БАКАЛАВР. ПРИКЛАДНОЙ КУРС

Н. Ю. Энатская

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА И СЛУЧАЙНЫЕ ПРОЦЕССЫ

УМО ВО РЕКОМЕНДУЕТ

ЮРАЙТ
издательство
biblio-online.ru

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА И СЛУЧАЙНЫЕ ПРОЦЕССЫ

ЭНАТСКАЯ Наталия Юрьевна

кандидат физико-математических наук, доцент Департамента прикладной математики Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики».

«Следует отметить удачный и логичный подбор тем и методическую проработку материала»

Г. И. Ивченко, доктор физико-математических наук, профессор общеинститутской кафедры высшей математики Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики»

«Широта охвата материала сочетается с доступностью»

В. Ф. Колчин, доктор физико-математических наук, профессор, ведущий научный сотрудник Математического института имени В. А. Стеклова Российской академии наук

В данном учебном пособии обсуждаются основные понятия и вопросы математической статистики, реально укладывающиеся в ограниченный учебным временем одноименный курс с выделением основных задач статистики: непараметрическая, параметрическая и проверка статистических гипотез. Цель пособия – дать представление об основных понятиях, структуре курса математической статистики и методах решения ее основных задач. Для иллюстрации материала приведено много решенных примеров и задач для самостоятельного решения. По характеру и объему изложения пособие предназначено для использования в учебном процессе бакалавриата.

В обучающий комплекс также входит учебное пособие «Теория вероятностей»:

ISBN 978-5-9916-9808-5

9 785991 698085

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ

Н. Ю. Энатская

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА И СЛУЧАЙНЫЕ ПРОЦЕССЫ

УЧЕБНОЕ ПОСОБИЕ
ДЛЯ ПРИКЛАДНОГО БАКАЛАВРИАТА

Рекомендовано Учебно-методическим отделом высшего образования в качестве учебного пособия для студентов высших учебных заведений, обучающихся по инженерно-техническим направлениям

Москва ▪ Юрайт ▪ 2017

УДК 519.2(075.8)
ББК 22.171я73
Э61

***Автор*:**

Энатская Наталия Юрьевна — кандидат физико-математических наук, доцент Департамента прикладной математики Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики».

***Рецензенты*:**

Ивченко Г. И. — доктор физико-математических наук, профессор кафедры высшей математики Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики»;

Колчин В. Ф. — доктор физико-математических наук, профессор, ведущий научный сотрудник Математического института им. В. А. Стеклова Российской академии наук.

Энатская, Н. Ю.

Э61 Математическая статистика и случайные процессы : учеб. пособие для прикладного бакалавриата / Н. Ю. Энатская. — М. : Издательство Юрайт, 2017. — 201 с. — Серия : Бакалавр. Прикладной курс.

ISBN 978-5-9916-9808-5

В данном учебном пособии обсуждаются основные понятия и вопросы математической статистики, реально укладывающиеся в ограниченный учебным временем одноименный курс с выделением основных задач математической статистики (непараметрическая, параметрическая и проверка статистических гипотез) и случайных процессов (теория марковских цепей, классификация ее состояний, ветвящиеся и пуассоновские процессы, числовые характеристики и линейные преобразования случайных процессов).

Цель пособия — дать представление об основных понятиях, структуре курсов математической статистики и случайных процессов и методах решения их основных задач. Для иллюстрации материала приведено много решенных примеров и задач для самостоятельного решения. По характеру и объему изложения учебник предназначен для использования в учебном процессе бакалавриата.

Соответствует актуальным требованиям Федерального государственного образовательного стандарта высшего образования.

Для студентов инженерно-технических специальностей.

УДК 519.2(075.8)
ББК 22.171я73

ISBN 978-5-9916-9808-5

Оглавление

Предисловие

Математическая статистика занимается разработкой и проведением научно обоснованных обработок наблюдений над случайной величиной с целью получения информации о ее вероятностном распределении.

Предметом изучения случайных процессов является изменение во времени случайных систем с определенными вероятностями поведения. Случайный процесс задается случайными величинами, меняющимися во времени.

Изучение этих вероятностных направлений базируется на знании общего курса теории вероятностей, которая вместе с математической статистикой и случайными процессами представляют базовые дисциплины вузовского вероятностного образования студентов инженерно-технических направлений.

В данном учебном пособии обсуждаются основные понятия и вопросы математической статистики и случайных процессов, реально укладывающиеся в ограниченные учебным временем одноименные курсы с выделением основных задач математической статистики (непараметрическая, параметрическая и проверка статистических гипотез) и случайных процессов (марковские цепи с дискретным временем, классификация их состояний, марковские цепи с непрерывным временем и конечным множеством состояний, широко используемые в теории надежности и теории массового обслуживания, числовые характеристики и линейные преобразования случайных процессов).

Цель пособия — дать представление об основных понятиях, структуре курсов математической статистики и случайных процессов и методах решения их основных задач. Для иллюстрации материала приведено много решенных примеров и задач для самостоятельного решения. По характеру и объему изложения учебник предназначен для использования в учебном процессе бакалавриата.

В результате изучения курса студенты должны:

знать

• основные понятия и основные задачи математической статистики: непараметрическую, параметрическую и проверку статистических гипотез;

• основы теории случайных процессов, цепи Маркова, пуассоновские процессы, ветвящиеся процессы; числовые характеристики случайных процессов;

уметь

• решать задачи математической статистики и теории случайных процессов;

владеть

• методами для решения конкретных задач математической статистики и случайных процессов.

Основные обозначения

СВ — случайная величина;
ч.т.д. — что и требовалось доказать;
нз — независимые;
нк — некоррелированные;
MX или EX — математическое ожидание СВ X;
DX — дисперсия СВ X;
K_{XY} — корреляция СВ X и Y;
r_{XY} — коэффициент корреляции СВ X и Y;
СВ $X \sim L(a)$ — означает, что СВ X имеет закон распределения $L(a)$, где $a = (a_1, ..., a_k)$ — k-мерный параметр распределения;
СВ $X \sim B(1, p)$ — бернуллиевское распределение;
СВ $X \sim B(n, p)$ — биномиальное распределение;
СВ $X \sim \pi(\lambda)$ — пуассоновское распределение;
СВ $X_i \sim G(p)$ — геометрическое распределение;
СВ $Y_i \sim$ сд$G(p)$ — сдвинутое геометрическое распределение;
СВ $Z_1 \sim$ Па(r, p) — распределение Паскаля;
СВ $Z_2 \sim OB(r, p)$ — отрицательное биномиальное распределение;
СВ $X \sim H(N, M, n)$ — гипергеометрическое распределение;
СВ $X \sim R[a; b]$ — равномерное распределение на отрезке $[a; b]$;
СВ $X_0 \sim R[0; 1]$ — равномерное распределение на отрезке $[0; 1]$;
СВ $X \sim N(a, \sigma)$ — нормальное распределение с параметрами (a, σ);
СВ $X_0 \sim N(0, 1)$ — нормальное распределение с параметрами $(0, 1)$;
СВ $X \sim \Gamma_{\alpha, \lambda}$ — гамма-распределение;
СВ $X \sim E(\lambda)$ — экспоненциальное распределение.

Раздел I

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

В результате изучения данного раздела студенты должны:

знать

- непараметрическую задачу статистики;
- параметрическую задачу статистики;
- основы проверки статистических гипотез;

уметь

- решать задачи математической статистики;

владеть

- методами математической статистики.

Глава 1
ОСНОВНЫЕ ПОНЯТИЯ СТАТИСТИКИ И НЕПАРАМЕТРИЧЕСКАЯ ЗАДАЧА

В результате изучения данной главы студенты должны:

знать

• основные понятия математической статистики;

• приемы первичной обработки наблюдений;

уметь

• решать задачи анализа известных вероятностных распределений и непараметрическую задачу математической статистики;

владеть

• методами моделирования значений случайных величин;

• методами решения непараметрической задачи статистики.

В данной главе обсуждаются некоторые начальные понятия математической статистики, которые используются в дальнейшем при рассмотрении в других разделах курса, такие, например, как порядковые статистики. Вопросы моделирования, обсуждаемые здесь, важны для создания статистического материала при изучении закономерностей и характеристик случайных процессов. Также рассматриваются приемы первоначальной обработки наблюдений над случайной величиной с целью распознавания вида ее вероятностного распределения.

1.1. Основные понятия математической статистики

Математическая статистика занимается разработкой, научным обоснованием и применением методов обработки данных наблюдений с целью получения информации о распределении изучаемой случайной величины.

Пусть X — наблюдаемая случайная величина (СВ).

Определение 1.1. Все значения, которые может принимать СВ X, называют ***выборочным пространством*** или ***генеральной совокупностью***; n возможных значений СВ X называются ***выборкой***

(n — объем выборки): $\vec{x} = (x_1, x_2, x_3, ..., x_n)$. Конкретные n наблюденных значений СВ X называются ***реализацией выборки***.

Для того чтобы изучение СВ по выборке имело смысл, необходимо, чтобы выборка не искажала, а отражала свойства генеральной совокупности. Такая выборка называется ***представительной*** (репрезентативной). Математически это означает, что элементы выборки должны быть *независимыми, одинаково распределенными СВ с тем же законом распределения*, что и у изучаемой СВ X.

Определение 1.2. Если упорядочить значения выборки по величине $X_{(1)} \leqslant X_{(2)} \leqslant ... \leqslant X_{(n)}$, то ряд $X_{(1)}, X_{(2)}, ..., X_{(n)}$ называется ***вариационным рядом*** или ***рядом порядковых статистик***, а элементы ряда называются ***вариантами*** или ***порядковыми статистиками***.

Элементы вариационного ряда не обладают теми же свойствами, которыми обладала выборка, т.е. они не являются независимыми, одинаково распределенными СВ и каждый имеет свой закон распределения, не совпадающий с законом распределения исходной СВ X (доказательство см. ниже).

Проиллюстрируем это на примере крайних порядковых статистик СВ X с функцией распределения $F(x)$: $x_{(1)} = \min(x_1, ..., x_n) = U$ и $x_{(n)} = \max(x_1, ..., x_n) = V$. Зависимость их очевидна.

Найдем закон распределения для СВ U и V. Обозначим $G_U(x)$ — функция распределения СВ U и $G_V(x)$ — функция распределения СВ V. Имеем: $G_U(x) = P\{U < x\} = P\{\min(x_1, ... , x_n) < x\}$ = (сформулируем эквивалентное событие в терминах элементов выборки, переходя к дополнительному событию и учитывая свойства элементов выборки) $= 1 - P\{x_i \geqslant x\} = 1 - \prod\limits_{i=1}^{n} P\{x_i \geqslant x\} = 1 - (1 - F(x))^n$.

Плотность распределения СВ U есть

$$g_U(x) = G_U'(x) = n(1 - F(x))^{n-1} f(x).$$

Аналогичными рассуждениями найдем закон распределения СВ V:

$$G_V(x) = P\{x_{(n)} < x\} = P\{x_1 < x, ..., x_n < x\} = \prod_{i=1}^{n} F(x).$$

Тогда плотность распределения СВ V есть

$$g_V(x) = G_V'(x) = nF(x)^{n-1} f(x).$$

Определение 1.3. ***Статистика*** — это любая функция от наблюдений, не зависящая от неизвестных параметров распределения.

Определение 1.4. ***Семейство распределений*** $F_\theta(x)$ — это множество распределений одного аналитического вида со всеми возможными значениями параметра $\theta = (\theta_1, ..., \theta_k)$.

Основные задачи математической статистики

1. Непараметрическая задача статистики.

Это первоначальная (грубая) обработка данных следующего вида: наблюдаемая СВ X и реализация выборки $\vec{x} = (x_1, \dots, x_n)$ с целью получения информации о виде распределения СВ X, т.е. нахождения семейства распределений, к которому оно относится. В результате решения задачи получается, что $L(X) \in F_\theta(x)$, где $L(X)$ — распределение СВ X, а $\theta = (\theta_1, ..., \theta_k)$ — неизвестный параметр распределения СВ X.

2. Параметрическая задача статистики.

Эта задача решается после получения решения непараметрической задачи (когда семейство распределений $F_\theta(x)$ для СВ X уже определено) и состоит в уточнении значения параметра θ или функции $\tau(\theta)$ от этого параметра в общем случае, где $\theta = (\theta_1, ..., \theta_k)$, т.е. в построении оценки $t(\vec{x}) = \hat{\tau}$ функции $\tau(\theta) = \tau$.

При решении этой задачи различают два подхода: точечное и доверительное оценивание.

А. Точечное оценивание.

Суть точечного оценивания в том, что для $\tau(\theta)$ строится одна статистика $t(x) = \hat{\tau}$, которая принимается за оценку $\tau(\theta)$, т.е. $t(x) = \hat{\tau}$.

«Хорошей» оценкой является такая оценка, которая наиболее близка к истинному значению $\tau(\theta)$, т.е. когда ее значения в каком-то смысле сконцентрированы вокруг истинного значения $\tau(\theta)$.

Математически это означает желательность следующих свойств оценки:

а) несмещенность: $Mt(x) = \tau(\theta)$, где $Mt(x)$ — среднее значение СВ X (при малых выборках это требование очень важно, при больших n достаточно выполнения свойства асимптотической несмещенности: $Mt(x) \neq \theta$, но $Mt(x) \to \theta$ при $n \to \infty$);

б) состоятельность: $t(x) \underset{n\to\infty}{\to} \tau(\theta)$, что означает

$$P\{|t(x) - \tau(\theta)| > \varepsilon \underset{n\to\infty}{\to} 0;$$

в) эффективность, которая задается только для несмещенных оценок и определяется дисперсией $Dt(x)$. Тогда значение дисперсии несмещенной оценки называется ее *эффективностью*, которая используется при сравнении качества несмещенных оценок: та оценка лучше, у которой дисперсия меньше, т.е. эффективность больше.

Определение 1.5. Оценка называется ***оптимальной***, если она несмещенная и имеет минимально возможную дисперсию.

Б. Доверительное оценивание.

Ограничимся сначала рассмотрением этого подхода в случае одномерного параметра. Тогда решение состоит в построении двух статистик: $t_1 = t_1(\vec{x})$ и $t_2 = t_2(\vec{x})$ таким образом, чтобы $P\{t_1(\vec{x}) < \tau(\theta) < t_2(\vec{x})\} = \gamma$, где γ — некоторое значение, называемое *доверительной вероятностью,* а интервал $I = (t_1(\vec{x}), t_2(\vec{x}))$ называется *доверительным интервалом*. Тогда γ — надежность оценивания, а длина интервала $(t_1(\vec{x}), t_2(\vec{x}))$ определяет точность оценивания, поэтому при заданном γ этот интервал желательно строить более коротким.

В случае большей размерности неизвестного параметра аналогично определяется доверительная область, содержащая истинное значение $\tau(\theta)$.

3. Проверка статистических гипотез.

Это проверка предположений (гипотез) о распределении изучаемой СВ. Если гипотеза H_0 описывает одно распределение, то она называется *простой*, а если несколько — то *сложной*. Та гипотеза H_0, которую нужно проверить, называется *нулевой* или *основной*; H_1 — *альтернативная* гипотеза.

Чтобы проверить гипотезу, нужно построить критерий, т.е. правило, по которому нулевая гипотеза принимается или отвергается. Это значит, что выборочное пространство нужно разделить границей на две зоны (рис. 1.1): T — область принятия гипотезы (основной гипотезы H_0) и R — критическая область, где она отвергается.

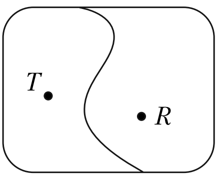


Рис. 1.1. **Зоны выборочного пространства**

Пусть X — изучаемая СВ, $x_1, ..., x_n$ — выборка ее наблюдаемых значений. Если точка попадает в зону T, то гипотеза принимается, если точка попадает в зону R, то гипотеза отвергается.

Как выбрать границу наилучшим образом?

Это делают на основе анализа возможных ошибок. При решении могут возникнуть ошибки двух родов.

Пусть α — вероятность ошибки первого рода (это вероятность того, что отвергли верную гипотезу (H_0): $\alpha = P\{x \in R/H_0\}$); β — вероятность ошибки второго рода (это вероятность того, что приняли неверную гипотезу (H_1): $\beta = P\{x \in R/H_1\}$). Величина α называется еще уровнем значимости или размером критерия.

Оказывается, что одновременно вероятности ошибок первого и второго рода (α и β) невозможно сделать сколько угодно маленькими. Поэтому два критерия, если они сравнимы, нужно сравнивать по качеству, определяемому вероятностями ошибок α и β следующим образом: фиксируется вероятность ошибки первого рода α для первого и второго критериев, по этому значению α строится граница и вычисляется вероятность ошибки второго рода β. Тогда тот критерий лучше, у которого меньше вероятность ошибки второго рода β. Говорят, что такой критерий более мощный: $W = 1 - \beta$ — мощность критерия.

Приведем пример, поясняющий смысл ошибок первого и второго рода.

Пример 1.1. Пусть проверяют партии деталей, упакованных по 1000 шт. Партию считают хорошей, если при контроле 20 наугад выбранных из нее деталей бракуют $X \leqslant 1$ шт. Пусть гипотеза H_0 состоит в том, что партия хорошая, а гипотеза H_1 — партия плохая. Тогда $\alpha = P(X > 1/H_0)$ — риск заказчика; $\beta = P(X \leqslant 1/H_1)$ — риск изготовителя.

1.2. Порядковые статистики

Порядковые статистики (их определение см. в параграфе 1.1) широко используются при решении многих задач статистики, поэтому должны быть заранее изучены в первую очередь по следующим направлениям:

- установление связи распределения изучаемой СВ X с законами распределений порядковых статистик разных размерностей;
- нахождение выражений моментов порядковых статистик через закон распределения СВ X;
- изучение закона распределения размаха выборки ($W = X_{(n)} - X_{(1)}$).

Все полученные формулы по указанным направлениям ниже проиллюстрированы на примере равномерного распределения.

Основные формулы для порядковых статистик

Через $F(x)$ и $f(x)$ будем обозначать соответственно функцию распределения и плотность распределения изучаемой СВ X, а через $F_{(k)}(x)$ и $f_{(k)}(x)$ — соответственно функцию распределения и плотность распределения k-й порядковой статистики.

1. Законы распределения крайних порядковых статистик $\boldsymbol{X_{(1)}}$, $\boldsymbol{X_{(n)}}$.

а) Для $X_{(1)}$ имеем

$$F_{(1)}(x) = P(X_{(1)} < x) = 1 - P(x_{(1)} \geqslant x) = 1 - P(x_1 \geqslant x, ..., x_n \geqslant x) =$$

$$= 1 - \prod_{i=1}^{n} P(x_i \geqslant x) = 1 - (P(x_1 \geqslant x))^n = 1 - (1 - F(x))^n ;$$

$$f_{(1)}(x) = F'_{(1)}(x), \text{ т.е. } f_{(1)}(x) = n(1 - F(x))^{n-1} f(x).$$

б) Для $X_{(n)}$ имеем

$$F_{(n)}(x) = P(X_{(n)} < x) = P(x_1 < x, ..., x_n < x) = \prod_{i=1}^{n} P(x_i < x) = F^n(x);$$

$$f_{(n)}(x) = nF^{n-1}(x)f(x).$$

2. Закон распределения *k*-й порядковой статистики $X_{(k)}$.

Пусть $f_{(k)}(x)$ — плотность распределения k-й порядковой статистики. Тогда $f_{(k)}(x)dx$ — элемент вероятности для k-й порядковой статистики, т.е. $f_{(k)}(x)dx = P(X_{(k)} \in (x; x + dx)) = P(A)$, где $A = \{X_{(k)} \in (x, x + dx)\}$.

Проинтерпретируем событие A в терминах элементов выборки в соответствии с рис. 1.2.

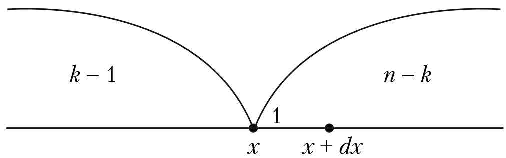


Рис. 1.2. **Интерпретация события $A = \{X_{(k)} \in (x, x + dx)\}$**

Событие A состоит в одновременном появлении следующих трех событий:

	Вероятность	Число вариантов
1. Какое-то наблюдение $\in (x; x + dx)$	$f(x)dx$	n
2. Какие-то $(k - 1)$ остальных наблюдений левее x	$F^{k-1}(x)$	C_{n-1}^{k-1}
3. Остальные $(n - k)$ наблюдений правее x	$(1 - F(x))^{n-k}$	1

Тогда

$$f_{(k)}(x)dx = P(x_{(k)} \in (x; x + dx)) = nC_{n-1}^{k-1}f(x)dxF^{k-1}(x)(1 - F(x))^{n-k}.$$

Отсюда плотность распределения k-й порядковой статистики

$$f_{(k)}(x) = nC_{n-1}^{k-1}f(x)F^{k-1}(x)(1 - F(x))^{n-k}.$$

Замечание 1.1. Если в эту формулу подставить $k = 1$, $k = n$, то получим законы распределения крайних порядковых статистик, найденные в п. 1.

3. Закон совместного распределения *k*-й и *l*-й порядковых статистик $(k < l)$.

Имеем $f_{(kl)}(u, v)dudv = P(X_{(k)} \in (u, u + du), X(l) \in (v, v + dv))$ — элемент вероятности СВ $(X_{(k)}, X_{(l)})$.

Пусть событие $A = \{X_{(k)} \in (u, u + du), X_{(l)} \in (v, v + dv)\}$. Переформулируем событие A в терминах элементов выборки в соответствии с рис. 1.3.

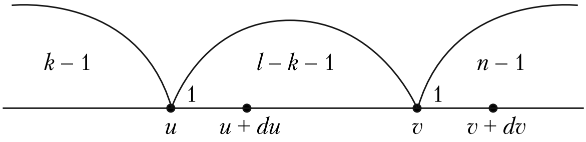


Рис. 1.3. **К закону совместного распределения *k*-й и *l*-й порядковых статистик**

Событие A состоит в одновременном появлении следующих событий:

	Вероятность	Число вариантов
1. Какое-то одно наблюдение $\in (u; u + du)$	$f(u)du$	N
2. Какое-то одно C_{n-2}^{k-1} наблюдение $\in (v; v + dv)$	$f(v)dv$	$n - 1$
3. Какие-то $(k - 1)$ из оставшихся наблюдений левее u	$F^{k-1}(u)$	C_{n-2}^{k-1}
4. Какие-то $(l - k - 1)$ из оставшихся наблюдений $\in (u; v)$	$(F(v) - F(u))^{l-k-1}$	C_{n-k-1}^{l-k-1}
5. Оставшиеся $(n - 1)$ наблюдений правее v	$(1 - F(v))^{n-1}$	1

Тогда

$$f_{(ki)}(u, v) = n(n-1)C_{n-k-1}^{l-k-1}f(u)f(v)F^{k-1}(u)(F(v) - F(u))^{l-k-1}(1 - F(v))^{n-1}.$$

Замечание 1.2. Подставляя $k = 1$ и $l = n$, получим закон распределения крайних порядковых статистик:

$$f_{(1n)}(u, v) = n(n-1)f(u)f(v)(F(v) - F(u))^{n-2}.$$

4. Закон совместного распределения *k*-й, *l*-й и *m*-й порядковых статистик ($k < l < m$).

Имеем $f_{(klm)}(u, v, w)dudvdw = P(X_{(k)} \in (u; u + du), X_{(l)} \in (v; v + dv), X_{(m)} \in (w; w + dw))$.

Пусть событие $A = (X_{(k)} \in (u; u + du), X_{(l)} \in (v; v + dv), X_{(m)} \in (w; w + dw))$.

Переформулируем событие A в терминах элемента выборки в соответствии с рис. 1.4.

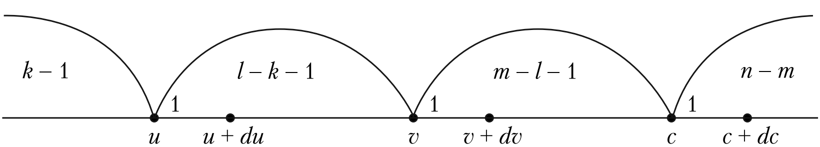


Рис. 1.4. **К закону совместного распределения *k*-й, *l*-й и *m*-й порядковых статистик**

Событие A состоит в одновременном появлении следующих событий:

	Вероятность	Число вариантов
1. Какое-то одно наблюдение $\in (u; u + du)$	$f(u)du$	n
2. Какое-то одно наблюдение $\in (v; v + dv)$	$f(v)dv$	$n - 1$
3. Какое-то одно наблюдение $\in (w; w + dw)$	$f(w)dw$	$n - 2$
4. Какие-то $(k - 1)$ из оставшихся левее u	$F^{k-1}(u)$	C_{n-3}^{k-1}
5. Какие-то $(l - k - 1)$ из оставшихся $\in (u; v)$	$(F(v) - F(u))^{l-k-1}$	C_{n-k-2}^{l-k-1}
6. Какие-то $(m - l - 1)$ из оставшихся $\in (v; w)$	$(F(w) - F(v))^{m-l-1}$	C_{n-l-1}^{m-l-1}
7. Оставшиеся $(n - m)$ правее w	$(1 - F(w))^{n-m}$	1

Тогда

$$f_{(klm)}(u, v, w) = n(n-1)(n-2)C_{n-3}^{k-1}C_{n-k-2}^{l-k-1}C_{n-l-1}^{m-l-1}f(u)f(v)f(w)F^{k-1}(u) \times$$
$$\times (F(v) - F(u))^{l-k-1}(F(w) - F(v))^{n-l-1}(1 - F(w))^{n-m}.$$

5. Совместное распределение первых r порядковых статистик ($r \leqslant n$).

Имеем $f_{(12...r)}(u_1, u_2, ..., u_r) = P(X_{(1)} \in (u_1, u_1 + du_1), X_{(2)} \in (u_2, u_2 + du_2), ..., X_{(r)} \in (u_r, u_r + du_r)) = P(A)$, где событие A в соответствии с рис. 1.5 состоит в одновременном появлении следующих событий:

	Вероятность	Число вариантов
1. Какое-то одно наблюдение $\in (u_1; u_1 + du_1)$	$f(u_1)du_1$	n
2. Какое-то одно из остальных наблюдений $\in (u_2; u_2 + du_2)$	$f(u_2)du_2$	$n - 1$
...	...	...
r. Какое-то одно из остальных наблюдений $\in (u_r; u_r + du_r)$	$f(u_r)du_r$	$n - r + 1$
$r + 1$. Остальные $(n - r)$ наблюдений правее u_r	$(1 - F(u_r))^{n-r}$	1

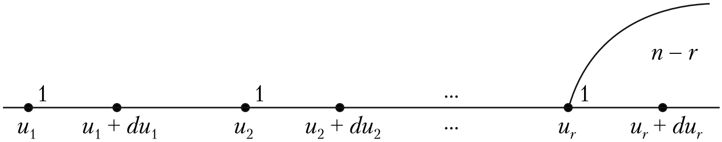


Рис. 1.5. **К совместному распределению первых *r* порядковых статистик**

Тогда

$$f_{(1\ldots r)}(u_1, \ldots, u_r) = n(n-1)(n-2)\ldots(n-r+1)f_{(u_1)}\ldots f_{(u_r)}(1-F(u_r))^{n-r}$$

— плотность распределения первых r порядковых статистик.

6. Размах выборки.

Назовем $W = X_{(n)} - X_{(1)}$ размахом выборки. Найдем закон распределения размаха выборки $H(t)$:

$$H(t) = P(W < t) = \iint\limits_{0 \leqslant v-u \leqslant t} f_{(1n)}(u,v)dudv =$$

$$= n(n-1)\int\limits_{-\infty}^{\infty} f(u)du \int\limits_{u}^{u+t} (F(v)-F(u))^{n-2} f(v)dv =$$

$$= n(n-1)\int\limits_{-\infty}^{\infty} f(u)\frac{(F(v)-F(u))^{n-1}}{n-1}\Bigg|_{u}^{u+t} du = n\int\limits_{-\infty}^{\infty} f(u)(F(u+t)-F(u))^{n-1}du.$$

7. Моменты порядковых статистик.

Имеем следующие формулы:

$$MX_{(k)} = \int\limits_{-\infty}^{\infty} xf_{(k)}(x)dx;\ MX_{(k)}^2 = \int\limits_{-\infty}^{\infty} x^2 f_{(k)}(x)dx; DX_{(k)} = MX_{(k)}^2 - (MX_{(k)})^2;$$

$$K_{X(k)X(l)} = MX_{(k)}X_{(l)} - MX_{(k)}MX_{(l)};\ MX_{(k)}X_{(l)} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} uvf_{kl}(u,v)dudv.$$

Порядковые статистики равномерного распределения

Пусть СВ $X \sim R[0; 1]$; СВ $Y \sim R[a; b]$, тогда $Y = (b-a)X + a$.

1. Связи распределений и моментов случайных величин $X_{(k)}$ и $Y_{(k)}$.

Очевидно, что $Y_{(k)} = (b-a)X_{(k)} + a$. Тогда

$$F_{Y_{(k)}}(x) = P\{Y_{(k)} < x\} = P\left\{X_{(k)} < \frac{x-a}{x-b}\right\} = F_{X_{(k)}}\left(\frac{x-a}{b-a}\right) \Rightarrow$$

$$\Rightarrow f_{Y_{(k)}}(x) = \frac{1}{b-a} f_{X_{(k)}}\left(\frac{x-a}{b-a}\right);$$

$$MY_{(k)} = (b-a)MX_{(k)} + a;\quad DY_{(k)} = (b-a)^2 DX_{(k)};$$

$$K_{Y_{(k)}Y_{(l)}} = (b-a)^2 K_{X_{(k)}X_{(l)}},\ 1 \leqslant k < l \leqslant n;$$

$$W_X = X_{(n)} - X_{(1)};\ \ W_Y = Y_{(n)} - Y_{(1)} = (b-a)(X_{(n)} - X_{(1)}) = (b-a)W_X;$$

$$MW_Y = (b-a)MW_X;\quad DW_Y = (b-a)^2 DW_X.$$

Замечание 1.3. На основании п. 1 достаточно изучать порядковые статистики случайной величины $X_{(k)}$.

2. Вычисление моментов порядковых статистик равномерного распределения.

Из результатов подпараграфа 1.2.1 для закона распределения порядковых статистик для СВ $X \sim R[0;\ 1]$ получаем следующие формулы:

$$f_k(x) = f_{X(k)}(x) = \begin{cases} 0, \ x \notin [0;1], \\ nC_{n-1}^{k-1} x^{k-1}(1-x)^{n-k}, x \in [0;1]; \end{cases}$$

$$F_k(x) = F_{X(k)}(x) = \begin{cases} 0, \ x < 0, \\ nC_{n-1}^{k-1} \int\limits_0^x t^{k-1}(1-t)^{n-k}\,dt, x \in [0;1], \\ 1, \ x > 1. \end{cases}$$

Далее для вычисления моментов используются гамма- и бета-функции, определенные в книге ранее:

$$MX_{(k)} = \int\limits_0^1 x f_k(x)dx = nC_{n-1}^{k-1} \int\limits_0^1 x^{k-1}(1-x)^{n-k}\,dx = nC_{n-1}^{k-1}\mathrm{B}(k+1, n-k+1) =$$

$$= nC_{n-1}^{k-1} \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} = \frac{n!}{(k-1)!(n-k)!} \frac{k!(n-k)!}{(n+1)!} = \frac{k}{n+1} \Rightarrow$$

$$\Rightarrow MX_{(1)} = \frac{1}{n+1}; \quad MX_{(n)} = \frac{n}{n+1} \Rightarrow$$

$$\Rightarrow MY_{(1)} = a + \frac{b-a}{n+1}; \quad MY_{(n)} = \frac{n(b-a)}{n+1} + a = b - \frac{b-a}{n+1};$$

$$DX_{(k)} = MX_{(k)}^2 - (MX_{(k)})^2;$$

$$MX_{(k)}^2 = \int\limits_0^1 x^2 f_k(x)dx = nC_{n-1}^{k-1} \int\limits_0^1 x^{k+1}(1-x)^{n-k}\,dx = nC_{n-1}^{k-1}\mathrm{B}(k+2, n-k+1) =$$

$$= nC_{n-1}^{k-1} \frac{\Gamma(k+2)\Gamma(n-k+1)}{\Gamma(n+3)} = \frac{n!}{(k-1)!(n-k)!} \frac{(k+1)!(n-k)!}{(n+2)!} =$$

$$= \frac{k(k+1)}{(n+1)(n+2)};$$

$$DX_{(k)} = \frac{k(k+1)}{(n+1)(n+2)} - \frac{k^2}{(n+1)^2} = \frac{k(n-k+1)}{(n+1)^2(n+2)} \Rightarrow$$

$$\Rightarrow DX_{(1)} = \frac{n}{(n+1)^2(n+2)} = DX_{(n)};$$

$$K_{X_{(1)}X_{(n)}} = MX_{(1)}X_{(n)} - MX_{(1)}MX_{(n)}; \quad MX_{(1)}X_{(n)} = \iint uv f_{1n}(u,v)\,du\,dv,$$

где $f_{1n} = \begin{cases} n(n-1)f(y)(F(y)-F(x))^{n-2}, x, y \in [0; 1], \\ 0, \text{ в противном случае} \end{cases} =$

$$= \begin{cases} n(n-1)(y-x)^{n-2}, x, y \in [0; 1], \\ 0, \text{ в противном случае.} \end{cases}$$

Далее,

$$MX_{(1)}X_{(n)} = n(n-1)\int_0^1\int_u^1 uv(v-u)^{n-2}\,du\,dv = n(n-1)\int_0^1 u\left(\int_u^1 v(v-u)^{n-2}\,dv\right)du;$$

$$I = \int_u^1 v(v-u)^{n-2}\,dv = \int_u^1 v(v-u)^{n-1}\,dv + u\int_u^1 (v-u)^{n-2}\,dv =$$

$$= \left.\frac{(v-u)^n}{n}\right|_u^1 + u\left.\frac{(v-u)^{n-1}}{n-1}\right|_u^1 = \frac{(1-u)^n}{n} + \frac{u(1-u)^{n-1}}{n-1} \Rightarrow$$

$$\Rightarrow MX_{(1)}X_{(n)} = (n-1)\int_0^1 u(1-u)^{n-1}\,du + n\int_0^1 u^2(1-u)^{n-1}\,du =$$

$$= (n-1)\mathrm{B}(2, n+1) + n\mathrm{B}(3, n) = (n-1)\frac{\Gamma(2)\Gamma(n+1)}{\Gamma(3+n)} + n\frac{\Gamma(3)\Gamma(n)}{\Gamma(n+3)} = \frac{1}{n+2} \Rightarrow$$

$$\Rightarrow K_{X_{(1)}X_{(n)}} = \frac{1}{n+2} - \left(\frac{n}{n+1}\right)^2 = \frac{1}{(n+1)^2(n+2)}.$$

Без доказательства приведем формулу для $K_{X_{(k)}X_{(l)}}$:

$$K_{X_{(k)}X_{(l)}} = \frac{k(n-l+1)}{(n+1)^2(n+2)}.$$

Моменты для $Y_{(k)}$ найдите самостоятельно.

3. Моменты размаха выборки из равномерного распределения.

$$W_X = X_{(n)} - X_{(1)};$$

$$MW_X = MX_{(n)} - MX_{(1)} = \frac{n}{n+1} - \frac{1}{n+1} = \frac{n-1}{n+1};$$

$$DW_X = MW_X^2 - (MW_X)^2,$$

где $MW_X^2 = M(X_{(n)} - X_{(1)})^2 = MX_{(n)}^2 + MX_{(1)}^2 - 2MX_{(1)}X_{(n)}$;

$$MX_{(1)}^2 = \frac{2}{(n+1)(n+2)}; MX_{(n)}^2 = \frac{n(n+1)}{(n+1)(n+2)}; MX_{(1)}X_{(n)} = \frac{1}{n+2} \Rightarrow$$

$$\Rightarrow DW_X = \frac{2+n(n+1)}{(n+1)(n+2)} - \frac{2}{n+2} - \left(\frac{n-1}{n+1}\right)^2 = \frac{2(n-1)}{(n+1)^2(n+2)}.$$

Можно вычислить DW_X по-другому:

$$DW_X = D(X_{(n)} - X_{(1)}) = DX_{(1)} + DX_{(n)} - 2K_{X_{(k)}X_{(l)}} =$$
$$= \frac{2n}{(n+1)^2(n+2)} - \frac{2}{(n+1)^2(n+2)} = \frac{2(n-1)}{(n+1)^2(n+2)}.$$

4. Распределение размаха выборки из равномерного распределения.

$$H_X(t) = P(W_X < t) = n\int_{-\infty}^{\infty}(F(u+t) - F(u))^{n-1} f(u)du.$$

Найдем $H_X(t)$ сначала для $t \in [0; 1]$. Обозначим $f(u) = f_X(t)$. Имеем

$$F(u) = \begin{cases} 0, u < 0, \\ u,\ u \in [0;1], \\ 1,\ u > 1; \end{cases} \quad F(u+t) = \begin{cases} 0, u+t < 0 \Leftrightarrow u < -t, \\ u+t,\ u+t \in [0;1] \Leftrightarrow -t \leqslant u \leqslant 1-t, \\ 1,\ u+t > 1 \Leftrightarrow u > 1-t. \end{cases}$$

Для дальнейшего вычисления удобнее представление $F(u)$ и $F(u + t)$ в следующем виде:

$$F(u) = \begin{cases} 0, u < 0, \\ 0,\ -t \leqslant u \leqslant 0, \\ u,\ 0 < u \leqslant 1-t, \\ u,\ 1-t < u \leqslant 1, \\ 1,\ u > 1; \end{cases} \quad F(u+t) = \begin{cases} 0, < -t, \\ u+t,\ -t \leqslant u \leqslant 0, \\ u+t,\ 0 < u \leqslant 1-t, \\ 1,\ 1-t < u \leqslant 1, \\ 1,\ u > 1. \end{cases}$$

Тогда

$$F(u+t) - F(u) = \begin{cases} 0, < -t, \\ u+t,\ -t \leqslant u \leqslant 0, \\ t,\ 0 < u \leqslant 1-t, \\ 1-u,\ 1-t < u \leqslant 1, \\ 0,\ u > 1; \end{cases}$$

$$H_X(t) = n\int_0^{1-t} t^{n-1}du + n\int_{1-t}^{1}(1-u)^{n-1}du = n(1-t)t^{n-1} + n\frac{t^n}{n} = n(1-t)t^{n-1} + t^n$$

для $t \in [0; 1] \Rightarrow$

$$H_X(t) = \begin{cases} 0, t < 0, \\ n(1-t)t^{n-1} + t^n, t \in [0;1], \\ 1, t > 1 \end{cases} \Rightarrow\ h_X(t) = \begin{cases} 0, t \notin [0;1], \\ n(n-1)(1-t)t^{n-1}, t \in [0;1]. \end{cases}$$

Для нахождения $H_Y(t)$ используем связь случайных величин X и Y:

$$H_Y(t) = P(W_Y < t) = P\left(W_X = X_{(n)} - X_{(1)} < \frac{t}{b-a}\right) = H_X\left(\frac{t}{b-a}\right);$$

$$h_Y(t) = \frac{1}{b-a} h_X\left(\frac{t}{b-a}\right).$$

Задание 1.1. Самостоятельно выпишите $H_Y(t)$ и $h_Y(t)$ в явном виде.

1.3. Моделирование выборок значений случайной величины с заданным законом распределения

Часто для исследования вероятностных характеристик СВ приходится создавать соответствующий статистический материал. Основные идеи и приемы такого моделирования подробно изложены, например, в работе [21], приведем здесь краткие сведения по этому вопросу.

СВ с любым распределением могут быть получены из равномерного распределения на отрезке [0; 1] чисел $\{r_i\}$, вырабатываемых датчиком случайных чисел.

Моделирование дискретной случайной величины (метод маркировки)

Для моделирования n значений дискретной СВ X с заданным рядом распределения

X	x_1	x_2	...	x_m
P	p_1	p_2	...	p_m

нужно отрезок [0; 1] разбить на m частичных отрезков: $[0; p_1]$, $[p_1; p_1 + p_2]$, ..., $[p_1 + p_2 + ... + p_{m-1}; 1]$ и промаркировать реализацию n случайных чисел $\{r_i\}$ по построенным интервалам. Тогда попадание случайного числа r_i в j-й интервал будет означать, что значение разыгрываемой СВ X равно x_j, $j = 1, ..., m$; $i = 1, ..., n$.

Задача 1.1. Разыграть (смоделировать) шесть значений СВ X, имеющей ряд распределения

X	1	3	8
p	0,34	0,25	0,41

Решение

С помощью датчика СВ получим шесть случайных чисел $\{r_i\}$: 0,43; 0,61; 0,21; 0,84; 0,70; 0,05.

Промаркируем эти случайные числа по отрезкам [0; 0,34], [0,34; 0,59], [0,59; 1]. Получим, что случайные числа $\{r_i\}$ соответственно попали в интервалы с номерами 2, 3, 1, 3, 3, 1 что отвечает следующей смоделированной последовательности $\{x_i\}$: 3, 8, 1, 8, 8, 1.

Моделирование непрерывной случайной величины (метод обратных функций)

В основе моделирования непрерывной СВ лежит следующий легко доказываемый факт.

Утверждение 1.1. Если $F(x)$ — непрерывная строго монотонная функция распределения и $F^{-1}(x)$ — обратная к ней функция, то СВ $X = F^{-1}(R)$, где СВ R равномерно распределена на отрезке [0; 1], имеет функцию распределения $F(x)$.

Доказательство. Действительно, $P\{F^{-1}(R) < x\} = P\{R < F(x)\} = F(x)$, ч.т.д.

Таким образом, если функцию $F^{-1}(x)$ можно явно вычислить, то моделирование n значений $\{x_i\}$ СВ X производится из n значений случайных чисел $\{r_i\}$ по формуле

$$x_i = F^{-1}(r_i). \tag{1.1}$$

Этот способ моделирования называется методом обратных функций (МОФ). Приведем примеры применения МОФ.

Задача 1.2. Смоделировать получение значений СВ X в следующих случаях: а) $L(X) = R[a; b]$; б) $L(X) = E(L)$.

Решение

а) $F(x) = (x - a)/(x - b)$ при $a < x < b$; по формуле (1.1) имеем

$$r_i = (x_i - a)/(b - a); \quad x_i = r_i(b - a) + a; \quad F(x) = 1 - \exp(-\lambda x);$$

б) $x \geqslant 0$; по формуле (1.1) имеем

$$r_i = 1 - \exp(-\lambda x_i) \Rightarrow x_i = -(1/\lambda)\ln(1 - r_i),$$

а так как СВ $(1 - r_i)$ распределена так же, как и СВ r_i, то возможное значение СВ x_i можно получить по более простой формуле $x_i = -(1/\lambda)\ln r_i$.

Замечание 1.4. Если $F(x)$ не строго монотонная функция, то обратную к ней функцию определяют при моделировании значений СВ X следующим выражением:

$$F^{-1}(r_i) = \sup\{x_i : F(x_i) < r_i\}.$$

Ограниченность применения МОФ связана с невозможностью явного вычисления обратной функции $F^{-1}(x)$. В таких ситуациях используются различные частные приемы моделирования.

Метод Неймана

Этот метод является достаточно общим для разыгрывания непрерывной СВ X, когда она определена на конечном отрезке $[a; b]$ и плотность ее ограничена: $f(x) \leqslant M$, где M — постоянная.

Разыгрывание возможного значения СВ X производят по следующему алгоритму:

а) получают две СВ, равномерно распределенных на отрезке $[0;1]$: r_1 и r_2;

б) вычисляют координаты случайной точки $Z = (z_1, z_2)$, где $z_1 = a + r_1(b - a)$, $z_2 = r_2M$;

в) если точка Z лежит под кривой $y = f(x)$ (т.е. $z_2 \leqslant f(z_1)$), то полагают, что возможное значение СВ X есть $x_i = z_1$; в противном случае точку отбрасывают и переходят к следующей точке, повторяя алгоритм с пункта а) до получения возможного СВ X.

В качестве примеров остановимся на практическом моделировании значений наиболее распространенных распределений.

Замечание 1.5. Если смоделированы возможные значения СВ, то для моделирования реализаций СВ $X_2 = aX_1 + b$ нужно применить это линейное преобразование к полученным возможным значениям СВ $X_1 : x_{2i} = ax_{1i} + b$.

Задача 1.3. Смоделировать получение значений СВ X в следующих случаях:

а) $L(X) = R[a; b]$; б) $L(X) = N(a, \sigma)$.

Решение

а) Эта задача решена выше по МОФ. Та же формула для моделирования получается на основании замечания 1.5: $x_i = r_i(b - a) + a$.

б) Если $L(X) = N(0, 1)$ и моделирование возможных значений x_{0i} СВ X_0 проведено, то на основании замечания 1.5 получение возможных значений x_i СВ X можно производить по формуле $x_i = \sigma x_{0i} + a$. Таким образом, задача сводится к моделированию СВ X_0.

Приведем два способа моделирования возможных значений стандартной нормальной СВ X_0.

1. По теореме Леви (ЦПТ) возможные значения СВ X_0 могут быть получены при достаточно большом n по формуле

$$x_{0i} = \frac{\sum_{i=1}^{n} r_i - n/2}{\sqrt{n/12}}.$$

Часто на практике достаточно брать $n = 12$. Тогда имеем следующую простую формулу моделирования таких СВ: $x_{0i} = \sum\limits_{i=1}^{12} r_i - 6$.

2. Две возможные реализации x_{01} и x_{02} вычисляются из двух значений чисел r_1 и r_2 по формулам

$$x_{01} = \sqrt{-2\ln r_2}\cos(2\pi r_1); \quad x_{02} = \sqrt{-2\ln r_2}\sin(2\pi r_1).$$

Первый из этих способов легче реализуется, но приводит к приближенно нормальным числам.

Задача 1.4. Смоделировать получение значений СВ X для биномиального распределения $B(n, p)$.

Решение

Возможное значение биномиально распределенной СВ $X \sim B(n, p)$ может быть получено по общей схеме с $p_i = P(X = i) = C_n^i p(1-p)^{n-1}$, где $i = 0, 1, ..., n$, или проще — методом браковки: берется n независимых равномерно распределенных случайных чисел $r_1, ..., r_n$, тогда количество чисел из них, меньших p, есть возможное значение x_i СВ $X \sim B(n, p)$. В частности, при $n = 1$ получаем возможное значение бернуллиевской СВ.

Задача 1.5. Смоделировать получение значений СВ X для геометрического распределения $G(p)$ ($P(X = i) = p(1-p)^{i-1}$; $i = 1, 2, ...$).

Решение

$x_i = [\ln r_i / \ln(1-p)]$, где операция [.] означает взятие целой части числа.

С помощью описанного приема для моделирования возможного значения геометрического распределения могут быть получены возможные значения СВ, распределенной:

а) по сдвинутому геометрическому закону сд$G(p)$ ($P(Y = i) = p(1-p)^i$, $i = 1, 2, ...$): $y_i = x_i - 1$;

б) закону Паскаля $\pi a(r, p)$ ($P(U = i) = C_{i-1}^{r-1} p^r (1-p)^{i-r}$, $i = r, r+1, ...$): $U = \sum\limits_{i=1}^{r} X_i$, где СВ $X_i \sim G(p)$;

в) отрицательному биномиальному закону ОВ(r, p) ($P(V = i) = C_{r+i-1}^{i} p^r (1-p)^i$, $i = 0, 1, ...$): $V = \sum\limits_{i=1}^{r} Y_i$, где СВ $Y_i \sim$ сд$G(p)$.

Задача 1.6. Смоделировать получение значений СВ X для пуассоновского распределения $\pi(\lambda)$ ($P(X = k) = (\lambda e^{-\lambda}/k!)$, $k = 0, 1, 2, ...$).

Решение

Возможное значение СВ X может быть получено по одному из соотношений:

$$x_i = \min\left\{n : \prod_{i=1}^{n+1} r_i < e^{-\lambda}\right\};\ x_i = \max\left\{n : \prod_{i=1}^{n} r_i \geqslant e^{-\lambda}\right\}.$$

Задача 1.7. Смоделировать получение значений СВ X для полиномиального распределения $P(n, p_1, ..., p_m, m)$, где n — число опытов, m — число возможных исходов в каждом опыте, а p_i, $i = 1, ..., m$, — вероятность i-го исхода.

Решение

Имеем

$$P(x_1 = v_1, ..., x_m = v_m) = n! \prod_{i=1}^{m} \frac{p_i}{v_i};\ \ \sum_{i=1}^{m} v_i = n;\ \ \sum_{i=1}^{m} p_i = 1.$$

Для моделирования:

а) разбивают отрезок на m частей: $\Delta_1 = [0; p_1]$, $\Delta_i = \left[\sum_{k=1}^{i-1} p_k, \sum_{k=1}^{i} p_k\right]$, $i = 2, ..., m$;

б) получают n чисел $r_1, r_2, ..., r_n$;

в) для каждого r_j, $j = 1, ..., n$, находят отрезок Δ_i, которому r_j принадлежит;

г) для каждого отрезка Δ_i, $i = 1, ..., m$, определяют число v_i попавших в него чисел из $r_1, r_2, ..., r_n$; тогда считают, что x_i приняла значение v_i, а СВ $X = (x_1, x_2, ..., x_m)$ — значение $(v_1, ..., v_m)$.

1.4. Непараметрическая задача статистики

Исходной информацией является случайная величина X и выборка наблюдений над ней $\vec{x} = (x_1, x_2, x_3, ..., x_n)$. Требуется найти вид распределения (семейство распределений), к которому относится распределение СВ X. Для решения этой (непараметрической) задачи существует целый ряд обработок или представлений данной выборки.

1. Статистический (или частотный) ряд — это таблица вида

x	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k
$\omega_i = \frac{n_i}{n}$	ω_1	ω_2	...	ω_k

где $x_1, ..., x_k$ — возрастающая последовательность разных наблюденных значений СВ X; n_i — число наблюденных значений СВ $X = x_i$, $i = 1, ..., n$; $\sum_{i=1}^{k} n_i = n$ — объем выборки; $\sum_{i=1}^{k} \omega_i = 1$.

2. Полигон частот (относительных частот) — это ломаная линия (строится для дискретной СВ), соединяющая точки (x_i, n_i) $((x_i, \omega_i))$ (рис. 1.6).

Полигоны частот и относительных частот отличаются друг от друга только масштабом по оси ординат. Полигон относительных частот является графическим статическим аналогом соответствующего представления ряда распределения для точек (x_i, p_i).

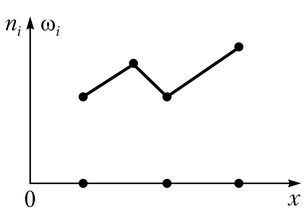


Рис. 1.6. **Полигон**

3. Гистограмма частот (относительных частот) — это ступенчатая фигура (строится для непрерывной СВ), сглаживающая кривая для которой является статистическим аналогом кривой распределения СВ X (рис. 1.7).

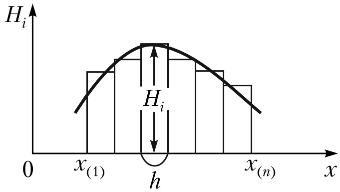


Рис. 1.7. **Гистограмма**

Гистограмма строится следующим образом: на оси абсцисс отмечают отрезок $(x_{(1)}, x_{(n)})$, его делят на равные отрезки длиной h таким образом, чтобы ни один отрезок не был пуст, а далее над i-м отрезком строят прямоугольники высотой H_i, где

$$H_i = \begin{cases} \dfrac{n_i}{h} \text{ – для гистограммы частот,} \\ \dfrac{\omega_i}{h} = \dfrac{n_i}{nh} \text{ – для гистограммы относительных частот.} \end{cases}$$

Площадь ступенчатой фигуры — гистограммы — равна

$$S = \begin{cases} \sum\limits_{i=1}^{k} \dfrac{n_i}{h} h = \sum\limits_{i=1}^{k} n_i = n \text{ — для гистограммы частот,} \\ \sum\limits_{i=1}^{k} \dfrac{n_i}{nh} h = \sum\limits_{i=1}^{k} \omega_i = 1 \text{ — для гистограммы относительных частот.} \end{cases}$$

Гистограммы частот и относительных частот отличаются только масштабом по оси ординат. Кривую распределения СВ X приближает сглаживающая кривая гистограммы относительных частот, так как в этом случае $S = \sum\limits_{i=1}^{k} \omega_i = 1$, как и площадь под графиком плотности распределения.

4. Эмпирическая функция распределения $F_n(x)$.

Эмпирическая функция распределения (ЭФР) определяется по формуле

$$F_n(x) = \frac{k}{n},$$

где n — объем выборки; k — число наблюдений, меньших x.

Задача 1.8. Пусть наблюденная выборка значений СВ X есть 3, 1, 4, 4, 3, 3, 3, 1, 3, 4. Построить ЭФР.

Решение

Построим статический ряд:

X	1	3	4
n_i	2	5	3

Имеем $\sum_{i=1}^{3} n_i = n = 10$.

Тогда ЭФР можно записать следующим образом (рис. 1.8).

$$F_n(x) = \begin{cases} \frac{0}{10} = 0,\ x \leqslant 1, \\ \frac{2}{10} = 0{,}2,\ \ 1 < x \leqslant 3, \\ \frac{2+5}{10} = 0{,}7,\ \ 3 < x \leqslant 4, \\ \frac{2+5+3}{10} = 1,\ \ x > 4. \end{cases}$$

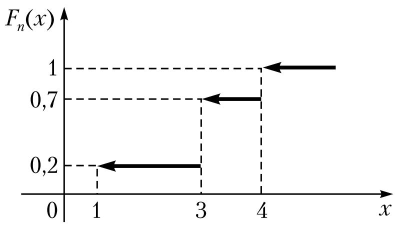


Рис. 1.8. **Эмпирическая функция распределения**

Свойства $F_n(x)$ совпадают со свойствами $F(x)$:

1) $0 \leqslant F_n(x) \leqslant 1$;

2)$F_n(-\infty) = 0$; $F_n(+\infty) = 1$;

3) $F_n(x)$ — неубывающая функция по x;

4) $F_n(x)$ непрерывна слева.

Вычислим распределение и моменты ЭФР.

Введем индикаторную функцию

$$J(z) = \begin{cases} 0,\ z \leqslant 0, \\ 1,\ z > 0. \end{cases}$$

Тогда

$$F_n(x) = \frac{k}{n} = \frac{1}{n} \sum_{i=1}^{n} J(x - x_i),$$

где при фиксированном значении x СВ $J(x - X) \sim B(1,\ p = F(x))$, так как

$J(x - X)$	0	1
P	$P(x - X \leqslant 0) = P(X \geqslant x) = 1 - P(X < x) = 1 - F(x)$	$P(x - X < 0) = P(X < x) = F(x)$

Известно, что тогда $Y = \sum_{i=1}^{n} J(x - x_i) \sim B(n, p)$, где $p = F(x)$,

$$MY = np = nF(x);\ DY = np(1-p) = nF(x)(1-F(x));$$

$$P\left\{F_n(x) = \frac{k}{n}\right\} = P\{Y = k\} = C_n^k (F(x))^k (1-F(x))^{n-k};$$

$$MF_n(x) = M\frac{Y}{n} = \frac{1}{n}MY = \frac{1}{n}nF(x) = F(x);$$

$$DF_n(x) = D\frac{Y}{n} = \frac{1}{n^2}DY = \frac{1}{n^2}nF(x)(1-F(x)) = \frac{F(x)(1-F(x))}{n} \to 0$$

при $n \to \infty$.

Таким образом, при фиксированном значении x $MF_n(x) = F(x)$, а это значит, что $F_n(x)$ — несмещенная оценка для MX; $DF_n(x) \to 0$ при $n \to \infty$, а это значит, что $F_n(x)$ — состоятельная оценка для $F(x)$.

Рассмотрим сходимость $F_n(x)$ к $F(x)$.

1. Обозначим $F_n(x) = F_n$; $F(x) = F$. Воспользуемся неравенством Чебышева:

$$P(|\xi - E\xi| \geqslant \varepsilon) \leqslant \frac{D\xi}{\varepsilon^2},$$

положим $\xi = F_n(x)$, тогда $E\xi = F_n(x)$ и

$$\xi_n = F_n D\xi = \frac{F(x)(1 - F(x))}{n}.$$

Получим $P(|F_n - F| \geqslant \varepsilon) \leqslant \dfrac{F(x)(1 - F(x))}{n\varepsilon^2} \to 0$ при $n \to \infty \Rightarrow$ $\Rightarrow F_n \xrightarrow{\text{п.в.}} F$ при $n \to \infty$.

2. Воспользуемся теоремой Колмогорова. Если $\xi_1, ..., \xi_n$ — последовательность независимых СВ и выполнено условие $\sum_{n=1}^{\infty} \dfrac{D\xi_n}{n^2} < \infty$, то $\xi_n \xrightarrow{\text{п.н.}} \xi$ при $n \to \infty$, т.е. $P(\xi_n > \xi) = 0$.

Положим $\xi_n = F_n$ и проверим условие теоремы Колмогорова:

$$\sum_{n=1}^{\infty} \frac{DF_n(x)}{n^2} = \sum_{n=1}^{\infty} \frac{F(x)(1 - F(x))}{n^3} < \infty \Rightarrow F_n \xrightarrow{\text{п.н.}} F \text{ при } n \to \infty.$$

Так как из сходимости почти наверное следует сходимость по вероятности, то $\xi_n \xrightarrow{\text{п.в.}} \xi$ при $n \to \infty$.

Замечание 1.6. Из доказанного утверждения следует, что так как при $n \to \infty$ $F_n \xrightarrow{\text{п.н.}} F$ по теореме Колмогорова, то $F_n \xrightarrow{\text{п.в.}} F$, что было установлено раньше непосредственно по неравенству Чебышева.

Проверим точность и надежность $F_n(x)$ для $F(x)$ при фиксированном значении x.

1. Пусть n велико. Тогда по следствию из интегральной теоремы Муавра — Лапласа

$$P = P(|F_n - F| \leqslant \varepsilon) = P\left(\left|\frac{k}{n} - P\right| \leqslant \varepsilon\right) \approx$$

$$\approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{F(1-F)}}\right) \geqslant 2\Phi(2\varepsilon\sqrt{n}) \to 1$$

при $n \to \infty$, так как $F(1 - F) \leqslant 4^{-1}$.

Здесь P — надежность оценки F_n для F с точностью ε, а $\geqslant 2\Phi(2\varepsilon\sqrt{n})$ — нижняя граница надежности.

2. При любом n по неравенству Чебышева имеем

$$P = P(|F_n - F| \leqslant \varepsilon) \geqslant 1 - \frac{DF_n}{\varepsilon^2} = 1 - \frac{F(1-F)}{n\varepsilon^2} \geqslant 1 - \frac{1}{4n\varepsilon^2}$$

— нижняя граница надежности оценки F_n для F с точностью ε.

Контрольные вопросы и задания

1. Дайте определения основных понятий математической статистики.

2. Дайте краткую характеристику основных задач математической статистики.

3. Чему равны плотность и функция распределения k-й порядковой статистики СВ $X \sim R[0; 1]$?

4. Как можно смоделировать возможные значения СВ $X \sim B(n, p)$?

5. Как можно смоделировать возможные значения СВ $X \sim N(a, \sigma)$?

6. Чему равны распределение и моменты эмпирической функции распределения $F_n(x)$?

7. Каковы точность и надежность эмпирической функции распределения для теоретической функции распределения $F_n(x)$ в каждой точке x^2?

Глава 2
ТЕОРИЯ ОЦЕНИВАНИЯ

В результате изучения данной главы студенты должны:

знать

• точечный и доверительный подходы к оцениванию неизвестных параметров распределения;

• свойства точечных и доверительных оценок;

• методы анализа и получения оценок;

уметь

• решать параметрическую задачу точечного и интервального оценивания;

• определять соответствие точных и выборочных моментов распределений;

владеть

• методами оценивания и сравнения качества оценок.

Данная глава посвящена решению одной из основных задач математической статистики — параметрической задачи. Исходными данными здесь являются изучаемая случайная величина X, выборка $\vec{x} = (x_1, ..., x_n)$ значений СВ X с точностью до параметра $\theta = (\theta_1, ..., \theta_n)$. Требуется оценить заданную функцию от θ: $\tau(\theta)$.

Рассматриваются два подхода к решению задачи: точечное и доверительное оценивание. Обсуждаются методы построения таких оценок.

Большое внимание отводится вопросу анализа качества полученных оценок с различных точек зрения и при различной природе неизвестного параметра распределения СВ X.

2.1. Выборочные моменты и их свойства

Начальные сведения

В качестве оцениваемых функций в параметрической задаче статистики часто встречаются выборочные моменты. Пусть X — наблюдаемая СВ, $\vec{x} = (x_1, ..., x_n)$ — выборка объема n.

$\widehat{\alpha}_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r$ — r-й начальный выборочный момент;

при $r = 1$ $\alpha_i = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ — выборочное среднее;

$\widehat{\mu}_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^r$ — r-й центральный выборочный момент;

при $r = 2$ $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ — выборочная дисперсия;

$\widehat{\sigma} = \widehat{\sigma}_x = \sqrt{S_1^2}$ — выборочное среднеквадратическое отклонение;

при известном $MX = m$ имеем

$\widehat{m}_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^r$ — r-й центральный выборочный момент.

Если выборки значений СВ X и Y есть соответственно $x_1, ..., x_n$; $y_1, ..., y_n$, то:

$\widehat{K}_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$ — выборочная корреляция;

$\widehat{r}_{xy} = \frac{\widehat{K}_{xy}}{\widehat{\sigma}_x \widehat{\sigma}_y}$ — выборочный коэффициент корреляции;

$\widehat{S}_k = \frac{\widehat{\mu}_3}{\widehat{\sigma}^3}$ — выборочный коэффициент асимметрии;

$\widehat{\varepsilon} = \frac{\widehat{\mu}_4}{\widehat{\sigma}^4} - 3$ — выборочный коэффициент эксцесса.

Свойства выборочных моментов

Сформулируем свойства в форме задач.

Задача 2.1. Исследовать $\widehat{\alpha}_r$ на несмещенность и состоятельность для α_r — r-го начального момента ($\alpha_r = MX^r$).

Решение

$$M\widehat{\alpha}_r = M\left(\frac{1}{n}\sum_{i=1}^{n} x_i^r\right) = \frac{1}{n}\cdot n\cdot MX^r = MX^r = \alpha_r,$$

т.е. $\widehat{\alpha}_r$ — несмещенная оценка для α_r; по теореме Хинчина при $\xi_i = x_i^r$ имеем: если $Ex_i^r < \alpha_r < \infty$, то для СВ $\xi_1, ..., \xi_n$ выполняется ЗБЧ, т.е. $\widehat{\alpha}_r$ — состоятельная оценка для α_r. В частности, при $r = 1$ $\overline{x}$ — несмещенная состоятельная оценка для MX.

Замечание 2.1. Отсюда получаем, например, что статистика $\overline{x}$ есть несмещенная состоятельная оценка для параметра λ распределения Пуассона $\pi(\lambda)$ и параметра a нормального распределения $N(a, \sigma)$.

Замечание 2.2. Если при исследовании смещенности оценки $T(x)$ получается линейная функция L от параметра θ, то для построения несмещенной оценки для θ нужно применить к оценке $T(x)$ преобразование L^{-1}.

Задача 2.2. Исследовать на несмещенность и состоятельность S_1^2 для DX.

Решение

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\left(\sum_{i=1}^{n}x_i^2 - \frac{2}{n}\overline{x}\sum_{i=1}^{n}x_i + \overline{x}^2\right) = \frac{1}{n}\left(\sum_{i=1}^{n}x_i^2 - \overline{x}^2\right);$$

$$M\overline{x} = MX;\ D\overline{x} = \frac{DX}{n} \Rightarrow M\overline{x}^2 = D\overline{x} + (M\overline{x})^2 = \frac{DX}{n} + (MX)^2;$$

$$MS_1^2 = \frac{1}{n}nMX^2 - M\overline{x}^2 = MX^2 - \frac{DX}{n} - (MX)^2 =$$

$$= \left(1 - \frac{1}{n}\right)DX = \frac{n-1}{n}DX \underset{n\to\infty}{\to} DX \Rightarrow$$

$\Rightarrow S_1^2$ — смещенная, но асимптотически несмещенная оценка для DX.

Смещение $b = MS_1^2 - DX = -\dfrac{DX}{n} < 0 \Rightarrow S_1^2$ в среднем занижает значение DX. Так как смещение линейно, можно его подправить:

$$S_2^2 = \frac{n}{n-1}S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2;$$

$$MS_2^2 = \frac{n}{n-1}MS_1^2 = DX \Rightarrow$$

$\Rightarrow S_2^2$ — несмещенная оценка для DX.

Состоятельность S_1^2 для DX следует из теоремы Хинчина для $\{\xi_i\} = \{x_1^2\}$ и состоятельности $\overline{x}$ для MX, так как

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2 \underset{n\to\infty}{\to} EX^2 - (EX)^2 = DX.$$

Задача 2.3. Исследовать оценку $\widehat{m}_r$ для m_r.

Решение

При $\xi_i = x_i - m$ получаем

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - M\xi_i\right| > \varepsilon\right\} = P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}(x_i - m)^r - m_r\right| > \varepsilon\right\} \underset{n\to\infty}{\to} 0.$$

Отсюда, в частности, следует, что при известном $MX = m$ $\widehat{m}_2$ — несмещенная и состоятельная оценка для DX.

Задача 2.4. Исследовать точность и надежность оценки $\overline{x}$ для $MX = m$.

Решение

Обозначим $L(Z)$ — закон распределения СВ Z.

Если $\gamma = P(|\overline{x} - m| > \varepsilon)$, то ε — точность, а γ — надежность оценки $\overline{x}$ для m; найдем распределение статистики $\overline{x}$. Обозначим

$$S_n = \overline{x} = \sum_{i=1}^{n} \frac{x_i}{n};\quad S_n^* = \frac{S_n - ES_n}{\sqrt{DS_n}};\quad ES_n = E\overline{x} = m;$$

$$DS_n = D\overline{x} = \frac{Dx}{n} = \frac{\sigma^2}{n};\quad S_n^* = \frac{\overline{x} - m}{\sigma/\sqrt{n}}.$$

Тогда по теореме Леви

$$L(S_n^x) = L\left(\frac{\overline{x} - m}{\sigma/\sqrt{n}}\right) \underset{n\to\infty}{\to} N(0,1);$$

$$\overline{x} = S_n^* \cdot \frac{\sigma}{\sqrt{n}} + m \Rightarrow L(\overline{x}) \underset{n\to\infty}{\to} N\left(m, \frac{\sigma}{\sqrt{n}}\right);$$

$$\gamma = P(|\overline{x} - m| > \varepsilon) = P(m - \varepsilon < \overline{x} < m + \varepsilon) =$$

$$= \Phi\left(\frac{m + \varepsilon - m}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{m - \varepsilon - m}{\sigma/\sqrt{n}}\right) = 2\Phi(\varepsilon\sqrt{n}/\sigma).$$

Таким образом, надежность $\overline{x}$ для MX с точностью ε есть $2\Phi(\varepsilon\sqrt{n}/\sigma)$ и может быть увеличена за счет выбора большего n.

Задача 2.5 (числовой пример). С какой вероятностью (надежностью) совершается ошибка, меньшая $\varepsilon = 0{,}3$, при оценке $MX = m$ через $\overline{x}$, если:

а) $n = 20$; $\overline{x} = 4{,}52$; $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = 2{,}35$;

б) $n = 100$; $\overline{x} = 4{,}52$; $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = 2{,}35$.

Решение

а) По задаче 2.2 при $D\overline{x} = \frac{DX}{n} \approx \frac{2{,}35}{20} = 0{,}1175$; $\sigma \approx \sqrt{\frac{2{,}35}{20}} = 0{,}343$;

$$\gamma = P(|\overline{x} - MX| < \varepsilon) \approx 2\Phi(0{,}3/0{,}343) \approx 0{,}618;$$

б) $\sigma \approx \sqrt{\frac{2{,}35}{100}} = 0{,}1533; \gamma = P(|\overline{x} - m| < 0{,}3) \approx 2\Phi(0{,}3/0{,}1533) \approx 0{,}9.$

При сравнении результатов а) и б) наглядно видно повышение надежности γ с ростом n.

Задача 2.6. Исследовать на несмещенность $\widehat{K}_{xy}$ для K_{xy}.

Решение

$$\widehat{K}_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}); \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i; \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i; MX = m_x, MY = m_y;$$

$$M\widehat{K}_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = Mx_iy_i - M\overline{x}y_i - M\overline{y}x_i + M\overline{x}\overline{y}.$$

$Mx_iy_i = K_{xy} + m_xm_y$, так как x_i и y_i независимы;

$$My_i\bar{x} = M\left(y_i \cdot \frac{1}{n}\sum_{j=1}^{n} x_j\right) = \frac{1}{n}Mx_iy_i + \frac{1}{n}\sum_{i\neq j} Mx_iy_i =$$

$$= \frac{K_{xy} + m_xm_y}{n} + \frac{n-1}{n}m_xm_y = \frac{K_{xy}}{n} + m_xm_y;$$

$$M\bar{x}\,\bar{y} = \frac{1}{n^2}M\left(\sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j\right) = \frac{1}{n^2}\sum_{i\neq j} Mx_iy_j + \frac{1}{n^2}\cdot nMx_iy_i =$$

$$= \frac{n(n-1)}{n^2}m_xm_y + \frac{1}{n}(K_{xy} + m_xm_y) =$$

$$= m_xm_y - \frac{1}{n}m_xm_y + \frac{K_{xy}}{n} + \frac{1}{n}m_xm_y = m_xm_y + \frac{K_{xy}}{n}.$$

Подставим все это в выражение для $M\widehat{K}_{xy}$:

$$M\widehat{K}_{xy} = K_{xy} + m_xm_y - \frac{K_{xy}}{n} - m_xm_y - \frac{K_{xy}}{n} - m_xm_y + \frac{K_{xy}}{n} + m_xm_y =$$

$$= \left(1 - \frac{1}{n}\right)K_{xy} = \frac{n-1}{n}K_{xy} \Rightarrow$$

$\Rightarrow \widehat{K}_{xy}$ — смещенная, но асимптотически несмещенная оценка для K_{xy}. Смещение оценки $\widehat{K}_{xy}$ есть $b = \left(\frac{n-1}{n} - 1\right)K_{xy} = -\frac{K_{xy}}{n}$ (линейное). Подправим смещение оценки $\widehat{K}_{xy}$:

$$\widetilde{K}_{xy} = \frac{n}{n-1}\widehat{K}_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x)(y_i - \bar{y})$$

— несмещенная оценка для K_{xy}.

Законы распределения и моменты статистик $\bar{x}$ и S_1^2 для $X \sim N(m, \sigma)$

Будем применять метод характеристических функций, обозначая, как и раньше, через $g_X(t)$ характеристическую функцию СВ X:

а) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $g_{x_i}(t) = \exp(itm - \sigma^2t^2/2)$;

$$g_{\bar{x}} = \left(g_{x_i}\left(\frac{t}{n}\right)\right)^n = e^{i\frac{t}{n}mn - \frac{n^2\sigma t^2}{2n^2}} = e^{itm - \frac{\left(\frac{\sigma}{\sqrt{n}}\right)^2 t^2}{2}} \Rightarrow$$

$$\Rightarrow \bar{x} \sim N\left(m, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow \begin{cases} E\bar{x} = m, \\ D\bar{x} = \dfrac{\sigma^2}{n} = \dfrac{DX}{n}; \end{cases}$$

б) $S_1^2 = \dfrac{1}{n}\sum_{i=1}^{n} x_i^2 - x^2.$

Подвергнем $\{x_i\}$ линейному преобразованию с матрицей преобразования следующего вида:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n} \\ \dfrac{1}{\sqrt{n}} & \dfrac{1}{\sqrt{n}} & \ldots & \dfrac{1}{\sqrt{n}} \end{pmatrix},$$

т.е.

$$\begin{cases} y_i = \sum\limits_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, n-1, \\ y_n = \dfrac{1}{\sqrt{n}} \sum\limits_{j=1}^{n} x_j = \sqrt{n}\,\overline{x} \end{cases} \tag{2.1}$$

при выполнении условий

$$\begin{cases} \sum\limits_{j=1}^{n} a_{ij}^2 = 1, \quad i = 1, \ldots, n-1, \\ \sum\limits_{k=1}^{n} a_{ik} a_{jk} = 0, \quad i = 1, \ldots, n-1; \\ \sum\limits_{j=1}^{n} a_{ij} = 0, \quad i = 1, \ldots, n-1. \end{cases} \tag{2.2}$$

Покажем, что условия (2.2) задают ортогональное преобразование для СВ $\{y_i\}$, т.е. для нее выполняются условия

$$\begin{cases} \sum\limits_{j=1}^{n} a_{ij}^2 = 1, \\ \sum\limits_{k=1}^{n} a_{ik} a_{jk} = 0, \quad i, j = 1, \ldots, n. \end{cases} \tag{2.3}$$

Нужно показать, что для данного преобразования (2.1) условия (2.3) следуют из условий (2.2). Действительно, имеем $\sum\limits_{j=1}^{n} a_{ij}^2 = 1$ (и при $i = n$ $\sum\limits_{j=1}^{n} a_{ij}^2 = n \cdot \left(\dfrac{1}{\sqrt{n}}\right)^2 = 1$), а из условия $\sum\limits_{k=1}^{n} a_{ik} a_{jk} = 0$ при $i = n$ имеем

$$\sum_{k=1}^{n} a_{nk}a_{jk} = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} a_{jk} = 0 \Rightarrow \sum_{k=1}^{n} a_{jk} = 0.$$

Таким образом, доказано, что для преобразования (2.1) выполнены условия (2.3), а значит, преобразование (2.1) — ортогональное преобразование, сохраняющее сумму квадратов:

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2 .$$

СВ $x_1, ..., x_n$ — независимые как элементы выборки, каждая из которых ~ $N(m, \sigma)$, значит, их совместное распределение нормально (с плотностью, равной произведению их одинаковых плотностей). Известно, что ортогональное преобразование переводит нормальный вектор $(x_1, ..., x_n)$ в другой нормальный вектор $(y_1, ..., y_n)$.

Найдем моменты СВ $\{y_i\}$, $i = 1, ..., n$, при условиях (2.2), обозначив $Mx_i = MX = m$, $Dx_i = DX = \sigma^2$, $i = 1, ..., n$:

$$My_i = M\left(\sum_{j=1}^{n} a_{ij}x_j\right) = \sum_{j=1}^{n} a_{ij}Mx_j = m\sum_{j=1}^{n} a_{ij} = 0,\ i = 1, ..., n-1;$$

$$My_n = \sqrt{n}M\overline{x} = m\sqrt{n};\ Dy_i = D\sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} a_{ij}^2 Dx_j = \sigma^2 \sum_{j=1}^{n} a_{ij}^2 = \sigma^2;$$

при $i \neq j$

$$K_{y_i y_j} = My_i y_j - My_i My_j = M\left(\sum_{k=1}^{n} a_{ik}x_k \sum_{l=1}^{n} a_{jl}x_l\right) - \sum_{k,l}^{n} a_{ik}a_{jl}Mx_k x_l =$$

$$= MX^2 \sum_{k=1}^{n} a_{ik}a_{jk} + (MX)^2 \sum_{k=1}^{n} a_{ik} \sum_{l=1}^{n} a_{jl} =$$

$$= MX^2 \sum_{k=1}^{n} a_{ik} \sum_{k=1}^{n} a_{jk} - (MX)^2 \sum_{k\neq 1} a_{ik}a_{jl} = 0,$$

а это означает для нормального вектора $(y_1, ..., y_n)$ независимость его компонент с учетом следующего утверждения (без доказательства).

Утверждение 2.1. Для гауссовского вектора некоррелированность его компонент эквивалентна их независимости.

Так как ортогональное преобразование сохраняет сумму квадратов и $y_n^2 = n\overline{x^2}$, то верно равенство

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{1}{n}y_n^2 = \frac{1}{n}\sum_{i=1}^{n-1} y_i^2 = \sum_{i=1}^{n-1}\left(\frac{y_i}{\sqrt{n}}\right)^2 = \sum_{i=1}^{n-1} (y_i^*)^2,$$

где $y_i^* = \dfrac{y_i}{\sqrt{n}};\ g_{y_i}(t) = e^{-\frac{\sigma^2 t^2}{2}};\ g_{y_i^*}(t) = g_{y_i}\left(\dfrac{t}{\sqrt{n}}\right) = e^{-\frac{\sigma^2 t^2}{2n}} \Rightarrow$

$$\Rightarrow y_i^* \sim N\left(0, \sigma^* = \frac{\sigma}{\sqrt{n}}\right) \Rightarrow \sum_{i=1}^{n-1} (y_i^*)^2 = S_1^2 \sim \Gamma_{\alpha,\lambda},$$

где $\alpha = (n-1)/2$, а $\lambda = \dfrac{1}{2\sigma^{*2}}$, что соответствует плотности гамма-распределения

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

$$ES_1^2 = \frac{\alpha}{\lambda} = \frac{(n-1)2\sigma^2}{2n} = \frac{n-1}{n}\sigma^2;$$

$$DS_1^2 = \frac{\alpha}{\lambda^2} = \frac{(n-1)4\sigma^4}{2n^2} = \frac{2(n-1)}{n^2}\sigma^4;$$

$$\frac{y_i}{\sigma^*} = \frac{y_i\sqrt{n}}{\sigma} \sim N(0,1) \Rightarrow \sum_{i=1}^{n^*=n-1} \left(\frac{y_i}{\sigma^*}\right) \sim \chi^2_{n-1};$$

$$\sum_{i=1}^{n-1}\left(\frac{\sqrt{n}y_i}{\sigma}\right) = \frac{nS_1^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow \begin{cases} M\dfrac{nS_1^2}{\sigma^2} = n-1, \\ D\dfrac{nS_1^2}{\sigma^2} = 2(n-1) \end{cases} \Rightarrow \begin{cases} MS_1^2 = \dfrac{n-1}{n}\sigma^2, \\ DS_1^2 = \dfrac{2(n-1)}{n^2}\sigma^4. \end{cases}$$

Замечание 2.3. Не нарушая общности, можно считать, что исходные СВ $X_i \sim N(0, \sigma)$, $i = 1, ..., n$, так как S_1^{*2} для $X^* \sim N(m, \sigma)$, и S_1^2 для $X_i \sim N(0, \sigma)$, $i = 1, ..., n$, совпадают. Покажем это:

$$x_i = x_i^* - m \Rightarrow x_i^* = x_i + m;\; S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2;$$

$$S_1^{*2} = \frac{1}{n}\sum_{i=1}^{n}(x_i^* - \overline{x^*})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i + m - \bar{x} - m)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = S_1^2,$$

так как $\bar{x}^* = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - m) = \bar{x} + m.$

Замечание 2.4. Для матрицы ортогонального преобразования $A = (a_{ij})$, $i, j = 1, ..., n$, $A^T = A^{-1}$, где A^T — транспонированная матрица; A^{-1} — обратная матрица.

Тогда условия ортогональности (2.3) эквивалентны условиям

$$\begin{cases} \sum_{k=1}^{n} a_{ki}a_{kj} = 0,\; i \neq j, \\ \sum_{k=1}^{n} a_{ki}^2 = 1. \end{cases} \tag{2.4}$$

Покажем, что с учетом замечания 2.3 и эквивалентности условий (2.2) и (2.3) для рассматриваемых преобразований (2.1) сохраняются суммы квадратов, т.е. $\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2$. Действительно:

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n-1}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 + (\sqrt{n}\,\bar{x})^2 =$$

$$= \sum_{i=1}^{n-1}\sum_{j=1}^{n} a_{ij}^2 x_j^2 + \sum_{i=1}^{n-1}\sum_{j_1}\sum_{j_2} a_{ij_1}a_{ij_2}x_{j_1}x_{j_2} + \frac{1}{n}(x_1 + \ldots + x_n)^2 =$$

$$= \sum_{j=1}^{n} x_j^2 \sum_{i=1}^{n-1} a_{ij}^2 + \sum_{j_1} x_{j_1} \sum_{j_2} x_{j_2} \sum_{i=1}^{n-1} a_{ij_1}a_{ij_2} + \frac{1}{n}\sum_{i=1}^{n-1} x_j^2 + \frac{1}{n}\sum_{j_1}\sum_{j_2} x_{j_1}x_{j_2},\ j_1 \neq j_2.$$

Отдельно вычисляем следующие суммы:

$$\sum_{i=1}^{n-1} a_{ij_1}a_{ij_2} = 0 - a_{nij_1}a_{nij_1} = -\frac{1}{\sqrt{n}}\cdot\frac{1}{\sqrt{n}} = -\frac{1}{n};$$

$$\sum_{i=1}^{n-1} a_{ij}^2 = \sum_{i=1}^{n} a_{ij}^2 - a_{nj}^2 = 1 - \frac{1}{\sqrt{n}}\cdot\frac{1}{\sqrt{n}} = 1 - \frac{1}{n},$$

поэтому

$$\sum_{i=1}^{n} y_i^2 = \sum_{j=1}^{n} x_j^2\left(1 - \frac{1}{n}\right) + 2\left(-\frac{1}{n}\right)\sum_{j_1<j_2} x_{j_1}x_{j2} + \frac{1}{n}\sum_{j=1}^{n} x_j^2 + \frac{2}{n}\sum_{j_1<j_2} x_{j_1}x_{j_2} =$$

$$= \left(1 - \frac{1}{n} + \frac{1}{n}\right)\sum_{j=1}^{n} x_j^2 = \sum_{j=1}^{n} x_j^2, \text{ ч.т.д.}$$

Замечание 2.5. Если $X \sim N(m, \sigma)$, то $\bar{x}$ и S_1^2 независимы.

Действительно, по формуле (2.1) $\bar{x} = \dfrac{y}{\sqrt{n}}$, а

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{y_n^2}{n} = \frac{1}{n}\sum_{i=1}^{n-1} y_i^2,$$

т.е. $\bar{x}$ и S_1^2 зависят от разных, независимых между собой СВ — соответственно y_n и $(y_1, \ldots, y_{n-1})$, а значит, $\bar{x}$ и S_1^2 независимы между собой.

2.2. Точечные оценки

Задача точечного оценивания

Исследуется случайная величина X, распределение которой относится к параметрическому множеству $F_\theta(x)$, где $\theta = (\theta_1, \ldots, \theta_k)$ — неизвестный k-мерный параметр. В дальнейшем будем обозначать это короче: $X \sim F_\theta(x)$.

Имеется выборка наблюденных значений случайной величины X: $\tau = \tau(\theta)$ объема n. Требуется построить точечную оценку (статистику) для данной функции $\tau = \tau(\theta)$ и исследовать качество данной оценки.

Замечание 2.6. При такой постановке задачи нет проблемы построения оценки для τ, так как требования к ее качеству (близости к истинному значению) не высказаны. Математическая задача возникает тогда, когда эти требования математически формализованы. А в связи с тем что они не все и не всегда выполнимы, будем называть их желательными свойствами оценок.

Простейшие свойства точечных оценок

Как уже указывалось в параграфе 1.1, свойствами точечных оценок являются несмещенность, состоятельность и эффективность.

Эти свойства или пожелания к качеству точечной оценки объединены стремлением достичь определенной степени концентрации возможных значений оценки $\hat{\tau} = t(\vec{x})$ вокруг истинного значения оцениваемой функции $\tau(\theta)$.

Одновременное выполнение этих желательных свойств не всегда возможно, поэтому представление о «хорошей» оценке зависит от цели и возможностей исследования, определяющих приоритетные свойства оценки. Так, для малых выборок часто важна несмещенность оценки, а для больших — асимптотическая несмещенность и состоятельность. А иногда, сознательно отказываясь от одних свойств оценок, добиваются выполнения других, более важных с точки зрения исследования свойств.

Напомним, что оценка называется оптимальной, если она является несмещенной и с минимальной дисперсией.

Теорема 2.1. Если оптимальная оценка существует, то она единственна.

Доказательство (от противного). Пусть не так, т.е. существуют две оптимальные оценки для $\tau = \tau(\theta)$: $t_1 = t_1(\vec{x})$ и $t_2 = t_2(\vec{x})$. Тогда $Mt_1 = Mt_2 = \tau$; $Dt_1 = Dt_2 = \sigma^2$ — минимальная возможная дисперсия несмещенных оценок для τ.

Построим оценку $t_3 = t_3(\vec{x}) = \dfrac{t_1 + t_2}{2}$ и изучим ее свойства. Имеем $Mt_3 =$ $= \tau \Rightarrow t_3$ — снова несмещенная оценка для τ. Дисперсия

$$Dt_3 = \frac{1}{4}(Dt_1 + Dt_2 + 2K_{t_1t_2}) = \frac{1}{4}(2\sigma^2 + 2K_{t_1t_2}) = \frac{\sigma^2 + K_{t_1t_2}}{2} \leqslant \sigma^2,$$

так как по неравенству Коши — Буняковского $|K_{t_1t_2}| \leqslant \sigma^2$. Но Dt_3 не может быть меньше σ^2 по условию. Отсюда следует, что $Dt_3 = \sigma^2$ (поэтому t_3 — оптимальная оценка для τ), а это значит, что $|K_{t_1t_2}| = \sigma^2$, или $|r_{t_1t_2}| = 1$ и t_1 и t_2

линейно зависимы, т.е. $t_2 = at_1 + b$, где a и b — const. Тогда $Mt_1 = aMt_2 + b$, или $\tau = a\tau + b \Rightarrow a = 1, b = 0 \Rightarrow t_1 = t_2$, ч.т.д.

Общий подход к сравнению оценок

Определение 2.1. ***Функция потерь*** $G(t(\vec{x}), \tau(\theta)) = G(t, \tau)$ — это любая неотрицательная функция, дающая потерю (ущерб) в результате того, что за $\tau = \tau(\theta)$ принята ее оценка $\hat{\tau} = t(\vec{x}) = t$.

Однако важными являются не единичные потери, а средние при многократном использовании оценки вместо истинного значения оцениваемой функции. Поэтому введем функцию риска $R(t, \tau) = M(G(t, \tau))$ — это средние потери относительно выбранной функции потерь.

Часто функция потерь выбирается в виде $G(t, \tau) = (t - \tau)^2$ и называется квадратичной функцией потерь, а соответствующий риск $R(t, \tau) = M(t - \tau)^2$ — квадратичным риском.

Для несмещенных оценок квадратичный риск

$$R(t, \tau) = M(t - \tau)^2 = M(t - Mt)^2 = Dt,$$

поэтому сравнение качества несмещенных оценок по квадратичному риску лежит в русле этого общего подхода и совпадает с их сравнением по эффективности.

Смещение оценки

Пусть $b = Mt - \tau$. Тогда b называется ***смещением*** оценки t для $Mt = \tau + b$. Если $b = 0$, оценка является несмещенной, если $b > 0$ ($b < 0$), то оценка в среднем завышает (занижает) истинное значение оцениваемой функции. В случае линейного смещения его легко устранить, т.е. подправить оценку по смещению. Пусть $Mt = a\tau + b$, где a и b — const, тогда получаем, что $M\left(\dfrac{t-b}{a}\right) = \tau \Rightarrow \hat{\tau} = \dfrac{t-b}{a}$ — несмещенная оценка для τ.

Пусть $Mt = \tau + b \Rightarrow \tau = Mt - b$; тогда можно установить *связь смещения, квадратичного риска и дисперсии оценки*:

$$R(t, \tau) = M(t - \tau)^2 = M(t - Mt + b)^2 =$$
$$= M(t - Mt)^2 + 2bM(t - Mt) + b^2 \Rightarrow R(t, \tau) = Dt + b^2$$

— эта формула часто упрощает вычисление квадратичного риска.

Достаточное условие состоятельности несмещенных и асимптотически несмещенных оценок

Теорема 2.2. Для состоятельности несмещенной оценки t_n достаточно, чтобы $Dt \underset{n\to\infty}{\to} 0$.

Доказательство. Пусть $Mt_n = \tau$ и $Dt \underset{n\to\infty}{\to} 0$. Воспользуемся неравенством Чебышева: $P(|\xi - M\xi| \geqslant \varepsilon) \leqslant \frac{D\xi}{\varepsilon^2}$, получив $\xi = t_n$, или, с учетом несмещенности t для τ: $P(|\xi - M\xi| \geqslant \varepsilon) \leqslant \frac{D\xi}{\varepsilon^2} \underset{n\to\infty}{\to} 0$, откуда и следует утверждение теоремы.

Теорема 2.3. Для состоятельности асимптотически несмещенной оценки t для τ достаточно, чтобы $Dt_n \underset{n\to\infty}{\to} 0$.

Доказательство. Пусть $Mt_n = \tau + \varepsilon_n$; ε_n, $Dt_n \to 0$; $\tau = Mt_n - \varepsilon_n$. Рассмотрим событие

$$A = \{|t_n - \tau| \geqslant \varepsilon\} = \{|t_n - Et_n + \varepsilon_n| \geqslant \varepsilon\} \subset \{|t_n - Et_n| + |\varepsilon_n| \geqslant \varepsilon\} =$$
$$= \{|t_n - Et_n| \geqslant \varepsilon - |\varepsilon_n|\} = B.$$

Из того что $A \subset B \Rightarrow P(A) \leqslant P(B)$, тогда по неравенству Чебышева имеем

$$P\{|t_n - \tau| \geqslant \varepsilon\} \leqslant P\{|t_n - Mt_n| \geqslant \varepsilon - |\varepsilon_n|\} \leqslant \frac{Dt_n}{(\varepsilon - |\varepsilon_n|)^2} \underset{n\to\infty}{\to} 0,$$

что и доказывает утверждение теоремы.

Замечание 2.7. Результаты приведенных теорем в указанных условиях часто упрощают установление состоятельности.

Далее рассмотрим примеры и задачи на анализ оценок.

Задачи на определение свойств оценок

Будем использовать обозначение $L(X)$ — закон распределения СВ X.

Задача 2.7. $L(X) = R[0; \theta]$. Проверить свойства оценки $\widehat{\theta} = 2x_1$ для θ.

Решение

Имеем $M\widehat{\theta} = 2Mx_1 = \frac{2\theta}{2} = \theta$; $\widehat{\theta} = 2x_1$ от объема выборки n не зависит, поэтому при $n \to \infty$ не сходится к θ, т.е. является несостоятельной оценкой для θ. Здесь мы имеем пример несмещенной и несостоятельной оценки для θ.

Задача 2.8. $L(X) = R[0; \theta]$. Проверить свойства оценки $\theta = 2x_{(n)}$ для θ. В случае смещенности подправить ее.

Решение

$$Mx_{(n)} = \int_0^\theta x^n \frac{n}{\theta^n} dx = \frac{n}{n+1}\theta \to \theta,$$

т.е. данная оценка $\widehat{\theta}$ является смещенной, но асимптотически несмещенной. Подправим ее. По замечанию 2.2 получаем, что оценка $\widetilde{\theta} = \frac{n+1}{n}\widehat{\theta} = \frac{n+1}{n}x_{(n)}$ — несмещенная оценка для параметра θ.

Для исследования состоятельности оценки $\hat{\theta}$ вычислим ее дисперсию:

$$Mx_{(n)}^2 = \int_0^{\theta} x^{n+1}\frac{n}{\theta^n}dx = \frac{n\theta^2}{n+2};$$

$$Dx_{(n)} = \frac{n\theta^2}{n+2} - (Mx_{(n)})^2 = \frac{\theta^2 n}{(n+1)^2(n+2)} \underset{n\to\infty}{\longrightarrow} 0,$$

откуда следует, что оценки $\hat{\theta}$ и $\tilde{\theta}$ являются состоятельными.

Задача 2.9. СВ X распределена по закону Коши:

$$f(x; \theta) = \frac{1}{n}[1 + (x - \theta)^2].$$

Состоятельна ли оценка $\hat{\theta} = \sum_{i=1}^{n} x_i = \bar{x}$ для θ?

Решение

Функция распределения закона Коши есть $F(x) = \frac{1}{2} + \frac{1}{\pi}\text{arctg}(x - \theta)$.

Характеристическая функции:
СВ X — $g_X(t) = \exp(itn\theta - n|t|)$;

СВ $Y = \sum_{i=1}^{n} x_i - g_Y(t) = \exp(itn\theta - n|t|)$;

СВ $\theta = \frac{\sum_{i=1}^{n} x}{n} - g_\theta(t) = g_Y\left(\frac{t}{n}\right) = \exp(itn\theta - |t|)$, т.е. СВ X и θ имеют одинаковое распределение (данное распределение Коши).

Тогда

$$P(|\hat{\theta} - \theta| > h) = P\{|x_i - \theta| > h) = 1 - P(-h + \theta < x_i < h + \theta) =$$
$$= 1 - (F(h+\theta) - F(-h+\theta)) = 1 - \left(\frac{1}{2} + \frac{1}{\pi}\text{arctg}\, h\right) = 1 - \frac{2}{\pi}\text{arctg}\, h \not\to 0.$$

При $n \to \infty$ это означает несостоятельность приведенной оценки $\hat{\theta}$ для θ.

Задача 2.10. Найти распределение и моменты статистики $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, если СВ $X \sim N(m, \sigma)$, $X_1, ..., X_n$ — выборка значений СВ X.

$g_{X_i}(t) = \exp\left(itm - \frac{\sigma^2 t^2}{2}\right)$.

Решение

$$g_{\bar{X}=\sum_{i=1}^{n}\frac{X_i}{n}}(t) = \left(g_{X_i}\left(\frac{t}{n}\right)\right)^n = \exp\left(i\frac{t}{n}mn - \frac{n\sigma^2 t^2}{2n^2}\right) = \exp\left(itm - \left(\frac{\sigma}{\sqrt{n}}\right)^2\frac{t^2}{2}\right) \Rightarrow$$

$$\Rightarrow \bar{X} \sim N\left(m, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow M\bar{X} = m;\ D\bar{X} = \frac{\sigma^2}{n}.$$

Задание 2.1. Показать, что статистика $S_3^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2$ при известном значении $MX = m$ является несмещенной для дисперсии.

Примеры, когда несмещенной оценки нет

Пример 2.1. СВ $X \sim R[0;\theta]$; $\tau(\theta) = \frac{1}{\theta}$. Пусть $t(x)$ — несмещенная оценка для $\tau(\theta) = \frac{1}{\theta}$, тогда $\frac{1}{\theta} = \int_0^{\theta}\frac{t(x)}{\theta}dx \Rightarrow \int_0^{\theta} t(x)dx = 1 \Rightarrow t(x) = \frac{1}{\theta}$, но это не статистика $\Rightarrow$ несмещенной оценки для $\tau(\theta) = \frac{1}{\theta}$ нет.

Пример 2.2. СВ $X \sim \pi(\theta)$; $\tau(\theta) = \frac{1}{\theta}$. Пусть $t(x)$ — несмещенная оценка для $\tau(\theta) = \frac{1}{\theta}$, тогда $\frac{1}{\theta} = \sum_{x=0}^{\infty} t(x)\frac{\theta^x e^{-\theta}}{x!}$, или $\sum_{E=0}^{\infty} t(x)\frac{\theta^{x+1}}{x!} = e^{\theta} = \sum_{x=0}^{\infty}\frac{\theta^x}{x!} \Rightarrow$ $\Rightarrow \tau(x) = \frac{1}{\theta}$, т.е. это не статистика $\Rightarrow$ несмещенной оценки для $\tau(\theta) = \frac{1}{\theta}$ нет.

Примеры бесполезных (осциллирующих) несмещенных оценок

Пример 2.3. СВ $X \sim G(\theta)$; $\tau(\theta) = \theta$; $P(X = x) = \theta(1-\theta)^{x-1}$, $\theta \in [0; 1]$, $x = 1, 2, \dots$. Пусть $t(x)$ — несмещенная оценка для $\tau(\theta) = \theta$, тогда $\sum_{x=1}^{\infty} t(x)\theta(1-\theta)^{x-1} = \theta \Rightarrow \sum_{x=1}^{\infty} t(x)(1-\theta)^{x-1} = 1 \Rightarrow t(x) = \begin{cases} 1, x = 1, \\ 0, x \geqslant 2 \end{cases}$ — это «плохая» оценка для θ, так как θ — вероятность успеха в одном опыте и не равна нулю по смыслу, кроме того, оценка $t(x)$ дает нулевую вероятность успеха в одном опыте, если успех не происходит в первом опыте, что не соответствует действительности.

Пример 2.4. СВ $X \sim \pi(\theta)$; $\theta \in [0; \infty)$; $\tau(\theta) = e^{-2\theta}$. Пусть $t(x)$ — несмещенная оценка для $\tau(\theta)$, тогда $\sum_{x=0}^{\infty} t(x)\frac{\theta^x e^{-\theta}}{x!} = e^{-2\theta} \Rightarrow \sum_{x=0}^{\infty} t(x)\frac{\theta^x}{x!} = e^{-\theta} \Rightarrow$ $\Rightarrow \sum_{x=0}^{\infty} t(x)\frac{\theta^x}{x!} = \sum_{x=0}^{\infty}\frac{(-1)^x\theta^x}{x!} \Rightarrow t(x) = \begin{cases} 1, x = 2k, \\ -1, x = 2k+1, \end{cases}$ $k = 0, 1, 2, \dots$ — это «плохая» оценка для $\tau(\theta) = e^{-2\theta}$, так как $e^{-2\theta} > 0$ и $t(x)$ реагирует только на четность значения СВ X.

Пример 2.5. СВ $X \sim B(m = \theta, p)$; $\tau(\theta) = \theta$; $\hat{\theta} = \frac{x_1}{p}$ — предлагаемая оценка для θ.

$M\hat{\theta} = M\frac{x_1}{p} = \frac{1}{p}\theta p = \theta \Rightarrow \hat{\theta}$ — несмещенная оценка для θ;

$$D\widehat{\theta} = M\widehat{\theta}^2 - (M\widehat{\theta})^2 = \sum_{x=0}^{n} \frac{x^2}{p^2} C_\theta^x p^x (1-p)^{\theta-x} - \theta^2 =$$

$$= \frac{1}{p^2}(\theta p(1-p) + \theta^2 p^2) - \theta^2 = \frac{\theta(1-p)}{p} + \theta^2 - \theta^2 = \frac{\theta(1-p)}{p}.$$

Если p близко к 1, то $D\widehat{\theta}$ мала и $\widehat{\theta}$ — «хорошая» оценка для θ, а при p малом $D\widehat{\theta}$ велика и $\widehat{\theta}$ — «плохая» оценка для θ. Значения $\widehat{\theta}$ нецелые, поэтому в качестве оценок для θ следует выбрать натуральные числа, ближайшие к $\frac{x_1}{p}$.

Примеры несостоятельных оценок

Пример 2.6. СВ $X \sim R[0;\theta]$; $\tau(\theta) = \theta$; $\widehat{\theta} = 2x_1$; $M\widehat{\theta} = 2Mx_1 = \theta \Rightarrow \widehat{\theta}$ — несмещенная оценка для $\tau(\theta)$, но несостоятельная оценка, так как не зависит от n.

Пример 2.7. СВ $X \sim \pi(\theta)$; $\widehat{\theta} = x_1$; $M\widehat{\theta} = \theta$, но $\widehat{\theta}$ — несостоятельная оценка, так как не зависит от n.

Пример 2.8. Пример несостоятельной оценки, зависящей от n. СВ $X \sim$ закон Коши с плотностью распределения $f(x, \theta) = \frac{1}{\pi[1 + (x-\theta)^2]}$; $\widehat{\theta} = \bar{x}$; $F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg}(x - \theta)$; характеристическая функция $g_x(t) = e^{it\theta - |t|}$, тогда

$$g_{\sum_{i=1}^{n} x_i}(t) = (g_{x_i}(t))^n = \exp(-itn\theta - n|t|);$$

$$g_{\bar{x}}(t) = g_{\sum_{i=1}^{n} x_i}\left(\frac{t}{n}\right) = e^{it\theta - |t|} = g_x(t) \Rightarrow$$

$$\Rightarrow P(|\widehat{\theta} - \theta| \geqslant \varepsilon) = P(|x_i - \theta| \geqslant \varepsilon) = 1 - P(-\varepsilon + \theta \leqslant x_i < \varepsilon + \theta) =$$

$$= 1 - (F(\varepsilon + \theta) - F(-\varepsilon + \theta)) = 1 - \frac{2}{\pi}\operatorname{arctg}\varepsilon \not\to 0$$

при $n \to \infty$, а это означает несостоятельность оценки $\widehat{\theta}$.

Задачи на сравнение качества оценок по среднему квадратичному риску

Задача 2.11. Пусть $X \sim R[0;\theta]$; даны оценки $\tau = \tau(\theta) = \frac{\theta}{2}$; $\widehat{\tau}_1 = x_1$; $\widehat{\tau}_2 = \bar{x}$; $\widehat{\tau}_3 = \frac{x_{(n)}}{2}$. Проверить эти оценки на несмещенность и сравнить их по квадратичным рискам.

Решение

Оценки $\widehat{\tau}_1$ и $\widehat{\tau}_2$ несмещены, так как

$$M\widehat{\tau}_1 = Mx_1 = Mx = \frac{\theta}{2}$$

и

$$M\hat{\tau}_2 = M\overline{x} = \frac{1}{n}\sum_{i=1}^{n} MX_i = n \cdot \frac{1}{n} \cdot \frac{\theta}{2} = \frac{\theta}{2}.$$

$$M\hat{\tau}_3 = M\frac{x_{(n)}}{2} = \int_0^Q \frac{x}{2} f_n(x)dx =$$

$$= \left\{ \begin{array}{c} F_n(x) = F^n(x) \to f_n(x) = F_n'(x) = nF^{n-1}(x)f(x); \\ f(x) = \frac{1}{\theta}; F(x) = \frac{x}{\theta} \text{ при } x \in (0; \theta] \end{array} \right\} =$$

$$= \frac{n}{2\theta^n}\int_0^\theta x^n dx = \frac{n\theta^{n+1}}{2\theta^n(n+1)} = \frac{n}{2(n+1)}\theta;$$

$$M\left(\frac{x_{(n)}}{2}\right)^2 = \frac{n}{2\theta^n}\int_0^\theta x^n dx = \frac{n\theta^{n+2}}{2\theta^n(n+2)} = \frac{n\theta^2}{4(n+2)};$$

$$D\hat{\tau}_3 = \frac{n\theta^2}{4(n+2)} - \frac{n^2\,\theta^2}{4(n+1)^2} = \frac{n\theta^2(n+1)^2 - n^2\theta^2(n+2)}{4(n+1)^2(n+2)} =$$

$$= \frac{n^3\theta^2 + 2n^2\theta^2 + n\theta^2 - n^3\theta^2 - 2n^2\theta^2}{4(n+1)^2(n+2)} = \frac{n\theta^2}{4(n+1)^2(n+2)}.$$

Найдем смещение b_{τ_3} оценки $\hat{\tau}_3$:

$$b_{\tau_3} = M\hat{\tau}_3 - \tau = \frac{n\theta}{2(n+1)} - \frac{\theta}{2} = -\frac{\theta}{2(n+1)} < 0;$$

$$R(\hat{\tau}_1, \tau) = R_1 = D\tau_1 = \frac{\theta^2}{48};\quad R(\hat{\tau}_2, \tau) = R_2 = D\tau_2 = \frac{\theta^2}{48n};$$

$$R(\hat{\tau}_3, \tau) = R_3 = D\hat{\tau}_3 + b_{\tau_3}^2 = \frac{n\theta^2}{4(n+1)^2(n+2)} + \frac{\theta^2}{4(n+1)^2} =$$

$$= \frac{\theta^2}{2(n+1)^2(n+2)}.$$

Теперь при $n = 1, 2, \ldots$ можно сравнивать риски данных оценок.

Задача 2.12. $X \sim N(a, \sigma)$. Сравнить по риску оценки для σ^2:

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \text{ и } S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Решение

Ранее было получено, что S_1^2 — смещенная оценка для σ^2 со смещением $b = -\frac{\sigma^2}{n}$; S_2^2 — несмещенная оценка ($MS_1^2 = \frac{n-1}{n}\sigma^2$; $MS_2^2 = \sigma^2$).

Так как элементы выборки $x_1, \ldots, x_n$ связаны равенством $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, то

$$\frac{nS_1^2}{\sigma^2} \sim X_{n-1}^2 \Rightarrow MX_{n-1}^2 = n-1,\ DX_{n-1}^2 = 2(n-1) \Rightarrow MS_1^2 = \frac{n-1}{n}\sigma^2 = \sigma^2 - \frac{\sigma^2}{n};$$

$$DS_1^2 = \frac{2\sigma^4(n-1)}{n};\ S_2^2 = \frac{n}{n-1}S_1^2 \ \Rightarrow MS_2^2 = \sigma^2;$$

$$DS_2^2 = R_2 = R(S_2^2, \sigma^2) = \frac{2\sigma^4(n-1)\cdot n^2}{n^2\cdot(n-1)^2} = \frac{2\sigma^4}{n-1} \underset{n\to\infty}{\longrightarrow} 0 \Rightarrow$$

$\Rightarrow S_2^2$ — состоятельная оценка для σ^2 (состоятельность S_1^2 для σ^2 определена раньше);

$$R_n = R_1(S_2^2, \sigma^2) = DS_1^2 + b^2 = \frac{2\sigma^4(n-1)}{n^2} + \frac{\sigma^4}{n^2} = \frac{2(n-1)\sigma^4}{n^2}.$$

Сравним риски R_1 и R_2: при $n > 1$ $R_1 < R_2$, т.е. S_1^2 по риску лучше, чем S_2^2.

Задачи на построение «лучшей» оценки из данного класса

При выборе оценок часто ограничиваются рассмотрением определенного класса оценок. Тогда в пределах этого класса при установлении приоритетных требований ищется «лучшая» в указанном смысле оценка.

Задача 2.13. СВ $X \sim N(a, \sigma^2 = \theta)$, θ — неизвестный параметр. В классе оценок $V(k) = \sigma^2(k) = \frac{k}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ построить оценку S_3^2 для σ^2 с наименьшим квадратичным риском. Найти $R(S_3^2, \sigma^2)$.

Решение

Ранее рассматривались следующие оценки для σ^2: $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ и $S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$. Получено, что S_1^2 — смещенная, но асимптотически несмещенная оценка для σ^2, S_2^2 — несмещенная оценка для σ^2 (см. задачу 2.2). Применительно к рассматриваемому классу получаем, что $S_1^2 = \sigma^2\left(k = \frac{n-1}{n}\right)$; $S_2^2 = \sigma^2\ (k = 1) \Rightarrow S_1^2$ и $S_2^2 \in V(k)$.

Найдем квадратичный риск оценки $\sigma^2(k)$:

$$R(\sigma^2(k), \sigma^2) = D\sigma^2(k) + b^2_{\sigma^2(k)};\ D\sigma^2(k) = D(kS_2^2) = \frac{2k^2\sigma^4}{n-1};$$

$$b^2_{\sigma^2(k)} = M\sigma^2(k) - \sigma^2 = M(kS_2^2) - \sigma^2 = (k-1)\sigma^2 \Rightarrow$$

$$\Rightarrow R(\sigma^2(k), \sigma^2) = \frac{2k^2\sigma^4}{n-1} + (k-1)^2\sigma^4 =$$

$$= \frac{\sigma^4}{n-1}(2k^2 + k^2(n-1) - 2k(n-1) + n - 1) =$$

$$= \frac{\sigma^4}{n-1}(k^2(n+1) - 2k(n-1) + n - 1).$$

Найдем $k = k_0(n) = k_0$, при котором $f(k) = k^2(n + 1) - 2k(n - 1) + n - 1$ достигает минимального значения и при этом значении $k = k_0$ риск оценки $\sigma^2(k)$ $R(\sigma^2(k), \sigma^2)$ будет минимален.

Искомое значение $k = k_0$ — абсцисса вершины параболы $f(k) \Rightarrow$ $\Rightarrow k_0 = \dfrac{n-1}{n+1}$, тогда искомая оценка из $V(k)$ с минимальным риском

$$S_3^2 = \sigma^2\left(k_0 = \frac{n-1}{n+1}\right) = \frac{1}{n+1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

При этом

$$R(S_3^2, \sigma^2) = R(\sigma^2(k), \sigma^2) = \frac{\sigma^4}{n-1}\left[\left(\frac{n-1}{n+1}\right)^2(n+1) - \frac{2(n-1)^2}{n+1} + n - 1\right] = \frac{2\sigma^4}{n+1}.$$

Задача 2.14. Пусть T_1 и T_2 — независимые оценки для $\tau(\theta) = \theta$: $MT_1 = \theta + d_1$, $MT_2 = \theta + d_2$, d_1 $d_2 \neq 0$ — const. В классе V: $\{T = aT_1 + bT_2$, $|a|, |b| < \infty\}$ построить несмещенную оценку для θ и найти ее дисперсию, если $DT_1 = \sigma_1^2$, $DT_2 = \sigma_2^2$.

Решение

Имеем $MT = (a + b)\theta + ad_1 + bd_2$; $DT = a^2\sigma_1^2 + b^2\sigma_2^2$; T — несмещенная оценка для θ, если

$$\begin{cases} a + b = 1, \\ ad_1 + bd_2 = 0, \\ d_1 \neq d_2 \end{cases} \Rightarrow \begin{cases} a^* = \dfrac{d_2}{d_1 - d_2}, \\ b^* = \dfrac{d_1}{d_1 - d_2}. \end{cases}$$

Тогда

$$T = a^*T_1 + b^*T_2 = \frac{d_2}{d_1 - d_2}T_1 + \frac{d_1}{d_1 - d_2}T_2;$$

$$R(T, \theta) = DT = \left(\frac{d_2}{d_1 - d_2}\right)^2\sigma_1^2 + \left(\frac{d_1}{d_1 - d_2}\right)^2\sigma_2^2,\ d_1 \neq d_2.$$

При $d_1 = d_2$ несмещенной оценки для θ в классе V не существует.

Замечание 2.8. Задача 2.14 дает алгоритм построения несмещенной оценки по двум любым смещенным с разным смещением.

Замечание 2.9. По задаче 2.14 при $d_1 = d_2$ можно строить примеры, когда не существует несмещенной оценки для θ в классе V.

Задача 2.15. Пусть T_1 и T_2 — независимые несмещенные оценки для $\tau(\theta) = \theta$; $DT_1 = \sigma_1^2$, $DT_2 = \sigma_2^2$. В классе V: $\{T = aT_2 + (1 - a)T_1$, $0 \leqslant a \leqslant 1\}$ найти несмещенную оценку T^* для θ с наименьшей возможной дисперсией DT^*.

Решение

$MT = a\theta + (1-a)\theta = \theta \Rightarrow T$ — несмещенная оценка для θ.

$DT = f(a) = a^2\sigma_2^2 + (1-a)^2\sigma_1^2$ достигает минимума при a, для которого $(DT)' = f'(a) = 0$ (так как коэффициент при a^2 положителен и равен $u = \sigma_1^2 + \sigma_2^2$):

$$f'(a) = 2a\sigma_2^2 + 2(1-a)\sigma_1^2 = 0 \Rightarrow a^* = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}.$$

Таким образом, искомая оценка

$$T^* = \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}T_2 + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}T_1;$$

$$DT^* = \left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2\sigma_2^2 + \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2\sigma_1^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}.$$

Задача 2.16. Пусть $x_1, ..., x_n; y_1, ..., y_m$ — две независимые выборки $\sim N(\mu, \sigma)$ (μ и σ — неизвестные параметры). В классе V: $\{a\overline{x} + (1-a)\overline{y}, 0 \leqslant a \leqslant 1\}$ ($\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i; \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$) найти оценку T^* для μ с минимальной дисперсией DT^*.

Решение

$MT = aM\overline{x} + (1-a)M\overline{y} = a\mu + (1-a)\mu = \mu \Rightarrow V$ — класс несмещенных оценок. Тогда мы находимся в условиях задачи 2.15, из которой следует, что единственной несмещенной оценкой в классе V с наименьшей дисперсией является

$$T^* = \frac{D\overline{x}}{D\overline{x}+D\overline{y}}\overline{y} + \frac{D\overline{y}}{D\overline{x}+D\overline{y}}\overline{x} = \left\{\begin{matrix} D\overline{x} = \dfrac{\sigma^2}{n}; \\ D\overline{y} = \dfrac{\sigma^2}{m} \end{matrix}\right\} =$$

$$= \frac{\frac{\sigma^2}{n}}{\frac{\sigma^2}{n}+\frac{\sigma^2}{m}}\overline{y} + \frac{\frac{\sigma^2}{m}}{\frac{\sigma^2}{n}+\frac{\sigma^2}{m}}\overline{x} = \frac{m\overline{y}-n\overline{x}}{n+m}.$$

Из задачи 2.15

$$DT^* = \frac{D\overline{x}D\overline{y}}{D\overline{x}+D\overline{y}}\overline{y} = \frac{\sigma^2}{n+m} \Rightarrow T^* \sim N\left(\mu, \sqrt{\frac{\sigma^2}{n+m}}\right).$$

Замечание 2.10. Задача 2.16 дает алгоритм построения лучшей по риску несмещенной оценки по данным для μ при $X \sim N(\mu, \sigma)$: $\overline{x}$ и $\overline{y}$.

Задача 2.17. СВ $X \sim B(1, p = \theta)$, θ — неизвестный параметр, $\tau(\theta) = \theta$. Сравнить квадратичные риски оценок $T_1 = \overline{x}$ и $T_2 = \overline{x} + \dfrac{0{,}5 - \overline{x}}{\sqrt{n}+1}$ (оценка Ходжеста — Лемана) для θ.

Решение

$MT_1 = \theta$; $MT_2 = \theta - \dfrac{\theta}{\sqrt{n}+1} + \dfrac{1}{2(\sqrt{n}+1)}$ $\Rightarrow T_1$ — несмещенная оценка, а T_2 — смещенная, но асимптотически несмещенная оценка для θ. Причем T_2 несмещенно оценивает лишь значение $\theta = 0{,}5$. Сравним T_1 и T_2 по квадратичным рискам:

$$R_1 = R_1(T_1, \theta) = DT_1 = \frac{\theta(1-\theta)}{n};\quad R_2 = R_2(T_2, \theta) = DT_2 + b_{T_2}^2,$$

где $b_{T_2}^2 = ET_2 - \theta = -\dfrac{\theta}{\sqrt{n}+1} + \dfrac{1}{2(\sqrt{n}+1)} = \dfrac{1-2\theta}{2(\sqrt{n}+1)}$; $DT_1 \underset{n\to\infty}{\longrightarrow} 0 \Rightarrow T_1$ — состоятельная оценка для θ.

$$T_2 = \bar{x} + \frac{0{,}5}{\sqrt{n}+1} + \frac{\bar{x}}{\sqrt{n}+1} = \frac{\sqrt{n}}{\sqrt{n}+1}\bar{x} + \frac{0{,}5}{\sqrt{n}+1};$$

$$DT_2 = \frac{n}{(\sqrt{n}+1)^2}\,\frac{Dx}{n} = \frac{\theta(1-\theta)}{(\sqrt{n}+1)^2} \underset{n\to\infty}{\longrightarrow} 0 \Rightarrow$$

$\Rightarrow T_2$ — состоятельная оценка для θ.

$$R_2 = \frac{\theta(1-\theta)}{2(\sqrt{n}+1)} + \frac{(1-2\theta)^2}{4(\sqrt{n}+1)^2} = \frac{1}{4(\sqrt{n}+1)^2},$$

т.е. R_2 не зависит от θ.

Изобразим R_1 и R_2 графически как функции от θ (рис. 2.1).

Из графика следует, что если есть основание считать, что значение θ близко к 0,5 (средняя зона A), то по риску лучше оценка T_2, если же значение близко к 0 или 1 (крайние зоны B), то лучше оценка T_1 для θ.

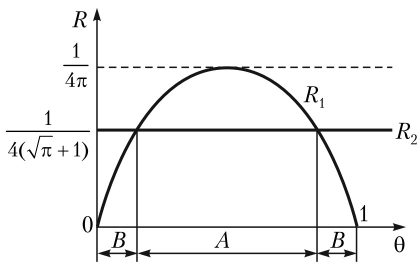


Рис. 2.1. **К задаче 2.17**

2.3. Достаточные статистики

Определение 2.2. Статистика $T = T(x)$ называется ***достаточной*** для семейства распределений $F_\theta(x)$ (θ — неизвестный параметр), если вероятность любой выборки $\vec{x} = (x_1, ..., x_n)$ при фиксированном значении T не зависит от значения неизвестного параметра θ (ее смысл: содержит в себе всю информацию о неизвестном параметре θ).

Определение 2.3. ***Функцией правдоподобия*** $L_\theta(\vec{x}) = L$ называется вероятность данной выборки $\vec{x} = (x_1, ..., x_n)$, т.е.

$$L = \begin{cases} \prod_{i=1}^{n} P(X = x_i), \text{ если СВ } X \text{ дискретна,} \\ \prod_{i=1}^{n} f(x_i), \text{ если СВ } X \text{ непрерывна } (f(x) \text{ — плотность ее распределения}). \end{cases}$$

Для опознавания и построения достаточной статистики (ДС) приведем критерий факторизации (КФ).

Теорема 2.4 (критерий факторизации). Для того чтобы $T(x)$ была достаточной статистикой для семейства распределения $F_\theta(x)$ (θ — неизвестный параметр), необходимо и достаточно, чтобы функция правдоподобия $L_\theta(x_1, ..., x_n)$ имела вид

$$L = \varphi(t(x) = T, \theta)h(x). \tag{2.5}$$

Из критерия следует, что, приводя к виду (2.5), получим вид ДС (по виду (2.5) ДС определяется неоднозначно, следовательно, достаточных статистик бесконечно много).

ДС существует, так как, например, выборка является ДС.

Доказательство теоремы 2.4

1. *Достаточность.* Пусть соотношение (2.5) выполнено, тогда покажем, что $T(x)$ является ДС, т.е. распределение любой выборки при ее фиксированном значении не зависит от параметра θ.

$$P\{X = \vec{x} / T(x) = T\} = \{\text{по теореме умножения вероятностей}\}=$$

$$= \frac{P(X = \vec{x}, T(x) = T)}{P(T(x) = T)} = \frac{L}{\sum_{\vec{y}: T(y) = T} L(\vec{y})} =$$

$$= \frac{\varphi(T(x) = T, \theta)h(x)}{\sum_{\vec{y}: T(y) = T} \varphi(T, \theta)h(y)} = \frac{h(x)}{\sum_{\vec{y}: T(y) = T} h(y)}$$

— не зависит от θ.

2. *Необходимость.* Пусть $T(x)$ — ДС. Тогда покажем, что (2.5) имеет место:

$$P(X = \vec{x} / T(x) = T) = \frac{P(X = \vec{x}, T(x) = T)}{P(T(x) = T)} = \frac{P(X = \vec{x})}{P(T(x) = T)} = h(x),$$

тогда по теореме умножения вероятностей имеем

$$P(X = \vec{x}) = L = P(X = \vec{x} / T(x) = T)P(T(x) = T) \equiv (6.5)$$

(так как $P(T(x) = T) = \varphi(T, \theta)$), а $P(X = \vec{x} / T(x) = T) = h(x)$ не зависит от θ и из определения ДС $T(x)$).

Рассмотрим пример применения критерия факторизации.

Задача 2.18. СВ $X \sim \pi(\lambda = \theta)$. Найти ДС для этого семейства.

Решение

$L = \prod_{i=1}^{n} P(X = x_i)$ — функция правдоподобия в дискретном случае;

$L = \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} = \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!} \Rightarrow T(x) = \sum_{i=1}^{n} x_i$ — ДС (так как $\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta} =$ $= \varphi(T(x), \theta)$) (проверка по определению проведена ниже).

Теорема 2.5 (Рао — Блэкуэлла — Колмогорова). Оптимальная оценка, если она существует, является функцией от достаточной статистики.

Доказательство. Пусть $T = T(x)$ — ДС для $F_\theta(x)$ и $T_1 = T(x)$ — несмещенная оценка (НО) для $\tau = \tau(\theta)$.

Рассмотрим статистику $H(t) = M(T_1/T)$ (при $T = t$ имеем $H(t) = M(T_1/T = t)$).

Тогда $MH(t) = M(M(T_1/T)) = MT_1 = \tau$ — НО для τ.

Теперь сравним $DH(t)$ с DT_1:

$$DT_1 = M(T_1 - MT_1)^2 = M(T_1 - \tau)^2 = M(T_1 - H(T) + H(T) - \tau)^2 =$$
$$= M(T_1 - H(T))^2 + 2M(T_1 - H(T))(H(T) - \tau) + M(H(T) - \tau)^2.$$

Если показать, что $M(T_1 - H(T))(H(T) - \tau) = 0$ (*), то, так как $M(H(T) - \tau)^2 \geqslant 0$, а $M(\tau - H(T))^2 = DH(T)$, получим $DT_1 \geqslant DH(T)$. Остается доказать (*):

$$M(T_1 - H(T))(H(T) - \tau) = M((T_1 - H(T))(H(T)) - \tau M(T_1 - H(T));$$
$$M(T_1 - H(T)) = MT_1 - MH(T) = \tau - \tau = 0.$$

По формуле полной вероятности при гипотезах $\{T = t\}$ получаем

$$M((T_1 - H(T))(H(T)) = \int_{-\infty}^{+\infty} (M(T_1/T = t) - H(t))H(t)g(t)dt = 0,$$

где $g(t)$ — плотность распределения статистики $H(t)$, так как $E(T/T = t) - H(t) = 0$.

Теорема доказана.

Полнота достаточной статистики

Определение 2.4. Достаточная статистика $T(x)$ называется ***полной***, если для любой функции $f(T(x)) = f(T)$ из того, что $Mf(T) = 0$ для всех θ, следует, что $f(T) = 0$ почти всюду (т.е. в отдельных областях она не нуль, но эти области имеют меру нуль).

Теорема 2.6. Если существует полная ДС, то любая функция от нее является оптимальной оценкой для своего математического ожидания.

Доказательство. Пусть $T = T(x)$ — полная ДС; $H(t)$ — произвольная функция от $T(x)$; $E_\theta H(T) = \tau(\theta)$.

Тогда из определения полноты ДС $T(x)$ следует единственность НО $H(T)$ для $\tau(\theta)$. Допустим, существует другая НО $H_1(T)$. Тогда $M(H(T) - H_1(T)) = 0 \Rightarrow H(T) = H_1(T)$ п.в. (из определения полноты ДС).

Из теоремы 2.5 следует, что оптимальную оценку следует искать в классе функций, зависящих от ДС, — $T(x)$. Но так как $H(T)$ — единственная НО для $\tau(\theta)$, зависимая от T, то она является оптимальной оценкой для $\tau(\theta)$.

Следствия из теорем

Пусть $T = T(x)$ — полная ДС; $\tau(\theta)$ — оцениваемая функция. Тогда:

1) когда оптимальная оценка существует, она является функцией ДС и однозначно определяется из уравнения несмещенности: $MH(T) = \tau(\theta)$;

2) оптимальная оценка, если она существует, ищется по формуле $\tau^* = H(T) = M(T_1/T = t)$, где T_1 — НО для $\tau(\theta)$.

Рассмотрим задачи.

Задача 2.19. Доказать, что:

а) любая взаимно однозначная функция $W(x)$ от ДС тоже является ДС;

б) любая выборка — ДС;

в) вариационный ряд — ДС;

г) эмпирическая функция распределения — ДС.

Решение (*доказательство*)

а) Пусть $T(x)$ — ДС, тогда $P(x/T(x) = t) = P(x/W(T(x)) = W(t))$, ч.т.д.;

б) $P\{x_1, ..., x_n / x_1', ..., x_n'\} = 1$, т.е. не зависит от θ, где под знаком вероятности стоит условная вероятность данной выборки $x_1, ..., x_n$, если она фиксирована: $x_1 = x_1', ..., x_n = x_n'$;

в) следует из а) и б);

г) следует из а) и в).

Задача 2.20. Для семейства распределений $\Pi(\theta)$ найти ДС и проверить ее по определению.

Решение

Функция правдоподобия имеет вид

$$L(x) = \frac{e^{-\theta}\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i}.$$

По КФ получим $T(x) = \sum_{i=1}^{n} x_i$ — ДС.

Проверка:

$$P\left\{\vec{x}/T(x)=\sum_{i=1}^{n}x_i=t\right\}=\frac{P\left\{X=x_1,\ ...,X=x_{n-1},X=t-\sum_{i=1}^{n-1}x_i\right\}}{P\left\{\sum_{i=1}^{n}x_i=t\right\}}=$$

$$=\frac{e^{-\theta}\theta^{x_1}}{x_1!}\cdots\frac{e^{-\theta}\theta^{t-\sum_{i=1}^{n-1}x_i}}{\left(t-\sum_{i=1}^{n-1}x_i\right)!}\frac{t!}{\theta^t n^t e^{-n\theta}}=\frac{t!}{n^t\left(t-\sum_{i=1}^{n-1}x_i\right)!\prod_{i=1}^{n-1}(x_i)!},$$

т.е. не зависит от θ.

Задача 2.21. Найти ДС для семейства распределений $R[0;\theta]$ и проверить ее по определению.

Решение

Функция правдоподобия имеет вид

$$L(x)=\underbrace{(1/\theta^n)W(\theta-x_{(n)})}_{\phi(T(x)=x_{(n)},\theta)}\underbrace{1}_{h(x)},$$

где $W(x)=\begin{cases}1, x\geqslant 0,\\ 0, x<0\end{cases}\Rightarrow x_{(n)}$ — ДС.

Проверка: $f(x_1, ..., x_n/x_{(n)}=t)=\dfrac{1}{t}$, т.е. не зависит от θ.

Задача 2.22. Найти ДС для семейства γ-распределений.

Решение

$$L(x)=\lambda^{n\alpha}\exp\left\{-\lambda\sum_{i=1}^{n}x_i\right\}(x_1,...,x_n)^{\alpha-1}\left(\frac{1}{\Gamma(\alpha)}\right)^n W(x_{(n)});$$

по КФ получим $T(x)=\sum_{i=1}^{n}x_i$ — ДС.

Определение 2.5. ***Экспоненциальным семейством*** (ЭС) называется семейство распределений с плотностью $f(x)$ или вероятностью $P(X=x)$ вида $\exp(A(\theta)B(x)+C(\theta)+D(x))$.

Задача 2.23. Найти ДС для ЭС.

Решение

$$L(x)=\exp\left\{A(\theta)\sum_{i=1}^{n}B(x_i)+nC(\theta)\right\}\exp\left\{\sum_{i=1}^{n}D(x_i)\right\},$$

по КФ получим $T(x)=\sum_{i=1}^{n}B(x_i)$ — ДС, так как

$$\exp\left\{A(\theta)\sum_{i=1}^{n}B(x_i)+nC(\theta)\right\}=\varphi(T(x=x,\theta));\exp\left\{\sum_{i=1}^{n}B(x_i)\right\}=h(x).$$

К ЭС относятся γ-распределение, биномиальное, пуассоновское, геометрическое, Паскаля и др. Путем приведения распределения к виду ЭС с учетом результата задачи 5 найдите ДС для перечисленных выше законов распределения.

Нахождение ДС путем приведения к экспоненциальному виду

Задача 2.24. СВ $X \sim \gamma(x;\, \alpha,\, \theta)$; найти ДС.

Решение

$$f(x;\alpha;\theta)=\begin{cases}\dfrac{\theta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-x\theta}, x\geqslant 0,\\ 0, x<0;\end{cases} \quad \theta \text{ — неизвестно.}$$

Воспользуемся введенной выше индикаторной функцией $J(z)$ и тем, что $\min\limits_{i=1,\ldots,n} x_i = x_{(1)}$:

$$L=\frac{\theta^{n\alpha}}{(\Gamma(\alpha))^n}(x_1,\ \ldots, x_n)^{\alpha-1}e^{-\theta\sum\limits_{i=1}^{n}x_i}J(x_{(1)})=$$

$$=\exp\left\{\ln\frac{\theta^{n\alpha}}{(\Gamma(\alpha))^n}(x_1,\ \ldots, x_n)^{\alpha-1}e^{-\theta\sum\limits_{i=1}^{n}x_i}J(x_{(1)})\right\}=$$

$$=\exp\left\{n\alpha\ \ln\theta-n\ln\Gamma(\alpha)+(\alpha-1)\sum_{i=1}^{n}\ln x_i-\theta\sum_{i=1}^{n}x_i+\ln J(x_{(1)})\right\}=$$

$$=\exp\left\{\underbrace{-\theta\sum_{i=1}^{n}x_i}_{A(Q)\sum B(x_i)}+\underbrace{n\alpha\ \ln\theta}_{C(Q)}+\underbrace{n\ln\ \Gamma(\alpha)+\ln X(x_{(1)})}_{D(X)}\right\}\Rightarrow$$

$\Rightarrow \sum\limits_{i=1}^{n} B(x_i) = \sum\limits_{i=1}^{n} x_i$ — ДС.

Задача 2.25. СВ $X \sim B(n,\, \theta)$; найти ДС.

Решение

$$f_0=P(X=x_i)=C_n^{x_i}\theta^{x_i}(1-\theta)^{n-x_i}=\exp[x\ln\theta+(n-x)\ln(1-\theta)+\ln C_n^{x_i}]=$$

$$=\exp\left\{\underbrace{x[\ln\theta-\ln(1-\theta)]}_{B(x)A(\theta)}+\underbrace{n\ln(1-\theta)}_{C(\theta)}+\underbrace{\ln C_n^{x_i}}_{D(x)}\right\}\Rightarrow\sum_{i=1}^{n}B(x_i)=\sum_{i=1}^{n}x_i-\text{ДС.}$$

Задача 2.26. СВ $X \sim \Gamma_{\alpha,\lambda}$ при $x \geqslant 0$; найти ДС.

Решение

$$f(x;\alpha;\theta)=\frac{\theta^{\alpha}x^{\alpha-1}e^{-x\theta}}{\Gamma(\alpha)}=\{x\geqslant 0\}=\exp[-\theta-(\alpha-1)\ln\theta+\alpha\ln\theta-\ln\,\Gamma(\alpha)]=$$

$$=\exp\left(\underbrace{-\theta x}_{A(\theta)B(x)}+\underbrace{\alpha\ln\theta}_{C(\theta)}-\underbrace{(\alpha-1)\ln x-\ln\,\Gamma(x)}_{D(x)}\right)\Rightarrow\sum_{i=1}^{n}x_i-\text{ДС}.$$

Задача 2.27. СВ $X \sim N(\theta, \sigma)$; найти ДС.

Решение

$$f_0(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(x-\theta)^2\right\}=$$

$$=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\underbrace{\frac{x^2}{2\sigma^2}}_{D(x)}+\underbrace{\frac{x\theta}{\sigma^2}}_{A(\theta)B(x)}-\underbrace{\frac{\theta^2}{2\sigma^2}}_{C(\theta)}\right\}\Rightarrow\sum_{i=1}^{n}x_i-\text{ДС}.$$

Задача 2.28. $X \sim \pi(\theta)$; найти ДС.

Решение

$$P(X=x)=\frac{e^{-\theta}\theta^{x}}{x!}=\{x=0,1,2,\ \ldots\}=$$

$$=\exp\left\{\underbrace{-\theta}_{C(\theta)}+\underbrace{x\ln\theta}_{B(x)A(\theta)}-\underbrace{\ln(x!)}_{D(x)}\right\}\Rightarrow\sum_{i=1}^{n}x_i-\text{ДС}.$$

Задача 2.29 (на полную ДС). СВ $X \sim B(1, \theta)$, где θ — неизвестно, $\theta \in (0; 1)$. Найти ДС и проверить ее на полноту.

Решение

Функция правдоподобия $L(x)$ имеет вид

$$L(x)=\underbrace{\theta^{\sum_{i=1}^{n}x_i}(1-\theta)^{n-\sum_{i=1}^{n}x_i}}_{\varphi(T(x)=\sum_{i=1}^{n}x_i,\theta)}\underbrace{1}_{h(x)}\Rightarrow r=\sum_{i=1}^{n}x_i-\text{ДС}.$$

СВ $r \sim B(n, \theta)$. Пусть $f(r)=f\left(\sum_{i=1}^{n}x_i\right)$ такова, что $M_\theta(f(r)) = 0$ при всех $\theta \in (0; 1)$, т.е. $M_\theta(f(r)) = \sum_{r=0}^{n} f(r)C_n^r\theta^r(1-\theta)^{n-r} = 0 \Rightarrow f(r) = 0$ почти всюду как коэффициенты многочлена, т.е. r — полная ДС.

Замечание 2.11. Необходимым и достаточным условием полной ДС, относящейся к ЭС, является совпадение размерности статистики и неизвестного параметра.

Задача 2.30 (на теорему 2.6). СВ $X \sim B(1, 0)$, $r = \sum_{i=1}^{n} x_i$. По теореме 2.6 установить, какие функции будут оптимальными оценками для θ: а) $\frac{r}{n}$; б) $\frac{r(r-1)}{n(n-1)}$.

Решение

а) $M\left(\frac{r}{n}\right) = M\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = \theta,$ т.е. для θ;

б) $M\left(\frac{r(r-1)}{n(n-1)}\right) = \frac{1}{n(n-1)} M\left[\left(\sum_{i=1}^{n} x_i\right)^2 - \left(\sum_{i=1}^{n} x_i\right)\right] = \frac{n(n-1)\theta^2}{n(n-1)} = \theta^2,$ так как

$$\left(\sum_{i=1}^{n} x_i\right)^2 = D\left(\sum_{i=1}^{n} x_i\right) + \left[M\left(\sum_{i=1}^{n} x_i\right)\right]^2 = n(1-\theta)\theta + n^2\theta^2 = n\theta - n\theta^2 + n^2\theta^2.$$

Задание 2.2. СВ $X \sim B(1, \theta)$, $r = \sum_{i=1}^{n} x_i$. По теореме 2.6 установить, будет ли оптимальной оценкой для θ функция $\frac{r(r-n)}{n(n-1)}$.

Задание 2.3. СВ $X \sim R[0; \theta]$ (θ — неизвестно). Показать, что $T(x) = x_{(n)}$ — полная ДС.

Задачи на полные и неполные достаточные статистики

Задача 2.31. СВ $X \sim R[\theta_1; \theta_2]$; $x_1, ..., x_n$ — выборка значений СВ X. Доказать, что $T(x) = (x_{(1)}, x_{(n)})$ — полная ДС.

Решение

$L = J(\theta_2 - x_{(n)}) \frac{J(x_{(1)} - \theta_1)}{(\theta_2 - \theta_1)^2}$, где $J(z) = \begin{cases} 0, z < 0, \\ 1, z \geqslant 0. \end{cases}$

По КФ получаем, что $T(x) = (x_{(1)}, x_{(n)})$ — ДС.

$f_{1,n}(u, v) = n(n-1)f(u)f(v)(F(v) - F(u))^{n-2}$ — плотность совместного распределения СВ $x_{(1)}$ и $x_{(n)}$; $F(x)$ — функция распределения СВ X:

$$F(x) = \frac{x - \theta_1}{(\theta_2 - \theta_1)^2} \text{ при } \theta_1 \leqslant x \leqslant \theta_2;$$

$$f_{1,n}(u, v) = \frac{n(n-1)}{(\theta_2 - \theta_1)^n}(v-u)^{n-2}, \theta_1 \leqslant u \leqslant v \leqslant \theta_2.$$

Пусть $g(t_1, t_2)$ такова, что для всех $\theta_1 < \theta_2$ выполняется соотношение

$$Mg(t_1,t_2) = \iint\limits_{D:\theta_1 \leqslant t_1 < t_2 \leqslant \theta_2} g(t_1,t_2) f_{1,n}(t_1,t_2)dt_1dt_2 = \iint\limits_{D} g(t_1,t_2)(t_2-t_1)^{n-2}dt_1dt_2 = 0$$

для всех $\theta_1 < \theta_2$.

Продифференцировав последнее выражение сначала по t_1, потом по t_2, получим $g(t_1, t_2) = 0$, откуда следует полнота ДС $T(x)$.

Задача 2.32. СВ $X \sim \pi(\lambda)$; $x_1, ..., x_n$ — выборка значений СВ X. Доказать, что статистика $T(x) = S = \left(S_{1l} = \sum_{i=1}^{l} x_i, S_{2l} = \sum_{i=l+1}^{n} x_i \right)$ — ДС, но неполная.

Решение

Функция правдоподобия $L(x)$ имеет вид

$$L = \prod_{i=1}^{n} \frac{\lambda^{x_i} \exp(-\lambda)}{x_i\,!} = \frac{\lambda^{S_{1l}} \lambda^{S_{2l}} \exp(-n\lambda)}{\prod_{i=1}^{n} x_i\,!} \Rightarrow T(x) \text{ — ДС по КФ.}$$

Докажем неполноту $T(x)$. Функция $U(S) = \dfrac{S_{1l}(n-1)}{1-S} \neq 0$, в то время как $MU(S) = 0 \Rightarrow T(x) = S$ — неполная ДС.

Задача 2.33. СВ $X \sim R[a; b] = R_1[a_1\theta + b_1, a_2\theta + b_2]$. Доказать, что статистика $T(x) = (x_{(1)}, x_{(n)})$ — достаточная, но неполная.

Решение

Ранее получено, что $Mx_{(1)} = a + \dfrac{b-a}{n+1}$, $Mx_{(n)} = b - \dfrac{b-a}{n+1}$. Тогда

$$Mx_{(1)} = \theta \frac{a_1 n + a_2}{n+1} + \frac{b_1 n + b_2}{n+1}; \; Mx_{(n)} = \theta \frac{a_2 n + a_1}{n+1} + \frac{b_2 n + b_1}{n+1}.$$

Выразим θ из двух последних выражений, заменяя $Mx_{(1)}$ на $x_{(1)}$, а $Mx_{(n)}$ на $x_{(n)}$. Тогда при взятии среднего от их разности, очевидно, получим нулевое значение:

$$W(T(x)) = \frac{(n+1)x_{(1)} - (b_1 n + b_2)}{a_1 n + a_2} - \frac{(n+1)x_{(n)} - (b_2 n + b_1)}{a_2 n + a_1} = 0,$$

в то время как $M(T(x)) = 0 \Rightarrow T(x)$ — неполная ДС.

Задание 2.4. СВ $X \sim B(1, \theta)$; $T(x) = \left(S_{1l} = \sum_{i=1}^{l} x_i, S_{2l} = \sum_{i=l+1}^{n} x_i \right)$. Доказать, что статистика $T(x) = (x_{(1)}, x_{(n)})$ — достаточная, но неполная.

Задание 2.5. СВ $X \sim R\left[\theta - \dfrac{1}{2}; \theta + \dfrac{1}{2}\right]$; $T(x) = (x_{(1)}, x_{(n)})$. Доказать, что статистика $T(x) = (x_{(1)}, x_{(n)})$ — достаточная, но неполная (решить с использованием задачи 15 и непосредственно).

Задание 2.6. СВ $X \sim R[0; \theta]$. Доказать полноту ДС $T(x) = x_{(n)}$.

Задание 2.7. СВ $X \sim R[\theta_1; \theta_2]$. Используя теорему 2.6 (о полных ДС), определить, для каких функций от неизвестных параметров θ_1, θ_2 будут оптимальными оценки: а) $\dfrac{x_{(1)} + x_{(n)}}{2}$; б) $(n+1)\dfrac{x_{(1)} - x_{(n)}}{n-1}$.

ДС могут быть использованы для улучшения имеющихся НО неизвестного параметра θ в соответствии с теоремой 2.4. Рассмотрим примеры.

Задача 2.34. СВ $X \sim \pi(\lambda)$; $\theta = x_1, x_2, ..., x_n$ — выборка значений СВ X. Улучшить оценку для θ.

Решение

$$MT_1 = Mx_1 = \theta \Rightarrow \theta = T_1 \text{ — НО.} \quad T(x) = \sum_{i=1}^{n} x_i \text{ — ДС.}$$

$$H(t) = \left(M \frac{x_1}{\sum_{i=1}^{n} x_i} = t \right) = \sum_{k=0}^{t} kP\left\{ x_1 = \frac{k}{\sum_{i=1}^{n} x_i} = t \right\} = \sum_{k=0}^{t} \frac{kP\left\{ x_1 = k, \sum_{i=2}^{n} x_i = t-k \right\}}{P\left\{ \sum_{i=1}^{n} x_i = t \right\}} =$$

$$= \sum_{k=0}^{t} k \frac{\theta^k \exp(-\theta)}{k!} \frac{\theta^{t-k}(n-1)^{t-k} \exp[-(n-1\theta)]}{(t-1)!} \frac{t!}{\theta^t n^t \exp[-n\theta]} =$$

$$= \sum_{k=0}^{t} kC_t^k \left(\frac{1}{n}\right)^k \left(1-\frac{1}{n}\right)^{t-k} = \frac{t}{n}.$$

Так как последняя сумма равна MZ, где $Z \sim B\left(t, \dfrac{1}{n}\right)$, то

$$DH(t) = D\left(\frac{t}{n}\right) = \frac{D\left(\sum_{i=1}^{n} x_i\right)}{n} = \frac{n\theta}{n^2} = \frac{\theta}{n}; \; DT_1 = Dx_1 = \theta \Rightarrow \begin{cases} DT_1 = DH(t), n = 1, \\ DT_1 > DH(t), n > 1. \end{cases}$$

Задание 2.8. Улучшить оценку для θ в случаях:
а) СВ $X \sim B(1, \theta)$; $T_1 = x_1$;
б) СВ $X \sim R[0; \theta]$; $T_1 = 2x_1$.

2.4. Неравенство Рао — Крамера

Сформулируем ***условия регулярности*** семейства распределений $x \sim F_\theta(x)$:

1) дифференцируемость L по θ;

2) множество $\{x: L \neq 0\}$ не зависит от неизвестного параметра θ.

Рассмотрим задачу точечного оценивания при $x \sim F_\theta(x)$, $\tau = \tau(\theta)$. Если выполняются условия регулярности и $t(x)$ — несмещенная оценка для $\tau = \tau(\theta)$, то дисперсия не может быть как угодно малой при построении разных оценок, т.е.

$$Dt(x) \geqslant \frac{(\tau')^2}{M\left(\dfrac{d \ln L}{d\theta}\right)^2}. \tag{2.6}$$

Это неравенство называется неравенством Рао — Крамера, и оно определяется следующей теоремой.

Теорема 2.7. Неравенство Рао — Крамера дает нижнюю грань дисперсий несмещенных оценок в регулярном случае.

Доказательство. По теореме о коэффициенте корреляции имеем $|r_{xy}| \leqslant 1$, причем $|r_{xy}| = 1$ тогда и только тогда, когда СВ X и Y линейно зависимы. Рассмотрим неравенство $|r_{xy}| \leqslant 1$ при условии $MX = 0$ и $MY = 0$. Получим $|r_{xy}| = \dfrac{|K_{xy}|}{\sqrt{DX \cdot DY}} \leqslant 1$ (по определению r_{xy}), тогда $|K_{xy}| \leqslant \sqrt{DX \cdot DY}$. Имеем $K_{xy} = MXY - MX \cdot MY = MXY$; $DX = MX^2 - (MX)^2 = MX^2$ и $DY = MY^2$. Подставляя эти значения, имеем:

$$|MXY| \leqslant \sqrt{MX^2 \cdot MY^2}, \text{ или } (MXY)^2 \leqslant MX^2 \cdot MY^2. \tag{2.7}$$

Эта запись является формой неравенства Коши — Буняковского (НКБ).

План доказательства:

1) выписать условие нормировки и продифференцировать его по θ;

2) для несмещенной оценки $t(x)$ для $\tau = \tau(\theta)$ выписать уравнение несмещенности и продифференцировать обе части этого уравнения по θ;

3) из результата п. 2 вычесть результат п. 1, умноженный на τ;

4) применить НКБ в форме (*).

Проводим доказательство НРК по плану.

1. $\int\limits_{R^n} L d\vec{x} = 1 \Rightarrow$

$$\Rightarrow 0 = (1)' = \int\limits_{R^n} \left(\frac{dL}{d\theta} \cdot \frac{1}{L} \right) \cdot L d\vec{x} = \int\limits_{R^n} \frac{d \ln L}{d\theta} \cdot L d\vec{x} = M\left(\frac{d \ln L}{d\theta} \right),$$

т.е.

$$M\left(\frac{d \ln L}{d\theta} \right) = \tau'. \tag{2.8}$$

2. $t = t(x)$ — несмещенная оценка для $\tau = \tau(\theta)$. Выпишем уравнение несмещенности: $\int\limits_{R^n} tL d\vec{x} = \tau$ и продифференцируем его по θ:

$$\tau' = \int\limits_{R^n} t \frac{dL}{d\theta} \frac{1}{L} L dx = \int\limits_{R^n} t \frac{d \ln L}{d\theta} L dx = M\left(t \frac{d \ln L}{d\theta} \right),$$

т.е.

$$M\left(t \frac{d \ln L}{d\theta} \right) = \tau'. \tag{2.9}$$

3. Обозначим $u = \dfrac{d \ln L}{d\theta}$ и вычтем из уравнения (2.9) уравнение (2.8), умноженное на τ, получим $M(ut) - \tau Mu = \tau'$. Воспользуемся свойствами математического ожидания MX:

$$M(ut - \tau u) = M(u(t - \tau)) = \tau',$$

возведем обе части равенства в квадрат:

$$(M(u(t-\tau)))^2 = (\tau')^2.$$

4. Воспользуемся НКБ в форме (2.7) (это возможно, так как $Mu = 0$ по формуле (2.8), а $M(t-\tau) = 0$, так как t — несмещенная оценка для τ):

$$(\tau')^2 \leqslant Mu^2 M(t-\tau)^2,$$

но так как $M(t-\tau)^2 = Mt^2 - (Mt^2) = Dt$, то $Dt \leqslant \dfrac{(\tau')^2}{Mu^2}$, но $u = \dfrac{d\ln L}{d\theta}$, поэтому $Dt \leqslant \dfrac{(\tau')^2}{M\left(\dfrac{d\ln L}{d\theta}\right)^2}$, ч.т.д.

Определение 2.6. Если в неравенстве (2.6) достигается равенство, то оценка $t(x)$ называется несмещенной оценкой с минимальной дисперсией (НОМД) или ***эффективной оценкой***.

Критерий НОМД (критерий эффективности)

Дисперсия будет наименьшей, если знак неравенства заменить знаком равенства, а это по НКБ возможно только в том случае, если $u = \dfrac{d\ln L}{d\theta}$ и $t - \tau$ линейно зависимы, т.е. искомый критерий формулируется следующим образом: $\dfrac{d\ln L}{d\theta} = a(\theta)(t-\tau) + b$. Покажем, что $b = 0$. Действительно, возьмем математическое ожидание от обеих частей последнего равенства. Тогда $M\dfrac{d\ln L}{d\theta} = 0$ по формуле (2.8), а $M(t-\tau) = 0$, так как t — несмещенная оценка для τ, откуда получаем, что $b = 0$, тогда критерий НОМД окончательно имеет вид

$$\frac{d\ln L}{d\theta} = a(\theta)(t-\tau). \tag{2.10}$$

В этом случае получим более простой вид дисперсии:

$$Dt = \frac{(\tau')^2}{M\left(\dfrac{d\ln L}{d\theta}\right)^2}. \tag{2.11}$$

Возведем обе части равенства (2.10) в квадрат и возьмем математическое ожидание от обеих частей:

$$M\left(\frac{d\ln L}{d\theta}\right)^2 = a^2(\theta)M(t-\tau)^2 = a^2(\theta)Dt.$$

Тогда из формулы (2.10) имеем $Dt = \dfrac{(\tau')^2}{a^2(\theta)Dt}$, или $(Dt)^2 = \dfrac{(\tau')^2}{a^2(\theta)}$, т.е.

$$Dt = \left| \frac{\tau'}{a(\theta)} \right|. \qquad (2.12)$$

Теорема 2.8. Если существует НОМД для θ, то НОМД существует для любой линейной функции от θ и не существует ни для какой другой функции от θ.

Доказательство. Пусть $t(x)$ — НОМД для θ, а $\tau = A\theta + B$. Используем критерий НОМД:

$$\frac{d\ln L}{d\theta} = a(\theta)(t-\theta) = \frac{a(\theta)}{A}(At - A\theta) = \frac{a(\theta)}{A}(At + B - A\theta - B) =$$
$$= a_1(\theta)\cdot[At + B - (A\theta + B)].$$

По критерию НОМД получим, что $At + B$ — НОМД для $\tau = A\theta + B$;

$$D(At+B) = A^2 Dt = A^2 \left| \frac{\tau'}{a(\theta)} \right|.$$

Теперь докажем отсутствие НОМД для любой другой функции.

Пусть $t = t(x)$ — НОМД для $\tau = \tau(\theta)$, $\tau_1 = \tau_1(\theta) = \varphi(\tau)$ — нелинейная функция от τ.

Предположим противное, т.е. t_1 — НОМД для τ_1. Тогда по формуле (2.10) имеем

$$\frac{d\ln L}{d\theta} = a(\theta)(t-\tau); \qquad (а)$$

$$\frac{d\ln L}{d\theta} = a_1(\theta)(t_1 - \tau_1) = a_1(\theta)(t_1 - \varphi(\tau)), \qquad (б)$$

где $a(\theta) = a_1(\theta)\psi(\theta)$, тогда из (б) получаем

$$\frac{d\ln L}{d\theta} = \left(\frac{a(\theta)}{\psi(\theta)}\right)(t_1 - \varphi(\tau)) = a(\theta)\left(\frac{t_1}{\psi(\theta)} - \frac{\varphi(\tau)}{\psi(\theta)}\right). \qquad (в)$$

Из (а) и (в) $\Rightarrow t - \tau = \dfrac{t_1}{\psi(\theta)} - \dfrac{\varphi(\tau)}{\psi(\theta)}$, поэтому, так как t и t_1 зависят только от $\vec{x}$, а τ и $\varphi(\tau)$ — только от θ, то $\psi(\theta) = C_1$ — const и $t = \dfrac{t_1}{C_1} + C_2$, где C_2 — const. Тогда

$$\frac{d\ln L}{d\theta} = a(\theta)\left(t - C_2 - \frac{\varphi(\tau)}{C_1}\right) = a(\theta)(t-\tau) \Rightarrow$$

$\Rightarrow C_1 t - C_1\tau - t - C_1C_2 = \varphi(t) \Rightarrow \varphi(t)$ — линейная функция от τ, а это противоречит предположению противного и значит, что утверждение теоремы доказано.

Практически это утверждение применяется в следующем случае: если получен вид (2.10) $\dfrac{d\ln L}{d\theta} = a(\theta)(t-\tau)$ и требуется ответить на вопрос о существовании НОМД для какой-либо функции от θ, отличной от линейной; ответ, очевидно, отрицательный.

Замечание 2.12. Если построена оценка $t(x)$ — НОМД для $\tau(\theta)$, то для $\tau_1(\theta) = A\tau(\theta) + B$ НОМД $t_1(x) = At(x) + B$.

В более общем случае при выводе критерия можно говорить о $\tau(\theta)$.

Замечание 2.13. Если для θ НОМД существует, то ее дисперсию можно получить из неравенства Рао — Крамера, заменяя знак неравенства равенством, или по критерию НОМД (формула (2.11)).

Если НОМД не существует, то смысл неравенства Рао — Крамера состоит в том, что оно дает нижнюю грань дисперсии, которая не достигается.

Задача 2.35. Проверить, является ли оценка $\widehat{\theta} = \dfrac{\overline{x}}{m} = \dfrac{\sum_{i=1}^{n} x_i}{nm}$ НОМД для СВ $X \sim B(m, \theta)$, где θ — вероятность успеха в первом опыте и $\tau(\theta) = \theta$.

Решение

Способ 1. Известно, что $MX = m\theta$, $DX = m\theta(1 - \theta)$, тогда найдем $D\widehat{\theta}$:

$$D\widehat{\theta} = \frac{D\left(\sum_{i=1}^{n} x_i\right)}{m^2 n^2} = \frac{nm\theta(1-\theta)}{m^2 n^2} = \frac{\theta(1-\theta)}{mn}.$$

Покажем несмещенность $\widehat{\theta}$: $M\widehat{\theta} = \dfrac{1}{mn} mn\theta = \theta$. Вычислим нижнюю грань дисперсии несмещенных оценок по неравенству Рао — Крамера:

$$L = \prod_{i=1}^{n} P(x = x_i) = \prod_{i=1}^{n} C_m^{x_i} \theta^{x_i} (1-\theta)^{m-x_i} = \left(\prod_{i=1}^{n} C_m^{x_i}\right) \cdot \theta^{\sum_{i=1}^{n} x_i} \cdot (1-\theta)^{nm - \sum_{i=1}^{n} x_i}.$$

$$\ln L = \sum_{i=1}^{n} \ln C_m^{x_i} + \sum_{i=1}^{n} x_i \ln\theta + \left(mn - \sum_{i=1}^{n} x_i\right) \ln(1-\theta),$$

тогда

$$\frac{d\ln L}{d\theta} = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{mn - \sum_{i=1}^{n} x_i}{1-\theta} = \frac{\sum_{i=1}^{n} x_i - \theta\sum_{i=1}^{n} x_i - mn\theta + \theta\sum_{i=1}^{n} x_i}{\theta(1-\theta)} = \frac{\sum_{i=1}^{n} x_i - mn\theta}{\theta(1-\theta)}.$$

Получим

$$M\left(\frac{d\ln L}{d\theta}\right)^2 = \frac{1}{\theta^2(1-\theta)^2} \cdot M\left(\sum_{i=1}^{n} x_i - nm\theta\right)^2 =$$

$$= M\left(\frac{d\ln L}{d\theta}\right)^2 = \frac{1}{\theta^2(1-\theta)^2}\left(M\left(\sum_{i=1}^{n} x_i\right)^2 - 2mn\theta M\sum_{i=1}^{n} x_i + n^2 m^2 \theta^2\right).$$

Пользуясь тем, что

$$M\left(\sum_{i=1}^{n} x_i\right)^2 = D\left(\sum_{i=1}^{n} x_i\right)^2 + \left(M\sum_{i=1}^{n} x_i\right)^2;$$

$$D\sum_{i=1}^{n} x_i = nm\theta(1-\theta);\ \left(M\sum_{i=1}^{n} x_i\right)^2 = n^2m^2\theta^2,$$

получим

$$M\frac{d\ln L}{d\theta} = \frac{1}{\theta^2(1-\theta)^2}(mn\theta - mn\theta^2 + m^2n^2\theta^2 - 2m^2n^2\theta^2 + m^2n^2\theta^2) = \frac{mn}{\theta(1-\theta)}.$$

Теперь, применяя неравенство Рао — Крамера и учитывая, что $\tau' = 1$, получим

$$Dt \geqslant \frac{\theta(1-\theta)}{mn}.$$

Видим, что $\hat{\theta}$ — НОМД, так как результат вычисления $D\hat{\theta}$ совпадает с нижней гранью из неравенства Рао — Крамера.

Способ 2. Получим эту же оценку по критерию НОМД.

$$\frac{d\ln L}{d\theta} = \frac{\sum_{i=1}^{n} x_i - mn\theta}{\theta(1-\theta)} \text{ (см. вычисления выше).}$$

Приведем это выражение к виду (2.10):

$$\frac{d\ln L}{d\theta} = \frac{mn}{\theta(1-\theta)}\left(\frac{\sum_{i=1}^{n} x_i}{mn} - \theta\right),$$

тогда $a(\theta) = \dfrac{mn}{\theta(1-\theta)}$ и $t(x) = \dfrac{\sum_{i=1}^{n} x_i}{mn}$ — НОМД. Учитывая что $\tau' = 1$, применим формулу (2.12) и получим, что

$$Dt = \left|\frac{\tau'}{a(\theta)}\right| = \frac{1}{a} = \frac{\theta(1-\theta)}{mn}.$$

Критерий НОМД для ЭС

Пусть $\left.\begin{matrix} f(x) \\ P(X)=x \end{matrix}\right\} = \exp(A(\theta) + B(x) + C(\theta) + D(x))$. Тогда

$$L = L_\theta(\vec{x}) = \exp\left\{A(\theta)\sum_{i=1}^{n} B(x_i) + nC(\theta) + \sum_{i=1}^{n} D(x_i)\right\};$$

$$\frac{d\ln L}{d\theta} = \sum_{i=1}^{n} B(x_i)\frac{dA(\theta)}{d\theta} + n\frac{dC(\theta)}{d\theta} = \sum_{i=1}^{n} B(x_i)A'(\theta) + nC'(\theta) =$$

$$= \frac{A'(\theta)}{n}\left(\frac{\sum B(x_i)}{n} - \frac{C'(\theta)}{A'(\theta)}\right),$$

следовательно, по критерию НОМД $t^*(\vec{x})=\frac{1}{n}\sum B(x_i)$ — НОМД для

$$\tau^*=\tau(\theta)=\frac{C'(\theta)}{A'(\theta)}, \tag{2.13}$$

тогда по формуле (2.12) $D\tau^*=\left|\frac{\tau^{*\prime} n}{A'(\theta)}\right|$, а значит, НОМД существует и для любой линейной функции от τ.

Теорема 2.9. Если НОМД для $\tau = \tau(\theta)$ существует, то распределение СВ X относится к экспоненциальному семейству.

Доказательство. Пусть $t = t(\vec{x})$ — НОМД для $\tau = \tau(\theta)$. Тогда $\frac{d\ln L}{d\theta}=$ $= a(\theta)(t-\tau)$. Проинтегрировав по θ обе части последнего равенства, получим

$$\ln L = A(\theta)t(\vec{x}) + C(\theta) + D(x), \text{ ч.т.д.}$$

Рассмотрим примеры функций $\tau = \tau(\theta)$, для которых существует НОМД при следующих распределениях.

1. СВ $X \sim B(1, \theta)$:

$$P(X=x_i)=\theta^{x_i}(1-\theta)^{1-x_i}=\exp[x_i\ln\theta+(1-x_i)\ln(1-\theta)]=$$

$$=\exp\left\{\underbrace{x_i}_{B(x_i)}\underbrace{(\ln\ \theta-\ln(1-\theta))}_{A(\theta)}+\underbrace{\ln(1-\theta)}_{C(\theta)}\right\}\Rightarrow$$

$$\Rightarrow\tau^*=\frac{C'(\theta)}{A'(\theta)}=-\frac{\frac{1}{1-\theta}}{\frac{1-\theta}{\theta}\frac{1}{(1-\theta)^2}}=-\theta;$$

$D\tau^*=\frac{1}{\theta(1-\theta)}\Rightarrow$ для $\tau = \tau(\theta) = A\theta + B$, где A и B — const, существует НОМД.

2. СВ $X \sim B(m, \theta)$:

$$P(X=x_i)=C_m^{x_i}\theta^{x_i}(1-\theta)^{m-x_i}=\exp[\ln C_m^{x_i}+x_i\ln\theta+(m-x_i)\ln(1-\theta)]=$$

$$=\exp\left\{\underbrace{x_i}_{B(x_i)}\underbrace{(\ln\ \theta-\ln(1-\theta))}_{A(\theta)}+\underbrace{m\ln(1-\theta)}_{C(\theta)}+\underbrace{C_m^{x_i}}_{D(x_i)}\right\}\Rightarrow$$

$$\Rightarrow\tau^*=\frac{C'(\theta\ )}{A'(\theta\ )}=-\frac{\frac{m}{1-\theta}}{\frac{1-\theta}{\theta}\frac{1}{(1-\theta)^2}}=-m\theta;$$

$D\tau^* = \dfrac{m}{\theta(1-\theta)} \Rightarrow$ для $\tau = \tau(\theta) = A\theta + B$, где A и B — const, существует НОМД.

3. СВ $X \sim \pi(\theta)$:

$$P(X = x_i) = \frac{\theta^{x_i} e^{-\theta}}{x_i!} = \exp\left\{ \underbrace{x_i}_{B(x_i)} \underbrace{\ln\,\theta}_{A(\theta)} - \underbrace{\theta}_{C(\theta)} - \underbrace{\ln(x_i!)}_{D(x_i)} \right\} \Rightarrow \tau^* = \frac{C'(\theta)}{A'(\theta)} = \theta;$$

$D\tau^* = \dfrac{1}{1/\theta} = \theta \Rightarrow$ для $\tau = A\theta + B$, где A и B — const, существует НОМД.

4. СВ $X \sim \Gamma(\theta)$:

$$P(X = x_i) = \theta(1-\theta)^{x_i - 1} = \exp[\ln\,\theta + (x_i - 1)\ln(1-\theta)] =$$

$$= \exp\left\{ \underbrace{x_i}_{B(x_i)} \underbrace{\ln(1-\theta)}_{A(\theta)} + \underbrace{\ln\,\theta - \ln(1-\theta)}_{C(\theta)} \right\} \Rightarrow \tau^* = \frac{C'(\theta)}{A'(\theta)} = \frac{1-\theta}{\theta} \cdot \frac{(1-\theta)}{(1-\theta)^2} = \frac{1}{\theta};$$

$D\tau^* = \dfrac{1}{\theta^2} \dfrac{1}{1-\theta} = \dfrac{1}{\theta^2(1-\theta)} \Rightarrow$ для $\tau = \dfrac{A}{\theta} + B$, где A и B — любые const, существует НОМД.

2.5. Методы получения точечных оценок

Постановка задачи. Пусть имеется некоторая выборка $\vec{x}$: x_1, x_2, ..., x_n, $X \sim F_\theta(x)$, $\theta \in \Theta$, где Θ — параметрическое пространство. Требуется построить оценку для θ или $\tau(\theta)$.

Различаются два подхода к решению этой задачи в зависимости от понимания природы неизвестного параметра:

- *1-й подход* реализуется в случае, когда θ является неизвестной постоянной, т.е. θ = const. В данной ситуации используются ***методы подстановки***;
- *2-й подход* реализуется в случае, когда θ является случайной величиной, т.е $\theta \sim H_\theta(t)$. В данной ситуации используются ***байесовские оценки***.

2.5.1. Методы подстановки

Суть методов. Выбирается некоторая мера расхождения теоретического и эмпирического распределения и строится некоторый функционал от этой меры. Оценка для неизвестного параметра ищется таким образом, чтобы этот функционал принимал некоторое значение, соответствующее минимуму расхождения теоретических и практических результатов.

Метод подстановки объединяет ряд конкретных методов, которые различаются по мере различия теории и практики:

1) метод моментов (ММ);
2) метод максимального правдоподобия (ММП);
3) метод минимального χ^2;
4) метод минимального расстояния.

Существуют и другие методы, но мы рассмотрим только эти, так как они используются наиболее часто.

Метод моментов

Совокупность неизвестных параметров будем рассматривать как k-мерный вектор $\theta = (\theta_1, \theta_2, ..., \theta_k)$. Тогда оценки метода моментов (ОММ) являются решениями системы k уравнений, составленных путем приравнивания k теоретических моментов соответственно k эмпирическим.

Достоинством ММ является простота его применения.

Недостатком ММ является то, что он не гарантирует качества оценок, хотя часто оценки, полученные этим методом, являются «удачными».

Рассмотрим примеры ОММ.

Пример 2.9. СВ $X \sim \pi(\theta)$. Тогда $\tilde{\theta}$ для неизвестного параметра θ: $\tilde{\theta} = EX = \overline{x}$.

Пример 2.10. СВ $X \sim N(a, \sigma)$, тогда $\theta = (\theta_1, \theta_2)$, где $\theta_1 = a$, $\theta_2 = \sigma^2$. Найдем ОММ для θ:

$$\begin{cases} \tilde{a} = \overline{x} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i, \\ \tilde{\sigma}^2 = S_2^1 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - \overline{x})^2, \end{cases}$$

отсюда $\tilde{\theta} = (\overline{x}, S_2^2)$.

Пример 2.11. СВ $X \sim B(m, p = \theta)$. Найдем ОММ для θ (m — известно):

$MX = m\theta = \overline{x}$, следовательно, $\tilde{\theta} = \dfrac{\overline{x}}{m} = \dfrac{\sum\limits_{i=1}^{n} x_i}{mn}$.

Пример 2.12. СВ $X \sim R[0; \theta]$. Найдем ОММ для θ:

$EX = \dfrac{\tilde{\theta}}{2} = \overline{x}$, тогда ОММ $\tilde{\theta} = 2\overline{x} = \dfrac{2\sum\limits_{i=1}^{n} x_i}{n}$.

Метод максимального правдоподобия

Оценка максимального правдоподобия (ОМП) $\hat{\theta}$ выбирается таким образом, чтобы функция правдоподобия $L = L_\theta(x)$ принимала наибольшее значение.

Различается два случая нахождения ОМП: регулярный и нерегулярный.

1. Регулярный случай.

Пусть выполнены условия регулярности, сформулированные нами в начале параграфа 2.4. Они обеспечивают достижение максимума функции $L(x)$, поэтому в регулярном случае ОМП ищется из следующей системы уравнений максимального правдоподобия (УМП):

$$\frac{\partial L}{\partial \theta_j} = 0, j = 1, ..., k,$$

или из системы уравнений

$$\frac{\partial \ln L}{\partial \theta_j} = 0, j = 1, ..., k,$$

так как функции L и $\ln L$ достигают максимума в тех же точках.

Пример 2.13. СВ $X \sim \pi(\theta)$. Найдем ОМП для θ:

$$P(X = x_i) = \frac{\theta^{x_i} e^{-\theta}}{x_i!},$$

тогда

$$L = \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!};\ \ln L = \sum_{i=1}^{n} x_i \ln \theta - n\theta - \sum_{i=1}^{n} \ln x_i!;\ \frac{d \ln L}{d\theta} = \frac{\sum_{i=1}^{n} x_i}{\theta} - n.$$

Решим УМП и получим, что $\hat{\theta} = \sum_{i=1}^{n} x_i = \overline{x}.$

В данном случае получили, что ОМП совпадает с ОММ.

Очевидно, что вычисления по ММП значительно сложнее, чем по ММ, но ММП теоретически обосновывает гарантии качества оценок.

Теорема 2.10. Если существует НОМД для $\tau(\theta)$, то в регулярном случае она является ОМП для $\tau(\theta)$.

Доказательство. Используем критерий НОМД: $\frac{\partial \ln L}{\partial \theta} = a(\theta)(t(x) - \tau(\theta))$, где $t(x)$ — НОМД для $\tau(\theta)$. Тогда если $\frac{\partial \ln L}{\partial \theta} = 0$, то $t(x)$ — ОМП для $\tau(\theta)$.

Приведем пример функции от неизвестного параметра θ, для которой не существует НОМД, а ОМП существует.

Пример 2.14. Пусть СВ $X \sim G(\theta)$, $\tau(\theta) = \theta$. Найдем ОМП для $\tau(\theta) = \theta$:

$$P(X = x_i) = \theta(1-\theta)^{x_i - 1},$$

тогда

$$L = \theta^n (1-\theta)^{\sum\limits_{i=1}^n x_i - n}; \quad \ln L - n \ln\theta + \left(\sum_{i=1}^n x_i - n\right)\ln(1-\theta);$$

$$\frac{d\ln L}{d\theta} = \frac{n}{\theta} - \frac{\sum\limits_{i=1}^n x_i - n}{1-\theta} = \frac{n - n\theta - \theta\sum\limits_{i=1}^n x_i + n\theta}{\theta(1-\theta)} = \frac{n - \theta\sum\limits_{i=1}^n x_i}{\theta(1-\theta)} =$$

$$= \frac{n}{1-\theta}\left(\frac{1}{\theta} - \frac{\sum\limits_{i=1}^n x_i}{n}\right) = \frac{n}{1-\theta}\left(\frac{1}{\theta} - \overline{x}\right).$$

Найдем ОМП из УМП:

$$\frac{n - \theta\sum\limits_{i=1}^n x_i}{\theta(1-\theta)} = 0,$$

откуда следует, что $n = \theta\sum\limits_{i=1}^n x_i$. Тогда $\widehat{\theta} = \dfrac{1}{x}$ — ОМП.

Воспользуемся критерием НОМД:

$$\frac{d\ln L}{d\theta} = \frac{n - \theta\sum\limits_{i=1}^n x_i}{\theta(1-\theta)} = \frac{n\theta}{\theta(1-\theta)} = \left(\frac{1}{\theta} - \frac{\sum\limits_{i=1}^n x_i}{n}\right) = -\frac{n\theta}{\theta(1-\theta)}\left(\overline{x} - \frac{1}{\theta}\right),$$

следовательно, по критерию НОМД имеем $\overline{x}$ — НОМД для $\dfrac{1}{\theta}$, тогда для $\tau(\theta) = \theta$ НОМД не существует.

По теореме 2.10 получаем, что если НОМД для $\tau(\theta)$ существует, то ОМП совпадает с ней. Это и является гарантией качества ОМП.

Теорема 2.11. Все решения УМП (в регулярном случае), т.е. ОМП, являются функциями достаточной статистики.

Доказательство. Пусть $\widehat{\theta}$ — достаточная статистика, $g_\theta(t)$ — плотность ее распределения, а $f(x_k)$ — плотность распределения исследуемой СВ в точке x_k. Имеем

$$L = \prod_{k=1}^n f(x_k) = \prod_{k=1}^n f(x_k / \widehat{\theta} = t\, g_\theta(t),$$

тогда

$$\ln L = \sum_{k=1}^n \ln f(x_k / \widehat{\theta} = t) + n \ln g_\theta(t),$$

следовательно,

$$\frac{d\ln L}{d\theta}=\frac{\partial\sum_{k=1}^{n}\ln f(x_k/\theta=t)}{\partial\theta}+n\frac{\partial\ln g_\theta(t)}{\partial\theta}=0.$$

УМП принимает вид $n\frac{d\ln g_\theta(t)}{d\theta}=0$, так как $\sum_{k=1}^{n}\ln f(x_k/\hat{\theta}=t)$ не зависит от θ, в силу того что $\hat{\theta}$ — достаточная статистика и производная от этой суммы равна нулю.

Решение полученного УМП будет находиться в терминах достаточной статистики, что и означает выполнение утверждения теоремы.

Теорема 2.12 (инвариантность ОМП). Если параметры θ и ν связаны непрерывной взаимно-однозначной зависимостью $\nu=\varphi(\theta)$ и $\hat{\theta}$ — ОМП для θ, то $\hat{\nu}=\varphi(\hat{\theta})$ — ОМП для ν.

Доказательство. Пусть $f(x)$ — плотность распределения изучаемой СВ X. Тогда $f_\theta(x)\leqslant f_{\hat{\theta}}(x)$ в некоторой окрестности точки $\hat{\theta}$; следовательно, так как φ непрерывна, $f_{\varphi^{-1}(\nu)}(x)\leqslant f_{\varphi^{-1}(\hat{\nu})}(x)$. Это неравенство означает утверждение теоремы в силу характера зависимости φ, т.е. $f_\nu(x)\leqslant f_{\hat{\nu}}(x)$ в некоторой окрестности и $\hat{\nu}=\varphi(\hat{\theta})$ является ОМП для ν.

Замечание 2.14. Если задача состоит в построении ОМП для некоторой $\tau(\theta)$ и $\tau(\theta)$ удовлетворяет условиям теоремы 2.12, то задача сводится к более простой — нахождению $\hat{\theta}$ — ОМП для θ, тогда $\hat{\tau}(\hat{\theta})=\tau(\hat{\theta})$ есть ОМП для $\tau(\theta)$.

Замечание 2.15. При достаточно общих условиях ОМП состоятельны, асимптотически нормальны и асимптотически эффективны при $n\to\infty$.

2. Особый интерес представляет нахождение ОМП в **нерегулярном случае**: ОМП находится из смысла метода ММП, т.е. как значение, при котором функция правдоподобия принимает наибольшие значения.

Пример 2.15. СВ $X\sim R[0;\theta]$. Найдем ОМП для θ, т.е. $\hat{\theta}$. Это нерегулярный случай. Имеем

$$L=\begin{cases}0, & x\notin[0;\theta],\\ \frac{1}{\theta^n}, & x\in[0;\theta].\end{cases}$$

Наибольшее значение L, равное $\frac{1}{\theta^n}$, получается при наименьшем возможном значении θ, если наблюдались значения x: x_1, ..., x_n, т.е. при $\theta=x_{(n)}$.

Пример 2.16 (иллюстрация того, что ОМП бесконечно много). СВ $X\sim R[0;1+\theta]$.

$$L = \begin{cases} 1, \theta \leqslant x_{(1)} \leqslant x_{(n)} \leqslant \theta + 1, \\ 0, \text{в противном случае.} \end{cases}$$

Очевидно, что $\max L = 1$, но

$$L = \begin{cases} \theta \leqslant x_{(1)}, \\ \theta \leqslant x_{(n)} - 1 \end{cases} \Rightarrow \widehat{\theta} \in [x_{(n)} - 1; x_{(1)}].$$

Точек на отрезке для $\widehat{\theta}$ бесконечно много, следовательно, значений $\widehat{\theta}$ бесконечно много. Остается только один вопрос: существует ли такой отрезок?

Из выражения для L имеем $\theta \leqslant x_{(1)} \leqslant x_{(n)} \leqslant \theta + 1$, значит, $x_{(1)} \geqslant x_{(n)} - 1$, т.е. отрезок существует.

Далее предлагаются к рассмотрению задачи на построение ОММ и ОМП.

Задача 2.36. СВ $X \sim \Gamma_{\alpha,\lambda}$, α, λ — const > 0. Найти ОММ для α и λ.

Решение

Составим систему уравнений для нахождения ОММ путем приравнивания теоретических моментов соответствующим выборочным в числе неизвестных оцениваемых параметров распределения:

$$\begin{cases} \dfrac{\alpha}{\lambda} = \overline{x}, \\ \dfrac{\alpha(\alpha+1)}{\lambda^2} = \dfrac{1}{n}\sum_{i=1}^{n} x_i^2. \end{cases}$$

Поделим второе уравнение на первое, возведенное в квадрат. Получаем

$$\frac{\alpha(\alpha+1)\lambda^2}{\lambda^2\alpha^2} = \frac{\sum_{i=1}^{n} x_i^{\,2}}{n\overline{x}^2} \Rightarrow \alpha\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) = nx \Rightarrow$$

$$\Rightarrow \widehat{\alpha} = \frac{n\overline{x}^2}{\sum_{i=1}^{n} x_i - n\overline{x}^2}; \widehat{\lambda} = \frac{\widehat{\alpha}}{\overline{x}} = \frac{n\overline{x}}{\sum_{i=1}^{n} x_i - n\overline{x}}.$$

Задача 2.37. СВ $X \sim R[m - \alpha; m + \alpha]$, α — известно, m — неизвестно. Найти ОМП для параметра m.

Решение

Вычислим функцию правдоподобия:

$$L = \begin{cases} 0, x \in [m - \alpha; m + \alpha], \\ \dfrac{1}{(2\alpha)^n}, x \in [m - \alpha; m + \alpha] \end{cases} \Rightarrow$$

$\Rightarrow$ максимальное значение $L = \dfrac{1}{(2\alpha)^n}$ достигается при всех $x_i = [m - \alpha;$ $m + \alpha]$, $i = 1, ..., n$. Отсюда следует, что множество ОМП для m есть пересечение всех отрезков $[x_i - \alpha; x_i + \alpha]$, $i = 1, ..., n$, т.е. это любое значение отрезка

$$[x_{(n)} - \alpha; x_{(1)} + \alpha] = [\max_{i=1,\dots,n} x_i - \alpha; \min_{i=1,\dots,n} x_i + \alpha],$$

при этом $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ может не принадлежать этому отрезку.

Задача 2.38. СВ $X \sim R[m - \alpha; m + \alpha]$, где параметры m и α оба неизвестны. Найти ОМП для параметров m и α.

Решение

По решению задачи 2 $\max L = \frac{1}{(2\alpha)^n}$ достигается при возможном минимальном значении α, совместном с выборкой, когда $m \in [x_{(n)} - \alpha; x_{(1)} + \alpha]$, т.е. когда

$$x_{(n)} - \alpha \leqslant x_{(1)} + \alpha \Rightarrow \alpha \geqslant \frac{x_{(n)} - x_{(1)}}{2} \Rightarrow \hat{\alpha} = \frac{x_{(n)} - x_{(1)}}{2}$$

— ОМП для α. Тогда получаем

$$\hat{m} \in [x_{(n)} - \hat{\alpha}; x_{(1)} + \hat{\alpha}] = \left[x_{(n)} - \frac{x_{(n)} - x_{(1)}}{2}; x_{(1)} + \frac{x_{(n)} - x_{(1)}}{2}\right] =$$

$$= \left[\frac{x_{(n)} + x_{(1)}}{2}; \frac{x_{(n)} + x_{(1)}}{2}\right] \Rightarrow$$

$\Rightarrow$ ОМП для параметра m: $\hat{m} = \frac{x_{(n)} + x_{(1)}}{2}$.

Задача 2.39. СВ $X \sim R[a\theta_1 + b, c\theta_2 + d]$, где a, b, c, d — известные константы; $\theta = (\theta_1, \theta_2)$ — неизвестный параметр распределения. Найти ОМП и ОММ для θ.

Решение

Вычислим функцию правдоподобия:

$$L = \begin{cases} \dfrac{1}{c\theta_2 - a\theta_1 + d - b}, \ a\theta_1 + b \leqslant x_{(1)} \leqslant x_{(n)} \leqslant c\theta_2 + d, \\ 0, \text{в противном случае.} \end{cases}$$

Она принимает наибольшее значение при минимальном значении θ_2 и максимальном значении θ_1, а так как $\theta_1 \leqslant \frac{x_{(1)} - b}{a}$, $\theta_2 \geqslant \frac{x_{(n)} - d}{c}$, то:

$$\text{ОМП: } \hat{\theta} = \begin{cases} \hat{\theta}_1 = \dfrac{x_{(1)} - b}{a}, \\ \hat{\theta}_2 = \dfrac{x_{(2)} - d}{c}; \end{cases} \quad \text{ОММ: } \begin{cases} \overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{a\theta_1 + b + c\theta_2 + d}{2}, \\ \dfrac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n} = \dfrac{(c\theta_2 + d - z\theta_1 - b)^2}{12}. \end{cases}$$

2.5.2. Байесовские оценки (решения)

Под решением здесь понимается выбор оценки неизвестного параметра распределения $\theta = (\theta_1, ..., \theta_k)$.

Решающей функцией $d = \delta(x)$ называется правило (функция), ставящее в соответствие каждому результату наблюдения некоторое решение d. Областью определения величины $\delta(x)$ является множество значений наблюдаемой СВ X, областью значений — множество решений D .

Чтобы функцию $\delta(x)$ выбрать наилучшим образом, нужно сравнить последствия использования различных функций $\delta(x)$. Для этого задается функция потерь $W(\theta, d)$, значение которой определяется выбранным решением, если СВ X имеет функцию распределения $F_\theta(x)$ с неизвестным параметром θ. Тогда при многократном применении функции $\delta(x)$ (при повторении опыта) определяются средние потери $MW(\theta, d) = R(\theta, d)$, называемые ***функцией риска***.

Цель состоит в выборе решающей функции $\delta(x)$, минимизирующей функцию риска $R(\theta, d)$.

Байесовский подход отличается от небайесовского тем, что неизвестный параметр θ считается не фиксированной постоянной, а СВ с известным распределением $\pi(\theta)$ — априорным распределением.

После завершения наблюдений над СВ X из априорного распределения неизвестного параметра можно получить его апостериорное распределение $\pi(\theta/X)$, и выбор байесовского решения естественно связывать с ожидаемыми потерями при этом апостериорном распределении.

Доказано, что при байесовском подходе минимальные средние потери (риск) по всем возможным значениям параметра θ и всем реализациям СВ X достигаются при той же решающей функции $\delta(x)$, при которой получаются минимальные средние потери при апостериорном распределении параметра θ.

Апостериорный риск вычисляется по формуле

$$R(d,x) = \begin{cases} \sum\limits_i W(\theta_i, d)\pi(\theta_i/x_j) \text{ в дискретном случае,} \\ \int\limits_\Theta W(\theta, d)\pi(\theta/x_j)d\theta \text{ в непрерывном случае;} \end{cases} \qquad (2.14)$$

$$\pi(\theta_i/x_j) = \frac{\pi(\theta_i)P(x_i/\theta_j)}{P(x_j)}; \; \pi(\theta/x_j) = \frac{\pi(\theta)P(x_j/\theta)}{P(x_j)}. \qquad (2.15)$$

Здесь $P(x_j/\theta_i) = P(X = x_j/\theta_i)$; $P(x_j/\theta) = f(x_j/\theta)$ — плотность распределения СВ X в непрерывном случае при $X = x_j$;

$$P(x_j)=\begin{cases}\sum\limits_i \pi(\theta_i)P(x_j/\theta_i) \text{ в дискретном случае,}\\ \int\limits_\Theta \pi(\theta)P(x_j/\theta)d\theta \text{ в непрерывном случае.}\end{cases} \quad (2.16)$$

Формула (2.15) — формула Байеса, а (2.16) — формула полной вероятности.

Формулы (2.14)—(2.16) даны для нахождения байесовского решения минимизацией апостериорного риска по одному наблюдению. Если же ставится задача сделать это по выборке, то в формулах (2.15) и (2.16) следует заменить вероятности $P(x_j/\theta_i)$ и $P(x_j/\theta)$ на соответствующие функции правдоподобия $L_\theta(x_1, ..., x_k)$.

Разберем подробно две задачи оценивания неизвестного параметра θ путем нахождения байесовского решения в дискретном и непрерывном случаях с использованием формул (2.14)—(2.16).

Задача 2.40. Пусть СВ X имеет бернуллиевское распределение $B(1, \theta)$, априорное распределение $\pi(\theta)$ задано рядом распределения

θ	1/4	1/2
P	1/3	1/3

а функция потерь $W(\theta_i, d_j)$ задается таблицей

$W(\theta_i, d_j)$	d_1	d_2
$\theta_1 = 1/4$	1	4
$\theta_2 = 1/2$	3	2

Здесь множество решений $D = (d_1, d_2)$ состоит из двух точек: d_1 ($\theta = 1/4$) и d_2 ($\theta = 1/2$). Найти байесовскую оценку для неизвестного параметра θ.

Решение

По формуле (2.14) $R(d, X_j) = \sum\limits_{i=1}^{2} W(\theta_i, d_j)\pi(\theta_i/X_j)$, где $j = 1, 2$ (так как СВ X принимает два значения: $X = 0$ и 1). По формуле (2.15) $\pi(\theta_i/X_j) = \dfrac{\pi(\theta)P(x_j, \theta_i)}{P(x_j)}$. По формуле (2.16) $P(x_j) = \sum\limits_{i=1}^{2}\pi(\theta_i)P(x_j/\theta_i)$, а по условию задачи $P(x_j/\theta_i) = \theta_i^{x_j}(1-\theta_i)^{1-x_j}$, откуда получаем

$$P(X_1=0)=\pi(\theta_1=1/4)P(x_1=0/\theta_1=1/4)+\pi(\theta_2=1/2)P(x_1=0/\theta_2=1/2)=$$
$$=(1/3)(3/4)+(2/3)(1/2)=7/12.$$

Аналогично получаем $P(X_2 = 1) = 3/12$; $\pi(\theta_1/X_1) = 3/7$; $\pi(\theta_2/X_1) = 4/7$; $\pi(\theta_1/X_2) = 1/5$; $\pi(\theta_2/X_2) = 4/5$.

Таким образом, найдено апостериорное распределение параметра θ.

Вычислим и сравним теперь апостериорные риски при каждом наблюдаемом значении СВ X ($X_1 = 0$ и $X_2 = 1$) для всех решений, и в качестве байесовского выберем то из них, при котором получается меньший риск:

а) пусть $X = X_1 = 0$, тогда по формуле (2.14)

$$R(d_1, X_1) = W(\theta_1, d_1)\pi(\theta_1/X_1) + W(\theta_2, d_1)\pi(\theta_2/X_1) =$$
$$= 1\cdot 3/7 + 2\cdot 4/7 = 11/7.$$

Аналогично получаем, что $R(d_2, X_1) = 4\cdot 3/7 + 2\cdot 4/7 = 20/7$;

б) пусть $X = X_2 = 1$. Тогда

$$R(d_1, X_2) = 1\cdot 1/5 + 3\cdot 4/5 = 13/5;$$
$$R(d_2, X_2) = 4\cdot 1/5 + 2\cdot 4/5 = 12/5.$$

При $X = X_1 = 0$ меньший апостериорный риск при решении d_1, а при $X = X_2 = 1$ — при d_2, таким образом, получаем байесовское правило:

$$\delta(x) = (\delta(0) = d_1, \delta(1) = d_2).$$

Задача 2.41. Пусть СВ X имеет экспоненциальное распределение с неизвестным параметром θ, априорное распределение которого есть гамма-распределение, а функция потерь $W(d, \theta) = (\theta - t)^2$, где решение $d : t = \hat{\theta}$ — оценка для θ. Найти байесовскую оценку для параметра θ:

а) по одному наблюдению X;

б) по выборке $x_1, ..., x_n$.

Решение

а) Найдем апостериорное распределение:

$$\pi(\theta/x) = \frac{\pi(\theta)P(x/\theta)}{\int\limits_{\Theta}\pi(\theta)P(x/\theta)d\theta} = \frac{r^k\theta^{k-1}\exp(-rx)\theta\exp(-\theta x)\Gamma(k)}{\Gamma(k)r^k\int\limits_0^\infty \theta^{k-1}\exp(-r\theta)Q\exp(-\theta x)d\theta} =$$

$$= \frac{\theta^k\exp[-\theta(x+r)]}{\int\limits_0^\infty \theta^k\exp[-\theta(r+x)]dt} =$$

$$= \frac{\theta^k\exp[-\theta(x+r)]}{(1/(x+r)^{k+1})\int\limits_0^\infty \theta^k(x+r)^k\exp[-\theta(r+x)]d(\theta(r+x))} =$$

$$= \frac{\theta^k\exp[-\theta(x+r)](x+r)^{k+1}}{\Gamma(k+1)}.$$

Это есть гамма-распределение с плотностью $g(\theta; k + 1, x + r)$.

Байесовскую оценку для θ находим, минимизируя апостериорный риск $R(d, x)$ по t:

$$R(d,x) = \int\limits_0^\infty (\theta - t)^2 g(\theta;\ k+1, x+r)d\theta.$$

Берем производную по t и приравниваем ее нулю. Получаем

$$\int\limits_0^\infty (\theta - t)g(\theta; k+1, x+r)d\theta = 0,$$

откуда

$$t = \int_0^\infty \theta g(\theta; k+1, x+r)d\theta =$$

$$= \frac{(x+r)^{k+1}}{\Gamma(k+1)} \int_0^\infty \theta^{k+1}(x+r)^{k+1} \exp[-\theta(x+r)] \frac{d(\theta(x+r))}{(x+r)^{k+2}} =$$

$$= \frac{(k+1)\Gamma(k+1)}{\Gamma(k+1)(x+r)} = \frac{k+1}{x+r},$$

т.е. $\widehat{\theta} = t = (k+1)/(x+r)$ — байесовская оценка для неизвестного параметра и по наблюденному значению x.

Апостериорный байесовский риск тогда вычисляется по формуле

$$R(\theta, x_j) = \int_0^\infty [\theta - (k+1)/(x+r)]^2 g(\theta; k+1, x+r)d\theta =$$

$$= \int_0^\infty \theta^2 g(\theta; k+1, x+r)d\theta - 2x\frac{k+1}{x+r}\int_0^\infty \theta g(\theta; k+1, x+r)d\theta + \left(\frac{k+1}{x+r}\right)^2,$$

где интегралы $\int_0^\infty \theta^2 g(\theta; k+1, x+r)d\theta$ и $\int_0^\infty \theta g(\theta; k+1, x+r)d\theta$ аналогично преобразуются, как было показано выше, выражаются через гамма-функцию и могут быть досчитаны самостоятельно.

б) Дана выборка $x_1, ..., x_n$. Найдем сначала апостериорное распределение $\pi(\theta / x_1, ..., x_n)$.

Аналогично случаю а) получаем байесовскую оценку в виде

$$t = \int_0^\infty \theta\pi(\theta / x_1, ..., x_n)d\theta =$$

$$= \frac{\left(\sum_{i=1}^n x_i + r\right)^{n+k}}{\Gamma(n+k)} \int_0^\infty \theta^{n+k} \exp\left\{-\theta\left(\sum_{i=1}^n x_i + r\right)\right\}d\theta =$$

$$= \frac{\left(\sum_{i=1}^n x_i + r\right)^{n+k}}{\Gamma(n+k)} \frac{\Gamma(n+k+1)}{\left(\sum_{i=1}^n x_i + r\right)^{n+k+1}} = \frac{n+k}{\sum_{i=1}^n x_i + r},$$

это байесовская оценка для параметра в случае б).

Апостериорный риск тогда в этом случае есть

$$R(d, x_1, ..., x_n) = \int_0^\infty \left[\theta - (n-k) / \left(\sum_{i=1}^n x_i + r\right)\right]^2 g\left(\theta; n+k, \sum_{i=1}^n x_i + r\right)d\theta,$$

вычисляется аналогично случаю а) и может быть получен самостоятельно.

2.6. Доверительное оценивание

Постановка задачи. $\vec{x} = x_1, ..., x_n$ — выборка объема n наблюдений над случайной величиной X, распределение которой относится к параметрическому семейству $F_\theta(x)$, где $\theta = (\theta_1, ..., \theta_k)$ и $\theta \in \Theta$ (Θ — параметрическое множество). Требуется оценить некоторую функцию $\tau = \tau(\theta)$. Доверительное оценивание τ означает нахождение k-мерной области, заключающей неизвестное значение функции τ с заданной доверительной вероятностью γ. Подробнее остановимся на рассмотрении случая $k = 1$ и $\tau(\theta) = \theta$. Тогда искомое доверительное множество становится доверительным интервалом и задача состоит в построении двух статистик $t_1 = t_1(\vec{x})$ и $t_2 = t_2(\vec{x})$ (концов доверительного интервала $J = (t_1, t_2)$, заключающего в себе неизвестное значение параметра θ с заданной доверительной вероятностью γ: $\gamma = P(t_1 < \theta < t_2)$).

При доверительном оценивании заданное значение γ (обычно близкое к единице) означает надежность оценивания $\tau(\theta)$ с точностью, определяемой размером доверительной области. При построении доверительного интервала для параметра θ его длина — точность оценивания, а γ — заданная надежность. Поэтому желательно строить кратчайший доверительный интервал, соответствующий наибольшей точности при данном γ.

Общий прием при нахождении доверительного интервала состоит в построении центральной статистики (ЦС) $Z = Z(\theta)$, т.е. такой статистики, распределение которой не зависит от неизвестного параметра θ. Если $Z(\theta)$ непрерывна и монотонна по θ, то это обеспечивает однозначную эквивалентность событий $\{t_1^* < Z < t_2^*\}$ и $\{t_1 < \theta < t_2\}$. Тогда если удалось найти $t_1^* = t_1^*(\theta)$ и $t_2^* = t_2^*(\theta)$ — нижнюю и верхнюю доверительные границы, то, решая неравенство $t_1^* < Z < t_2^*$ относительно θ, находим значения t_1 и t_2 — искомые границы доверительного интервала для неизвестного параметра θ.

Остается обсудить две проблемы: построение центральной статистики $Z = Z(\theta)$ и нахождение значений t_1^* и t_2^* из уравнения

$$\gamma = P(t_1^* < Z < t_2^*). \tag{2.17}$$

1. Путями построения ЦС могут быть следующие:

1) замена исходной СВ на новую, зависящую от неизвестного параметра θ, распределение которой не зависит от θ;

2) стандартизация имеющейся точечной оценки;

3) использование результатов ЦПТ или асимптотической нормальности ОМП (для построения асимптотических доверительных интервалов).

2. Для нахождении значений t_1^* и t_2^* требуется сформулировать дополнительное ограничение на t_1^* и t_2^*, чтобы это было возможно,

т.е. чтобы уменьшить число неизвестных в уравнении (2.17) с двух до одной.

При решении этой проблемы различают обычно два случая: регулярный и нерегулярный (см. ранее определение регулярного случая требованиями дифференцируемости функции правдоподобия L и независимости области, в которой $L \neq 0$, от неизвестного параметра θ).

А. Регулярный случай. Строят центральный доверительный интервал (ЦДИ). Определим ЦДИ. Пусть $g_\theta(x)$ — кривая распределения неизвестного параметра θ (рис. 2.2).

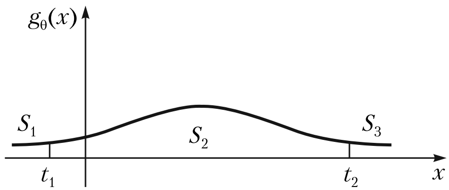


Рис. 2.2. **Кривая распределения неизвестного параметра θ**

Тогда $J = (t_1, t_2)$ — ЦДИ, если площади S_1 и S_3 одинаковы.

Определим $(x_p - p)$ — ***квантиль*** распределения $F(x)$, если x_p — корень уравнения $F(x) = p$.

Теперь выразим значения t_1 и t_2 в терминах квантили распределения параметра θ с функцией распределения $G(\theta)$: $S_1 + S_2 + S_3 = 1$.

Пусть $S_2 = \gamma$, тогда $S_1 = S_3 = \dfrac{1-\gamma}{2}$, откуда следует, что $t_1 = \dfrac{1-\gamma}{2}$, а $t_2 = \left(\dfrac{1-\gamma}{2} + \gamma\right) = \dfrac{1+\gamma}{2}$, т.е. t_1 и t_2 — соответственно $\dfrac{1-\gamma}{2}$- и $\dfrac{1+\gamma}{2}$-квантили распределения $G(\theta)$. Значит, ЦДИ — интервал между $\dfrac{1-\gamma}{2}$- и $\dfrac{1+\gamma}{2}$-квантилями распределения $G(\theta)$. Таким образом, требование построения ЦДИ и есть необходимое дополнительное требование в уравнении (2.17).

Б. Нерегулярный случай. В качестве искомого доверительного интервала в уравнении (2.17) выбирают крайнюю зону значений неизвестного параметра. Тогда число неизвестных в уравнении (2.17) уменьшается до одного.

Рассмотрим примеры построения точных доверительных интервалов (ДИ).

Задача 2.42. СВ $X \sim R[2\theta - 1; 3\theta + 4]$. Построить ДИ с уровнем доверия, т.е. надежностью, γ для неизвестного параметра θ.

Решение

Введем новую СВ $Y = X - 2\theta + 1$, тогда СВ $Y \sim R[0; \theta + 5] = R[0; \theta^*]$, где $\theta^* = \theta + 5$. Построим ДИ с уровнем доверия γ для параметра θ.

Обозначим СВ $Z = Y(n)/\theta^*$, тогда $F_Z(Y) = P\{Y(n) < \theta^* y\} = y^n$ — это функция распределения максимума выборки n равномерно распределенных на интервале [0; 1] значений $y_1, ..., y_n$. Распределение СВ Z не зависит от θ, поэтому Z — ЦС:

$$P(t_\gamma < Y_{(n)}/\theta^* < 1) = \gamma = P(t_\gamma < Z < 1) = F_z(1) - F_z(t_\gamma) = 1 - t_\gamma^n,$$

откуда

$$\gamma = P(Y_{(n)} < \theta^* = \theta + 5 < Y_{(n)} \big/ \sqrt[n]{1-\gamma}) =$$
$$= P\left(x(n) - 2\theta - 4 < \theta < \frac{x_{(n)} - 2\theta + 1}{\sqrt[n]{1-\nu}} - 5 \right),$$

или с вероятностью γ выполнены неравенства

$$\begin{cases} \dfrac{x_{(n)} + 1 - 5\sqrt[n]{1-\nu}}{\sqrt[n]{1-\nu} - 2} > \theta, \\ (x_{(n)} - 4)/3 < \theta, \end{cases}$$

т.е.

$$P\left(\frac{x_{(n)} - 4}{3} < \theta < \frac{x_{(n)} + 1 - 5\sqrt[n]{1-\gamma}}{\sqrt[n]{1-\gamma} - 2} \right) = \gamma.$$

Требуемый интервал (t_1, t_2) для θ построен:

$$t_1 = \frac{x_{(n)} - 4}{3};\ t_2 = \frac{x_{(n)} + 1 - 5\sqrt[n]{1-\gamma}}{\sqrt[n]{1-\gamma} - 2}.$$

Задача 2.43. СВ $X \sim E(a\theta + b)$. Построить ДИ для неизвестного параметра θ с уровнем доверия γ, где a и b — const, $a > 0$.

Решение

Обозначим $\theta^* = a\theta + b$ и построим сначала ДИ для параметра θ с уровнем доверия γ.

Так как $X \sim E(\theta^*)$, имеем $F(x) = 1 - \exp(-\theta^* x)$, $x \geqslant 0$,

$$F_{(1)}(x) = 1 - (1 - F(x))^n = 1 - \exp(-\theta^* n x).$$

Введем новую СВ $Y = \theta^* x_{(1)}$. Для нее $F_Y(x) = P(x_{(1)} < x/\theta^*\} = F_{(1)}(x/\theta^*) = 1 - \exp(-nx)$ — это функция распределения минимума выборки экспоненциально распределенных значений $y_1, ..., y_n$ СВ $Y \sim E(1)$, поэтому Y — ЦС и

$$\gamma = P(0 < Y < \theta^* x_{(1)} < t_\gamma) = P(0 < \theta^* < t_\gamma / x_{(1)}) =$$
$$= F_Y(t_\gamma) - F_Y(0) = 1 - \exp(-n t_\gamma),$$

откуда $t_\gamma = -(\ln(1 - \gamma))/n$. Тогда

$$P\left(0 < \theta^* < -\frac{\ln(1-\gamma)}{n x_{(1)}} \right) = P\left(0 < a\theta + b < -\frac{\ln\ (1-\gamma)}{n x_{(1)}} \right) =$$
$$= P\left(-b/a < \theta^* < -\frac{\ln\ (1-\gamma)}{a n x_{(1)}} - b/a \right),$$

т.е. требуемый интервал (t_1, t_2) для θ построен:

$$t_1 = -\frac{b}{a}; \quad t_2 = -\frac{\ln(1-\gamma)}{anx_{(1)}} - \frac{b}{a}.$$

Задача 2.44. Построить доверительный интервал для m: СВ $X \sim N(m; \sigma)$, где σ — известный параметр, m — неизвестно.

Решение

Построим статистику $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ и найдем ее распределение методом характеристических функций. СВ $X \sim N(m; \sigma) \Rightarrow$

$$\Rightarrow g_{x_i}(t) = e^{itm - \frac{\sigma^2 t^2}{2}} \Rightarrow g_x(t) = \left(g_{x_i}\left(\frac{t}{n}\right)\right)^n = \left(e^{i\frac{t}{n}m - \frac{\sigma^2 t^2}{2n^2}}\right)^n = e^{itm - \frac{\left(\frac{\sigma}{\sqrt{n}}\right)^2 t^2}{2}} \Rightarrow$$

$$\Rightarrow \bar{x} \sim N\left(m; \frac{\sigma}{\sqrt{n}}\right) \Rightarrow M\bar{x} = m; \; D\bar{x} = \frac{\sigma^2}{n},$$

откуда следует, что СВ $T = \frac{\bar{x} - M\bar{x}}{\sqrt{D\bar{x}}} = \frac{(\bar{x} - m)\sqrt{n}}{\sigma} \sim N(0,1)$, а значит, T — центральная статистика $\Rightarrow$

$$\Rightarrow \gamma = P(|\bar{x} - m| < \delta) = P(|\bar{x} - M\bar{x}| < \delta) = P\left(|T| < \frac{\delta\sqrt{n}}{\sigma} = t_\gamma\right) = 2\Phi(t_\gamma) \Rightarrow$$

$\Rightarrow t_\gamma$ находим по таблице функции Лапласа. Тогда с заданной вероятностью γ должно выполняться неравенство

$$\frac{|\bar{x} - m|\sqrt{n}}{\sigma} = |T| < t_\gamma, \text{ или } |\bar{x} - m| < \frac{t_\gamma \sigma}{\sqrt{n}}, \text{ или } \bar{x} - \frac{t_\gamma \sigma}{\sqrt{n}} < m < \bar{x} + \frac{t_\gamma \sigma}{\sqrt{n}} \Rightarrow$$

$\Rightarrow$ искомый доверительный интервал $Y = (t_1, t_2) = \left(\bar{x} \mp \frac{t_\gamma \sigma}{\sqrt{n}}\right)$ — ЦДИ. Если $f_x(x)$ — плотность распределения случайной величины $X \sim N(m, \sigma)$, то (рис. 2.3)

$$t_1 = x_{\frac{1-\gamma}{2}} = \bar{x} - \frac{t_\gamma \sigma}{\sqrt{n}}; \quad t_2 = x_{\frac{1+\gamma}{2}} = \bar{x} + \frac{t_\gamma \sigma}{\sqrt{n}}.$$

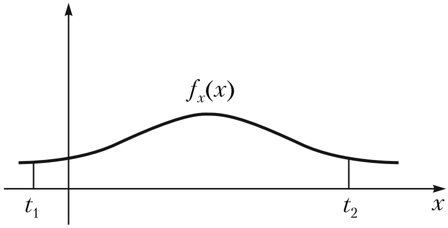


Рис. 2.3. **Доверительный интервал для нормального распределения**

Рассмотрим два числовых примера.

Задача 2.45. СВ $X \sim N(m; \sigma = 3)$. Найти доверительный интервал J для неизвестного параметра m c доверительной вероятностью $\gamma = 0{,}95$.

Решение

t_γ находим по таблице функции Лапласа из условия $\Phi(t_\gamma) = \dfrac{\gamma}{2} =$ $= 0{,}475 \Rightarrow$

$$\Rightarrow t_\gamma = 1{,}96;\ Y = \left(\bar{x} \mp \frac{1{,}96 \cdot 3}{\sqrt{36}}\right) = (\bar{x} \mp 0{,}98).$$

Например, при $\bar{x} = 4$ имеем $J = (3{,}02; 4{,}98)$.

Задача 2.46. В предыдущей задаче найти минимальный объем выборки n, при котором с заданной точностью $\delta = 0{,}98$ и надежностью $\gamma = 0{,}95$ можно оценить параметр m.

Решение

$$\delta = \frac{t_\gamma \sigma}{\sqrt{n}} \Rightarrow n = \frac{t_\gamma^2 \sigma^2}{\delta^2} = \frac{1{,}96^2 \cdot 3^2}{0{,}98^2} \approx 36.$$

Следующая задача представляет собой пример построения асимптотических доверительных интервалов по выборке для некоторых распределений.

Задача 2.47. Построить асимптотические доверительные интервалы J для неизвестного параметра θ с уровнем доверия γ по выборке $x_1, x_2, ..., x_n$, где n — объем выборки, для следующих распределений: а) СВ $X \sim B(N, \theta)$; б) СВ $X \sim \pi(\theta)$; в) СВ $X \sim N(\theta = m, \sigma)$, σ — неизвестно; г) СВ $X \sim N$(и $= m$, у), m — неизвестно.

Решение

а) Рассмотрим статистику

$$\hat{\theta} = \frac{1}{nN}\sum_{i=1}^{n} x_i;\ \ M\hat{\theta} = \theta;\ D\hat{\theta} = \frac{1}{n^2N^2} nN\theta(1-\theta) = \frac{\theta(1-\theta)}{nN};$$

тогда

$$L\left(\frac{\hat{\theta} - M\hat{\theta}}{\sqrt{D\hat{\theta}}}\right) \underset{n\to\infty}{\to} N(0,1) \Rightarrow T = \frac{\hat{\theta} - M\hat{\theta}}{\sqrt{D\hat{\theta}}} = \frac{(\hat{\theta} - \theta)\sqrt{nN}}{\sqrt{\theta(1-\theta)}} - \text{ЦС} \Rightarrow$$

$$\Rightarrow \gamma = P(|\hat{\theta} - \theta| < \delta) = P(|\hat{\theta} - M\hat{\theta}| < \delta) = P\left(|T| < \frac{\delta\sqrt{nN}}{\sqrt{\theta(1-\theta)}} = t_\gamma\right) = 2\Phi(t_\gamma) \Rightarrow$$

$\Rightarrow$ значение $t\gamma$ находим по таблице функции Лапласа. Тогда с заданной вероятностью γ должно выполняться неравенство

$$\frac{|\hat{\theta} - \theta|\sqrt{nN}}{\sqrt{\theta(1-\theta)}} < t_\gamma \Rightarrow \frac{(\hat{\theta} - \theta)^2 nN}{(1-\theta)\theta} < t_\gamma^2 \Rightarrow$$

$$\Rightarrow f(\theta) = (\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2)nN - t_\gamma^2\theta(1-\theta) < 0 \Rightarrow$$

$$\Rightarrow \theta^2(nN + t_\gamma^2) - 2\theta(\hat{\theta}nN + t_\gamma^2) + \hat{\theta}^2 nN < 0.$$

Решая неравенство относительно θ, получим искомый доверительный интервал $J = (\theta_1, \theta_2)$, где θ_1 и θ_2 — корни уравнения $f(\theta) = 0$.

б) Рассмотрим статистику

$$\hat{\theta} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i;\ \ M\hat{\theta} = \theta;\ \ D\hat{\theta} = \frac{DX}{n} = \frac{\theta}{n};$$

тогда

$$L\left(\frac{\overline{x} - M\overline{x}}{\sqrt{Dx}}\right) \underset{n\to\infty}{\rightarrow} N\left(0,1\right) \Rightarrow T = \frac{\overline{x} - M\overline{x}}{\sqrt{Dx}} = \frac{(\overline{x}-\theta)\sqrt{n}}{\sqrt{\theta}} - \text{ЦС} \Rightarrow$$

$$\Rightarrow \gamma = P(|\overline{x} - \theta| < \delta) = P(|\overline{x} - M\overline{x}| < \delta) = P\left(|T| < \frac{\delta\sqrt{n}}{\sqrt{\theta}} = t_\gamma\right) = 2\Phi(t_\gamma),$$

где значение t_γ находят по таблице функции Лапласа.

Имеем $\overline{x} \sim N\left(\theta, \dfrac{\sigma}{\sqrt{n}}\right)$; $\dfrac{nS_1^2}{\sigma^2} \sim \chi^2_{n-1}$ (было показано раньше). Тогда с заданной вероятностью γ должно выполняться неравенство

$$\frac{(\overline{x}-\theta)\sqrt{n}}{\sqrt{\theta}} < t_\gamma \Rightarrow \frac{(\overline{x}-\theta)^2 n}{\theta} < t_\gamma^2 \Rightarrow (\overline{x}^2 - 2\theta\overline{x} + \theta^2)n - t_\gamma^2\theta = f(\theta) < 0.$$

Отсюда находят искомый доверительный интервал $J = (\theta_1, \theta_2)$, где θ_1 и θ_2 — корни уравнения $f(\theta) = 0$.

в) Имеем $\overline{x} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i$; $\hat{\sigma} = \sqrt{S_1^2} = \sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}$. Покажем, что случайная величина $T = \dfrac{\overline{x}-\theta}{\dfrac{\hat{\sigma}}{\sqrt{n}}} = \dfrac{(\overline{x}-\theta)\sqrt{n}}{\hat{\sigma}}$ — ЦС. Известно, что случайная величина $\dfrac{Z\sqrt{n}}{\sqrt{V}}$ имеет распределение Стьюдента с плотностью

$$S_{n-1}(t) = \frac{\Gamma\left(\dfrac{n}{2}\right)}{\sqrt{\pi(n-1)\Gamma\left(\dfrac{n-1}{2}\right)}} \cdot \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}}$$

— четная табличная функция, где СВ $Z \sim N(0, 1)$, СВ $V \sim \chi^2_{n-1}$. Покажем, что случайная величина T распределена по Стьюденту, а значит, является ЦС.

Действительно,

$$T = \frac{(\overline{x}-\theta)\sqrt{n}}{\hat{\sigma}} \frac{\sigma}{\sqrt{n}\hat{\sigma}}\sqrt{n},$$

где $Z = \dfrac{(\overline{x}-\theta)\sqrt{n}}{\hat{\sigma}} \sim N(0,1)$, а $\dfrac{\sigma}{\sqrt{n}\hat{\sigma}} = \dfrac{1}{\sqrt{V}} \Rightarrow V = \dfrac{n\hat{\sigma}^2}{\sigma^2} = \dfrac{nS_1^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow$ $L(T)$ — распределение Стьюдента. Тогда

$$\gamma = P(|T| < t_\gamma) = P\left(\frac{|\overline{x} - \theta|\sqrt{n}}{\hat{\sigma}} < t_\gamma\right) = P\left(|\overline{x} - \theta\,| < \frac{t_\gamma \hat{\sigma}}{\sqrt{n}} = \sigma\right),$$

где t_γ находится по таблицам $S_{n-1}(t)$. Так как $S_{n-1}(t)$ — четная функция, то

$$P(|T| < t_\gamma) = 2\int_0^{t_\gamma} S_{n-1}(t)dt = \gamma.$$

Поэтому $J = \left(\overline{x} \mp \frac{t_\gamma \hat{\sigma}}{\sqrt{n}}\right)$ — искомый доверительный интервал.

г) Известно, что статистика $\chi^2_{n-1} = \frac{S_1^2}{\sigma^2}$, где $S_1^2 = \frac{1}{10}\sum_{i=1}^{n}(x_i - \overline{x})^2$, имеет распределение χ^2_{n-1} с плотностью $k_n(x) = f(x) = \dfrac{x^{\frac{n-3}{2}} e^{-\frac{x}{2}}}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}$ (это табличная функция).

Построим новую случайную величину $Y = \sqrt{X^2} = \dfrac{S_1\sqrt{n}}{\sigma}$, где $S_1 = \sqrt{S_1^2}$. Найдем плотность распределения $g(y)$ случайной величины Y.

Воспользуемся следующим определением. Пусть $Y = \varphi(x)$; $f(x)$ — плотность распределения случайной величины X; $g(y)$ — плотность распределения случайной величины Y; $\varphi(x)$ — монотонная функция; $x = \psi(y)$. Тогда $g(y) = f(\psi(y))|\psi'(y)|$.

По определению $y = \sqrt{x} = \varphi(x)$; $x = y^2 = \psi(y)$; $\psi'(y) = 2y$;

$$g(y) = \frac{(y^2)^{\frac{n-3}{2}} e^{-\frac{y^2}{2}}}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} 2y = \frac{y^{n-2} e^{-\frac{y^2}{2}}}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}$$

— это табличная функция χ («хи»).

По данному γ ищем доверительный интервал J для ψ из условия

$$\gamma = P(|\sigma - S_1| < \sigma = P(S_1 - \delta) = P\{S_1(1 - q) < \sigma < S_1(1 + q)\},$$

где $q = \delta/S_1$.

При $q < 1$ имеем

$$\gamma = P\left\{\frac{1}{S_1(1+q)}\right\} < \frac{1}{q} < \frac{1}{S_1(1-q)} = P\left\{\frac{\sqrt{n}}{1+q} < \frac{S_1\sqrt{n}}{\sigma} < \frac{\sqrt{n}}{1-q}\right\} = P(\chi_1 < \chi < \chi_2),$$

где $\chi_1 = \dfrac{\sqrt{n}}{1+q}$; $\chi_2 = \dfrac{\sqrt{n}}{1-q}$, так что $\gamma = \int_{x_1}^{x_2} g(y)dy$.

По таблице распределения «хи» по γ и n находим q, а S_1 вычисляем по выборке $\Rightarrow J_\sigma = \{S_1(1 - q) < \sigma < S_1(1 + q)\}$.

При $q > 1$ имеем

$$\gamma = P\{S_1(1-q) < \sigma < S_1(1+q)\} = P\{0 < \sigma < S_1(1+q)\},$$

так как $\sigma \geqslant 0 \Rightarrow J = \{0, S_1(1+q)\}$, так как

$$\gamma = P\{\sigma < S_1(1+q)\} = P\left\{\frac{1}{\sigma} > \frac{1}{S_1(1+q)}\right\} = P\left\{\frac{S_1\sqrt{n}}{\sigma} > \frac{\sqrt{n}}{1+q}\right\} = P(\chi > \chi_1),$$

где $\chi_1 = \dfrac{\sqrt{n}}{1+q}$, так что $\gamma = \int\limits_{\gamma_1}^{\infty} g(y)dy$.

В итоге получаем

$$J_\sigma = \begin{cases} \{S_1(1-q), S_1(1+q)\}, \text{ при } q < 1, \\ \{0, S_1(1+q)\}, \text{ при } q > 1. \end{cases}$$

Приведем числовые примеры.

Задача 2.48. Дана выборка $\vec{x} = (2{,}5;\ 2;\ -2{,}3;\ 1{,}9;\ -2{,}1;\ 2{,}4;\ 2{,}3;\ -2{,}5;\ 1{,}5;\ 1{,}7)$; $\gamma = 0{,}95$. Построить доверительный интервал для MX.

Решение

$n = 10$; $\overline{x} = \dfrac{1}{10}\sum\limits_{i=1}^{10} x_i = 0{,}4$; $S_1^2 = \dfrac{1}{10}\sum\limits_{i=1}^{n}(x_i - \overline{x})^2 = 5$; $\hat{\sigma} = \sqrt{5}$; $n - 1 = 9$; $t_\gamma =$ $= 2{,}26$;

$$\sigma = \frac{t_\gamma}{\sqrt{n}} = \frac{2{,}26\cdot\sqrt{5}}{\sqrt{10}} \approx 1{,}58 \Rightarrow J = (0{,}4 \mp 1{,}58) = (1{,}58;\ 1{,}98).$$

Задача 2.49. $X \sim N(m, \sigma)$, m — неизвестно. Найти J для σ:

а) $n = 25$; $\sqrt{S_1^2} = S_1 = 0{,}8$; $\gamma = 0{,}95$;

б) $n = 10$; $S_1 = 0{,}16$; $\gamma = 0{,}99$;

Решение

а) По таблице распределения χ находим значение $q = 0{,}32 \Rightarrow$ искомый ДИ J для σ есть $J = \{0{,}8\cdot(1 \mp 0{,}32)\} = \{0{,}544;\ 1{,}056\}$.

б) По таблице распределения χ находим значение $q = 1{,}8$ $(q > 1) \Rightarrow$ $\Rightarrow$ искомый ДИ для σ есть $J = \{0;\ 0{,}16(1+1{,}8)\} = \{0;\ 0{,}448\}$.

Задача 2.50. СВ $X \sim N(m, \sigma)$, γ— заданный уровень доверия. Найти J_σ и J_{σ^2} ДИ для σ и $\sigma^2 = DX$ в случаях:

а) $MX = m$ — известно;

б) $MX = m$ — неизвестно.

Решение

а) Построим статистику $S_0^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - m)^2 = \sum\limits_{i=1}^{n}\left(\dfrac{x_i - m}{\sqrt{n}}\right)^2$. Найдем распределение S_0^2 методом характеристических функций:

$$g_{\frac{x_i - m}{\sqrt{n}}}(t) = g_{\frac{x_i}{\sqrt{n}} - \frac{m}{\sqrt{n}}}(t) = \left\{g_{ax+b(x)} = e^{itb} g_x(at)\right\} = e^{it\frac{m}{\sqrt{n}}} e^{im\frac{t}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2n}} = e^{\frac{\left(\frac{\sigma}{\sqrt{n}}\right)^2 t}{2}} \Rightarrow$$

$$\Rightarrow \frac{x_i - m}{\sqrt{n}} \sim N\left(0,\ \sigma^* = \frac{\sigma}{\sqrt{n}}\right) \Rightarrow S_0^2 = (\sigma^*)^2 u,$$

где $u \sim \chi_n^2 \Rightarrow \dfrac{S_0^2}{(\sigma^*)^2} \sim \chi_n^2$, или $\dfrac{S_0^n n}{\sigma^2} \sim \chi_n^2 \Rightarrow$

$$\Rightarrow \int_{\chi_1^2}^{\chi_2^2} k_n(x)dx = \gamma,$$

где $k_n(x)$ — плотность распределения χ_n^2, а значения χ_1^2 и χ_2^2 находятся по таблице из условий

$$\begin{cases} \displaystyle\int_0^{\chi_1^2} k_n(x)dx = \frac{1-\gamma}{2}, \\ \displaystyle\int_{\chi_2^2}^{\infty} k_n(x)dx = \frac{1-\gamma}{2}. \end{cases}$$

Тогда

$$\gamma = P\left\{\chi_1^2 < \frac{nS_0^2}{\sigma^2} < \chi_2^2\right\} = P\left\{\frac{nS_0^2}{\chi_2^2} < \sigma^2 < \frac{nS_0^2}{\chi_1^2}\right\} \Rightarrow$$

$$\Rightarrow J_{DX} = \left\{\frac{nS_0^2}{\chi_2^2}; \frac{nS_0^2}{\chi_1^2}\right\} \Rightarrow J_\sigma \left\{\frac{\sqrt{n}\cdot S_0}{\chi_1}; \frac{\sqrt{n}\cdot S_0}{\chi_2}\right\}.$$

б) Построим статистику $S_1^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - \overline{x})^2$. Известно, что $\dfrac{nS_1^2}{\sigma^2} \sim \chi_{n-1}^2$ с плотностью распределения $k_{n-1}(x)$, γ задано. Тогда из условий $\displaystyle\int_{\chi_1^2}^{\chi_2^2} k_{n-1}(x)dx = \gamma$ и $\begin{cases} \displaystyle\int_0^{\chi_1^2} k_{n-1}(x)dx = \frac{1-\gamma}{2}, \\ \displaystyle\int_{\chi_2^2}^{\infty} k_{n-1}(x)dx = \frac{1-\gamma}{2} \end{cases}$ находим по таблицам χ^2 значения χ_1^2 и χ_2^2 и получаем

$$P\left\{\chi_1^2 < \frac{nS_1^2}{\sigma^2} < \chi_2^2\right\} = P\left\{\frac{nS_1^2}{\chi_2^2} < \sigma^2 < \frac{nS_1^2}{\chi_1^2}\right\} \Rightarrow$$

$$\Rightarrow J_{DX} = \left\{\frac{nS_1^2}{\chi_2^2}; \frac{nS_1^2}{\chi_1^2}\right\} \Rightarrow J_\sigma = \left\{\frac{\sqrt{n}\cdot S_1}{\chi_1}; \frac{\sqrt{n}\cdot S_1}{\chi_2}\right\},$$

где $S_1 = \sqrt{S_1^2}$, а χ_1 и χ_2 имеют табличное распределение χ («хи»).

Числовой пример: $n = 22$; $\overline{x} = 3{,}031$; $\widehat{\sigma}^2 = S_1^2 = 0{,}03$; $\gamma = 0{,}98$; $\chi_1^2 = 8{,}897$; $\chi_2^2 = 38{,}932$ (по таблицам χ^2). Тогда

$$\frac{nS_1^2}{\chi_1^2} = 0{,}0715;\ \frac{nS_1^2}{\chi_1^2} = 0{,}0163;\quad J_{DX} = \{0{,}0163; 0{,}0715\}.$$

Контрольные вопросы и задания

1. Укажите свойства оценок выборочного среднего $\overline{x}$ и выборочной дисперсии $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ соответственно для MX и DX.

2. Сравните качество оценок по квадратичному риску для оценок $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ и $S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x - \overline{x})^2$.

3. Приведите пример сравнения качества оценок в данном классе оценок.

4. Приведите пример использования теоремы Рао — Блэкуэлла — Колмогорова для улучшения качества несмещенной оценки.

5. Как осуществляется нахождение достаточных статистик для распределений экспоненциального типа? Приведите примеры.

6. Как используется полнота достаточных статистик для построения оптимальных оценок определенных функций неизвестного параметра распределения? Приведите примеры.

7. Приведите примеры существования и несуществования эффективной оценки и объясните свои выводы.

8. Приведите пример бесконечного числа оценок максимального правдоподобия для неизвестного параметра распределения.

9. Какие есть проблемы при построении доверительных интервалов для неизвестного параметра распределения и каковы подходы к их решению?

10. Приведите примеры центральных статистик при построении доверительных интервалов.

Глава 3
ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

В результате изучения данной главы студенты должны:

знать

• основные понятия теории проверки статистических гипотез;

• наиболее практически значимые критерии согласия Пирсона, Колмогорова и параметрический критерий Неймана — Пирсона;

уметь

• применять изученные критерии для проверки гипотез по конкретным наблюдениям, а также определять их мощность;

владеть

• методами математической статистики для решения задач;

• методами проверки статистических критериев.

Материал данной главы дает представление о широте применения статистических методов в практических исследованиях, которая определяется неограниченностью выбора проверяемых гипотез.

Рассматривается подход к решению задачи принятия или неприятия проверяемой гипотезы путем построения критерия (правила) ее проверки с оценкой вероятностей ошибок первого и второго рода, т.е. ошибочного отвержения или принятия проверяемой гипотезы и общего подхода сравнения качества разных сравнимых критериев проверки гипотез по их мощностям.

В качестве конкретных примеров здесь рассматриваются часто встречаемые на практике случаи проверяемых гипотез о согласии наблюдаемой выборки предполагаемому закону распределения генеральной совокупности (критерии согласия) и проверки гипотез о конкретном значении параметра распределения наблюдаемой случайной величины (критерий Неймана — Пирсона).

Одной из *основных* задач математической статистики является проверка статистических гипотез и исследование качества этой проверки. Определим основные понятия, связанные с процедурой проверки и сравнения гипотез о распределении изучаемой СВ X.

Пусть $\vec{x} = (x_1, x_2, ..., x_n)$ — результат независимых наблюдений над СВ X. Пусть H_0 — основная проверяемая гипотеза о распределении этой СВ, а H_1 — альтернативная (конкурирующая) гипотеза. Требуется по наблюденной выборке $\vec{x}$ принять решение о правильности основной гипотезы H_0, т.е. установить правило (критерий), по которому наблюденная выборка согласуется или нет с высказанной гипотезой H_0. Это соответствует делению выборочного пространства на две области: область принятия гипотезы (H_0) T и критическую область R (где гипотеза H_0 отвергается).

В результате проверки гипотезы H_0 могут возникать ошибки, состоящие в том, что будет отвергнута верная гипотеза (H_0) (ошибка первого рода), или в том, что будет принята неверная гипотеза (H_1) (ошибка второго рода). Вероятность ошибки первого рода α называется еще ***уровнем значимости критерия***, или ***объемом критерия***. Поэтому для области принятия гипотезы T (или критерия T) часто используется обозначение T_α.

Вероятность ошибки второго рода обозначим через β. Тогда

$$\alpha = P(\vec{x} \in R/H_0); \quad \beta = P(\vec{x} \in T/H_1).$$

Определение 3.1. ***Мощностью критерия*** W называется вероятность отвергнуть гипотезу H_1, если она неверна, т.е. это есть

$$W = 1 - \beta = P(\vec{x} \in R/H_1) = W(T_\alpha/H_1).$$

При фиксированном значении вероятности ошибки первого рода α из двух критериев T_1 и T_2 лучше тот, у которого меньше вероятность ошибки второго рода β, т.е. у которого больше мощность W.

Определение 3.2. Если для любого критерия T $W(T^*/H_1) > W(T/H_1)$, то критерий T^* называется ***равномерно наиболее мощным*** (РНМК) для проверки основной гипотезы H_0 против альтернативы H_1.

Определение 3.3. Гипотеза называется ***простой***, если соответствующее ей множество состоит из одного элемента. В противном случае это ***сложная*** гипотеза.

В качестве примеров рассмотрим критерии двух типов: критерии согласия и РНМК Неймана — Пирсона.

3.1. Критерии согласия

Критерии согласия позволяют судить о соответствии наблюденных значений выборки $\vec{x}$ предполагаемому в гипотезе H_0 закону распределения. Если все параметры этого закона известны, то гипотеза H_0 — простая, так как она представлена одним конкретным распределением, в случае же неизвестных параметров гипотеза H_0 — сложная.

Критерии согласия основаны на выборе определенной меры D расхождения между теоретическим (гипотетическим — определенным гипотезой H_0) и эмпирическим распределениями СВ X. В зависимости от способа задания этой меры D получаются различные критерии для проверки гипотезы H_0.

Для построения критерия согласия выбирается ***мера расхождения*** теоретического и эмпирического распределений D, задается уровень значимости критерия (вероятность ошибки первого рода) α, в случае сложной гипотезы H_0 неизвестные параметры определенным образом оцениваются по выборке (подробнее об этом будет сказано ниже) и ищется значение D_α такое, что $P(D \geqslant D_\alpha) = \alpha$. Тогда D_α называется ***пределом значимости критерия***. Таким образом критерий построен, т.е. все выборочное пространство разделено на две области T, где $D < D_\alpha$ (т.е. верна гипотеза H_0), и R, где $D \geqslant D_\alpha$.

Для определения значения D_α при заданном уровне значимости α необходимо знать распределение статистики D, выяснение чего часто затруднительно. Поэтому критерии согласия обычно применяют при больших объемах n выборок, позволяющих заменять точное распределение статистики D на ее предельное (при $n \to \infty$).

Рассмотрим теперь для примера некоторые наиболее распространенные способы задания меры D и соответствующие им критерии согласия.

Критерий согласия χ^2 (критерий Пирсона)

Этот критерий применяется при больших выборках, так как использует предельное распределение (при $n \to \infty$) меры D. Он имеет вид

$$\chi^2 = \sum_{i=1}^{r} \frac{(m_i - np_i)^2}{np_i}, \tag{3.1}$$

где r — число интервалов, на которые разбита вся область возможных значений СВ X; m_i — число значений выборки, попавших в i-й интервал ($i = 1, ..., r$); p_i — теоретическая вероятность (т.е. вероятность по гипотезе H_0) того, что значение наблюдаемой СВ X попадет в i-й интервал. Если теоретическое распределение по гипотезе H_0 соответствует дискретной СВ X, то интервалы разбиения области ее наблюденных значений заменяются фиксированными СВ X.

Покажем, что статистика χ^2, заданная в виде формулы (3.1), представляет из себя меру D расхождения теоретического и эмпирического распределений, для чего представим статистику (3.1) в виде

$$\chi^2 = \sum_{i=1}^{r} (n/p_i)(m_i/n - p_i)^2, \tag{3.2}$$

где m_i/n — частота попадания СВ X в i-й интервал, т.е. m_i/n — приращение эмпирической функции распределения, а p_i — приращение теоретической функции распределения на i-м интервале, поэтому вид (3.2), включающий разность этих приращений, представляет собой меру D.

К. Пирсоном доказано, что если проверяемая гипотеза H_0 простая, то статистика критерия χ^2 (3.1) имеет при $n \to \infty$ распределение χ^2 с $k = r - 1$ степенями свободы с функцией распределения

$$F_{r-1}(x) = \frac{1}{2^{\frac{r-1}{2}} \Gamma\left(\frac{r-1}{2}\right)} \int_0^x u^{\frac{r-1}{2}} \exp(-u/2)\, du, \; x \geqslant 0.$$

Случай простой гипотезы H_0 или полностью определенного гипотетического распределения встречается на практике очень редко.

Значительно чаще распределение содержит некоторое количество неизвестных параметров, которые могут быть оценены по выборке, тогда в формуле (3.1) $p_i = p_i(\theta)$, где θ — неизвестный (возможно, векторный) параметр теоретического распределения.

Р. Фишер доказал, что при весьма общих условиях существуют методы оценок параметров теоретического распределения СВ X, при которых предельным распределением статистики χ^2 (3.1) при $n \to \infty$ также является распределение χ^2 с числом степеней свободы $k = r - s - 1$, где s — число оцениваемых параметров.

Одним из вышеупомянутых методов оценивания неизвестных параметров теоретического распределения является метод максимального правдоподобия, основанный на частотах $m_1, m_2, ..., m_r$, т.е. когда в качестве оценки для векторного параметра θ в критерии (3.1), где $p_i = p_i(\theta)$, используют мультиномиальную оценку максимального правдоподобия (МОМП) (см., например, [7]).

Суть этого метода состоит в следующем: с любым способом группировки связано соответствующее полиномиальное (или мультиномиальное) распределение вектора частот $m = (m_1, m_2, ..., m_r)$ с параметрами $n = m_2 + ... + m_r$ и $p_i = p_i(\theta)$, $i = 1, ..., r$ (при $r = 2$ имеем биномиальное распределение). Введем функцию

$$L(m, \theta) = n! \prod_{i=1}^{r} p_i^{m_i}(\theta) / m_i!.$$

Тогда МОМП для параметра θ дает в качестве оценки ее значение, максимизирующее функцию $L(m, \theta)$ по θ. Система уравнений для отыскания этой оценки имеет вид

$$d \ln L(m, \theta)/d\theta_j = 0, \quad j = 1, ..., s,$$

где s — размерность параметра θ.

Таким образом, асимптотическое поведение статистики χ^2 становится известным как при простой, так и при сложной гипотезе H_0 — оно есть распределение χ^2 с соответствующим числом степеней свободы и не зависит от предполагаемого в гипотезе H_0 распределения СВ X, т.е. критерий χ^2 является асимптотически непараметрическим.

При достаточно большом объеме исходной выборки $x_1, x_2, ..., x_n$ и выполнении определенных требований к ее статистической обработке, о которой будет сказано ниже, распределение статистики χ^2 становится близким к ее предельному распределению.

Для обоснованного применения критерия согласия χ^2 при гипотезе H_0 предполагается, что биномиальное распределение частоты m_i, $i = 1, ..., r$, может быть приближено нормальным. Такой предельный переход осуществляется достаточно быстро, если $np_i \geqslant 10$. Для выполнения этого требования при вычислении статистики χ^2 необходимо объединять интервалы области наблюденных значений СВ X. Кроме этого, общее число интервалов должно быть не меньше восьми.

Предельное поведение статистики χ^2 табулировано и зависит от числа степеней свободы k. Оно показывает, какое количество наблюденных частот может принимать произвольные значения с учетом числа независимых линейных связей, которым частоты подчиняются. Такие линейные связи обусловливаются числом наблюдений над СВ X и числом неизвестных оцениваемых параметров теоретического распределения, заданного гипотезой H_0. Это объясняет приведенные выше выражения для чисел степеней свободы в случаях простой и сложной нулевой гипотезы H_0.

В таблицах обычно приводятся значения статистики $\chi^2_{1-\alpha}$ границы критической области R размера α (вероятности ошибки первого рода) в зависимости от уровня значимости критерия α и числа степеней свободы k в соответствии с формулой

$$\alpha = P(\chi^2 \geqslant \chi^2_{1-\alpha}) = 1/(\Gamma(k/2)2^{k/2})) \int\limits_{\chi^2_{1-\alpha}}^{\infty} x^{k/2-1} \exp(-x/2)du.$$

Отсюда $1 - \alpha = P(\chi^2 < \chi^2_{1-\alpha})$, т.е. граница критической области размера α есть $(1 - \alpha)$ — квантиль распределения χ^2 с k степенями свободы, чем и объясняется ее обозначение $\chi^2_{1-\alpha}$.

Статистика χ^2 из формулы (3.1) легко преобразуется к виду

$$\chi^2 = \sum_{i=1}^{r} (m_i / np_i) - n \tag{3.3}$$

следующим образом:

$$\chi^2 = \sum_{i=1}^{r}(m_i/np_i)/np_i = \sum_{i=1}^{r}(m_i/np_i) - 2\sum_{i=1}^{r} m_i + n\sum_{i=1}^{r} p_i = \sum_{i=1}^{r}(m_i/np_i) - n,$$

так как $\sum_{i=1}^{r} m_i = n; \sum_{i=1}^{r} p_i = 1.$

Выражение (3.3) для статистики χ^2 может быть использовано как для непосредственного вычисления значения χ^2, так и для проверки вычисления по формуле (3.1).

Правило проверки гипотезы H_0 по критерию χ^2 состоит в следующем: по заданной вероятности ошибки первого рода α и вычисленному числу степеней свободы k по таблицам распределения χ^2 определяют границу $\chi^2_{1-\alpha}$ критической области R, и гипотеза H_0 отвергается, если $\chi^2 < \chi^2_{1-\alpha}$; если же $\chi^2 > \chi^2_{1-\alpha}$, то считаем отклонение эмпирического распределения от теоретического случайным и гипотеза H_0 принимается.

Иногда исходными данными таблицы распределения χ^2 является вычисленное значение статистики χ^2 и число степеней свободы k, а выходным — значение α^* — вероятности того, что полученное значение $\chi^2 \geqslant \chi^2_{1-\alpha}$ при гипотезе H_0. Тогда гипотеза H_0 принимается, если заданный уровень значимости критерия α оказывается меньше α^*, и отвергается — в противном случае.

Таким образом, алгоритм пошагового применения критерия χ^2 для проверки гипотезы H_0 состоит в следующем.

1. Весь диапазон наблюденных значений исследуемой СВ X разбивают на r интервалов (не обязательно одинаковой длины) так, что их число не меньше 8, а в каждом интервале не менее 10 наблюденных значений СВ X.

2. Если параметры гипотетического распределения (по гипотезе H_0) неизвестны, то они оцениваются по выборке, как описано выше.

3. Для каждого интервала группирования h_i определяют частоту m_i, $i = 1, ..., r$, попавших в него значений выборки наблюдений над СВ X.

4. Вычисляют значение p_i, $i = 1, ..., r$, как вероятность попадания значения СВ X в интервал при гипотезе H_0 по формуле

$$p_i = P(x_i \leqslant X \leqslant x_{i+1}) = F(x_{i+1}) - F(x_i),$$

где неизвестные параметры гипотетического распределения (если они есть) заменены их оценками.

5. Вычисляют значение статистики χ^2 по формуле (3.1) или (3.2).

6. Определяют число степеней свободы k предельного распределения χ^2.

7. По таблице распределения χ^2 с k степенями свободы находят значение границы критерия $\chi^2_{1-\alpha}$ с данным уровнем значимости α

(или, в случае других таблиц, — значение α^* допустимого уровня значимости α по вычисленному значению статистики χ^2).

8. По описанному выше правилу делают вывод о согласии или несогласии выборочных данных с предполагаемым в гипотезе H_0.

В. И. Романовский предложил правило, значительно облегчающее применение критерия χ^2, основанное на оценке расхождения между эмпирическим и теоретическим распределениями. Оно состоит в следующем: если $Y = |\chi^2 - k|/\sqrt{2k} \geqslant 3$, то расхождение считается существенным, в противном случае — случайным. Это правило основывается на том, что математическое ожидание и дисперсия СВ с распределением χ^2 соответственно есть k и $2k$, откуда следует, что СВ Y при большом объеме наблюдений n имеет приближенно нормальное распределение $N(0, 1)$ и поэтому вероятность отклонения СВ X от своего среднего менее чем на $3\sigma_{\chi^2} = 3\sqrt{2k}$ в ту и другую сторону близка к 1. Этот способ можно применять для предварительного анализа данных опыта, т.е. для быстрой грубой проверки вывода, получаемого по критерию χ^2.

Пример 3.1. Рассмотрим пример применения критерия χ^2 для проверки с уровнем значимости $\alpha = 0{,}05$ гипотезы H_0 о согласии данных наблюдений над СВ X с законом распределения Пуассона.

Рекомендуется для удобства наблюденные значения СВ X и результаты счета записывать в следующую таблицу:

№ п/п	x_i	m_i	p_i	np_i	$m_i - np_i$	$(m_i - np_i)^2$	$\frac{(m_i - np_i)^2}{np_i}$
1	0	57	0,021	54,8	2,2	4,84	0,088
2	1	203	0,081	211,2	–8,2	67,24	0,318
3	2	383	0,156	406,8	–23,8	566,44	1,392
4	3	525	0,201	524,2	0,8	0,64	0,001
5	4	532	0,195	508,6	23,4	547,56	1,007
6	5	408	0,151	393,7	14,2	201,64	0,022
7	6	273	0,097	253,0	20,0	400,00	1,581
8	7	139	0,054	140,8	–1,8	3.24	0,023
9	8	45	0,026	67,8	–22,8	519,84	7,667
10	9	27	0,011	28,7	–1,7	2,89	0,101
11	10	16	0,007	18,3	–2,3	5,29	0,289
—	—	2608	1,000	—	—	—	13,05

Здесь $k = 11 - 1 - 1 = 9$; $P\{X = x_i\} = \exp(-\lambda)\lambda^{x_i}/x_i!$, оценка для неизвестного параметра λ по [1] $\tilde{\lambda} = (1/n)\sum_{i=1}^{11} x_i m_i = (1/2608)\,(0\cdot 57 + 1\cdot 203 + \ldots + 10\cdot 16) = 3{,}870$; $p_i = P_L(X = x_i) = \exp(-3{,}87)3{,}87^{x_i}/x_i!$ (эти значения p_i находятся по таблице распределения Пуассона).

Далее по таблице распределения χ^2 при $k = 9$ и $\chi^2 = 13{,}05$ находим значение вероятности $P(\chi^2 \geqslant \chi^2_{1-\alpha}) = 0{,}166 > \alpha = 0{,}05$, поэтому считаем, что отклонение наблюденных данных от закона Пуассона незначимо и гипотезу H_0 не отвергаем.

В рассмотренном примере способ Романовского дает значение дроби $|\chi^2 - k|/\sqrt{2k} = (13{,}65 - 9)/\sqrt{18} \approx 1{,}1 < 3$, т.е. расхождение между наблюдениями и теоретическим распределением можно считать случайным, что совпадает с ранее полученным выводом.

Критерий согласия Колмогорова

Этот критерий применяется при большом объеме выборки n, когда гипотетическая (по гипотезе H_0) функция распределения $F(x)$ СВ X полностью определена и непрерывна. Здесь мера отклонения D эмпирической функции распределения F_n(x) выборки $x_1, ..., x_n$ от $F(x)$ определяется следующим образом:

$$D = D(F_n(x), F(x)) = \max|F_n(x) - F(x)|, \qquad (3.4)$$

где

$$F(x) = \begin{cases} 0, & x \leqslant x_{(1)}, \\ k/n, & x_{(k)} < x \leqslant x_{(k+1)}, \\ 1, & x > x_{(n)}. \end{cases} \qquad (3.5)$$

При $n \to \infty$ в предположении правильности гипотезы H_0 доказано, что закон распределения СВ $L = D/\sqrt{n}$ (независимо от закона распределения СВ X) стремится к закону распределения Колмогорова, т.е.

$$P\{D\sqrt{n} < x\} \underset{n\to\infty}{\rightarrow} K(x) = \begin{cases} \sum\limits_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2/x^2), x > 0, \\ 0, x \leqslant 0. \end{cases}$$

Функция распределения K(x) табулирована, в таблице даны вероятности $\alpha = P\{D\sqrt{n} > L_\alpha\} = 1 - K(L_\alpha)$, где L_α — заданное число, зависящее от α — уровня значимости критерия. Значение L_α при данном α также может быть определено по таблице распределения Колмогорова.

Применение критерия согласия Колмогорова состоит в выполнении следующих шагов.

1. По выборке $x_1, ..., x_n$ находят по формуле (3.5) $F_n(x)$.

2. Вычисляют значение D по формуле (3.4) по всем точкам выборки.

3. Вычисляют значение $L = D\sqrt{n}$.

4. По таблице распределения $K(x)$ находят по заданному значению α значение L_α.

5. Если $L < L_\alpha$, то гипотезу H_0 считают не противорячащей выборке $x_1, \ldots, x_n$.

Иногда таблицы распределения Колмогорова используют по-другому, по вычисленному значению L определяют значение $1 - K(L)$. Тогда гипотеза H_0 не отвергается, если при заданном $dK(L) \geqslant d$.

Пример 3.2. Рассмотрим пример применения критерия согласия Колмогорова для проверки гипотезы H_0 о нормальности СВ X (по группированной выборке $x_1, \ldots, x_n$) с $m = MX = 100{,}25$ и $\sigma = 1$ при $n = 1000$ и уровне значимости $\alpha = 0{,}05$. Данные опыта и результаты вычислений представим в виде таблицы:

i	x_i	m_i	$z_i = x_i - \overline{x}$	$\Phi(z_i)$	$F(x_i)$	$F_n(x_i)$	$\lvert F_n(x_i) - F(x_i)\rvert$
1	98,0	21	–2,25	0,0123	0,0123	0,105	0,0018
2	98,5	47	–1,75	0,0401	0,0401	0,0445	0,0044
3	99,0	87	–1,25	0,1056	0,1056	0,1115	0,0059
4	99,5	158	–0,75	0,2266	0,2266	0,2340	0,0074
5	100,1	181	–0,25	0,4013	0,4013	0,4035	0,0022
6	100,1	201	0,25	0,5987	0,5987	0,5945	0,0042
7	101,0	142	0,75	0,7734	0,7734	0,7660	0,0074
8	101,5	97	1,25	0,8944	0,8944	0,8855	0,0089
9	102,0	41	1,75	0,9599	0,9599	0,9545	0,0054
10	102,5	25	2,25	0,9877	0,9877	0,9875	0,0002

Здесь $P(x) = \Phi(x - m)$, так как $\sigma = 1$; $\Phi(x) = \dfrac{1}{\sqrt{2n}} \int\limits_{-\infty}^{x} \exp(-u^2/2)du$;

$$F_n(x) = 0{,}001\left(\sum_{j=1}^{i-1} m_j + 0{,}5m\right).$$

По таблицам распределения Колмогорова $L_\alpha = 1{,}358$. По нашей таблице получаем $L = 0{,}0089\sqrt{1000} = 0{,}281 < 1{,}358$. Отсюда следует, что гипотетическое нормальное распределение можно считать согласующимся с данной выборкой.

3.2. Критерий Неймана — Пирсона

Этот критерий задает процедуру различения двух простых гипотез о параметре закона распределения изучаемой случайной величины. Обсудим его суть.

Лемма Неймана — Пирсона при простых гипотезах H_0 и H_1 дает достаточно общее решение при нахождении РНМК в виде ***критерия отношения правдоподобия***:

$$T = X_{1,\alpha} = \{x : L(\vec{x}/H_1)/L(\vec{x}/H_0) \geqslant K_\alpha\},$$

где $L = L(\vec{x}/H)$ — функция правдоподобия, а значение K_α определяется из таблиц.

Пример 3.3. Рассмотрим пример построения РНМК (по Нейману — Пирсону) по выборке $x_1, \ldots, x_n$ из генеральной совокупности, распределенной по закону Пуассона $\pi(\lambda)$ для проверки простой основной (нулевой) гипотезы H_0, состоящей в том, что $\lambda = \lambda_0$, против альтернативы H_1 о том, что $\lambda = \lambda_1$. Здесь критерий Неймана — Пирсона имеет вид

$$L(\vec{x}/H_1)/L(\vec{x}/H_0) = \frac{\lambda_1^{\sum_{i=1}^{n} x_i} \exp(-\lambda_1 n)}{\lambda_0^{\sum_{i=1}^{n} x_i} \exp(-\lambda_0 n)} = (\lambda_1/\lambda_0)^{\sum_{i=1}^{n} x_i} \exp[n(\lambda_0 - \lambda_1)] \geqslant K_\alpha,$$

или

$$(1/n)\sum_{i=1}^{n} x_i(\ln\lambda_1 - \ln\lambda_0) + (\lambda_0 - \lambda_1) \geqslant (1/n)\ln K_\alpha,$$

откуда получаем

$$(1/n)\sum_{i=1}^{n} x_i \geqslant n[(1/n)\ln K_\alpha - (\lambda_0 - \lambda_1)]/(\ln\lambda_1 - \ln\lambda_0) = K'_\alpha.$$

Тогда критическая область определяется следующими соотношениями:

1) $\sum_{i=1}^{n} x_i \geqslant K_\alpha$ при $\lambda_0 < \lambda_1$;

2) $\sum_{i=1}^{n} x_i < K_\alpha$ при $\lambda_0 > \lambda_1$.

Статистика $\sum_{i=1}^{n} x_i$ имеет распределение Пуассона с параметром $n\lambda$, если слагаемые имеют распределение Пуассона с параметром λ.

Пусть, например, $d = 0{,}02$; $n = 20$; $\lambda_0 = 0{,}1$; $\lambda_1 = 0{,}2$. Здесь $\lambda_0 < \lambda_1$, поэтому критическая область определяется соотношением $\sum_{i=1}^{n} x_i \geqslant K_\alpha = X_\alpha$.

Значение X_α определяется из выражения для ошибки первого рода α:

$$P\left\{\sum_{i=1}^{20} x_i < X_\alpha / H_0 : \lambda_0 = 0{,}1\right\} = 0{,}02.$$

Статистика $\sum_{i=1}^{n} x_i$ распределена здесь по Пуассону с параметром $n\lambda$, поэтому формула перепишется в виде

$$1 - \sum_{k=0}^{X_\alpha} (\lambda_0 n)^k \exp(-\lambda_0 n)/k! = \alpha,$$

или

$$1 - \sum_{k=0}^{X_\alpha} (0{,}1 \cdot 20)^k \exp(-0{,}1 \cdot 20) = 0{,}02,$$

или

$$1 - \sum_{k=0}^{X_\alpha} 2^k \exp(-2)/k! = 0{,}98.$$

Отсюда по таблице распределения Пуассона находим $X_\alpha = 5$.

Вероятность ошибки второго рода здесь вычисляется из выражения

$$\beta = P\left\{\sum_{i=1}^{n} x_i < 5 / H_1 : \lambda_1 = 0{,}2\right\} = 1 - \sum_{k=0}^{5} (0{,}2 \cdot 20)^k \exp(-0{,}2 \cdot 20) / k! =$$

$$= 1 - \sum_{k=0}^{5} 4^k \exp(-4) / k! \approx 0{,}37.$$

Мощность критерия $W = 1 - \beta \approx 0{,}63$.

Задание 3.1. Самостоятельно применить критерий Неймана — Пирсона для различения простых гипотез и определить мощность критерия в следующих случаях распределений генеральных совокупностей и различаемых гипотез:

1) $N(m, l)$; H_0: $m = m_0$; H_1: $m = m_1$ ($m_0 = 2$; $m_1 = 2{,}5$);
2) $B(n, p)$; H_0: $p = p_0$; H_1: $p = p_1$ ($p_0 = 0{,}4$; $p_1 = 0{,}5$);
3) $\Gamma(p)$; H_0: $p = p_0$; H_1: $p = p_1$ ($p_0 = 0{,}7$; $p_1 = 0{,}8$);
4) $N(0, \sigma^2)$; H_0: $\sigma^2 = \sigma_0^2$; H_1: $\sigma^2 = \sigma_1^2$ ($\sigma_0 = 3$; $\sigma_1 = 2$).

Контрольные вопросы и задания

1. Выпишите аналитические формулы для вероятностей ошибок 1-го и 2-го рода и мощности критерия. Как сравниваются качества сравнимых критериев?
2. Каковы ограничения в применении критерия Пирсона (хи-квадрат) и вычисления числа степеней свободы статистики критерия Пирсона?
3. В чем заключается смысл статистики критерия Пирсона?
4. Каковы ограничения в применении критерия Колмогорова? Укажите вид и смысл статистики критерия.
5. Какова критическая область критерия Неймана — Пирсона?
6. Как осуществляется проверка двух простых гипотез о параметре a по выборке из нормального распределения $N(a, \sigma)$ при известном значении параметра σ?
7. Как осуществляется проверка двух простых гипотез о параметре по выборке из распределения Пуассона $\Pi(\lambda)$?

Раздел II

СЛУЧАЙНЫЕ ПРОЦЕССЫ

В результате изучения данного раздела студенты должны:

знать

• основы теории случайных процессов;

• цепи Маркова, марковские, пуассоновские процессы;

• ветвящиеся процессы;

• числовые характеристики случайных процессов;

уметь

• решать задачи теории случайных процессов;

• применять анализ марковских процессов к решению задач теории массового обслуживания;

владеть

• методами анализа случайных процессов.

В данном разделе приводятся общие сведения по теории случайных процессов, а также подробно обсуждаются свойства марковских случайных процессов, и в частности цепей Маркова (марковских цепей). Рассмотрена классификация состояний цепи Маркова с установлением многочисленных связей свойств состояний. Представлены основы теории ветвящихся процессов. Подробно изучены свойства пуассоновских процессов.

Глава 4
ЦЕПИ МАРКОВА

В результате изучения данной главы студенты должны:

знать

• основные сведения о случайных процессах и свойствах цепей и процессов Маркова;

уметь

• проверять марковское свойство;

• рассчитывать вероятности состояний цепи Маркова;

владеть

• навыками анализа цепей Маркова и моделированием их значений.

В главе даются основные понятия теории случайных процессов и их классификация. В качестве одного из важнейших объектов начального курса подробно рассматриваются марковские процессы с дискретным временем. Подробно объяснена их специфика, свойства и задание через матрицу переходных вероятностей и вектора начальных вероятностей процесса. Изучаются свойства характеристической матрицы. Приведено много примеров цепей Маркова с проверкой марковского свойства и с расчетом вероятностей состояний за фиксированное число шагов.

4.1. Начальные сведения о случайных процессах

Любая меняющаяся система, находящаяся под влиянием случайных факторов, представляет собой случайный процесс. Иначе: ***случайным процессом*** называется семейство случайных величин, зависящих от параметра t, пробегающего произвольное множество T. В частности, если множество T состоит из одной точки, мы имеем случайную величину. Значение $X(t)$ случайной величины при фиксированном значении параметра t называется ***значением случайного процесса*** в данный момент времени. $X(t)$ может быть векторной, скалярной величиной, функцией и т.д.

Чтобы задать случайный процесс $X(t)$, обычно для каждого натурального значения n и любых возможных значений $t_1, ..., t_n$ параметра t задают n-мерную функцию распределения вектора $X(t_1), ..., X(t_n)$ (это конечномерные распределения):

$$F(x_1, ..., x_n; t_1, ..., t_n) = P\{X(t_1) < x_1, ..., X(t_n) < x_n\}.$$

При этом на $F(x_1, ..., x_n; t_1, ..., t_n)$ накладываются два условия:

а) условие симметрии (инвариантности): для любой перестановки индексов

$$\begin{matrix} 1 & ... & n \\ i_1 & ... & i_n \end{matrix}$$

выполняется равенство

$$F(x_1, ..., x_n; t_1, ..., t_n) = F(x_{i_1}, ..., x_{i_n}; t_{i_1}, ..., t_{i_n});$$

б) условие согласованности: при $m < n$ для любых $t_{m+1}, t_{m+2}, ..., t_n$ выполняется равенство

$$F(x_1, x_2, ..., x_m, \infty, ..., \infty; t_1, t_2, ..., t_m, t_{m+1}, t_{m+2}, ..., t_n) =$$
$$= F(x_1, ..., x_m; t_1, ..., t_m).$$

Основные признаки, по которым различаются случайные процессы, касаются природы пространства состояний S; временного параметра T; отношения зависимости между временем наблюдения системы и ее состоянием, а именно:

а) S — пространство возможных состояний, т.е. пространство всех возможных значений случайного процесса $X(t) = x_t$. Если $S =$ $= 0, 1, 2, ...$, то это целочисленный процесс, если $S \in (-\infty, \infty)$, то это действительный процесс, если S — евклидово k-мерное пространство, то это k-мерный процесс;

б) если $T = 0, 1, 2, ...$, то $\{x_t\}$ — случайный процесс с дискретным временем; если $T \in (0, \infty)$, то $\{x_t\}$ — случайный процесс с непрерывным временем;

в) отношение зависимости определяется семейством конечномерных распределений $\{F(x_1, ..., x_n; t_1, ..., t_n)\} = \{F_n\}$.

Случайный процесс полностью задан, если определены его S, T и $\{F_n\}$.

4.2. Определения цепи Маркова

Определение 4.1. Процесс (протекающий в физической системе) называется ***марковским***, если для каждого момента времени вероятность любого состояния системы в будущем зависит от состояния системы в настоящий момент и не зависит от того, каким образом система пришла в это состояние. Цепь Маркова — это целочисленный марковский процесс с дискретным временем.

Подробнее: пусть G — некоторый эксперимент, имеющий конечное или счетное множество исходов $\{E_1, E_2, ..., E_n, ...\}$.

Будем повторять эксперимент. Номер исхода n-го эксперимента обозначим x_n. Если в n-м опыте реализовалось событие E_{i_n}, то будем считать, что $x_n = i_n$ (в n-м событии E_{i_n}).

Приведем два определения цепи Маркова.

Определение 4.2 (первое определение цепи Маркова). Последовательность случайных величин $\{X_n\}$ образует ***цепь Маркова***, если

$$\begin{gathered} P\{x_n = j_n / x_1 = i_1, x_2 = i_2, ..., x_{n-1} = i_{n-1}\} = \\ = P\{x_n = j_n = j / x_{n-1} = i_{n-1} = i\} = P_{ij}(n), \end{gathered} \tag{4.1}$$

где $P_{ij}(n)$ — вероятность перехода на n-м шаге из состояния $E_{i_{n-1}} = E_j$ в состояние $E_{i_n} = E_j$.

Событие, стоящее в условии (4.1), будем называть ***предысторией*** события в момент n.

Определение 4.3 (второе определение цепи Маркова). Последовательность случайных величин $\{x_n\}$ образует ***цепь Маркова***, если для любого набора целых чисел $n_1 < n_2 < ... < n_r < n$ верно, что

$$\begin{gathered} P\{x_n = j / x_{n_1} = i_1, x_{n_2} = i_2, ..., x_{n_{r-1}} = i_{r-1}, x_{n_r} = i\} = \\ = P\{x_n = j / x_{n_r} = i\}. \end{gathered} \tag{4.2}$$

Докажем эквивалентность приведенных здесь двух определений цепи Маркова.

Доказательство. Очевидно, что из второго определения следует первое, если в качестве набора чисел $n_1, n_2, ..., n_r$ взять числа 1, 2, 3, ..., n.

Чтобы доказать, что из первого определения следует второе, воспользуемся методом математической индукции. Для этого проверим выполнение условия (4.2), когда выполнено условие (4.1), сначала при любом конкретном значении $n_r = n - i$. Очевидно, что при $i = 1$ из условия (4.1) следует условие (4.2), так как если к предыстории правой части условия (4.1) добавим информацию, от которой событие не зависит, в виде события $\{x_{n-2} = i_{n-2}, ..., x_1 = i_1\}$, равенство останется верным.

Теперь предполагаем, что условие (4.2) следует из условия (4.1) для $n_r = n - k$ при $k = 1, ..., n - 1$, т.е. верно, что

$$P\{x_n = i_n = j / x_{n_1} = i_1, ..., x_{n-k} = i_{n-k} = i\} = P\{x_n = j / x_{n-1} = i\}.$$

Тогда докажем, что справедливо аналогичное равенство при $i = k + 1$, т.е. при $n_r = n - k - 1$:

$$\begin{gathered} P\{x_n = i_n = j / x_{n_1} = i_1, ..., x_{n-k-1} = i_{n-k-1} = i\} = \\ = \frac{P\{x_n = j, x_{n-k-1} = i_1, x_{n_{r-1}} = i_{r-1}, ..., x_{n_1} = i_1\}}{P\{x_{n-k-1} = i, x_{n_{r-1}} = i_{r-1}, ..., x_{n_1} = i_1\}} = \end{gathered}$$

$$= \frac{\sum_l P\{x_n = j, x_{n-k} = l, x_{n-k-1} = i,\ \dots, x_{n_1} = i_1\}}{P\{x_{n-k-1} = i, x_{n_{r-1}} = i_{r-1},\ \dots, x_{n_1} = i_1\}} =$$

$$= \frac{\sum_l P\{x_n = j / x_{n-k} = l, x_{n-k-1} = i\} P\{x_{n-k} = l, x_{n-k-1} = i, \dots, x_{n_1} = i_1\}}{P\{x_{n-k-1} = i, x_{n_{r-1}},\ \dots, x_{n_1} = i_1\}} =$$

$$= \sum_l P\{x_n = j / x_{n-k} = l, x_{n-k-1} = i\} P\{x_{n-k} = l / x_{n-k-1} = i\} =$$

$$= \sum_l \frac{P\{x_n = j, x_{n-k} = l, x_{n-k-1} = i\} P\{x_{n-k} = l, x_{n-k-1} = i\}}{P\{x_{n-k} = l, x_{n-k-1} = i\} P\{x_{n-k-1} = i\}} =$$

$$= \sum_l \frac{P\{x_n = j, x_{n-k} = l, x_{n-k-1} = i\}}{P\{x_{n-k-1} = i\}} =$$

$$= \frac{1}{P\{x_{n-k-1} = i\}} \sum_l P\{x_n = j, x_{n-k} = l,\ x_{n-k-1} = i\} =$$

$$= \frac{P\{x_n = j, x_{n-k-1} = i\}}{P\{x_{n-k-1} = i\}} = P\{x_n = j / x_{n-k-1} = i\}, \text{ ч.т.д.}$$

4.3. Свойства траекторий цепей Маркова

Под траекторией понимается последовательность состояний цепи Маркова. Рассмотрим свойства траекторий.

1. Фиксированная траектория в будущем при полной информации о прошлом.

Докажем, что

$$P\{x_{n+k} = i_{n+k}, \dots, x_n = i_n / x_{n-1} = i_{n-1}, \dots, x_1 = i_1\} =$$
$$= P\{x_{n+k} = i_{n+k}, \dots, x_n = i_n / x_{n-1} = i_{n-1}\}.$$

Доказательство

$$P\{x_{n+k} = i_{n+k}, \dots, x_n = i_n / x_{n-1} = i_{n-1},\ \dots, x_1 = i_1\} =$$

$$= \frac{P\{x_{n+k} = i_{n+k}, \dots, x_{1=i_1}\}}{P\{x = i_{n-1}, \dots, x_{1=i_1}\}} = \frac{P\{x_{n+k} = i_{n+k}, \dots, x_1 = i_1\}}{P\{x_{n+k-1} = i_{n+k-1}, \dots, x_1 = i_1\}} \times$$

$$\times \frac{P\{x_{n+k-1} = i_{n+k-1}, \dots, x_1 = i_1\}}{P\{x_{n+k-2} = i_{n+k-2}, \dots, x_1 = i_1\}} \dots \frac{P\{x_n = i_{n+k}, \dots, x_1 = i_1\}}{P\{x_{n-1} = i_{n-1}, \dots, x_1 = i_1\}} =$$

$$= P\{x_{n+k} = i_{n+k} / x_{n+k-1} = i_{n+k-1}, \dots, x_{n-1} = i_{n-1}\} \times$$

$$\times P\{x_{n+k-1} = i_{n+k-1} / x_{n+k-2} = i_{n+k-2}, \dots, x_{n-1} = i_{n-1}\} \cdots P\{x_n\} =$$

$$= \{i_n / x_{n-1} = i_{n-1}\} = \frac{P\{x_{n+k} = i_{n+k}, \dots, x_{n-1} = i_{n-1}\}}{P\{x_{n+k-1} = i_{n+k-1}, \dots, x_{n-1} = i_{n-1}\}} =$$

$$= \frac{P\{x_{n+k-1} = i_{n+k-1}, \dots, x_{n-1} = i_{n-1}\}}{P\{x_{n+k-2} = i_{n+k-2}, \dots, x_{n-1} = i_{n-1}\}} \cdots \frac{P\{x_n = i_n, x_{n-1} = i_{n-1}\}}{P\{x_{n-1} = i_{n-1}\}} =$$

$$= \frac{P\{x_{n+k} = i_{n+k}, \dots, x_{n-1} = i_{n-1}\}}{P\{x_{n-1} = i_{n-1}\}} = P\{x_{n+k} = i_{n+k}, \dots, x_n = i_n / x_{n-1} = i_{n-1}\}, \text{ ч.т.д.}$$

2. Фиксированная траектория в будущем при неполной информации о прошлом.

Докажем, что при $1 \leqslant n_1 < n_2 < ... < n_r \leqslant n$, $r \leqslant n$

$$P\{x_{n+k} = i_{n+k}, ..., x_n = i_n / x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_2}\} =$$
$$= P\{x_{n+k} = i_{n+k}, ..., x_n = i_n / x_{n_r} = i_{n_r}\}.$$

Доказательство

$$P\{x_{n+k} = i_{n+k}, ..., x_n = i_n / x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_2}\} =$$
$$= \frac{P\{x_{n+k} = i_{n+k}, ..., x_n = i_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}}{P\{x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}} =$$
$$= \frac{P\{x_{n+k} = i_{n+k}, ..., x_n = i_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}}{P\{x_{n+k-1} = i_{n+k-1}, ..., x_n = i_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}} \times$$
$$\times \frac{P\{x_{n+k-1} = i_{n+k-1}, ..., x_n = i_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}}{P\{x_{n+k-2} = i_{n+k-2}, ..., x_n = i_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}} \cdots$$
$$\cdots \frac{P\{x_n = i_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}}{P\{x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}} =$$
$$= P\{x_{n+k} = i_{n+k} / x_{n+k-1} = i_{n+k-1}\} \times$$
$$\times P\{x_{n+k-1} = i_{n+k-1} / x_{n+k-2} = i_{n+k-2}\} \cdots P\{x_n = i_n / x_{n_r} = i_{n_r}\} =$$
$$= P\{x_{n+k} = i_{n+k} / x_{n+k-1} = i_{n+k-1}, ..., x_n = i_n, x_{n-1} = i_{n-1}, ..., x_{n_r}\} =$$
$$= P\{i_{n_r}\} P\{x_{n+k-1} = i_{n+k-1} / x_{n+k-2} = i_{n+k-2}, ..., x_n = i_n, x_{n-1}\} =$$
$$= P\{i_{n-1}, ..., x_{n_r} = i_{n_r}\} ... P\{x_n = i_n / x_{n_r} = i_{n_r}\} =$$
$$= \frac{P\{x_{n+k} = i_{n+k}, ..., x_n = i_n, x_{n_r} = i_{n_r}\}}{P\{x_{n+k-1} = i_{n+k-1}, ..., x_n = i_n, x_{n_r} = i_{n_r}\}} \times$$
$$\times \frac{P\{x_{n+k-1} = i_{n+k-1}, ..., x_n = i_n, x_{n_r} = i_{n_r}\}}{P\{x_{n+k-2} = i_{n+k-2}, ..., x_n = i_n, x_{n_r} = i_{n_r}\}} \cdots$$
$$\cdots \frac{P\{x_n = i_n, x_{n_r} = i_{n_r}\}}{P\{x_{n_r} = i_{n_r}\}} = \frac{P\{x_{n+k} = i_{n+k}, ..., x_n = i_n, x_{n_r} = i_{n_r}\}}{P\{x_{n_r} = i_{n_r}\}} =$$
$$= P\{x_{n+k} = i_{n+k}, ..., x_n = i_n / x_{n_r} = i_{n_r}\}, \text{ ч.т.д.}$$

3. Любая траектория в будущем при полной информации о прошлом.

Докажем, что

$$P\{x_{n+k} \in A_k, ..., x_n \in A_0 / x_{n-1} = i_{n-1}, ..., x_1 = i_1\} =$$
$$= P\{x_{n+k} \in A_k, ..., x_n \in A_0 / x_{n-1} = i_{n-1}\}.$$

Доказательство

$$P\{x_{n+k} \in A_k, ..., x_n \in A_0 \,/\, x_{n-1} = i_{n-1}, ..., x_1 = i_1\} =$$
$$= \frac{P\{x_{n+k} \in A_k, ..., x_n \in A_0, x_{n-1} = i_{n-1}, ..., x_1 = i_1\}}{P\{x_{n-1} = i_{n-1}, ..., x_1 = i_1\}} =$$
$$= \frac{\sum\limits_{j_s \in A_s} P\{x_{n+k} = j_k, \ldots, x_n = j_0, x_{n-1} = i_{n-1}, ..., x_1 = i_1\}}{P\{x_{n-1} = i_{n-1}, ..., x_1 = i_1\}} =$$
$$= \sum_{j_s \in A_s} P\{x_{n+k} = j_k, ..., x_n = j_n \,/\, x_{n-1} = i_{n-1}, ..., x_1 = i_1\} =$$
$$= \sum_{j_s \in A_s} P\{x_{n+k} = j_{n+k}, ..., x_n = j_n \,/\, x_{n-1} = i_{n-1}\} =$$
$$= \sum_{j_s \in A_s} \frac{P\{x_{n+k} = j_{n+k}, ..., x_n = j_0, x_{n-1} = i_{n-1}\}}{x_{n-1} = i_{n-1}} =$$
$$= P\{x_{n+k} \in A_k, \ldots, x_n \in A_0 \,/\, x_{n-1} = i_{n-1}\}, \text{ ч.т.д.}$$

4. Любая траектория в будущем при неполной информации о прошлом.

Докажем, что при $1 \leqslant n_1 < n_2 < ... < n_r \leqslant n$, $r \leqslant n$

$$P\{x_{n+k} \in A_k, ..., x_n \in A_0 / x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\} =$$
$$= P\{x_{n+k} \in A_k, ..., x_n \in A_0 / x_{n_r} = i_{n_r}\}.$$

Доказательство

$$P\{x_{n+k} \in A_k, ..., x_n \in A_0 \,/\, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\} =$$
$$= \frac{P\{x_{n+k} \in A_k, ..., x_n \in A_n, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}}{P\{x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}} =$$
$$= \frac{\sum\limits_{j_s \in A_s} P\{x_{n+k} = j_k, ..., x_n = j_0, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}}{P\{x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\}} =$$
$$= \sum_{j_s \in A_s} P\{x_{n+k} = j_k, ..., x_n = j_n \,/\, x_{n_r} = i_{n_r}, ..., x_{n_1} = i_{n_1}\} =$$
$$= \sum_{j_s \in A_s} P\{x_{n+k} = j_n, ..., x_n = j_0 \,/\, x_{n_r} = i_{n_r}\} =$$
$$= P\{x_{n+k} \in A_k, ..., x_n \in A_0 \,/\, x_{n_r} = i_{n_r}\}, \text{ ч.т.д.}$$

5. Любая траектория в будущем при любой траектории в прошлом при фиксированном настоящем.

Докажем, что при $A = \{x_{n+k} \in A_k, ..., x_n \in A_0\}$; $A = \{x_{n-2} \in B_{n-2}, ..., x_n \in B_1\}$ и $x_{n-1} = i_{n-1}$ верно равенство

$$P\{A / x_{n-1} = i_{n-1}, B\} = P\{A / x_{n-1} = i_{n-1}\}.$$

Доказательство

$$P\{A / x_{n-1} = i_{n-1}, B\} = \frac{P\{A, x_{n-1} = i_{n-1}, B\}}{P\{x_{n-1} = i_{n-1}, B\}} =$$

$$= \frac{\sum_{j_s \in B} P\{A, x_{n-1} = i_{n-1}, x_{n-2} = j_{n-2}, ..., x_1 = j_1\}}{P\{x_{n-1} = i_{n-1}, B\}} =$$

$$= \frac{\sum_{j_s \in B} P\{A / x_{n-1} = i_{n-1}\} P\{x_{n-1} = i_{n-1}, x_{n-2} = j_{n-2}, ..., x_1 = j_1\}}{P\{x_{n-1} = i_{n-1}, B\}} =$$

$$= \frac{P\{A / x_{n-1} = i_{n-1}\} P\{x_{n-1} = i_{n-1}, B\}}{P\{x_{n-1} = i_{n-1}, B\}} = P\{A / x_{n-1} = i_{n-1}\}, \text{ ч.т.д.}$$

6. Независимость любых будущего и прошлого при фиксированном настоящем.

Докажем, что любые события $A = \{x_{n+k} = i_{n+k}, ..., x_n = i_n\}$ и $B = \{x_{n-2} = i_{n-2}, ..., x_1 = i_1\}$ при фиксированном настоящем независимы, т.е.

$$P\{A, B / x_{n-1} = i_{n-1}\} = P\{A / x_{n-1} = i_{n-1}\} P\{B / x_{n-1} = i_{n-1}\}.$$

Доказательство

$$P\{A, B / x_{n-1} = i_{n-1}\} = \frac{P\{A, B, x_{n-1} = i_{n-1}\}}{P\{x_{n-1} = i_{n-1}\}} =$$

$$= \frac{P\{A, B, x_{n-1} = i_{n-1}\}}{P\{B, x_{n-1} = i_{n-1}\}} \frac{P\{B, x_{n-1} = i_{n-1}\}}{P\{x_{n-1} = i_{n-1}\}} =$$

$$= P\{A / x_{n-1} = i_{n-1}\} P\{B / x_{n-1} = i_{n-1}\}, \text{ ч.т.д.}$$

4.4. Матрица переходных вероятностей

Определение 4.4. Если вероятность $p_{ij}(n)$ — вероятность перехода из состояния E_i в состояние E_j на n-м шаге — не зависит от n, то цепь Маркова называется ***однородной***.

При $n = 1$ обозначим $p_{ij}(n) = p_{ij}$, а матрицу (p_{ij}) — матрицу одношаговых переходных вероятностей — через P, т.е. $P = (p_{ij})$, где $p_{ij} = P\{x_n = j / x_{n-1} = i\}$.

Свойства матрицы P:

1) $p_{ij} > 0$;

2) $\sum_j p_{ij} = 1$ для любого i.

Определение 4.5. Если для матрицы P выполняются условия 1) и 2), то матрица называется ***стохастической***. Если матрица стохастическая и $\sum_i p_{ij} = 1$ для любого j, то она называется ***дважды стохастической***.

Цепь Маркова *задана*, если заданы матрица ее переходных вероятностей P за один шаг (одношаговая) и вектор начальных вероятностей.

Матрица переходных вероятностей цепи Маркова за n шагов

Утверждение 4.1. Матрица переходных вероятностей за n шагов есть n-я степень матрицы переходных вероятностей за один шаг.

Доказательство. Это следует из равенства

$$p_{ij}(n) = \sum_k p_{ik}p_{kj}(n-1), \tag{4.3}$$

где $p_{ij}(m)$ — вероятность перехода из i-го состояния в j-е за m шагов. Действительно, положив в формуле (4.3) $n = 2$, получим $p_{ij}(2) = \sum\limits_k p_{ik}p_{kj}$, что соответствует матричному равенству $P(2) = P^2$, где $P(n)$ — матрица переходных вероятностей за n шагов

Тогда, предполагая по индукции, что утверждение верно для n шагов, и положив $r = n$, докажем аналогичное утверждение для $(n + 1)$-го шага:

$$P(n+1) = P(n)P = P^nP = P^{n+1},$$

что и утверждалось. Остается доказать равенство (4.3):

$$p_{ij}(n) = P\{x_n = j / x_0 = i\} = \frac{P\{x_n = j, x_0 = i\}}{P\{x_0 = i\}} = \frac{\sum\limits_k P\{x_n = j, x_1 = k, x_0 = i\}}{P\{x_0 = i\}} =$$

$$= \frac{\sum\limits_k P\{x_n = j / x_i = k, x_0 = i\}P\{x_1 = k, x_0 = i\}}{P\{x_0 = i\}} =$$

$$= \sum_k P\{x_n = j / x_i = k\}P\{x_1 = k / x_0 = i\} = \sum_k p_{ik}(n-r)p_{ik}(r),$$

что и доказывает равенство (4.3).

Замечание 4.1. Матрица переходных вероятностей для цепи Маркова за n шагов может быть получена непосредственно по вероятностному графу (графическому изображению стрелками переходов в цепи Маркова из любого состояния в любое на каждом шаге с указанием вероятностей этих переходов на стрелках). Ниже это будет показано на конкретных примерах (иногда такой путь для нахождения $P(n)$ является практически единственным).

4.5. Примеры цепей Маркова

Пример 4.1 (блуждание с поглощающими экранами). Рассматриваем блуждание по целочисленным точкам $[0; a]$ по схеме, изображенной на рис. 4.1, *а* (здесь $p + q = 1$, стрелками указаны возможные переходы из состояния в состояние за один шаг).

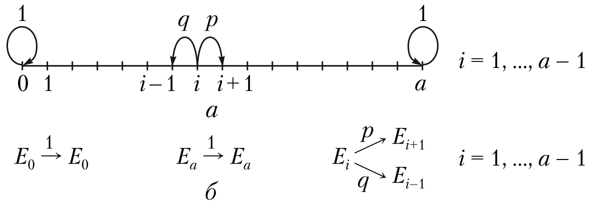


Рис. 4.1. **Блуждание с поглощающими экранами:**

а — схема; *б* — вероятностный граф переходов за один шаг

Состояние системы определяется положением точки на отрезке $[0; a]$. Это цепь Маркова, так как состояние E в момент времени n определяется только состоянием системы E в момент $(n - 1)$; $0 < i, j < a$.

Цепь Маркова однородная, так как по условию

$$P\{x_m = j / x_{m-1} = i\} = P\{x_n = j / x_{n-1} = i\}.$$

Матрица переходных вероятностей, составленная по рис. 4.1, *б*, имеет вид

$$P = (P_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ q & 0 & p & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & q & 0 & p & 0 & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Для полного задания цепи Маркова выберем начальное распределение, т.е. вектор начальных вероятностей, например, в одном из видов:

а) $(P_0^{(0)}, ..., P_a^{(0)}) = (0, ..., 0, 1, 0, ..., 0)$ — блуждание начинается с i-й точки;

б) $P_i^{(0)} = \dfrac{1}{a+1}$, $i = 0, 1, ..., a$ — блуждание начинается с произвольной точки.

Приведем аналитическое представление данного марковского процесса.

Пусть x_n — значение процесса в n-й момент: y_n — результат n-го опыта, тогда если определить

$$y_n = \begin{cases} 1, \text{при шаге вправо}, \\ 0, \text{при шаге влево}, \end{cases}$$

то $x_n = y_{n-1} + y_n$, где x_n — положение точки на отрезке $[0; a]$.

Формулу для x_n можно рассматривать как иллюстрацию марковости, отражающую зависимость состояния процесса от его состояния в предшествующий момент и результата последнего опыта.

Найдем теперь матрицу переходных вероятностей за два шага — $P(2)$.

Здесь это проще сделать непосредственно по вероятностным графам из любого состояния за два шага (рис. 4.2).

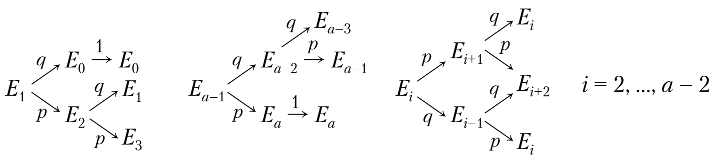


Рис. 4.2. **Вероятностный граф переходов за два шага**

Получаем

$$P(2) = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & 0 & 0 \\ q & pq & 0 & p^2 & \dots & \dots & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & \dots & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \dots & q^2 & 0 & pq & p \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Пример 4.2 (блуждание с отражающими экранами). Рассматриваем блуждание по целочисленным точкам $[0; a]$ по схеме, изображенной на рис. 4.3, *а* ($p + q = 1$).

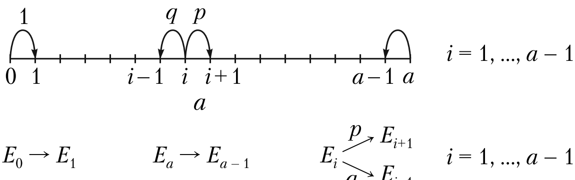


Рис. 4.3. **Блуждание с отражающими экранами:**

а — схема; *б* — вероятностный граф переходов за один шаг

Состояние системы определяется положением точки на отрезке $[0; a]$.

Это цепь Маркова, так как состояние E_i в момент времени n определяется только состоянием системы E_i в момент $(n - 1)$; $0 < i, j < a$.

Цепь Маркова однородная, так как по условию

$$P\{x_m = j / x_{m-1} = i\} = P\{x_n = j / x_{n-1} = i\}.$$

Матрица переходных вероятностей, составленная по рис. 4.3, *б*, имеет вид

$$P = (P_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ q & 0 & p & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & q & 0 & p & 0 & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Для полного задания цепи Маркова выберем начальное распределение, т.е. вектор начальных вероятностей, например, в одном из видов:

а) $(P_0^{(0)}, ..., P_a^{(0)}) = (0, ..., 0, 1, 0, ..., 0)$ — блуждание начинается с i-й точки;

б) $P_i^{(0)} = \dfrac{1}{a+1}$, $i = 0, 1, ..., a$ — блуждание начинается с произвольной точки.

Приведем аналитическое представление данного марковского процесса.

Пусть x_n — значение процесса в n-й момент: y_n — результат n-го опыта, тогда если определить

$$y_n = \begin{cases} 1, \text{ при шаге вправо,} \\ 0, \text{ при шаге влево,} \end{cases}$$

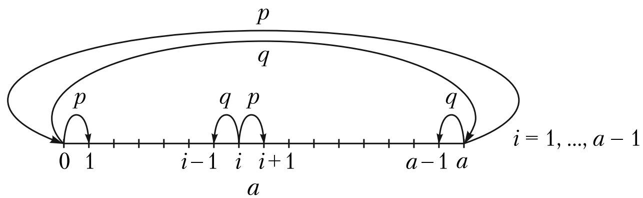
то

$$x_n = \begin{cases} x_{n-1} + y, x_n = 1, 2, ..., a-1, \\ x_{n-1} + 1, x_n = 0, \\ x_{n-1} - 1, x_n = a. \end{cases}$$

Задание 4.1. С помощью вероятностных графов найдите $P(2)$ в примере 4.2 аналогично примеру 4.1.

Задание 4.2 (циклическое блуждание). Рассматривается блуждание по целочисленным точкам $[0; a]$ по схеме, изображенной на рис. 4.4, *а* ($p + q = 1$), с вероятностным графом, изображенным на рис. 4.4, *б*.

Состояние системы определяется положением точки на отрезке $[0; a]$. Проведите анализ процесса.



$$E_0 \begin{matrix} \xrightarrow{p} E_1 \\ \xrightarrow[q]{} E_a \end{matrix} \qquad E_a \begin{matrix} \xrightarrow{p} E_1 \\ \xrightarrow[q]{} E_{a-1} \end{matrix} \qquad E_i \begin{matrix} \xrightarrow{p} E_{i+1} \\ \xrightarrow[q]{} E_{i-1} \end{matrix} \qquad i = 1, ..., a-1$$

б

Рис. 4.4. **Циклическое блуждание:**

а — схема; *б* — вероятностный граф переходов за один шаг

Задание 4.3 (обобщенное блуждание). Рассматривается блуждание по целочисленным точкам $[0; a]$ так, что в общем случае возможны переходы из каждого состояния в каждое, т.е. заданы $q_0, q_1, ..., q_{n-1}$ — вероятности перейти соответственно на i шагов (единиц) вправо по циклу, где $i = 0, 1, ..., a-1$ и $\sum\limits_{i=1}^{a-1} q_i = 1$.

Состояние системы определяется положением точки на отрезке $[0; a]$. Проведите анализ процесса.

Пример 4.3 (цепь Маркова, связанная с испытаниями Бернулли). Пусть состояние системы определяется результатами двух последних испытаний Бернулли, т.е. УУ; УН; НУ; НН, где У и Н означают соответственно успех и неуспех испытания Бернулли. Проиллюстрируем марковость и напишем матрицы переходных вероятностей P и $P(2)$ (из вероятностных соображений и вероятностного графа состояний с проверкой путем возведения матрицы P в квадрат).

Выписываем матрицу P:

$$P = \begin{pmatrix} p & q & 0 & 0 \\ 0 & 0 & p & q \\ p & q & 0 & 0 \\ 0 & 0 & p & q \end{pmatrix}.$$

Матрица получена по вероятностным графам (рис. 4.5).

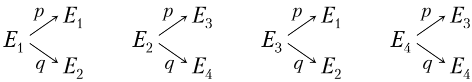


Рис. 4.5. **Вероятностные графы для матрицы переходных вероятностей *P***

Иллюстрация марковости (пусть y_n — результат последнего опыта):

$$y_n = \begin{cases} 0, \text{ если в последнем опыте Н}, \\ 1, \text{ если в последнем опыте У}. \end{cases}$$

$x_n = 2|x_{n-1} + 1|_2 + 1 + y_n$, где $|B|_a$ (здесь и в дальнейшем) есть остаток при делении числа B на число а.

Найдем матрицу переходных вероятностей $P(2)$ двумя способами — возведением в квадрат матрицы P и по вероятностным графам (рис. 4.6).

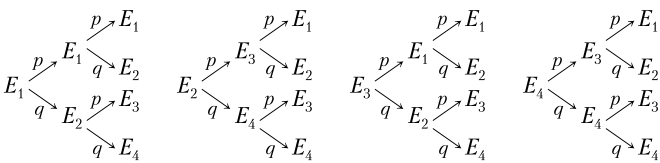


Рис. 4.6. **Вероятностные графы для матрицы переходных вероятностей *P*(2)**

По вероятностным графам получаем

$$P(2) = \begin{pmatrix} p^2 & pq & pq & q^2 \\ p^2 & pq & pq & q^2 \\ p^2 & pq & pq & q^2 \\ p^2 & pq & pq & q^2 \end{pmatrix}.$$

Легко получить тот же результат возведением матрицы P в квадрат.

Пример 4.4 (движение точки по оси *Ox* со скоростью ±1). Состояние системы определяется скоростью точки в данный момент времени (состояния E_1 и E_2 означают, что точка имеет скорость соответственно $+1$ и -1). Проиллюстрируем марковость и выпишем матрицы вероятностей P и $P(2)$ (рис. 4.7).

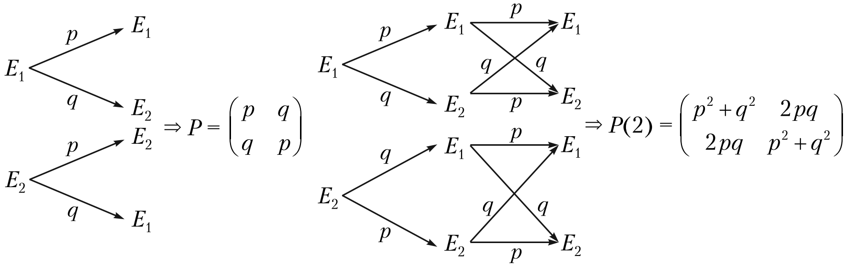


Рис. 4.7. **Вероятностные графы для движения точки по оси**

Определим

$$y_n = \begin{cases} 0, \text{ направление сохраняется на } n\text{-м шаге}, \\ 1, \text{ в противном случае.} \end{cases}$$

Тогда если x_n — индекс состояния на n-м шаге, то $x_{n+1} = x_n + (-1)^{x_n+1} y_n$.

Пример 4.5. Случайное размещение частиц по α ячейкам. Пусть сначала все ячейки пусты, это значит, что начальное распределение (или вектор начальных вероятностей) имеет вид

$$(P_0^{(0)}, ..., P_a^{(0)}) = (1, 0, ..., 0).$$

Состояние системы определяется числом занятых (непустых) ячеек. Выпишем матрицу одношаговых переходных вероятностей с помощью вероятностных графов (рис. 4.8).

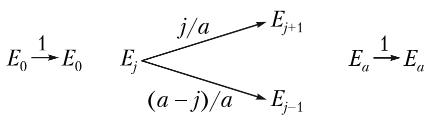


Рис. 4.8. **Вероятностные графы для размещения частиц по ячейкам**

Тогда очевидно, что при j занятых ячейках вероятность перехода в то же состояние E_j есть $p_{ij} = j/a$, а перехода в состояние E_{j+1} — $p_{i,j+1} = (a-j)/a$. Получаем

$$P = (p_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1/a & 1-1/a & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2/a & 1-2/a & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Определим

$$y_n = \begin{cases} 0, \text{ если частица попадает в непустую ячейку}, \\ 1, \text{ если частица попадает в пустую ячейку.} \end{cases}$$

Тогда $x_{n+1} = x_n + y_n$, $n = 0, 1, ..., a$. Это пример неоднородной цепи Маркова.

Пример 4.6 (цепь Маркова со счетным множеством состояний). Пусть в примере 4.4 состояние системы определяется положением точки на оси Ox. Тогда матрица переходных вероятностей за один шаг P имеет вид

$$P = \begin{pmatrix} \dots & -2 & -1 & 0 & 1 & 2 & 3 & \dots \\ \dots & 0 & q & 0 & p & 0 & 0 & \dots \\ \dots & 0 & 0 & q & 0 & p & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Абсциссы точек, соответствующих состоянию системы: ..., E_{-2}, E_{-1}, E_0, E_1, E_2, E_3, ...

Чтобы записать матрицу переходных вероятностей в привычной форме, состояния системы расположим в порядке E_0, E_{-1}, E_1, E_{-2}, E_2, ...

Тогда матрица P запишется в виде

$$P = \begin{pmatrix} 0 & q & p & 0 & 0 & 0 & \dots \\ p & 0 & 0 & q & 0 & 0 & \dots \\ q & 0 & 0 & 0 & p & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

4.6. Свойства матрицы $P(2)$

Сначала сформулируем пару вопросов и ответим на них.

Вопрос 1. Всякая ли стохастическая матрица может быть матрицей вероятностей перехода за два шага некоторой цепи Маркова?

Ответ. Нет, например $P(2)$ не может иметь вид $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, так как из состояния 1 в состояние 2 нельзя перейти за два шага, не оставаясь за один шаг в том же состоянии, а это противоречит тому, что дано $p_{11}(2) = 0$ и $p_{22}(2) = 0$.

Вопрос 2. Мы выше установили, что цепь Маркова однозначно определяется начальным распределением и матрицей одношаговых переходных вероятностей P. Можно ли задавать цепь Маркова начальным распределением и матрицей переходных вероятностей за два шага $P(2)$?

Ответ. Нет, так как, например, если $P(2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, то ей соответствуют разные матрицы одношаговых вероятностей:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ и } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Утверждение 4.2. Стохастическая матрица второго порядка является матрицей вероятностей перехода за два шага некоторой цепи Маркова тогда и только тогда, когда сумма ее диагональных элементов больше либо равна единице.

Доказательство. Пусть

$$P=\begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix};\ A=P^2=\begin{pmatrix} c & 1-c \\ 1-d & d \end{pmatrix}.$$

Тогда

$$\begin{pmatrix} c & 1-c \\ 1-d & d \end{pmatrix}=\begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}\begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}=$$
$$=\begin{pmatrix} a^2+(1-a)(1-b) & (a+b)(1-a) \\ (a+b)(1-b) & b^2+(1-a)(1-b) \end{pmatrix}.$$

Необходимость. Пусть $A = P(2)$, тогда

$$c = a^2 + (1 - a)(1 - b),\ d = b^2 + (1 - a)(1 - b).$$

Отсюда следует:

$$c + d = a^2 + b^2 + 2 - 2a - 2b + 2ab = (a + b - 1)^2 + 1 \geqslant 1, \text{ ч.т.д.}$$

Достаточность. Пусть $c + d \geqslant 1$. Тогда докажем: при

$$\begin{cases} c = a^2 + (1 - a)(1 - b), \\ d = b^2 + (1 - a)(1 - b) \end{cases}$$

существуют корни этой системы уравнений a и b такие, что $a, b \in (0; 1)$, а это и значит существование $P=\begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}$, такой что $P^2=\begin{pmatrix} c & 1-c \\ 1-d & d \end{pmatrix}$. Остается решить данную систему. Проводим выкладки:

$$c + d = (a + b)^2 - 2(a + b) + 2 \Rightarrow c + d - 1 = (a + b - 1)^2 \Rightarrow$$
$$\Rightarrow a + b = \sqrt{c+d-1}+1.$$

Из первого уравнения системы получаем

$$c=a(a+b)-(a+b)+1=a(\sqrt{c+d-1}+1)-\sqrt{c+d-1}\Rightarrow a=\frac{c+\sqrt{c+d-1}}{\sqrt{c+d-1}+1},$$

откуда при $c + d \geqslant 1$ следует, что $0 < a < 1$. Аналогично из второго уравнения системы получаем $b=\dfrac{d+\sqrt{c+d-1}}{\sqrt{c+d-1}+1}$, т.е. при $c + d \geqslant 1$ $0 < b < 1$, ч.т.д.

4.7. Определение безусловных вероятностей состояний цепи Маркова

Пусть задана цепь Маркова, т.е. заданы вектор начальных вероятностей $(P_1^{(0)}, ..., P_m^{(0)})$ и матрица одношаговых переходных вероятностей P:

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \dots & \dots & \dots & \dots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix}.$$

Найдем безусловные вероятности всех состояний цепи на n-м шаге.

Обозначим через $P_i^{(k)}$ безусловную вероятность i-го состояния на k-м шаге. По формуле полной вероятности, считая гипотезами состояния системы на предыдущем $(k-1)$-м шаге, имеем

$$P_1^{(1)} = P_1^{(0)} P_{11} + P_2^{(0)} P_{21} + \ldots + P_m^{(0)} P_{m1},$$
$$P_2^{(1)} = P_1^{(0)} P_{12} + P_2^{(0)} P_{22} + \ldots + P_m^{(0)} P_{m2},$$
....
$$P_m^{(1)} = P_1^{(0)} P_{1m} + P_2^{(0)} P_{2m} + \ldots + P_m^{(0)} P_{mm}.$$

При записи в матричной форме это означает

$$(P_1^{(1)}, P_2^{(1)}, \ldots, P_m^{(1)}) = (P_1^{(0)}, P_2^{(0)}, \ldots, P_m^{(0)})P.$$

Аналогично получаем формулу для определения безусловных вероятностей на 2-м шаге:

$$(P_1^{(2)}, P_2^{(2)}, \ldots, P_m^{(2)}) = (P_1^{(1)}, P_2^{(1)}, \ldots, P_m^{(1)})P = (P_1^{(0)}, P_2^{(0)}, \ldots, P_m^{(0)})P^2.$$

Таким же образом для вектора безусловных вероятностей всех m состояний цепи Маркова на n-м шаге получаем формулу

$$(P_1^{(n)}, P_2^{(n)}, \ldots, P_m^{(n)}) = (P_1^{(0)}, P_2^{(0)}, \ldots, P_m^{(0)})P^n. \qquad (4.4)$$

Замечание 4.2. Пусть в начальный момент времени система находится в состоянии E_i, т.е. вектор начальных вероятностей имеет вид

$$\vec{P}^{(0)}(P_1^{(0)} = 0, \ldots, P_{i-1}^{(0)} = 0, P_i^{(0)} = 1, P_{i+1}^{(0)} = 0, \ldots, P_m^{(0)} = 0).$$

Тогда вектор безусловных вероятностей по формуле (4.4) совпадает с i-й строкой матрицы P переходных вероятностей цепи. Это дает возможность интерпретировать i-ю строку матрицы переходных вероятностей цепи Маркова за n шагов ($i = 1, \ldots, n$) как вектор безусловных вероятностей всех состояний системы, при условии что процесс начался с i-го состояния.

Задача 4.1. Даны матрица P одношаговых переходных вероятностей цепи Маркова и вектор начальных вероятностей:

$$P = \begin{pmatrix} 0,1 & 0,5 & 0,4 \\ 0,6 & 0,2 & 0,2 \\ 0,3 & 0,4 & 0,3 \end{pmatrix}; \ \vec{P}^{(0)} = (0,7; 0,2; 0,1).$$

Требуется:

а) найти вектор безусловных вероятностей в момент времени $t = 2$ (т.е. на втором шаге);

б) найти вероятность $P(X)$ того, что в моменты $t = 0; 1; 2; 3$ состояниями системы будут соответственно состояния с индексами 1; 3; 3; 2;

в) найти вероятность $P(X)$ того, что в моменты $t = 2; 3; 4; 5$ состояниями системы будут соответственно состояния с индексами 1; 1; 3; 2;

г) найти вероятность $P(X)$ того, что в моменты $t = 2; 4; 5$ состояниями системы будут соответственно состояния с индексами 2; 2; 3.

Решение

а) По формуле (4.4): $\vec{P}^{(2)} = \vec{P}^{(0)}P(2)$;

$$P(2) = P^2 = \begin{pmatrix} 0,43 & 0,31 & 0,26 \\ 0,24 & 0,42 & 0,34 \\ 0,36 & 0,35 & 0,29 \end{pmatrix};$$

$$\vec{P}^{(2)} = (0,7; 0,2; 0,1)\begin{pmatrix} 0,43 & 0,31 & 0,26 \\ 0,24 & 0,42 & 0,34 \\ 0,36 & 0,35 & 0,29 \end{pmatrix} = (0,385; 0,336; 0,279);$$

б) $P(X) = P_1^{(0)} p_{13} p_{33} p_{32} = 0,7 \cdot 0,4 \cdot 0,3 \cdot 0,4 = 0,0336$;

в) $P(X) = P_1^{(2)} p_{11} p_{13} p_{32} = 0,385 \cdot 0,1 \cdot 0,4 \cdot 0,4 = 0,00616$;

г) $P(X) = P_1^{(2)} p_{22}^{(2)} p_{23} = 0,385 \cdot 0,42 \cdot 0,2 = 0,03234$.

4.8. Проверка на марковость

Для проверки марковости (под марковостью понимается выполнение определений для цепи Маркова) нужно проверять соответствие с первым или вторым определением цепи Маркова (определения 4.2 и 4.3).

Для доказательства того, что изучаемый случайный процесс не является цепью Маркова, достаточно привести пример определенной последовательности состояний, при которой определение для цепи Маркова не выполняется.

Задача 4.2 (суммы случайных величин). Пусть $X_1, X_2, ..., X_n$ — взаимно независимые, одинаково распределенные СВ с распределением $P(X_i = k) = p$, $i = 1, ..., n$. Пусть $Y_n = X_1 + X_2 + ... + X_n$. Определить, образует ли последовательность СВ $\{Y_n\}$ цепь Маркова.

Решение

Если последовательность СВ $\{Y_n\}$ образует цепь Маркова, то для любого набора чисел $i_1, i_2, ..., i_n$ выполняется соотношение

$$P\{Y_n = i_n = j / Y_1 = i_1,\ Y_2 = i_2,\ ...,\ Y_{n-1} = i_{n-1} = i\} = P\{Y_n = j / Y_{n-1} = i\},$$

что соответствует первому определению цепи Маркова. Проверим это для $\{Y_n\}$. Будем преобразовывать отдельно левую и правую части проверяемого равенства. Предварительно события, выраженные через СВ Y_i, заменим на эквивалентные, записанные в терминах исходных СВ X_i, для того чтобы в дальнейшем воспользоваться их независимостью и данной информацией об их распределениях:

$$\{Y_1 = X_1 = i_1, Y_2 = X_1 + X_2 = i_2, ..., Y_n = X_1 + X_2 + ... + X_n = i_n\} =$$
$$= \{X_1 = i_1,\ X_2 = i_2 - i_1, ..., X_n = i_n - i_{n-1}\};$$
$$\{Y_n = i_n, Y_{n-1} = X_1 + X_2 + ... + X_{n-1} = i_{n-1}\} =$$
$$= \{X_n = i_n - i_{n-1}, ..., X_1 + X_2 + ... + X_{n-1} = i_{n-1}\}.$$

Воспользуемся теперь теоремой умножения вероятностей:

$$P\{Y_n = j / Y_1 = i_1,\ Y_2 = i_2, ..., Y_{n-1} = i\} =$$
$$= \frac{P\{Y = j / Y_1 = i_1,\ Y_2 = i_2, ..., Y_{n-1} = i\}}{P\{Y_1 = i_1, Y_2 = i_2, ..., Y_{n-1} = i\}} =$$
$$= \frac{P\{X_1 = i_1, X_2 = i_2 - i_1, ..., X_n = i_n - i_{n-1}\}}{P\{X_1 = i_1,\ X_2 = i_2 - i_1, ..., X_{n-1} = i_{n-1} - i_{n-2}\}} =$$
$$= \frac{P\{X_n = i_n - i_{n-1}\}P\{X_{n-1} = i_{n-1} - i_{n-2}\}\cdots P\{X_1 = i_1\}}{P\{X_{n-1} = i_{n-1} - i_{n-2}\}\cdots P\{X_1 = i_1\}} = P\{X_n = i - j\};$$
$$P\{Y_n = j / Y_{n-1} = i\} = \frac{P\{Y_n = j,\ Y_{n-1} = i\}}{P\{Y_{n-1} = i\}} =$$
$$= \frac{P\{X_n = j - 1\}P\{X_1 + X_2 + ... + X_{n-1} = i\}}{P\{X_1 + X_2 + ... + X_{n-1} = i\}} = P\{X_n = i_n - i_{n-1}\}.$$

Из совпадения результатов преобразований левой и правой частей следует, что последовательность СВ $\{Y_n\}$ является цепью Маркова.

Задача 4.3 (урновая схема Пойя). В урне находятся a красных и b черных шаров. Наугад извлеченный шар всегда возвращается в урну, причем в нее добавляется еще c шаров того же цвета, что и извлеченный. Пусть СВ

$$X = \begin{cases} 1, \text{если при } n\text{-м извлечении получили черный шар,} \\ 0, \text{в противном случае.} \end{cases}$$

Показать немарковский характер последовательности $\{X_n\}$.

Решение

Для установления немарковости $\{X_n\}$ достаточно привести пример конкретной последовательности значений элементов проверяемой последовательности, для которой определение цепи Маркова не выполняется.

В качестве такой числовой последовательности рассмотрим единичные значения в первые три последовательных момента, т.е. сравним вероятности событий $P\{X_3 = 1/X_2 = 1, X_1 = 1\}$ и $P\{X_3 = 1/X_2 = 1\}$ (обозначим результаты извлечений шаров разного цвета соответствующей последовательностью букв «Ч» и «К» в соответствующем порядке):

$$P\{X_3 = 1 / X_2 = 1\} = \frac{P\{X_3 = 1,\ X_2 = 1\}}{P\{X_2 = 1\}} = \frac{P(\text{ЧЧЧ}) + P(\text{ЧЧК})}{P(\text{ЧК}) + P(\text{ЧЧ})} =$$

$$= \frac{\frac{b}{b+a}\cdot\frac{b+c}{b+a+c}\cdot\frac{b+2c}{b+a+2c}+\frac{a}{b+a}\cdot\frac{b}{b+a+c}\cdot\frac{b+c}{b+a+2c}}{\frac{a}{b+a}\cdot\frac{b}{b+a+c}+\frac{b}{b+a}\cdot\frac{b+c}{b+a+c}} = \frac{b+c}{a+b+c};$$

$$P\{X_3 = 1/X_2 = 1, X_1 = 1\} = \frac{P(\text{ЧЧЧ})}{P(\text{ЧЧ})} =$$

$$= \frac{\frac{b}{b+a}\cdot\frac{b+c}{b+a+c}\cdot\frac{b+2c}{b+a+2c}}{\frac{b}{b+a}\cdot\frac{b+c}{b+a+c}} = \frac{b+2c}{b+a+2c}.$$

Из несовпадения этих результатов следует немарковость последовательности $\{X_n\}$.

Задача 4.4 (бернуллиевские испытания). Свяжем состояния системы с последними двумя исходами испытаний Бернулли, обозначив буквами «У» и «Н» успех и неуспех в соответствующем порядке в двух последних испытаниях, т.е. пусть система находится в состояниях E_1 и E_2, если соответственно два последних исхода есть УУ или любые другие. Определим СВ

$$X = \begin{cases} 1, \text{ если система после } n\text{-го шага находится в состоянии } E_1, \\ 0, \text{ если система после } n\text{-го шага находится в состоянии } E_2. \end{cases}$$

Показать, что последовательность СВ $\{X_n\}$ не является цепью Маркова.

Решение

Достаточно привести пример конкретных значений СВ X, для которых не выполняется определение цепи Маркова. Сравним вероятности двух событий: $P\{X_3 = 1 / X_2 = 2, X_1 = 1\}$ и $P\{X_3 = 1 / X_2 = 2, X_1 = 2\}$. Если окажется, что они не совпадают, то это и будет доказательством немарковости $\{X_n\}$, так как при одинаковых значениях в будущем (в момент 3 состояние E_3) и в настоящем (в момент 2 состояние E_2) они отличаются прошлым (в момент 1 в первом случае состояние E_1, а во втором — E_2). Очевидно, что первая вероятность равна нулю, так как указанная последовательность событий $E_1E_2E_1$ невозможна (при возможной последовательности событий E_1E_2), а вторая не равна нулю, так как указанная последовательность событий $E_2E_2E_1$ возможна.

Задача 4.5 (скользящие средние). Пусть последовательность СВ $\{Y_n\}$ — последовательность взаимно независимых СВ с распределением

Y_n	–1	1
P	0,5	0,5

Пусть $X_n = Y_n + Y_{n+1}$. Является ли последовательность $\{X_n\}$ цепью Маркова?

Решение

Вычислим, например, совпадают ли вероятности $P\{X_3 = 1 / X_2 = 0, X_1 = 1\}$ и $P\{X_3 = 1 / X_2 = 0, X_1 = -1\}$. Если вероятности не совпадут, то это значит, что $\{X_n\}$ не является цепью Маркова. Получаем $P\{X_3 = 1 / X_2 = 0, X_1 = 1\} = 0$, так как это вероятность невозможного события ни при каком наборе значений СВ Y_1, Y_2, Y_3, Y_4, что легко проверяется непосредственно. В то же время $P\{X_3 = 1 / X_2 = 0, X_1 = -1\} = P\{Y_3 = Y_4 = 1 / Y_1 = -1, Y_2 = -1, Y_3 = 1\} = P\{Y_1 = -1, Y_2 = -1, Y_3 = 1, Y_4 = 1\}/P\{Y_1 = -1, Y_2 = -1, Y_3 = 1\} = 1/2$.

Таким образом, доказано, что $\{X_n\}$ не является цепью Маркова.

Задача 4.6. Пусть $X_1, X_2, ...$ — последовательность бернуллиевских СВ с вероятностью успеха, равной p ($q = 1 - p$). $Y_n = X_nX_{n+1}$. Является ли последовательность $\{Y_n\}$ цепью Маркова?

Решение

Сравним, например, следующие условные вероятности:

$$P\{Y_n = 1 / Y_{n-1} = 1, Y_{n-2} = 1\} \text{ и } P\{Y_n = 1 / Y_{n-1} = 1\}.$$

Если эти вероятности не совпадут, то это значит, что $\{Y_n\}$ не является цепью Маркова. Имеем

$$P\{Y_n = 1 / Y_{n-1} = 1, Y_{n-2} = 1\} = \frac{P\{X_nX_{n+1} = 1, X_nX_{n-1} = 1, X_{n-2}X_{n-1} = 1\}}{P\{X_{n-1}X_n = 1, X_{n-2}X_{n-1} = 1\}} =$$

$$= \frac{P\{X_{n-2} = X_{n-1} = X_n = X_{n+1} = 1\} + P\left\{X_n = X_{n-1} = X_{n-2} = X_{n+1} = -1\right\}}{P\left\{X_{n-2} = X_{n-1} = X_n = 1\right\} + P\left\{X_{n-2} = X_{n-1} = X_n = -1\right\}} = \frac{p^4 + q^4}{p^3 + q^3};$$

$$P\{Y_n = 1 / Y_{n-1} = 1\} = \frac{P\{X_{n-1} = X_n = X_{n+1} = 1\} + P\{X_{n-1} = X_n = X_{n+1} = -1\}}{P\{X_n = X_{n-1} = 1\} + P\{X_n = X_{n-1} = -1\}} =$$

$$= \frac{p^3 + q^3}{p^2 + q^2}.$$

В общем случае вычисленные вероятности интересующих нас событий не совпадают, однако их равенство происходит при следующих частных значениях параметра p: 0; 1; 1/2. Это значит, что при этих значениях параметра p последовательность $\{Y_n\}$ может быть цепью Маркова. Ответ на вопрос о марковости требует дополнительного исследования при $p = 1/2$. При $p = 0$ или $p = 1$ марковость очевидна. Докажем марковость $\{Y_n\}$ при $p = 1/2$. Для этого введем вспомогательные СВ a_i и b_i, принимающие значения 1 и –1 с равными вероятностями, и СВ $c_i = -b_i$, $i = 1, 2, ...$. Будем сравнивать вероятности:

$$P\{Y_n = a / Y_1 = a_1, ...Y_{n-1} = a_{n-1}\} = \frac{(1/2)^{n+1} + (1/2)^{n+1}}{(1/2)^n + (1/2)^n} = \frac{1}{2};$$

$$P\{Y_n = a / Y_{n-1} = a_{n-1}\} = \frac{P\{X_nX_{n+1} = a_n, X_{n-1}X_{n+1} = a_{n-1}\}}{P\{X_{n-1}X_n = a_{n-1}\}} =$$

$$= \frac{P\{X_{n+1} = 1, X_n = a_n, X_{n-1} = b_{n-1}\} + P\{X_{n+1} = -1, X_n = -a_n, X_{n-1} = c_n\}}{P\{X_{n-1} = 1, X_n = a_{n-1}\} + P\{X_{n-1} = -1, X_n = -a_{n-1}\}} =$$

$$= \frac{(1/2)^3 + (1/2)^3}{(1/2)^2 + (1/2)^2} = \frac{1}{2}.$$

Из совпадения этих вероятностей следует марковость $\{Y_n\}$ при $p = q = 1/2$.

Задание 4.4. Пусть $X_1, X_2, ...$ — последовательность бернуллиевских СВ с вероятностью успеха, равной p ($q = 1 - p$), $Y_n = X_n + X_{n+1}$. Проверьте, представляют ли из себя последовательности СВ $\{X_n\}$ и $\{Y_n\}$ цепи Маркова.

4.9. Моделирование цепи Маркова

1. Моделирование случайной цепи Маркова с *n* состояниями.

Шаги моделирования:

1) $n + 1$ раз генерируем по $n - 1$ базовой случайно величины (БСВ)[1]: $\vec{r} = (r_1, ..., r_{n-1})$;

2) строим вариационный ряд из $\vec{r}$: $\vec{r}_{(.)} = (r_{i_1}, ..., r_{i_{n-1}})$;

3) вычисляем длины отрезков, на которые $\vec{r}_{(.)}$ делит отрезок $[0; 1]$, — получаем вектор $\vec{p} = (p_1, ..., p_n)$;

4) первый из $n + 1$ результат считаем вектором начальных вероятностей $\vec{P}^{(0)}$, а остальные n — строками матрицы переходных вероятностей $P = P(1)$ (за один шаг) цепи Маркова.

2. Моделирование значения цепи Маркова на *m*-м шаге и траектории переходов за *m* шагов (состояние цепи назовем ее значением).

Шаги моделирования:

1) генерируем одно значение БСВ — r_1 и по $\vec{P}^{(0)}$, разыгрываем начальное состояние E_{k_1}, где $k_1 \in \{1, ..., n\}$;

2) генерируем еще одно состояние БСВ — r_2 и по k_1-й строке матрицы P разыгрываем состояние цепи E_{k_2};

3) заменяя значение k_1 на значение k_2, повторяем п. 2 m раз, тогда полученные последовательные результаты считаем траекторией цепи Маркова за m шагов, а последнее состояние — значением цепи на m-м шаге.

Замечание 4.3. Для моделирования N значений цепи Маркова за m шагов и траекторий переходов повторяем процедуру 1—3 N раз.

Замечание 4.4. При большом N можно частотно оценить вектор безусловных вероятностей цепи Маркова за m шагов по результатам N-кратного проведения процедуры 1—3.

Замечание 4.5. Для разыгрывания значения цепи Маркова по вектору $\vec{p} = (p_1, ..., p_n)$, где $\sum_{i=1}^{n} p_i = 1$, нужно разделить $[0; 1]$ точками $\sum_{i=1}^{l} p_i$, $l = 1, ..., n$. Тогда попадание БСВ r в j-й отрезок, $j = 1, ..., n$, означает приход цепи Маркова в состояние E_j .

Задача 4.7. Пусть разыграна цепь Маркова с тремя состояниями:

$\vec{P}^{(0)} = (0{,}2;\ 0{,}3;\ 0{,}5)$; $P = \begin{pmatrix} 0{,}1 & 0{,}4 & 0{,}5 \\ 0{,}2 & 0{,}1 & 0{,}7 \\ 0{,}6 & 0{,}2 & 0{,}2 \end{pmatrix}$. Требуется разыграть случайную траекторию цепи Маркова из шести последовательных состояний со случайными значениями в четырех состояниях, включая последнее.

Решение

а) Разыгрываем случайные моменты, предшествующие шестому, т.е. совершаем случайный выбор без возвращения из чисел 1, 2, 3, 4, 5 объемом 3, упорядоченный по возрастанию. Для этого генерируем пять БСВ: $\vec{r} = (0{,}41;\ 0{,}73;\ 0{,}15;\ 0{,}34;\ 0{,}62)$; строим из $\vec{r}$ вариационный ряд $\vec{r}_{(.)} = (0{,}15;\ 0{,}34;\ 0{,}41;\ 0{,}62;\ 0{,}73)$; выбираем первые три числа из $\vec{r}$: 0,41; 0,73; 0,15 и находим их номера в $\vec{r}_{(.)}$ в порядке просмотра $\vec{r}_{(.)}$ — получаем упорядоченную по возрастанию последовательность моментов (1; 3; 5). Дополняя ее заранее выбранным шестым моментом, окончательно получаем следующие моменты: (1; 3; 5; 6).

б) Разыгрываем случайные значения цепи Маркова в полученные моменты. Для этого делим отрезок [0; 1] на три равные части, подряд их нумеруя; генерируем 4 БСВ; $\vec{r} = (0{,}54;\ 0{,}47;\ 0{,}82;\ 0{,}13)$ и последовательно выписываем номера частей отрезка, куда они попали, получаем числа (2; 2; 3; 1). Считаем их номерами состояний цепи Маркова в смоделированные выше моменты времени.

в) Теоретически вычисляем вероятность смоделированной в п. а) и б) траектории цепи Маркова, т.е. находим вероятность $p = p_2(1)p_{22}(2)p_{23}(2)p_{31}(1)$, где $p_2(1)$ — безусловная вероятность состояния E_2 на 1-м шаге, а $p_{ij}(k)$ — вероятности перехода из состояния E_i в состояние E_j за k шагов, которые нужно предварительно вычислить:

$$p_2(1) = 0{,}2\cdot 0{,}4 + 0{,}3\cdot 0{,}1 + 0{,}5\cdot 0{,}2 = 0{,}21;$$

$$P(2) = P^2 = \begin{pmatrix} 0{,}1 & 0{,}4 & 0{,}5 \\ 0{,}2 & 0{,}1 & 0{,}7 \\ 0{,}6 & 0{,}2 & 0{,}2 \end{pmatrix}^2 = \begin{pmatrix} 0{,}39 & 0{,}18 & 0{,}43 \\ 0{,}46 & 0{,}23 & 0{,}31 \\ 0{,}32 & 0{,}30 & 0{,}41 \end{pmatrix} \Rightarrow$$

$$\Rightarrow p = 0{,}21\cdot 0{,}23\cdot 0{,}31\cdot 0{,}6 = 0{,}004278.$$

Задание 4.4. Отработайте следующие пункты.

1. Смоделировать случайную цепь Маркова с $n = 3$ состояниями.

2. Смоделировать: а) 5 значений цепи Маркова после 1-го шага; б) 3 значения цепи Маркова после 5 шагов и выписать траекторию переходов за 5 шагов.

3. Смоделировать 1000 значений цепи Маркова за 2 шага и приближенно вычислить вектор безусловных вероятностей ее состояний за 2 шага.

4. Вычислить точно вектор безусловных вероятностей цепи Маркова за 2 шага и сравнить его с результатом 3).

5. Составить компьютерную программу вычисления вектора безусловных вероятностей для цепи Маркова с n состояниями за m шагов.

6. Смоделировать для цепи Маркова, полученной в 1), случайную траекторию из шести состояний со случайными значениями в четырех состояниях, включая последнее, и теоретически вычислить вероятность смоделированной траектории.

Контрольные вопросы и задания

1. Каково место цепей Маркова в классификации случайных процессов?

2. Приведите примеры задания цепей Маркова.

3. Дайте определения цепи Маркова и докажите их эквивалентность.

4. Как можно вычислить вероятности состояний цепи Маркова за n шагов?

5. Каковы свойства траекторий цепи Маркова?

6. Приведите примеры проверки марковости с положительным и отрицательным выводами.

Глава 5
КЛАССИФИКАЦИЯ СОСТОЯНИЙ ЦЕПЕЙ МАРКОВА

В результате изучения данной главы студенты должны:

знать

• определения и смысл понятий в классификации состояний цепи Маркова;

• связи свойств состояний;

уметь

• проводить классификацию состояний цепи Маркова с теоретическим обоснованием;

владеть

• навыками анализа свойств состояний цепи Маркова.

В главе приводится подробная классификация состояний цепей Маркова. Даны определения основных понятий, относящихся к свойствам состояний. Все эти свойства изучаются во взаимной связи с теоретическим обоснованием. Задаются критерии определения ряда свойств состояний цепей Маркова. Все теоретические исследования проиллюстрированы многочисленными разобранными примерами.

5.1. Определение основных понятий

Для более детального и глубокого изучения свойств цепи Маркова приведем описание характеристик ее состояний. Будем изучать асимптотическое поведение цепей Маркова при неограниченном росте числа шагов. Это оказывается возможным на основании предварительного проведения классификации ее состояния.

Приведем основные понятия.

Определение 5.1. Назовем состояние E_j ***достижимым*** из состояния E_i, если существует такое целое число n, что $P(n) > 0$.

Определение 5.2. Два состояния E_i и E_j называются ***сообщающимися*** или связанными между собой, если они достижимы друг из друга.

Определение 5.3. Состояние E_i называется ***несущественным***, если для него существует другое состояние E_j такое, что при некотором n $P_{ij}(n) > 0$, но не существует числа шагов m, при котором $P_{ji}(m) > 0$. Все остальные состояния называются ***существенными***.

Введем вероятность $f_i(n)$ — вернуться в состояние E_i за n шагов впервые. Тогда $F_i = \sum_{n=1}^{\infty} f_i(n)$ — вероятность вернутся в состояние E_i когда-нибудь.

Определение 5.4. Если $F_i = 1$, то состояние E_i называется ***возвратным***, а если $F_i < 1$ — ***невозвратным***.

Для возвратных состояний введем среднее время (число шагов) до возвращения в данное состояние E_i и обозначим его через m_i: $m_i = \sum_{n=1}^{\infty} nf_i(n)$.

Определение 5.5. Будем называть состояние E_i ***нулевым***, если $m_i = \infty$. Все остальные возвратные состояния (и все невозвратные состояния) будем называть ***ненулевыми***.

Если $P_{ij}(n)$ — элемент матрицы переходных вероятностей цепи Маркова за n шагов, то $P_{ij}(n) = \sum_{k=0}^{n-1} f_i(n-k)P_{ij}(k)$.

Определение 5.6. ***Периодом*** состояния E_i называется наибольший общий делитель d_i таких целых чисел n, для которых $P_{ii}(n) > 0$. При $d_i > 1$ состояние E_i называется периодическим с периодом d_i. При $d_i = 1$ состояние E_i называется непериодическим. Понятие периода определяется только для существенных состояний.

Определение 5.7. Возвратные ненулевые непериодические состояния называются ***эргодическими***.

Состояния цепи Маркова образуют замкнутый класс, если никакое состояние внутри этого класса не может быть достигнуто ни из какого состояния вне этого класса. Если какое-то состояние образует замкнутый класс, то оно называется поглощающим. Из определения замкнутого класса следует, что несущественные состояния не образуют замкнутого класса.

Определение 5.8. Цепь Маркова называется ***неприводимой*** (неразложимой), если в ней нет никаких замкнутых классов, кроме множества всех состояний. В остальных случаях цепь приводима (разложима).

Критерий неприводимости цепи Маркова. *Цепь Маркова неприводима тогда и только тогда, когда любое ее состояние может быть достигнуто из любого другого состояния.*

Критерий приводимости цепи Маркова. *В случае приводимой цепи Маркова из нее может быть выделен один или несколько замкнутых классов состояний (в случае одного замкнутого класса в приводимой цепи он не совпадает с множеством всех состояний).*

В общем случае путем перестановки строк и столбцов матрицы одношаговых переходных вероятностей P (это соответствует перенумерации состояний цепи) она может быть приведена к следующему виду:

$$P = \begin{pmatrix} A_{12} & 0 & 0 & \cdots & 0 \\ A_{21} & A_{22} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & A_{n3} & \cdots & A_{nn} \end{pmatrix}. \tag{5.1}$$

Здесь $A_{11}, A_{22}, ..., A_{nn}$ — квадратные матрицы, размерностями совпадающие с численностями состояний соответствующих им замкнутых классов. Подматрицы A_{ij} с $i \neq j$ являются прямоугольными и соответствуют состояниям, не входящим ни в один из замкнутых классов. Поэтому правый верхний угол матрицы P занят нулями. Таким образом, любую квадратную матрицу A_{ij} можно рассматривать как матрицу одношаговых переходных вероятностей неприводимой цепи Маркова, состоящую из состояний i-го замкнутого класса (так как условие $P_{ij} > 0$ и стохастичность матрицы A_{ii} имеют место).

Если C — замкнутое множество состояний, то $P_{jk}(n) = 0$ при всех j и k таких, что $E_i \in C$, а $E_k \notin C$ при любом n. Если в матрице P вычеркнуть все строки и столбцы, соответствующие состояниям, не входящим в замкнутые множества C, то полученная таким образом матрица соответствует неприводимой цепи Маркова, так как для нее выполняются свойства $P_{ij} \geqslant 0$ и $\sum\limits_{j=1} P_{ij} = 1$.

Сформулируем две теоремы.

Теорема 5.1 (теорема солидарности 1). *В неприводимой цепи все состояния — состояния одного типа, т.е. или все — невозвратные, или все — возвратные ненулевые, или все — эргодические* (*возвратные ненулевые непериодические*). *Если одно состояние — периодическое с периодом d, то все состояния — периодические с тем же периодом d.*

Теорема 5.2 (о классификации состояний в конечной цепи Маркова). *Конечная цепь Маркова не содержит нулевых состояний и не может состоять только из невозвратных состояний.*

Поясним приведенные понятия и теоремы на задачах.

Задача 5.1. По данному графу перехода из состояния в состояние (рис. 5.1) выписать матрицу одношаговых переходных вероятностей P

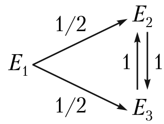


Рис. 5.1. **Направления и вероятности переходов для задачи 5.1**

цепи Маркова и провести ее классификацию или классификацию ее состояний.

Решение

Выпишем матрицу одношаговых переходных вероятностей, предварительно переставив 1-ю и 3-ю строки, т.е. перенумеровав 1-е и 3-е состояния. Получим матрицу P в виде

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

Очевидно, что в старой нумерации состояния E_2 и E_3 — существенные сообщающиеся, а состояние E_1 — несущественное. Цепь Маркова разложима. Состояния E_2 и E_3 образуют замкнутый класс, состояние E_1 не образует замкнутого класса. Состояния E_2 и E_3 — состояния одного типа (по теореме солидарности). Поэтому достаточно исследовать одно из этих состояний. Состояние E_2 возвратно, так как $f_2(2) = 1$, $f_2(n) = 0$ при $n \neq 2$, следовательно, $F_2 = 1$. Состояние E_2 — ненулевое, так как $m_2 = 2f_2(2) = 2 < \infty$ (или это следует из теоремы 5.2, по которой в ней нет нулевых состояний). Состояние E_2 — периодическое с периодом 2, а значит, оно неэргодическое. Состояние E_1 — невозвратное, так как $F_1 = 0$.

Задача 5.2. По графу состояний (рис. 5.2) дать классификацию цепи Маркова.

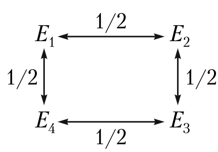


Рис. 5.2. **Граф состояний для задачи 5.2**

Решение

Выпишем матрицу переходов:

$$P = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}.$$

Все состояния — связанные между собой, существенные, образуют один замкнутый класс состояния; цепь неприводима; все состояния — периодические с периодом 2; все состояния — неэргодические.

Задача 5.3. По графу состояний (рис. 5.3) дать классификацию цепи Маркова.

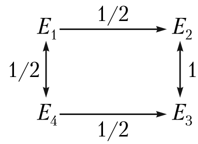


Рис. 5.3. **Граф состояний для задачи 5.3**

Решение

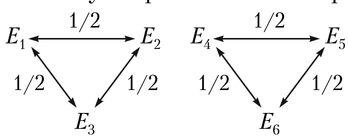
Выпишем матрицу переходов: $P=\begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$.

Получаем: E_1 и E_4 — несущественные сообщающиеся состояния; E_2 и E_3 — существенные сообщающиеся состояния; цепь разложима; перенумерацией первого и третьего состояний она приводит к виду

$$P=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}.$$

Этот вид соответствует форме (5.1).

Задание 5.1. По графам состояний дайте классификацию цепи Маркова. В случае разложимости приведите к виду (5.1).

а) 

б) $E_1 \xrightarrow{\;1\;} E_2 \xleftrightarrow{\;1\;} E_3 \xleftrightarrow{\;1\;} E_4$

Задача 5.4. В матрице переходных вероятностей цепи Маркова за один шаг указаны ненулевые элементы (знаком \$):

$$P=\begin{pmatrix} 0 & 0 & 0 & \$ & 0 & 0 & 0 & 0 & \$ \\ 0 & \$ & \$ & 0 & \$ & 0 & 0 & 0 & \$ \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \$ & 0 \\ \$ & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \$ & 0 & 0 & 0 & 0 \\ 0 & \$ & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \$ & 0 & 0 & 0 & \$ & \$ & 0 & 0 \\ 0 & 0 & \$ & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \$ & 0 & 0 & 0 & 0 & \$ \end{pmatrix}.$$

Найти все замкнутые множества состояний и их периоды.

Решение

Получаем: E_5 — поглощающее состояние; E_3 и E_8 — замкнутый класс состояний, $d = 2$; E_1, E_4 и E_9 — замкнутый класс состояний, $d = 1$, так как $P_{44}(2) > 0$ и $P_{44}(3) > 0$; E_2, E_6, E_7 — несущественные состояния (все состояния каждого замкнутого класса представляют собой неприводимую цепь Маркова и по теореме солидарности 1 являются состояниями одного типа).

Задание 5.2. Дайте классификацию цепи Маркова и ее состояний по следующим одношаговым матрицам переходных вероятностей: а) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; б) $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$; в) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Задача 5.5. Дать классификацию цепи Маркова и ее состояний по следующим одношаговым матрицам переходных вероятностей:

а) $\begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix}$; б) $\begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}$.

Решение

а) E_2 — поглощающее состояние, а значит, оно эргодическое; E_1 — несущественное состояние, $f_1(1) = 1/2$, $f_1(n) = 0$ при $n \neq 1$, следовательно, $F_1 < 1$, значит, E_1 — невозвратное состояние, $P_{11}(1) = 1/2 > 0$, поэтому E_1 — непериодическое состояние.

б) E_1 и E_2 — сообщающиеся состояния, поэтому цепь неразложима, это состояния одного типа (по теореме солидарности 1), и достаточно исследовать одно из них, например E_2:

$$f_2(1) = 0,\ f_2(2) = 1/2,\ ...,\ f_2(n) = (1/2)^{n-1};$$

$$F_2 = \sum_{n=1}^{\infty} f_2(n) = 1/2 + (1/2)^2 + ... + (1/2)^n + ... = \frac{1/2}{1-1/2} = 1,$$

следовательно, E_1 и E_2 — возвратные состояния; $P_{11}(1) = 1/2 > 0$, поэтому E_1 и E_2 — непериодические состояния.

Исследуем для состояния E_2 среднее время возвращения:

$$M_2 = \sum_{n=1}^{\infty} n f_2(n) = \sum_{n=2}^{\infty} \frac{n}{2^{n-1}}.$$

По признаку Даламбера получаем, что этот ряд сходится, так как

$$\lim_{n \to \infty} \frac{n+1}{2^n} = \frac{1}{2} < 1,$$

следовательно, E_1 и E_2 — ненулевые, а значит, оба состояния — эргодические.

Замечание 5.1. В данном случае проще было исследовать на ненулевое состояние E_1. Действительно:

$$M_1 = \sum_{n=1}^{\infty} n f_1(n) = 1 \cdot 0{,}5 + 2 \cdot 0{,}5 = 1{,}5 < \infty,$$

т.е. состояние E_1, а значит, и E_2 оба ненулевые.

Задача 5.6. Доказать, что конечная цепь Маркова не содержит нулевых состояний и не может состоять только из невозвратных состояний.

Решение (*доказательство*)

Проводим от противного.

Рассмотрим сначала случай *неприводимой* цепи Маркова. Допустим противное, т.е. что все состояния — невозвратные и хотя бы одно — возвратное нулевое. По теореме солидарности 1 последнее предположение равносильно тому, что все состояния — нулевые. Тогда для любой фиксированной пары индексов j и k (в соответствии с приведенными необходимыми и достаточными условиями для нулевого и невозвратного состояний) это значит, что $P_{jk}(n) \to \infty$ при $n \to \infty$. А это невозможно, так как при конечном числе состояний цепи нарушается стохастичность матрицы переходных вероятностей за n шагов.

Таким образом, утверждение доказано для любого замкнутого класса разложимой цепи Маркова. Остается рассмотреть оставшиеся после выделения замкнутых классов несущественные состояния. Они невозвратны, так как из любого несущественного состояния (по определению) с положительной вероятностью можно уйти и не вернуться никогда, значит, вероятность возвращения когда-нибудь для этого состояния $E_1F_1 < 1$, что и означает его невозвратность. А невозвратные состояния по определению являются ненулевыми, так как нулевые состояния — это возвратные состояния.

Осталось показать, что конечная цепь Маркова не может состоять из одних несущественных состояний. Допустим противное, т.е. что все состояния — несущественные. Покажем, что это приводит к противоречию.

Для удобства перенумеруем состояния так, что соответственно для состояний с номерами 1, 2, ..., $n - 1$ теми состояниями, куда можно уйти, а вернуться нельзя, будут состояния с номерами 2, 3, ..., n. А тогда для n-го состояния невозможно указать такое состояние, куда можно пойти, а вернуться нельзя, потому что (в силу конечности цепи) иначе мы вернемся в одно из предшествующих состояний, что противоречит предположению их несущественности. Утверждение полностью доказано.

Задача 5.7. Дана матрица одношаговых переходных вероятностей цепи Маркова:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Число состояний цепи равно n. Дать классификацию цепи и ее состояний. Выписать любую ее степень.

Решение

Все состояния связанные, значит, цепь неприводима, а все ее состояния принадлежат одному замкнутому классу. По теореме солидарности 1 все состояния — состояния одного типа. По примеру 6 все состояния — возвратные ненулевые. Очевидно, что все они — периодические с периодом n. При возведении в любую k-ю степень каждая строка матрицы циклически сдвигается вправо на k шагов.

Задача 5.8. Рассмотрим случайное блуждание на бесконечной прямой в соответствии с графом на рис. 5.4. Имеем $p + q = 1$. Дать классификацию цепи Маркова и ее состояний.

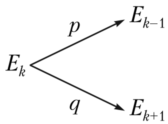


Рис. 5.4. **Граф состояний для задачи 5.8**

Решение

Все состояния — сообщающиеся, значит, цепь неразложима и все ее состояния одного типа. Будем исследовать нулевое состояние. Получаем $P_\infty = 0$ при $n = 2k + 1$ и $P_\infty = C_{2k}^{k}p^{kqk}$ при $n = 2k$.

Следовательно, все состояния — периодические с периодом $d = 2$. Для проверки необходимых и достаточных условий возвратных нулевых и невозвратных состояний воспользуемся формулой Стирлинга: $n! \to \sqrt{2\pi n}n^n e^{-n}$. Получим при $n = 2k$:

$$P_\infty(n) = \frac{(2k)!\,p^k q^k}{k!k!};\ P_\infty(n) \underset{n\to\infty}{\to} \frac{\sqrt{\pi k}(2k)^{2k}\,p^k q^k}{\pi k k^{2k}} = \frac{(4pq)^k}{\sqrt{\pi k}} \underset{n\to\infty}{\to} 0.$$

Следовательно, все состояния — нулевые. Далее составим ряд

$$\sum_{k=1}^{\infty} P_\infty(2k) \underset{k\to\infty}{\to} \sum_{k=1}^{\infty} \frac{(4pq)^k}{\sqrt{\pi k}}.$$

Ряд сходится при $p = q = 1/2$, так как $4pq = 1$, тогда все состояния — возвратные; ряд сходится при $p \neq q$, так как $4pq < 1$ (по признаку Даламбера), и тогда все состояния — невозвратные.

Замечание 5.2. В отличие от случая конечной цепи Маркова бесконечная цепь может содержать нулевые состояния и может состоять из невозвратных состояний.

Задача 5.9. Рассмотрим симметричное случайное блуждание частицы на плоскости в соответствии с графом на рис. 5.5. Дать классификацию цепи и провести анализ состояний на возвратность.

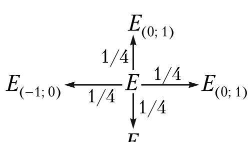


Рис. 5.5. **Граф состояний для задачи 5.9**

Решение

Все состояния сообщающиеся, одного типа, цепь неприводима. Будем исследовать начало координат. Возвращение в начало координат возможно только тогда, когда количества шагов соответственно в положительном и отрицательном направлениях осей Ox и Oy равны между собой. Поэтому при нечетном n $P_\infty(n) = 0$. При четном n ($n = 2k$) выпишем вероятность возвращения в начало координат:

$$P_\infty(2n) = \frac{\frac{1}{4^{2n}}\sum_{k=0}^{2n}(2n)!}{k!k!(n-k)!(n-k)!} = \frac{\frac{1}{4^{2n}}\sum_{k=0}^{2n}(2n)!(n!)^2}{(k!)^2[(n-k)!]^2(n!)^2} = \frac{C_{2n}^n}{4^{2n}}\sum_{k=0}^{2n}C_n^k = \frac{(C_{2n}^n)^2}{4^{2n}}.$$

По формуле Стирлинга полученная дробь стремится при $n \to \infty$ к выражению

$$\frac{2\pi 2n2n\exp(-4n)}{4^{2n}2^2\pi^2 4^{2+4n}\exp(-4n)} = \frac{1}{n},$$

поэтому ряд $\sum_{k=0}^{2n} P_\infty(2k)$ расходится. Следовательно, все состояния — возвратные.

Задание 5.3. Аналогично покажите, что при симметричном случайном блуждании в трехмерном пространстве и пространстве большей размерности все состояния невозвратны.

Рассмотрим ряд теоретических задач-утверждений.

Утверждение 5.1. Неприводимая цепь Маркова, у которой хотя бы один диагональный элемент $P_{ii} > 0$, является непериодической.

Доказательство следует из теоремы солидарности 1 и из того, что при $P_{ii} > 0$ период d_i состояния E_i равен 1.

Утверждение 5.2. Для любой конечной цепи Маркова с n состояниями всякое состояние $E_к$ достижимо из состояния E_j, или недостижимо вообще, или достижимо не более чем за n шагов.

Доказательство. Если состояние E_k достижимо из состояния $E_{j,}$ то можно указать цепочку переходов, связывающих состояния E_k и E_j, причем поскольку каждое состояние цепочки участвует в ней по одному разу, то всех переходов не более n. Если же состояние E_k из состояния E_j недостижимо, то и тогда снова утверждение верно.

Утверждение 5.3. В конечной цепи множества возвратных и существенных состояний совпадают.

Доказательство. Пусть состояние E_j существенно, следовательно, оно принадлежит одному из замкнутых классов состояний, каждый из которых может быть рассмотрен как неприводимая цепь Маркова, в которой все состояния одного типа. Тогда по результату примера 6 все состояния возвратны.

Пусть теперь состояние E_j возвратно. Предположим, что оно несущественно. Тогда найдется состояние E_k и целое I такое, что $P_{jk}(I) > 0$, а при любом m $P_{kj}(m) = 0$, т.е. с положительной вероятностью система за конечное число шагов перейдет из состояния E_j в состояние E_k и никогда не вернется в состояние E_j. А это значит, что вероятность вернуться когда-нибудь в состояние $E_jF_j < 1$, что означает невозвратность состояния E_j и противоречит исходному условию, откуда и следует доказываемое утверждение.

Замечание 5.3. Условием конечности цепи здесь мы пользовались только в первой части доказательства, откуда следует, что возвратное состояние всегда существенно. Обратное же установлено лишь для конечных цепей. Можно привести пример бесконечной цепи, в которой это не так.

Задание 5.4. Докажите следующие теоретические утверждения.

1. Если состояние E_i возвратно и сообщается с состоянием E_j, то и состояние E_j тоже возвратно. (Доказывать это утверждение без использования теоремы солидарности 1.)
2. Замкнутые классы состояний или не пересекаются, или совпадают.

5.2. Стационарность и эргодичность цепи Маркова

Смысл *стационарности* цепи Маркова состоит в независимости распределения ее состояний от времени течения процесса.

Смысл *эргодичности* цепи Маркова состоит в утрате зависимости вероятности состояния системы с ростом времени от ее начального состояния и существовании предельного распределения ее состояний.

Дадим теперь точные определения этим понятиям.

Определение 5.9. Цепь Маркова ***эргодична***, если $P_{ik}(n) \underset{n\to\infty}{\longrightarrow} \pi_k$, где $\sum_{k=1}^{m} \pi_k = 1$. Тогда $\{\pi_k\}$, $k = 1, 2, ..., m$, — финальные вероятности всех m состояний цепи.

В матричной форме эргодичность цепи означает, что

$$P^n \underset{n\to\infty}{\longrightarrow} \pi = \begin{pmatrix} \pi_1 & \dots & \pi_m \\ \dots & \dots & \dots \\ \pi_1 & \dots & \pi_m \end{pmatrix}; \quad \sum_{k=1}^{m} \pi_k = 1;$$

π — матрица финальных вероятностей, составленная из n векторов $\overline{\pi} = (\pi_1, ..., \pi_m)$.

Выведем уравнение для нахождения финальных вероятностей. Известно уравнение Колмогорова — Чепмана:

$$p_{ik}(n + 1) = \sum_{j=1}^{m} p_{ij}(n) p_{jk},$$

которое при $n \to \infty$ для эргодической цепи приводит к равенству $\pi_k = \sum_{j=1}^{m} \pi_j p_{jk}$, или в матричной форме $\overline{\pi} = \overline{\pi}P$. Это уравнение для финальных вероятностей.

Докажем следующие утверждения.

Утверждение 5.4. Для эргодической цепи Маркова $p_k(n) \to \pi_k$ при $n \to \infty$.

Доказательство. Известно, что для вектора безусловных вероятностей $\vec{p}(n)$ на n-м шаге верно равенство $\vec{p}(n) = \vec{p}(1)P^{n-1}$, где P — матрица одношаговых переходных вероятностей цепи Маркова. По условию при $n \to \infty$ $p_k(n) \to \pi_k$, отсюда следует, что

$$\vec{p}(n) = \vec{p}(1)\begin{pmatrix} \pi_1 & \dots & \pi_m \\ \dots & \dots & \dots \\ \pi_1 & \dots & \pi_m \end{pmatrix},$$

или покомпонентно $p_k(n) = p_1(1)\pi_k + p_2(1)\pi_k + ... + p_m(1)\pi_k = \pi_k(p_1(1) + p_2(1) + + ... + p_m(1)) = \pi_k$, $k = 1, ..., n$, ч.т.д.

Утверждение 5.5. Если $p_k(n) \underset{n\to\infty}{\to} \pi_k$, $k = 1, ..., m$, то цепь Маркова эргодична.

Доказательство. Известно, что $\vec{p}(n) = \vec{p}(n-1)P$, откуда по условию утверждения следует, что при $n \to \infty$ $\overline{\pi} = \overline{\pi}P$, а это покомпонентно эквивалентно равенству $\pi_k = \sum_{j=1}^{m} \pi_j p_{jk}$, $k = 1, ..., m$, т.е. по определению имеем, что $\{\pi_k\}$ — финальные вероятности, существование которых и означает эргодичность цепи Маркова, ч.т.д.

Замечание 5.4. На основании утверждений 5.4 и 5.5 можно дать второе определение эргодичности цепи Маркова.

Определение 5.10. Цепь Маркова ***эргодична***, если для любого целого k

$$p_k(n) \underset{n\to\infty}{\to} \pi_k, \ \sum_{k=1}^{m} \pi_k = 1.$$

Теперь рассмотрим понятие стационарности.

Определение 5.11. Распределение вероятностей $\{\nu_1, \nu_2, ..., \nu_m\}$ называется ***стационарным***, если $\nu_j = \sum_{k=1}^{m} \nu_i p_{ij}$; $\sum_{k=1}^{m} \nu_k = 1$, или в матричной форме $\overline{\nu} = \overline{\nu}P$.

Физический смысл стационарности — это макроскопическое равновесие, поддерживаемое большим числом случайных переходов (отражение действия закона больших чисел).

Теорема 5.3 (теорема солидарности 2). *Неприводимая непериодическая цепь Маркова принадлежит к одному из следующих классов:*

а) или все состояния невозвратные, или все состояния нулевые; в этом случае $p_{jk}(n) \underset{n\to\infty}{\longrightarrow} 0$ *для любой пары индексов* j *и* k *и не существует стационарного распределения;*

б) или же все состояния эргодические, т.е. $\lim\limits_{n\to\infty} p_{jk}(n) = \pi_k > 0$, *где* π_k *— величина, обратная среднему времени возвращения в* E_k. *В этом случае* $\{\pi_k\}$ *— стационарное распределение и не существует никаких других стационарных распределений.*

Прокомментируем смысл последней теоремы в форме замечаний.

Замечание 5.5. Для неприводимой цепи Маркова если все состояния эргодические, то и цепь — эргодическая; в приводимой цепи Маркова это не так. Приведем пример. $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ — матрица одношаговых переходных вероятностей цепи Маркова, соответствующая приводимой цепи, состоящей из двух замкнутых классов, представляющих собой по одному поглощающему состоянию. Оба эти состояния — эргодические, однако финальных вероятностей нет, так как с течением времени не утрачивается зависимость от начального состояния, из которого система никогда не выходит.

Замечание 5.6. В случае б) теоремы цепь Маркова является эргодической и финальные вероятности совпадают со стационарными вероятностями.

Замечание 5.7. Стационарное распределение существует в условиях теоремы только в эргодическом случае (т.е. когда все состояния — эргодические, или, что то же самое по замечанию 2, когда цепь — эргодическая).

Замечание 5.8. В конечной цепи Маркова стационарное распределение существует всегда (в том числе и для периодических, а значит, неэргодических цепей).

Пример 5.1. $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ — матрица одношаговых переходных вероятностей, соответствующая периодической цепи Маркова ($d = 2$). Уравнения стационарности в этом случае примут вид

$$\begin{cases} \pi_1 = \pi_2, \\ \pi_2 = \pi_1, \\ \pi_1 + \pi_2 = 1 \end{cases} \Rightarrow$$

$\Rightarrow \pi_1 = 1/2$; $\pi_2 = 1/2$ — стационарное распределение; финальных вероятностей для данной матрицы нет, как было указано в замечании 5.5.

Замечание 5.9. В случае а) теоремы финальных вероятностей нет, так как все нули — это не распределение вероятностей. Таким образом, в неприводимой непериодической цепи Маркова эргодичности состояний и цепи следуют друг из друга.

Замечание 5.10. В приводимой цепи Маркова стационарных распределений столько, сколько замкнутых классов содержит цепь Маркова.

Замечание 5.11. В приводимой и периодической цепи Маркова нет финальных вероятностей, так как с течением времени зависимость от начального состояния не пропадет.

Замечание 5.12. Случай а) теоремы относится только к цепям Маркова с бесконечным числом состояний. Это следует из утверждения примера 6 из параграфа 5.1 (конечная цепь Маркова не содержит нулевых состояний и не может состоять только из невозвратных состояний).

Задача 5.10. Дана матрица одношаговых переходных вероятностей цепи Маркова

$$P=\begin{pmatrix} 1/2 & 1/4 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}.$$

Исследовать цепь на стационарность и эргодичность и найти соответствующие распределения, если они есть.

Решение

Имеем конечную неприводимую цепь Маркова; все ее состояния — возвратные ненулевые непериодические. Следовательно, по теореме солидарности 2 (см. замечание 3) существуют финальные и стационарные вероятности, которые совпадают и находятся из системы уравнений

$$p_j = \sum_i p_i p_{ij},\ j = 1, 2, 3,$$

т.е.

$$\begin{cases} p_1 = 0{,}5p_1 + 0{,}25p_2 + 0{,}25p_3, \\ p_2 = 0{,}25p_1 + 0{,}5p_2 + 0{,}25p_3, \\ p_3 = 0{,}25p_1 + 0{,}5p_2 + 0{,}25p_3, \\ p_1 + p_2 + p_3 = 1. \end{cases}$$

Отсюда имеем вектор финальных вероятностей, совпадающий с вектором стационарных вероятностей $\{1/3, 1/3, 1/3\}$. Тогда матрица финальных вероятностей π имеет вид

$$\pi = \lim_{n\to\infty} P^n = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

Докажем часть последней теоремы в форме утверждения.

Утверждение 5.6. Если цепь Маркова эргодична, то у системы уравнений для финальных вероятностей $(p_1, ..., p_m)P = (p_1, ..., p_m)$ всегда есть решение, имеющее вероятностный смысл.

Доказательство. Для безусловных вероятностей верно, что

$$(p_1(n-1), ..., p_m(n-1))P = (p_1(n), ..., p_m(n)).$$

Если цепь эргодична, то отсюда при $n \to \infty$ следует, что

$$(\pi_1, ..., \pi_m)P = (\pi_1, ..., \pi_m),$$

причем $\sum_{i=1}^{m} \pi_i = 1$, значит, решение $(\pi_1, ..., \pi_m)$ дает финальные вероятности.

Признаки эргодичности цепи Маркова

Пусть E — k-мерная единичная матрица. Тогда характеристической матрицей для цепи Маркова P с k состояниями будем называть матрицу вида

$$(P-\lambda E)=\begin{pmatrix} p_{11}-\lambda & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22}-\lambda & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn}-\lambda \end{pmatrix}.$$

Этой матрице соответствует характеристическое уравнение

$$|P-\lambda E| = 0, \tag{5.2}$$

его корни — характеристические числа.

Очевидно, что $\lambda = 1$ — всегда характеристическое число. Это проверяется при непосредственной подстановке в формулу (5.2):

$$\begin{vmatrix} p_{11}-1 & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22}-1 & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn}-1 \end{vmatrix} = 0.$$

Равенство становится очевидным, если к первому столбцу прибавить все остальные, и тогда он становится нулевым.

Теорема 5.4. *Если все корни характеристического уравнения (5.2), кроме одного, меньше 1, то существует*

$$\lim_{n\to\infty} P^n = \pi = \begin{pmatrix} \pi_1 & \dots & \pi_k \\ \dots & \dots & \dots \\ \pi_1 & \dots & \pi_k \end{pmatrix},$$

т.е. тогда цепь Маркова — эргодическая.

Замечание 5.13. Если цепь Маркова содержит несколько классов существенных состояний, то она неэргодична, так как для нее $L = 1$ является корнем, кратность которого больше 1 (его кратность совпадает с числом классов существенных состояний цепи) (см. критерий приводимости цепи Маркова).

Замечание 5.14. Конечная цепь Маркова эргодична тогда и только тогда, когда она представляет один класс существенных состояний.

Задача 5.11. Проверить (по теореме 5.4) эргодичность цепи Маркова с матрицей переходных вероятностей P из примера 2.

Решение

Составим характеристическое уравнение

$$\begin{vmatrix} 0{,}5-\lambda & 0{,}25 & 0{,}25 \\ 0{,}25 & 0{,}5-\lambda & 0{,}25 \\ 0{,}25 & 0{,}5 & 0{,}25-\lambda \end{vmatrix} = 0,$$

которое сводится к уравнению $\lambda(\lambda^2 - 1{,}25\lambda + 0{,}25) = 0$, откуда $\lambda_1 = 0$; $\lambda_2 = 0{,}25$, $\lambda_3 = 1$.

Таким образом, условия из теоремы 5.4 выполнены, и цепь Маркова эргодична.

Теорема 5.5 (эргодическая теорема Маркова — Бернштейна). *Пусть у матрицы P (одношаговых переходных вероятностей цепи Маркова) имеется один столбец, состоящий из положительных элементов, тогда цепь — эргодическая.*

Теорема 5.6 (обобщенная эргодическая теорема). *Если в матрице P нет столбца из строго положительных элементов, но он есть в любой степени матрицы P^k (k — целое положительное число), то цепь — эргодическая.*

Для иллюстрации теорем 5.4—5.6 рассмотрим следующую задачу.

Задача 5.12. Проверить на эргодичность цепь Маркова, описывающую случайное блуждание, задаваемое схемой (рис. 5.6).

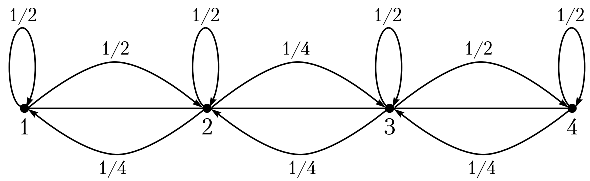


Рис. 5.6. **Цепь Маркова для задачи 5.12**

Решение

Здесь матрица P имеет вид

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 \\ 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Матрица P не имеет столбца из строго положительных элементов. Покажем, что такой столбец есть в P^2 и что это, например, второй столбец.

В этом можно убедиться непосредственным возведением в квадрат матрицы P. Сделаем это проще. Конкретные значения элементов матрицы P не важны, важна лишь их положительность. Поэтому достаточно просто показать, что за два шага из любого состояния возможно перейти во второе, т.е. надо указать пути переходов из всех состояний во второе состояние. Укажем их: $1 \to 2 \to 2$; $2 \to 2 \to 2$; $3 \to 2 \to 2$; $4 \to 3 \to 2$.

Убедимся теперь в эргодичности цепи по характеристическим числам, которые здесь принимают следующие значения: $\lambda_1 = 1$; $\lambda_2 = 3/4$; $\lambda_3 = 1/4$; $\lambda_4 = 0$, что по теореме 2 доказывает эргодичность цепи.

Найдем финальные вероятности, совпадающие со стационарными, из системы уравнений

$$\begin{cases} \pi_1 = 0{,}5\pi_1 + 0{,}25\pi_2, \\ \pi_2 = 0{,}5\pi_1 + 0{,}5\pi_2 + 0{,}25\pi_3, \\ \pi_3 = 0{,}25\pi_2 + 0{,}5\pi_3 + 0{,}5\pi_4, \\ \pi_4 = 0{,}25\pi_3 + 0{,}5\pi_4, \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1. \end{cases}$$

Задание 5.5. Дана цепь Маркова с матрицей P:

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Найти финальные и стационарные вероятности, если они есть.

Задание 5.6. Эргодичны ли цепи Маркова со следующими матрицами переходных вероятностей P:

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};\ P = \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix};\ P = \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix};\ P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/4 & 3/4 \end{pmatrix}.$$

Задание 5.7. Проверить по теореме 5.6 на эргодичность цепь Маркова с данной матрицей P переходных вероятностей:

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 \\ 0 & 1/4 & 3/4 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 0 & 0 & 1/2 & 1/4 \end{pmatrix}.$$

Задача 5.13 (модель Эренфеста для процесса диффузии). По двум резервуарам B и C распределено A молекул. На каждом шаге наугад выбранная молекула перемещается в другой резервуар. Состояние системы определяется числом молекул в резервуаре B. Исследовать на эргодичность и найти стационарное распределение, если оно есть (эту модель можно интерпретировать как случайное блуждание по целочисленным точкам отрезка с отражающими экранами).

Решение

Цепь — периодическая с периодом $d = 2$. Поэтому финальных вероятностей нет. Выпишем матрицу переходных вероятностей и систему уравнений для нахождения стационарного распределения:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 1/A & 0 & 1-1/A & 0 & \ldots & 0 & 0 & 0 \\ 0 & 2/A & 0 & 1-2/A & \ldots & 0 & 0 & 0 \\ 0 & 0 & 3/A & 0 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 1-1/A & 0 & 1/A \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix};$$

$$\begin{cases} \nu_0 = \nu_1/A, \\ \nu_1 = \nu_0 + 2\nu_2/A, \\ \nu_2 = (1 - 1/A)\nu_1 + 3\nu_3/A, \\ \nu_3 = (1 - 2/A)\nu_2 + 4\nu_4/A, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \nu_k = [1 - (k-1)/A]\nu_{k-1} + (k+1)\nu_{k+1}/A, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \nu_A = \nu_{A-1}/A. \end{cases}$$

Непосредственной подстановкой легко проверить, что $\nu_k = C_A^k/2^A$.

Это биномиальное распределение, поэтому полученный результат может быть интерпретирован следующим образом: каково бы ни было начальное число молекул в резервуаре B, через достаточно долгое время вероятность нахождения в нем ровно k молекул такая же, как если бы A молекул были распределены случайно, причем вероятность для каждой молекулы попасть в резервуар B равна 1/2.

Задача 5.14. Имеется по m белых и черных шаров, которые случайным образом разложены по двум урнам, по m шаров в каждой. На каждом шаге два шара, взятые из разных урн, меняются местами. Состояние системы определяется числом белых шаров в первой урне. Найти финальные вероятности, если они есть.

Решение

Выпишем граф переходов из состояния E_i за один шаг:

$i = 1, ..., m - 1$: $E_0 \xrightarrow{i} E_1$; $E_m \xrightarrow{i} E_{m-1}$;

$$\nearrow E_{i-1} \text{ с вероятностью } P_{i,i-1} = P(\text{БЧ}) = \frac{i}{m} \cdot \frac{i}{m} = \left(\frac{i}{m}\right)^2;$$

$$E_i \to E_i \text{ с вероятностью } P_{i,i} = P(\text{ББ}) + P(\text{ЧЧ}) = 2\left(\frac{i}{m}\right) \cdot \frac{m-i}{m} = \frac{2i(m-i)}{m^2};$$

$$\searrow E_{i+1} \text{ с вероятностью } P_{i,i+1} = P(\text{ЧБ}) = \left(\frac{m-i}{m}\right)^2.$$

Тогда матрица одношаговых переходных вероятностей P имеет вид

$$P=\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & \dots & 0 \\ \left(\frac{1}{m}\right)^2 & \frac{2(m-1)}{m^2} & \left(\frac{m-1}{m}\right)^2 & 0 & \dots & \dots & \dots \\ 0 & \left(\frac{2}{m}\right)^2 & \frac{2\cdot 2(m-2)}{m^2} & \left(\frac{m-2}{m}\right)^2 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \left(\frac{m-1}{m}\right)^2 & \frac{2(m-1)}{m^2} & \left(\frac{1}{m}\right)^2 \\ 0 & 0 & \dots & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Все состояния сообщающиеся — цепь Маркова неприводимая, непериодическая, все состояния — возвратные ненулевые, т.е. эргодические. Тогда по теореме солидарности 2 существуют финальные вероятности, определяемые из системы уравнений

$$\begin{cases} \pi_0=\dfrac{1}{m}, \\ \pi_1=\pi_0+2\left(\dfrac{1}{m}\right)\left(\dfrac{m-1}{m}\right)\pi_1+\left(\dfrac{2}{m}\right)^2\pi_2, \\ \pi_2=\left(\dfrac{m-1}{m}\right)^2\pi_1+2\left(\dfrac{2}{m}\right)\left(\dfrac{m-2}{m}\right)\pi_2+\left(\dfrac{3}{m}\right)^2\pi_3, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \pi_k=\pi_{k-1}\left(\dfrac{m-k+1}{m}\right)^2+\pi_k\cdot 2\cdot\dfrac{k}{m}\cdot\dfrac{m-k}{m}+\pi_{k+1}\left(\dfrac{k+1}{m}\right)^2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \sum\limits_{i=1}^{m}\pi_i=1. \end{cases}$$

Нетрудно (по индукции) проверить, что $\pi_k = (C_m^k)^2\pi_0$. Тогда $\pi_0\sum\limits_{i=0}^{m}(C_m^i)^2 = 1$.

Отсюда (использовано свойство $\sum\limits_{i=0}^{m}(C_m^i)^2 = C_{2m}^m$): $\pi_0 = \dfrac{1}{C_{2m}^m}$.

Следовательно, $\pi_k = \dfrac{C_m^k C_m^{m-k}}{C_{2m}^m}$. Таким образом, финальная вероятность π такова же, какова вероятность того, что в выборке ровно k белых шаров, если бы из $2m$ шаров (среди которых по m белых и черных) извлекли наугад m шаров, т.е. финальные вероятности имеют гипергеометрическое распределение и совпадают со стационарными.

Задача 5.15. Задано случайное блуждание на отрезке $[1; a]$ по схеме, изображенной на рис. 5.7.

Найти финальные и стационарные вероятности, если они есть.

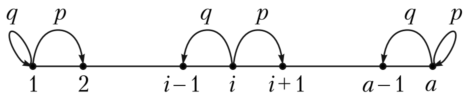


Рис. 5.7. **Схема случайного блуждания для задачи 5.15**

Решение

В $P^a = P(a)$ первый столбец состоит из строго положительных элементов, так как за a шагов можно из любого состояния попасть в первое. Следовательно, цепь Маркова — эргодическая. Все состояния этой неприводимой цепи непериодические, возвратные, ненулевые, т.е. эргодические.

Финальные и стационарные вероятности совпадают и находятся из системы уравнений

$$\begin{cases} \pi_1 = \pi_1 q + \pi_2 q, \\ \pi_2 = \pi_1 p + \pi_3 q, \\ \dots\dots\dots\dots\dots\dots \\ \pi_k = \pi_{k-1} p + \pi_{k+1} q, \\ \dots\dots\dots\dots\dots\dots \\ \pi_a = \pi_{a-1} p + \pi_a p, \\ \sum\limits_{i=1}^{a} \pi_i = 1. \end{cases}$$

По индукции нетрудно проверить, что $\pi_k = \left(\dfrac{p}{q}\right)^{k-1} \pi_1$, поэтому

$$\pi_1 \left[1 + \frac{p}{q} + \left(\frac{p}{q}\right)^2 + \dots + \left(\frac{p}{q}\right)^{a-1} \right] = 1,$$

отсюда $\pi_1 = \dfrac{1 - \left(\dfrac{p}{q}\right)^2}{1 - \dfrac{p}{q}}$, следовательно, $\pi_k = \left(\dfrac{p}{q}\right)^{k-1} \cdot \dfrac{1 - \left(\dfrac{p}{q}\right)^a}{1 - \dfrac{p}{q}}$, а значит, $\{\pi_k\}$ — искомые финальные и стационарные вероятности.

Задание 5.8. Рассмотреть аналогичное случайное блуждание по схеме, изображенной на рис. 5.8.

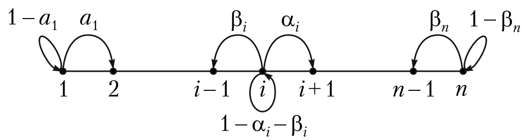


Рис. 5.8. **Схема к заданию 5.8**

Контрольные вопросы и задания

1. Приведите критерий приводимости цепи Маркова.
2. Какова классификация состояний конечной цепи Маркова?
3. Приведите доказательство уравнения Колмогорова — Чепмана.
4. Приведите пример применения теоремы солидарности 1.
5. Приведите пример применения теоремы солидарности 2.
6. Каковы признаки эргодичности цепи Маркова? Приведите примеры.
7. Приведите примеры нахождения финальных и стационарных вероятностей.

Глава 6
ВЕТВЯЩИЕСЯ И ПУАССОНОВСКИЕ ПРОЦЕССЫ

В результате изучения данной главы студенты должны:

знать

- основные понятия теории ветвящихся процессов и теоремы о них;
- определения пуассоновского процесса и его практическое использование;

уметь

- решать задачи определения матрицы переходных вероятностей ветвящегося процесса и вероятности его вырождения;
- находить характеристики пуассоновского процесса (потока) и вероятности его состояний;

владеть

- методом производящих функций для анализа ветвящихся процессов;
- навыками анализа пуассоновских потоков.

В главе вводится понятие ветвящегося процесса как вида марковского процесса. Определяется производящая функция в качестве инструмента изучения ветвящихся процессов. В ее терминах находятся вид матрицы его переходных вероятностей и моменты (математическое ожидание и дисперсия) численности n-го поколения ветвящегося процесса. Исследуется вопрос об условии вырождения процесса. Разбирается много примеров исследования ветвящихся процессов. Даны два определения пуассоновского процесса, и доказана их эквивалентность. Обсуждены его определяющие свойства, такие как независимость приращений, однородность (стационарность) и ординарность; рассмотрены некоторые другие свойства в форме разобранных теоретических задач.

6.1. Ветвящиеся процессы

Определение 6.1. *Ветвящийся процесс* — целочисленный марковский процесс с дискретным временем — процесс размножения и гибели частиц, в котором частицы развиваются независимо друг от друга.

Пусть количества порождений j-й частицы $V = V_j$ в любом поколении независимы и одинаково распределены с производящей функцией

$$\varphi_{v_j}(S) = \varphi(s) = \sum_{k=0}^{\infty} P(V_j = V = k)s^k,$$

где $V_j = V$ — случайное число порождений одной j-й частицы; $p_k = P(V = k) \geqslant 0$, $k = 0, 1, ...;$ $\sum\limits_{k=0}^{\infty} p_k = 1$.

Состояние процесса на n-м шаге определяется численностью n-го поколения X_n; $X_{n+1} = \sum\limits_{j=1}^{X_n} V_j$, т.е. состояние процесса на $(n + 1)$-м шаге зависит только от состояния процесса на n-м шаге и значений СВ $V_j, j = 1, 2, ...$.

Введем производящую функцию $\varphi_n(S) = \sum\limits_{k}^{\infty} P(X_n = k)s^k$ для численности n-го поколения X_n.

1. Докажем марковость ветвящегося процесса.

Доказательство

$$P\{X_{n+1} = j = i_{n+1} / X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i_n = i\} =$$

$$= \frac{P\{X_{n+1} = j, X_n = i_n = i, ..., X_0 = i_0\}}{P\{X_n = i_n = i, ..., X_0 = i_0\}} =$$

$$= \frac{P\{V_1 + ... + V_{i_n} = j, V_1 + ... + V_{i_{n-1}} = i_n = i, ..., V_1 = X_0 = i_0\}}{P\{V_1 + ... + V_{i_{n-1}} = 1_n = i, ..., V_1 + ... + V_{i_0} = i_1, V_1 = X_0 = i_0\}} =$$

$= \{$Обозначим события $V_1 + ... V_{i_{k-1}} = 1$ через A_k, $k = 1, 2, ...;$ $A_0 : \{V_1 = i_{i_0}\}$, тогда

$$A_{n+1} \subset A_n \subset ... \subset A_1 \subset A_0\} =$$

$$= \frac{P\{A_{n+1}, A_n, ..., A_1, A_0\}}{P\{A_n, ..., A_1, A_0\}} = \frac{P\{A_{n+1}, A_n\}}{P\{A_n\}} =$$

$$= \frac{P\{V_1 + ... + V_i = j,\ V_1 + ... + V_{i_{n-1}} = i\}}{P\{V_1 + ... + V_{i_{n-1}} = i\}} = P\{X_{n+1} = j \ / \ X_n = i\},$$

что и доказывает марковость процесса.

2. Найдем матрицу переходных вероятностей P ветвящегося процесса, если известно, что $X_0 = 1$ и $P\{V_j = k\} = p_k = P\{X_1 = 1\}$; $p_k \geqslant 0$.

Имеем

$$\sum_{k=0}^{\infty} p_k = 1;\ p_{ij}(X_n = j / X_{n-1} = i) = P\{V_1 + ... + V_i = j\} = P\{X_1 = j / X_0 = i\},$$

а это означает однородность марковского процесса.

Нулевая строка искомой матрицы P переходных вероятностей ветвящегося процесса состоит из вероятности перехода частиц из поглощающего состояния 0, и поэтому есть $(1, 0, ..., 0)$. Первая строка этой матрицы дает вероятности перехода частиц из состояния 1 во все другие состояния и получается из разложения в нуле (ряд Маклорена) по степеням s производящей функции $\varphi(s) = \varphi_1(s)$:

$$\varphi(s)=\varphi(0)+\frac{\varphi'(0)s}{1!}+\frac{\varphi''(0)s^2}{2!}+\ldots+\frac{\varphi^{(j)}(0)s^j}{j!}+\ldots.$$

Отсюда получаем, что $p_{1j} = \dfrac{\varphi^{(j)}(0)}{j!}$.

Вторая строка матрицы P дает вероятность перехода частиц из состояния 2 во все остальные состояния и получается аналогично первой строке из производящей функции $\varphi_2(s) = \varphi_{\nu_1}(s)\varphi_{\nu_2}(s) = \varphi^2(s)$ (так как $X = V_1 + V_2$, где V_1 и V_2 — независимые одинаково распределенные СВ):

$$\varphi_2(s)=\varphi^2(s)=\varphi^2(0)+\frac{(\varphi^2(0))'s}{1!}+\frac{(\varphi^2(0))''s^2}{2!}+\ldots+\frac{(\varphi^2(0))^{(j)}s^j}{j!}+\ldots.$$

Отсюда получаем, что $p_{2j} = \dfrac{(\varphi^2(0))^{(j)}}{j!}$, и из аналогичных соображений находим выражение для элемента i-й строки матрицы P: $p_{ij} = \dfrac{(\varphi^i(0))^{(j)}}{j!}$, т.е. матрица P найдена.

Замечание. Из аналогичных соображений нетрудно получить матрицу переходных вероятностей ветвящегося процесса за n шагов, если $X_0 = 1$, исходя из равенств

$$\varphi_1(s)=\varphi(s);\ \varphi_2(s)=\varphi(\varphi(s));\ \varphi_3(s)=\varphi(\varphi(\varphi(s)));\ \ldots;$$
$$\varphi_n(s)=\underbrace{\varphi(\varphi(\varphi(\ldots\varphi(s)))\ldots)}_{n\text{ раз}}.$$

Отсюда получаем, что $p_{ij}(n) = \dfrac{(\varphi_n^i(0))^{(j)}}{j!}$.

3. Докажем формулу

$$\varphi_{n+1}(s) = \varphi_n(\varphi(s)). \tag{6.1}$$

Доказательство

$$\varphi_{n+1}(s)=\sum_{k=0}^{\infty}P\{X_{n+1}=k\}s^k=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}P\{X_{n+1}=k\,/\,X_n=j\}P\{X_n=j\}s^k=$$
$$=\sum_{j=0}^{\infty}P\{X_n=j\}\sum_{k=0}^{\infty}P\{V_1+\ldots+V_j=k\}s^k=\sum_{j=0}^{\infty}P\{X_n=j\}(\varphi(s))^j=\varphi_n(\varphi(s)).$$

Формула (6.1) верна при любом значении начальной популяции.

4. Докажем, что при $X_0 = 1$ верна формула

$$\varphi_{n+1}(s) = \varphi(\varphi_n(s)). \tag{6.2}$$

Доказательство. Проводим по индукции. Сначала проверим правильность формулы при некоторых частных значениях n:

При $n = 1$ по формуле (6.1) имеем

$$\varphi_2(s) = \varphi_1(\varphi(s)) = \varphi(\varphi_2(s)) = \varphi(\varphi_1(s)),$$

так как $\varphi(s) = \varphi_1(s)$ при $X_0 = 1$.

При $n = 2$ по формуле (6.1) и результату предыдущей строки получаем

$$\varphi_3(s) = \varphi_2(\varphi(s)) = \varphi_1(\varphi(\varphi(s))) = \varphi(\varphi_1(\varphi(s))) = \varphi(\varphi_2(s)).$$

Пусть теперь формула (6.2) верна до $n = k$, т.е. $\varphi_k(s) = \varphi(\varphi_{k-1}(s))$, тогда отсюда, используя формулу (6.1), имеем

$$\varphi_{k+1}(s) = \varphi_k(\varphi(s)) = \varphi(\varphi_{k-1}(\varphi(s))) = \varphi(\varphi_k(s)).$$

Таким образом, формула (6.2) установлена в случае $X_0 = 1$, что было использовано при ее выводе. Поэтому при других начальных условиях зависимость производящих функций численностей поколений ветвящегося процесса должна быть получена отдельно.

5. Докажем, что если X_0 — СВ с производящей функцией $\psi(s)$, то верна формула

$$\varphi_{n+1}(s) = \psi(\varphi(\psi^{-1}(\varphi_n(s)))). \tag{6.3}$$

Отсюда, в частности, при $X_0 = i_0$ $(\psi(s) = s^{i_0})$ будет следовать, что

$$\varphi_{n+1}(s) = (\varphi(\varphi_n(s))^{1/i_0})^{i_0}.$$

Доказательство. Для доказательства формулы (6.3) установим равенства

$$\varphi_0(s) = \psi(s);\ \ \varphi_1(s) = \varphi_0(\varphi(s)) = \psi(\varphi(s)).$$

Теперь с использованием их и формулы (6.1) получаем

$$\varphi_2(s) = \varphi_1(\varphi(s)) = \psi(\varphi(\varphi(s))).$$

По индукции докажем, что $\varphi_{n+1}(s) = \psi(\underbrace{\varphi(\varphi ... \varphi(s)...)}_{(n+1)\text{ раз}})$, предполагая, что $\varphi_n(s) = \psi(\underbrace{\varphi(\varphi ... \varphi(s)...)}_{n\text{ раз}})$. С учетом этого предположения и формулы (6.1) получаем

$$\varphi_{n+1}(s) = \varphi_n(\varphi(s)) = \psi(\varphi(\underbrace{\varphi(...\varphi(s)...)}_{n\text{ раз}})) = \psi(\underbrace{\varphi(\varphi ... \varphi(s)...)}_{(n+1)\text{ раз}}).$$

Тогда $\underbrace{\varphi(\varphi ... \varphi(s)...)}_{n\text{ раз}} = \psi^{-1}(\varphi_n(s))$, откуда и следует формула (6.3).

6. Математическое ожидание и дисперсия.

В общем случае задания начального состояния, рассмотренном в п. 5, получим выражения для математического ожидания и дисперсии численности n-го поколения, обозначенных соответственно как MX и DX.

В силу независимости развития частиц численность n-го поколения

$$X_n = \sum_{k=0}^{X_0} X_{n_k}, \qquad (6.4)$$

где X_{n_k} — численность n-го поколения, начавшегося с одной k-й частицы, и $\{X_{n_k}\}$ — независимые, одинаково распределенные СВ. Представление (6.4) позволяет легко найти моменты численности n-го поколения как суммы случайного числа независимых одинаково распределенных случайных слагаемых методом производящих функций.

Рассмотрим отдельно конструкцию $Z = \sum_{i=0}^{Y} X_i$, где $\{X_i\}$ — независимые целочисленные одинаково распределенные СВ; Y — целочисленная СВ.

Пусть $\varphi_X(s)$, $\varphi_Y(s)$ — производящие функции соответственно СВ $\{X_i\}$ и Y. Тогда

$$\varphi_z(s) = \sum_{m=0}^{\infty} P\{Z=m\}s^m = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} P\{Z=m/Y=k\}P\{Y=k\}s^m =$$

$$= \sum_{k=0}^{\infty} P\{Y=k\}\sum_{m=0}^{\infty} P\left\{\sum_{i=1}^{k} X_i = m\right\}s^m = \sum_{k=0}^{\infty} P\{Y=k\}(\varphi_x(s))^k = \varphi_Y(\varphi_X(s)).$$

Отсюда получаем

$$MZ = \varphi'_Z(1) = \varphi''_Y(\varphi_X(s))\varphi''_X(s)\Big|_{s=1} = \varphi'_Y(1)\varphi''_X(1) = MY \cdot MX; \qquad (6.5)$$

$$DZ = \varphi''_Z(1) + \varphi''_X(1) - \varphi(\varphi'_Z(1))^2;\ \varphi''_Z(s) = \varphi''_Y(\varphi_X(s))\varphi''_X(s);$$

$$\varphi''_Z(s) = \varphi''_Y(\varphi_X(s))(\varphi'_X(s))^2 + \varphi'_Y(\varphi_X(s))\varphi''_X(s);$$

$$\begin{cases} DX = \varphi''_X(1) + \varphi'_X(1) - (\varphi'_X(1))^2, \\ DY = \varphi''_Y(1) + \varphi'_Y(1) - (\varphi'_Y(1))^2 \end{cases} \Rightarrow \begin{cases} \varphi''_X(1) = DX - MX + (MX)^2, \\ \varphi''_Y(1) = DY - MY + (MY)^2; \end{cases}$$

$$\varphi''_Z(1) = \varphi''_Y(1)(MX)^2 + MY''_X(1) =$$

$$= (DY - MY + (MY)^2)(MX)^2 + MY(DX - MX + (MX)^2) =$$

$$= DY(MX)^2 + (MY)^2(MX)^2 + MY \cdot DX - MY \cdot MX;$$

$$DZ = \varphi''_Z(1) + MZ - (MZ)^2 = DY(MX)^2 + DX \cdot MY. \qquad (6.6)$$

Формулы (6.5) и (6.6) сводят задачу нахождения моментов численности n-го поколения, начавшегося со случайного числа X индивидуумов, к аналогичной задаче, когда $X_0 = 1$. Поэтому найдем MX_n и DX_n, когда $X_0 = 1$. Обозначим $MX_1 = m$; $DX_1 = \sigma^2$. Тогда

$$MX_{n+1}=\varphi'_{n+1}(1)=\varphi_n(\varphi(s))\big|_{s=1}=\varphi'_n(1)\varphi'(1)=MX_n\cdot m. \qquad (6.7)$$

При $n = 0$ $MX_1 = MX_0 = m$; при $n = 1$ $MX_2 = MX_1 \cdot m = m^2$; предположим, что $MX_n = m^n$. Тогда с учетом (6.7) $MX_{n+1} = MX_n \cdot m = m^{n+1}$, т.е. доказано, что при $X_0 = 1$ $MX_n = m^n$. Дисперсия

$$DX_{n+1}=\varphi''_{n+1}(1)+\varphi_{n+1}(1)-(\varphi'_{n+1}(1))^2;$$

обозначим $L = \varphi''(1)$; тогда

$$\varphi'_{n+1}(s)=\varphi'_n\varphi(s)\varphi'(s);\quad \varphi''_{n+1}(s)=\varphi''_n(\varphi(s))(\varphi'(s))^2+\ \varphi'_n(\varphi(s))\varphi''(s);$$

$$\varphi''_{n+1}(1)=\varphi''_n(1)(\varphi'(1))^2+\varphi'_n(1)\varphi''(1)=$$

$$=\varphi''(1)m^2+m^n(\sigma^2-m+m^2)=Lm^n+m^2\varphi''(1)=$$

$$=Lm^n+m^2(Lm^{n-1}+m^2\varphi''_{n-1}(1))=\ldots=L(m^n+m^{n+1}+\ldots+m^{2n});$$

$$DX_{n+1}=\sigma_n^2=\varphi^n_{n+1}(1)+\varphi'_{n+1}(1)-(\varphi'_{n+1}(1))^2=$$

$$=L(m^n+m^{n+1}+\ldots+m^{2n})+m^{n+1}-m^{2n+2}=$$

$$=\sigma^2(m^n+m^{n+1}+\ldots+m^{2n})=\begin{cases}\dfrac{\sigma^2m^n(m^{n+1}-1)}{m-1},\ \text{если}\ m\neq 1,\\ \sigma^2(n+1),\ \text{если}\ m=1.\end{cases}$$

Теперь по формулам (6.5) и (6.6), используя представление (6.4) для X_n, в случае, когда X_0 — СВ с заданной производящей функцией $\psi(s)$, можно получить MX_n и DX_n:

$$MX_n=MX_0MX_{n_k}=\psi'(1),\ \varphi'_n(1)=\psi'(1)m^n;$$

$$DX_n=DX_0MX_{n_k}+MX_0DX_{n_k}=$$

$$=\begin{cases}[\psi''(1)+\psi'(1)-(\psi'(1))^2]m^{2n}+\psi'(1)\sigma^2m^{n-1}(m^n-1)(m-1),\ \ m\neq 1,\\ [\psi''(1)+\psi'(1)-(\psi'(1))^2]+\psi'(1)\sigma^2m,\ \ m=1.\end{cases}$$

Вероятность вырождения

Пусть X_0 — СВ с производящей функцией $\psi(s) = \varphi_0(s)$; $g_n = P\{X_n = 0\}$ — вероятность того, что вырождение (т.е. первое появление поколения нулевой численности) произойдет не позже n-го поколения. Тогда $g_n = \varphi_n(0)$ и по формуле (6.3)

$$g_{n+1}=\varphi_{n+1}(0)=\psi(\varphi(\psi^{-1}(\varphi_n(0))))=\psi(\varphi(\psi^{-1}(g_n))). \qquad (6.8)$$

Последовательность $\{g_n\}$ при $n \to \infty$ имеет предел как неубывающая (это последовательность вероятностей вложенных событий) и ограниченная единицей. Переходя к пределу в формуле (6.8), получим уравнение для вероятности вырождения

$$h = \psi(\varphi(\psi^{-1}(h))), \tag{6.9}$$

где $h = \lim\limits_{n\to\infty} g_n$.

При $X_0 = i_0$ уравнение для вероятности вырождения примет вид

$$h = \left(\varphi\left(h^{\frac{1}{i_0}}\right)\right)^{i_0}. \tag{6.10}$$

Осталось решить вопрос о том, какой из корней уравнения (6.9) является вероятностью вырождения. Оказывается, что это наименьший положительный корень данного уравнения. Докажем это.

Естественно считать, что $g_0 = P\{X_0 = 0\} = 0$.

Пусть s_0 — любой положительный корень уравнения (6.9). Положительный корень уравнения (6.9) всегда есть, так как непосредственной подстановкой убеждаемся, что $h = 1$ — всегда корень уравнения (6.9). Утверждение будет доказано, если для любого положительного корня s уравнения (6.9) будет установлено, что $g_n \leqslant s_0$.

Из формулы (6.8) имеем $g_{n+1} = \psi(\varphi(\psi^{-1}(g_n)))$, а так как последовательность $\{g_n\}$ — неубывающая, то и функция $\psi(\varphi(\psi^{-1}(g_n)))$ — тоже неубывающая, тогда

$$g_1 = \psi(\varphi(\psi^{-1}(g_0))) \leqslant \psi(\varphi(\psi^{-1}(s_0))) = s_0.$$

Предположив, что $g_n \leqslant s_0$, докажем по индукции, что и $g_{n+1} \leqslant s_0$:

$$g_{n+1} = \psi(\varphi(\psi^{-1}(g_n))) \leqslant \psi(\varphi(\psi^{-1}(s_0))) = s_0, \text{ ч.т.д.}$$

Приведем (без доказательства) следующую важную теорему.

Теорема 6.1 (о вырождении). *Если $m \geqslant 1$, то вероятность того, что процесс оборвется до n-го поколения, $h \xrightarrow[n\to\infty]{} 1$ (т.е. вероятность вырождения $h = 1$ при $m \leqslant 1$). Если $m > 1$, то существует единственный корень, меньший единицы, уравнения $\varphi(s) = s$.*

Замечание. $1 - h$ — вероятность бесконечного продолжения процесса.

Определение 6.2. Ветвящиеся процессы называются ***докритическими***, если $m = 1$, $\varphi^n(1) > 0$ (последнее условие означает допредельную невырожденность процесса) и ***надкритическими***, если $m > 1$.

Задачи

Приведем задачи на анализ конкретных ситуаций, ограничиваясь случаем $X_0 = 1$, что позволит без громоздких выкладок проиллюстрировать суть приведенных теоретических результатов.

Переход к более общему случаю задания закона распределения размера начальной популяции после всего сказанного не должен вызывать принципиальных трудностей.

Задача 6.1. Дано: $X_0 = 1$;

X	0	2
P	0,5	0,5

Найти первые три строки матрицы $P = P(1)$, первые две строки матрицы $P(2)$, MX_n, DX_n и вероятность вырождения h из уравнения вырождения.

Решение

$$\varphi(s) = 0{,}5(1 + s^2);\ \varphi_2(s) = \varphi(\varphi(s)) = 0{,}5(1 + 0{,}25(1 + s^2)^2;$$

$$m = \varphi'(1) = 1;\ MX_n = m^n = 1;\ \sigma_n^2 = DX_n = n\sigma^2,$$

где $\sigma^2 = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2 = 1 + 1 - 1 = 1 \Rightarrow DX_n = n$.

Вероятность вырождения находится из уравнения $s = \varphi(s)$ (см. формулу (6.10) при $i_0 = 1$): $0{,}5(1 + s^2) = s \Rightarrow s_{1,2} = 1 \Rightarrow h = 1$ (тот же результат для h имеем из теоремы 6.1);

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0\ldots \\ 0{,}5 & 0 & 0{,}5 & 0 & 0\ldots \\ 0{,}25 & 0 & 0 & 0{,}25 & 0\ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix};$$

$$P(2) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0\ldots \\ 0{,}625 & 0 & 0{,}25 & 0 & 0{,}125 & 0\ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

Задача 6.2. Дано: $X_0 = 1$; $b = 1 - a$;

$X_1 = \varphi$	0	k
P	a	b

Найти $\varphi(s)$, $\varphi_2(s)$, MX_n, DX_n, уравнение вырождения и вероятность вырождения h.

Решение

$$\varphi(s) = a + bs^k;\ \varphi_2(s) = \varphi(\varphi(s)) = a + b(a + bs^k)^k;$$

$$m = \varphi'(1) = kb;\ \varphi''(1) = k(k-1)b;\ \ \sigma^2 = \varphi''(1) + \varphi'(1) - (\varphi'(1))^2 = k^2b(1 - b).$$

а) $m = kb = 1$; $MX_n = m^n = 1$; $\sigma^2 = k - 1$; $DX = n(k - 1)$;
уравнение вырождения $bs^k - s + a = 0 \Rightarrow h = 1$.

б) $m = kb > 1$; $MX = (kb)^n \underset{n\to\infty}{\longrightarrow} \infty$; $\sigma^2 = k^2 b(1-b)$;

$$DX = \frac{\sigma^2 m^n (m^n - 1)}{m(m-1)} = \frac{k^2 b(1-b)(kb)^n (kb-1)}{kb(kb-1)} \underset{n\to\infty}{\longrightarrow} \infty;$$

уравнение вырождения $bs^k - s + a = 0$.

Найдем, в частности, вероятность вырождения при $k = 2$. В этом случае

$$s^k - \frac{s}{b} + \frac{a}{b} = (s-1)\left(s - \frac{a}{b}\right) = 0 \Rightarrow h = \min\left(1; \frac{a}{b}\right).$$

в) $m = kb < 1$; $MX_n = (kb)^n \underset{n\to\infty}{\longrightarrow} 0$; $DX_n \underset{n\to\infty}{\longrightarrow} 0$; $h = 1$ (по теореме 6.1).

Задача 6.3. Дано: $X_0 = 1$; $X_1 = V$ распределено по Пуассону с параметром λ.

Найти $\varphi(s)$, $\varphi_2(s)$, MX_n, DX_n, уравнение и вероятность вырождения h.

Решение

$$P\{X_1 = k\} = \frac{\exp(-\lambda)L^k}{k!};\ \varphi(s) = \exp[-L(1-s)];$$

$$\varphi_2(s) = \exp\{-\lambda[1 - \exp[-1(1-s)]]\};\ m = \varphi'(1) = \{\exp[-\lambda(1-s)]\}' = \lambda.$$

а) $\lambda = 1$; $MX_n = 1$; $DX_n = nL$; $h = 1$ (по теореме 6.1).

б) $\lambda \neq 1$; $MX_n = \lambda^n$; $DX_n = \dfrac{\lambda^2\lambda^n(\lambda^n - 1)}{\lambda(\lambda - 1)}$; при $\lambda > 1$ $MX_n \underset{n\to\infty}{\longrightarrow} \infty$; $DX_n \underset{n\to\infty}{\longrightarrow} \infty$; $h < 1$. Так как h есть наименьший положительный корень уравнения вырождения $\varphi(s) = s$, она находится из уравнения $s = \exp[-\lambda(1-s)]$, т.е. приходится решать (трансцендентное) уравнение вырождения (проще искать корень графически).

В случае $L < 1$ $h = 1$.

Задача 6.4. Дано: $\varphi(s) = as^2 + bs + c$, $a, b, c > 0$, $a + b + c = 1$, $h < 1$. Доказать, что $h = \dfrac{c}{a}$.

Решение (*доказательство*)

$\varphi(1) = m > 1$, или $2a + b > 1$, или $a + b > 1 - a \Rightarrow 1 - a < 1 - c \Rightarrow \dfrac{c}{a} < 1$.

Уравнение вырождения: $s = as^2 + bs + c$, или $as^2 + (b-1)s + c = (s-1)(as + c) = 0 \Rightarrow$ (так как доказано, что $\dfrac{c}{a} < 1$) $h = \dfrac{c}{a}$.

Задача 6.5. Доказать, что при $X_0 = 1$ и $m < 1$ $M\sum\limits_{k=1}^{\infty} X_k = \dfrac{m}{1-m}$.

Решение (*доказательство*)

$$MX_n = m^n;\ M\sum_{k=1}^{\infty} X_k = M\sum_{k=1}^{\infty} m^k = \frac{m}{1-m}.$$

Задача 6.6. Дано: V — бернуллиевская СВ и $X_0 = 1$. Найти: а) $\varphi_n(s)$; б) вероятность вырождения h и распределение времени до вырождения; в) производящую функцию общего числа R частиц в первых n поколениях.

Решение

а) $\Phi_1(s) = \varphi_n(s)$;

$$\varphi_n(s) = \varphi_{n-1}(\varphi(s)) = \varphi_{n-2}(\varphi(\varphi(s))) = \underbrace{\varphi(\varphi(...\varphi(s)))}_{n \text{ раз}};$$

$$\varphi_2(s) = \varphi(\varphi(s)) = p^2 s - p^2 + 1; \quad \varphi_3(s) = \varphi_2(\varphi(s)) = p^3 s - p^3 + 1;$$

предположим по индукции, что $\varphi_{n-1}(s) = p^{n-1}s - p^{n-1} + 1$, тогда

$$\varphi_n(s) = \varphi_{n-1}(\varphi(s)) = p^{n-1}(ps - p + 1) - p^{n-1} + 1 = p^n s - p^n + 1.$$

б) Выпишем уравнение для вероятности выражения: $ph + 1 - p = h \Rightarrow$ $\Rightarrow h = 1$ — вероятность вырождения; $t = \min\{n: X_n = 0\}$, V распределена по $B(1, \text{p})$, отсюда следует, что $P\{t = n\} = p(1 - p)^n$.

в) $P\{R = k\} = p^{n-1}q$ при $k \leqslant n$; $P\{R = n + 1\} = p^n$, где $q = 1 - p$, поэтому искомая производящая функция

$$\varphi_R(s) = \sum_{k=0}^{n+1} P\{R = k\} s^k = qs[(ps)^0 + (ps)^1 + (ps)^2 + ... + (ps)^{n-1}] + p^n s^{n+1} =$$

$$= \frac{qs(1 - p^n s^n)}{1 - ps} + p^n s^{n+1}.$$

6.2. Пуассоновские процессы (потоки)

Рассмотрим случайный процесс $\xi(t)$ с непрерывным временем $t \in T = (0, \infty)$ с дискретным множеством состояний S. При фиксированном t $\xi(t)$ — значение процесса — число наступлений некоторого события за время t (от нулевого момента).

6.2.1. Определения пуассоновского процесса

Дадим два определения пуассоновског процесса.

Определение 6.3. ***Пуассоновский процесс*** — это целочисленный процесс с непрерывным временем, обладающий следующими свойствами:

1) числа наступлений события в непересекающихся интервалах времени независимы (т.е. *отсутствие последствия или независимость приращений*);

2) СВ $\xi(t + t_0) - \xi(t)$ зависит от t и не зависит от t_0 (т.е. *однородность или стационарность*);

3) если $p(\Delta t)$, где Δt — малый промежуток времени, — вероятность того, что за время Δt произойдет хотя бы одно событие, то:

а) $p(\Delta t) = \lambda\Delta t + o(\Delta t)$ при $\Delta t \to 0$, где λ — const > 0;

б) $P(\xi(t + \Delta t) - \xi(t) \geqslant 2) = o(\Delta t)$ (т.е. *ординарность*).

Ординарность потока событий означает, что число событий, происходящих за малое время, пропорционально этому времени, а вероятность появления за малое время более одного события есть малое более высокого порядка, чем это время.

Определение 6.4. ***Пуассоновский процесс*** — это целочисленный процесс с непрерывным временем, обладающий следующими свойствами:

1) независимость приращений;

2) однородность;

3) $P(\xi(t) = m) = \dfrac{e^{-\lambda t}(\lambda t)^m}{m!}$.

Покажем **эквивалентность** определений 6.3 и 6.4 пуассоновского процесса.

1. Пусть выполнено определение 6.3. Тогда $n - 1$ точками разобьем интервал $(0; t)$ на n равных частей $\Delta_1, ..., \Delta_n$ так, чтобы Δ_i были малы (n — велико). $\xi(t) = \sum\limits_{i=1}^{n} \xi(\Delta_i)$, где $\xi(\Delta_i)$ независимы, так как Δ_i не пересекаются, $\Delta_i = \dfrac{t}{n}$. Найдем $\varphi(s)$ — производящую функцию СВ $\xi(t)$ при фиксированном значении t.

По свойствам 1—3 определения 6.3 $\varphi(s) = (\varphi_n(s))^n$, где

$$\varphi_n(s) = \left(1 - \lambda\frac{t}{n}\right)s^0 + \lambda\frac{t}{n}s + o\left(\frac{t}{n}\right) = 1 + \frac{\lambda t}{n}(s-1) + o\left(\frac{t}{n}\right),$$

следовательно,

$$\varphi(s) = \left(1 + \frac{\lambda t}{n}(s-1) + o\left(\frac{t}{n}\right)\right)^n \underset{n\to\infty}{\longrightarrow} e^{\lambda t(s-1)}$$

— это производящая функция пуассоновского замена с параметром λt, что и доказывает выполнение определения 6.4.

Найдем $\varphi(t)$ и $g(t)$ — характеристическую функцию для пуассоновского закона с параметрами λt:

$$\varphi(s) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k e^{-\lambda t}}{k!} s^k = e^{-\lambda t}\sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} s^k = e^{-\lambda t} e^{\lambda t s} = e^{\lambda t(s-1)};$$

$$g(t) = \varphi(e^{it}) = e^{\lambda t(e^{it}-1)}.$$

Пусть выполнено определение 6.4. Тогда для выполнения условий определения 6.3 достаточно показать, что имеет место ординарность потока. Используя построенные выше Δ_i, имеем (по свойству 3 определения 6.4):

$$P(\xi(t_i) - \xi(t_{i-1}) = 0) = P(\xi(\Delta_i) = 0) = \frac{e^{-\lambda\Delta_i}(\lambda\Delta_i)^0}{0!} = e^{-\lambda\Delta_i} \approx$$

$$\approx 1 - \lambda(\Delta_i) + o(\Delta_i);$$

$$P(\xi(\Delta_i)=1)=\frac{e^{\lambda\Delta_i}\lambda\Delta_i}{1!}=\lambda\Delta_i e^{-\lambda\Delta_i}\approx\lambda\Delta_i+o(\Delta_i);$$
$$P(\xi(\Delta_i)\geqslant 2)=1-P(\xi(\Delta_i)=0)-P(\xi(\Delta_i)=1)=o(\Delta_i);$$
$$P(\xi(\Delta_i)\geqslant 1)=1-P(\xi(\Delta_i)=0)=\lambda\Delta_i+o(\Delta_i),$$

что означает ординарность потока событий.

6.2.2. Свойства пуассоновских процессов

1. Докажем следующее утверждение.

Утверждение 6.1. Время между двумя соседними событиями в пуассоновском потоке имеет экспоненциальное распределение.

Доказательство. Пусть T_i — время нахождения пуассоновского процесса в состоянии E_i, где E_i — i событий за время $(0; t)$ в потоке. Тогда

$$G(t)=P(T_i<t)=P(\xi(t)\geqslant 1)=1-P(\xi(t)=0)=$$
$$=\begin{cases}1-\dfrac{(\lambda t)^0 e^{\lambda t}}{0!}, \ t\geqslant 0,\\ 0, \ t<0\end{cases} = \begin{cases}1-e^{\lambda t}, \ t\geqslant 0,\\ 0, \ t<0,\end{cases}$$

а это функция распределения экспоненциального закона, ч.т.д.

Замечание 6.1. Из того, что $T_i \sim E(\lambda)$, следует, что $MT_i = \dfrac{1}{\lambda}$ — среднее время пребывания в каждом состоянии пуассоновского процесса, а $\omega = 1/MT_i = \lambda$ — среднее число событий потока в единицу времени.

При $\lambda = 0$ процесс навсегда остается в данном состоянии (это состояние — поглощающее), при $\lambda = \infty$ процесс мгновенно уходит из данного состояния. Поэтому для пуассоновского процесса считаем, что $0 < \lambda < \infty$.

2. Прошлое пуассоновского процесса при условии настоящего. Дано $u < t$, $k < n$, $\xi(t)$ — пуассоновский процесс. Тогда

$$P(\xi(u)=k/\xi(t)=n)=\frac{P(\xi(u)=k,\xi(t-u)=n-k)}{P(\xi(t)=n)}=$$
$$=\frac{\dfrac{e^{-\lambda u}(\lambda u)^k}{k!}\,\dfrac{e^{-\lambda(t-u)}(\lambda(t-u))^{n-k}}{(n-k)!}}{\dfrac{e^{-\lambda t}(\lambda t)^n}{n!}}=C_n^k\left(\frac{u}{t}\right)^k\left(1-\frac{u}{t}\right)^{n-k},$$

а это биномиальное распределение $B(n, p = u/t)$.

3. Проверка: является ли процесс пуассоновским?

Задача 6.7. Дано $\xi_1 = \xi_1(t)$ и $\xi_2 = \xi_2(t)$ — два независимых пуассоновских процесса с параметрами λ_1 и λ_2 соответственно. Определить, является ли процесс $\eta = \eta(t) = \xi_1(t) + \xi_2(t)$ пуассоновским.

Решение

Проверим выполнение условий, например, определения 11.2 пуассоновского процесса. Имеем

$$\varphi_\eta(s) = \varphi_{\xi_1}(s)\varphi_{\xi_2}(s) = e^{\lambda_1 t(s-1)}e^{\lambda_2 t(s-1)} = e^{(\lambda_1+\lambda_2)t(s-1)},$$

что означает выполнение условия 3.

Проверим постулат о независимости приращений. Возьмем три близких момента времени t_1, t_2, t_3. Тогда $\xi(\Delta_1) = \xi(t_2) - \xi(t_1)$ — приращение процесса на интервале Δ_1, а $\xi(\Delta_2) = \xi(t_3) - \xi(t_2)$ — приращение процесса на интервале Δ_2.

По построению интервалы Δ_1 и Δ_2 не пересекаются. Покажем, что

$$P(\xi(\Delta_1) = l), \xi(\Delta_2) = k) = P(\xi(\Delta_1) = l)P(\xi(\Delta_2) = k).$$

Имеем

$$\xi(\Delta_1) = \xi_1(\Delta_1) + \xi_2(\Delta_1);\ \xi(\Delta_2) = \xi_1(\Delta_2) + \xi_2(\Delta_2) \Rightarrow$$

$$\Rightarrow P(\xi(\Delta_1) = l, \xi(\Delta_2) = k) = P(\xi_1(\Delta_1) + \xi_2(\Delta_1) = l, \xi_1(\Delta_2) + \xi_2(\Delta_2) = k) =$$

$$= \sum_{k_1+k_2=k}\sum_{l_1+l_2=l} P(\xi_1(\Delta_1) = l_1, \xi_2(\Delta_1) = l_2, \xi_1(\Delta_2) = k_1, \xi_2(\Delta_2) = k_2) \underset{\xi_1 \text{ и } \xi_2 - \text{НЗ}}{=}$$

$$= \sum_{\substack{k=k_1+k_2\\ l=l_1+l_2}} P(\xi_1(\Delta_1) = l_1, \xi_1(\Delta_2) = k_1)P(\xi_2(\Delta_1) = l_2, \xi_2(\Delta_2) = k_2) \underset{\text{нз приращений}}{=}$$

$$= \sum_{\substack{k_1+k_2=k\\ l_1+l_2=l}} P(\xi_1(\Delta_1) = l_1)P(\xi_1(\Delta_2) = k_1)P(\xi_2(\Delta_1) = l_1)P(\xi_2(\Delta_2) = l_2) =$$

$$= \sum_{\substack{k_1+k_2=k\\ l_1+l_2=l}} P(\xi_1(\Delta_1) = l_1, \xi_2(\Delta_1) = l_2)P(\xi_1(\Delta_2) = k_1, \xi_2(\Delta_2) = k_2) =$$

$$= P(\xi(\Delta_1) = l)P(\xi(\Delta_2) = k), \text{что и проверялось.}$$

Стационарность процесса $\xi(s)$ очевидна, так как он имеет распределение Пуассона с параметром λt и определяется только длиной интервала $(0; t)$, а не его расположением на оси времени.

Задача 6.8. Являются ли пуассоновскими: а) разность двух пуассоновских процессов $\xi_1(t)$ и $\xi_2(t)$ с параметрами λ_1 и λ_2; б) сумма пуассоновского процесса с константой $\xi_1(t) + k$?

Решение

а) Найдем производящую функцию СВ $\xi(t) = \xi_1(t) - \xi_2(t)$:

$$\varphi_{\xi_1(t)-\xi_2(t)}(s) = \varphi_{\xi_1(t)}(s)\varphi_{\xi_2(t)}(-s) = e^{\lambda_1 t(s-1)}e^{\lambda_2 t(-s-1)} = e^{(\lambda_2-\lambda_1)t(s-1)} = e^{2\lambda_2 ts},$$

а это не производящая функция пуассоновского закона $\Rightarrow \xi(t)$ — не пуассоновский процесс.

б) $\xi(t) = \xi_1(t) + k \Rightarrow \varphi_{\xi(t)}(s) = \varphi_{\xi_1(t)+k}(s) = s^k e^{\lambda_1 t(s-1)}$, а это не производящая функция пуассоновского закона $\Rightarrow \xi(t)$ — не пуассоновский процесс.

4. Прореживание пуассоновского потока.

Задача 6.9. Пусть $\xi(t)$ — пуассоновский поток с параметром λt; с вероятностью p событие оставляем в потоке и с вероятностью $q = 1 - p$ — выбрасываем. Найти для результирующего потока $\eta(t)$ $P(\eta(t) = m)$.

Решение

$$P(\eta(t)=m)=\sum_{k=m}^{\infty}P(\eta(t)=m\,/\,\xi(t)=k)P(\xi(t)=k)=$$

$$=\sum_{k=m}^{\infty}C_k^m p^m(1-p)^{k-m}\,\frac{(\lambda t)^k e^{-\lambda t}}{k!}=\frac{p^m e^{-\lambda t}}{m!}\sum_{k=m}^{\infty}\frac{k!}{(k-m)!}(1-p)^{k-m}\,\frac{(\lambda t)^k}{k!}=$$

$$=\{\text{замена индекса суммирования: } k-m=1;\, k=m+1;\, m\leqslant k\leqslant\infty;\, 0\leqslant 1\leqslant\infty\}=$$

$$=\frac{(\lambda tp)^m e^{-\lambda t}}{m!}\sum_{l=0}^{\infty}\frac{[\lambda t(1-p)]^l}{l!}=\frac{(\lambda tp)^m}{m!}e^{-\lambda t}e^{\lambda t(1-p)}=\frac{(\lambda tp)^m}{m!}e^{-\lambda tp}\Rightarrow$$

$\Rightarrow \eta(t)$ — пуассоновский поток с параметром λtp (исходный пуассоновский поток $\xi(t)$ имел параметр λt).

Задача 6.10. Пусть $\xi_1(t)$ и $\xi_2(t)$ — независимые пуассоновские процессы с параметрами соответственно λ_1 и λ_2. Найти $P(X) = P(\xi_1(t) = k\,/\,\xi_1(t) + \xi_2(t) = n)$.

Решение

$$P(X)=\frac{P(\xi_1(t)=k,\xi_1(t)+\xi_2(t)=n)}{P(\xi_1(t)+\xi_2(t)=n)}=\frac{P(\xi_1(t)=k,\,\xi_2(t)=n-k)}{P(\xi_1(t)+\xi_2(t)=n)}=$$

$$\overset{\text{нз}}{=}\frac{P(\xi_1(t)=k)P(\xi_2(t)=n-k)}{P(\xi_1(t)+\xi_2(t)=n)}=\frac{(\lambda_1 t)^k e^{-\lambda_1 t}}{k!}\,\frac{(\lambda_2 t)^k e^{-\lambda_2 t}}{(n-k)!}\,\frac{n!}{(\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)t}}=$$

$$=C_n^k\,\frac{\lambda_1^k\lambda_2^{n-k}}{(\lambda_1+\lambda_2)^n}=C_n^k\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(1-\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-k},$$

а это биномиальная вероятность $B\left(n, p=\dfrac{\lambda_1}{\lambda_1+\lambda_2}\right)$ (при $\lambda_1 = \lambda_2$ имеем СВ $X \sim B(n, 0{,}5)$).

5. Характеризационное свойство пуассоновского потока — нестарение.

Утверждение 6.2. Если промежуток времени T между событиями пуассоновского потока, распределенный по закону $E(\lambda)$, уже длится некоторое время τ, то распределение оставшегося промежутка $(T - \tau)$ имеет такое же распределение, как и СВ T.

Доказательство

$$P\{T-\tau<|t|\,/\,T>\tau\}=\frac{P\{T-\tau<|t|,T>\tau\}}{P\{T>\tau\}}=\frac{P(\tau<T<\tau+t)}{1-F(\tau)}=$$

$$=\left|F(t)=\begin{cases}0,\, t<0,\\ 1-e^{-\lambda t},\, t\geqslant 0\end{cases}\right|=\frac{F(\tau+t)-F(\tau)}{1-F(\tau)}=\frac{e^{-\lambda\tau}-e^{-\lambda(\tau+t)}}{e^{-\lambda\tau}}=1-e^{-\lambda t},\ t\geqslant 0,$$

т.е. тот же закон распределения, что и для T.

6. Отсутствие последействия.

Дано: СВ $T \sim E(\lambda)$, т.е. $F_T(t) = \begin{cases} 0, t < 0, \\ 1 - e^{-\lambda t}, t \geqslant 0. \end{cases}$ Тогда

$$P(T > x + y / T > y) = \frac{P(T > x + y,\ T > y)}{P(T > y)} = \frac{P(T > x + y)}{P(T > y)} =$$

$$= \frac{1 - F_T(x + y)}{1 - F_T(y)} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} = e^{-\lambda x} = P(T > x);$$

сравнивая соотношения, получаем, что:

1) $P(T > x + y / T > y) = P(T > x)$;

2) $P(T > x + y) = P(T > y)P(T > x)$.

Эти соотношения имеют смысл отсутствия последствия при экспоненциальном распределении промежутков между событиями потока.

7. Настоящее (в пуассоновском процессе) при условии прошлого. Пусть $u < t$, $k < n$, тогда

$$\Pi(\lambda(t-u))P(\xi(t) = n / \xi(u) = k) = \frac{P(\xi(t) = n,\ \xi(u) = k)}{P(\xi(u) = k)} =$$

$$= \frac{P(\xi(t-u) = n-k, \xi(u) = k)}{P(\xi(u) = k)} \overset{\text{нз прира-}\atop\text{щений}}{=} \frac{P(\xi(t-u) = n-k)P(\xi(u) = k)}{P(\xi(u) = k)} =$$

$$= P(\xi(t-u) = n-k) = \frac{e^{-\lambda(t-u)}(\lambda(t-u))^{n-k}}{(n-k)!},$$

т.е. снова получаем пуассоновское распределение.

Задача 6.11. На телефонную станцию поступает пуассоновский поток вызовов. Среднее число вызовов за час равно 30. Найти $P(X)$ — вероятность того, что за минуту поступит не менее двух вызовов.

Решение

$$\lambda = 30; t = \frac{1}{60};\ P(\xi(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!};$$

$$P(X) = 1 - P(\xi(t) = 0) - P(\xi(t) = 1) = 1 - e^{-\frac{1}{2}} - \frac{1}{2}e^{-\frac{1}{2}} = 1 - \frac{3}{2}e^{-\frac{1}{2}} \approx 0{,}09.$$

Задача 6.12. Среднее число телеграмм в час равно трем. а) Какова вероятность того, что за 4 ч подряд не поступит ни одной телеграммы? б) Каково распределение момента T поступления первой телеграммы?

Решение

а) $\lambda = 3$; $t = 4$; $\lambda t = 12$; $P\{\xi(t) = 0\} = e^{-\lambda t} = e^{-12}$;

б) известно, что $T = \sim E(\lambda) \Rightarrow$

$$\Rightarrow f_T(t) = \begin{cases} 3e^{-3t}, t \geqslant 0, \\ 0, t < 0. \end{cases}$$

Контрольные вопросы и задания

1. Дайте определение ветвящегося процесса.
2. Приведите доказательство марковости ветвящегося процесса.
3. Выпишите матрицу переходных вероятностей ветвящегося процесса (в общем виде).
4. Чему равно вероятностное распределение популяции на n-м шаге при заданной численности первого поколения?
5. Чему равны моменты численности n-го поколения?
6. Выведите уравнение вероятности вырождения.
7. Докажите теорему о вырождении процесса.
8. Приведите примеры исследования ветвящихся процессов.
9. Докажите эквивалентность двух определений пуассоновского процесса.
10. Каково место пуассоновских процессов в классификации случайных процессов?
11. Каково распределение времени ожидания между событиями пуассоновского потока?
12. Каково прошлое пуассоновского процесса при фиксированном настоящем?
13. Приведите примеры проверки, является ли процесс пуассоновским.
14. Что представляет собой прореживание пуассоновского процесса?
15. Опишите характеризационное свойство пуассоновского потока (нестарение).

Глава 7
ОДНОРОДНЫЕ ЦЕПИ МАРКОВА С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ И КОНЕЧНЫМ МНОЖЕСТВОМ СОСТОЯНИЙ

В результате изучения данной главы студенты должны:

знать

• основные понятия теории массового обслуживания и анализа различных систем массового обслуживания (СМО);

уметь

• проводить расчеты характеристик СМО и стационарного и достационарного распределений ее состояний;

владеть

• методами анализа СМО.

Для однородных цепей Маркова с непрерывным временем и конечным множеством состояний в главе вводятся основные понятия, такие как интенсивность потока событий, матрица переходных вероятностей для однородных потоков событий, системы массового обслуживания (СМО), системы уравнений Колмогорова для вероятностей ее состояний в достационарном и стационарном режимах. Приведено много примеров расчетов для различных СМО.

7.1. Основные понятия

Поток событий — это последовательность событий, наступающих в случайные моменты времени. *Пуассоновский поток событий* — это стационарный ординарный поток без последствия (см. гл. 6).

Будем рассматривать системы — случайные процессы $\xi(t)$, которые с течением времени t могут менять свои состояния, число которых конечно.

Будем предполагать, что если в момент времени s система находится в состоянии E_i, $i = 1, ..., n$, то независимо от своего поведения до этого момента s она через время t с вероятностью $p_{ij}(t)$ переходит в состояние E_j, т.е.

$$p_{ij}(t) = P\{\xi(s+t) = E_j / \xi(s) = E_i\}, \tag{7.1}$$

где $\xi(t)$ — состояние процесса в момент времени t. Тогда все потоки событий, переводящие систему из одного состояния в другое, являются пуассоновскими. Эта модель является однородным марковским процессом, вероятности $p_{ij}(t)$ — переходные вероятности процесса, а $\{P_i^{(0)}\} = \{P(\xi(0) = E_i)\}$, $i = 1, 2, ...$ — его начальные вероятности.

Соотношение (7.1) называется еще стационарностью переходных вероятностей.

Интенсивность потока событий

Определение 7.1. *Плотностью*, или ***интенсивностью***, потока (процесса) называется среднее число событий в единицу времени.

В частности, пуассоновский (или простейший) поток стационарен и все его характеристики не зависят от времени, т.е. его интенсивность постоянна. Найдем интенсивность $\lambda(t)$ пуассоновского потока $\xi(t)$, т.е. $\dfrac{MX}{t}$, где X — число событий за время t:

$$\lambda(t) = \frac{MX}{t} = \frac{1}{t}\sum_{k=1}^{\infty} k\frac{(\lambda k)^k}{k!}e^{-\lambda t} = (e^{-\lambda t}/t)\sum_{k=1}^{\infty}\frac{(\lambda k)^k}{(k-1)!} = (e^{-\lambda t}/t)\lambda t e^{\lambda t} = \lambda.$$

Интенсивность пуассоновского потока можно найти по-другому. Известно (см. параграф 6.2), что интервалы времени Y между событиями пуассоновского потока распределены по экспоненциальному закону с математическим ожиданием $MY = 1/\lambda$, откуда следует, что $1/MY$ — среднее число событий в единицу времени, тогда это и есть интенсивность (плотность) потока $\lambda(t) = 1/MY = \lambda$.

Если события образуют пуассоновский поток, то число X событий, попадающих на любой участок времени $(t_0, t_0 + \tau)$, распределено по закону Пуассона

$$p_m = P(X = m) = \frac{\lambda^m}{m!}e^{-\lambda},$$

где λ — математическое ожидание числа событий на интервале $(t_0, t_0 + \tau)$: $\lambda = \int_{t_0}^{t_0+\tau}\lambda(t)dt$, $\lambda(t)$ — плотность потока событий. Если $\lambda(t) = \text{const}$, то число событий на участке длиной τ есть $\lambda\tau$.

Матрица переходных вероятностей $P(t)$ однородного потока событий

Составим матрицу переходных вероятностей вида $P(t) = (p_{ij}(t))$, где $p_{ij}(t)$ определены по формуле (7.1).

Свойства матрицы $P(t)$:

1) $p_{ij}(t) \geqslant 0$;

2) $\sum p_{ij}(t) = 1$;

3) $p_{ij}(s + t) = \sum_{k=1}^{n} p_{ik}(s) p_{kj}(t)$, или в матричной форме $P(s + t) = P(s)P(t)$ — это уравнение Колмогорова — Чепмена (мультикативность);

4) непрерывность в нуле:

$$\lim_{t\to 0} p_{ij}(t) = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases} \to P(0) = E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix};$$

5) дифференцируемость в нуле:

$$P'(0) = \lim_{\Delta t\to 0} \frac{P(\Delta t) - P(0)}{\Delta t} = \lim_{\Delta t\to 0} \frac{P(t) - E}{\Delta t} = \Lambda = \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \dots & \dots & \dots \\ \lambda_{n1} & \dots & \lambda_{nn} \end{pmatrix},$$

где Λ — инфинитезимальная матрица, т.е. матрица, у которой $\sum_{j=1}^{n} \lambda_{ij} = 0$, — это следует из стохастичности ($\sum_j p_{ij}(t) = 1$) матрицы $P(t)$. Найдем, в частности, вид матрицы Λ пуассоновского потока:

$$P'(0) = \lim_{\Delta t\to 0} \frac{P(\Delta t) - E}{\Delta t} =$$

$$= \lim_{\Delta t\to 0} \frac{1}{\Delta t} \begin{pmatrix} 1-\lambda\Delta t - 1 & \lambda\Delta t & \dots & \dots \\ \dots & 1-\lambda\Delta t - 1 & \lambda\Delta t & \dots \\ \dots & \dots & \ddots & \dots \\ \dots & \dots & 1-\lambda\Delta t - 1 & \lambda\Delta t \\ \dots & \dots & \dots & 1-\lambda\Delta t - 1 \end{pmatrix} =$$

$$= \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -\lambda & \lambda & 0 & \vdots \\ \vdots & \dots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\lambda & \lambda \\ 0 & \dots & 0 & 0 & -\lambda \end{pmatrix}.$$

Любой недиагональный элемент инфинитезимальной матрицы $\Lambda(p_{ij}(t))$ при $i \neq j$ (интенсивность или точность потока, переводящего систему из состояния E_i в состояние E_j) есть $\lambda_{ij,i\neq j} = \dfrac{p_{ij}(t)}{\Delta t}$.

Системы массового обслуживания

Часто приходится иметь дел с системами, предназначенными для многоразового использования при решении однотипных задач. Возникающие при этом процессы получили название процессов обслуживания, а системы — систем массового обслуживания (СМО). Главная их особенность — наличие случайности.

Определение 7.2. ***Система массового обслуживания*** — совокупность пунктов (каналов, станций, приборов), на которые в случайные моменты времени поступают заявки на обслуживание, подлежащие удовлетворению. При этом время их удовлетворения также случайно. Примерами таких систем служат телефонная сеть, магазин, парикмахерская и т.д.

В СМО обслуживаемый объект называют ***требованием***. В общем случае под требованием обычно понимают запрос на удовлетворение некоторой потребности, например разговор с абонентом, покупку билета, получение материалов. Средства, обслуживающие требования, называются ***каналами обслуживания***. Например, к ним относятся каналы телефонной связи, посадочные полосы, билетные кассиры. Предметом теории СМО является установление зависимости между факторами, определяющими функциональные возможности системы массового обслуживания, и эффективностью ее функционирования.

Основной задачей теории СМО является изучение режима функционирования обслуживающей системы и исследование явлений, возникающих в процессе обслуживания. Одной из важнейших характеристик обслуживающей системы является *время пребывания требования в очереди*. Очевидно, что это время можно сократить за счет увеличения количества обслуживающих устройств. Однако каждое дополнительное устройство требует определенных материальных затрат. Таким образом, в теории СМО возникают задачи оптимизации: каким образом достичь определенного уровня обслуживания (максимального сокращения очереди или потерь требований) при минимальных затратах, связанных с простоем обслуживающих устройств.

Основными элементами СМО являются:

- входящий поток требований;
- очередь требований;

- обслуживающие устройства (каналы);
- выходящий поток требований.

Входящий поток требований представляет собой совокупность требований, которые поступают в систему и нуждаются в обслуживании. Он изучается с целью установления закономерностей этого потока и дальнейшего улучшения качества обслуживания. В большинстве случаев входящий поток неуправляем и зависит от ряда случайных факторов. Число требований, поступающих в единицу времени, — случайная величина. Случайной величиной является также интервал времени между соседними поступающими требованиями. Однако среднее количество требований, поступивших в единицу времени, и средний интервал времени между соседними поступающими требованиями предполагаются заданными.

Среднее число требований, поступающих в систему массового обслуживания за единицу времени, является интенсивностью потока событий.

Для многих реальных процессов поток требований достаточно хорошо описывается законом распределения Пуассона.

По характеру обслуживания СМО классифицируются следующим образом: *системы с отказами* (требование, поступившее в момент, когда все каналы заняты, получает отказ, покидает систему и в дальнейшем процессе обслуживания не участвует; другое название — системы с потерями); *системы с очередью* (с ожиданием — упорядоченным и неупорядоченным, случайным и т.д.). Такие системы делятся далее на системы с неограниченным ожиданием и ограниченным (предельной длиной очереди, временем и др.) ожиданием.

По количеству обслуживающий устройств СМО делятся на одноканальные и многоканальные.

Уравнение Колмогорова

Определение 7.3. ***Уравнением Колмогорова*** называется уравнение

$$P'(t) = \lim_{\Delta t \to 0} \frac{P(t + \Delta t) - P(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{P(t)P(\Delta t) - P(t)}{\Delta t} =$$

$$= \lim_{\Delta t \to 0} \frac{P(\Delta t) - E}{\Delta t} P(t) = \Lambda P(t). \qquad (7.2)$$

Это матричная запись соответствующей системы уравнений для вероятностей состояний СМО. Оно решается при начальном условии $P(0) = E$ стандартными методами теории обыкновенных дифференциальных уравнений.

Если в уравнении (7.2) положить $P'(t) = 0$, то из него находится стационарное распределение вероятностей всех состояний системы.

Замечено, что составление уравнений (7.2) можно производить по графу состояний, на ребрах которых указаны интенсивности переходов и их направления в соответствии с правилом: в левой части k-го уравнения стоит P_k', а в правой — столько слагаемых, сколько состояний с ним связано в графе, причем со знаком «+», если состояние E_k переходит в него, и со знаком «–», если оно переходит в состояние E_k. Каждое слагаемое имеет вид произведения интенсивности перехода по данной стрелке на вероятность состояния, из которого эта стрелка исходит.

7.2. Решение задач по системам массового обслуживания

Задача 7.1. Дан граф переходов из состояния в состояние с указанием интенсивности переходов в системе массового обслуживания (СМО) с пуассоновским потоком событий (рис. 7.1).

Составить уравнения (7.2).

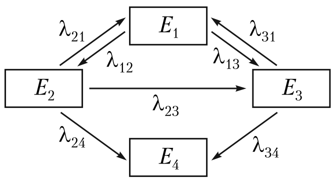


Рис. 7.1. **Граф переходов для СМО задачи 7.1**

Решение

Покажем, как составить уравнения (7.2) по смыслу с использованием формулы полной вероятности ФПВ и по правилу, приведенному выше.

По ФПВ имеем для 1-го уравнения системы (7.2)

$$P_1(t+\Delta t) = P_1(t)(1-\lambda_{12}\Delta t + o(\Delta t)) + P_1(t)(1-\lambda_{13}\Delta t + o(\Delta t)) + \\ + P_2(t)(\lambda_{21}\Delta t + o(\Delta t)) + P_3(t)(\lambda_{31}\Delta t + o(\Delta t));$$

отсюда получаем

$$\lim_{\Delta t\to 0}\frac{P_1(t+\Delta t)-P_1(t)}{\Delta t} =$$

$$= P_1(t)(-\lambda_{12}) + P_1(t)(-\lambda_{13}) + P_2(t)\lambda_{21} + P_3(t)\lambda_{31} = P_1'(t),$$

или окончательно $P_1'(t) = -(\lambda_{12}+\lambda_{13})P_1(t) + \lambda_{21}P_2(t) + \lambda_{31}P_3(t)$.

Аналогично остальные уравнения из (7.2) получаем в виде

$$\begin{cases} P_2'(t) = \lambda_{12}P_1(t) - (\lambda_{23}+\lambda_{24}+\lambda_{21})P_2(t), \\ P_3'(t) = \lambda_{13}P_1(t) + \lambda_{23}P_2(t) - (\lambda_{31}+\lambda_{34})P_3(t), \\ P_4'(t) = \lambda_{24}P_2(t) + \lambda_{34}P_3(t), \\ P_1(t)+P_2(t)+P_3(t)+P_4(t) = 1. \end{cases}$$

Ту же систему уравнений получаем сразу по правилу, что непосредственно проверяется.

Задача 7.2. Для иллюстрации составления и решения системы уравнений (7.2) и нахождения стационарных вероятностей приведем одноканальную СМО с отказами, с пуассоновскими потоками событий, описываемую следующим графом переходов (рис. 7.2), где E_0 — система свободна, E_1 — система занята.

Выписать уравнения (7.2) и найти вероятности состояний.

$$\boxed{E_0} \underset{\mu}{\overset{\lambda}{\leftrightarrows}} \boxed{E_1}$$

Рис. 7.2. **Граф переходов для СМО задачи 7.2**

Решение

Система уравнений (7.2) легко выписывается по правилу:

$$\begin{cases} P_0'(t) = -\lambda P_0(t) + \mu P_1(t), \\ P_1'(t) = \lambda P_0(t) + \mu P_1(t), \\ P_0(t) + P_1(t) = 1. \end{cases}$$

Решим полученную систему дифференциальных уравнений для нахождения вероятностей всех состояний системы:

$$P_0'(t) = -(\lambda + \mu)P_0(t) + \mu, \text{ так как } P_0(t) + P_1(t) = 1;$$

$$\frac{dP_0(t)}{-(\lambda+\mu)P_0(t)+\mu} = dt \Rightarrow \frac{d(-(\lambda+\mu)P_0(t))}{-(\lambda+\mu)P_0(t)+\mu} = -(\lambda+\mu)dt \Rightarrow$$

$$\Rightarrow d\ln(-(\lambda+\mu)P_0(t)+\mu) = -(\lambda+\mu)dt \Rightarrow$$

$$\Rightarrow \ln(-(\lambda+\mu)P_0(t)+\mu) = -(\lambda+\mu)t + c \Rightarrow$$

$$\Rightarrow -(\lambda+\mu)P_0(t)+\lambda = e^{c-(\lambda+\mu)t} \Rightarrow P_0(t) = \frac{\mu}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \Rightarrow$$

$$\Rightarrow P_1(t) = 1 - P_0(t) = \frac{\lambda}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t};$$

$$\lim_{t\to\infty} P_0(t) = \frac{\mu}{\lambda+\mu}; \quad \lim_{t\to\infty} P_1(t) = \frac{\lambda}{\lambda+\mu}$$

— это стационарные вероятности состояний E_0 и E_1: (π_0, π_1).

Стационарное распределение также можно получить приравниванием $P_0'(t)$ и $P_1'(t)$ нулю, т.е. если π — стационарное распределение, а Λ — матрица интенсивности переходов из состояния в состояние, то матричное уравнение для определения π имеет вид $\pi(t)\Lambda = 0$. Действительно, имеем в примере 7.2:

$$\Lambda = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}, \ \pi(t) = (\pi_0, \pi_1) \Rightarrow \begin{cases} 0 = -\lambda\pi_0 + \mu\pi_1, \\ 0 = \lambda\pi_0 - \mu\pi_1, \quad \Rightarrow \\ \pi_0 + \pi_1 = 1 \end{cases}$$

$$\Rightarrow \pi(t)\Lambda = (\pi_0, \pi_1)\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} = 0 \Rightarrow$$

$$\Rightarrow \pi_0 = \frac{\mu}{\lambda}\pi_1;\ \pi_1\left(1 - \frac{\mu}{\lambda}\right) = 1 \Rightarrow \pi_0 = \frac{\mu}{\lambda+\mu};\ \pi_1 = \frac{\lambda}{\lambda+\mu},$$

что совпадает с полученным выше результатом.

Задача 7.3. Дана одноканальная СМО с пуассоновским потоком событий и графом переходов, изображенным на рис. 7.3. Составить систему уравнений для нахождения вероятностей всех состояний и найти стационарное распределение (матрицу Λ).

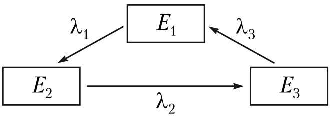


Рис. 7.3. **Граф переходов для СМО задачи 7.3**

Решение

Уравнения (7.2) имеют вид

$$\begin{cases} P_1'(t) = -\lambda_1 P_1(t) + \lambda_3 P_3(t), \\ P_2'(t) = \lambda_1 P_1(t) - \lambda_2 P_2(t), \\ P_3'(t) = \lambda_2 P_2(t) - \lambda_3 P_3(t), \\ P_1(t) + P_2(t) + P_3(t) = 1, \end{cases}$$

отсюда следует

$$\begin{cases} 0 = -\lambda_1\pi_1 + \lambda_3\pi_3, \\ 0 = \lambda_1\pi_1 - \lambda_2\pi_2, \\ 0 = \lambda_2\pi_2 - \lambda_3\pi_3, \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \Rightarrow \pi_3 = \frac{\lambda_1}{\lambda_3}\pi_1,\ \ \pi_2 = \frac{\lambda_1}{\lambda_2}\pi_1 \Rightarrow \pi_1\left(1 + \frac{\lambda_1}{\lambda_2} + \frac{\lambda_1}{\lambda_3}\right) = 1 \Rightarrow$$

$$\Rightarrow \begin{cases} \pi_1 = \dfrac{1}{1 + \dfrac{\lambda_1}{\lambda_2} + \dfrac{\lambda_1}{\lambda_3}} = \dfrac{\lambda_2\lambda_3}{\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2}, \\ \pi_2 = \dfrac{\lambda_1\lambda_2\lambda_3}{\lambda_3(\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2)} = \dfrac{\lambda_2\lambda_3}{\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2}, \\ \pi_3 = \dfrac{\lambda_1\lambda_2\lambda_3}{\lambda_2(\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2)} = \dfrac{\lambda_2\lambda_3}{\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2}. \end{cases}$$

Проверка: $\pi_1 + \pi_2 + \pi_3 = 1$.

Матрица $\Lambda = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 \\ 0 & -\lambda_2 & \lambda_2 \\ \lambda_3 & 0 & -\lambda_3 \end{pmatrix}$.

Задание 7.1. Дана одноканальная СМО (обобщенный циклический процесс) с пуассоновскими потоками событий и графом переходов (рис. 7.4). Составить систему уравнений для нахождения вероятностей всех состояний и найти стационарное распределение (матрицу Λ).

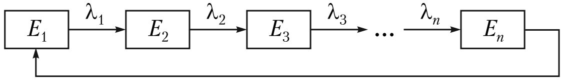


Рис. 7.4. **Граф переходов к заданию 7.1**

Задача 7.4. Дана одноканальная СМО с ожиданием, пуассоновским потоком событий (λ — интенсивность потока заявок, μ — интенсивность потока обслуживания), с графом переходов, изображенным на рис. 7.5. Составить систему уравнений для нахождения вероятностей всех состояний и найти стационарное распределение (матрицу Λ). Очередь имеет m мест, а остальные заявки пропадают.

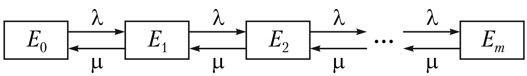


Рис. 7.5. **Граф переходов для СМО задачи 7.4**

Решение

Очередь имеет m мест. Обозначим $\lambda + \mu = \nu$. Тогда

$$\Lambda = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \mu & -\nu & \lambda & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & \mu & -\nu & \lambda & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \mu & -\nu & \lambda \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & \mu & -\mu \end{pmatrix}.$$

Система дифференциальных уравнений (7.2) имеет следующий вид (нулевое состояние — отсутствие в очереди заявок, $(m + 1)$-е — в очередь поступает $(m + 1)$-я заявка):

$$\begin{cases} P_0'(t) = -\lambda P_0(t) + \mu P_1(t), \\ P_1'(t) = \lambda P_0(t) - \nu P_1(t) + \mu P_2(t), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ P_k'(t) = \lambda P_{k-1}(t) - \nu P_k(t) + \mu P_{k+1}(t), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ P_{m+1}'(t) = \lambda P_m(t) - \mu P_{m+1}(t), \\ P_0(t) + P_1(t) + \dots + P_{m+1}(t) = 1, \end{cases}$$

отсюда следует система для нахождения стационарного распределения

$$\begin{cases} 0 = -\lambda\pi_0 + \mu\pi_1, \\ 0 = \lambda\pi_0 - \nu\pi_1 + \mu\pi_2, \\ \dots \\ 0 = \lambda\pi_m - \mu\pi_{m+1}, \\ \sum_{i=0}^{m+1} \pi_1 = 1. \end{cases}$$

Обозначим $\rho = \dfrac{\lambda}{\mu} \Rightarrow$

$$\Rightarrow \begin{cases} \pi_1 = \rho\pi_0, \\ \pi_2 = \rho\pi_1, \\ \ldots \\ \pi_{k+1} = \rho\pi_k, \\ \ldots \\ \pi_{m+1} = \rho\pi_m, \\ \sum\limits_{i=0}^{m+1}\pi_{m+1} = \pi_0(1+\rho+\ldots+\rho^{m+1}) \end{cases} \Rightarrow$$

$\Rightarrow$ стационарное распределение

$$\begin{cases} \pi_0 = \dfrac{1}{1+\rho+\ldots+\rho^{m+1}} = \dfrac{1-\rho}{1-\rho^{m+2}}, \\ \pi_k = \pi_0\rho^k. \end{cases}$$

Задача 7.5. Молекула, испытавшая столкновение в момент времени $t = 0$ с другой молекулой и не имевшая других столкновений до момента t, испытывает столкновение в интервале $(t, t + \Delta t)$ с вероятностью, равной $\lambda\Delta t + o(\Delta t)$. Найти вероятность того, что время между двумя столкновениями больше t.

Решение

Имеем два состояния системы: E_0 — за время t молекула не имела столкновений, и E_1 — за время t молекула имела столкновения. Плотность перехода из состояния E_0 в состояние E_1 задана и равна λ. Граф перехода представлен на рис. 7.6.

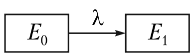


Рис. 7.6. **Граф переходов для задачи 7.5**

Система уравнений (7.2) имеет вид

$$\begin{cases} P_0'(t) = -\lambda P_0(t), \\ P_0(0) = 1 \text{ — начальное условие} \end{cases} \Rightarrow$$

$\Rightarrow P_0(0) = e^{-\lambda 0} + c = 1 + c = 1 \Rightarrow c = 0 \Rightarrow P_0(t) = e^{-\lambda t}$.

Задача 7.6. Вероятность бактерии разделиться надвое за время Δt ($\Delta t \to 0$) равна $a\Delta t + o(\Delta t)$ и не зависит от числа бактерий и предшествующих делений. Составить систему уравнений для нахождения вероятности того, что в момент t будет i бактерий, если в начальный момент была одна бактерия, т.е. $P_0(1)$.

Решение

Составим дифференциальные уравнения (7.2) для вероятностей состояний — чисел бактерий. $P_i(t)$ — вероятность того, что к моменту t будет i бактерий.

По ФПВ имеем:

$$P_i(t+\Delta t) = P_i(t)(1-\ a\Delta t)^i + P_{i-1}(t)C_{i-1}^1(1-a\Delta t)^{i-1} +$$

$$+P_{i-2}(t)C_{i-2}^2(a\Delta t)^2(1-a\Delta t)^{i-2} + \ldots \approx P_i(t)(1-ia\Delta t)+(i-1)a\Delta t P_{i-1}(t);$$

$$P_1'(t) = \lim_{\Delta t \to 0} \frac{P(t+\Delta t)-P(t)}{\Delta t} =$$

$$= \lim_{\Delta t \to 0} \frac{P_i(t)(1-a\Delta t)+(i-1)a\Delta t P_{i-1}(t)-P_i(t)}{\Delta t} = -iaP_i(t)+(i-1)aP_i(t).$$

Иначе, имеем систему уравнений (7.2)

$$\begin{cases} P_i(t) = -iaP_i(t) + (i-1)aP_i(t), \\ P_1(0) = 1;\ P_2(0) = 0,\ i \neq 1. \end{cases}$$

‖ **Задание 7.2.** Численно решите задачу 7.6 для $i = 3$.

Контрольные вопросы и задания

1. В чем заключается понятие потока событий и его интенсивности?
2. Объясните смысл уравнения Колмогорова — Чепмена.
3. Каковы свойства матрицы переходных вероятностей однородного потока событий?
4. Приведите определение и классификацию СМО.
5. Приведите правило составления системы уравнений для вероятностей состояний СМО по графику переходов из состояния в состояние.
6. Опишите проведение анализа СМО с пуассоновским потоком событий (заявок и обслуживания). Укажите стационарное распределение вероятностей состояний системы.
7. Опишите проведение анализа СМО с пуассоновскими потоками событий с ожиданием. Укажите стационарное распределение вероятностей состояний системы.

Глава 8
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ И ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

В результате изучения данной главы студенты должны:

знать

- основные определения и смысл числовых характеристик случайных процессов;
- канонические и спектральные разложения случайных процессов;

уметь

- вычислять числовые характеристики случайных процессов;
- анализировать процессы на их стационарность в широком и узком смысле;

владеть

- методами анализа случайных процессов по их числовым характеристикам.

В главе вводятся и обсуждаются свойства основных числовых характеристик случайных процессов (математического ожидания, дисперсии и корреляции). Определяются понятия стационарности процесса в широком и узком смысле в связи со значениями его числовых характеристик. Обсуждается каноническое разложение случайного процесса, его получение и практический смысл. Разбирается большое число числовых примеров вычисления характеристик процессов, получение канонических разложений.

Изучаются линейные однородные преобразования случайных процессов, отвечающие реальным практическим ситуациям, с вычислением их числовых характеристик. Вводится понятие спектрального разложения случайного процесса.

8.1. Определение основных характеристик случайных процессов и их свойства

Пусть $X(t)$ — случайная функция от t или случайный процесс, t — время. При фиксированном t $X(t)$ — случайная величина с законом распределения $F(x, t) = P\{X(t) < x, t$ — фиксировано$\}$.

Характеристики случайных функций

Обозначим через $[X(t)]$ значения случайной функции $X(t)$ для фиксированного значения t.

Определение 8.1. ***Математическим ожиданием случайной функции*** $X(t)$ называется неслучайная функция $M_X(t)$, которая при каждом значении аргумента t равна математическому ожиданию соответствующего сечения случайной функции, т.е. значения $M_X(t)$ при фиксированном значении t: $M_X(t) = M[X(t)]$.

Определение 8.2. ***Дисперсией случайной функции*** $X(t)$ называется неслучайная функция $D_X(t)$, значение которой для каждого значения аргумента t равно дисперсии соответствующего сечения функции: $D_X(t) = D[X(t)]$; $\sigma_X(t) = \sqrt{D_X(t)}$.

Определение 8.3. ***Корреляционной функцией случайной функции*** $X(t)$ называется неслучайная функция двух аргументов $K_X(t_1, t_2)$, которая равна корреляционному моменту соответствующего сечения случайной функции:

$$K_X(t_1, t_2) = M[X(t_1) - M_X(t_1)(X(t_1) - M_X(t_2)] .$$

При $t_1 = t_2$ $K_X(t_1, t_1) = D_X(t_1)$.

Нормированная корреляционная функция:

$$r_X(t_1, t_2) = \frac{K_X(t_1, t_2)}{\sigma_X(t_1)\sigma_X(t_2)}.$$

Свойства характеристик случайных функций

1. Если $Y(t) = X(t) + \varphi(t)$, где $\varphi(t)$ — неслучайная функция, то

$$M_Y(t) = M_X(t) + \varphi(t);\ K_Y(t_1,t_2) = K_X(t_1, t_2).$$

2. Если $Y(t) = \varphi(t)X(t)$, где $\varphi(t)$ — неслучайная функция, то

$$M_Y(t) = \varphi(t)M_X(t);\ K_Y(t_1,t_2) = \varphi(t_1)\varphi(t_2)K_X(t_1, t_2),$$

из чего следует, что $D_Y(t) = \varphi^2 D_X(t)$. Если $\varphi(t) = c$ — const, то $K_Y(t) = c^2K_X(t_1, t_2)$, из чего следует, что $D_Y(t) = c^2D_X(t)$.

3. Центральный момент определяется по формуле $\mu = \mu(t) = X(t) - M_X(t)$, тогда

$$M_\mu(t) = 0,\ K_\mu(t_1, t_2) = K_X(t_1, t_2).$$

4. Нормировка $X_N = X_N(t)$ определяется по формуле

$$X_N(t) = \frac{X(t) - M_X(t)}{\sigma_X(t)}.$$

Тогда $K_{X_N}(t_1, t_2) = \dfrac{K_X(t_1, t_2)}{\sigma_X(t_1)\sigma_X(t_2)} = r_X(t_1, t_2).$

8.2. Стационарность случайных процессов

Определение 8.4. Случайная функция $X(t)$ называется ***стационарной*** (или ***стационарной в широком смысле***), если ее математическое ожидание $M_X(t) = m_X$ — const, а корреляционная функция зависит только от разности между своими аргументами t_1 и t_2, т.е. от $\tau = t_2 - t_1$: $K_X(t_1, t_2) = k_X(t_2 - t_1) = k_X(\tau)$.

Свойства характеристик стационарной функции:

1) из симметрии $K_X(t_1, t_2) = K_X(t_2, t_1)$ следует, что $k_X(\tau) = k_X(-\tau)$, т.е. корреляционная функция стационарной случайной функции четна относительно значения τ;

2) $D_X(t) = R_X(t, t) = k_X(0) = \text{const}$.

Определение 8.5. Случайная функция или процесс $X(t)$ называется ***стационарным в узком смысле***, если все ее n-мерные законы $F_n(x_1, ..., x_n; t_1, ..., t_n)$ не изменяются от сдвига по оси t всех временных моментов $t_1, ..., t_n$ на одну и ту же величину, т.е. зависят только от взаимного расположения этих моментов.

У стационарного в узком смысле процесса $X(t)$ математическое ожидание и дисперсия — const, а корреляция $K_X(t_1, t_2)$ зависит лишь от $\tau = t_1 - t_2$.

Из определений 8.4 и 8.5 следует, что из *стационарности в узком смысле следует стационарность в широком смысле*, но не наоборот.

Замечание. Для стационарных нормальных процессов понятия стационарности в широком и узком смысле совпадают.

8.3. Задачи на нахождение числовых характеристик случайного процесса

Задача 8.1. Пусть $\xi(t) = X\cos\lambda t + Y\sin\lambda t$, где X и Y — некоррелированные СВ, λ — const, $MX = MY = 0$, $DX = DY = 1$. Найти $M\xi(t)$, $D\xi(t)$, $K_\xi(t_1, t_2)$. Проанализировать стационарность процесса $\xi(t)$ в широком смысле.

Решение

$$M\xi(t) = MX\cos\lambda t + MY\sin\lambda t = 0;\ D\xi(t) = M(\xi(t) - M\xi(t))^2 =$$

$$= M(\xi(t))^2 = M(X^2\cos^2\lambda t + 2XY\cos\lambda t\sin\lambda t + Y^2\sin^2\lambda t) =$$

$$= \cos^2\lambda t MX^2 + 2\cos\lambda t\sin\lambda t MXY + \sin^2\lambda t MY^2 = DX\cos^2\lambda t + DY\sin^2\lambda t = 1;$$

$$K_\xi(t_1, t_2) = M(\xi(t_1), \xi(t_2)) =$$
$$= M[X^2\cos\lambda t_1\cos\lambda t_2 + XY(\cos\lambda t_1\sin\lambda t_2 + \cos\lambda t_2\sin\lambda t_1) +$$
$$+ Y^2\sin\lambda t_1\sin\lambda t_2] = \cos\lambda t_1\cos\lambda t_2 MX^2 +$$
$$+ (\cos\lambda t_1\sin\lambda t_2 + \cos\lambda t_2\sin\lambda t_1)MXY + \sin\lambda t_1\sin\lambda t_2 MY^2 = \cos\lambda(t_1 - t_2),$$

т.е. $K_\xi(t_1, t_2)$ зависит лишь от разности $t_1 - t_2$, значит, процесс стационарен в широком смысле.

Задача 8.2. Пусть $\xi(t) = \sum_{k=1}^{n} b_k\xi_k(t)$; $\xi_k(t) = X_k\cos\lambda_k t + Y_k\sin\lambda_k t$; $\sum_{k=1}^{n} b_k^2 = 1$; $MX_k = MY_k = 0$; $DX_k = DY_k = 1$; $k = 1, ..., n$; $MX_iX_j = MY_iY_j = 0$, $i, j = 1, ..., n$; при $i \neq j$ $MX_iX_j = 0$. Найти $M\xi(t)$, $D\xi(t)$, $K\xi(t_1, t_2)$.

Решение

$$M\xi(t) = \sum_{k=1}^{n} b_k M\xi_k(t) = 0; \; D\xi(t) = \sum_{k=1}^{n} b_k^2 D\xi_k(t) = \sum_{k=1}^{n} b_k^2 = 1;$$

$$K\xi(t_1, t_2) = M\left(\sum_{k=1}^{n} b_k(X_k\cos\lambda_k t_1 + Y_k\sin\lambda_k t_1)\sum_{k=1}^{n} b_k(X_k\cos\lambda_k t_2 + Y_k\sin\lambda_k t_2)\right) =$$
$$= \sum_{k=1}^{n} b_k^2\cos\lambda_k t_1\cos\lambda_k t_2 MX_k^2 + \sum_{k=1}^{n} b_k^2\sin\lambda_k t_1\sin\lambda_k t_2 MY_k^2 =$$
$$= \sum_{k=1}^{n} b_k^2\cos\lambda_k(t_1 - t_2),$$

так как $MX_iX_j = MY_iY_j = 0$; $i \neq j$, $MX_iX_j = 0$, т.е. $K_\xi(t_1, t_2)$ зависит лишь от разности $(t_1 - t_2)$), значит, процесс стационарен в широком смысле.

Задача 8.3. Найти $M\xi(t)$, $D\xi(t)$, $K\xi(t_1, t_2)$, где $\xi(t) = Xt + Y\sin t$, $MX = 1$; $MY = -1$, $K_{XY} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$.

Решение

$$M\xi(t) = tMX + \sin t\, MY = t - \sin t;$$
$$K\xi(t_1, t_2) = M(\xi(t_1) - M\xi(t_1))(\xi(t_2) - M\xi(t_2)) =$$
$$= M(Xt_1 + Y\sin t_1 - t_1MX - \sin t_1MY)(Xt_2 + \sin t_2 - t_2MX - \sin t_2MY) =$$
$$= M((X - MX)t_1 + (Y - MY)\sin t_1)(X - MX)t_2 + (Y - MY)\sin t_2) =$$
$$= t_1t_2M(X - MX)^2 + t_1\sin t_2M(X - MX)(Y - MY) +$$
$$+ t_2\sin t_1M(X - MX)(Y - MY) + \sin t_1\sin t_2M(Y - MY)^2 =$$
$$= 2t_1t_2 + 3\sin t_1\sin t_2 + t_1\sin t_2 + t_2\sin t_1;$$
$$D\xi(t) = K\xi(t, t) = 2t^2 + 3\sin^2 t + 2t\sin t.$$

Задача 8.4. Найти $M\xi(t)$, $D\xi(t)$, $K\xi(t_1, t_2)$, если $\xi(t)$ — пуассоновский процесс с параметром λ.

Решение

$$P\{\xi(t)=n\}=\frac{(\lambda t)^n e^{-\lambda t}}{n!};\ n=0,1,2,...;$$

$$M\xi(t)=\sum_{k=0}^{\infty}\frac{k(\lambda t)^k e^{-\lambda t}}{k!}=e^{-\lambda t}\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{(k-1)!}=e^{-\lambda t}\lambda t e^{\lambda t}=\lambda t;$$

$$K\xi(t_1,t_2)=M(\xi(t_1)-M\xi(t_1))(\xi(t_2)-M\xi(t_2))=$$
$$=M\xi(t_1)\xi(t_2)-\lambda t_2 M\xi(t_1)-\lambda t_1 M\xi(t_2)+\lambda t_1 t_2=$$
$$=M\xi(t_1)\xi(t_2)-2\lambda^2 t_1 t_2+\lambda^2 t_1 t_2=M\xi(t_1)\xi(t_2)-\lambda^2 t_1 t_2.$$

Остается вычислить $M\xi(t_1)\xi(t_2)$. Пусть $t_1 < t_2$, тогда $P(\xi(t_2)=m, \xi(t_1)=n)=0$ при $m=n$; а при $m>n$

$$P(\xi(t_2)=m,\xi(t_1)=n)=P(\xi(t_1)=n,\xi(t_2)=m)=$$
$$=P(\xi(t_1)=n)P(\xi(t_2-t_1)=m-n);$$

$$M\xi(t_1)\xi(t_2)=\sum_{n=0}^{\infty}nP(\xi(t_1)=n)\sum_{m=n}^{\infty}mP(\xi(t_2-t_1)=m-n)=$$
$$=\sum_{n=0}^{\infty}n\frac{(\lambda t_1)^n}{n!}e^{-\lambda t_1}\sum_{m=n}^{\infty}m\frac{[\lambda(t_2-t_1)]^{m-n}}{(m-n)!}e^{-\lambda(t_2-t_1)}=$$
$$=\lambda t_1\sum_{n=1}^{\infty}\frac{(\lambda t_1)^{n-1}}{(n-1)!}e^{-\lambda t_1}\sum_{k=0}^{\infty}(k+n)\frac{[\lambda(t_2-t_1)]^k}{k!}e^{-\lambda(t_2-t_1)}=$$
$$=\lambda t_1 e^{-\lambda t_1}e^{-\lambda(t_2-t_1)}\sum_{n=1}^{\infty}\frac{(\lambda t_1)^{n-1}}{(n-1)!}\left(\sum_{k=1}^{\infty}k\frac{[\lambda(t_2-t_1)]^k}{k!}+n\sum_{k=0}^{\infty}\frac{[\lambda(t_2-t_1)]^k}{k!}\right)=$$
$$=\lambda t_1 e^{-\lambda t_2}\sum_{n=1}^{\infty}\frac{(\lambda t_1)^{n-1}}{(n-1)!}[\lambda(t_2-t_1)e^{\lambda(t_2-t_1)}+ne^{\lambda(t_2-t_1)}]=$$
$$=\lambda t_1 e^{\lambda t_2}e^{\lambda(t_2-t_1)}\left(e^{\lambda t_1}\lambda(t_2-t_1)+\sum_{n=1}^{\infty}n\frac{(\lambda t_1)^{n-1}}{(n-1)!}\right)=$$
$$=\lambda t_1 e^{-\lambda t_1}\left(e^{\lambda t_1}\lambda(t_2-t_1)+\sum_{n=2}^{\infty}\frac{(\lambda t_1)^{n-1}}{(n-2)!}+e^{\lambda t_1}\right)=$$
$$=\lambda t_1 e^{-\lambda t_1}[e^{\lambda t_1}\lambda(t_2-t_1)+\lambda t_1 e^{\lambda t_1}+e^{\lambda t_1}]=\lambda t_1(\lambda t_2-\lambda t_1+\lambda t_1+1)=$$
$$=\lambda t_1(\lambda t_2+1)=\lambda^2 t_1 t_2+\lambda t_1;$$

$$K_\xi(t_1,t_2)=\lambda^2 t_1 t_2+\lambda t_1-\lambda^2 t_1 t_2=\lambda t_1 \text{ при } t_1<t_2.$$

Аналогично при $t_2 < t_1$ $K_\xi(t_1, t_2) = \lambda t_2$, отсюда получаем общую формулу: $K_\xi(t_1, t_2) = \lambda\min(t_1, t_2)$. Тогда $D\xi(t) = K_\xi(t, t) = \lambda\min(t, t) = \lambda t$.

Задача 8.5. Пусть $\xi(t)$ — случайный процесс, который принимает значения +1 и –1 попеременно в моменты наступления событий в пуассоновском потоке с параметром λ. Найти $M\xi(t)$, $D\xi(t)$, $K\xi(t_1, t_2)$.

Решение

$M\xi(t) = 0$ очевидно по условию, так как сечение случайной функции ξ имеет закон распределения

ξ	–1	1
p	1/2	1/2

Тогда $D\xi(t) = (-1)^2\frac{1}{2} + 1^2\frac{1}{2} = 1$; $K_\xi(t_1, t_2) = M\xi(t_1)\xi(t_2)$.

Пусть $t_1 < t_2$, тогда $\xi(t_1)\xi(t_2) = -1$.

Обозначим $\tau = t_2 - t_1$, P_1 — вероятность четного, а P_2 — нечетного чисел событий за время τ. Тогда

$$P_1 = \sum_{n=0}^{\infty}\frac{(\lambda\tau)^{2n}}{(2n)!}e^{-\lambda\tau} = e^{-\lambda\tau}\frac{e^{\lambda\tau} + e^{-\lambda\tau}}{2};\ P_2 = \sum_{n=0}^{\infty}\frac{(\lambda\tau)^{2n+1}}{(2n+1)!}e^{-\lambda\tau} = e^{-\lambda\tau}\frac{e^{\lambda\tau} - e^{-\lambda\tau}}{2}.$$

Из этого следует, что

$$K_\xi(t_1, t_2) = M\xi(t_1)\xi(t_2) = P_1 - P_2 = e^{-\lambda\tau}\frac{2e^{-\lambda\tau}}{2} = e^{-2\lambda\tau},$$

так как СВ $\xi(t_1)\xi(t_2)$ имеет в сечении распределение

$\xi(t_1)\xi(t_2)$	$+1$	-1
P	P_1	P_2

Аналогично, при $t_2 < t_1$ $K_\xi(t_1, t_2) = e^{-2\lambda(-\tau)}$. Объединяя случаи $t_1 < t_2$ и $t_2 < t_1$, получаем окончательно, что $K_\xi(t_1, t_2) = e^{-2\lambda|\tau|}$. Отсюда следует, что $D\xi(t) = K_\xi(t, t) = e^{-2\lambda 0} = 1$, что было получено ранее.

Задача 8.6. Найти математическое ожидание, дисперсию и корреляционную функцию синусоиды постоянной частоты ω со случайной амплитудой X, если $MX = 1$, $DX = 0{,}2$, $X(t) = X\sin\omega t$.

Решение

$$MX = \sin\omega t;$$

$$K_X(t_1, t_2) = M((X(t_1) - MX(t_1))(X((t_2) - MX((t_2)) =$$

$$= M(X\sin\omega t_1 - \sin\omega t_1)(X\sin\omega t_2 - \sin\omega t_2) = \sin\omega t_1 \sin\omega t_2 M(X - MX) =$$

$$= \sin\omega t_1 \sin\omega t_2 DX = 0{,}2\sin\omega t_1 \sin\omega t_2.$$

Задача 8.7. Пусть $X(t)$ и $Y(t)$ — две случайные функции: $X(t) = V_1\cos\omega_1 t + V_2\sin\omega_2 t$; $Y(t) = U_1\cos\omega_2 t + U_2\sin\omega_2 t$; $MU_1 = MU_2 = MV_1 = MV_2 = 0$, $DV_1 = DV_2 = 1$; $DU_1 = DU_2 = 4$; для случайных величин V_1, V_2, U_1, U_2, $(r) = \begin{pmatrix} 1 & 0 & 0{,}5 & 0 \\ & 1 & 0 & -0{,}5 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}$.

Найти $K_{XY}(t_1, t_2)$ при $t_1 = 0$, $t_2 = 1$.

Решение

$$K_{XY}(t_1, t_2) = M(X(t_1) - MX(t_1))(Y(t_2) - MY(t_2)) =$$

$$= M(V_1\cos\omega_1 t_1 + V_2\sin\omega_2 t_2)(U_1\cos\omega_2 t_2 + U_2\sin\omega_2 t_2)) =$$

$$= \cos\omega_1 t_1 \cos\omega_2 t_2 MV_1U_1 + \cos\omega_1 t_1 \sin\omega_2 t_2 MV_1U_2 +$$

$$+ \sin\omega_1 t_1 \cos\omega_2 t_2 MV_2U_1 + \sin\omega_1 t_1 \sin\omega_2 t_2 MV_2U_2 =$$

$$= \cos\omega_1 t_1 \cos\omega_2 t_2 - \sin\omega_1 t_1 \sin\omega_2 t_2 = \cos(\omega_1 t_1 + \omega_2 t_2),$$

так как

$$MV_1U_1 = K_{13} = r_{23}\sqrt{DV_1DU_1} = 2r_{13} = 1;\ MV_1U_2 = K_{14} = r_{14}\sqrt{DV_1DU_2} = 0;$$

$$MV_2U_1 = K_{23} = r_{23}\sqrt{DV_1DU_1} = 0;\ MV_2U_2 = K_{24} = r_{24}\sqrt{DV_2DU_2} = -1.$$

При $t_1 = 0$, $t_2 = 1$ $K_{XY}(t_1, t_2) = \cos\omega_2$.

8.4. Каноническое разложение случайного процесса

Определение 8.6. *Простейшей* случайной функцией (процессом) называется функция вида $X(t) = X\varphi(t)$, где X — случайная величина, а $\varphi(t)$ — неслучайная функция. Все реализации $X(t)$ могут быть совмещены путем изменения масштаба по оси ординат (при этом ось абсцисс — одна из реализаций $X(t)$). $MX(t) = MX\varphi(t) = \varphi(t)MX$.

При $MX = 0$ простейшая случайная функция называется ***элементарной*** (ЭСФ). Определим корреляцию ЭСФ:

$$K_X(t_1, t_2) = M(X(t_1)X(t_2)) = \varphi(t_1)\varphi(t_2)MX^2 = \varphi(t_1)\varphi(t_2)DX.$$

Определение 8.7. ***Каноническим разложением*** случайной функции $X(t)$ называется всякое представление $X(t)$ в виде суммы ее математического ожидания и взаимно некоррелированных элементарных функций, т.е. в виде

$$X(t) = MX(t) + \sum_{k=1}^{m} U_k\varphi_k(t), \tag{8.1}$$

где U_k — случайные величины с дисперсиями D_k — коэффициенты канонического разложения; $\varphi_k(t)$ — координатные функции.

Канонические разложения случайных функций удобны для выполнения различных операций над ними, так как в равенстве (8.1) участвуют лишь случайные величины в качестве коэффициентов разложения, а зависимость от t определяется только неслучайными функциями. Таким образом, операции над случайной функцией $X(t)$ по формуле (8.1) сводятся к соответствующим операциям над обычными в математическом анализе неслучайными функциями $\varphi_k(t)$.

Определение 8.8. ***Комплексной случайной функцией*** называется функция

$$Z(t) = X(t) + iY(t),$$

где $X(t)$ и $Y(t)$ — действительные случайные функции от действительного аргумента t. Тогда математическое ожидание

$$MZ(t) = MX(t) + iMY(t);$$

дисперсия

$$DZ(t) = M(Z(t) - MZ(t))^2;$$

если случайные функции $X(t)$ и $Y(t)$ некоррелированы, то $DZ(t) = DX(t) + DY(t)$; корреляция случайной функции $Z(t)$

$$K_Z(t_1, t_2) = K_X(t_1, t_2) + K_Y(t_1, t_2) + i(K_X(t_1, t_2) + K_Y(t_1, t_2));$$

если случайные функции $X(t)$ и $Y(t)$ некоррелированы, то

$$K_Z(t_1, t_2) = K_X(t_1, t_2) + K_Y(t_1, t_2).$$

Если случайная функция $X(t)$ имеет каноническое разложение (8.1) и *действительна*, то корреляционная функция имеет вид

$$K_X(t_1, t_2) = \sum_{k=1}^{m} D_k \varphi_k(t_1) \varphi_k(t_2). \tag{8.2}$$

Если случайная функция $X(t)$ имеет каноническое разложение (8.1) и *комплексна*, то корреляционная функция имеет вид

$$K_X(t_1, t_2) = \sum_{k=1}^{m} D_k \varphi_k(t_1) \overline{\varphi_k(t_2)}, \tag{8.3}$$

где $\overline{\varphi_k(t_2)}$ — комплексная сопряженная величина с $\varphi_k(t_2)$. Разложения (8.2) и (8.3) называются ***каноническими разложениями корреляционной функции*** в указанных случаях.

Рассмотрим задачи, использующие каноническое разложение.

Задача 8.8. Случайная функция $X(t)$ задана каноническим разложением $X(t) = t - 3\cos t + u(t + \cos t) + v\cos 2t$, $Du = 1$; $Dv = 2$. Найти $MX(t)$, $DX(t)$, $K_X(t_1, t_2)$.

Решение

Из разложения (8.1) следует, что $MU = MV = 0$, $MX(t) = t - 3\cos t$; по формуле (8.2)

$$DX(t) = \sum_{k=1}^{2} D_k \varphi_k^2(t) = 1(1 + \cos t)^2 + 2(\cos 2t)^2 = (1 + \cos t)^2 + 2\cos^2 2t;$$

$$K_X(t_1, t_2) = (t_1 + \cos t_1)(t_2 + \cos t_2) + 2\cos 2t_1 \cos 2t_2.$$

Задача 8.9. Задана случайная функция $X(t) = X_1 t + X_2 \sin t$, где случайный вектор (X_1, X_2) имеет математическое ожидание $(+1, -1)$ и корреляционную матрицу $\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$. Построить каноническое разложение $X(t)$, найти $MX(t)$, $DX(t)$ и $K_X(t_1, t_2)$.

Решение

Имеем $MX(t) = MX_1 t + MX_2 \cos 2t = t - \sin t$. Найдем каноническое разложение случайной функции $X(t)$. Представим $X(t)$ в виде

$$X(t) = (X_1 - 1)t + (X_2 + 1)\sin t + t - \sin t = Ut + V\sin t + MX(t),$$

где $X_1 - 1 = U$, $X_2 + 1 = aU + V$, $MX(t) = t - \sin t$, причем $MU = MV = 0$, $DU = DX_1 = 2$. Подберем также значение a, чтобы $MUV = 0$:

$$M(x_1 - 1)(x_2 + 1) = M(U\,(aU + V)) = aMU^2 + MUV =$$
$$= a(DU + (MU)^2) + MUV = 2a,$$

так как $MU = MUV = 0$, или

$$MX_1X_2 - MX_2 + MX_1 - 1 = K_{x_1x_2} + MX_1X_2 + 1 + 1 - 1 = 1 - 1 + 1 + 1 - 1 = 2a,$$

откуда $a = \dfrac{1}{2}$, т.е. $X_1 - 1 = U$; $X_2 + 1 = \dfrac{1}{2}U + V$. Тогда

$$X(t) = t - \sin t + Ut + \left(\frac{U}{2} + V\right)\sin t = t - \sin t + U\left(t + \frac{1}{2}\sin t\right) + V\sin t$$

— искомое каноническое разложение, из которого по формуле (8.2) имеем

$$K_X(t_1, t_2) = DU\left(t_1 + \frac{1}{2}\sin t_1\right)\left(t_2 + \frac{1}{2}\sin t_2\right) + DV\sin t_1 \sin t_2,$$

где $DU = 2$,

$$DV = D\left(x_2 + 1 - \frac{u}{2}\right) = D\left(x_2 + 1 - \frac{x_1 + 1}{2}\right) = D\left(x_2 - \frac{x_1}{2} + \frac{1}{2}\right) =$$
$$= D\left(x_2 - \frac{x_1}{2}\right) = Dx_2 + D\frac{x_1}{2} - 2\frac{1}{2}K_{X_1X_2} = 3 + \frac{1}{4}2 - 1 = 2{,}5,$$

поэтому получаем, что

$$K_X(t_1, t_2) = 2\left(t_1 + \frac{1}{2}\sin t_1\right)\left(t_2 + \frac{1}{2}\sin t_2\right) + 2{,}5\sin t_1 \sin t_2,$$

откуда

$$DX(t) = K_X(t, t) = 2\left(t_1 + \frac{1}{2}\sin t_1\right)^2 + 2{,}5\sin^2 t.$$

Задача 8.10. Дана случайная функция $X(t) = t - 3\cos t + u(t + \cos t) + v\cos 2t$, $DU = 1$, $DV = 2$. Найти $MX(t)$, $DX(t)$ и $K_X(t_1, t_2)$.

Решение

По формуле (12.1) $MX(t) = t - 3\cos t$, $K_{UV} = 0$. Тогда по формуле (8.2)

$$DX(t) = (t + \cos t)^2 + 2(\cos 2t)^2;$$
$$K_X(t_1, t_2) = (t_1 + \cos t_1)(t_2 + \cos t_2) + 2\cos t_1 \cos t_2.$$

8.5. Линейные однородные преобразования

На практике часто возникает задача нахождения характеристик преобразованной случайной функции $X(t)$, т.е. новой случайной функции $Y(t) = A\{X(t)\}$, где A — оператор преобразования случайной функции. Оператор A может иметь любой вид: умножение на неслучайную функцию, дифференцирование, интегрирование и т.д.

Операторы A делятся на линейные и нелинейные. Линейные преобразования делятся на однородные и неоднородные.

Определение 8.9. ***Линейным однородным преобразованием*** (обозначается буквой $L0$) называем преобразование, обладающее следующими свойствами:

1) $L_0\{X_1(t) + X_2(t)\} = L_0\{X_1(t)\} + L_0\{X_2(t)\}$;

2) $L_0\{CX(t)\} = CL_0\{X(t)\}$; $L_0(0) = 0$.

Приведем примеры линейных однородных операторов:

а) умножение на неслучайную функцию $f_1(t)$: $Y(t) = f_1(t)X(t)$;

б) дифференцирование: $Y(t) = \dfrac{dX(t)}{dt}$;

в) интегрирование: $Y(t) = \int\limits_0^t X(t)dt$;

г) интегрирование с заданным весом: $Y(t) = \int\limits_0^t X(t)\varphi(t)dt$.

Линейные операторы неоднородны (обозначаются буквой L), если они состоят из линейных операторов с прибавлением неслучайной функции, т.е.

$$L\{X(t)\} = L_0\{X(t)\} + \varphi(t).$$

Примеры линейных неоднородных операторов:

а) $Y(t) = f_1(t)X(t) + f_2(t)$;

б) $Y(t) = \dfrac{dx(t)}{dt} + f(t)$;

в) $Y(t) = \int\limits_0^t X(t)dt + f(t)$.

Числовые характеристики при линейных преобразованиях

1. Умножение случайной функции на неслучайную.

Пусть случайная функция $X(t)$ имеет $MX(t)$ и $K_X(t_1, t_2)$, $Y = f(t)X(t)$, где $f(t)$ — неслучайная функция. Найдем $MY(t)$, $K_Y(t_1, t_2)$:

$$MY(t) = f(t)MX(t);$$

$$K_Y(t_1, t_2) = M((X(t_1)f(t_1) - f(t_1)MX(t_1))(X(t_2)f(t_2) - f(t_2)MX(t_2))) =$$
$$= f(t_1)f(t_2)K_X(t_1, t_2).$$

Если $f(t) = C$ — const, то $MY(t) = CMX(t)$; $K_Y(t_1, t_2) = C^2K_X(t_1, t_2)$.

Определение 8.10. Случайная функция $X(t)$ называется ***непрерывной***, если для любого значения аргумента t и для любого сколь угодно малого значения $\varepsilon > 0$ выполняется равенство

$$\lim_{\Delta t \to 0} P(|X(t + \Delta t) - X(t)| > \varepsilon) = 0.$$

2. Производная от случайной функции.

Пусть случайная функция $X(t)$ непрерывна, для нее существует производная и она имеет $MX(t)$ и $K_X(t_1, t_2)$, $Y = \dfrac{dX(t)}{dt}$. Найдем $MY(t)$ и $K_Y(t_1, t_2)$:

$$MY(t) = \lim_{\Delta t \to 0} M\frac{X(t + \Delta t) - X(t)}{\Delta t} =$$

$$= \lim_{\Delta t \to 0} M\frac{MX(t + \Delta t) - MX(t)}{\Delta t} = \frac{dMX(t)}{dt},$$

т.е. операции математического ожидания и дифференцирования можно менять местами;

$$K_Y(t_1, t_2) = \frac{\partial^2 K_X(t_1, t_2)}{\partial t_1 \partial t_2}.$$

Задача 8.11. Случайная функция $X(t)$ имеет $MX(t) = A\sin t$ и $K_X(t_1, t_2) = DX(t)e^{-\beta(t_1 - t_2)^2}$; $Y(t) = \dfrac{dx(t)}{dt}$. Найти $MY(t)$, $DY(t)$, $K_Y(t_1, t_2)$.

Решение

$$MY(t) = \frac{d}{dt}MX(t) = \alpha A\cos \alpha t;$$

$$K_Y(t_1, t_2) = \frac{\partial^2 K_X(t_1, t_2)}{\partial t_1 \partial t_2} = 2DX(t)\beta e^{-\beta(t_1 - t_2)^2}[1 - 2\beta(t_1 - t_2)];$$

$$DY(t) = 2\beta DX(t).$$

3. Интеграл от случайной функции.

Пусть случайная функция $X(t)$ имеет $MX(t)$ и $K_X(t_1, t_2)$, $Y(t) = \int_0^t X(t)dt$. Тогда

$$MY(t) = \int_0^t MX(t)dt,$$

т.е. операции интегрирования и математического ожидания можно менять местами;

$$K_Y(t_1, t_2) = \int_0^{t_1}\int_0^{t_2} K_X(s_1, s_2)ds_1 s_2.$$

Задача 8.12. Случайная функция $X(t)$ имеет $MX(t) = 2t + 3$, $K_X(t_1, t_2) = \cos t_1 \cos t_2$, $Y(t) = \frac{1}{t}\int_0^t X(s)ds$. Найти $MY(t)$ и $K_Y(t_1, t_2)$.

Решение

$Y(t)$ — случайная функция — результат применения умножения на неслучайную функцию и интегрирования, тогда

$$MY(t) = \frac{1}{t}\int_0^t (2s+3)ds = \frac{1}{t}\left(\frac{2s^2}{2} + 3s\right)\Bigg|_0^t = t + 3;$$

$$K_Y(t_1, t_2) = \frac{1}{t_1}\frac{1}{t_2}\int_0^{t_1}\int_0^{t_2} \cos s_1 \cos s_2 ds_1 ds_2 = \frac{1}{t_1 t_2}\sin t_1 \sin t_2;$$

$$DY(t) = \frac{1}{t^2}\sin^2 t.$$

4. Сложение случайных функций.

Пусть $X(t)$ и $Y(t)$ — случайные функции, имеющие математические ожидания $MX(t)$ и $MY(t)$ и корреляции $K_X(t_1, t_2)$ и $K_Y(t_1, t_2)$, $Z(t) = X(t) + Y(t)$. Тогда

$$MZ(t) = MX(t) + MY(t);$$

$$K_Z(t) = K_X(t_1, t_2) + K_Y(t_1, t_2) + K_{XY}(t_1, t_2) + K_{XY}(t_2, t_1).$$

Если случайные функции $X(t)$ и $Y(t)$ некоррелированы, то

$$K_Z(t_1, t_2) = K_X(t_1, t_2) + K_Y(t_1, t_2).$$

Если случайная функция $X(t)$ задана каноническим разложением $X(t) = MX(t) + \sum_{k=1}^{m} U_k\varphi_k(t)$ и $Y = L\{X(t)\}$, то случайную функцию $Y(t)$ тоже можно представить в виде канонического разложения

$$Y(t) = MY(t) + \sum_{k=1}^{m} U_k\psi_k(t),$$

где $MY(t) = L\{MX(t)\}$, $\psi_k(t) = L_0\{\varphi_k(t)\}$.

Задача 8.13. Случайная функция $X(t)$ имеет $MX(t) = t^2 - 1$ и $K_X(t_1, t_2) = 2e^{-\alpha(t_2-t_1)^2}$. Найти $MY(t)$ и $K_Y(t_1, t_2)$, где $Y(t) = tX(t) + t^2 + 1$, $MZ(t)$ и $K_Z(t_1, t_2)$, где $Z(t) = 2t\frac{dX(t)}{dt} + (1 - t)^2$ и $MU(t)$ и $K_U(t_1, t_2)$, где $U(t) = \frac{d^2X(t)}{dt^2} + 1$.

Решение

$$MY(t) = tMX(t) + t^2 + 1 = t(t^2 - 1) + t^2 + 1 = t^3 + t^2 - t + 1;$$

$$K_Y(t_1, t_2) = t_1 t 2e^{-\alpha(t_2-t_1)^2};\quad MZ(t) = 2t\frac{dMX(t)}{dt} + (1 - t)^2 = 1 - 2t + 5t^2;$$

$$K_Z(t_1, t_2) = 4t_1t_2\frac{\partial K_X(t_1, t_2)}{\partial t_1 \partial t_2} = -\alpha 16 t_1 t_2 [(t_2 - t_1)e^{-\alpha(t_2-t_1)^2} - 2\alpha(t_2 - t_1) -$$

$$- e^{-\alpha(t_2-t_1)}) = 16\alpha t_1 t_2 e^{-\alpha(t_2-t_1)^2}(-2\alpha(t_2 - t_1)^2 + 1);$$

$$MU(t) = \frac{dMX^2(t)}{dt^2} + 1 = 3;$$

$$K_U(t_1, t_2) = \frac{\partial^4 K_X(t_1, t_2)}{\partial t_1^2 \partial t_2^2} = \frac{\partial^2}{\partial t_1 \partial t_2} 4\alpha e^{-\alpha(t_2-t_1)^2}(1 - 2\alpha(t_2 - t_1)^2) =$$

$$= 8\alpha^2 e^{-\alpha(t_2-t_1)^2}[3 + 4\alpha^2(t_2 - t_1)^4 - 12\alpha(t_2 - t_1)^2].$$

Задача 8.14. Случайная функция $X(t)$ имеет характеристики $MX(t) = 0, K_X(t_1, t_2) = \dfrac{1}{1 + (t_2 - t_1)^2}, Y(t) = \int\limits_0^t X(s)ds$. Найти $MY(t), K_Y(t_1, t_2)$.

Решение

$$MY(t) = \int\limits_0^t MX(s)ds = 0;$$

$$K_Y(t_1, t_2) = \int\limits_0^{t_1} ds_1 \int\limits_0^{t_2} K_X(s_1, s_2)ds_2 = \int\limits_0^{t_1}\left(\int\limits_0^{t_2} \frac{ds_2}{1 + (s_2 - s_1)^2}\right)ds_1 =$$

$$= \int\limits_0^{t_1} [\mathrm{arctg}(t_2 - s_1) - \mathrm{arctg}(-s_1)]ds_1 =$$

$$= t_1 \mathrm{arctg}\, t_1 + t_2 \mathrm{arctg}\, t_2 - (t_1 - t_2)\mathrm{arctg}(t_1 - t_2) - \frac{1}{2}\frac{(1 + t_1^2)(1 + t_2^2)}{1 - (1 - t_2)^2}.$$

Задача 8.15. Заданы две случайные функции $X(t)$ и $Y(t)$ с характеристиками $MX(t) = t^2$; $MY(t) = 1$; $K_X(t_1, t_2) = e^{\alpha_1(t_1+t_2)}$; $K_Y(t_1, t_2) = e^{\alpha_2(t_2-t_1)^2}$. Найти характеристики случайной функции $Z(t) = X(t) + tY(t) + t^2$ в случаях:

а) $K_{XY}(t_1, t_2) = 0$;

б) $K_{XY}(t_1, t_2) = ae^{-\alpha(t_2-t_1)}$.

Решение

а) $MZ(t) = MX(t) + tMY(t) + t^2 = 2t^2 + t$;

$K_Z(t_1, t_2) = K_X(t_1, t_2) + t_1t_2K_Y(t_1, t_2) = e^{\alpha_1(t_1+t_2)} + t_1t_2e^{\alpha_2(t_2-t_1)^2}$;

б) $MZ(t)$ то же;

$K_Z(t_1, t_2) = K_X(t_1, t_2) + K_Y(t_1, t_2)t_1t_2 + t_2K_{XY}(t_1, t_2) + t_1K_{XY}(t_1, t_2) =$
$= e^{\alpha_1(t_1+t_2)} + t_1t_2e^{\alpha_2(t_2-t_1)^2} + a(t_1 + t_2)e^{-\alpha|t_2-t_1|}$.

8.6. Спектральное разложение случайной функции

Определение 8.11. ***Спектральным разложением*** называется каноническое разложение стационарной случайной функции вида

$$X(t) = MX(t) + \sum_{k=0}^{\infty}(U_k\cos\omega_k t + V_k\sin\omega_k t), \tag{8.4}$$

где U_k, V_k, $k = 0, 1, ...,$ — центрированные (т.е. $MU_k = MV_k = 0$), некоррелированные случайные величины с попарно равными дисперсиями, т.е. $DU_k = DV_k = D_k$.

Спектральному разложению (8.4) случайной функции $X(t)$ соответствует следующее разложение в ряд ее корреляционной функции:

$$K_X(\tau) = K_X(t_1 - t_2) = \sum_{k=0}^{\infty} D_k\cos\omega_k(t_1 - t_2) =$$

$$= \sum_{k=0}^{\infty} D_k(\cos\omega_k t_1 \cos\omega_k t_2 + \sin\omega_k t_1 \sin\omega_k t_2);$$

$$DX(t) = \sum_{k=0}^{\infty}(\cos^2\omega_k t + \sin^2\omega_k t)D_k = \sum_{k=0}^{\infty} D_k,$$

т.е. дисперсия стационарной случайной функции равна сумме дисперсий всех случайных гармоник ее спектрального разложения.

Определение 8.12. ***Спектральной плотностью*** $S_X(\omega)$ стационарной функции $X(t)$ называется предел отношения дисперсии, приходящейся на данный интервал частот (т.е. разброс частот внутри данного интервала или средний квадрат отклонения частот из данного интервала от своего среднего значения), к длине этого интервала при его стремлении к нулю. Спектральная плотность и корреляция связаны преобразованием Фурье вида

$$S_X(\omega) = \frac{2}{\pi}\int_0^{\infty} K_X(\tau)\cos\omega\tau\, d\tau,$$

где $\tau = t_2 - t_1$; $K_X(\tau)$ — корреляционная функция, $K_X(\tau) = \int_0^{\infty} S_X\cos\omega\tau\, d\omega$ $\left(DX(t) = K_X(0) = \int_0^{\infty} S_X(\omega)d\omega\right)$ [4].

Таким образом, $S_X(t)$ и $K_X(\tau)$ связаны между собой преобразованиями Фурье. В комплексной форме преобразования Фурье, связывающие спектральную плотность и корреляционную функцию, имеют вид

$$S_X^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} K_X(\tau)e^{-i\omega\tau}d\tau; \quad K_X(\tau) = \int_{-\infty}^{\infty} S_X^*(\omega)e^{i\omega\tau}d\omega,$$

где $S_X^*(\omega) = \frac{1}{2}S_X(\omega)$ ($S_X(\omega)$ и $S_X^*(\omega)$ — действительные неотрицательные четные функции).

Задача 8.16. Случайная функция $X(t)$ имеет $MX(t) = 0$; $K_X(\tau) = D_X e^{-\alpha|\tau|}$. Найти ее спектральную плотность $S_X^*(\omega)$.

Решение

$$S_X^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} K_X(\tau)e^{-i\omega^2}d\tau = \frac{1}{\pi}\mathrm{Re}\int_0^{\infty} D_X e^{-(\alpha^2+i\omega t)}d\tau = \frac{D_X}{\pi}\frac{\alpha}{\alpha^2+\omega^2}.$$

Задача 8.17. $X(t)$ — стационарный случайный процесс на отрезке $[0; a]$.

$$K_X(\tau) = \begin{cases} D_X\left(1-\dfrac{\tau}{a}\right), \ 0 \leqslant \tau \leqslant a, \\ 0, \ \tau > 0. \end{cases}$$

Найти спектральную плотность.

Решение

$$S_X(\omega) = \frac{2}{\pi}\int_0^{\infty} K_X(\tau)\cos\omega\tau d\tau = \frac{2}{\pi}\int_0^{a} D_X\left(1-\frac{\tau}{a}\right)\cos\omega\tau d\tau =$$

$$= \frac{2}{\pi}\left(\frac{1}{\omega}\sin\omega\tau\Big|_0^a - \frac{2}{\pi}\frac{D_X}{\omega a}i\sin\omega\tau\Big|_0^a + \frac{2D_X}{\pi\omega a}\int_0^a \sin\omega\tau d\tau\right) = \frac{4D_X}{\pi a\omega^2}\sin^2\frac{\omega a}{2}.$$

Задача 8.18. Случайная функция (телеграфный сигнал) $X(t)$ имеет корреляционную функцию $K_X(\tau) = a^2 e^{-2\lambda|\tau|}$. Найти ее спектральную плотность $S_X(\omega)$.

Решение

$$S_X(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} K_X(\tau)e^{-i\omega\tau}d\tau = \frac{1}{\pi}\int_{-\infty}^{\infty} a^2 e^{-2\lambda|\tau|}e^{-i\omega\tau}d\tau =$$

$$= \frac{a^2}{\pi}\left(\int_{-\infty}^{0} e^{2\lambda\tau - i\omega\tau}d\tau + \int_0^{\infty} e^{-2\lambda\tau - i\omega\tau}d\tau\right) =$$

$$= \frac{a^2}{\pi}\left(\frac{1}{2\lambda - i\omega}e^{2(\lambda - i\omega)\tau}\Big|_{-\infty}^{0} - \frac{1}{2\lambda + i\omega}e^{-2(\lambda - i\omega)\tau}\Big|_0^{\infty}\right) =$$

$$= \frac{a^2}{\pi}\left(\frac{1}{2\lambda - i\omega} + \frac{1}{2\lambda + i\omega}\right) = \frac{4\lambda a^2}{\pi(4\lambda^2 + \omega^2)}.$$

Задача 8.19. Стационарный случайный процесс $X(t)$ имеет корреляционную функцию

$$K_X(\tau) = \begin{cases} 1 - \dfrac{|\tau|}{\tau_0}, |\tau| \leqslant \tau_0, \\ 0, \ |\tau| > \tau_0. \end{cases}$$

Найти спектральную плотность процесса $X(t)$.

Решение

$$S_X(\omega)=\frac{1}{\pi}\int_{-\infty}^{\infty}K_X(\tau)e^{-i\omega\tau}d\tau=\frac{1}{\pi}\int_{-\infty}^{\infty}\left(1-\frac{|\tau|}{\tau_0}\right)e^{-i\omega\tau}d\tau=$$

$$=\frac{1}{\pi}\int_{-\tau_0}^{\tau_0}\left(1-\frac{|\tau|}{\tau_0}\right)e^{-i\omega\tau}d\tau=\frac{1}{\pi}\int_{-\tau_0}^{\tau_0}e^{-i\omega\tau}d\tau-\frac{1}{\pi}\left(\int_{-\tau_0}^{0}\frac{\tau}{\tau_0}e^{-i\omega t}d\tau-\int_{0}^{\tau_0}\frac{\tau}{\tau_0}e^{-i\omega t}d\tau\right)=$$

$$=\frac{1}{-i\pi\omega}e^{-2i\omega\tau_0}-\frac{1}{\pi\tau_0}\left(\int_{\tau_0}^{0}\tau e^{-i\omega\tau}d\tau-\int_{0}^{\tau_0}\tau e^{-i\omega\tau}d\tau\right)=\frac{2}{\tau_0\omega^2}(1-\cos\omega\tau_0).$$

Контрольные вопросы и задания

1. Что такое каноническое разложение случайной функции и в чем преимущество такого ее представления?

2. Как рассчитываются моменты комплексной случайной функции? Приведите примеры.

3. Что такое линейное однородное преобразование случайной функции? Приведите примеры линейных однородных операторов.

4. Как рассчитываются числовые характеристики при линейных преобразованиях? Приведите примеры.

5. Что такое спектральное разложение случайной функции и как определяются ее моменты?

6. Какова связь корреляционной функции и спектральной плотности случайной функции?

Литература

1. *Большев, Л. Н.* Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. — М., 1983.

2. *Боровков, А. А.* Теория вероятностей / А. А. Боровков. — М. : Наука, 1976.

3. *Вентцель, Е. С.* Теория вероятностей и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. — М. : Наука, 1970.

4. *Вентцель, Е. С.* Теория вероятностей / Е. С. Вентцель, Л. А. Овчаров. — М. : Наука, 1973.

5. *Дейвид, Г.* Порядковые статистики / Г. Дейвид. — М. : Наука,1979.

6. *Емельянов, Г. В.* Задачник по теории вероятности и математической статистике / Г. В. Емельянов, В. П. Скитович. — Л. : Изд-во Ленинградского ун-та, 1967.

7. *Ивченко, Г. И.* Задачи с решениями по математической статистике / Г. И. Ивченко, Ю. И. Медведев, В. П. Чистяков. — М., 2007.

8. *Ивченко, Г. И.* Математическая статистика / Г. И. Ивченко, Ю. И. Медведев. — М. : URSS, 2014.

9. *Ивченко, Г. И.* Основные вероятности распределения / Г. И. Ивченко, Н. Ю. Энатская. — М. : МИЭМ, 2008.

10. *Козлов, М. В.* Введение в математическую статистику / М. В. Козлов, А. В. Прохоров. — М. : Изд-во МГУ им. М. В. Ломоносова, 1987.

11. *Крамер, Г.* Математические методы статистики / Г. Крамер. — М. : Мир, 1975.

12. *Козлов, М. В.* Электронные теории вероятности в примерах и задачах / М. В. Козлов. — М. : Изд-во МГУ им. М. В. Ломоносова, 1990.

13. *Леман, Э.* Проверка статистических гипотез / Э. Леман. — М., 1974.

14. *Прохоров, А. В.* Задачи по теории вероятностей / А. В. Прохоров, В. Г. Ушаков, Н. Г. Ушаков. — М. : Наука, 1986.

15. *Розанов, Ю. А.* Теория вероятностей, случайные процессы и математическая статистика / Ю. А. Розанов. — М. : Наука, 1985.

16. *Свешников, А. А.* Сборник задач по теории вероятностей, математической статистике и теории случайных фунций / А. А. Свешников [и др.]. — 5-е изд. — М. : URSS, 2013.

17. *Севастьянов, Б. А.* Сборник задач по теории вероятностей / Б. А. Севастьянов, А. Н. Зубков, В. П. Чистяков. — М. : Наука, 1980.

18. *Смирнов, Н. В.* Курс теории вероятностей и математической статистики для технических приложений / Н. В. Смирнов, И. В. Дунин-Барховский. — М. : Наука, 1965.

19. *Уилкс, С.* Математическая статистика / С. Уилкс. — М. : Наука, 1967.

20. *Феллер, В.* Введение в теорию вероятностей и ее приложения. Т. 1 / В. Феллер. — М., 1984.

21. *Энатская, Н. Ю.* Стохастическое моделирование / Н. Ю. Энатская, Е. Р. Хаммуллин. — М. : МИЭМ, 2012.

Новые издания по дисциплине «Математическая статистика» и смежным дисциплинам

1. *Андрухаев, Х. М.* Теория вероятностей и математическая статистика. Сборник задач : учеб. пособие для прикладного бакалавриата / Х. М. Андрухаев. — 3-е изд., испр. и доп. — М. : Издательство Юрайт, 2016.
2. *Гмурман, В. Е.* Руководство к решению задач по теории вероятностей и математической статистике : учеб. пособие для прикладного бакалавриата / В. Е. Гмурман. — 11-е изд., перераб. и доп. — М. : Издательство Юрайт, 2016.
3. *Гмурман, В. Е.* Теория вероятностей и математическая статистика : учебник для прикладного бакалавриата / В. Е. Гмурман. — 12-е изд. — М. : Издательство Юрайт, 2016.
4. *Далингер, В. А.* Теория вероятностей и математическая статистика с применением Mathcad : учебник и практикум для прикладного бакалавриата / В. А. Далингер, С. Д. Симонженков, Б. С. Галюкшов. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2016.
5. *Ивашев-Мусатов, О. С.* Теория вероятностей и математическая статистика : учебник и практикум для академического бакалавриата / О. С. Ивашев-Мусатов. — 3-е изд., испр. и доп. — М. : Издательство Юрайт, 2016.
6. *Калинина, В. Н.* Теория вероятностей и математическая статистика : учебник для академического бакалавриата / В. Н. Калинина. — 2-е изд., перераб. и доп. — М. : Издательство Юрайт, 2015.
7. *Кацман, Ю. Я.* Теория вероятностей и математическая статистика. Примеры с решениями : учебник для прикладного бакалавриата / Ю. Я. Кацман. — М. : Издательство Юрайт, 2016.
8. *Кремер, Н. Ш.* Теория вероятностей и математическая статистика : учебник и практикум для академического бакалавриата / Н. Ш. Кремер. — 4-е изд., перераб. и доп. — М. : Издательство Юрайт, 2016.
9. *Кремер, Н. Ш.* Теория вероятностей и математическая статистика : в 2 ч. Часть 1. Теория вероятностей : учебник и практикум для академического бакалавриата / Н. Ш. Кремер. — 4-е изд., перераб. и доп. — М. : Издательство Юрайт, 2016.
10. *Кремер, Н. Ш.* Теория вероятностей и математическая статистика : в 2 ч. Часть 2. Математическая статистика : учебник и практикум для академического бакалавриата / Н. Ш. Кремер. — 4-е изд., перераб. и доп. — М. : Издательство Юрайт, 2016.
11. *Попов, А. М.* Теория вероятностей и математическая статистика : учебник для бакалавров / А. М. Попов, В. Н. Сотников. — М. : Издательство Юрайт, 2015.
12. *Сидняев, Н. И.* Теория вероятностей и математическая статистика : учебник для бакалавров / Н. И. Сидняев. — М. : Издательство Юрайт, 2015.
13. *Энатская, Н. Ю.* Теория вероятностей и математическая статистика для инженерно-технических направлений : учебник и практикум для прикладного бакалавриата / Н. Ю. Энатская, Е. Р. Хакимуллин. — М. : Издательство Юрайт, 2016.

Приложение 1

Задачи для самостоятельного решения

Раздел I

Глава 1

1. Случайная точка $(X, Y) \sim R[D]$ (рис. П.1). Найти уравнения линий регрессий.

2. Случайная точка $(X, Y) \sim R[D]$ (рис. П.2). Найти уравнения линий регрессий.

3. Случайная точка $(X, Y) \sim R[D]$ (рис. П.3). Найти уравнения линий регрессий.

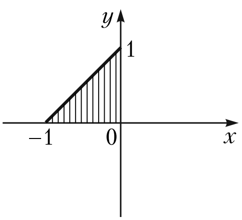


Рис. П.1. **Область 1 для случайной точки**

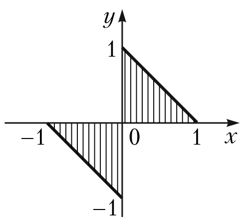


Рис. П.2. **Область 2 для случайной точки**

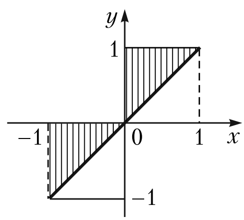


Рис. П.3. **Область 3 для случайной точки**

4. $P(\xi = k) = P(\eta = k) = pk^{k-1}$, $k = 1, 2, ..., p + q = 1$ и СВ ξ, η независимы. Найти $P(\xi = k/\xi > \eta)$. (Указание: $P(\xi > \eta) = P(\xi < \eta)$.)

5. $P(\xi = n) = P(\eta = n) = \dfrac{1}{2^n}$; СВ ξ, η независимы. Найти $P(\xi = n/\xi > \eta)$.

6. СВ $X_1 \sim R[0; 1]$; при $X = x_1$ СВ $X_2 \sim R[0; x_1]$, ...; при $X_{n-1} = x_{n-1}$ СВ $X_n \sim R[0; x_{n-1}]$. Найти MX_n.

7. СВ ξ, η независимы. $P(\xi = k) = P(\eta = k) = pq^k$, $k = 0, 1, 2, ...$, $p + q = 1$. Найти $P(\xi = k/\xi + \eta = l)$ и $M(\xi/\xi + \eta = l)$.

8. $f(x, y) = R\exp\left(-\dfrac{x^2}{4} - \dfrac{xy}{3} + \dfrac{2y^2}{9}\right)$ — плотность двумерного нормального закона. Найти R и уравнения линий регрессии.

9. СВ $X_1 \sim E(\lambda)$; при $X_1 = x_1$ СВ X_2 $E(\lambda/x_1)$, ...; при $X_{n-1} = x_{n-1}$ СВ $X_n \sim E(\lambda/x_{n-1})$. Найти MX_n.

10. СВ X и Y независимы, $X \sim N(0, \sigma = 1/\sqrt{2\pi})$, $Y \sim R[0; 1]$. Найти $f(x, y)$, $F(x, y)$ и уравнения линий регрессии.

11. СВ $X \sim R[a; a + 1]$, $a > 0$; при $X = x$ СВ $Y \sim R[x_1; 1]$. Найти регрессию СВ X по Y.

12. СВ $\xi_1, ..., \xi_n$ независимы и одинаково распределены по закону $B(n, p)$. Найти $M(\xi_1 + ... + \xi_k / \xi_1 + ... + \xi_N = M)$ при $k < N$. (Указание: если $P(X = l) = C_a^l C_b^{k-p} / C_{a+b}^k$, то $MX = \dfrac{a}{a+b} k$.)

13. СВ $X \sim R[0; 1]$; при $X = x$ СВ $Y \sim R[x; 2x]$. Найти уравнения линий регрессии.

14. СВ $\xi_1, ..., \xi_n$ независимы и одинаково распределены по закону $\Pi(\lambda)$. Найти $P(\xi_1 + ... + \xi_n = m / \xi_1 + ... + \xi_N = K)$ при $n < N$.

15. По мишени произведено два независимых выстрела. СВ X — число попаданий, СВ Y — число промахов. Найти $F(x, y)$ и $M(X/Y)$.

16. Найти плотность совместного распределения k-й, l-й и m-й ($1 \leqslant k < l < m \leqslant n$) порядковых статистик.

17. СВ $Y \sim R[a; b]$. Найти функцию распределения и плотность распределения размаха выборки.

18. Найти плотность совместного распределения всех четных порядковых статистик при четном объеме выборки n.

19. Найти совместное распределение первых r порядковых статистик: а) СВ $X \sim R[a; b]$; б) СВ $X \sim E(\lambda)$.

20. Случайные величины X и Y имеют соответственно функции распределения $F_1(x)$ и $F_2(y)$; $Y = \varphi(x)$, где $\varphi(x)$ — монотонно: а) возрастающая, б) убывающая функция. Найти вид функции $\varphi(x)$.

21. X — СВ непрерывного типа с функцией распределения $F(x)$. Найти закон распределения СВ $Y = F(x)$.

22. Смоделировать 20 бросаний игральной кости. Построить статистический ряд, вариационный ряд числа выпавших очков, построить полигон частот, эмпирическую функцию распределения, выборочную среднюю и выборочную дисперсию.

23. Смоделировать 20 раз по 20 значений СВ $X \sim R[3; 8]$. Построить эмпирическую функцию распределения третьей порядковой статистики.

24. Смоделировать 50 значений СВ $X \sim N(5, 2)$. Найти выборочные моменты $\overline{x}$ и S_1^2.

Глава 2

25. СВ $X \sim R[0; \theta]$. Для $\tau(\theta) = \theta$ сравнить по квадратичному риску следующие оценки для $\tau(\theta)$: $t_1 = 2x_1$, $t_2 = 2\overline{x}$, $t_3 = x_{(n)}$. Исследовать оценки на несмещенность. В случае линейности смещения подправить оценки.

26. Исследовать на несмещенность оценки $\dfrac{x_{(1)} + x_{(n)}}{2}$ и $x_{(n)} + x_{(1)}$ соответственно для середины и длины отрезка $[a; b]$, если СВ $X \sim R[a; b]$. Смещенные оценки подправить.

27. СВ $X \sim \Gamma_{\theta,\alpha,}$ где θ неизвестно, а α — известно, с плотностью распределения $f(x; \alpha, \theta) = \dfrac{\theta^\alpha x^{\alpha-1} e^{-\theta x}}{\Gamma(\alpha)}$, $x \geqslant 0, \alpha, \theta > 0$. В качестве оценки для θ предлагается $\tilde{\theta} = \dfrac{n\alpha}{\sum\limits_{i=1}^{n} x_1}$. Исследовать оценку $\tilde{\theta}$ на смещенность. В случае смещенности подправить ее. (Указание: $\sum\limits_{i=1}^{n} x_i \sim \Gamma_{\theta, n\alpha}$.)

28. СВ $X \sim R[\theta; 1]$, где θ неизвестно. Смещена ли оценка $\tilde{\theta} = x_{(1)}$ для θ? В случае смещения подправить оценку $\tilde{\theta}$.

29. $\vec{x} = (x_1, \dots, x_n)$ — выборка значений СВ X. Для дисперсии СВ X взята оценка $\tilde{\sigma}^2 = \dfrac{1}{2(n-1)} \sum\limits_{j=1}^{n-1} (x_{j+1} - x_j)^2$. Смещена ли она?

30. СВ X имеет плотность распределения $f(x) = e^{\theta - x}$, $x > \theta$. Проверить оценку $\tilde{\theta} = x_{(1)}$ для параметра θ на смещенность. В случае смещенности подправить ее.

31. В качестве оценки для MX^2 взята $\tilde{m}^2 = \left(\dfrac{1}{n} \sum\limits_{i=1}^{m} x_i\right)^2$. Смещена ли она? В случае смещенности подправить ее.

32. В качестве оценки для дисперсии взята оценка $\tilde{\sigma}^2 = k \sum\limits_{j=1}^{n-1} (x_{j+1} - x_j)^2$. Каким должно быть k, чтобы оценка $\tilde{\sigma}^2$ была несмещенной, если математическое ожидание $MX = m$ известно.

33. Найти достаточные статистики двумя способами: по критерию факторизации и по параметрам экспоненциального семейства для следующих семейств распределений: а) СВ $X \sim B(1, \theta)$; б) СВ $X \sim G(\theta)$; в) СВ $X \sim N(\theta, \sigma)$; СВ $X \sim \Gamma_{\alpha,\theta}$; д) СВ $X \sim N(a, \theta = \sigma)$, где θ — неизвестный параметр распределений.

34. СВ X имеет плотность распределения $f(x) = \theta(\theta x)^{k-1} e^{-\theta x}/(k-1)!$, $x > 0, \theta > 0$. Найти достаточную статистику для данного семейства распределений с неизвестным параметром θ.

35. СВ $X \sim R[-3\theta + 1; 3\theta + 3]$. Доказать, что статистика $T = (x_{(1)}, x_{(n)})$ достаточна, но не полна.

36. СВ $X \sim B(1, \theta)$, где θ — неизвестный параметр. Показать, что статистика $S = \left(s_1 = x_1; s_2 = \sum\limits_{i=2}^{l} x_i; s_3 = \sum\limits_{i=l+1}^{n} x_i\right)$ достаточна, но неполна.

37. СВ X имеет плотность распределения $f(x) = \dfrac{\lambda e^{-\lambda x}}{1 - e^{-\lambda\theta}}$, $0 < x < \theta$, θ — неизвестный параметр. Найти достаточную статистику.

38. СВ $X \sim B(1, \theta)$, где параметр θ неизвестен. Найти достаточную статистику и проверить ее достаточность по определению.

39. СВ $X \sim N(\mu, \theta = \sigma)$, где θ — неизвестный параметр,

$$S = \left\{ s_1 = \sum_{i=1}^{2}(x_i - \mu)^2;\ s_2 = \sum_{i=3}^{6}(x_i - \mu)^2;\ s_3 = \sum_{i=7}^{n}(x_i - \mu)^2 \right\}.$$

Показать, что статистика S достаточна, но неполна.

40. Для семейств распределений $B(1, \theta)$, где θ — неизвестный параметр, улучшить оценку $\tilde{\theta} = x_1$ для $\tau(\theta) = \theta$ по теореме Блэкуэлла.

41. Для семейств распределений $B(1, \theta)$, где θ — неизвестный параметр, найти, для каких функций от θ оптимальными оценками являются следующие: а) $2\dfrac{r(r-1)}{n(n-1)}$; б) $\dfrac{r(n-r)}{2n(n-1)}$, где $r = \sum_{i=1}^{n} x_i$.

42. Для каких функций от неизвестных параметров θ_1 и θ_2 при СВ $X \sim R[\theta_1; \theta_1]$ будут оптимальными оценки: а) $\dfrac{x_{(1)} + x_{(n)}}{2}$ и б) $\dfrac{n+1}{n-1}(x_{(n)} - x_{(1)})$?

43. СВ $X \sim N(\mu, \theta)$, где θ — неизвестно, μ — известно. По критерию найти эффективную оценку для $\tau(\theta) = \theta^2$ и ее дисперсию.

44. СВ $X \sim B(m, \theta)$, где θ — неизвестно, m — известно. Найти нижнюю границу дисперсий несмещенных оценок для $\tau(\theta) = e^{-\theta}$ по неравенству Рао — Крамера.

45. СВ $X \sim Ge(\theta)$, где θ — неизвестный параметр. Найти нижнюю границу дисперсий несмещенных оценок для $\tau(\theta) = \theta^2$ по неравенству Рао — Крамера.

46. СВ X имеет плотность распределения $f(x) = e^{\theta - x}$, $x \geqslant 0$. Состоятельна ли оценка $\tilde{\theta} = x_{(1)} - \dfrac{1}{n}$ для $\tau(\theta) = \theta$?

47. СВ $X \sim B(m, \theta)$, где θ — неизвестно, m — известно. По критерию эффективности найти эффективную оценку для $\tau(\theta) = 2\theta + 1$ и ее дисперсию.

48. СВ $X \sim B(1, \theta)$, где θ — неизвестно. Найти нижнюю границу дисперсий несмещенных оценок для $\tau(\theta) = \theta(1 - \theta)$ по неравенству Рао — Крамера.

49. СВ а) $X \sim B(1, \theta)$ и б) $X \sim \Pi(\theta)$, где θ — неизвестно. Найти дисперсии эффективных оценок для $\tau(\theta) = \theta$ сведением к виду ЭС по формуле для ЭС.

50. СВ $X \sim B(m, \theta)$, где θ — неизвестно, m — известно. Для функций параметра θ вида $\tau_1 = \theta$, $\tau_2 = 3\theta - 2$, $\tau_3 = 3 - \frac{2}{\theta}$ существуют ли эффективные оценки? Если да, вычислить их дисперсии.

51. СВ $X \sim R[\theta + 1; 3\theta + 4]$. Найти ОММ и ОМП для параметра θ.

52. СВ $X \sim \Gamma_{\alpha,\lambda}$. Считая параметр λ известным, найти ОММ и ОМП для параметра λ.

53. СВ $X \sim R[2\theta_1 - 1; 5\theta_2 + 3]$. Найти ОММ и ОМП для параметра $\theta = (\theta_1, \theta_2)$.

54. СВ $X \sim R[\theta - m; \theta + m]$, где θ — неизвестно, m — известно. Найти ОММ и ОМП для параметра θ.

55. Построить доверительный интервал для параметра θ с уровнем доверия γ в случаях: а) СВ $X \sim R[a\theta - b; c\theta + d]$; б) $X \sim E(a\theta + b)$.

56. Найти минимальный объем выборки, при котором с надежностью 0,975 точность математического ожидания a по выборочной средней $\overline{x}$, $\delta = 0{,}4$, если $X \sim N(a, \sigma)$, $\sigma = 1{,}5$.

Глава 3

57. СВ $X \sim B(1, \theta)$. Проводится четыре наблюдения значений СВ X $\vec{x} = (x_1, x_2, x_3, x_4)$. Проверяется основная гипотеза H_0: $\theta = 0{,}1$ против альтернативы H_1: $\theta = 0{,}4$. Пусть построен критерий $\rho(x) = \begin{cases} H_0, \sum_{i=1}^{4} x_i \leqslant 1, \\ H_1, \sum x_i > 1. \end{cases}$

Найти вероятности ошибок первого и второго рода (α и β) и мощность критерия W. При $n = 10$ и $\alpha = 0{,}05$ построить аналогичный критерий и найти β и мощность.

58. СВ $X \sim R[a; b]$. Предложить критерий проверки гипотез H_0: $(a = 0, b = 2)$ против альтернативы H_1: $(a = 1, b = 3)$, для которого вероятности ошибок первого и второго родов стремятся к нулю при $n \to \infty$ (n — объем выборки).

59. Рассчитать мощность критерия Неймана — Пирсона для различения простых гипотез о параметрах законов распределения наблюдаемой СВ X при заданном размере критерия $\alpha = 0{,}05$ в случаях: а) СВ $X \sim \Pi(\lambda)$; H_0: $(\lambda = \lambda_0 = 3)$; H_1: $(\lambda = \lambda_1 = 4)$; б) СВ $X \sim N(\mu, \sigma = 1)$; H_0: $(\mu = \mu_0 = 2)$; H_1: $(\mu = \mu_1 = 1{,}8)$; в) СВ $X \sim B(1, p)$; H_0: $(p = p_0 = 0{,}4)$; H_1: $(p = p_1 = 0{,}5)$; г) СВ $X \sim B(m, p)$, $m = 5$; H_0: $(p = p_0 = 0{,}6)$; H_1: $(p = p_1 = 0{,}7)$; д) СВ $X \sim N(0, \sigma)$; H_0: $(\sigma = \sigma_0 = 1)$; H_1: $(\sigma = \sigma_1 = 2)$.

Раздел II

Глава 4

1. N нумерованных шаров случайным образом размещены по двум урнам. На каждом шаге шар со случайно выбранным номером

перекладывают в другую урну. Состояние процесса определяется числом шаров в первой урне. Написать матрицу переходных вероятностей P процесса. Для $N = 3$ выписать матрицы P и $P(2)$ из вероятностных соображений.

2. Из таблицы чисел от 1 до m выбирают наугад числа по одному с возвращением. Система находится в состоянии E_k, если k — наибольшее выбранное к этому шагу число. Построить матрицу переходных вероятностей P процесса. Для $m = 4$ выписать матрицы P и $P(2)$. Объяснить $P(2)$ из вероятностных соображений.

3. Шары случайным образом размещают по одному по a ящикам равновероятно. Система находится в состоянии E_k на i-м шаге, если после размещения i шаров окажется k занятых ящиков. Найти матрицу переходных вероятностей P процесса. Для $a = 3$ выписать матрицы P и $P(2)$.

4. На стоянку такси равномерно прибывают машины через некоторую единицу времени. Если пассажиров нет, машина уезжает. Система находится в состоянии E_k, если в очереди k пассажиров. Вероятность того, что за единицу времени в очередь встанут еще k пассажиров: $P(\xi = k) = a_k$, $a_k \neq 0$; $\sum\limits_{k=0}^{\infty} a_k = 1$. Написать матрицу P переходных вероятностей процесса. Показать, что это цепь Маркова.

5. Система находится в состоянии E_k в момент времени n, если последняя неудача была в $(n - k)$-м испытании. В каждом опыте вероятность успеха есть p, а неуспеха q; $p + q = 1$. Написать матрицу P переходных вероятностей процесса, а также $P(2)$ и $P(3)$.

6. Автомашина перевозит грузы между $2m$ пунктами по кольцевой трассе. Грузы перевозят только в соседние пункты: в предыдущий с вероятностью q и в последующий с вероятностью p $(p + q = 1)$. Показать, что движение машины управляется цепью Маркова с матрицей P. Выписать матрицы P и $P(2)$ при $m = 2$.

Глава 5

7. Цепь Маркова задана матрицей переходных вероятностей

$$P = P(1) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Дать классификацию состояний цепи. Найти предельную матрицу, если она существует.

8. Блуждание частицы происходит по схеме, изображенной на рис. П.4. Написать матрицу переходных вероятностей P. Дать клас-

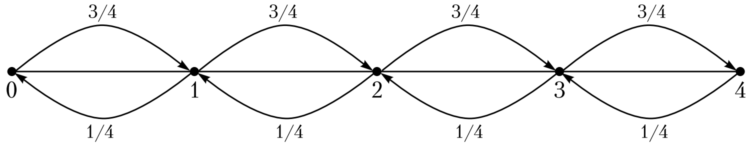


Рис. П.4. **Схема блуждания частицы**

сификацию состояний цепи. Найти матрицу финальных вероятностей, если она есть.

9. Матрица переходных вероятностей

$$P = \begin{pmatrix} 1/2 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Дать классификацию состояний цепи.

10. Цепь Маркова задана матрицей переходных вероятностей

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Дать классификацию состояний цепи. Найти предельную матрицу, если она существует.

11. Матрица переходных вероятностей

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 1 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Эргодична ли цепь Маркова с матрицей P? Найти предельную матрицу, если она существует.

12. Привести пример цепи Маркова со счетным множеством состояний, у которой все состояния невозвратные нулевые.

13. Доказать, что для конечной неприводимой непериодической цепи Маркова с дважды стохастической матрицей вероятностей перехода финальные вероятности равны $\left(\frac{1}{m}, ..., \frac{1}{m}\right)$, где m — порядок матрицы.

14. Показать, что в неприводимой конечной цепи Маркова нет нулевых состояний.

15. Показать, что $\sum_{n=1}^{\infty} p_{ij}(n)$ и $\sum_{n=1}^{\infty} p_{jj}(n)$ сходятся или расходятся одновременно ($p_{ij}(n)$ — вероятность перехода из состояния E_i в состояние E_j за n шагов в цепи Маркова).

16. Доказать, что все состояния конечной неприводимой периодической цепи Маркова: а) можно разбить на d непересекающихся подклассов $S_0, S_1, ..., S_{d-1}$; б) с вероятностью 1 из класса S_k переходим за один шаг в класс S_{k+1}, $k = 0, 1, ..., d-2$, из класса S_{d-1} — в S_0.

17. Показать, что для периодической конечной цепи Маркова последовательность степеней матрицы вероятностей перехода P, P^2, P^3, ..., P^n можно разбить на d последовательностей, таких что каждая подпоследовательность имеет предельную матрицу.

18. Показать, что периодическая конечная цепь Маркова не будет эргодической.

19. Показать, что разложимая цепь Маркова не будет эргодической, если число классов существенных сообщающихся состояний больше одного.

20. Показать, что поглощающая цепь Маркова будет эргодической, когда она содержит только одно поглощающее состояние.

21. Можно ли определить эргодическую цепь как цепь, все состояния которой эргодические?

22. Всегда ли у эргодической цепи все состояния эргодические?

23. Когда конечная эргодическая цепь Маркова будет эргодической?

24. Показать, что $\sum_{n=1}^{\infty} p_{jj}(n)$ есть математическое ожидание числа возвращений в состояние E_j.

25. Доказать, что для непериодичности неприводимой цепи Маркова достаточно, чтобы хотя бы один диагональный элемент матрицы переходных вероятностей был больше нуля.

26. Доказать, что для эргодической цепи Маркова условные вероятности перехода в состояние E_j из состояния E_i за n шагов с ростом n стремятся к безусловной вероятности через n шагов быть в состоянии E_j.

27. Доказать, что неприводимая периодическая цепь Маркова не может быть эргодической.

28. Доказать, что для приводимой конечной цепи Маркова поглощение одним из замкнутых классов есть достоверное событие.

29. Доказать, что возвращенное состояние всегда существенно.

30. Доказать, что в эргодической цепи финальные вероятности для несущественных состояний равны нулю.

31. Доказать, что несущественные состояния всегда невозвратны.

32. Доказать, что конечная цепь Маркова должна содержать хотя бы один замкнутый класс состояний.

33. Доказать, что любой замкнутый класс состояний приводимой цепи Маркова можно рассматривать как неприводимую цепь Маркова.

34. Доказать, что в неприводимой цепи Маркова нет нулевых состояний и все состояния не могут быть невозвратными.

35. Доказать, что в конечной цепи Маркова существенное состояние всегда возвратно.

36. Доказать, что в конечной цепи Маркова невозвратные состояния — несущественные.

37. Доказать, что цепь Маркова, состоящая из более чем одного замкнутого класса, не может быть эргодической.

38. К какому виду может быть приведена матрица переходных вероятностей приводимой цепи Маркова, состоящей из m замкнутых классов и l несущественных состояний?

39. Доказать, что если состояние E_i возвратно и состояние E_j достижимо из состояния E_i, то состояния E_i и E_j — сообщающиеся состояния.

40. Доказать, что если состояние E_j возвратное, то все состояния, достижимые из него, образуют замкнутый класс состояний.

41. Доказать, что для неприводимой непериодической конечной цепи Маркова найдется некоторое n, такое что все элементы матрицы $P(n)$ больше нуля.

42. Доказать, что если для неприводимой конечной цепи Маркова существует такое n, что все элементы матрицы $P(n)$ строго положительны, то цепь непериодическая.

43. Доказать, что состояния замкнутого класса существенных состояний имеют одинаковый период.

44. Доказать, что в конечной цепи Маркова порядка n любое состояние E_i или может быть достигнуто из состояния E_j за не более чем n шагов, или вообще не может быть достигнуто.

45. Доказать, что в бесконечной дважды стохастической цепи Маркова все состояния или нулевые, или невозвратные.

Глава 6

46. $\{\xi_i\}$ — ветвящийся процесс, $\xi_0 = 1$; $P\{\nu = 0\} = \dfrac{1}{4}$; $P\{\nu = 3\} = \dfrac{3}{4}$. Найти $\varphi_1(s)$, $\varphi_2(s)$, $M\xi_n$, $D\xi_n$ и вероятность вырождения процесса. Найти первые три строки матрицы P и первые две строки матрицы $P(2)$.

47. $\{\xi_i\}$ — ветвящийся процесс, $\xi_0 = 1$; $\nu \sim B(n, p)$, $n = 4$, $p = \dfrac{1}{4}$. Найти $\varphi_1(s)$, $\varphi_2(s)$, $M\xi_n$, $D\xi_n$, вероятность вырождения, первые две строки матриц P и $P(2)$.

48. $\xi_1(t)$ и $\xi_2(t)$ — пуассоновские процессы. Является ли пуассоновским процесс: а) $\xi(t) = \xi_1(t) - \xi_2(t)$; б) $\xi(t) = \xi_1(t) + C$, где C — const?

49. Среднее число вызовов, поступающих в час на телефонную станцию, равно 30. Считая поток вызовов пуассоновским, определить вероятность того, что за минуту поступит менее двух вызовов. Выписать плотность распределения в потоке.

50. В пуассоновском потоке телеграмм среднее их число в час равно 3. Найти вероятность того, что с 8:00 до 12:00 часов не поступит ни одной телеграммы. Выписать плотность распределения в потоке.

Глава 7

51. Физическая система проходит через последовательность состояний $E_0 \to E_1 \to E_2 \to ...$, причем время пребывания в состоянии E_k имеет плотность распределения $\lambda_k e^{-\lambda_k x}$, $x \geqslant 0$ ($\lambda_k \neq \lambda_i$ при $i \neq k$). Написать систему дифференциальных уравнений для состояний системы. Найти вероятность $\Pi_0(t)$ того, что в момент времени t система находится в состоянии E_0.

52. Телефонный узел имеет m каналов. Моменты поступления вызовов образуют пуассоновский поток с параметром λ. Интенсивность обслуживания — μ. Составить систему дифференциальных уравнений вероятностей чисел занятых каналов.

53. Автоматическая телефонная станция имеет четыре линии связи, плотность потока заявок — λ, а потока обслуживания — μ. Написать систему дифференциальных уравнений для вероятностей чисел занятых каналов, найти их стационарное распределение.

54. По двум линиям связи в один пункт поступают два независимых пуассоновских потока телеграмм. Найти вероятность того, что за время t в пункт поступит n телеграмм, если параметры потоков λ_1 и λ_2.

Глава 8

55. $Z(t) = \xi \cos \lambda t + \eta \sin \lambda t$ — случайный процесс; ξ и η — некоррелированные СВ, $m_\xi = m_\eta = 0$; $D_\xi = D_\eta = 1$, λ — const. Найти числовые характеристики процесса $Z(t)$. Будет ли процесс $Z(t)$ стационарным в широком смысле?

56. $\xi(t) = \xi_1(t) + i \sin \xi_2(t)$ — случайный процесс; а) $\xi_1(t)$ и $\xi_2(t)$ — некоррелированные случайные функции с $m_{\xi_1} = t^2$, $m_{\xi_2} = 1$, $K_{\xi_1}(t, s) = e^{-\alpha_1 (t-s)^2}$, $K_{\xi_2}(t, s) = e^{2\alpha_2 (t+s)}$. Найти $m_\xi(t)$, $D_\xi(t)$ и $K_\xi(t, s)$; б) $K_{\xi_1 \xi_2}(t, s) = e^{\beta (t+s)^2}$. Будет ли $\xi(t)$ стационарен в широком смысле?

57. $X(t) = Ut + Vt^2$; а) U и V некоррелированы, $MU = 3$, $MV = 0{,}5$; $DU = 1$; $DV = 0{,}05$. Найти $MX(t)$, $DX(t)$, $K_X(t, s)$. Стационарен ли $X(t)$? б) U и V имеют корреляционную матрицу $\begin{pmatrix} 1 & 3 \\ & 2 \end{pmatrix}$. Найти $MX(t)$, $DX(t)$, $K_X(t, s)$.

58. Комплексная случайная функция $Z(t) = \xi(t) + i\eta(t)$, где $\xi(t) = (a + v)e^{-\alpha t}$; $\eta(t) = (b + u)e^{-\beta t}$, a и b — const. Случайный вектор (U, V) имеет математическое ожидание $(0, 0)$ и корреляционную матрицу $\begin{pmatrix} 1 & 1 \\ & 2 \end{pmatrix}$. Найти характеристики случайной функции $Z(t)$.

59. Случайная функция $\xi(t)$ имеет характеристики $m_\xi(t) = t^2 - 1$, $K_\xi(t, s) = 2e^{-\alpha(s-t)^2}$. Определить характеристики случайной функции $\eta(t) = t\xi(t) + t^2 + 1$; $Z(t) = 2t\dfrac{d\xi(t)}{dt} + (1 - t)^2$. Являются ли $\eta(t)$ и $Z(t)$ стационарными в широком смысле?

Приложение 2

Работа в статистическом пакете программ «Статистика»

Моделирование выборок и их статистическая обработка

Задание 1. Смоделировать выборки $m \times n$ (m — число столбцов (*vairs*), n — число строк (*cases*)) с заданным законом распределения.

Выполнение

Путь: *File New Data* → таблица 10×10.

Если ее нужно расширить по строкам или столбцам — щелкнуть левой кнопкой «мыши» по *cases* или *vairs* → *add* → в окне *how many* ввести количество добавлений — *OK* (другой путь для расширения таблицы: *copy* → *File* → *Nove* → указать число строк и столбцов) → → зафиксировать заполняемый столбец левой кнопкой «мыши» по названию *vari* (он станет черным) → левой кнопкой «мыши» по названию щелкнуть два раза → в окне *Name* задать название столбца с пульта (обычно сокращенное название моделируемого распределения с номером выборки) → в нижнем окне *Long name* (*Functions*) набрать операцию моделирования (см. [21]):

а) методом обратных функций задать название моделируемого распределения (*Rnd* (1) и параметры моделируемого распределения)

Пример: = *V Normal* (*Rnd* (1), 1; 2) — это распределение $N(1, 2)$;

б) методом связи распределений (мер). Тогда формула моделирования выражается через ранее сформулированные выборки.

Пример: = $N1 \wedge 2 + N2 \wedge 2$ — получение выборки из распределения χ^2 с двумя степенями свободы, где $N1$ и $N2$ — выборки из $N(0, 1)$.

Задание 2. Построить вариационный ряд. Вычислить среднее и стандартное отклонения.

Выполнение

Выбрать столбец наблюдений (щелкнув левой кнопкой «мыши» по заголовку) → *Graph of Input Data* → *values / stats* → в окне *values / stats* : *variable* появится вариационный ряд, среднее (*mean*) и стандартное отклонение (*SD*).

Задание 3. Построить гистограмму частот.

Выполнение

Вариант 1. Путь: *Graphs* → *2D Graphs* → *Histograms* → *Quick* → → *Variables* → *Select Variables of Histograms* → выделить столбец (щелчком левой кнопкой «мыши» по названию) → *OK* → *Advenced* → *Graph Type* → *Regular* → *Fit Type* → *off* → *Intrevals* → *Auto* – ставим галку (автоматический выбор числа интервалов) или *Categories* – число интервалов → *OK*.

Вариант 2. Щелчком левой кнопкой «мыши» по названию выделяем столбец → *Graphs* → *Histograms* → *Advenced* → распределение, которое ожидается, например *Gamma* → *Standart* → *OK*.

Задание 4. Построить эмпирическую функцию распределения.

Выполнение

Вариант 1. Путь: *Graphs* → *2D Graphs* → *Histograms* → *2D Histograms* → *Quick* → *Variables* → *Select Variables of Histograms* → → выбираем столбец → *OK* → *Advenced* → *Graph Type* → *Regular* → → в *Fit Type off* или нужный нам тип → в *Intrevals* ставим галку против *Auto* или число интервалов в *Categories* → (в меню) *Showing Type cumulative* → *OK*.

Вариант 2. Щелчком левой кнопкой «мыши» по названию выделяем столбец → *Graphs* → *Histograms* → *Advenced* → ожидаемое распределение, например *Gamma* → *Standart* → *cumulative* → *OK*.

Задание 5. Получить выборочные характеристики.

Выполнение

Путь: *Statistica* → *Basic Statistic / Tables* → *Discriptive Statistics* в формах галками выбираем нужные характеристики, т.е. в *Lactions*, *valid N*; *variation*, *moments*; *Percentiles*, *ranges* – *mean* – среднее; *Conf. Limits for means* – 95% доверительные границы верхняя и нижняя (с уровнем доверия 0,95), *Sum* – сумма, *varance* – дисперсия и т.д. → *Summary* – получаем таблицу требуемых характеристик.

Задание 6. Проверить согласие выборки с данным законом распределения по критерию χ^2 Пирсона с уровнем доверия $1 - \alpha$.

Выполнение

Путь: *Statistic Distribution Fitting* → выбираем тип распределения → → в окне *Fitting Continuous Distributions* указываем имя столбца (проверяемого) → *Plot of observed und expected distributions* → появляется рисунок гистограммы, значение χ^2 и уровень доверия γ, с которым это значение $\chi^2 <$ критического. Тогда если $\gamma < 1 - \alpha$, то гипотеза о согласии выборки с данным распределением принимается с уровнем доверия $1 - \alpha$.

Учебное издание

Энатская Наталия Юрьевна

**МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
И СЛУЧАЙНЫЕ ПРОЦЕССЫ**

Учебное пособие для академического бакалавриата

Формат 60×90 1/16.
Гарнитура «PetersburgC». Печать офсетная.
Усл. печ. л. 00,00.

ООО «Издательство Юрайт»
111123, г. Москва, ул. Плеханова, д. 4а.
Тел.: (495) 744-00-12. E-mail: izdat@urait.ru, www.urait.ru