a. Следующее поколение состоит из $2N$ гамет, выбранных биномиальным образом из родительской популяции (с замещением).

Если в каком-либо фиксированном поколении необходимо построить популяцию генотипов, это делается путем случайного выбора N пар (без замещения) гамет из популяции, состоящей из $2N$ гамет.

Доказать следующую теорему.

Теорема. *Если родительская популяция состоит из i, j и k генотипов AA, Aa, aa соответственно, то вероятность того, что следующее поколение имеет l, m, n соответствующих генотипов, одинакова для модели случайного спаривания генотипов и модели случайного слияния зигот* $(i+j+k=l+m+n=N)$.

18. Выписать точный вид матрицы переходных вероятностей (5.8) для частного случая, когда

$$f_1(u) = f_2(u) = f_3(u) = e^{a(u-1)}, \quad a > 0.$$

19. В предыдущей задаче найти собственные значения матрицы переходных вероятностей.

20. Рассмотрим марковскую цепь Райта (4.13) при $\lambda = \mu$ и $a = b = 0$. Чему равны собственные значения этой матрицы?

Ответ:

$$\lambda_r = (1-\alpha_1-\alpha_2)^r \frac{N(N-1)\ldots(N-r+1)}{N^r}, \qquad r = 0, 1, \ldots, N.$$

21. Рассмотрим марковскую модель, учитывающую мутационное давление, описанную в задаче 20. Пусть $\pi = \{\pi_i\}_{i=0}^N$ — стационарное распределение процесса. Найти его первый и второй моменты.

Указание: Использовать тот факт, что $\sum_{j=0}^N P_{ij}j$ линейна по i, а $\sum_{j=0}^N P_{ij}j^2$ — квадратичная функция i. См. также стр. 430.

ЛИТЕРАТУРА

1. Moran P. A. P., The Statistical Processes of Evolutionary Theory, Oxford Univ. Press (Clarendon), London and New York, 1962.
2. Fisher R. A., The Genetical Theory of Natural Selection, Oxford Univ. Press (Clarendon), London and New York, 1930.
3. Malécot G., Les mathématiques de l'hérédité, Masson, Paris, 1948.
4. Li C. C., Population Genetics, McGraw-Hill, New York, 1956.

Глава 14

ПРОЦЕССЫ МАССОВОГО ОБСЛУЖИВАНИЯ[1])

§ 1. ОБЩЕЕ ОПИСАНИЕ

В гл. 3 мы обсудили два простых примера процессов массового обслуживания с дискретным временем. В этой главе мы более подробно рассмотрим идеи и методы, используемые при анализе некоторых процессов массового обслуживания с непрерывным временем. Общая модель обслуживания имеет следующую структуру. На некоторое устройство в случайные моменты времени поступают требования[2]) и требуют какого-либо рода обслуживания. Процессы массового обслуживания классифицируются по следующим признакам.

(1) Входящий поток требований, для задания которого необходимо знать распределение моментов поступления требований, или, в более частном случае, распределение интервалов между моментами поступления.

(2) Распределение времени обслуживания требования.

(3) Дисциплина обслуживания (очереди), число обслуживающих приборов и организация очереди и процесса обслуживания. В большинстве моделей принята дисциплина «первый пришел — первый обслужен» — прямой порядок обслуживания, когда обслуживание требования начинается, как только оно достигает начала очереди. Все модели, которые будут рассмотрены в этой главе, такого типа. (Другие типы дисциплин обслуживания рассматриваются в задачах 2—4 настоящей главы.)

Мы повсюду будем предполагать, что входящий поток требований не зависит от размера очереди и что интервалы между последовательными моментами поступления требований являются независимыми одинаково распределенными положительными слу-

---

[1]) В оригинале «queueing processes», т. е. «процессы образования очередей». При переводе мы используем принятый в отечественной литературе термин «процессы массового обслуживания», который был введен А. Я. Хинчиным и более полно отражает существо рассматриваемых задач. — *Прим. перев.*

[2]) Под термином «требования» можно понимать, например, клиентов, требующих обслуживания, корабли, входящие в порт, поток сообщений в некоторую контору, неисправные машины, ожидающие ремонта, и т. д. (Вместо термина «требование» используют также термины «клиент», «вызов», «заявка», а вместо термина «обслуживающее устройство» — «прибор», «линия», «канал». Употребление того или иного термина часто определяется характером решаемой прикладной задачи. — *Перев.*)

чайными величинами. Такие потоки иногда называют рекуррентными, или потоками восстановления.

Термин «простейший поток» используется иногда в случае, когда моменты поступления требований образуют пуассоновский поток, т. е. когда интервалы между моментами поступления распределены экспоненциально.

Будет также предполагаться, что длительности обслуживания отдельных требований — независимые одинаково распределенные случайные величины, не зависящие от входящего потока.

§ 2. ПРОСТЕЙШИЕ ПРОЦЕССЫ ОБСЛУЖИВАНИЯ ($M/M/1$)[1]

Наиболее простыми и наиболее полно изученными являются процессы обслуживания с пуассоновским входящим потоком и экспоненциальным распределением времени обслуживания. Эти процессы уже были описаны и было показано, что процесс изменения длины очереди является процессом рождения и гибели (см. пример 2, § 6 гл. 7)[2].

Вновь рассмотрим случай одного обслуживающего прибора. Функция распределения интервалов между моментами поступления равна

$$F(t) = 1 - e^{-\lambda t}, \qquad \lambda > 0,$$

а функция распределения длительности обслуживания

$$G(t) = 1 - e^{-\mu t}, \qquad \mu > 0.$$

В силу «отсутствия последействия» (теорема 2.2 гл. 7) у экспоненциального распределения очевидно, что процесс $X(t)$ (длина очереди в момент t) — однородный по времени марковский процесс рождения и гибели. Пусть $P_{ij}(t)$ — его переходная вероятностная функция. Тогда $P_{i,\,i+1}(h)$ — вероятность того, что за время h поступит одно новое требование и не закончится обслуживание ни одного требования. При малых h

$$P_{i,\,i+1}(h) = \lambda h + o(h), \qquad i \geqslant 0.$$

---

[1]) Для обозначения простых процессов обслуживания в литературе широко используются стандартные сокращения. Дадим для справки сокращенные обозначения процессов, которые будут изучаться ниже. В записи $(A/B/c)$ c — число обслуживающих приборов, а A и B указывают типы распределений интервалов между моментами поступлений требований и длительностей обслуживания соответственно. На первых двух местах используются следующие символы: G или GI, когда относительно распределений не делается никаких частных предположений; M, когда соответствующее распределение экспоненциальное; E_k, когда соответствующие интервалы имеют гамма-распределение порядка k (эрланговское распределение) (так что $E_1 = M$); D (детерминированное), когда указанные интервалы имеют фиксированные длительности.

[2]) Под «длиной очереди» здесь и далее понимается общее число требований, находящихся на обслуживании и ждущих его. — *Прим. перев.*

Аналогично находим

$$P_{i,\,i-1}(h) = \mu h + o(h), \qquad i \geqslant 1,$$

и

$$P_{ii}(h) = 1 - (\lambda + \mu)\,h + o(h), \qquad i \geqslant 1,$$
$$P_{00}(h) = 1 - \lambda h + o(h).$$

Инфинитезимальная матрица равна

$$\mathbf{A} = \begin{Vmatrix} -\lambda & \lambda & 0 & 0 \ldots \\ \mu & -(\lambda+\mu) & \lambda & 0 \ldots \\ 0 & \mu & -(\lambda+\mu) & \lambda \ldots \\ \vdots & \vdots & \vdots & \vdots \end{Vmatrix}.$$

В § 6 гл. 7 показано, что

$$\lim_{t\to\infty} P_{ij}(t) = p_j = \begin{cases} \left(\frac{\lambda}{\mu}\right)^j \left(1 - \frac{\lambda}{\mu}\right), & \lambda < \mu, \\ 0, & \lambda \geqslant \mu. \end{cases}$$

Отсюда получаем ответ на многие вопросы, включая стационарность. Если процесс развивался достаточно долгое время и $\lambda < \mu$, то вероятность того, что поступившее требование начнет немедленно обслуживаться (обслуживающий прибор свободен, т. е. длина очереди равна нулю), равна

$$p_0 = 1 - \frac{\lambda}{\mu}.$$

В случае $\lambda < \mu$ можно также найти распределение времени ожидания в стационарном режиме. Если поступающее требование застает очередь длины n, то его время T пребывания в системе складывается из длительностей обслуживания его самого и требований, стоящих перед ним. Все эти величины распределены экспоненциально с параметром μ, и, поскольку длительности обслуживания не зависят от размера очереди, T имеет гамма-распределение порядка $n+1$ с масштабным параметром μ:

$$\mathrm{P}\{T \leqslant t \mid \text{длина очереди равна } n\} = \int_0^t \frac{\mu^{n+1}\tau^n e^{-\mu\tau}}{\Gamma(n+1)}\,d\tau. \tag{2.1}$$

В силу формулы полной вероятности имеем

$$\mathrm{P}\{T \leqslant t\} = \sum_{n=0}^{\infty} \mathrm{P}\{T \leqslant t \mid \text{длина очереди равна } n\} \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right),$$

поскольку $\left(\frac{\lambda}{\mu}\right)^n\left(1-\frac{\lambda}{\mu}\right)$ — вероятность того, что в стационарном режиме поступающее требование застанет очередь длины n. Учитывая (2.1), находим

$$\begin{aligned}\mathrm{P}\{T \leqslant t\} &= \sum_{n=0}^{\infty}\int_0^t \frac{1}{\Gamma(n+1)}\mu^{n+1}\tau^n e^{-\mu\tau}\left(\frac{\lambda}{\mu}\right)^n\left(1-\frac{\lambda}{\mu}\right)d\tau = \\ &= \int_0^t \mu e^{-\mu\tau}\left(1-\frac{\lambda}{\mu}\right)\sum_{n=0}^{\infty}\frac{1}{\Gamma(n+1)}\tau^n\lambda^n\,d\tau = \\ &= \int_0^t \left(1-\frac{\lambda}{\mu}\right)\mu\exp\left\{-\tau\mu\left(1-\frac{\lambda}{\mu}\right)\right\}d\tau = \\ &= 1-\exp\left[-t\mu\left(1-\frac{\lambda}{\mu}\right)\right].\end{aligned}$$

Вновь получили экспоненциальное распределение.

Если мы хотим исследовать неустановившийся режим, следует прежде найти $P_{ij}(t)$ для всех t. Это существенно более сложная задача, но она решена. Подробности этого решения выходят за рамки данной книги, и мы отсылаем интересующегося читателя к любой из специальных книг по теории массового обслуживания, перечисленных в конце главы.

Если имеются два обслуживающих прибора, то для тех же законов поступления требований и обслуживания, когда в очереди имеется не меньше двух требований, среднее время μ_n^{-1} до завершения очередного акта обслуживания вдвое меньше, чем при одном приборе. Таким образом, $\mu_n = 2\mu$, $n \geqslant 2$. Если же $n = 1$, то один прибор пустует, и $\mu_1 = \mu$. Инфинитезимальная матрица этого процесса рождения и гибели имеет вид

$$\mathbf{A} = \left\|\begin{matrix} -\lambda & \lambda & 0 & 0 & \ldots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \ldots \\ 0 & 2\mu & -(\lambda+2\mu) & \lambda & \ldots \\ 0 & 0 & 2\mu & -(\lambda+2\mu) & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix}\right\|.$$

§ 3. НЕКОТОРЫЕ ОБЩИЕ МОДЕЛИ ОБСЛУЖИВАНИЯ ОДНИМ ПРИБОРОМ

Обсудим некоторые аспекты трех методов анализа частных видов системы обслуживания $(GI/G/1)$. Первый метод, известный под названием метода интегрального уравнения, сводит задачу нахождения предельного распределения времени ожидания начала обслуживания n-м требованием ($n \to \infty$) к задаче решения интегрального уравнения типа Винера — Хопфа.

Если входящий поток является пуассоновским, то при втором методе исследования рассматривается длина очереди в моменты окончания актов обслуживания. Можно показать, что этот *вложенный процесс* является цепью Маркова (см. ниже § 4). Если распределение времени обслуживания является экспоненциальным, а входящий поток определяется общим распределением, то *вложенная цепь Маркова* получается при рассмотрении длины очереди в моменты новых поступлений. Результирующий процесс является цепью Маркова специального вида.

С помощью третьего метода исследуются свойства случайной величины $W(t)$, равной времени до начала обслуживания, которое пришлось бы ожидать требованию, если бы оно поступило в момент t независимо от того, поступило оно на самом деле в этот момент или нет. Эта величина называется *виртуальным временем ожидания* в момент t.

Рассмотрим сначала метод интегрального уравнения, а затем перейдем к моделям более частного вида, к которым применим метод вложенных цепей Маркова[1]. Некоторые вопросы, относящиеся к третьему методу, обсуждаются в § 8.

А. Метод интегрального уравнения[2]

Введем величины

W_r — время ожидания r-м поступившим требованием начала обслуживания,
S_r — длительность обслуживания r-го требования,
T_r — длительность интервала между поступлением r-го и $(r+1)$-го требований,

где $W_0 = S_0 = T_0 = 0$. Условие $W_0 = 0$ означает, что первое поступающее требование застает обслуживающий прибор свободным. Предположим, что оно поступает в момент $t = 0$.

Очевидно, $W_r + S_r$ — время пребывания r-го требования в системе. Следовательно, если $T_r > W_r + S_r$, то $(r+1)$-е требование застанет обслуживающий прибор свободным, т. е. в этом случае $W_{r+1} = 0$. Если $T_r \leqslant W_r + S_r$, то длительность времени ожидания $(r+1)$-го требования равна, очевидно, $W_r + S_r - T_r$. Следовательно,

$$W_{r+1} = \begin{cases} W_r + S_r - T_r, & \text{если } W_r + S_r - T_r \geqslant 0, \\ 0, & \text{если } W_r + S_r - T_r < 0. \end{cases}$$

Обозначим

$$U_r = S_r - T_r.$$

---

[1]) Заметим, что при выводе интегрального уравнения типа Винера — Хопфа также сначала строится вложенная цепь Маркова $\{W_r\}_{r=1}^{\infty}$ (см. ниже), у которой пространством состояний является полупрямая $[0, \infty)$. Поэтому отличие первого метода от второго относится по существу к типам вложенных цепей Маркова и, следовательно, к методам их анализа. — *Прим. перев.*

[2]) При первом чтении оставшуюся часть материала данного параграфа можно опустить.

Тогда, очевидно, $\{U_r\}_{r=1}^{\infty}$ — последовательность независимых одинаково распределенных д. с. в.[1]). Пусть $F_r(x)$ — функция распределения величины W_r, а $g(x) = G'(x)$ — плотность величины U_r, которая по предположению одинакова для всех r[2]). Поскольку W_r и U_r независимы, то при $x \geqslant 0$

$$F_{r+1}(x) = \mathrm{P}\{W_{r+1} \leqslant x\} = \mathrm{P}\{\max(W_r + U_r,\ 0) \leqslant x\} = \mathrm{P}\{W_r + U_r \leqslant x\} =$$

$$= \int_{-\infty}^{x} \mathrm{P}\{W_r + U_r \leqslant x \mid U_r = y\}\, g(y)\, dy = \int_{y \leqslant x} F_r(x-y)\, g(y)\, dy \qquad (r \geqslant 1). \qquad (3.1)$$

Далее, поскольку первое требование поступает в момент $t = 0$ и застает обслуживающий прибор свободным, то

$$F_1(x) = \begin{cases} 1 & \text{при} \quad x \geqslant 0, \\ 0 & \text{при} \quad x < 0, \end{cases}$$

и так как при $x < 0$ все $F_i(x) = 0$, то

$$F_1(x) - F_2(x) \geqslant 0, \qquad -\infty < x < \infty.$$

Но

$$F_r(x) - F_{r+1}(x) = \int_{y \leqslant x} [F_{r-1}(x-y) - F_r(x-y)]\, g(y)\, dy,$$

и отсюда по индукции следует, что при любом r

$$F_r(x) - F_{r+1}(x) \geqslant 0, \qquad -\infty < x < \infty.$$

Следовательно, при каждом x функция $F_r(x)$ убывает с ростом r.

Поскольку $F_r(x) \geqslant 0$, то отсюда следует, что $F_r(x)$ сходится, скажем, к функции $F(x)$. Переходя к пределу при $r \to \infty$ в равенстве (3.1), получим

$$F(x) = \int_{y \leqslant x} F(x-y)\, g(y)\, dy\ ^{3)},$$

или, если положить $z = x - y$,

$$F(x) = \int_{0}^{\infty} F(z)\, g(x-z)\, dz.$$

Теперь нужно исследовать вопрос о том, когда предел $F(x)$ является собственным распределением. Очевидно, что $F(x)$ — неубывающая функция, но может оказаться, что $\lim\limits_{x \to \infty} F(x) < 1$, а не $\lim\limits_{x \to \infty} F(x) = 1$. Первый случай можно интерпретировать как возможность того, что время ожидания n-го требования ($n \to \infty$) стремится к ∞ с положительной вероятностью (или длина очереди стремится к ∞ с положительной вероятностью).

---

[1]) Заметим, что множество значений д. с. в. U_r, $r = 1, 2, \ldots$, — вся действительная прямая $(-\infty, \infty)$. — *Прим. перев.*

[2]) Ограничение, что $G(x)$ имеет плотность, наложено для простоты рассмотрения и не существенно для последующих рассуждений.

[3]) Обоснование возможности перехода к пределу под знаком интеграла требует знания теории интеграла Лебега. Если читатель еще не знаком с ней, то ему следует принять соответствующее утверждение на веру.

Получим сначала другое выражение для $F(x)$. Поскольку

$$F_1(x) = \begin{cases} 1, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

то $F_2(x) = \int\limits_{u \leqslant x} g(u)\,du = \mathrm{P}\{U_1 \leqslant x\}$, $x \geqslant 0$. Далее,

$$F_3(x) = \int\limits_{u \leqslant x} F_2(x-u)\,g(u)\,du = \int\limits_{u \leqslant x}\left[\int\limits_{v \leqslant x-u} g(v)\,dv\right] g(u)\,du =$$

$$= \int\limits_{u_2 \leqslant x,\ u_1+u_2 \leqslant x} g(u_2)\,g(u_1)\,du_2\,du_1 =$$

$$= \mathrm{P}\{U_2 \leqslant x,\ U_1 + U_2 \leqslant x\} = \mathrm{P}\{U_1 \leqslant x,\ U_1 + U_2 \leqslant x\},$$

где использован тот факт, что U_1 и U_2 — независимые одинаково распределенные д. с. в. По индукции непосредственно получаем, что

$$F_{r+1}(x) = \mathrm{P}\{U_r \leqslant x,\ U_r + U_{r-1} \leqslant x,\ \ldots,\ U_r + \ldots + U_1 \leqslant x\} =$$

$$= \mathrm{P}\{U_1 \leqslant x,\ U_1 + U_2 \leqslant x,\ \ldots,\ U_1 + \ldots + U_r \leqslant x\},$$

поскольку $U_1, U_2, \ldots, U_r$ одинаково распределены. Таким образом, если $\tilde{U}_r = \sum_{i=1}^{r} U_i$, то

$$F_{n+1}(x) = \mathrm{P}\{\tilde{U}_r \leqslant x,\ r = 1, \ldots, n\}, \qquad x \geqslant 0.$$

Очевидно, $F_n(x)$ монотонно убывает с ростом n (это было также доказано выше), и поэтому

$$F(x) = \mathrm{P}\{\tilde{U}_r \leqslant x \text{ при всех } r\}, \qquad x \geqslant 0.$$

Если $x < 0$, то $F_i(x) = 0$, и тривиальным образом получаем $F(x) = 0$, $x < 0$.

Используя полученный результат, можно определить, когда $F(x)$ является собственным распределением. Предположим, что $\mathrm{M}(S) < \infty$ и $\mathrm{M}(T) < \infty$, т. е. д. с. в. S и T имеют конечные математические ожидания. Тогда справедлива следующая

Теорема 3.1. (1) *Если* $\mathrm{M}(U) \geqslant 0$, *то* $F(x) \equiv 0$. (2) *Если* $\mathrm{M}(U) < 0$, *то* $\lim\limits_{x\to\infty} F(x) = 1$.

Интуитивно этот результат очевиден. Он утверждает, что если средняя длительность интервала между моментами поступления меньше средней длительности обслуживания, то очередь растет бесконечно и $W_r \to \infty$ с вероятностью 1. Доказательство разбивается на три части.

Доказательство.

(1) $\mathrm{M}(U) > 0$.

В силу усиленного закона больших чисел

$$\lim_{n\to\infty} \frac{\tilde{U}_n}{n} = \mathrm{M}(U) \text{ с вероятностью } 1,$$

следовательно, для почти всех реализаций последовательности $\tilde{U}_1, \tilde{U}_2, \tilde{U}_3, \ldots$, т. е. с вероятностью 1 имеем

$$\tilde{U}_n \geqslant \frac{n}{2} \mathrm{M}(U) \tag{3.2}$$

при достаточно больших n, где выбор n зависит от реализации. Событие $\{\tilde{U}_r \leqslant x$ при всех $r\}$ является частью события, дополнительного к событию (3.2). Следовательно, его вероятность равна нулю.

(2) $\mathrm{M}(U) < 0$.

Вновь в силу усиленного закона больших чисел для любых $\varepsilon > 0$ и $\delta > 0$ существует целое $N_{\varepsilon, \delta}$, такое, что при $n \geqslant N_{\varepsilon, \delta}$

$$\mathrm{P}\left\{\left|\frac{1}{n} \tilde{U}_n - \mathrm{M}(U)\right| \leqslant \varepsilon \text{ при всех } n \geqslant N_{\varepsilon, \delta}\right\} \geqslant 1 - \delta.$$

Выберем ε таким малым, что $\mathrm{M}(U) + \varepsilon < 0$. Тогда для любого $\delta > 0$ существует N_δ, такое, что

$$1 - \delta \leqslant \mathrm{P}\{\tilde{U}_n \leqslant n(\mathrm{M}(U) + \varepsilon) \text{ при всех } n \geqslant N_\delta\} \leqslant$$
$$\leqslant \mathrm{P}\{\tilde{U}_n \leqslant 0 \text{ при всех } n \geqslant N_\delta\}.$$

Далее, поскольку $G(x)$ — собственное распределение, для указанных δ и N_δ можно выбрать достаточно большое x, такое, что

$$\mathrm{P}\{B\} = \mathrm{P}\{\tilde{U}_r \leqslant x \text{ при всех } 1 \leqslant r \leqslant N_\delta - 1\} \geqslant 1 - \delta.$$

Смысл события B очевиден из приведенного равенства.

Пусть $A = \{\tilde{U}_n \leqslant 0$ при всех $n \geqslant N_\delta\}$. Событие $\{\tilde{U}_r \leqslant x$ при всех $r\}$ содержит пересечение двух событий A и B и его вероятность не меньше $\mathrm{P}\{A \cap B\} = \mathrm{P}\{A\} + \mathrm{P}\{B\} - \mathrm{P}\{A \cup B\} \geqslant 2 - 2\delta - 1 = 1 - 2\delta$. Следовательно, $F(x) \geqslant 1 - 2\delta$ при достаточно больших x, или

$$\lim_{x \to \infty} F(x) \geqslant 1 - 2\delta.$$

Но, поскольку δ произвольно, это означает, что $\lim\limits_{x \to \infty} F(x) = 1$.

(3) $\mathrm{M}(U) = 0$.

В этом случае утверждение следует из довольно глубокой теоремы, относящейся к возвратности сумм независимых случайных величин, которая лежит за пределами данной книги[1]). Дискретным аналогом теоремы является теорема 3.3 гл. 6.

Б. Возвратность события, заключающегося в том, что время ожидания поступающего требования равно нулю

При анализе случайной последовательности $\{W_n\}_{n=0}^{\infty}$ с пространством состояний $[0, \infty)$ возникает вопрос о возвратности события A, заключающегося в том, что некоторое требование застанет систему свободной. Формально скажем, что событие A наступило на n-м шаге, если $W_n = 0$. Предложение «событие A происходит» будет в дальнейшем означать, что оно происходит на некотором конечном шаге. Заметим, что, если событие A происходит, процесс начинается с очередного значения W, равного нулю.

---

[1]) См. Спицер Ф., Принципы случайного блуждания, «Мир», М., 1968.

Теорема 3.2. (1) *Если* $\mathrm{M}(U)>0$, *то событие* A — *невозвратное (т. е. вероятность события* A *меньше* 1).

(2) *Если* $\mathrm{M}(U)\leqslant 0$, *то событие* A — *возвратное (т. е.* $\mathrm{P}\{A\}=1$).

(3) *Если* $\mathrm{M}(U)<0$, *то событие* A — *положительное возвратное (т. е. среднее время до наступления* A *конечно).*

Доказательство.

(1) Используя те же обозначения, что и прежде, заметим, что

$$W_{n+1}\geqslant \tilde{U}_n.$$

(Напомним, что $\tilde{U}_n=U_1+\ldots+U_n=S_1+\ldots+S_n-T_1-\ldots-T_n$ — разность между суммарным временем обслуживания первых n требований и временем поступления $(n+1)$-го требования.) При этом равенство выполняется лишь в том случае, если обслуживающий прибор был все это время занят (т. е. до момента поступления $(n+1)$-го требования не было перерыва в его работе).

Таким образом, если все $\tilde{U}_n>0$, то событие A не наступает. В силу усиленного закона больших чисел для любых ε, $\delta>0$ существует число N, такое, что

$$\mathrm{P}\left\{\left|\frac{\tilde{U}_n}{n}-\mathrm{M}(U)\right|\leqslant\varepsilon \text{ при всех } n\geqslant N\right\}\geqslant 1-\delta.$$

Таким образом, если $\mathrm{M}(U)>0$ и выбрать достаточно малое ε ($\varepsilon<\mathrm{M}(U)$), то существует число N, такое, что

$$\mathrm{P}\{\tilde{U}_n>0 \text{ при всех } n\geqslant N\}>0.$$

Это означает, что с положительной вероятностью A может происходить лишь конечное число раз, т. е. событие $\{W_r=0$ лишь для конечного набора индексов $r\}$ имеет положительную вероятность. Но событие является возвратным тогда и только тогда, когда вероятность его осуществления бесконечное число раз равна 1 (см. теорему 7.1 гл. 2). Следовательно, если $\mathrm{M}(U)>0$, то A невозвратно.

(2) Если $\mathrm{M}(U)<0$, то для произвольных ε, $\delta>0$ существует число N, такое, что

$$\mathrm{P}\left\{\left|\frac{\tilde{U}_n}{n}-\mathrm{M}(U)\right|<\varepsilon \text{ для всех } n\geqslant N\right\}\geqslant 1-\delta.$$

Отсюда, если ε достаточно мало, получаем

$$\mathrm{P}\{\tilde{U}_n\leqslant 0 \text{ при всех } n\geqslant N\}\geqslant 1-\delta$$

и в силу произвольности δ, усиливая неравенство, имеем

$$\mathrm{P}\{\tilde{U}_n\leqslant 0 \text{ при некотором } n\}=1.$$

Но если $\tilde{U}_n\leqslant 0$, то некоторая величина $W_i=0$; в частности, если $\tilde{U}_k$ — первая из последовательности $\{\tilde{U}_n\}$ величина, которая $\leqslant 0$, то $W_k=0$. Поэтому если $\mathrm{M}(U)<0$, то A — возвратное событие.

Если $\mathrm{M}(U)=0$, то соответствующее доказательство является довольно тонким и мы не будем его проводить, а отошлем читателя к цитируемой в конце главы литературе.

(3) Если $\mathrm{M}(U)<0$, мы утверждаем, что событие A является возвратным положительным. Доказательство этого мы опускаем.

§ 4. МЕТОД ВЛОЖЕННЫХ ЦЕПЕЙ МАРКОВА ПРИМЕНИТЕЛЬНО К МОДЕЛИ ОБСЛУЖИВАНИЯ ($M/G/1$)

Рассмотрим частный случай одноканальной системы с пуассоновским входящим потоком (с параметром λ). Предположим, что длительность обслуживания V является положительной случайной величиной с произвольным распределением $B(v)$, для которого $\mathrm{M}(V) < \infty$. Для простоты изложения предположим, что $B(v)$ имеет плотность $b(v)$. Исследуем этот процесс с помощью вложенной цепи Маркова, определяемой следующим образом.

Пусть $Z(t)$ — число требований в очереди в момент t $(t \geqslant 0)$. Предположим, что процесс $Z(t)$ наблюдается в моменты окончаний операций обслуживания[1]. При этом получается последовательность целых чисел

$$Z(t_1), \quad Z(t_2), \quad Z(t_3), \ \ldots, \tag{4.1}$$

где $t_1, t_2, t_3, \ldots$ — последовательные моменты окончаний операций обслуживания. Последовательность $\{Z(t_n)\}$ образует процесс с дискретным временем

$$X_0 = 0, \quad X_n = Z(t_n), \quad n = 1, 2, \ldots .^{2)} \tag{4.2}$$

Ниже будет показано, что в силу «пуассоновости» входящего потока последовательность (4.2) является цепью Маркова.

Дадим несколько более интуитивное описание данного процесса с дискретным временем. Его переходы происходят только в те моменты времени, когда заканчиваются операции обслуживания требований. Состоянием процесса является число требований в очереди (включая и требование, которое начало обслуживаться, если таковое имеется), оставшееся после того, как обслуженное требование покинуло систему.

Легко видеть, что этот процесс является марковским, поскольку если X_n — состояние системы в момент n, то

$$X_{n+1} = \begin{cases} X_n - 1 + N, & \text{если} \quad X_n \geqslant 1, \\ N, & \text{если} \quad X_n = 0, \end{cases} \qquad n = 0, 1, \ldots, \tag{4.3}$$

где N — число требований, поступающих за время V обслуживания $(n+1)$-го требования. Но случайная величина V по предположению не зависит от предыдущих длительностей операций обслуживания и длины очереди. В силу стационарного характера пуассоновского потока число поступлений N за время обслуживания

[1]) Именно если операция обслуживания оканчивается в момент t, то система наблюдается в момент $t + 0$ (после ухода обслуженного требования). — *Прим. перев*

[2]) Условие $X_0 = 0$ говорит о том, что рассматривается система, в которой в момент $t = 0$ отсутствуют требования. — *Прим. перев.*

зависит только от V и не зависит ни от длины очереди, ни от момента начала обслуживания. Отсюда следует, что $\{X_n\}$ — цепь Маркова.

Вероятностное распределение величины N можно найти, налагая условие на значение V и применяя формулу полной вероятности

$$\mathrm{P}\{N=n\}=\int_0^\infty \mathrm{P}\{N=n \mid V=v\}\, b(v)\, dv.$$

Далее, число требований, поступающих за время v, является случайной величиной с пуассоновским распределением, имеющим параметр λv. Следовательно,

$$\mathrm{P}\{N=n \mid V=v\}=e^{-\lambda v}\frac{1}{n!}(\lambda v)^n,$$

поэтому

$$\begin{aligned} P_{ij} &= \mathrm{P}\{X_{n+1}=j \mid X_n=i\}=\mathrm{P}\{N=j-i+1\}= \\ &= \int_0^\infty \mathrm{P}\{N=j-i+1 \mid V=v\}\, b(v)\, dv=\int_0^\infty e^{-\lambda v}\frac{(\lambda v)^{j-i+1}}{(j-i+1)!}\, b(v)\, dv, \end{aligned}$$

$$j \geqslant i-1 \quad (i \geqslant 1),$$

$$P_{ij}=0, \quad j<i-1 \quad (i \geqslant 1).$$

Если требование покидает систему, оставляя ее свободной, то ее состояние остается нулевым до тех пор, пока не поступит и не начнет обслуживаться очередное требование. Таким образом, нулевое и первое состояния эквивалентны относительно переходов в другие состояния.

Если обозначить

$$k_r=\int_0^\infty e^{-\lambda v}\frac{(\lambda v)^r}{r!}\, b(v)\, dv, \qquad r=0,\ 1,\ 2,\ \ldots$$

(k_r — вероятность того, что за время обслуживания одного требования поступит r требований), то

$$\mathbf{P}=\begin{Vmatrix} k_0 & k_1 & k_2 \ldots \\ k_0 & k_1 & k_2 \ldots \\ 0 & k_0 & k_1 \ldots \\ 0 & 0 & k_0 \ldots \\ \vdots & \vdots & \vdots \end{Vmatrix}. \tag{4.4}$$

Детальный анализ цепи Маркова (4.4) уже был проведен в § 5 гл. 3. Там было доказано, что цепь является возвратной положительной, возвратной нулевой или невозвратной в зависимости от того, $\rho < 1$, $\rho = 1$ или $\rho > 1$ соответственно, где

$$\rho = \sum_{r=0}^{\infty} r k_r = \int_0^{\infty} e^{-\lambda v} \sum_{r=0}^{\infty} \frac{r(\lambda v)^r}{r!} b(v)\,dv = $$
$$= \lambda \int_0^{\infty} v b(v)\,dv = \lambda \mathrm{M}(V).$$

Величина

$$\rho = \frac{\text{среднее время обслуживания одного требования}}{\text{средний интервал между моментами поступления требования}}$$

называется *нагрузкой системы*. Найдем теперь предельное распределение для данной цепи Маркова в случае $\rho < 1$.

А. Стационарное распределение вложенной цепи Маркова

Найдем вектор

$$\pi = (\pi_0, \pi_1, \ldots), \qquad \pi_i > 0, \qquad \sum_{i=0}^{\infty} \pi_i = 1,$$

такой, что

$$\sum_{i=0}^{\infty} \pi_i P_{ij} = \pi_j, \qquad j = 0, 1, 2, \ldots,$$

где матрица $\|P_{ij}\| = \mathbf{P}$ определена равенством (4.4). Если выразить эти уравнения через величины k_i, то они примут вид

$$\pi_i = \pi_0 k_i + \sum_{r=1}^{i+1} \pi_r k_{i-r+1}, \qquad i = 0, 1, 2, \ldots.$$

Найдем производящую функцию

$$\pi(s) = \sum_{i=0}^{\infty} \pi_i s^i$$

через функцию

$$K(s) = \sum_{i=0}^{\infty} k_i s^i.$$

Умножая записанное выше равенство на s^i, получим

$$s^i \pi_i = \pi_0 k_i s^i + \frac{1}{s} \sum_{r=0}^{i+1} \pi_r k_{i-r+1} s^{i+1} - \frac{\pi_0 k_{i+1} s^{i+1}}{s}, \qquad i = 0, 1, 2, \ldots.$$

Суммируя по i и учитывая, что $\sum_{r=0}^{i+1} \pi_r k_{i+1-r}$ является сверткой, получим

$$\sum_{i=0}^{\infty} \pi_i s^i = \pi(s) = \pi_0 K(s) + \frac{1}{s}[K(s)\pi(s) - \pi_0 k_0] - \frac{\pi_0}{s}[K(s) - k_0].$$

Отсюда

$$\pi(s) = \frac{\pi_0 K(s)(s-1)}{s - K(s)}.$$

Эта формула определяет производящую функцию стационарного распределения с точностью до постоянного множителя π_0. Поскольку

$$K(1) = \sum_{i=0}^{\infty} k_i = 1,$$

то дробь

$$\frac{s - K(s)}{s-1} = \frac{s-1}{s-1} - \frac{1-K(s)}{1-s},$$

стоящая в знаменателе выражения для $\pi(s)$, стремится при $s \to 1$ к величине $1 - K'(1)$.

Найдем $K'(1)$ — среднее число требований, поступающих за время обслуживания одного требования:

$$K'(1) = \sum_{r=1}^{\infty} r k_r = \lambda \mathrm{M}(V) = \frac{\mathrm{M}(V)}{\mathrm{M}(A)} = \rho$$

(ρ, очевидно, является нагрузкой системы), где $\frac{1}{\lambda} = \mathrm{M}(A)$ — средняя длительность интервала между моментами поступления. Поскольку $\rho < 1$, стационарное распределение существует и, следовательно, $\pi(1) = 1$. Но из полученной формулы следует

$$\pi(1) = \frac{\pi_0}{1-\rho},$$

поэтому

$$\pi_0 = 1 - \rho.$$

Таким образом производящая функция стационарного распределения равна

$$\pi(s) = (1-\rho)\frac{K(s)(s-1)}{s-K(s)}\,[1].$$

Величина $1 - \rho$ является стационарной вероятностью того, что система свободна.

---

[1]) Данное выражение для $\pi(s)$ называют формулой Полячека — Хинчина. — *Прим. перев.*

Б. Средняя длина очереди для системы $(M/G/1)$ в стационарном режиме

В заключение данного параграфа найдем среднюю длину очереди и среднее время ожидания поступившим требованием начала обслуживания для системы, находящейся в стационарном режиме.

Дифференцируя $\pi(s)$, нелегко непосредственно получить выражение для $\mathrm{M}(q)$, где q — состояние вложенной цепи Маркова, находящейся в стационарном режиме. Изберем для получения $\mathrm{M}(q)$ другой метод.

Если q — число требований в очереди после ухода некоторого требования, а q' — число требований в очереди после следующего ухода, то

$$q' = q - 1 + \delta + N,$$

где N — число поступивших за время обслуживания требований, а

$$\delta = \begin{cases} 1, & \text{если } q = 0, \\ 0, & \text{если } q > 0. \end{cases}$$

В стационарном режиме величина q' имеет такое же распределение, что и q. Таким образом, $\mathrm{M}(q') = \mathrm{M}(q)$ и

$$\mathrm{M}(\delta) = 1 - \mathrm{M}(N) = 1 - \rho. \tag{4.5}$$

Из этого же выражения, поскольку $\delta^2 = \delta$, имеем

$$q'^2 = q^2 + \delta + (N-1)^2 + 2q\delta + 2\delta(N-1) + 2q(N-1).$$

Так как $q\delta = 0$ (см. определение δ), то

$$q'^2 = q^2 + \delta + N(N-1) + (1-N) + 2\delta(N-1) + 2q(N-1).$$

Но величина N (число требований, поступивших за время обслуживания) не зависит от q и, следовательно, от δ.

Усредняя полученное равенство и учитывая (4.5), получаем

$$\mathrm{M}(q'^2) = \mathrm{M}(q^2) + 1 - \rho + \mathrm{M}[N(N-1)] + \\ + 1 - \rho + 2(1-\rho)(\rho - 1) + 2\mathrm{M}(q)(\rho - 1).$$

Поскольку $\mathrm{M}(q'^2) = \mathrm{M}(q^2)$ в силу предположения о стационарности, то

$$\mathrm{M}(q) = \rho - \frac{\mathrm{M}[N(N-1)]}{2(\rho - 1)}.$$

Далее,

$$K(s) = \sum_{n=0}^{\infty} k_n s^n = \int_0^{\infty} e^{-\lambda v(1-s)} b(v)\, dv,$$

$$\mathrm{M}[N(N-1)] = \sum_{n=0}^{\infty} n(n-1) k_n = K''(1)$$

и

$$K''(s) = \lambda^2 \int_0^\infty v^2 e^{-\lambda v(1-s)} b(v)\, dv.$$

Следовательно,

$$M[N(N-1)] = \lambda^2 \int_0^\infty v^2 b(v)\, dv = \lambda^2 \{\sigma^2(V) + [M(V)]^2\} = \sigma^2(\lambda V) + \rho^2,$$

поскольку

$$\rho = \lambda M(V).$$

Таким образом

$$M(q) = \rho + \frac{\sigma^2(\lambda V) + \rho^2}{2(1-\rho)}. \qquad (4.6)$$

В. Среднее время ожидания

В условиях стационарности можно найти среднее время ожидания требования. Предположим, что требование ожидает время W начала своего обслуживания, которое продолжается время V. Предположим, что, когда требование покидает систему, в очереди остается q требований. Это означает, что за время $W + V$ поступило q требований пуассоновского потока. В силу стационарности сумма

(среднее время ожидания) + (среднее время обслуживания требования) = $M(W) + M(V)$

должна равняться среднему числу требований, поступивших за этот период, умноженному на среднюю длительность интервала между требованиями, т. е. величине $\frac{1}{\lambda} M(q)$. Но поскольку рассматривается стационарный режим, то в силу (4.6)

$$M(W) + M(V) = \frac{1}{\lambda}\left[\rho + \frac{\sigma^2(\lambda V) + \rho^2}{2(1-\rho)}\right].$$

Деля на $\mu^{-1} = M(V)$ и вспоминая, что $\frac{\lambda}{\mu} = \rho$, из последнего соотношения получим

$$\frac{M(W)}{M(V)} = \frac{\sigma^2(\lambda V) + \rho^2}{2\rho(1-\rho)},$$

или

$$M(W) = \rho\, \frac{\sigma^2(\mu V) + 1}{2(1-\rho)} M(V). \qquad (4.7)$$

Из формул (4.6) и (4.7) следует один несколько неожиданный факт. Именно: при заданных средних интервалах между поступлениями и длительностях обслуживания можно уменьшить средние длину очереди и время ожидания, уменьшая дисперсию времени обслуживания. Очевидно, что наилучшим в этом отношении случаем является постоянное время обслуживания.

Г. Распределение времени ожидания

В тех же предположениях, что и выше, найдем преобразование Лапласа распределения времени ожидания. Пусть $\{\pi_i\}$ — равновесные вероятности, производящая функция которых $\pi(s)$ была найдена выше. Если требование ожидает время W начала обслуживания и обслуживается время V, то вероятность того, что после его ухода останется q требований, равная π_q, совпадает с вероятностью поступления q требований за время $W+V$. Поскольку входящий поток — пуассоновский с параметром λ, то

$$\pi_q = \int_0^\infty e^{-\lambda t}\frac{1}{q!}(\lambda t)^q\,dC(t),$$

где $C(t)$ — функция распределения величины $W+V$. Отсюда

$$\pi(s) = \sum_{q=0}^{\infty}\pi_q s^q = \int_0^\infty e^{-\lambda t(1-s)}\,dC(t) = \tilde{C}(\lambda(1-s)),$$

где $\tilde{C}(s)$ — преобразование Лапласа функции $C(t)$ [1]. Но $C(t)$ — функция распределения суммы независимых случайных величин W и V, функции распределения которых равны $W(t)$ и $B(t)$ соответственно. Преобразование Лапласа для суммы W и V равно произведению соответствующих преобразований. Следовательно,

$$\pi(s) = \tilde{W}(\lambda(1-s))\,\tilde{B}(\lambda(1-s)),$$

или

$$\tilde{W}(u) = \frac{\pi\left(\frac{\lambda-u}{\lambda}\right)}{\tilde{B}(u)}.$$

§ 5. ЭКСПОНЕНЦИАЛЬНО РАСПРЕДЕЛЕННОЕ ВРЕМЯ ОБСЛУЖИВАНИЯ ($G/M/1$)

Другой моделью, которая может быть изучена с помощью метода вложенных цепей Маркова, является модель с произвольным распределением $H(u)$ интервалов между моментами поступлений и экспоненциально распределенным (с параметром μ) временем обслуживания.

---

[1]) Такое преобразование часто называют преобразованием Лапласа — Стильтьеса. — *Прим. перев.*

В этом случае примем, что переходы вложенной цепи Маркова определяются моментами поступления новых требований, а ее состояние до следующего перехода — длина очереди перед вновь поступившим требованием.

Если q — состояние системы после некоторого момента поступления, а q' — состояние после следующего момента поступления, то

$$q' = q + 1 - N, \tag{5.1}$$

где N — число обслуженных требований за рассматриваемый отрезок времени. В силу свойства «отсутствия последействия» у экспоненциального распределения (см. теорему 2.2 гл. 7) число N требований, обслуженных за время между моментами поступлений, зависит только от длины этого интервала и величины q и не зависит от времени, в течение которого уже обслуживалось очередное требование. Интервалы между моментами поступлений являются, конечно, независимыми случайными величинами. В силу указанных фактов заключаем, что соотношение (5.1) определяет цепь Маркова. Вычислим ее матрицу перехода $\|P_{ij}\|$.

Поскольку $N \geqslant 0$, то $P_{ij} = 0$ при $j > i + 1$. Если $i + 1 \geqslant j \geqslant 1$, то $i + 1 - j$ требований было обслужено до поступления очередного требования. Обозначим вероятность этого события через a_{i+1-j}. Очевидно, если $i + 1 \geqslant j \geqslant 1$, то $P_{ij} = a_{i+1-j}$.

Целесообразно найти выражение для a_k через распределение интервалов между моментами поступления и распределение времени обслуживания. Для этого заметим, что если длина интервала между моментами поступления равна ξ, то вероятность того, что завершится обслуживание в точности k требований, равна

$$F^{(k)}(\xi) - F^{(k+1)}(\xi), \tag{5.2}$$

где $F(\xi) = 1 - e^{-\mu\xi}$ — распределение времени обслуживания, а $F^{(k)}(\xi)$ — k-кратная свертка $F(\xi)$. Действительно, пусть $\Xi_1, \Xi_2, \ldots$ $\ldots, \Xi_r, \ldots$ — длительности первой, второй и т. д. операций обслуживания. Величины Ξ_i независимы и одинаково распределены по закону $F(\xi)$. Вероятность того, что за время ξ закончатся по меньшей мере k операций обслуживания, совпадает с вероятностью того, что временной интервал до окончания k-го акта обслуживания не превышает ξ, т. е.

$$P\{\Xi_1 + \Xi_2 + \ldots + \Xi_k \leqslant \xi\} = F^{(k)}(\xi).$$

Следовательно, вероятность завершения за время ξ в точности k операций обслуживания равна P {время, необходимое для завершения по крайней мере k операций обслуживания $\leqslant \xi$} — P {время, необходимое для завершения по крайней мере $k + 1$ операций обслуживания $\leqslant \xi\}$. Отсюда следует формула (5.2). Точное выражение для $F^{(k)}(\xi)$ следующее:

$$F^{(k)}(\xi) = \int_0^{\xi} e^{-\mu t} \frac{t^{k-1}\mu^k}{\Gamma(k)}\, dt.$$

Интегрируя соответствующую формулу для $F^{(k+1)}(\xi)$ по частям, получим

$$F^{(k)}(\xi) - F^{(k+1)}(\xi) = e^{-\mu\xi} \frac{\xi^k \mu^k}{\Gamma(k+1)}.$$

В силу формулы полной вероятности имеем

$$a_k = \int_0^{\infty} e^{-\mu\xi} \frac{\xi^k \mu^k}{\Gamma(k+1)}\, dH(\xi),$$

где $H(\xi)$ — функция распределения интервалов между моментами поступления. Формула для a_k может быть выведена непосредственным образом. Однако метод, приведенный выше, имеет самостоятельную ценность и может использоваться при решении других задач.

Наконец, величина P_{i0} — вероятность того, что все имевшиеся i требований были обслужены, равна вероятности того, что при наличии более чем i требований по крайней мере i из них были бы обслужены. Лучше всего записать $P_{i0} = 1 - \sum_{j=1}^{\infty} P_{ij}$; тогда

$$\mathbf{P} = \left\| \begin{matrix} r_0 & a_0 & 0 & 0 & 0 \dots \\ r_1 & a_1 & a_0 & 0 & 0 \dots \\ r_2 & a_2 & a_1 & a_0 & 0 \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix} \right\|, \tag{5.3}$$

где $r_i = 1 - a_0 - a_1 - \dots - a_i$.

Цепь Маркова с матрицей переходных вероятностей (5.3) была рассмотрена в § 6 гл. 3 и там был проведен довольно подробный анализ, касающийся таких свойств, как положительная возвратность и невозвратность. В частности, было доказано, что если

$$\sum_{k=0}^{\infty} k a_k > 1,$$

то цепь Маркова является возвратной положительной и предельное распределение имеет вид

$$\pi_i = (1 - \xi_0)\xi_0^i, \qquad i = 0,\ 1,\ 2,\ \dots,$$

где ξ_0 — единственное решение уравнения $f(\xi_0)=\xi_0\,(0<\xi_0<1)$, а

$$f(\xi)=\sum_{k=0}^{\infty} a_k\xi^k.$$

В силу определения a_k имеем

$$f'(1)=\sum_{k=0}^{\infty} ka_k=\frac{\text{средняя длина интервала между моментами поступления}}{\text{среднее время обслуживания}}=\frac{1}{\rho}.$$

Следовательно, процесс является возвратным положительным тогда и только тогда, когда $\rho<1$.

Время ожидания

Если $f'(1)>1$ и функция распределения длины очереди является стационарной, найдем функцию распределения времени ожидания W начала обслуживания.

Вероятность того, что требование не будет ожидать в очереди, равна

$$\pi_0=1-\xi_0.$$

Если требование поступает и застает впереди себя $n\geqslant 1$ требований, то оно должно ожидать время, равное сумме n независимых одинаково (экспоненциально) распределенных длительностей обслуживания, прежде чем поступит на обслуживающий прибор. Эта сумма обладает гамма-распределением порядка n с параметром μ. Таким образом,

$$\mathrm{P}\{W\leqslant t\mid\text{длина очереди равна } n\}=\int_0^t \frac{\mu^n\tau^{n-1}e^{-\mu\tau}}{\Gamma(n)}\,d\tau,\qquad n\geqslant 1.$$

Следовательно, в силу того, что

$$\mathrm{P}\{\text{длина очереди равна } n\}=\pi_n=(1-\xi_0)\,\xi_0^n,$$

имеем

$$W(t)=\mathrm{P}\{W\leqslant t\}=\sum_{n=1}^{\infty}\mathrm{P}\{W\leqslant t\mid\text{ длина очереди равна } n\}\times$$

$$\times\,\mathrm{P}\{\text{длина очереди равна } n\}+\pi_0=$$

$$=(1-\xi_0)\int_0^t\sum_{n=1}^{\infty}\frac{\mu^n\tau^{n-1}}{\Gamma(n)}e^{-\mu\tau}\xi_0^n\,d\tau+(1-\xi_0)=$$

$$=(1-\xi_0)+\xi_0\,(1-\exp[-\mu t\,(1-\xi_0)]).$$

Это распределение является комбинацией экспоненциального с параметром $\mu(1-\xi_0)$ и вырожденного (сосредоточенного в точке 0). Последнее имеет вес $1-\xi_0$, который равен вероятности того, что поступившее требование не будет ожидать начала обслуживания. Условная функция распределения времени ожидания при условии занятости обслуживающего устройства равна при этом

$$\Omega(t) = 1 - \exp[-\mu t(1-\xi_0)].$$

§ 6. ГАММА-РАСПРЕДЕЛЕНИЕ ИНТЕРВАЛОВ МЕЖДУ ПОСТУПЛЕНИЯМИ И ОБОБЩЕНИЯ ($E_k/M/1$)

Это частный случай предыдущей модели, который может быть исследован с помощью изящного приема, широко применяемого также и в других задачах[1]). Рассмотрим одноканальную систему с экспоненциально (с параметром μ) распределенным временем обслуживания и с интервалами между моментами поступления, имеющими гамма-распределение $H(u)$ порядка k с плотностью

$$h(u) = \begin{cases} \frac{1}{\Gamma(k)} \lambda^k u^{k-1} e^{-\lambda u}, & u > 0, \\ 0, & u \leqslant 0. \end{cases}$$

Функцию распределения $H(u)$ можно считать распределением суммы k независимых случайных величин, каждая из которых распределена экспоненциально с параметром λ. Следовательно, можно свести задачу к анализу марковского процесса, считая, что каждое поступление состоит из k *фаз* $0, 1, \ldots, k-1$, в каждой из которых требование проводит экспоненциально (с параметром λ) распределенное время, а затем переходит к следующей фазе. Действительное поступление требования в систему соответствует его достижению k-й фазы. В любой момент времени лишь одно требование находится в одной из фаз $0, 1, \ldots, k-1$, причем новое требование поступает на фазу 0 в момент, когда предшествующее требование покидает фазу $k-1$.

Состояние системы определяется как сумма соответствующих всем требованиям фаз. Таким образом, если система находится в состоянии $nk+l$, $l<k$, это означает, что n требований находится либо в очереди, либо на обслуживании (что соответствует k-й фазе), а еще одно требование находится на «l-й фазе поступления». Когда какое-либо требование заканчивает обслуживаться, состояние системы убывает на k.

---

[1]) Излагаемый ниже прием получил название *метода фиктивных фаз Эрланга*. — *Прим. перев.*

Тем самым определена цепь Маркова с непрерывным временем, инфинитезимальная матрица которой равна

$$
\mathbf{A} = \begin{Vmatrix}
\multicolumn{7}{c}{\overbrace{\hphantom{-\lambda \quad \lambda \quad 0 \quad 0 \quad \cdots \quad \cdots \quad 0}}^{k}} & & & & \\
-\lambda & \lambda & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & 0 \ldots \\
0 & -\lambda & \lambda & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & 0 \ldots \\
0 & 0 & -\lambda & \lambda & \cdots & \cdots & 0 & 0 & 0 & 0 & 0 \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \cdots & 0 & -\lambda & \lambda & 0 & 0 & 0 \ldots \\
\mu & 0 & 0 & \cdots & \cdots & 0 & 0 & -(\mu+\lambda) & \lambda & 0 & 0 \ldots \\
0 & \mu & 0 & \cdots & \cdots & 0 & 0 & 0 & -(\mu+\lambda) & \lambda & 0 \ldots \\
0 & 0 & \mu & \cdots & \cdots & 0 & 0 & 0 & 0 & -(\mu+\lambda) & \lambda \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{Vmatrix}.
$$

Можно найти равновесные свойства данной цепи Маркова с непрерывным временем. Мы не будем этим заниматься, поскольку, как упомянуто выше, рассматриваемый пример является частным случаем системы $(G/M/1)$, которая рассматривалась в § 5. Преимущество описанного приема состоит в том, что, используя марковский характер процесса, можно найти его переходные характеристики. Обсуждение этого вопроса выходит за рамки данной книги.

А. Гамма-распределение времени обслуживания и произвольный (рекуррентный) входящий поток [1])

Можно использовать приемы, изложенные в последних нескольких параграфах, для нахождения стационарных характеристик процесса обслуживания с одним обслуживающим прибором с произвольным распределением $H(v)$ интервалов между моментами поступления и гамма-распределением порядка k с параметром μ времени обслуживания.

Будем считать, что обслуживание состоит из k фаз $1, 2, \ldots, k$, в каждой из которых требование проводит экспоненциально (с параметром μ) распределенное случайное время. После того как обслуживание на k-й фазе завершилось, требование покидает систему.

Построим вложенную цепь Маркова, переходы которой обусловлены поступлением требований. Пусть состояние цепи между моментами поступления равно $kq - p + 1$, где q — число требований в системе, а p — номер фазы обслуживания, на которой находилось в момент последнего поступления обслуживаемое требование. Поскольку $k - p + 1$ — число фаз, которое осталось пройти обслуживаемому требованию, то состояние системы $kq - p + 1 = k(q - 1) +$

---

[1]) При первом чтении текст, напечатанный петитом, можно опустить без потери целостности изложения.

$+(k-p+1)$ можно интерпретировать как число экспоненциально распределенных интервалов времени, которые должны завершиться прежде, чем вновь поступившее требование начнет обслуживаться. Если $q=0$, то состояние системы полагается равным 0.

Вероятности перехода за один шаг для данной цепи имеют следующий вид. При всех i

(1) если $j > i+k$, то $P_{ij}=0$,

(2) если $j \leqslant i+k$, $j \neq 0$, то в одном интервале между моментами поступления содержится $i+k-j$ экспоненциально распределенных отрезков и

$$P_{ij}=\eta_r=\int\limits_0^{\infty} \frac{1}{r!} (\mu v)^r e^{-\mu v}\, dH(v),$$

где $r=i+k-j$. Вывод этого равенства идентичен выводу выражения для a_k на стр. 477,

(3) наконец, P_{i0} есть вероятность того, что сумма S_{i+k} $i+k$ экспоненциально распределенных отрезков времени не превышает длину интервала между моментами поступления:

$$P_{i0}=\int\limits_0^{\infty} \mathrm{P}\{S_{i+k} \leqslant v\}\, dH(v)=\int\limits_0^{\infty}\int\limits_0^{v} \frac{\xi^{i+k-1}\mu^{i+k}e^{-\xi\mu}}{(i+k-1)!}\, d\xi\, dH(v).$$

Б. Стационарные вероятности

Если нагрузка системы

$$\rho=\frac{\mathrm{M}\,(\text{время обслуживания})}{\mathrm{M}\,(\text{время между моментами поступления})}=$$

$$=\frac{k}{\mu \mathrm{M}\,(\text{время между моментами поступления})}$$

меньше 1, то мы ожидаем, что вероятностное распределение состояний системы стремится к предельному. Такое стационарное распределение пропорционально неотрицательной сходящейся последовательности $\mathbf{x}=(x_0, x_1, \ldots)$, удовлетворяющей равенству

$$\mathbf{x}=\mathbf{x}\mathbf{P}. \tag{6.1}$$

По аналогии с предыдущими моделями выберем пробное решение в виде

$$x_i=\lambda^i,$$

где λ — некоторое действительное число. При $j>k$ из (6.1) получаем

$$\lambda^j=\sum_{i=0}^{\infty} \lambda^i P_{ij}=\sum_{i=j-k}^{\infty} \lambda^i \eta_{i+k-j}=\sum_{r=0}^{\infty} \lambda^{j-k+r}\eta_r=\lambda^{j-k}\sum_{r=0}^{\infty}\lambda^r\eta_r, \tag{6.2}$$

или

$$\lambda^k=F(\lambda),$$

где

$$F(\lambda)=\sum_{r=0}^{\infty}\lambda^r\eta_r=\int\limits_0^{\infty} e^{-\mu v(1-\lambda)}\, dH(v).$$

С помощью теоремы Руше можно доказать, что функция $\lambda^k - F(\lambda)$ при $|\lambda| < 1$ имеет k корней (считая и кратные).

В частности, теорема Руше утверждает, что если $f(z)$ и $g(z)$ — аналитические функции в области D и $|f(z)| > |g(z)|$ при значениях z, принадлежащих границе D, то функции $f(z)$ и $f(z) + g(z)$ имеют в D одинаковое количество нулей. (Доказательство этой теоремы можно найти в любом учебнике по теории функций комплексного переменного.) Применим теперь теорему Руше для случая $D = \{z \mid |z| \leqslant 1 - \delta, \delta > 0\}$, $f(z) = z^k$, $g(z) = -F(z)$. В самом деле, при $|z| = 1 - \delta$ имеем $|z^k| = (1-\delta)^k = 1 - k\delta + o(\delta)$. Далее, при $|z| = 1 - \delta$ $|F(z)| \leqslant F(1-\delta)$ (поскольку $F(z)$ — степенной ряд с неотрицательными коэффициентами). Но при $\delta \to 0$

$$F(1-\delta) = F(1) - \delta F'(1) + o(\delta).$$

Кроме того, непосредственный подсчет показывает, что

$$F'(1) = \mu \int_0^\infty v\, dH(v) = \frac{k}{\rho} > k,$$

поскольку по предположению $\rho < 1$. Отсюда следует, что при достаточно малых δ $|z^k| > |F(z)|$, $|z| = 1 - \delta$. В силу теоремы Руше заключаем, что функции z^k и $z^k - F(z)$ имеют в области $\{z \mid |z| \leqslant 1 - \delta\}$ одинаковое количество нулей.

Если обозначить k корней через $\lambda_1, \lambda_2, \ldots, \lambda_k$, то последовательность $\{x_n = \lambda_r^n\}_{n=0}^\infty$ будет удовлетворять уравнению (6.2) при любом r $(r = 1, 2, \ldots, k)$. Можно попытаться найти линейную комбинацию

$$\pi_n = \alpha_1\lambda_1^n + \alpha_2\lambda_2^n + \ldots + \alpha_k\lambda_k^n, \qquad \sum_{r=1}^k \alpha_r = 1,$$

такую, что система уравнений $\mathbf{x} = \mathbf{xP}$ удовлетворяется при $x_n = \pi_n$, $n = 0, 1, \ldots$. Действительно, уравнения $x_j = \sum_{i=0}^\infty x_i P_{ij}$ $(j = 0, 1, \ldots)$ при $j \geqslant k$ удовлетворяются при любом выборе коэффициентов α_r, поскольку любая последовательность $\{\lambda_r^n\}$, $r = 1, \ldots, k$, является решением и их линейная комбинация вновь является решением. Остается найти постоянные $\alpha_1, \alpha_2, \ldots, \alpha_k$, такие, чтобы первые k уравнений системы $\mathbf{x} = \mathbf{xP}$ также удовлетворялись. В случае если все λ_r различны, совершая некоторые алгебраические преобразования, получим точное решение $\left(\text{в котором произведена нормировка } \sum_{n=0}^\infty \pi_n = 1\right)$

$$\pi_n = \frac{1}{\sum_{i=1}^k \frac{a_i}{1-\lambda_i}} \sum_{i=1}^k a_i \lambda_i^n,$$

где

$$a_i = \prod_{\substack{j \neq i \\ j=1}}^k \left(\frac{\lambda_i}{\lambda_i - \lambda_j}\right).$$

Изменения, которые необходимо внести в случае кратных корней, а также соответствующие подробности утомительны и будут опущены.

В. Время ожидания

Выше было указано, что состояние системы определялось как число экспоненциально (с параметром μ) распределенных отрезков времени, которые составляют время ожидания вновь поступившего требования. Следовательно, если состояние системы равно $n > 0$, то время ожидания только что поступившего требования имеет гамма-распределение порядка n с параметром μ. Если $n = 0$, время ожидания равно 0. Следовательно,

$$W(\xi) = P\{W \leqslant \xi\} = \int_0^{\xi} \sum_{j=1}^{\infty} \frac{1}{(j-1)!} \mu^j w^{j-1} e^{-\mu w} \pi_j dw + \pi_0.$$

В случае различных корней можно подставить выражение для π_j, приведенное выше, и получить

$$W(\xi) = \frac{1}{\sum_{i=1}^{k} \frac{\alpha_i}{1-\lambda_i}} \left\{ \int_0^{\xi} \sum_{j=1}^{\infty} \frac{\mu^j w^{j-1} e^{-\mu w}}{(j-1)!} \sum_{i=1}^{k} \alpha_i \lambda_i^j dw + 1 \right\} =$$

$$= \frac{1}{\sum_{i=1}^{k} \frac{\alpha_i}{1-\lambda_i}} \left\{ 1 + \int_0^{\xi} \sum_{i=1}^{k} \alpha_i \mu \lambda_i \exp\left[-\mu w (1-\lambda_i)\right] dw \right\} =$$

$$= \frac{1}{\sum_{i=1}^{k} \frac{\alpha_i}{1-\lambda_i}} \left\{ 1 + \sum_{i=1}^{k} \frac{\alpha_i \lambda_i}{(1-\lambda_i)} \left(1 - \exp\left[-\mu \xi (1-\lambda_i)\right]\right) \right\} =$$

$$= 1 - \frac{\sum_{i=1}^{k} \frac{\alpha_i \lambda_i}{1-\lambda_i} \exp\left[-\mu \xi (1-\lambda_i)\right]}{\sum_{i=1}^{k} \frac{\alpha_i}{1-\lambda_i}}.$$

§ 7. ЭКСПОНЕНЦИАЛЬНОЕ ОБСЛУЖИВАНИЕ s ПРИБОРОВ ($GI/M/s$)

Обобщим рассмотренные выше для случая одного обслуживающего прибора приемы на случай системы, состоящей из s приборов, в которую требования поступают через интервалы, имеющие распределение $H(v)$, а распределение времени обслуживания экспоненциально с параметром μ. Предположим, что распределения длительностей обслуживания для всех приборов совпадают.

Как и прежде, процесс обслуживания не является марковским, но можно построить вложенную цепь Маркова. Пусть переходы цепи определяются моментами поступления новых требований, и пусть q, состояние системы, есть число требований, ожидающих или обслуживаемых, которые находились в системе в момент поступления последнего требования.

Вероятность P_{ij} можно найти следующим образом:

(1) Если $j > i+1$, то $P_{ij} = 0$ при всех $i = 0, 1, 2, \ldots$.

(2) Если $j \leqslant i+1 \leqslant s$, то все требования находились на обслуживании и до следующего момента поступления закончилось $i-j+1$ актов обслуживания Вероятность того, что за время t фиксированное требование обслужится, равна $1-e^{-\mu t}$. Таким образом,

$$P_{ij} = \int_0^\infty \mathrm{P}\{i+1-j \text{ актов обслуживания закончится за время } t \mid \text{ в момент 0 в системе было } i+1 \text{ требований}\}\, dH(t) =$$

$$= \int_0^\infty \binom{i+1}{j} (1-e^{-\mu t})^{i+1-j} e^{-\mu t j}\, dH(t). \qquad (7.1)$$

Подинтегральная функция является вероятностью из биномиального распределения, соответствующей $i+1-j$ успехам (завершениям актов обслуживания до следующего момента поступления).

(3) Если $i+1 \geqslant j \geqslant s$ и $i \geqslant s$, то все приборы до следующего момента поступления заняты обслуживанием. Следовательно,

$$P_{ij} = \mathrm{P}\{\text{окончилось } i+1-j \text{ актов обслуживания}\} =$$

$$= \int_0^\infty \mathrm{P}\{\text{за время } t \text{ окончилось } i+1-j \text{ актов обслуживания}\}\, dH(t) =$$

$$= \int_0^\infty \frac{1}{(i+1-j)!} e^{-\mu s t} (\mu s t)^{i+1-j}\, dH(t).$$

(Вывод последнего равенства совпадает с выводом, приведенным после формулы (5.2). В нашем случае функция распределения времени до завершения очередного акта обслуживания является экспоненциальной с параметром $s\mu$, поскольку в рассматриваемом случае заняты все s приборов.)

(4) Если $i+1 \geqslant s > j$, то в начале рассматриваемого интервала $m = i-s+1$ требований будут ожидать, а s требований обслуживаться. В конце интервала будет $n = s-j$ свободных приборов. Пусть v — время до того момента, когда не останется ожидающих требований, т. е. время до окончания обслуживания m требований всеми s работающими приборами. Поскольку интервалы между моментами окончаний актов обслуживания распре-

делены экспоненциально с параметром $s\mu$, то величина v имеет гамма-распределение порядка m с параметром $s\mu$. Предположим, что оставшиеся n требований обслуживаются время $u-v$, где u — длительность интервала между соседними моментами поступления. Тогда

$$P_{ij}=\int_0^\infty \mathrm{P}\{\text{за время } u \text{ обслужилось } m+n \text{ требований}\}\, dH(u)=$$

$$=\int_0^\infty\Big[\int_0^\infty \mathrm{P}\{\text{за время } u-v \text{ обслужилось } n \text{ требований}\}\times$$

$$\times\frac{1}{(m-1)!}e^{-s\mu v}(s\mu)^m v^{m-1}\,dv\Big]dH(u)=$$

$$=\frac{(s\mu)^m}{(m-1)!}\int_0^\infty\int_0^u v^{m-1}e^{-s\mu v}\binom{s}{n}e^{-\mu(u-v)(s-n)}\left(1-e^{-\mu(u-v)}\right)^n dv\,dH(u),$$

где последнее равенство получено с помощью биномиального распределения, аналогично равенству (7.1).

А. Стационарные вероятности

Естественно ожидать, что если нагрузка системы <1, т. е.

$$\rho=\frac{\text{M (длина интервала между окончаниями актов обслуживания, когда все приборы заняты)}}{\text{M (длина интервала между моментами поступлений)}}=$$

$$=\frac{1}{s\mu \text{M (длина интервала между моментами поступлений)}}<1,$$

то спустя достаточно долгое время вероятности пребывания в каждом состоянии должны стабилизироваться. Найдем положительный вектор $\mathbf{x}=(x_0, x_1, x_2, \ldots)$, удовлетворяющий соотношениям $\Sigma x_i<\infty$ и $\mathbf{x}=\mathbf{xP}$. Сравнивая с рассмотренным ранее частным случаем одноканальной системы, приходим к рассмотрению пробного решения вида

$$\mathbf{x}=(\beta_0, \beta_1, \ldots, \beta_{s-2}, 1, \alpha, \alpha^2, \ldots).$$

В системе $\mathbf{x}=\mathbf{xP}$ j-е $(j>s)$ уравнение имеет вид

$$x_j=\alpha^{j-s+1}=\sum_{i=0}^\infty x_iP_{ij}=\sum_{i=j-1}^\infty x_iP_{ij}=\sum_{i=j-1}^\infty \alpha^{i-s+1}P_{ij}=$$

$$=\alpha^{j-s}\int_0^\infty e^{-\mu su(1-\alpha)}dH(u).$$

Это уравнение вида $\alpha = F(\alpha)$, где

$$F(\alpha) = \int_0^\infty e^{-\mu s u(1-\alpha)}\, dH(u)$$

— выпуклая возрастающая на $(0, 1)$ функция, $F(0) > 0$ и $F(1) = 1$. Выпуклость функции F можно проверить двойным дифференцированием. Следовательно, решение α в $(0, 1)$ существует тогда и только тогда, когда $F'(1) > 1$. Поскольку

$$F'(1) = \mu s \int_0^\infty u\, dH(u),$$

то это условие как раз совпадает с условием $\rho < 1$. Найдя α, можно найти и остальные компоненты $\beta_{s-2}, \beta_{s-3}, \ldots, \beta_0$ из рекуррентных соотношений

$$\beta_j = \sum_{i=j-1}^{s-2} \beta_i P_{ij} + \sum_{i=s-1}^{\infty} \alpha^{i-s+1} P_{ij}, \qquad j = 0, 1, \ldots, s-2,$$

или

$$\beta_{j-1} = (P_{j-1,\, j})^{-1} \left[\beta_j - \sum_{i=j}^{s-2} \beta_i P_{ij} - \sum_{i=s-1}^{\infty} \alpha^{i-s+1} P_{ij}\right], \quad j = 1, 2, \ldots, s-1,$$

где $\beta_{s-1} = 1$. Нормализуя, получим финальные вероятности

$$\pi_j = \frac{x_j}{(1-\alpha)^{-1} + \sum_{i=0}^{s-2} \beta_i}.$$

Б. Время ожидания в стационарном режиме

Вероятность того, что поступившее требование не будет ожидать начала обслуживания, совпадает с вероятностью того, что в момент поступления по крайней мере один прибор свободен, и равна

$$W(0) = \mathrm{P}\{q \leqslant s-1\} = \sum_{i=0}^{s-1} \pi_i = \frac{1 + \sum_{i=0}^{s-2} \beta_i}{(1-\alpha)^{-1} + \sum_{i=0}^{s-2} \beta_i} = A \sum_{i=0}^{s-1} \beta_i,$$

где

$$A = \left[(1-\alpha)^{-1} + \sum_{i=0}^{s-2} \beta_i\right]^{-1}, \qquad \beta_{s-1} = 1.$$

Если состояние системы равно $n \geqslant s$, то вновь поступившее требование должно ожидать начала обслуживания до тех пор, пока не обслужатся $n-s+1$ требований, стоящих перед ним. Но поскольку работают все s приборов, интервалы между окончаниями актов обслуживания распределены экспоненциально с параметром μs. Таким образом, время ожидания такого требования имеет гамма-распределение порядка $n-s+1$ с параметром μs, и

$$P\{W \leqslant \xi\} = W(\xi) = W(0) + A \int_0^{\xi} \sum_{n=s}^{\infty} \frac{1}{(n-s)!} (\mu s)^{n-s+1} w^{n-s} e^{-\mu s w} \alpha^{n-s} dw =$$

$$= A \left[\sum_{i=0}^{s-1} \beta_i + \int_0^{\xi} \mu s e^{-\mu s w (1-\alpha)} dw \right] =$$

$$= 1 - \frac{1}{1 + (1-\alpha) \sum_{i=0}^{s-2} \beta_i} e^{-\mu s \xi (1-\alpha)}.$$

§ 8. ВИРТУАЛЬНОЕ ВРЕМЯ ОЖИДАНИЯ И ПЕРИОД ЗАНЯТОСТИ

В этом параграфе рассматривается другой подход к задаче о времени ожидания для одноканальной системы обслуживания с пуассоновским входящим потоком и произвольным распределением времени обслуживания $(M/G/1)$. Будут также получены некоторые результаты, относящиеся к периоду занятости данной системы. Для этой цели будет использован материал гл. 9, § 3.

Поскольку процесс обслуживания в целом является немарковским, мы выше рассматривали вложенную цепь Маркова и анализировали времена ожидания требований, используя ее. Однако если рассмотреть величину $\eta(t)$ ($\eta(t)$ иногда называют *виртуальным временем ожидания*), которая равна времени, в течение которого должно было бы ожидать требование, если бы оно поступило в момент t, то $\eta(t)$ определяет марковский процесс с непрерывным временем. Так, если t_n и v_n — момент поступления и время обслуживания n-го требования соответственно, то при $t_n < t < t_{n+1}$ имеем

$$\eta(t) = [\eta(t_n+) - (t - t_n)]_+ \text{[1]}$$

и

$$\eta(t_n+) = \eta(t_n-) + v_n,$$

$$\left(\eta(t+) = \lim_{\varepsilon \to 0} \eta(t+\varepsilon), \quad \eta(t-) = \lim_{\varepsilon \to 0} \eta(t-\varepsilon)\right).$$

---

[1]) Символ $[x]_+$ определяется равенством

$$[x]_+ = \begin{cases} x, & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Обозначим через λ параметр входящего потока (который по предположению пуассоновский), а через $H(v)$ — распределение времени обслуживания. Типичная реализация процесса $\eta(t)$ показана на рис. 1. Очевидно, что будущее поведение траектории $\eta(t)$ не зависит от ее предыстории до попадания в текущее состояние. В самом деле, поскольку значения t_j являются моментами наступления событий пуассоновского потока, то время для следующего поступления требования не зависит от того, когда поступило предыдущее требование.

Можно дать величине $\eta(t)$ другую интерпретацию как времени, необходимому для завершения обслуживания всех требований,

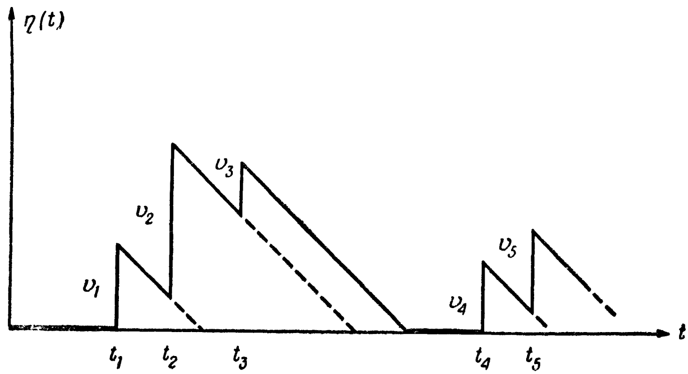


Рис. 1.

имеющихся в системе в момент t. Действительное время ожидания n-го требования равно $\eta(t_n) = \eta(t_n -)$. Обозначим

$$F(t,\ x) = \mathrm{P}\{\eta(t) \leqslant x\}.$$

Можно вывести интегро-дифференциальное уравнение, которому удовлетворяет $F(t, x)$. Это уравнение может быть проанализировано с целью нахождения свойств функции $F(t, x)$. В § 3 было доказано, что функция распределения времени ожидания n-го требования

$$F_n(x) = F(t_n,\ x)$$

сходится к предельному распределению при $n \to \infty$. С помощью соответствующих теорем восстановления можно доказать также сходимость функции $F(t, x)$ и показать, что ее предел совпадает с пределом функций $F_n(x)$ [1]). Дальнейшие подробности выходят за рамки данной книги (см. литературу в конце данной главы).

[1]) Заметим, что последний факт не является очевидным в силу того, что моменты t_n случайны и совпадают с моментами скачков процесса $\eta(t)$, т. е. последовательность $\{t_n\}$ и функция $\eta(t)$ не являются независимыми. — *Прим. перев.*

В оставшейся части главы будут рассмотрены различные представляющие интерес случайные величины, связанные с системой $(M/G/1)$.

Заметим, что если $\eta(t) > 0$, то обслуживающий прибор в момент t занят, а если $\eta(t) = 0$, то свободен. Пусть

$$P_0(t) = \mathrm{P}\{\eta(t) = 0\},$$

т. е. $P_0(t)$ — вероятность того, что прибор в момент t свободен.

Период занятости определяется как такой временной интервал, в течение которого прибор постоянно занят. Если $\eta(0) > 0$, т. е. обслуживающий прибор в момент $t = 0$ занят, то существует начальный период занятости, который заканчивается, когда $\eta(t)$ в первый раз обращается в 0. Обозначим через $\widehat{G}(x)$ вероятность того, что длительность начального периода занятости $\leqslant x$. Следующие за начальным периодом занятости (если таковой имеется) периоды незанятости и занятости чередуются. Длительности периодов занятости, следующих за начальным, являются независимыми одинаково распределенными случайными величинами, поскольку каждый последующий период занятости начинается в одних и тех же условиях. Обозначим через $G(x)$ вероятность того, что длительность периода занятости (отличного от начального) $\leqslant x$. Периоды незанятости также являются независимыми одинаково распределенными случайными величинами, функция распределения которых экспоненциальна с параметром λ.

Нашей первоочередной задачей является доказательство теоремы 8.1, сформулированной ниже. Это доказательство опирается на один результат, полученный в гл. 9, § 2, который мы для удобства приведем здесь в виде леммы 8.1.

Лемма 8.1. *Пусть $\chi_1, \chi_2, \ldots, \chi_n$ — неотрицательные перестановочные. случайные величины с суммой $\chi_1 + \chi_2 + \ldots + \chi_n = y$, пусть $\tau_1, \tau_2, \ldots, \tau_n$ — координаты (расположенные в порядке возрастания) n точек, распределенных равномерно и независимо друг от друга на интервале $(0, t)$. Если $\{\chi_k\}$ и $\{\tau_k\}$ — независимые последовательности, то*

$$\mathrm{P}\{\chi_1 + \ldots + \chi_k \leqslant \tau_k, \quad k = 1, 2, \ldots, n\} = \begin{cases} 1 - \dfrac{y}{t}, & \text{если } 0 \leqslant y \leqslant t, \\ 0, & \text{если } y > t. \end{cases} \tag{8.1}$$

Доказательство леммы 8.1 и ее применения даны в гл. 9 (см., в частности, § 2). Приведем еще одну лемму.

Лемма 8.2. *Пусть $\chi_1, \chi_2, \ldots, \chi_n$ — неотрицательные перестановочные случайные величины с суммой $\chi_1 + \chi_2 + \ldots + \chi_n = t$, и пусть $\tau_1, \tau_2, \ldots, \tau_{n-1}$ — координаты (расположенные в порядке* *возрастания) $n-1$ точек, распределенных равномерно и независимо друг от друга на интервале $(0, t)$. Если $\{\chi_k\}$ и $\{\tau_k\}$ — независимые последовательности, то*

$$P\{\chi_1 + \ldots + \chi_k \leqslant \tau_k,\ k = 1, 2, \ldots, n-1\} = \frac{1}{n}. \tag{8.2}$$

Доказательство. В силу леммы 8.1

$$P\{\chi_1 + \ldots + \chi_k \leqslant \tau_k,\ k = 1, 2, \ldots, n-1 \mid \chi_1 + \ldots + \chi_{n-1} = y\} = \begin{cases} 1 - \frac{y}{t}, & \text{если } 0 \leqslant y \leqslant t, \\ 0, & \text{если } y > t. \end{cases}$$

Далее, в силу формулы полной вероятности

$$P\{\chi_1 + \ldots + \chi_k \leqslant \tau_k,\ k = 1, 2, \ldots, n-1\} =$$
$$= \int_0^t P\{\chi_1 + \ldots + \chi_k \leqslant \tau_k,\quad k = 1, 2, \ldots, n-1 \mid \chi_1 + \ldots + \chi_{n-1} = y\} \times$$
$$\times\, dP\{\chi_1 + \ldots + \chi_{n-1} \leqslant y\} =$$
$$= \int_0^t \left(1 - \frac{y}{t}\right) dP\{\chi_1 + \ldots + \chi_{n-1} \leqslant y\} = 1 - \frac{1}{t} M(\chi_1 + \ldots + \chi_{n-1}) =$$
$$= 1 - \frac{1}{t}\left(\frac{n-1}{n} t\right),$$

поскольку $\chi_1, \ldots, \chi_n$ перестановочны и их сумма равна t. Отсюда, очевидно, получаем

$$P\{\chi_1 + \ldots + \chi_k \leqslant \tau_k,\quad k = 1, 2, \ldots, n-1\} = \frac{1}{n}. \ \blacksquare$$

Теорема 8.1. *Если $\eta(0) = c$ (c — постоянная), то вероятность того, что начальный период занятости имеет длительность $\leqslant x$, равна*

$$\hat{G}(x) = \sum_{n=0}^{\infty} \frac{1}{n!} c\lambda^n \int_0^{x-c} e^{-\lambda(c+y)} (c+y)^{n-1}\, dH_n(y), \quad \text{если } x \geqslant c,$$
$$\hat{G}(x) = 0, \qquad \text{если } x < c, \tag{8.3}$$

где $H_n(y)$ — n-кратная свертка функции $H(y)$ ($H_0(y)$ — распределение, имеющее единичный скачок в точке 0).

Доказательство. Проведем вычисления, налагая условие на число поступивших требований. Число поступлений за начальный период занятости может принимать значения $n = 0, 1, 2, \ldots$. Если $n = 0$, то начальный период занятости имеет длительность c, а вероятность того, что за время $(0, c)$ не поступит ни одного требования, равна $e^{-\lambda c}$. Таким образом, получен первый член $(n = 0)$ суммы (8.3). Если $n \geqslant 1$, то обозначим через $\tau_1, \tau_2, \ldots, \tau_n$ моменты поступлений требований, а через $\chi_1, \chi_2, \ldots, \chi_n$ — длительности их обслуживания. Эти величины должны удовлетворять условиям

$$\tau_j \leqslant \chi_1 + \ldots + \chi_{j-1} + c, \qquad j = 1, 2, \ldots, n, \tag{8.4}$$

где сумма по пустому множеству индексов полагается равной нулю. Действительно, соотношение (8.4) утверждает, что суммарное время обслуживания $j-1$ требований, поступивших после момента 0, и требований, имевшихся в системе в момент 0, превышает момент поступления j-го требования, $j = 1, 2, \ldots, n$. Это условие, очевидно, гарантирует, что обслуживающий прибор занят по крайней мере до завершения обслуживания n-го поступившего требования. Конечно, $P\{\chi_1 + \chi_2 + \ldots + \chi_n \leqslant x\} = H_n(x)$.

Если $\chi_1 + \ldots + \chi_n = y$, то длительность начального периода занятости равна $c + y$, а вероятность того, что в интервале $(0, c+y)$ поступит в точности n требований, равна $\frac{1}{n!}[\lambda(c+y)]^n e^{-\lambda(c+y)}$. Моменты поступления можно считать координатами (расположенными в возрастающем порядке) n точек, распределенных равномерно и независимо друг от друга в интервале $(0, c+y)$ (см. стр. 206, гл. 7). Далее, $\chi_1, \chi_2, \ldots, \chi_n$ — неотрицательные перестановочные случайные величины.

Вычитая неравенства (8.4) из $y + c$, получим эквивалентные соотношения

$$y + c - \tau_j \geqslant y - \chi_1 - \chi_2 - \ldots - \chi_{j-1}, \qquad j = 1, 2, \ldots, n. \tag{8.5}$$

Пусть $\tau^*_{n+1-j} = y + c - \tau_j$, и, поскольку $\chi_1 + \ldots + \chi_n = y$, неравенства (8.5) можно переписать в виде

$$\tau^*_{n+1-j} \geqslant \chi_n + \chi_{n-1} + \ldots + \chi_j, \qquad j = 1, 2, \ldots, n. \tag{8.6}$$

Но величины τ^*_{n+1-j}, $j = 1, \ldots, n$, очевидно, вновь распределены как n порядковых статистик равномерного распределения на $(0, c+y)$. Чтобы убедиться в этом, достаточно посмотреть на значения τ_j в обратном порядке, проходя интервал от точки $c+y$ к точке 0. Требуемое утверждение получается из соображений симметрии. Кроме того, величины $\chi_n, \chi_{n-1}, \ldots, \chi_{n+1-j}$ имеют совместное распределение, такое же, что и величины $\chi_1, \chi_2, \ldots, \chi_j$, в силу перестановочности. Следовательно, событие (8.6) имеет ту же вероятность, что и событие

$$\tau_j \geqslant \chi_1 + \chi_2 + \ldots + \chi_j, \qquad j = 1, 2, \ldots, n. \tag{8.7}$$

Обращаясь к лемме 8.1, мы заключаем, что вероятность события (8.7) равна

$$1-\frac{y}{c+y}=\frac{c}{c+y}.$$

Учитывая все приведенные факты, с помощью формулы полной вероятности получаем

$$\hat{G}(x)=\sum_{n=0}^{\infty}\int_{0}^{x-c} \mathrm{P}\,\{\text{начальный период занятости} \leqslant x \mid \eta(0)=c,\ n \text{ требований, поступивших за это время, имеют суммарное время обслуживания } y\} \times \mathrm{P}\,\{\text{за период занятости поступило } n \text{ требований} \mid \eta(0)=c, \text{ суммарное время обслуживания } n \text{ требований равно } y\}\, dH_n(y) =$$

$$=\sum_{n=0}^{\infty}\int_{0}^{x-c} \frac{c}{c+y} e^{-\lambda(c+y)} \frac{1}{n!}(c+y)^{n}\lambda^{n}\, dH_n(y)=$$

$$=\sum_{n=0}^{\infty} c\,\frac{\lambda^{n}}{n!}\int_{0}^{x-c} e^{-\lambda(c+y)}(c+y)^{n-1}\, dH_n(y). \qquad \blacksquare$$

В следующей теореме выводится функция распределения периода занятости, отличного от начального.

Теорема 8.2. *Вероятность того, что период занятости, отличный от начального, имеет длительность $\leqslant x$, равна*

$$G(x)=\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{n!}\int_{0}^{x} e^{-\lambda y} y^{n-1}\, dH_n(y), \qquad x \geqslant 0. \tag{8.8}$$

Доказательство. Если предположить, что период занятости состоит из n, $n=1, 2, \ldots$, актов обслуживания, то его длительность равна $\chi_1+\chi_2+\ldots+\chi_n$, где $\chi_1, \chi_2, \ldots, \chi_n$ — независимые одинаково распределенные случайные величины с функцией распределения $\mathrm{P}\{\chi_i \leqslant x\}=H(x)$, $i=1, 2, \ldots, n$. В этом случае за период занятости поступит в точности $n-1$ требований. Будем отсчитывать время от начала периода занятости и обозначим через $\tau_1, \tau_2, \ldots, \tau_{n-1}$ моменты поступлений. Они должны удовлетворять условиям

$$\tau_j \leqslant \chi_1+\ldots+\chi_j, \qquad j=1, 2, \ldots, n-1. \tag{8.9}$$

Если $\chi_1+\ldots+\chi_n=y$, то период занятости имеет длительность y, а моменты поступлений можно считать расположенными в порядке возрастания координатами n точек, равномерно и независимо друг от друга распределенных в интервале $(0, y)$. Далее, $\chi_1, \ldots, \chi_n$ — неотрицательные перестановочные случайные величины. Если

$\chi_1 + \ldots + \chi_n = y$, то событие (8.9) имеет ту же вероятность, что и событие

$$\chi_1 + \ldots + \chi_k \leqslant \tau_k, \quad k = 1, 2, \ldots, n-1, \tag{8.10}$$

поскольку в (8.9) можно заменить χ_j на χ_{n+1-j} и τ_j на $y - \tau_{n-j}$, $j = 1, 2, \ldots, n-1$, не изменяя вероятности события. По лемме 8.2 вероятность события (8.9) (и (8.10)) равна $\frac{1}{n}$. Поскольку $\mathrm{P}\{\chi_1 + \ldots + \chi_n \leqslant y\} = H_n(y)$, а вероятность того, что за интервал времени $(0, y)$ поступит в точности $n-1$ требований, равна $\frac{1}{(n-1)!}(\lambda y)^{n-1} e^{-\lambda y}$, то, применяя формулу полной вероятности, получим

$$G(x) = \sum_{n=1}^{\infty} \frac{1}{n} \int_0^x e^{-\lambda y} \frac{(\lambda y)^{n-1}}{(n-1)!}\, dH_n(y), \tag{8.11}$$

что и требовалось доказать. ■

ЗАДАЧИ

1. Показать, что стационарное распределение длины очереди $\{p_n,\ n = 0, 1, 2, \ldots\}$ для системы $(M/M/s)$ равно

$$p_0 = \left\{ \frac{(s\rho)^s}{s!\,(1-\rho)} + \sum_{i=0}^{s-1} \frac{(s\rho)^i}{i!} \right\}^{-1},$$

$$p_n = \begin{cases} p_0 \dfrac{(s\rho)^n}{n!}, & 1 \leqslant n \leqslant s, \\ p_0 \rho^n \dfrac{s^s}{s!}, & s < n < \infty, \end{cases}$$

где $\rho = \frac{\lambda}{s\mu} < 1$. Пусть $Q = \max(n - s, 0)$ $(n = 0, 1, 2, \ldots)$ — длина очереди без обслуживаемых требований. Показать, что

(1) $$\gamma = \mathrm{P}\{Q = 0\} = \left\{ \sum_{i=0}^{s} \left[\frac{(s\rho)^i}{i!} \right] + \left[\frac{(s\rho^s \rho}{s!\,(1-\rho)} \right] \right\}^{-1} \sum_{i=0}^{s} \frac{(s\rho)^i}{i!},$$

(2) $$\mathrm{M}(Q) = \frac{1-\gamma}{1-\rho}.$$

2. Сравнить системы типа $(M/M/1)$ с прямым и обратным порядками обслуживания (обратный порядок возникает, скажем, когда статьи из стопки берутся сверху). Насколько отличаются (если это отличие есть) распределения длин очередей, времен ожидания и периодов занятости?

Ответ: Распределения длин очередей и периодов занятости не отличаются. Отличаются распределения времен ожидания. Почему это так?

3. Рассмотрим систему $(M/M/1)$ с обратным порядком обслуживания. Пусть $X(t)$ — длина очереди в момент t. Показать, что процесс $\{X(t),\ t \geqslant 0\}$ является процессом рождения и гибели, и найти его параметры.

Ответ: $\lambda_n = \lambda$, $\mu_n = \mu$.

4. Рассмотрим систему с бесконечным числом обслуживающих приборов и экспоненциально (параметр μ) распределенным временем обслуживания. Предположим, что требования поступают группами, а интервалы между моментами поступлений распределены экспоненциально с параметром λ. Число требований в каждой группе случайно и имеет геометрическое распределение с параметром ρ $(0<\rho<1)$, т. е. P{число требований в группе равно $k\}=\rho^{k-1}(1-\rho)$ $(k=1, 2, \ldots)$.

Представить этот процесс в виде цепи Маркова с непрерывным временем и найти ее инфинитезимальную матрицу.

Ответ: $\mathbf{Q}=\|q_{ij}\|$, где

$$\begin{aligned} &q_{i,\,i-1}=i\mu, \quad i\geqslant 1,\\ &q_{ij}=\lambda\rho^{j-i-1}(1-\rho), \quad j>i,\\ &q_{ij}=0, \quad j<i-1,\\ &q_{ii}=-\sum_{j\neq i} q_{ij}.\end{aligned}$$

5 (продолжение). Найти производящую функцию $\pi(s)$ равновесного распределения процесса.

Ответ:

$$\pi(s)=\left[1+\frac{\rho}{1-\rho}(1-s)\right]^{-\frac{\lambda}{\rho\mu}}.$$

6. (Система с ограничениями) [1]). Требования поступают в систему по пуассоновскому потоку с параметром λ. Длительности их обслуживания — независимые одинаково распределенные д. с. в. с распределением $H(x)$. Требование, которое застает прибор занятым, присоединяется к очереди с вероятностью p $(0<p<1)$. Найти переходные вероятности вложенной цепи Маркова, построенной по моментам ухода требований из системы. Найти предельное распределение длины очереди.

Ответ:

$$\mathbf{P}=\left\|\begin{matrix} p_0 & p_1 & p_2\ldots\\ p_0 & p_1 & p_2\ldots\\ 0 & p_0 & p_1\ldots\\ 0 & 0 & p_0\ldots\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot \end{matrix}\right\|, \qquad p_j=\int_0^\infty e^{-\lambda p x}\frac{(\lambda p x)^j}{j!}\,dH(x),$$

$$K(s)=\sum_{j=0}^{\infty} p_j s^j, \qquad \pi(s)=\frac{(1-\rho)K(s)(s-1)}{s-K(s)},$$

где

$$\rho=\lambda\alpha p<1, \qquad \alpha=\int_0^\infty x\,dH(x)<\infty.$$

7. Рассмотрим систему, описанную в задаче 6. Пусть $H(x)=1-e^{-\mu x}$. Описать эту модель как случайный процесс рождения и гибели.

---

[1]) В оригинале «queueing with balking», т. е. дословно «процесс образования очереди с препятствиями». — *Прим. перев.*

Ответ:

$$\lambda_n = \begin{cases} \lambda, & n = 0, \\ \lambda_p, & n > 0, \end{cases} \qquad \mu_n = \begin{cases} 0, & n = 0, \\ \mu, & n > 0. \end{cases}$$

8. Следующие два процесса рождения и гибели (см. § 4 гл. 7) можно интерпретировать как модели обслуживания с ограничениями.

(а) $\lambda_n = \lambda q^n$, $0 < q < 1$, $\lambda > 0$, $n = 0, 1, 2, \ldots$,
$\mu_n = \mu$, $\mu > 0$,
$\mu_0 = 0$;

(б) $\lambda_n = \dfrac{\lambda}{n+1}$, $\mu_n = \mu$, $n = 1, 2, \ldots$,
$\mu_0 = 0$.

Для каждого случая найти стационарное распределение.
Ответ:

(а) $p_m = p_0 \left(\dfrac{\lambda}{\mu}\right)^m q^{\frac{m(m-1)}{2}}$, $m \geqslant 0$,

(б) $p_m = p_0 \left(\dfrac{\lambda}{\mu}\right)^m \dfrac{1}{m!}$, $m \geqslant 0$, $p_0 = e^{-\frac{\lambda}{\mu}}$.

9. Рассмотрим задачу о пешеходах, желающих перейти улицу с односторонним движением в заданной точке. Предположим, что автомобили (нулевой длины), которые движутся без остановок, проезжают мимо данной точки в моменты, образующие пуассоновский поток с параметром μ. Все ожидающие пешеходы перейдут дорогу, как только временно́е «окно» между автомобилями составит по крайней мере T секунд. Чему равны (1) распределение времени ожидания пешехода, который подошел к дороге в произвольный (не зависящий от движения автомобилей) момент времени, (2) распределение времени между моментом окончания перехода улицы некоторым пешеходом и следующим моментом начала возможного перехода? Ответ дать в терминах преобразований Лапласа. Найти среднее время ожидания пешехода.

Ответ: Распределения (1) и (2) имеют одно и то же преобразование Лапласа

$$L(s) = \left[(\mu + s)\, e^{-\mu T}/(s + \mu e^{-(\mu+s)T})\right];$$

$$\text{среднее время ожидания} = \frac{e^{\mu T} - (1 + \mu T)}{\mu}.$$

10. (Система $(M/G/\infty)$). Предположим, что имеется бесконечное число обслуживающих приборов и, следовательно, требования не ждут в очереди. Мы интересуемся числом занятых приборов. Моменты поступления требований образуют пуассоновский поток с параметром λ. Длительности обслуживания требований независимы и одинаково распределены по закону $H(x)$.

Найти

(1) $P_k(t) = \mathrm{P}\{$в момент t обслуживается ровно k требований$\}$,

(2) $\lim\limits_{t\to\infty} P_k(t) = P_k$ при условии, что в начальный момент в системе нет требований и $\alpha = \displaystyle\int_0^\infty x\, dH(x) < \infty$.

Указание: Использовать формулу полной вероятности и тот факт, что при условии поступления n требований за время t моменты поступления распределены как порядковые статистики размера n для равномерного распределения в $(0, t)$.

Ответ:

$$(1)\ P_k(t) = \exp\left(-\lambda \int_0^t (1 - H(x))\,dx\right)\left\{\lambda \int_0^t (1 - H(x))\,dx\right\}^k \frac{1}{k!},$$

$$(2)\ P_k = e^{-\lambda\alpha}\frac{1}{k!}(\lambda\alpha)^k,\quad \alpha = \int_0^\infty x\,dH(x).$$

11. В системе $(M/G/\infty)$ моменты поступления требований образуют пуассоновский поток с параметром λ, а функция распределения времени обслуживания равна $H(x)$. В начальный момент в системе нет ни одного требования. Показать, что вероятность того, что за время t будет обслужено n требований, равна

$$\frac{1}{n!}\left\{\lambda \int_0^t H(u)\,du\right\}^n \exp\left\{-\lambda \int_0^t H(u)\,du\right\}.$$

12. Рассмотрим процесс, описанный в задаче 11. Показать, что вероятность $\varphi(t, T)$ того, что в интервале $(t,\ t + T)$ ни одно требование не покинет систему, удовлетворяет рекуррентному соотношению

$$\varphi(t, T) = \int_0^t \lambda e^{-\lambda(t-\tau)}[H(\tau) + 1 - H(\tau + T)]\,\varphi(\tau, T)\,d\tau + \\ + \int_t^{t+T} \lambda e^{-\lambda\tau}[1 - H(T + t - \tau)]\,\varphi(0, T + t - \tau)\,d\tau + e^{-\lambda(t+T)}. \qquad (*)$$

Указание: Рассмотреть возможности, возникающие в момент поступления первого требования.

13. Используя результат задачи 12, доказать, что

$$\varphi(t, T) = \exp\left(-\lambda \int_t^{t+T} H(\xi)\,d\xi\right).$$

Указание: Вывести дифференциальное уравнение первого порядка (по переменной t) относительно функции $\varphi(t, T)$ и решить его.

14 (продолжение). Пусть $\varphi_n(t, T)$ — вероятность того, что в интервале $(t, t + T)$ n требований покинули систему. Получить интегральное соотношение типа (*) между φ_n и φ_{n-1}. Затем показать, что

$$\varphi_n(t, T) = \frac{1}{n!}\left\{\lambda \int_t^{t+T} H(\xi)\,d\xi\right\}^n \exp\left\{-\lambda \int_t^{t+T} H(\xi)\,d\xi\right\}.$$

15. В задаче **10** рассматривалась система с бесконечным числом обслуживающих приборов и пуассоновским входящим потоком. Рассмотрим «двойственную» систему $(GI/M/\infty)$, у которой интервалы между моментами поступления независимы и одинаково распределены с плотностью $h(x)$, а длительности обслуживания независимы и распределены по экспоненте с параметром μ. Число обслуживающих приборов бесконечно. Найти матрицу переходных вероятностей вложенной цепи Маркова, состояние которой η_n в момент n есть число занятых приборов в момент n-го поступления.

Ответ:

$$\mathrm{P}\{\eta_{n+1} = j \mid \eta_n = i\} = P_{ij} = \binom{i+1}{j} \int_0^\infty e^{-j\mu x} \left(1 - e^{-\mu x}\right)^{i+1-j} h(x)\, dx.$$

***16.** Рассмотрим систему $(M/G/1)$. Пусть $B_1, B_2, \ldots$ — последовательность независимых одинаково распределенных случайных величин с функцией распределения $B(x)$, которая равна функции распределения периода занятости системы. Предположим, что длительность обслуживания первого (за период занятости) требования равна X (с функцией распределения $H(x) = \mathrm{P}\{X \leqslant x\}$ и средним $\alpha < \infty$) и что за время его обслуживания поступают n других требований. Показать, что

$$B(x) = \mathrm{P}\{X + B_1 + B_2 + \ldots + B_n \leqslant x\}.$$

Отсюда получить, что преобразование Лапласа

$$\tilde{B}(\theta) = \int_0^\infty e^{-\theta x}\, dB(x)$$

удовлетворяет функциональному уравнению

$$\tilde{B}(\theta) = \psi\left(\theta + \lambda\left(1 - \tilde{B}(\theta)\right)\right),$$

где

$$\psi(\theta) = \int_0^\infty e^{-\theta x}\, dH(x).$$

Использовать этот результат для нахождения средней продолжительности периода занятости.

Указание: Период занятости не зависит от порядка обслуживания. Предположим (поскольку случай $n = 0$ тривиален), что за время первого акта обслуживания поступают $n > 0$ требований. Обслуживанием первого из этих требований начнем новый период занятости. По завершении его вернемся ко второму из указанных требований и начнем другой период занятости и т. д. n раз.

Ответ:

Средняя продолжительность периода занятости $= \dfrac{\alpha}{1 - \lambda\alpha}$,

$$\alpha = \int_{x=0}^\infty x\, dH(x).$$

***17.** В условиях задачи 3 при $\lambda < \mu$ рассмотрим систему в момент поступления требования. Найти вероятность того, что за время ожидания этого требования обслужатся в точности n требований при условии, что поступившее требование не застает систему свободной.

Указание: Применить метод решения задачи 16 для доказательства того, что вероятностная производящая функция $g(s)$ числа требований, обслуженных за период занятости, удовлетворяет функциональному уравнению

$$g(s) = \mu s \left(\mu + \lambda - \lambda g(s)\right)^{-1}.$$

Ответ:

$$g(s)=\sum_{r=1}^{\infty}\binom{1/2}{r}(-1)^{r+1}\left(\frac{4\lambda\mu}{(\lambda+\mu)^2}\right)^r\left(\frac{\lambda+\mu}{2\lambda}\right)s^r.$$

18. Рассмотрим систему обслуживания, в которой требования поступают регулярно в моменты $\frac{n}{\lambda}$, $n=0, 1, 2, \ldots$. Предположим, что время обслуживания X_j j-го требования распределено экспоненциально с параметром μ. Пусть $\lambda>\mu$. Найти вероятность того, что прибор никогда не освободится, если в момент 0 в системе имеется одно требование.

Указание: Показать, что искомая вероятность равна

$$\mathrm{P}\{Y_1+Y_2+\ \ldots\ +Y_i\geqslant i,\quad i=1, 2, \ldots\},$$

где Y_i — независимые д. с. в., имеющие экспоненциальное распределение с параметром $\frac{\lambda}{\mu}$ (использовать лемму 8.1).

Ответ: $1-\frac{\mu}{\lambda}$.

19. (Очередь с абсолютными приоритетами.) Рассмотрим одноканальную систему, в которую поступают пуассоновские потоки двух типов требований (приоритетных и неприоритетных) с параметрами λ_1 и λ_2 ($\lambda_1+\lambda_2=1$). Длительности обслуживания требований этих типов распределены экспоненциально с параметрами μ_1 и μ_2 соответственно. Внутри требований одного типа поддерживается прямой порядок обслуживания и процесс обслуживания приоритетных требований никогда не прерывается. Если приоритетное требование поступает во время обслуживания неприоритетного требования, то обслуживание последнего немедленно прерывается и начинает обслуживаться приоритетное требование. Требование, процесс обслуживания которого был прерван, вновь поступает на прибор, когда в системе не остается приоритетных требований. Пусть $p_{m,n}$ — стационарная вероятность того, что в системе имеется m приоритетных и n неприоритетных требований. Стационарный режим существует при $\rho_1+\rho_2<1$ $\left(\rho_1=\frac{\lambda_1}{\mu_1},\quad \rho_2=\frac{\lambda_2}{\mu_2}\right)$. Доказать, что величины $p_{m,n}$ удовлетворяют системе уравнений

$$\{\lambda_1+\lambda_2+\mu_1(1-\delta_{m0})+\mu_2(1-\delta_{n0})\,\delta_{m0}\}\,p_{m,n}=$$
$$=\lambda_1 p_{m-1,n}+\lambda_2 p_{m,n-1}+\mu_1 p_{m+1,n}+\mu_2\delta_{m0}p_{m,n+1}\qquad (m,\ n=0, 1, 2, \ldots),$$

где δ_{ij} — символ Кронекера и где принимается, что величина p с отрицательным индексом есть 0. Используя это уравнение, показать, что среднее число неприоритетных требований равно

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}np_{m,n}=\frac{\rho_2}{1-\rho_1-\rho_2}\left\{1+\frac{\mu_2\rho_1}{\mu_1(1-\rho_1)}\right\}.$$

20. Показать, что для системы $(M/M/1)$, находящейся в стационарном режиме, распределение времени между последовательными моментами ухода требований из системы совпадает с (экспоненциальным) распределением времени между моментами поступления (см. также задачу 33 гл. 7).

21. Требования поступают в систему в соответствии с произвольным рекуррентным потоком. Проанализировать структуру очереди в моменты регенерации для следующих двух систем:

(1) Имеется s приборов с одинаковым экспоненциальным распределением времени обслуживания на каждом из них.

(2) Имеется один прибор с эрланговским распределением времени обслуживания.

22. Рассмотрим следующее обобщение системы обслуживания $(GI/G/1)$ с распределениями $A(t)$ интервала между моментами поступления и $B(t)$ времени обслуживания, имеющими средние a и b соответственно. Требование, поступающее и застающее прибор свободным, ожидает случайное время с функцией распределения $V(t)$, а затем начинает обслуживаться.

Пусть $F_n(x)$ — функция распределения времени ожидания n-го требования. Показать, что предел $F(x) = \lim\limits_{n\to\infty} F_n(x)$ существует и

(1) если $b-a>0$, то $F(x)=0$,

(2) если $b-a<0$, то $F(x)$ — собственная функция распределения.

*23. Обобщим идею задачи 9 на случай двух односторонних движений по пересекающимся дорогам A и B. Движение по дороге A имеет абсолютный приоритет. На дороге B имеется останавливающий движение сигнал. Как и прежде, автомобили, движущиеся по дороге A, проезжают перекресток в моменты, образующие пуассоновский поток с параметром μ. Автомобили, движущиеся по дороге B, подъезжают к перекрестку в моменты, образующие пуассоновский поток с параметром λ, и выстраиваются в очередь, ожидая возможности пересечь его. Когда автомобиль (на дороге B) становится первым в очереди, он ожидает, пока между автомобилями, движущимися по дороге A, не образуется временнóе «окно» длительностью по крайней мере T, и тогда пересекает перекресток за время T. Другие автомобили из очереди не начинают движения, пока он не пересечет перекрестка. Длина автомобиля равна нулю. Найти производящую функцию распределения числа автомобилей в очереди на дороге B в стационарном режиме и среднюю длину очереди.

Указание: Это пример системы $(M/G/1)$, и достаточно найти распределение «времени обслуживания» автомобилей на дороге B.

Ответ:

$$\pi(s) = \frac{(1-\rho)(s-1)K(s)}{s-K(s)}, \qquad K(s) = \tilde{B}(\lambda - \lambda s).$$

$$\tilde{B}(\theta) = \int_0^\infty e^{-\theta x}\, dH(x) = \frac{(\mu+\theta)\, e^{-(\mu+\theta)T}}{\theta + \mu e^{-(\mu+\theta)T}}.$$

ЗАМЕЧАНИЯ

Литература по теории массового обслуживания обширна. Прекрасной монографией, в которой дается обзор этой теории с приложениями, является книга Кокса и Смита [1].

Мы также рекомендуем читателю более сложные книги Такача [2] и Риордана [3]. Многие результаты по теории массового обслуживания приведены в книге Саати [4]. В ней также имеется большая библиография.

Применения к транспортным задачам и задачам телефонии можно найти в книге Сиски [5].

В монографии Бенеша [6] развиты некоторые специальные математические вопросы теории массового обслуживания.

ЛИТЕРАТУРА

1. Кокс Д. Р., Смит В. Л., Теория очередей, «Мир», М., 1966.
2. Takács L., Introduction to the Theory of Queues, Oxford Univ. Press, London and New York, 1962.
3. Риордан Дж., Вероятностные системы обслуживания, «Сов. радио», М., 1966.
4. Саати Т. Л., Элементы теории массового обслуживания с применениями, «Сов. радио», М., 1965.
5. Syski E., Congestion Theory, Wiley, New York, 1960.
6. Beneš V. E., General Stochastic Processes in the Theory of Queues, Addison — Wesley, Reading, Massachusetts, 1963.

ПРИЛОЖЕНИЕ

§ 1. СПЕКТРАЛЬНАЯ ТЕОРЕМА

А. Вводные понятия, линейная независимость и базис[1])

Множество всех n-векторов $\mathbf{x} = (x_1, \ldots, x_n)$, где x_i — комплексные числа, образует n-мерное векторное пространство. Сумма двух векторов $\mathbf{x} = (x_1, \ldots, x_n)$ и $\mathbf{y} = (y_1, \ldots, y_n)$ определяется как $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$, а произведение вектора $\mathbf{x}$ на комплексное число λ по формуле $\lambda\mathbf{x} = (\lambda x_1, \ldots, \lambda x_n)$.

Векторы $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(r)}$ *называются линейно независимыми*, если из равенства

$$c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \ldots + c_r\mathbf{x}^{(r)} = \mathbf{0}$$

следует, что $c_1 = c_2 = \ldots = c_r = 0$; в противном случае эти векторы называются *линейно зависимыми*. Например, векторы $(1, 0, \ldots, 0)$, $(0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$ являются, очевидно, линейно независимыми. В n-мерном векторном пространстве не может быть более, чем n линейно независимых векторов, или другими словами, любое множество, состоящее более чем из n векторов, линейно зависимо.

Пусть $\varphi_1, \ldots, \varphi_r$, $r < n$, — линейно зависимые векторы. Тогда существует вектор φ_{r+1}, не являющийся линейной комбинацией векторов $\varphi_1, \ldots, \varphi_r$, или, что то же самое, не представимый в виде $c_1\varphi_1 + \ldots + c_r\varphi_r$. Это означает, как легко видеть, что $\varphi_1, \ldots, \varphi_{r+1}$ линейно независимы. Рассуждая далее точно так же, мы получим множество $\varphi_1, \ldots, \varphi_n$ из n линейно независимых векторов, построенное пополнением множества векторов $\varphi_1, \ldots, \varphi_r$ векторами $\varphi_{r+1}, \ldots, \varphi_n$. Поскольку никакое линейно независимое множество не может состоять более чем из n векторов, для каждого вектора $\mathbf{y}$ и любого линейно независимого множества векторов $\varphi_1, \ldots, \varphi_n$ мы можем определить (и притом единственным образом) константы $c_1, \ldots, c_n$, такие, что $c_1\varphi_1 + \ldots + c_n\varphi_n = \mathbf{y}$.

Аналогичные результаты имеют место для любого линейного подпространства $\mathfrak{M}$, т. е. для любого множества векторов $\mathfrak{M}$, такого, что если $\mathbf{x}, \mathbf{y} \in \mathfrak{M}$, то $a\mathbf{x} + b\mathbf{y} \in \mathfrak{M}$ для любых комплексных чисел a и b. Каждое линейное подпространство характеризуется целым числом m, $0 \leqslant m \leqslant n$, называемым *размерностью подпро-*

---

[1]) Некоторые утверждения приводятся нами без доказательства; читателю будет полезно провести их самостоятельно.

странства, которое равно максимальному числу векторов, все еще образующих линейно независимое множество. Если $\varphi_1, \ldots, \varphi_r$, $r < m$, — линейно независимые векторы из $\mathfrak{M}$, то существует вектор $\varphi_{r+1} \in \mathfrak{M}$, который нельзя представить в виде линейной комбинации векторов $\varphi_1, \ldots, \varphi_r$. Как и ранее, легко показать, что существуют векторы $\varphi_{r+1}, \ldots, \varphi_m$, такие, что $\varphi_1, \ldots, \varphi_m$ образуют линейно независимое множество векторов. Более того, для любого вектора $\mathbf{y} \in \mathfrak{M}$ существуют (и единственны) константы $c_1, \ldots, c_m$, такие, что $c_1\varphi_1 + \ldots + c_m\varphi_m = \mathbf{y}$. Заметим, что если размерность подпространства $\mathfrak{M}$ равна нулю ($\dim \mathfrak{M} = 0$), то это означает, что $\mathfrak{M}$ состоит лишь из нулевого элемента; если же $\dim \mathfrak{M} = n$, то $\mathfrak{M}$ совпадает с исходным векторным пространством. Если $\dim \mathfrak{M} = m$, то любое линейно независимое множество из m векторов, принадлежащих $\mathfrak{M}$, называется *базисом подпространства* $\mathfrak{M}$. Мы будем пользоваться термином «базис» (без указания подпространства) для обозначения любого множества из n линейно независимых векторов.

Б. Скалярное произведение

Скалярное произведение двух векторов $\mathbf{x}$ и $\mathbf{y}$ определяется формулой

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i \bar{y}_i,$$

где $\bar{y}_i$ — числа, комплексно сопряженные с y_i. Отметим следующие легко доказываемые свойства скалярного произведения:

(i) $(\mathbf{x}, \mathbf{x}) \geqslant 0$, причем равенство имеет место тогда и только тогда, когда $\mathbf{x} = (0, \ldots, 0) = \mathbf{0}$;

(ii) $(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$, где λ — комплексное число;

(iii) $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$.

Из последнего соотношения следует, что $(\mathbf{x}, \lambda\mathbf{y}) = \bar{\lambda}(\mathbf{x}, \mathbf{y})$.

Говорят, что векторы $\mathbf{x}$ и $\mathbf{y}$ ортогональны, если $(\mathbf{x}, \mathbf{y}) = 0$. Норма $\|\mathbf{x}\|$ вектора $\mathbf{x}$ определяется как $\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2}$.

Набор $\{a_{ij}\}$, $i, j = 1, \ldots, n$, комплексных чисел образует (квадратную) матрицу порядка n, обычно обозначаемую $\mathbf{A} = \|a_{ij}\|$, $i, j = 1, \ldots, n$. Квадратная матрица порядка n определяет преобразование (или оператор) в n-мерном пространстве, ставящее в соответствие вектору $\mathbf{x}$ из этого пространства вектор $\mathbf{y} = \mathbf{A}\mathbf{x}$, где $y_i = \sum_{j=1}^{n} a_{ij} x_j$, $i = 1, \ldots, n$, либо вектор $\mathbf{z} = \mathbf{x}\mathbf{A}$, где $z_j = \sum_{i=1}^{n} x_i a_{ij}$, $j = 1, \ldots, n$.

Из этого определения непосредственно следует, что

$$\mathbf{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{A}\mathbf{x} + \beta\mathbf{A}\mathbf{y}, \qquad (\alpha\mathbf{x} + \beta\mathbf{y})\mathbf{A} = \alpha\mathbf{x}\mathbf{A} + \beta\mathbf{y}\mathbf{A}$$

для любых векторов $\mathbf{x}$ и $\mathbf{y}$ и любых констант α и β. Кроме того, $(\mathbf{x}, \mathbf{A}\mathbf{y}) = (\mathbf{x}\bar{\mathbf{A}}, \mathbf{y})$, где $\bar{\mathbf{A}}$ — квадратная матрица порядка n с элементами $\bar{a}_{ij}$.

B. Собственные значения и собственные векторы

Комплексное число λ называется собственным значением (или характеристическим числом) матрицы $\mathbf{A}$, если существует такой вектор $\mathbf{x}^{(\lambda)} \neq 0$, что $\mathbf{A}\mathbf{x}^{(\lambda)} = \lambda\mathbf{x}^{(\lambda)}$. Пусть λ — собственное значение матрицы $\mathbf{A}$; множество $\mathfrak{M}_\lambda$, состоящее из всех векторов $\mathbf{x}$, таких, что $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, называется правым собственным подпространством матрицы $\mathbf{A}$, соответствующим собственному значению λ. Элементы подпространства $\mathfrak{M}_\lambda$ называют при этом правыми собственными (или характеристическими) векторами, соответствующими (отвечающими или принадлежащими) собственному значению λ. Очевидно, что из $\mathbf{y}$, $\mathbf{z} \in \mathfrak{M}_\lambda$ следует, что $a\mathbf{y} + b\mathbf{z} \in \mathfrak{M}_\lambda$ для любых констант a и b. Размерность $\mathfrak{M}_\lambda$ называют кратностью собственного значения λ.

Если $\mathfrak{M}_1, \ldots, \mathfrak{M}_r$ — собственные подпространства оператора $\mathbf{A}$, соответствующие различным собственным значениям, то ненулевые векторы $\varphi_1, \ldots, \varphi_r$, принадлежащие подпространствам $\mathfrak{M}_1, \ldots, \mathfrak{M}_r$ соответственно, линейно независимы. В самом деле, предположив противное, обозначим через m наименьшее целое число, для которого существуют векторы $\varphi_{i_1} \in \mathfrak{M}_{i_1}, \ldots, \varphi_{i_m} \in \mathfrak{M}_{i_m}$ и константы $c_1, \ldots, c_m$, такие, что $c_1\varphi_{i_1} + \ldots + c_m\varphi_{i_m} = \mathbf{0}$, где не все c_i равны 0 и все i_j, $1 \leqslant j \leqslant m$, различны. Очевидно, $m \geqslant 2$. Применяя $\mathbf{A}$ к обеим частям предыдущего равенства, получаем $\lambda_1 c_1\varphi_{i_1} + \ldots + \lambda_m c_m\varphi_{i_m} = \mathbf{0}$. Если хотя бы одно из собственных значений λ_i равно нулю, то $m-1$ векторов φ_{i_j} линейно зависимы, что противоречит определению числа m. Если же все λ_i отличны от нуля, то, умножая равенство $c_1\varphi_{i_1} + \ldots + c_m\varphi_{i_m} = \mathbf{0}$, например, на λ_1 и вычитая из результата равенство $\lambda_1 c_1\varphi_{i_1} + \ldots + \lambda_m c_m\varphi_{i_m} = \mathbf{0}$, получаем

$$(\lambda_2 - \lambda_1) c_2\varphi_{i_2} + \ldots + (\lambda_m - \lambda_1) c_m\varphi_{i_m} = \mathbf{0}.$$

Последнее также противоречит определению m.

Пусть $\varphi_1^{(i)}, \ldots, \varphi_{m_i}^{(i)}$ — базис собственного подпространства $\mathfrak{M}_i$, $i = 1, \ldots, r$; легко видеть, что векторы

$$\varphi_1^{(1)}, \ldots, \varphi_{m_1}^{(1)}, \varphi_1^{(2)}, \ldots, \varphi_{m_2}^{(2)}, \ldots, \varphi_1^{(r)}, \ldots, \varphi_{m_r}^{(r)}$$

линейно независимы. Из вышесказанного следует, что $\mathbf{A}$ может иметь лишь конечное число собственных значений и собственных подпространств. В том важном случае, когда сумма размерностей собственных подпространств равна n, в n-мерном векторном пространстве существует базис из собственных векторов матрицы $\mathbf{A}$. Матрица $\mathbf{A}$, обладающая этим свойством, называется *приводимой к диагональной форме* (или *диагонализируемой*).

Точно так же, как в предыдущих рассуждениях мы отправлялись от уравнения $\mathbf{Ax} = \lambda\mathbf{x}$, за основу можно принять уравнение $\mathbf{xA} = \lambda\mathbf{x}$. Оказывается, что значения λ, при которых уравнение $\mathbf{xA} = \lambda\mathbf{x}$ имеет ненулевое решение, суть собственные значения матрицы $\mathbf{A}$, определенные ранее. Размерность подпространства, образуемого векторами, удовлетворяющими уравнению $\mathbf{xA} = \lambda\mathbf{x}$ (левыми собственными векторами), равна кратности собственного значения λ. (Читателю следует убедиться в этом самостоятельно.) Как и в предыдущем случае, левые собственные векторы, отвечающие различным собственным значениям, линейно независимы.

Отметим, что собственные значения матрицы $\mathbf{A}$ являются корнями алгебраического уравнения n-й степени $\det\|\mathbf{A} - \lambda\mathbf{I}\| = 0$, где $\mathbf{I}$ — единичная матрица. Отсюда следует, что если

$$\mathbf{A} = \begin{Vmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{B} & \mathbf{A}_2 \end{Vmatrix},$$

где $\mathbf{A}_1$, $\mathbf{A}_2$ — квадратные матрицы, то λ является собственным значением матрицы $\mathbf{A}$ тогда и только тогда, когда λ — собственное значение по крайней мере одной из матриц $\mathbf{A}_1$ или $\mathbf{A}_2$. Действительно, так как

$$\det\|\mathbf{A} - \lambda\mathbf{I}\| = \det\|\mathbf{A}_1 - \lambda\mathbf{I}\| \det\|\mathbf{A}_2 - \lambda\mathbf{I}\|,$$

где через $\mathbf{I}$ обозначены единичные матрицы соответствующих порядков, то утверждение очевидно.

(а) *Спектральное представление*

В этом пункте мы будем предполагать, что матрица $\mathbf{A}$ действительная, т. е. что ее элементы — действительные числа. Пусть матрица $\mathbf{A}$ такова, что ее правые собственные векторы образуют базис в n-мерном векторном пространстве. В этом случае, как легко убедиться, левые собственные векторы также образуют базис. Если элементы матрицы $\mathbf{A}$ действительны, указанные базисы можно сделать биортогональными. Покажем это. Пусть $\varphi_1, \ldots, \varphi_n$ и $\psi_1, \ldots, \psi_n$ — базисы из правых и левых собственных векторов соответственно. Биортогональность базисов означает, что

$$(\varphi_i, \psi_j) = \delta_{ij} = \begin{cases} 1, & \text{если} \quad i = j, \\ 0, & \text{если} \quad i \neq j. \end{cases}$$

Прежде всего отметим, что если $\mathbf{Ax}_i = \lambda_i\mathbf{x}_i$ и $\mathbf{y}_j\mathbf{A} = \mu_j\mathbf{y}_j$, то

$$\mu_j(\mathbf{y}_j, \mathbf{x}_i) = (\mu_j\mathbf{y}_j, \mathbf{x}_i) = (\mathbf{y}_j\mathbf{A}, \mathbf{x}_i) = (\mathbf{y}_j, \mathbf{Ax}_i) = (\mathbf{y}_j, \lambda_i\mathbf{x}_i) = \overline{\lambda}_i(\mathbf{y}_j, \mathbf{x}_i).$$

Таким образом, если $\mu_j = \bar{\lambda}_i$, то $(\mathbf{y}_j, \bar{\mathbf{x}}_i) = 0$. Далее, из $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ следует, что $\mathbf{A}\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$, где $\bar{\mathbf{x}} = (\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_n)$. Это означает, что $\bar{\lambda}$ является собственным значением матрицы $\mathbf{A}$ той же кратности, что и λ. Пусть

$$\lambda_1,\ \bar{\lambda}_1,\ \lambda_2,\ \bar{\lambda}_2,\ \dots,\ \lambda_r,\ \bar{\lambda}_r,\ \lambda_{r+1},\ \lambda_{r+2},\ \dots,\ \lambda_m$$

— собственные значения матрицы $\mathbf{A}$, причем $\lambda_1, \dots, \lambda_r$ — комплексные, а $\lambda_{r+1}, \dots, \lambda_m$ — действительные. Обозначим соответствующие правые собственные подпространства через $\mathfrak{M}_1$, $\bar{\mathfrak{M}}_1, \dots, \mathfrak{M}_r$, $\bar{\mathfrak{M}}_r$, $\mathfrak{M}_{r+1}, \dots, \mathfrak{M}_m$, а левые собственные подпространства через $\mathfrak{L}_1$, $\bar{\mathfrak{L}}_1, \dots, \mathfrak{L}_r$, $\bar{\mathfrak{L}}_r$, $\mathfrak{L}_{r+1}, \dots, \mathfrak{L}_m$.

Как мы уже показали, каждый вектор из $\mathfrak{L}_1$ ортогонален любому правому собственному вектору, за исключением принадлежащих подпространству $\bar{\mathfrak{M}}_1$. Наша задача теперь состоит в выборе базиса $\psi_1, \dots, \psi_d$ подпространства $\mathfrak{L}_1$ и базиса $\varphi_1, \dots, \varphi_d$ (d — кратность собственного значения λ_1) подпространства $\bar{\mathfrak{M}}_1$, векторы которых обладают тем свойством, что $(\psi_i, \varphi_j) = \delta_{ij}$. Чтобы решить эту задачу, предположим, что $\varphi_1, \dots, \varphi_d$ — произвольный базис в $\bar{\mathfrak{M}}_1$, а $\mathbf{y}_1, \dots, \mathbf{y}_d$ — произвольный базис в $\mathfrak{L}_1$. Нам нужно найти такие константы $c_1, \dots, c_d$, что $\psi_1 = c_1\mathbf{y}_1 + \dots + c_d\mathbf{y}_d$ и $(\psi_1, \varphi_i) = \delta_{1i}$, $i = 1, 2, \dots, d$, т. е.

$$\begin{aligned}
c_1(\mathbf{y}_1, \varphi_1) + c_2(\mathbf{y}_2, \varphi_1) + \dots + c_d(\mathbf{y}_d, \varphi_1) &= 1,\\
c_1(\mathbf{y}_1, \varphi_2) + c_2(\mathbf{y}_2, \varphi_2) + \dots + c_d(\mathbf{y}_d, \varphi_2) &= 0,\\
&\ \ \vdots\\
c_1(\mathbf{y}_1, \varphi_d) + c_2(\mathbf{y}_2, \varphi_d) + \dots + c_d(\mathbf{y}_d, \varphi_d) &= 0.
\end{aligned}$$

Если эта система линейных уравнений относительно $c_1, \dots, c_d$ не имеет решения, то вектор $(1, 0, \dots, 0)$ не представим в виде линейной комбинации векторов

$$\mathbf{f}_1 = ((\mathbf{y}_1, \varphi_1), \dots, (\mathbf{y}_1, \varphi_d)), \dots, \mathbf{f}_d = ((\mathbf{y}_d, \varphi_1), \dots, (\mathbf{y}_d, \varphi_d)).$$

Отсюда следует, что эти векторы линейно зависимы, т. е. существуют константы $a_1, \dots, a_d$, не все равные нулю, такие, что $a_1\mathbf{f}_1 + \dots + a_d\mathbf{f}_d = \mathbf{0}$. Но это значит, что

$$(a_1\mathbf{y}_1 + \dots + a_d\mathbf{y}_d, \varphi_i) = 0, \qquad i = 1, \dots, d.$$

Ранее нами было доказано, что векторы $\mathbf{y}_1, \dots, \mathbf{y}_d$ (а следовательно, и любая их линейная комбинация) ортогональны любому правому собственному подпространству, за исключением $\bar{\mathfrak{M}}_1$. Теперь мы видим, что $a_1\mathbf{y}_1 + \dots + a_d\mathbf{y}_d$ ортогонален всем правым собственным векторам и, разумеется, любой их линейной комбинации. По предположению правые собственные векторы образуют базис. Поэтому вектор $a_1\mathbf{y}_1 + \ldots + a_d\mathbf{y}_d$ ортогонален самому себе и, следовательно, равен нулю. Это противоречит линейной независимости векторов $\mathbf{y}_1, \ldots, \mathbf{y}_d$. Таким образом, искомый вектор ψ_1 существует. Остальные векторы, $\psi_2, \ldots, \psi_d$, строятся точно так же. Остается показать, что векторы $\psi_1, \ldots, \psi_d$ линейно независимы. Предположим, что $a_1\psi_1 + \ldots + a_d\psi_d = \mathbf{0}$; тогда

$$\begin{aligned} 0 &= (a_1\psi_1 + \ldots + a_d\psi_d, \varphi_1) = a_1, \\ 0 &= (a_1\psi_1 + \ldots + a_d\psi_d, \varphi_2) = a_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ 0 &= (a_1\psi_1 + \ldots + a_d\psi_d, \varphi_d) = a_d, \end{aligned}$$

а это и означает, что векторы ψ_i линейно независимы.

Итак, мы показали, что действительная диагонализируемая матрица имеет базис из правых собственных векторов $\varphi_1, \ldots, \varphi_n$ и базис из левых собственных векторов $\psi_1, \ldots, \psi_n$, векторы которых биортогональны, т. е. $(\psi_i, \varphi_j) = \delta_{ij}$. Мы воспользуемся этим результатом при построении канонического представления матрицы $\mathbf{A}$, называемого спектральным. Пусть $\lambda_1, \ldots, \lambda_n$ — собственные значения матрицы $\mathbf{A}$, а $\varphi_1, \ldots, \varphi_n$ и $\psi_1, \ldots, \psi_n$ — соответствующие им правые и левые собственные векторы. Пусть

$$\varphi_i = (\varphi_{i1}, \ldots, \varphi_{in}), \qquad \psi_i = (\psi_{i1}, \ldots, \psi_{in}),$$

$$\boldsymbol{\Phi} = \left\| \begin{matrix} \varphi_{11} & \cdots & \varphi_{n1} \\ \vdots & & \vdots \\ \varphi_{1n} & \cdots & \varphi_{nn} \end{matrix} \right\|, \qquad \boldsymbol{\Psi} = \left\| \begin{matrix} \psi_{11} & \cdots & \psi_{1n} \\ \vdots & & \vdots \\ \psi_{n1} & \cdots & \psi_{nn} \end{matrix} \right\|,$$

$$\boldsymbol{\Lambda} = \left\| \begin{matrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \lambda_n \end{matrix} \right\|.$$

Из биортогональности базисов $\varphi_1, \ldots, \varphi_n$ и $\psi_1, \ldots, \psi_n$ сразу же следует, что $\boldsymbol{\Psi\Phi} = \mathbf{I}$, где $\mathbf{I}$ — единичная матрица. Кроме того, непосредственным вычислением легко установить, что $\boldsymbol{\Phi\Lambda\Psi}\varphi_i = \lambda_i\varphi_i$, $i = 1, \ldots, n$. Поскольку $\mathbf{A}\varphi_i = \lambda_i\varphi_i$, $i = 1, \ldots, n$, и векторы φ_i образуют базис во всем пространстве, то

$$\mathbf{A} = \boldsymbol{\Phi\Lambda\Psi} \quad \text{и} \quad \boldsymbol{\Phi\Psi} = \mathbf{I},$$

Легко видеть, что $\mathbf{A}^m = \mathbf{\Phi\Lambda\Psi\Phi\Lambda\Psi} \ldots \mathbf{\Phi\Lambda\Psi} = \mathbf{\Phi\Lambda}^m\mathbf{\Psi}$. Так как

$$\Lambda^m = \begin{Vmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n^m \end{Vmatrix},$$

то $\mathbf{A}^m$ относительно легко вычисляется, если известно спектральное представление матрицы $\mathbf{A}$.

(б) *Сходимость*

Нам понадобится понятие сходимости для последовательностей векторов и матриц.

Говорят, что последовательность $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$ векторов n-мерного векторного пространства сходится к вектору $\mathbf{x}^{(0)}$, если

$$\lim_{j \to \infty} \mathbf{x}_i^{(j)} = \mathbf{x}_i^{(0)}, \qquad i = 1, \ldots, n.$$

Точно так же последовательность квадратных матриц $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots$ порядка n сходится к матрице $\mathbf{A}^{(0)}$, если

$$\lim_{h \to \infty} a_{ij}^{(h)} = a_{ij}^{(0)}, \qquad i,\ j = 1, \ldots, n.$$

Из этих определений сразу же следует, что если $\lim\limits_{h\to\infty} \mathbf{A}^{(h)} = \mathbf{A}^{(0)}$ и $\lim\limits_{j\to\infty} \mathbf{x}^{(j)} = \mathbf{x}^{(0)}$, то $\lim\limits_{h\to\infty} \mathbf{A}^{(h)}\mathbf{x}^{(h)} = \mathbf{A}^{(0)}\mathbf{x}^{(0)}$. Более того, если $\{\mathbf{A}^{(j)}\}$ — последовательность матриц, для которой существует матрица $\mathbf{A}^{(0)}$ и базис $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}$, такие, что

$$\lim_{h \to \infty} \mathbf{A}^{(h)}\mathbf{x}^{(i)} = \mathbf{A}^{(0)}\mathbf{x}^{(i)}, \qquad i = 1, \ldots, n,$$

то $\lim\limits_{h\to\infty} \mathbf{A}^{(h)} = \mathbf{A}^{(0)}$. Действительно, в этом случае $\lim\limits_{h\to\infty} \mathbf{A}^{(h)}\mathbf{y} = \mathbf{A}^{(0)}\mathbf{y}$ для любого $\mathbf{y}$, так как $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}$ образуют базис, а, следовательно, $\mathbf{y} = c_1\mathbf{x}^{(1)} + \ldots + c_n\mathbf{x}^{(n)}$.

§ 2. ТЕОРИЯ ФРОБЕНИУСА ПОЛОЖИТЕЛЬНЫХ МАТРИЦ

Теория положительных матриц, развитая Фробениусом, играет важную роль во многих разделах теории вероятностей, в частности в анализе матриц переходных вероятностей цепей Маркова.

Предварительные замечания

Если каждый элемент a_{ij} некоторой квадратной матрицы $\mathbf{A} = \|a_{ij}\|$, $i,\ j = 1, \ldots, n$, неотрицателен, то мы будем писать

$\mathbf{A} \geqslant \mathbf{0}$; если $\mathbf{A} \geqslant \mathbf{0}$ и по крайней мере один элемент $a_{ij} > 0$, то мы будем писать $\mathbf{A} > \mathbf{0}$ и называть $\mathbf{A}$ положительной матрицей. Если все $a_{ij} > 0$, то этот факт мы будем обозначать $\mathbf{A} \gg \mathbf{0}$. Такими же обозначениями мы будем пользоваться и для векторов. Именно если $\mathbf{x} = (x_1, \ldots, x_n)$, то $\mathbf{x} \geqslant \mathbf{0}$ означает, что $x_i \geqslant 0$ для всех $i = 1, \ldots, n$; $\mathbf{x} > \mathbf{0}$ означает, что $\mathbf{x} \geqslant \mathbf{0}$ и $x_i > 0$ хотя бы для одного i, и, наконец, $\mathbf{x} \gg \mathbf{0}$ означает, что $x_i > 0$ для всех $i = 1, \ldots, n$. Мы также будем писать $\mathbf{x} \geqslant \mathbf{y}$, если $\mathbf{x} - \mathbf{y} \geqslant \mathbf{0}$, и т. п. Ясно, что из $\mathbf{A} \geqslant \mathbf{0}$ и $\mathbf{x} \geqslant \mathbf{y}$ следует, что $\mathbf{Ax} \geqslant \mathbf{Ay}$, а из $\mathbf{A} \gg \mathbf{0}$ и $\mathbf{x} > \mathbf{y}$ следует, что $\mathbf{Ax} \gg \mathbf{Ay}$.

Пусть $\mathbf{A} \geqslant \mathbf{0}$; рассмотрим множество Λ всех действительных чисел λ, каждому из которых соответствует вектор $\mathbf{x} = (x_1, \ldots, x_n)$, такой, что

$$\sum_{i=1}^{n} x_i = 1, \qquad \mathbf{x} > \mathbf{0}, \qquad \mathbf{Ax} \geqslant \lambda \mathbf{x}.$$

Пусть $\lambda_0 = \sup_\Lambda \lambda$; легко показать, что λ_0 конечно, а в случае когда $\mathbf{A} \gg \mathbf{0}$, еще и положительно. В самом деле, если $M = \max\limits_{1 \leqslant i,\, j \leqslant n} a_{ij}$, то из условий $\sum\limits_{i=1}^{n} x_i = 1$ и $\mathbf{x} > \mathbf{0}$ следует, что $\sum\limits_{j=1}^{n} a_{ij} x_j \leqslant M \sum\limits_{j=1}^{n} x_j = M$, $i = 1, \ldots, n$, а также что $x_j \geqslant 1/n$ хотя бы для одного j. Отсюда получаем неравенство $\lambda_0 \leqslant nM$. Если $\mathbf{A} \gg \mathbf{0}$ и $0 < \delta = \min\limits_{1 \leqslant i,\, j \leqslant n} a_{ij}$, то $\sum\limits_{j} a_{ij} \geqslant \delta n$; это означает, что число $\delta n \in \Lambda$. (Ему соответствует вектор $\mathbf{x} = (1/n, \ldots, 1/n)$.)

Предположим теперь, что $\lambda_0 = 0$ для матрицы $\mathbf{A} > \mathbf{0}$. Для $\mathbf{x} \gg \mathbf{0}$ соотношение $\mathbf{Ax} \gg \mathbf{0}$ невозможно, так как $\lambda_0 = 0$. С другой стороны, из равенства $\mathbf{Ax} = \mathbf{0}$ для некоторого $\mathbf{x} \gg \mathbf{0}$ следует, что $\mathbf{A} = \mathbf{0}$. Таким образом, существует вектор $\mathbf{x} \gg \mathbf{0}$, такой, что $\mathbf{Ax} > \mathbf{0}$. Пусть C_1 — множество индексов положительных элементов вектора $\mathbf{Ax}$; очевидно, C_1 не зависит от выбора вектора $\mathbf{x} \gg \mathbf{0}$. Пусть вектор $\mathbf{y} = (y_1, \ldots, y_n) > \mathbf{0}$ таков, что $y_i > 0$, если $i \in C_1$, и $y_i = 0$, если $i \notin C_1$. Определим C_2 как множество индексов положительных координат вектора $\mathbf{Ay}$; это множество не зависит от выбора вектора $\mathbf{y}$ и, кроме того, $C_2 \subseteq C_1$. Поскольку $\lambda_0 = 0$, то $C_2 \neq C_1$. Рассуждая точно так же, легко убедиться, что

$$C_1 \supset C_2 \supset \ldots \supset C_m = C_{m+1} = \ldots = \varnothing,$$

причем здесь все включения являются строгими. Теперь уже не составляет труда показать, что $\mathbf{A}^m = \mathbf{0}$. Действительно, $\mathbf{A}^m \mathbf{x} = \mathbf{0}$ для $\mathbf{x} \gg \mathbf{0}$, но так как любой вектор может быть представлен как разность двух строго положительных векторов, то $\mathbf{A}^m \mathbf{z} = \mathbf{0}$ для любого $\mathbf{z}$, что эквивалентно равенству нулю матрицы $\mathbf{A}^m$.

Первая теорема Фробениуса

Теорема 2.1. *Если* $\mathbf{A} \gg \mathbf{0}$, *то* (а) *существует вектор* $\mathbf{x}^0 \gg \mathbf{0}$, *такой, что* $\mathbf{A}\mathbf{x}^0 = \lambda_0\mathbf{x}^0$; (б) $|\lambda| < \lambda_0$ *для всех собственных значений* λ *матрицы* $\mathbf{A}$, *отличных от* λ_0; (в) *правые собственные векторы матрицы* $\mathbf{A}$, *соответствующие* λ_0, *образуют одномерное подпространство, т. е.* $\dim \mathfrak{M}_{\lambda_0} = 1$.

Доказательство. (а) По определению λ_0 существует последовательность $\gamma_1, \gamma_2, \ldots \to \lambda_0$ и векторы $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$, такие, что

$$\mathbf{x}^{(i)} > \mathbf{0}, \quad \mathbf{A}\mathbf{x}^{(i)} \geqslant \gamma_i \mathbf{x}^{(i)}, \quad \mathbf{x}_1^{(i)} + \ldots + \mathbf{x}_n^{(i)} = 1. \qquad (\Pi.2.1)$$

Так как значения компонент всех векторов $\mathbf{x}^{(i)}$ принадлежат отрезку $[0, 1]$, то с помощью процесса диагонализации мы можем построить последовательность положительных целых чисел $n_1 < n_2 < n_3 < \ldots$, такую, что

$$\lim x_r^{(n_j)} \to x_r^0, \qquad r = 1, \ldots, n, \qquad (\Pi.2.2)$$

где $x_r^0 \in [0, 1]$, $r = 1, 2, \ldots, n$. Из (2.1) следует, что $x_1^0 + \ldots + x_n^0 = 1$ и $\mathbf{x}^0 = (x_1^0, \ldots, x_n^0) > \mathbf{0}$. Кроме того, если мы заменим в неравенстве $\mathbf{A}\mathbf{x}^{(i)} \geqslant \gamma_i\mathbf{x}^{(i)}$ индекс i на n_j и устремим j к ∞, то придем к неравенству $\mathbf{A}\mathbf{x}^0 \geqslant \lambda_0\mathbf{x}^0$. Мы покажем, что на самом деле $\mathbf{A}\mathbf{x}^0 = \lambda_0\mathbf{x}^0$. Действительно, если это не так, то $\mathbf{A}\mathbf{x}^0 > \lambda_0\mathbf{x}^0$. Умножая обе части этого неравенства слева на $\mathbf{A}$, получаем неравенство $\mathbf{A}\mathbf{y}^0 \gg \lambda_0\mathbf{y}^0$, где $\mathbf{y}^0 = \mathbf{A}\mathbf{x}^0 \gg \mathbf{0}$. Это неравенство сохранится и тогда, когда мы прибавим к λ_0 достаточно малое положительное число ε: $\mathbf{A}\mathbf{y}^0 \gg (\lambda_0 + \varepsilon)\mathbf{y}^0$. Нормируя вектор $\mathbf{y}^0$ так, чтобы сумма его координат равнялась единице, мы убеждаемся в том, что $\lambda_0 + \varepsilon$ принадлежит множеству Λ. Но последнее противоречит определению λ_0. Таким образом, $\mathbf{A}\mathbf{x}^0 = \lambda_0\mathbf{x}^0$. Поскольку $\mathbf{x}^0 > \mathbf{0}$, а $\mathbf{A} \gg \mathbf{0}$, то $\lambda_0\mathbf{x}^0 \gg \mathbf{0}$, т. е. $\mathbf{x}^0 \gg \mathbf{0}$. Этим завершается доказательство утверждения (а).

(б) Пусть $\lambda \neq \lambda_0$ и $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$, где $\mathbf{z} \neq \mathbf{0}$. Координатная запись равенства $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$ имеет вид

$$\sum_{j=1}^{n} a_{ij} z_j = \lambda z_i, \qquad i = 1, \ldots, n.$$

Отсюда для абсолютных величин имеем

$$\sum_{j=1}^{n} a_{ij} |z_j| \geqslant |\lambda| |z_i|, \qquad i = 1, \ldots, n,$$

т. е.

$$\mathbf{A}|\mathbf{z}| \geqslant |\lambda| |\mathbf{z}|, \quad \text{где} \quad |\mathbf{z}| = (|z_1|, \ldots, |z_n|).$$

Нормируя вектор $|\mathbf{z}|$ так, чтобы сумма его координат равнялась единице (напомним, что $\mathbf{z} \neq \mathbf{0}$), мы видим, что $|\lambda|$ принадлежит Λ. Тогда по определению λ_0 имеем $|\lambda| \leqslant \lambda_0$. Чтобы показать, что $|\lambda| < \lambda_0$, введем матрицу $\mathbf{A}_\delta = \mathbf{A} - \delta\mathbf{I}$, где $\mathbf{I}$ — единичная матрица, а δ — положительное число, малое настолько, что $\mathbf{A}_\delta \gg \mathbf{0}$. Поскольку λ_0 — наибольшее по абсолютной величине положительное собственное значение матрицы $\mathbf{A}$, то $\lambda - \delta$ является таковым для матрицы $\mathbf{A}_\delta$. Повторяя те же рассуждения, что и при доказательстве неравенства $|\lambda| \leqslant \lambda_0$, для матрицы $\mathbf{A}_\delta$ и ее собственного значения $\lambda - \delta$, получаем неравенство $|\lambda - \delta| \leqslant \lambda_0 - \delta$. Но, с другой стороны,

$$|\lambda| = |\lambda - \delta + \delta| \leqslant |\lambda - \delta| + \delta \leqslant \lambda_0,$$

так что из $|\lambda| = \lambda_0$ следует $|\lambda| = |\lambda - \delta| + \delta$. Последнее означает, что λ действительно и положительно. Таким образом, имеет место равенство $\lambda = |\lambda| = \lambda_0$, что противоречит исходному предположению ($\lambda \neq \lambda_0$).

(в) Предположим, что $\mathbf{A}\mathbf{y} = \lambda_0\mathbf{y}$ и не существует такой константы c, что $\mathbf{y} = c\mathbf{x}^0$. Поскольку $\mathbf{A}$ — действительная матрица, векторы $\mathbf{u}$ и $\mathbf{v}$, координатами которых служат соответственно действительные и мнимые части координат вектора $\mathbf{y}$, являются собственными векторами матрицы $\mathbf{A}$, соответствующими λ_0. Поскольку $\mathbf{y} \neq c\mathbf{x}^0$ для всех констант c, то по крайней мере один из векторов $\mathbf{u}$ или $\mathbf{v}$ не представим в виде $c\mathbf{x}^0$. Поэтому мы можем считать вектор $\mathbf{y}$ действительным. Так как $\mathbf{x}^0 \gg \mathbf{0}$, можно выбрать такое число μ, что $\mathbf{x}^0 - \mu\mathbf{y} > \mathbf{0}$, но не $\gg \mathbf{0}$; при этом $|\mu| = \min\limits_{y_i \neq 0} \{x_i^0/|y_i|\}$. Отсюда $\mathbf{A}(\mathbf{x}^0 - \mu\mathbf{y}) = \lambda_0(\mathbf{x}^0 - \mu\mathbf{y})$, и, как и в доказательстве утверждения (а), с необходимостью получаем: $(\mathbf{x}^0 - \mu\mathbf{y}) \gg \mathbf{0}$. Это противоречит выбору μ. ■

Отметим попутно несколько простых фактов. Если $\mathbf{A} \gg \mathbf{0}$, то можно утверждать, что существует вектор $\mathbf{f}^0 \gg \mathbf{0}$, такой, что $\mathbf{f}^0\mathbf{A} = \lambda_0\mathbf{f}^0$, а подпространство левых собственных векторов, соответствующих λ_0, одномерно. Следуя доказательству теоремы 2.1, положим $\lambda' = \sup_{\Lambda'} \lambda$, где $\Lambda' = \{\lambda |\ \mathbf{f}\mathbf{A} \geqslant \lambda\mathbf{f}$ для некоторого $\mathbf{f} > \mathbf{0}\}$. Легко убедиться в существовании вектора $\mathbf{f}^0 \gg \mathbf{0}$, такого, что $\mathbf{f}^0\mathbf{A} = \lambda'\mathbf{f}^0$, в одномерности подпространства левых собственных векторов, отвечающих λ', а также в том, что если λ — любое собственное значение матрицы $\mathbf{A}$, не равное λ', то $|\lambda| < \lambda'$. Но отсюда следует, что $|\lambda_0| < \lambda'$, если $\lambda_0 \neq \lambda'$, тогда как по теореме 2.1 $|\lambda'| < \lambda_0$, если собственное значение λ' отлично от λ_0. Следовательно, $\lambda' = \lambda_0$.

Теорема 2.2. *Если матрица $\mathbf{A} > \mathbf{0}$ такова, что $\mathbf{A}^m \gg \mathbf{0}$ для некоторого целого числа $m > 0$, то для нее остаются справедливыми утверждения предыдущей теоремы.*

Доказательство. Как и в доказательстве теоремы 2.1, мы можем найти вектор $\mathbf{x}^0 > \mathbf{0}$, такой, что $\mathbf{A}\mathbf{x}^0 \geqslant \lambda_0\mathbf{x}^0$. Если

$\mathbf{A}\mathbf{x}^0 = \lambda_0\mathbf{x}^0$, то $\mathbf{A}\mathbf{x}^0 > \lambda_0\mathbf{x}^0$. Умножая обе части этого неравенства слева на $\mathbf{A}^m$, приходим к соотношению $\mathbf{A}\mathbf{y} \gg \lambda_0\mathbf{y}$, где $\mathbf{y} = \mathbf{A}^m\mathbf{x}^0 \gg \mathbf{0}$. Это противоречит определению λ_0 (см. доказательство теоремы 2.1); следовательно, $\mathbf{A}\mathbf{x}^0 = \lambda_0\mathbf{x}^0$. Отсюда получаем соотношение $\mathbf{A}^m\mathbf{x}^0 = \lambda_0^m\mathbf{x}^0$. Так как $\mathbf{A}^m \gg \mathbf{0}$ и $\mathbf{x}^0 > \mathbf{0}$, то $\lambda_0^m\mathbf{x}^0 \gg \mathbf{0}$, т. е. $\mathbf{x}^0 \gg \mathbf{0}$.

Для доказательства неравенства $|\lambda| \leqslant \lambda_0$ в пункте (б) теоремы 2.1 нам было бы достаточно положительности (не строгой) матрицы $\mathbf{A}$. Поэтому это неравенство выполняется и в условиях настоящей теоремы. Предположим, что $|\lambda| = \lambda_0$ и $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$ для некоторого вектора $\mathbf{z} \neq \mathbf{0}$. Тогда $\mathbf{A}^m\mathbf{z} = \lambda^m\mathbf{z}$, $\mathbf{A}^m\mathbf{x}^0 = \lambda_0^m\mathbf{x}^0$ и $|\lambda^m| = \lambda_0$. Если мы покажем, что λ_0^m является наибольшим по модулю положительным собственным значением матрицы $\mathbf{A}^m$, то теорема 2.1 сразу же приведет нас к противоречию. Так как $\mathbf{A}^m \gg \mathbf{0}$, то $\mathbf{A}^m$ имеет наибольшее по модулю положительное собственное значение по теореме 2.1, которому соответствует собственный вектор с положительными компонентами. В таком случае, если λ_0^m не является наибольшим по модулю положительным собственным значением матрицы $\mathbf{A}^m$, то матрица $\mathbf{A}^m$ имеет по крайней мере два положительных собственных значения $\lambda_1 > \lambda_2$, которым соответствуют собственные векторы $\mathbf{x}_1, \mathbf{x}_2 \gg \mathbf{0}$. Но это невозможно; действительно, пусть $\mu > 0$ таково, что $\mathbf{x}_2 - \mu\mathbf{x}_1 > \mathbf{0}$, но не $\gg \mathbf{0}$. Тогда $\mathbf{A}^m(\mathbf{x}_2 - \mu\mathbf{x}_1) \gg \mathbf{0}$. С другой стороны, $\mathbf{A}^m(\mathbf{x}_2 - \mu\mathbf{x}_1) = \lambda_2\mathbf{x}_2 - \mu\lambda_1\mathbf{x}_1 = \lambda_2(\mathbf{x}_2 - \mu\mathbf{x}_1) - (\lambda_1 - \lambda_2)\mu\mathbf{x}_1$. Поскольку вектор $(\mathbf{x}_2 - \mu\mathbf{x}_1)$ имеет по крайней мере одну нулевую компоненту, в то время как все компоненты вектора $\mathbf{x}_1$ положительны, мы приходим к противоречию.

Тот факт, что все собственные векторы матрицы $\mathbf{A}$ являются собственными векторами матрицы $\mathbf{A}^m$, сводит доказательство утверждения (в) к повторению доказательства утверждения (в) теоремы 2.1. ■

Продолжая изучение матриц $\mathbf{A} > \mathbf{0}$, для которых $\mathbf{A}^m \gg \mathbf{0}$ при некотором целом $m > 0$, введем матрицу ранга 1 по формуле

$$\mathbf{P} = \| x_i^0 f_j^0 \|,$$

где $\mathbf{x}^0$ — правый собственный вектор, введенный в доказательстве теоремы 2.1, а $\mathbf{f}^0 \gg \mathbf{0}$ — левый собственный вектор, принадлежащий λ_0 и нормированный условием $\sum_{i=1}^{n} x_i^0 f_i^0 = 1$. Такая матрица обладает следующими свойствами:

(I) $\mathbf{P}\mathbf{x} = (\mathbf{x}, \mathbf{f}^0)\,\mathbf{x}^0$, $\mathbf{f}\mathbf{P} = (\mathbf{f}, \mathbf{x}^0)\,\mathbf{f}^0$ для любых векторов $\mathbf{x}$ и $\mathbf{f}$ и, в частности, $\mathbf{P}\mathbf{x}^0 = \mathbf{x}^0$, $\mathbf{f}^0\mathbf{P} = \mathbf{f}^0$;

(II) $\mathbf{P}^2 = \mathbf{P}$;

(III) $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{A} = \lambda_0\mathbf{P}$.

Первые два свойства проверяются непосредственно; что же касается третьего, то заметим, что для любого вектора x имеем

$$\mathbf{APx} = \mathbf{A}(\mathbf{x}, \mathbf{f}^0)\mathbf{x}^0 = (\mathbf{x}, \mathbf{f}^0)\mathbf{Ax}^0 = (\mathbf{x}, \mathbf{f}^0)\lambda_0\mathbf{x}^0 = \lambda_0\mathbf{Px},$$

так что $\mathbf{AP} = \lambda_0\mathbf{P}$; аналогично $\mathbf{fPA} = \mathbf{f}\lambda_0\mathbf{P}$, т. е. $\mathbf{PA} = \lambda_0\mathbf{P}$.

Приведем без доказательства следующий факт. Пусть $\mathbf{B}$ — квадратная матрица; положим

$$r = \varlimsup_{n\to\infty} \sqrt[n]{\max_{i,j} |b_{ij}^{(n)}|},$$

где $b_{ij}^{(n)}$ — элементы матрицы $\mathbf{B}^n$. Матрица $\mathbf{B}$ имеет собственное значение λ^*, такое, что $|\lambda^*| = r$; для всех других собственных значений λ матрицы $\mathbf{B}$ имеет место неравенство $|\lambda| < r$.

r часто называют спектральным радиусом матрицы. Мы воспользуемся этим понятием при доказательстве следующей теоремы.

Теорема 2.3. *Если* $\mathbf{A} > 0$ *и* $\mathbf{A}^m \gg 0$ *для некоторого целого* $m > 0$, *то*

$$\frac{1}{\lambda_0^n}\mathbf{A}^n \to \mathbf{P} \quad \textit{при} \quad n \to \infty,$$

где λ_0 *и* $\mathbf{P}$ *имеют прежний смысл.*

Доказательство. Покажем сначала, что если λ является собственным значением матрицы $\mathbf{B} = \mathbf{A} - \lambda_0\mathbf{P}$, то $|\lambda| < \lambda_0$. Действительно, предположим, что $\mathbf{Bz} = \lambda\mathbf{z}$ при некотором $\mathbf{z} \neq \mathbf{0}$; тогда

$$\lambda\mathbf{Pz} = \mathbf{PBz} = \mathbf{P}(\mathbf{A} - \lambda_0\mathbf{P})\mathbf{z} = (\lambda_0\mathbf{P} - \lambda_0\mathbf{P}^2)\mathbf{z} = \lambda_0(\mathbf{P} - \mathbf{P})\mathbf{z} = \mathbf{0},$$

откуда следует, что всякое собственное значение матрицы $\mathbf{B}$ является собственным значением матрицы $\mathbf{A}$; то же можно сказать о собственных векторах. Из теоремы 2.2 следует, что либо $\lambda = \lambda_0$, либо $|\lambda| < \lambda_0$. Предположим, что $\lambda = \lambda_0$, тогда в силу одномерности собственного подпространства, соответствующего λ_0, матрицы $\mathbf{A}$, а следовательно, и матрицы $\mathbf{B}$ имеет место соотношение $\mathbf{Bx}^0 = \lambda_0\mathbf{x}^0$. С другой стороны, $\mathbf{Bx}^0 = (\mathbf{A} - \lambda_0\mathbf{P})\mathbf{x}^0 = \mathbf{Ax}^0 - \lambda_0\mathbf{Px}^0 = \lambda_0\mathbf{x}^0 - \lambda_0\mathbf{x}^0 = 0$. Получаем противоречие. Значит, спектральный радиус r матрицы $\mathbf{B}$ удовлетворяет неравенству $r < \lambda_0$. Пусть ρ — положительное число, такое, что $r < \rho < \lambda_0$; поскольку

$$r = \varlimsup_{n\to\infty} \sqrt[n]{\max_{i,j} |b_{ij}^{(n)}|} < \rho,$$

то $\max\limits_{i,j}|b_{ij}^{(n)}| < \rho^n$ для достаточно больших n. Используя свойства (II) и (III) матрицы $\mathbf{P}$, получаем

$$\mathbf{B}^m = \mathbf{A}^m - \lambda_0^m\mathbf{P},$$

или

$$\frac{\mathbf{A}^m}{\lambda_0^m} = \frac{\mathbf{B}^m}{\lambda_0^m} + \mathbf{P}.$$

Так как $\max_{i,j} |b_{ij}^{(n)}| < \rho^n$ для достаточно больших n, то

$$\left| \frac{b_{ij}^{(m)}}{\lambda_0^m} \right| \leqslant \left(\frac{\rho}{\lambda_0} \right)^m \to 0,$$

а следовательно, $\mathbf{B}^m/\lambda_0^m \to \mathbf{0}$. ■

Вторая теорема Фробениуса

Теорема 2.4. *Пусть* $\mathbf{A} > \mathbf{0}$, *а* λ_0 *имеет тот же смысл, что и в теореме* 2.1; *тогда*: (а) λ_0 *является собственным значением матрицы* $\mathbf{A}$ *и ему соответствует собственный вектор* $\mathbf{x}^0 > \mathbf{0}$; (б) *если* λ — *любое другое собственное значение матрицы* $\mathbf{A}$, *то* $|\lambda| \leqslant \lambda_0$; (в) *среднее*

$$\frac{1}{m} \sum_{i=1}^{m} \frac{\mathbf{A}^i}{\lambda_0^i}$$

сходится, если $\mathbf{x}^0 \gg \mathbf{0}$; (г) *если* λ — *собственное значение матрицы* $\mathbf{A}$ *и* $|\lambda| = \lambda_0$, *то* $\eta = \lambda/\lambda_0$ *является корнем из единицы, а* $\eta^m \lambda_0$, $m = 0, 1, 2, \ldots$, *суть собственные значения матрицы* $\mathbf{A}$.

Доказательство. (а) Пусть $\mathbf{E}$ — матрица, все элементы которой равны единице; тогда $\mathbf{A} + \delta\mathbf{E} \gg \mathbf{0}$ при любом $\delta > 0$. Пусть $0 < \delta_1 < \delta_2$. Выберем $\mathbf{x} = (x_1, \ldots, x_n) > \mathbf{0}$ таким, что $\sum_{i=1}^{n} x_i = 1$, тогда из $(\mathbf{A} + \delta_1\mathbf{E})\mathbf{x} \geqslant \lambda\mathbf{x}$ следует, что

$$(\mathbf{A} + \delta_2\mathbf{E})\,\mathbf{x} = (\mathbf{A} + \delta_1\mathbf{E})\,\mathbf{x} + (\delta_2 - \delta_1)\,\mathbf{E}\mathbf{x} \geqslant [\lambda + (\delta_2 - \delta_1)]\,\mathbf{x}.$$

Таким образом, если $\lambda_0(\delta)$ — значение λ_0, соответствующее матрице $\mathbf{A} + \delta\mathbf{E}$, то $\lambda_0(\delta)$ является возрастающей функцией аргумента δ. Заметим, что $\lambda_0(0)$ есть значение λ_0, соответствующее самой матрице $\mathbf{A}$. По теореме 2.1 существует вектор $\mathbf{x}(\delta) \gg \mathbf{0}$, нормированный условием $\sum_{i=1}^{n} x_i(\delta) = 1$, такой, что

$$(\mathbf{A} + \delta\mathbf{E})\,\mathbf{x}(\delta) = \lambda_0(\delta)\,\mathbf{x}(\delta).$$

Пусть $\delta_1 > \delta_2 > \ldots$ — последовательность положительных чисел, сходящаяся к нулю. Как и при доказательстве теоремы 2.1, мы можем выделить последовательность целых чисел $n_1, n_2, \ldots$, такую, что $\lim\limits_{j\to\infty} \mathbf{x}(\delta_{n_j}) \to \mathbf{x}^0$, где вектор $\mathbf{x}^0 > \mathbf{0}$ и $\sum\limits_{i=1}^{n} \mathbf{x}_i^0 = 1$. Ясно, что $\mathbf{A} + \delta_{n_j}\mathbf{E} \to \mathbf{A}$ и $\lambda_0(\delta_{n_j}) \to \lambda' \geqslant \lambda_0$. Поскольку

$$(\mathbf{A} + \delta_{n_j}\mathbf{E})\,\mathbf{x}(\delta_{n_j}) = \lambda_0(\delta_{n_j})\,\mathbf{x}(\delta_{n_j}),$$

в пределе при $j \to \infty$ получаем равенство $\mathbf{A}\mathbf{x}^0 = \lambda'\mathbf{x}^0$. Но из доказательства пункта (б) теоремы 2.1 следует, что $\lambda_0 \geqslant \lambda'$; значит, $\lambda_0 = \lambda'$.

Доказательство утверждения (б) совпадает с доказательством для случая $\mathbf{A} \gg \mathbf{0}$.

Доказывая утверждения (в) и (г), мы можем предположить без потери общности, что $\lambda_0 = 1$, так как если $\lambda_0 \neq 1$, то можно разделить каждый элемент матрицы $\mathbf{A}$ на λ_0.

(в) Так как $\mathbf{A}\mathbf{x}_0 = \mathbf{x}_0$, то $\mathbf{A}^m\mathbf{x}_0 = \mathbf{x}_0$. Записав это соотношение в координатной форме, мы сразу же найдем, что

$$0 \leqslant a_{ij}^{(m)} \leqslant \frac{\max\limits_i x_i^0}{\min\limits_i x_i^0}.$$

Таким образом, элементы матриц $\mathbf{A}^m$ равномерно ограничены.

Пусть $\mathbf{L} = \{\mathbf{x}\colon \mathbf{A}\mathbf{x} = \mathbf{x}\}$ и $\mathbf{K} = \{\mathbf{y}\colon \mathbf{y} = (\mathbf{I} - \mathbf{A})\mathbf{x}$ для некоторого $\mathbf{x}\}$, т. е. $\mathbf{L}$ есть линейное пространство неподвижных точек матрицы $\mathbf{A}$, а $\mathbf{K}$ — линейное пространство, совпадающее с областью значений матрицы $\mathbf{I} - \mathbf{A}$. Положим еще

$$\mathbf{S}_m = \frac{\mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^m}{m}.$$

Ясно, что $\mathbf{L}$ является замкнутым линейным пространством, таким, что для каждого $\mathbf{x}$ из $\mathbf{L}$

$$\mathbf{S}_m\mathbf{x} = \frac{\mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^m}{m}\,\mathbf{x} = \mathbf{x},$$

и следовательно, $\lim\limits_{m\to\infty} \mathbf{S}_m\mathbf{x} = \mathbf{x}$. Мы покажем также, что $\mathbf{S}_m\mathbf{x}$ сходится для каждого $\mathbf{x}$ из $\mathbf{K}$ и что n-мерное векторное пространство можно представить в виде прямой суммы пространств $\mathbf{L}$ и $\mathbf{K}$. Этим мы докажем утверждение (в).

Покажем сначала, что $\lim\limits_{m\to\infty} \mathbf{S}_m\mathbf{y} = \mathbf{0}$ для всякого $\mathbf{y} \in \mathbf{K}$. Поскольку $\mathbf{y} = (\mathbf{I} - \mathbf{A})\mathbf{x}$ при некотором $\mathbf{x}$,

$$\mathbf{S}_m\mathbf{y} \equiv \frac{\mathbf{A}\mathbf{y} + \ldots + \mathbf{A}^m\mathbf{y}}{m} = \frac{\mathbf{A}\mathbf{x} - \mathbf{A}^{m+1}\mathbf{x}}{m}$$

стремится к 0 при $m \to \infty$ в силу того, что элементы матриц $\mathbf{A}^m$ равномерно ограничены.

Чтобы показать, что каждый вектор $\mathbf{x}$ можно представить в виде суммы вектора из $\mathbf{L}$ и вектора из $\mathbf{K}$, разложим $\mathbf{x}$ в сумму вида

$$\mathbf{x} = (\mathbf{x} - \mathbf{S}_m\mathbf{x}) + \mathbf{S}_m\mathbf{x} = \mathbf{y}_m + \mathbf{z}_m.$$

Так как элементы матриц $\mathbf{A}^m$ равномерно ограничены, то ограничены и компоненты векторов $\mathbf{y}_m$ и $\mathbf{z}_m$. Поэтому существует последовательность положительных целых чисел $m_1 < m_2 < \ldots$ и вектор $\mathbf{z}^0$, такие, что

$$\lim_{i\to\infty} \mathbf{z}_{m_i} = \mathbf{z}_0.$$

Поскольку

$$\mathbf{z}_{m_i} - \mathbf{A}\mathbf{z}_{m_i} = \frac{\mathbf{A} - \mathbf{A}^{m_i+1}}{m_i}\mathbf{x} \to 0 \quad \text{при} \quad i \to \infty,$$

то

$$\mathbf{z}_0 = \lim_{i\to\infty} \mathbf{z}_{m_i} = \lim_{i\to\infty} \mathbf{A}\mathbf{z}_{m_i} = \mathbf{A} \lim_{i\to\infty} \mathbf{z}_{m_i} = \mathbf{A}\mathbf{z}_0,$$

т. е. $\mathbf{z}_0 \in \mathbf{L}$.

Кроме того,

$$\mathbf{y}_m = \mathbf{x} - \mathbf{S}_m\mathbf{x} = \frac{1}{m}[(\mathbf{x} - \mathbf{A}\mathbf{x}) + (\mathbf{x} - \mathbf{A}^2\mathbf{x}) + \ldots + (\mathbf{x} - \mathbf{A}^m\mathbf{x})] =$$
$$= (\mathbf{I} - \mathbf{A})\left[\frac{\mathbf{x}}{m} + \frac{(\mathbf{I}+\mathbf{A})\,\mathbf{x}}{m} + \frac{(\mathbf{I}+\mathbf{A}+\mathbf{A}^2)\,\mathbf{x}}{m} + \ldots \right.$$
$$\left. \ldots + \frac{(\mathbf{I}+\mathbf{A}+\mathbf{A}^2+\ldots+\mathbf{A}^{m-1})\,\mathbf{x}}{m}\right],$$

откуда следует, что $\mathbf{y}_m \in \mathbf{K}$. Так как $\mathbf{K}$ — замкнутое линейное пространство и элементы векторов $\mathbf{y}_m$ равномерно ограничены, то $\mathbf{y}_{m_i} \to \mathbf{x} - \mathbf{z}_0 \in \mathbf{K}$ при $i \to \infty$. Итак, $\mathbf{x} = (\mathbf{x} - \mathbf{z}_0) + \mathbf{z}_0$, где $\mathbf{x} - \mathbf{z}_0 \in \mathbf{K}$, $\mathbf{z}_0 \in \mathbf{L}$, и доказательство утверждения (в) закончено.

(г) Мы знаем, что существует вектор $\mathbf{f}^0 > 0$, такой, что $\mathbf{f}^0\mathbf{A} = \mathbf{f}^0$. Предположим сначала, что $\mathbf{f}^0 \gg 0$. Пусть теперь $\lambda \neq 1$, $|\lambda| = 1$ и $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ при некотором $\mathbf{x} \neq \mathbf{0}$; тогда

$$\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i, \qquad i = 1, 2, \ldots, n,$$

и, таким образом,

$$\sum_{j=1}^{n} a_{ij}|x_j| \geqslant |x_i|, \quad \text{или} \quad \mathbf{A}|\mathbf{x}| \geqslant |\mathbf{x}|.$$

Но если $\mathbf{A}|\mathbf{x}| > |\mathbf{x}|$, то $(\mathbf{f}^0, |\mathbf{x}|) < (\mathbf{f}^0, \mathbf{A}|\mathbf{x}|) = (\mathbf{f}^0\mathbf{A}, |\mathbf{x}|) = (\mathbf{f}^0, |\mathbf{x}|)$. Итак, $\mathbf{A}|\mathbf{x}| = |\mathbf{x}|$; следовательно,

$$\sum_{j=1}^{n} a_{ij}|x_j| = |x_i| = \left|\sum_{j=1}^{n} a_{ij}x_j\right|, \qquad i = 1, 2, \ldots, n.$$

Это означает, что существуют такие константы $\mu_1, \ldots, \mu_n$, $|\mu_i|=1$, что

$$a_{ij}x_j = a_{ij}|x_j|\mu_i \quad \text{при всех } i \text{ и } j. \qquad (*)$$

Обозначим через $\mathbf{x}\cdot\mathbf{y}$ вектор $(x_1y_1, \ldots, x_ny_n)$. Умножая соотношение (*) на μ_j^r и суммируя по j, получаем

$$\mathbf{A}(\mathbf{x}\cdot\boldsymbol{\mu}^r) = \boldsymbol{\mu}\cdot\mathbf{A}(|\mathbf{x}|\cdot\boldsymbol{\mu}^r).$$

В то же время суммирование по i приводит к равенству

$$\mathbf{A}\mathbf{x} = \boldsymbol{\mu}\cdot\mathbf{A}|\mathbf{x}|,$$

откуда следует, что

$$\lambda\mathbf{x} = \boldsymbol{\mu}\cdot|\mathbf{x}|.$$

Далее,

$$\mathbf{A}(\mathbf{x}\cdot\boldsymbol{\mu}^r) = \boldsymbol{\mu}\cdot\mathbf{A}(\lambda\mathbf{x}\cdot\boldsymbol{\mu}^{r-1}) = \lambda\boldsymbol{\mu}\cdot\mathbf{A}(\mathbf{x}\cdot\boldsymbol{\mu}^{r-1}), \qquad r=1, 2, \ldots,$$

откуда следует, что $\mathbf{A}(\mathbf{x}\cdot\boldsymbol{\mu}^r) = \lambda^{r+1}(\boldsymbol{\mu}^r\cdot\mathbf{x})$. Итак, λ^r является собственным значением матрицы $\mathbf{A}$ при $r=1, 2, \ldots$. Поскольку число собственных значений матрицы $\mathbf{A}$ конечно, λ должно быть корнем из единицы.

Пусть теперь $\mathbf{f}^0 > \mathbf{0}$, но не $\gg \mathbf{0}$. Проведя в случае надобности перенумерацию строк и столбцов матрицы $\mathbf{A}$, мы можем считать, что $\mathbf{f}^0 = (f_1^0, \ldots, f_r^0, 0, \ldots, 0)$, где $f_i^0 > 0$, $i=1, \ldots, r$. Так как $\mathbf{A} > \mathbf{0}$, то из соотношения $\mathbf{f}^0\mathbf{A} = \mathbf{f}^0$ вытекает, что

$$\mathbf{A} = \begin{Vmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{B} & \mathbf{A}_2 \end{Vmatrix},$$

где порядки квадратных матриц $\mathbf{A}_1$ и $\mathbf{A}_2$ равны соответственно r и $n-r$. Вектор $(f_1^0, \ldots, f_r^0)$ является левым собственным вектором матрицы $\mathbf{A}_1$. Пусть λ — собственное значение матрицы $\mathbf{A}$; тогда, если λ является собственным значением матрицы $\mathbf{A}_1$, то задача сводится к рассмотренному выше случаю. Если же λ не является собственным значением матрицы $\mathbf{A}_1$, то оно должно быть собственным значением матрицы $\mathbf{A}_2$. Но собственные значения матрицы $\mathbf{A}_2$ являются собственными значениями матрицы $\mathbf{A}$ и поэтому не превышают по модулю единицы. В то же время, так как $\mathbf{A}_2 > 0$, то она имеет наибольшее положительное собственное значение, являющееся верхней границей для абсолютных величин всех остальных ее собственных значений. Так как $|\lambda| = 1$, то наибольшее собственное значение матрицы $\mathbf{A}_2$ равно единице. Теперь предыдущие рассуждения применимы к матрице $\mathbf{A}_2$, именно: либо $\mathbf{A}_2$ имеет левый собственный вектор со всеми положительными координатами, соответствующий собственному значению 1, либо $\mathbf{A}_2$

имеет вид (после соответствующей перенумерации строк и столбцов)

$$\mathbf{A}_2 = \begin{Vmatrix} \mathbf{A}_3 & \mathbf{0} \\ \mathbf{B}_1 & \mathbf{A}_4 \end{Vmatrix}.$$

Продолжая таким образом, мы придем за конечное число шагов к ситуации, когда существует левый собственный вектор $\gg 0$, соответствующий собственному значению 1. ■

Приводимые ниже следствия содержат полезную информацию относительно спектрального радиуса $\lambda_0(\mathbf{A})$ положительной матрицы $\mathbf{A}$. Первое из них представляет собой другую формулировку утверждений (а) и (б) теоремы 2.4.

Следствие 2.1. *Если $\mathbf{A} > \mathbf{0}$, то наибольшее по абсолютной величине собственное значение $\lambda_0 = \lambda_0(\mathbf{A})$ является действительным неотрицательным и может быть охарактеризовано как $\lambda_0 = \max\limits_{\Lambda} \lambda$, где*

$$\Lambda = \{\lambda \mid \mathbf{A}\mathbf{x} \geqslant \lambda \mathbf{x},\ \mathbf{x} > \mathbf{0}\}.$$

Следствие 2.2. *Если $\mathbf{A} > \mathbf{0}$ и существует вектор $\mathbf{x}^0 \gg \mathbf{0}$, такой, что $\mathbf{A}\mathbf{x}^0 \leqslant \mu \mathbf{x}^0$, то μ является верхней границей для $\lambda_0(\mathbf{A})$.*

Доказательство. Умножая слева на $\mathbf{A}$ обе части неравенства $\mathbf{A}\mathbf{x}^0 \leqslant \mu\mathbf{x}^0$, получаем $\mathbf{A}^2\mathbf{x}^0 \leqslant \mu\mathbf{A}\mathbf{x}^0 \leqslant \mu^2\mathbf{x}^0$. Легко видеть, что и в общем случае

$$\mathbf{A}^n\mathbf{x}^0 \leqslant \mu^n\mathbf{x}^0, \qquad n = 1, 2, \ldots,$$

откуда сразу же следует неравенство

$$a_{ij}^{(n)} \leqslant \mu^n \frac{\max\limits_i x_i^0}{\min\limits_i x_i^0}.$$

Это неравенство приводит к оценке

$$\lambda_0(\mathbf{A}) = \varlimsup_{n\to\infty} \sqrt[n]{\max_{i,j} |a_{ij}^{(n)}|} \leqslant \mu. \ ■$$

Следствие 2.3. *Если $\mathbf{A} \geqslant \mathbf{B} \geqslant \mathbf{0}$, то $\lambda_0(\mathbf{B}) \leqslant \lambda_0(\mathbf{A})$.*

Доказательство. Так как из $\mathbf{A} \geqslant \mathbf{B} \geqslant \mathbf{0}$ следует, что $\mathbf{A}^n \geqslant \mathbf{B}^n \geqslant \mathbf{0}$, $n = 1, 2, \ldots$, то

$$\lambda_0(\mathbf{A}) = \varlimsup_{n\to\infty} \sqrt[n]{\max_{i,j} a_{ij}^{(n)}} \geqslant \varlimsup_{n\to\infty} \sqrt[n]{\max_{i,j} b_{ij}^{(n)}} = \lambda_0(\mathbf{B}). \ ■$$

РАЗЛИЧНЫЕ ЗАДАЧИ

1. Рассмотрим популяцию из n пар, в которой у i-й пары с вероятностью p_i рождается мальчик, а среднее число детей в семье равно c. Предположим, что вероятности p_i остаются постоянными во времени для всех пар, при каждых родах рождается не более одного ребенка, а пол рожденного ребенка не зависит от пола детей, рожденных ранее и тех, что появятся на свет после него. Введем характеристику распределения детей по полу

$$S = \frac{\text{среднее число мальчиков, рожденных в популяции из } n \text{ пар}}{\text{среднее число детей, рожденных в популяции из } n \text{ пар}}.$$

Найти S для этой популяции.

Ответ: $S \equiv S_0 = \left(\sum_{i=1}^{n} p_i\right)\Big/ n.$

2 (продолжение). Показать, что если все пары решают рождать детей до тех пор, пока не будет рожден мальчик, после чего детей больше не имеют, то

$$S = S_1 = \frac{n}{\sum_{i=1}^{n} \frac{1}{p_i}} \leqslant S_0.$$

3 (продолжение). Предположим, что все пары в зависимости от того, мальчик или девочка их первый ребенок, продолжают рождать детей, пока не будет рождена девочка или, во втором случае, мальчик, после чего детей больше не имеют. Найти S для этого случая.

Ответ:

$$S = S_2 = \frac{\left\{\sum_{i=1}^{n} 1/q_i - \sum_{i=1}^{n} p_i\right\}}{\sum_{i=1}^{n} [p_i q_i]^{-1} - n}, \qquad q_i = 1 - p_i.$$

4 (продолжение). Пусть все пары, если их первый ребенок — мальчик, продолжают рождать детей, пока не родится девочка. Если же первый ребенок — девочка, то пары ограничиваются одним ребенком. Найти S.

Ответ:

$$S = S_3 = 1 - \left[\frac{n}{\sum_{i=1}^{n} 1/q_i}\right].$$

5 (продолжение). Показать, что в зависимости от значений $p_1, \ldots, p_n$ может быть как $S_2 > S_0$, так и $S_2 \leqslant S_0$.

6 (продолжение). а) Предположим, что вероятность осложнений при родах для i-й пары равна p_i, $i = 1, \ldots, n$, а наступление осложнений приводит к тому, что пара решает больше детей не иметь. Пусть s — число детей в отдельно взятой семье при условии отсутствия родов с осложнениями, и вероятность родов с осложнениями p_i не изменяется во времени. Показать, что

$$R_1 = \frac{\text{среднее число родов с осложнениями}}{\text{среднее число рожденных детей}} = \frac{\sum\limits_{i=1}^{n} 1 - q_i^s}{\sum\limits_{i=1}^{n} \left(1 - q_i^s\right)/p_i}.$$

б) Предположим, что пара решает больше не иметь детей после двух (а не одних, как ранее) родов с осложнениями. Показать, что в этом случае

$$R_2 = \frac{\sum\limits_{i=1}^{n} \left\{2\left(1 - q_i^s\right) - sp_i q_i^{s-1}\right\}}{\sum\limits_{i=1}^{n} \left\{\left[2\left(1 - q_i^s\right)/p_i\right] - sq_i^{s-1}\right\}}.$$

7. Пусть X и Y — пара независимых неотрицательных целочисленных с. в., обладающих тем свойством, что

$$P\{X = x \mid X + Y = x + y\} = \frac{\binom{m}{x}\binom{n}{y}}{\binom{m+n}{x+y}}$$

для всех неотрицательных целых чисел x и y (m и n — заданные положительные целые числа). Предположим, что вероятности $P\{X = 0\}$ и $P\{Y = 0\}$ строго положительны. Показать, что как X, так и Y распределены по биномиальному закону с одним и тем же параметром p, причем другими параметрами служат m и n соответственно.

8. (а) Пусть X и Y — независимые случайные величины, такие, что $P\{X = i\} = f(i)$, $P\{Y = i\} = g(i)$, где $f(i) > 0$ и $g(i) > 0$, $\sum\limits_{i=0}^{\infty} f(i) = \sum\limits_{i=0}^{\infty} g(i) = 1$, $i = 0, 1, 2, \ldots$. Пусть, кроме того,

$$P\{X = k \mid X + Y = l\} = \begin{cases} \binom{l}{k} p^k (1-p)^{l-k}, & 0 \leqslant k \leqslant l, \\ 0, & k > l. \end{cases}$$

Доказать, что

$$f(i) = e^{-\theta\alpha} \frac{(\theta\alpha)^i}{i!}, \quad g(i) = e^{-\theta} \frac{\theta^i}{i!}, \qquad \alpha = 0, 1, 2, \ldots,$$

где $\alpha = p/(1-p)$ и $\theta > 0$ произвольно.

(б) Показать, что p определяется из уравнения

$$G\left(\frac{1}{1-p}\right) = \frac{1}{f(0)},$$

$$G(s) = \sum g(i) s^i.$$

Указание: Сначала установить соотношение

$$F(u)\,F(v)=F(vp+(1-p)\,u)\,G(vp+(1-p)\,u),$$

где $F(s)=\sum f(i)\,s^i$.

9. Пусть X — неотрицательная целочисленная с. в. с производящей функцией $f(s)=\sum\limits_{n=0}^{\infty}a_n s^n$. Предположим, что после наблюдения с в. X проводится X биномиальных испытаний с вероятностью успеха, равной p. Пусть Y обозначает результирующее число успехов.

(а) Найти производящую функцию с. в. Y.

(б) Найти производящую функцию условного распределения X при условии, что X биномиальных испытаний закончились успехом.

Ответ: (а) $f(1-p+qs)$; (б) $f(ps)/f(p)$.

10 (продолжение). Предположим, что при каждом p $(0<p<1)$ производящие функции вероятностей (а) и (б) совпадают. Доказать, что в этом случае с. в. X распределена по закону Пуассона, т. е. $f(s)=e^{\lambda(s-1)}$ с некоторым параметром $\lambda>0$.

11. Рассмотрим последовательность $X_1, X_2, \ldots$ независимых одинаково распределенных с. в., функция распределения которых $F(x)$ непрерывна. Значение X_k называется «рекордным», если

$$X_k>\max(X_1, X_2, \ldots, X_{k-1}).$$

(По определению значение X_1 является рекордным.)

Пусть N_n есть число рекордных значений в последовательности $X_1, X_2, \ldots, X_n$, а V_k — индекс k-го рекордного значения (см. задачи 21—28 гл. 9). Доказать соотношение

$$\mathrm{P}\{V_k>n\}=\mathrm{P}\{N_n<k+1\}.$$

12 (продолжение). Доказать формулу

$$\mathrm{P}\{N_n=r\}=\frac{1}{n}\sum_{1<k_1<k_2<\ldots<k_{r-1}<n}\frac{1}{(n-k_{r-1})}\,\frac{1}{(n-k_{r-2})}\cdot\ldots\cdot\frac{1}{(n-k_1)}.$$

13. Пусть $X_1, X_2, \ldots, X_m$ и $Y_1, Y_2, \ldots, Y_n$ — выборки соответственно из m и n случайных величин с одной и той же непрерывной функцией распределения вероятностей $F(x)$. Пусть N — число значений в выборке из второго набора, не превышающих X_k^*, k-й порядковой статистики первого набора. Показать, что распределение с. в. N имеет вид

$$\mathrm{P}\{N=l\}=\frac{\binom{m}{k}\binom{n}{l}}{\binom{m+n}{k+l}}\,\frac{k}{k+l}\quad\text{для}\quad l=0,\ 1,\ \ldots,\ n.$$

Указание: Воспользоваться результатом задачи 7 гл. 9.

14. Рассмотрим выборку объема n из популяции с функцией плотности вероятностей $f(x)$. Из этой же популяции берется другая выборка того же объема, не зависящая от первой. Пусть U_r^n — число значений во второй выборке, превышающих r-е наименьшее значение в первой выборке, а V_s^n — число значений во второй выборке, больших s-го наибольшего значения в первой выборке. Доказать, что

$$\mathrm{P}\left\{U_r^n = x\right\} = \frac{\binom{n-x+r-1}{r-1}\binom{n-r+x}{x}}{\binom{2n}{n}} = \frac{1}{2}\,\frac{\binom{n-1}{r-1}\binom{n}{x}}{\binom{2n-1}{n-r+x}},$$

$x = 0, 1, \ldots, n$, и, кроме того, что

$$\mathrm{P}\left\{U_r^n = x\right\} = \mathrm{P}\left\{V_s^n = x\right\}, \quad \text{где} \quad s = n - r.$$

Указание: Воспользоваться результатом задачи 7 гл. 9.

15. Пусть $X_1, X_2, \ldots, X_m$ и $Y_1, Y_2, \ldots, Y_n$ — независимые наблюдения над случайными величинами с распределениями $F(x)$ и $G(y)$ соответственно. Предположим, что F и G — непрерывные функции и

$$F = G^\delta, \qquad \delta > 0.$$

Пусть

$$X_{1,\,m} \leqslant X_{2,\,m} \leqslant \ldots \leqslant X_{m,\,m}$$

и

$$Y_{1,\,n} \leqslant Y_{2,\,n} \leqslant \ldots \leqslant Y_{n,\,n}$$

суть порядковые статистики соответствующих выборок. Найти

$$\mathrm{P}\{X_{m-i,\,m} \leqslant Y_{n,\,n} < X_{m-i+1,\,m} \mid Y_{n,\,n} = t\}.$$

Ответ:

$$\binom{m}{i}[G(t)]^{\delta(m-i)}(1-[G(t)]^\delta)^i.$$

16 (продолжение). Найти вероятность того, что ровно i значений из первой выборки больше всех значений из второй выборки.

Ответ:

$$\binom{m}{i}\frac{n}{\delta}\,\frac{\Gamma((n/\delta)+m-i)\,\Gamma(i+1)}{\Gamma((n/\delta)+m+1)}.$$

17 (продолжение). Найти вероятность того, что ровно i значений из первой выборки меньше всех значений из второй выборки.

Ответ:

$$\binom{m}{i} n! \sum_{r=0}^{m-i} (-1)^r \binom{m-r}{r} \frac{1}{[\delta(i+r)+1]\,[\delta(i+r)+2]\cdot\ \ldots\ \cdot[\delta(i+r)+n]}.$$

18. Пусть $X_1, X_2, \ldots, X_m$ и $Y_1, Y_2, \ldots, Y_n$ — независимые случайные выборки для распределений $F(x)$ и $G(y)$ соответственно. Предположим, что $F(x)$ и $G(y)$ — непрерывные строго возрастающие функции, и положим

$$p = \mathrm{P}\{Y < X\},$$

где X и Y — независимые с. в. с распределениями F и G соответственно. Положим

$$U = \text{число пар } (X_i, Y_j), \quad \text{для которых} \quad Y_j < X_i.$$

Показать, что

$$\mathrm{M}(U) = mnp.$$

Указание: Представить U в виде

$$U = \sum_{j=1,\, i=1}^{m,\, n} U_{ij}, \quad \text{где} \quad U_{ij} = \begin{cases} 1, & \text{если} \quad Y_j < X_i, \\ 0, & \text{если} \quad Y_j \geqslant X_i. \end{cases}$$

19 (продолжение). Показать, что

$$\sigma^2(U) = mn\{(m-1)\alpha + (n-1)\beta - (m+n-1)p^2 + (2m-1)p - (m-1)\},$$

где

$$\alpha = \int_0^1 L[(t)]^2\,dt, \quad \beta = 1 - 2\int_0^1 tL(t)\,dt,$$

$L(t) = F(G^{-1}(t))$ и G^{-1} — функция, обратная к G.

20. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — совокупность из n независимых равномерно распределенных на $[0, t]$ с. в. Найти среднее и дисперсию их минимума.

Ответ: Среднее равно $t/n+1$; дисперсия равна $nt^2/[(n+1)^2(n+2)]$.

21. Пусть $X_1, X_2, \ldots, X_n$ — независимые с. в. с плотностями $f_1(x), f_2(x), \ldots$ $\ldots, f_n(x)$ и распределениями $F_1(x), F_2(x), \ldots, F_n(x)$ соответственно. Положим $Z = \min(X_1, X_2, \ldots, X_n)$ и N — первый индекс, при котором достигается $\min(X_1, X_2, \ldots, X_n)$. Пусть

$$H_k(z) = \mathrm{P}\{N = k,\ Z \leqslant z\}.$$

Доказать, что

$$H_k(z) = \int_{-\infty}^{z} \left[\prod_{j \neq k} \{(1 - F_j(t))\}\right] f_k(t)\,dt.$$

22. Группа из 2^n игроков, в равной степени искусных, принимает участие в следующей незамысловатой игре. Игроки разбиваются на пары случайным образом и проводят партию; вероятность выигрыша для каждого из них равна 1/2. В следующей партии 2^{n-1} победителей опять разбиваются на пары случайным образом и т. д. Игра продолжается до тех пор, пока не останется один победитель. Пусть игрок А и игрок Б принимают участие в игре; какова вероятность этим игрокам когда-нибудь оказаться противниками?

Ответ:

$$p_n = \left(\frac{1}{2}\right)^{n-1}.$$

23. Игроки А и Б договариваются провести N бросаний правильной монеты на условиях, соответствующих безобидной игре. Какова вероятность p_N того, что во время игры они ни разу не окажутся при своих интересах?

Ответ:

$$p_N = \begin{cases} \dbinom{N-1}{n}\dfrac{1}{2^N-1}, & \text{если} \quad N = 2n+1, \\[2ex] \dbinom{N}{n}\Big/2^N, & \text{если} \quad N = 2n. \end{cases}$$

24. Пусть $X_1, X_2, \ldots$ — последовательность с. в., такая, что X_k равномерно распределена в интервале $(X_{k-1}, 1)$ и $X_0 \equiv 0$. Доказать, что

$$\prod_{i=1}^{\infty} \mathrm{M}(X_i) = \prod_{i=1}^{\infty}\left(1 - \left(\frac{1}{2}\right)^i\right).$$

Указание: По индукции доказать, что плотность вероятностей с. в. X_i имеет вид

$$f(x)=\frac{(-\ln(1-x))^{i-1}}{(i-1)!}, \qquad i\geqslant 2.$$

25 (продолжение). Показать, что $-\sum_{i=1}^{n} \mathrm{M}(\ln X_i)$ ограничены равномерно по n.

Указание: Показать, что

$$-\sum_{i=1}^{n} \mathrm{M}(\ln X_i)=-\int_0^\infty \ln(1-e^{-\omega})\,e^{-\omega}\sum_{i=0}^{n-1}\frac{\omega^i}{i!}\,d\omega\leqslant-\int_0^\infty \ln(1-e^{-\omega})\,d\omega=\sum_{k=1}^{\infty}\frac{1}{k^2}.$$

26 (продолжение). Показать, что

$$\mathrm{M}\left(\prod_{i=k}^{\infty} X_i \mid X_{k-1}=\xi\right)=e^{\xi-1}.$$

Указание: Положим $\varphi(\xi)=\mathrm{M}\left(\prod_{i=k}^{\infty} X_i \mid X_{k-1}=\xi\right)$. (Почему φ не зависит от k?) Воспользоваться результатом предыдущего упражнения для того, чтобы показать, что $\varphi(\xi)\neq 0$ при $0\leqslant\xi<1$. Затем вывести функциональное уравнение

$$\varphi(x)=\frac{1}{1-x}\int_x^1 \xi\varphi(\xi)\,d\xi, \quad \varphi(1)=1,$$

и найти его решения. Это уравнение имеет два решения, одно из которых непрерывно в точке 1, а другое разрывно. Показать, что решение $\varphi(\xi)=0$ $(0\leqslant\xi<1)$, $\varphi(1)=1$ не является искомым.

27 (продолжение). Показать, что $\mathrm{M}\left(\prod_{i=1}^{\infty} X_i\right)\neq\prod_{i=1}^{\infty}\mathrm{M}(X_i)$, и указать, какая из этих величин больше.

Ответ:

$$\mathrm{M}\left(\prod_1^\infty X_i\right)>\prod_1^\infty \mathrm{M}(X_i).$$

28. Урна содержит n различных занумерованных шаров. Нам ничего не известно о том, каковы эти номера. Мы извлекаем шары один за одним и останавливаемся тогда, когда так или иначе приходим к заключению, что номер последнего извлеченного шара является наибольшим среди номеров всех n шаров. Наша цель — максимизировать вероятность того, что это заключение справедливо.

(а) Рассмотрим следующие стратегии остановки: сначала извлечь m шаров; продолжать извлечение до тех пор, пока номер последнего извлеченного шара не будет больше номеров первых m шаров. Какова вероятность того, что номер последнего извлеченного шара наибольший среди всех?

(б) Выберем p между 0 и 1 и пусть $m=[np]$ (наибольшее целое число, меньшее, чем np). Каково значение p, асимптотически (при $n\to\infty$) наилучшее в смысле максимизации указанной вероятности?

Решение. (а) После извлечения шара, номер которого мы считаем наибольшим среди номеров всех n шаров, продолжим извлечение и выясним, какой же на самом деле наибольший номер. Тогда

$$\mathrm{P}\{\text{выбранный номер наибольший}\} =$$
$$= \sum_{k=m}^{n-1} \mathrm{P}\left\{\begin{matrix}(k+1)\text{-й является наибольшим во всей совокупности и} \\ \text{наибольший среди первых } k \text{ содержится среди первых } m\end{matrix}\right\} =$$
$$= \sum_{k=m}^{n-1} \mathrm{P}\{(k+1)\text{-й номер наибольший}\} \cdot \mathrm{P}\left\{\begin{matrix}\text{наибольший среди первых } k \\ \text{содержится среди первых } m\end{matrix}\right\} =$$
$$= \sum_{k=m}^{n-1} \frac{1}{n} \frac{m}{k} = \frac{m}{n} \sum_{k=m}^{n-1} \frac{1}{k}.$$

Другое доказательство состоит в следующем. Предположим, что номера на шарах представляют собой выборку объема n при некотором известном нам непрерывном распределении F. Покажем, что результат не зависит от F. Действительно,

$$\mathrm{P}\{\text{выбранный номер наибольший}\} =$$
$$= \sum_{k=m}^{n-1} \int_{-\infty}^{\infty} \mathrm{P}\{\max(X_{m+1}, X_{m+2}, \ldots, X_k) < x\} \times$$
$$\times \mathrm{P}\{X_{k+1} = \max(X_{k+1}, X_{k+2}, \ldots, X_n) > x\} \frac{d}{dx} \mathrm{P}\{\max(X_1, \ldots, X_m) > x\}.$$

В силу независимости правая часть этого выражения равна

$$\sum_{k=m}^{n-1} \int_{-\infty}^{\infty} F^{k-m}(x) \frac{1}{n-k} [1 - F^{n-k}(x)]\, mF^{m-1}(x)\, dF(x) =$$
$$= \sum_{k=m}^{n-1} \frac{m}{n-k} \int_0^1 \xi^{k-1} (1 - \xi^{n-k})\, d\xi = \sum_{k=m}^{n-1} \frac{m}{n-k} \left[\frac{1}{k} - \frac{1}{n}\right] = \frac{m}{n} \sum_{k=m}^{n-1} \frac{1}{k},$$

здесь $\xi = F(x)$.

(б) Воспользовавшись найденным решением, получаем

$$\frac{m}{n} \sum_{k=m}^{n-1} \frac{1}{k} = \frac{[np]}{n} \sum_{k=[np]}^{n-1} \frac{1}{k} \sim p \int_{np}^{n} \frac{1}{x}\, dx = -p \ln p.$$

$d/dp[-p \ln p] = 1 - \ln p$ является убывающей функцией p при $0 < p \leqslant 1$. Максимум достигается при $\ln p = -1$, т. е. при $p = e^{-1}$.

29. Задача Банаха о спичечных коробках формулируется следующим образом. Некто носит с собой две коробки спичек А и Б, в которых первоначально было M и N спичек соответственно. Когда ему нужна спичка, он берет ее из коробки А с вероятностью p $(0 < p < 1)$ или из коробки Б с вероятностью $1 - p$. Найти вероятность u_r того, что в момент, когда одна из коробок окажется пустой, в другой будет r спичек.

Ответ:

$$u_r = \binom{M+N-r}{M} p^{M+1} q^{N-r} + \binom{M+N-r}{N} p^{M-r} q^{N+1}.$$

30 (продолжение). Предположим, что всякий раз, когда выбор падает на коробку А, из нее извлекают две спички, а когда выбор падает на коробку Б, из нее по-прежнему извлекают одну спичку. Найти u_r.

Ответ:

$$u_r = \binom{[M/2]+N-r}{[M/2]} p^{[M/2]+1} q^{N-r} + \binom{N+(M-r)/2}{N} p^{(M-r)/2} q^{N+1} \delta_r,$$

где $[x]$ есть целая часть x и $\delta_r = \begin{cases} 1, & \text{если} \quad M-r \text{ четно}, \\ 0, & \text{если} \quad M-r \text{ нечетно}. \end{cases}$

31. Пусть $\{X_i(t);\ t \geqslant 0\}$, $i = 1, 2,$ — независимые пуассоновские процессы с параметрами λ_1 и λ_2 соответственно. Предположим, что $X_1(0) = m$, $X_2(0) = N-1$ и $m < N$. (а) Найти вероятность того, что процесс X_2 достигнет состояния N раньше, чем процесс X_1. (б) Решить эту же задачу для случая, когда $X_2(0) = n$, $n < N$.

Ответ:

$$\text{(б)} \quad \sum_{r=0}^{N-m-1} \binom{N-n+r-1}{r} p^r q^{N-n}, \quad p = \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

32. Проводятся последовательные независимые наблюдения случайной величины, плотность вероятностей которой имеет вид

$$f(x) = \begin{cases} xe^{-x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Наблюдения проводятся до тех пор, пока сумма наблюденных значений не превысит числа t. Пусть $N+1$ — число требующихся для этого наблюдений. Показать, что

$$\mathrm{P}\{N=n\} = \frac{t^{2n+1}e^{-t}}{\Gamma(2n+2)} + \frac{t^{2n}e^{-t}}{\Gamma(2n+1)}.$$

33. Рассмотрим процесс восстановления с функцией распределения $F(x)$ (см. задачи 8—9 гл. 8). Предположим, что каждое событие с вероятностью $1-q$ не регистрируется («стирается»). Изменим масштаб времени на множитель $1/q$. Показать, что результирующая последовательность событий представляет собой процесс восстановления с функцией распределения интервала времени между двумя последующими событиями вида

$$\sum_{n=1}^{\infty} (1-q)^{n-1} qF^{(n)}\left(\frac{x}{q}\right) = F_q(x),$$

где $F^{(n)}$, как обычно, обозначает n-кратную свертку распределения F.

34 (продолжение). Пусть $\varphi(s)$ — преобразование Лапласа функции $F(x)$. Найти преобразование Лапласа функции $F_q(x)$.

Ответ:

$$\varphi_q(s) = \frac{q\varphi(sq)}{1-(1-q)\varphi(sq)}.$$

35 (продолжение). Если распределение F имеет два первых конечных момента, то показать, что

$$\varphi_q(s) \to \frac{1}{1+\lambda s} \quad \text{при} \quad q \to 0+ \quad \text{для всех } s, \quad \operatorname{Re} s \geqslant 0,$$

где $\lambda^{-1} = \int\limits_0^\infty x dF(x)$.

36 (продолжение). С помощью теоремы о сходимости из § 1 гл. 1 показать, что

$$F_q(x) \to 1 - e^{-\lambda x} \quad \text{при} \quad q \to 0+.$$

37. Рассмотрим симметричное случайное блуждание в пространстве r измерений (см. стр. 43). Пусть χ_n есть число траекторий, состоящих из n переходов и не содержащих ни одного состояния дважды. Доказать простое неравенство

$$\chi_{n+m} \leqslant \chi_n \chi_m$$

и с его помощью показать, что предел $\lim_{n \to \infty} \chi^{1/n}$ существует.

Указание: Воспользоваться тем фактом, что функция $\psi_n = \ln \chi_n$ является полуаддитивной, т. е. $\psi_{n+m} \leqslant \psi_n + \psi_m$; см. также стр. 244.

38. Рассмотрим дискретный во времени марковский процесс, пространством состояний которого служит единичный интервал. Если процесс в настоящий момент находится в состоянии p $(0 < p < 1)$, то с вероятностью p в следующий момент он перейдет в состояние $\alpha + \beta p$ и с вероятностью $1 - p$ в состояние βp; здесь α, $\beta > 0$ и $\alpha + \beta = 1$. Таким образом, процесс описывается соотношением

$$X_{n+1} = \begin{cases} \alpha + \beta X_n & \text{с вероятностью } X_n, \\ \beta X_n & \text{с вероятностью } 1 - X_n. \end{cases}$$

Показать, что этот процесс является мартингалом.

***39.** Задача о баллотировке. Кандидат А получил a голосов, а кандидат Б — b голосов $(a > b)$. Предположим, что голоса подсчитывались один за другим, и пусть α_r и β_r представляют число голосов, поданных за кандидата А и за кандидата Б соответственно в момент, когда было рассмотрено r избирательных бюллетеней. Пусть $\Delta_{a,b}$ — число раз, когда кандидат А лидировал при подсчете голосов, т. е. $\Delta_{a,b}$ — число индексов r, при которых $\alpha_r > \beta_r$ $(r = 1, 2, \ldots, a+b)$. Через $\Delta^*_{a,b}$ обозначим число индексов, для которых $\alpha_r \geqslant \beta_r$ $(r = 1, 2, \ldots, a+b)$. Показать, что

$$P\{\Delta_{a,b} = a - b + r\} = \begin{cases} \dfrac{a-b-1}{\dbinom{a+b}{a}} \displaystyle\sum_{m=0}^{[r/2]} \frac{\dbinom{2b-2m}{b-m}\dbinom{a-b-2+2m}{m}}{(b-m+1)(a-b+m-1)}, & a > b+1, \\ \dfrac{1}{2b+1}, & a = b+1, \end{cases}$$

и

$$\mathrm{P}\left\{\Delta^*_{a,\,b}=a-b+r\right\}=$$

$$=\begin{cases}\dfrac{a-b+1}{\binom{a+b}{a}}\displaystyle\sum_{m=0}^{[(r-1)/2]}\frac{\binom{2b-2m-2}{b-m-1}\binom{a-b+2m}{m}}{(b-m)(a-b+m+1)}, & r=1,\ 2,\ \ldots,\ 2b-1,\\ \dfrac{a-b+1}{a+1}, & r=2b,\\ 0 & \text{во всех остальных случаях.}\end{cases}$$

Следующая группа задач связана с преобразованиями пуассоновских процессов.

40. Пусть $0=\tau_0<\tau_1<\tau_2\ldots$ — последовательность времен наступления событий пуассоновского процесса с параметром λ. Предположим, что наступление событий (независимо от моментов других наступлений) фиксируется с вероятностью p и не фиксируется с вероятностью $1-p$. Показать, что $X_1(t)$, число зарегистрированных наступлений, и $X_2(t)$, число незарегистрированных наступлений, представляют собой пуассоновские процессы с параметрами λp и $\lambda(1-p)$ соответственно.

Указание: Рассмотреть производящую функцию совместных вероятностей, а именно:

$$f(z_1,\ z_2)=\mathrm{M}\left\{z_1^{X_1(t)}z_2^{X_2(t)}\right\}=$$

$$=\mathrm{M}\left[\mathrm{M}\left\{z_1^{X_1(t)}z_2^{X_2(t)}\,\middle|\,X_1(t)+X_2(t)\right\}\right]=$$

$$=\mathrm{M}\,(pz_1+qz_2)^{X_1(t)+X_2(t)}=$$

(так как при заданной сумме $X_1(t)+X_2(t)$ величина $X_1(t)$ подчиняется биномиальному распределению)

$$=e^{\lambda t\,(pz_1+qz_2-1)}=e^{\lambda pt\,(z_1-1)}e^{\lambda qt\,(z_2-1)}.$$

41 (продолжение). Пусть каждое событие, независимо от других, объявляется принадлежащим одной из k категорий с вероятностью p_i, $i=1,\ 2,\ \ldots,\ k$, $\sum_{i=1}^{k}p_i=1$, и пусть $X_i(t)$ — число событий, попавших в i-ю категорию за время t. Показать, что $X_1(t)$, $X_2(t)$, ..., $X_k(t)$ являются пуассоновскими процессами с параметрами λp_i, $i=1,\ 2,\ \ldots,\ k$, соответственно.

42. Рассмотрим пуассоновский процесс с параметром λ на положительной полуоси $(0,\infty)$. Предположим, что событие, наступившее в момент t, относится к одной из k категорий с вероятностями $p_i(t)$, $i=1,\ 2,\ \ldots,\ k$, $\sum_{i=1}^{k}p_i(t)=1$. Вероятности $p_i(t)$ будем считать непрерывными функциями времени. Пусть $X_i(t)$ обозначает число событий, наступивших за время t и отнесенных к категории i. Показать, что при каждом i процесс $\{X_i(t);\ t\geqslant 0\}$ является нестационарным пуассоновским (см. задачу 13 гл. 7).

Указание: Показать, что $\mathrm{P}\,\{$событие категории i наступает в интервале $(t,t+h)\}=\lambda p_i(t)h+o(h)$ и что событие в левой части этого равенства не зависит от значений $X_i(\tau)$ при $\tau<t$.

43 (продолжение). Показать, что процессы $\{X_i(t);\ t \geqslant 0\}$, $i = 1, 2, \ldots, k$, независимы.

Указание: Пусть (a_i, b_i), $i = 1, 2, \ldots, k$, — произвольные k интервалов на полуоси $[0, \infty)$. Положим $Y_i = X_i(b_i) - X_i(a_i)$. Показать, что $Y_1, Y_2, \ldots, Y_k$ — независимые с. в. Для этого обозначим через $N(T)$ число событий всех категорий, наступивших в интервале $(0, T)$, где $T = \max\limits_{1 \leqslant i \leqslant k} b_i$ Если значение $N(T)$ задано, то моменты наступления событий всех категорий в интервале $(0, T)$ можно представить как n независимых наблюдений случайной величины, равномерно распределенной на $(0, T)$. При условии $N(T) = n$ совокупность $\left(Y_1, Y_2, \ldots, Y_k, n - \sum\limits_{i=1}^{k} Y_i\right)$ подчиняется полиномиальному распределению с вероятностями

$$p_i = \frac{1}{T} \int\limits_{a_i}^{b_i} p_i(t)\, dt, \qquad i = 1, 2, \ldots, k,$$

$$p_{k+1} = 1 - \sum_{i=1}^{k} p_i.$$

Следовательно,

$$\mathrm{M}\left(z_1^{Y_1} z_2^{Y_2} \ldots z_k^{Y_k} \,\middle|\, N(T) = n\right) = (p_1 z_1 + p_2 z_2 + \ldots + p_k z_k + 1 - p_1 - \ldots - p_k)^n.$$

Но $N(T)$ распределена по закону Пуассона со средним λT, поэтому

$$\mathrm{M}\left(z_1^{Y_1} \ldots z_k^{Y_k}\right) = \exp\left[\mu T\left(\sum_{i=1}^{k} p_i(z_i - 1)\right)\right] = \prod_{i=1}^{k} \exp\left[\lambda_i (z_i - 1)\right],$$

где

$$\lambda_i = \lambda \int\limits_{a_i}^{b_i} p_i(t)\, dt.$$

Эти же доводы справедливы и в том случае, если каждый интервал (a_i, b_i) заменить любым конечным объединением интервалов.

44 (продолжение). Пусть $\int\limits_0^\infty p_i(t)\, dt < \infty$ при $i = 1, 2, \ldots, r$, где $r \leqslant k - 1$. Доказать, что совместное распределение набора с. в. $(X_1(t), \ldots, X_r(t))$ стремится в пределе при $t \to \infty$ к совместному распределению r независимых с. в., распределенных по закону Пуассона со средними $\lambda \int\limits_0^\infty p_i(t)\, dt < \infty$, $i = 1, 2, \ldots, r$.

45. Пусть $0 = \tau_0 < \tau_1 < \tau_2 \ldots$ — моменты наступления событий пуассоновского процесса с параметром λ. В момент $t = 0$ частица начинает совершать броуновское движение ($\sigma^2 = 1$) из начального положения τ_i. Событие, наступившее в момент τ_i, не регистрируется, если броуновская частица, отправившаяся из состояния τ_i, находится слева от $-\alpha\,(\alpha > 0)$ в момент t_0; в противном случае это событие регистрируется. Показать, что зарегистрированные и незарегистрированные события образуют пуассоновские процессы с параметрами соответственно $\lambda[1-\Phi((t+\alpha)/\sqrt{t_0})]$ и $\lambda\Phi((t+\alpha)/\sqrt{t_0})$, где Φ — функция нормального распределения.

46. Пусть $0=\tau_0<\tau_1<\ldots$ — моменты наступления событий пуассоновского процесса с параметром λ. Пусть $\varphi(x, y)$ — действительная функция, $x\geqslant 0$; рассмотрим последовательность $\tilde{\tau}_i=\varphi(\tau_i, \xi(\tilde{\tau}_i))$, где $\{\xi(t);\ t\geqslant 0\}$ — случайный процесс, не зависящий от исходного процесса, такой, что для любого n и любых $t_1, t_2, \ldots, t_n$ с. в. $\xi(t_1), \ldots, \xi(t_n)$ независимы. Положим $F(x, t)=\mathrm{P}\{\varphi(t, \xi(t))<x\}$. Для каждого $t>0$ и любых r непересекающихся интервалов (a_i, b_i), $i=1, 2, \ldots, r$, на оси $(-\infty, +\infty)$ положим $X_i(t)$ равным числу моментов на интервале $(0, t)$, таких, что $a_i<\tilde{\tau}_i<b_i$. Показать, что $\{X_i(t);\ t\geqslant 0\}$, $i=1, 2, \ldots, r$, являются независимыми неоднородными пуассоновскими процессами. Показать, что $\mathrm{M}(X_i(t))=\lambda\int\limits_0^t[F(b_i, u)-F(a_i, u)]\,du$, предполагая, что при каждом x функция $F(x, u)$ интегрируема по u на любом конечном интервале.

Указание: Воспользоваться результатами и методами доказательства задач 42 и 43.

47. В предыдущей задаче предположить, что $\int\limits_0^\infty F(x, u)\,du<\infty$ при каждом x. Показать, что функция распределения с. в. $X_i(t)$ стремится при $t\to\infty$ к функции распределения пуассоновской с. в. Показать также, что $\{N(a);\ -\infty<a<\infty\}$, где $N(a_2)-N(a_1)$ есть число величин $\tilde{\tau}_j$ в интервале (a_1, a_2), является процессом с независимыми, но не обязательно стационарными приращениями.

48. В предыдущей задаче предположить, что существует функция $h(x)$, такая, что для любых $x_1\leqslant x_2$

$$\int\limits_0^\infty F(x_2, u)\,du-\int\limits_0^\infty F(x_1, u)\,du=\int\limits_{x_1}^{x_2}h(t)\,dt.$$

Показать, что процесс $\{N(a);\ -\infty<a<\infty\}$ является пуассоновским процессом, не обязательно однородным, с параметром $\lambda h(t)$.

49. В предыдущей задаче положить $h(t)\equiv\mu$. Показать, что $\{N(a);\ -\infty<a<\infty\}$ является одномерным пространственным пуассоновским процессом с параметром $\lambda\mu$.

50. Пусть процесс $\{\xi(t); t>0\}$, определенный в задаче 46, таков, что при любом n и любых $t_1, t_2, \ldots, t_n\geqslant 0$ случайные величины $\xi(t_1), \ldots, \xi(t_n)$ положительны, независимы и обладают общим распределением $G(x)$. Получить результаты задач 46—49 для следующих частных случаев:

(а) $\varphi(x, y)=xy$;

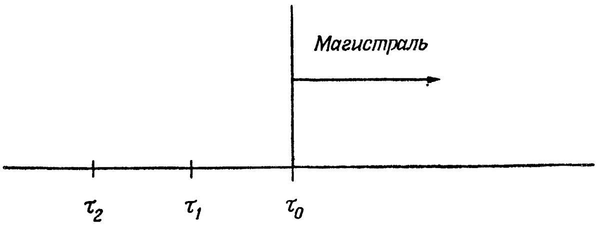
(б) $\varphi(x, y)=x+y$.

В (а) предположить, что $1/w=\int\limits_0^\infty(dG(v)/v)<\infty$.

51. Из результата задачи 50 (б) вывести, что для процесса обслуживания $(M/G/\infty)$ (см. задачу 10 гл. 14) со стационарным пуассоновским входящим потоком и произвольным распределением времени обслуживания $G(x)$ выходной поток $\{U(t), t \geqslant 0\}$ (т. е. число обслуженных заявок к моменту t) является нестационарным пуассоновским потоком при условии, что в момент $t=0$ система свободна от заявок. Параметром этого процесса служит $\lambda G(t)$.

52. Предположим, что в моменты $0=\tau_0<\tau_1<\tau_2<\ldots$ промежутки между которыми распределены по закону Пуассона с параметром λ, на магистраль с одного ее конца въезжают автомобили и движутся в положительном направлении. Автомобиль, въезжающий на магистраль в момент τ_k, приобретает скорость v_k и сохраняет эту скорость неизменной. Предположим, что скорости v_k являются независимыми с. в. с общим распределением $F(x)$. Показать, что в стационарном случае, т. е. по прошествии бесконечного времени, пространственное распределение автомобилей на магистрали будет описываться однородным пуассоновским процессом при условии, что $1/w=\int\limits_0^\infty dF(v)/v<\infty$. Найти параметр процесса.

Указание: Пусть (a_1, b_1) и (a_2, b_2) — два непересекающихся интервала на полуоси $(0, \infty)$. Представим себе ось времени, простирающуюся от настоящего момента времени к бесконечному прошлому, и будем теперь отмерять положение автомобилей от въезда на магистраль. Моменты въезда автомобилей на магистраль образуют пуассоновский процесс на этой временной оси. Будем относить τ_i к категории 1, если $a_1 \leqslant \tau_1 v_1 \leqslant b_1$, к категории 2, если $a_2 \leqslant \tau_i v_i \leqslant b_2$, и к категории 3 во всех остальных случаях. Тогда, в обозначениях задачи 43, имеем



$$p_1(t)=F(b_1/t)-F(a_1/t)$$
$$p_2(t)=F(b_2/t)-F(a_2/t)$$
$$p_3(t)=1-p_1(t)-p_2(t)$$

Воспользоваться теперь результатами задач 42—44. Отметим, что $\int\limits_0^\infty F(x/t)\,dt=x/w$, где $w^{-1}=\int\limits_0^\infty dF(v)/v$. Подобные же доводы применимы и для случая любого конечного числа неперекрывающихся интервалов.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, издательство «Мир».