УКРАИНСКАЯ ГОСУДАРСТВЕННАЯ АКАДЕМИЯ
ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА

ДОНЕЦКИЙ ИНСТИТУТ ЖЕЛЕЗНОДОРОЖНОГО
ТРАНСПОРТА

Кафедра высшей математики

Ю.М. ВОЛЧЕНКО

КОНСПЕКТ ЛЕКЦИЙ

по дисциплине

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Донецк

2007

Волченко Ю. М. Теория вероятностей и математическая статистика. Конспект лекций. – Донецк: ДонИЖТ, 1998-2007. – 325 с.

Методическое пособие содержит материал лекций, которые автор читал студентам специальности «Автоматика и автоматизация на транспорте», начиная с 1983 года. Поэтому и теоретические вопросы, и примеры решения задач в первую очередь ориентированы на упомянутую специальность. Исключение составляют лекции, посвященные элементам теории массового обслуживания, которые слушались студентами специальности «Организация перевозок и управление на железнодорожном транспорте». Пособие рассчитано на уровень подготовки студентов железнодорожного вуза, содержит необходимый теоретический материал, снабженный в большинстве случаев доказательствами, примеры и используемые в расчетах таблицы.

Ил. 53, табл. 36, библиогр. 28:

Конспект лекций рассмотрен и рекомендован к публикации на заседаниях кафедры высшей математики и физики 15 сентября 1999 г., протокол № 1; 2 октября 1999 г., протокол № 2; 15 сентября 2000 г., протокол № 1.

Рецензенты:

Проф., д. ф.-м.н. О.О. Игнатьев,
доц., к.э.н. А.М.Кремлина,

Содержание

I Теория вероятностей **9**

1 Случайные события и их вероятности **10**
1.1 Предмет теории вероятностей 10
1.2 Случайные события 13
1.3 Вероятности случайных событий 16
1.4 Классическое определение вероятности 19

2 Геометрическая и условная вероятности **23**
2.1 Геометрическая вероятность 23
2.2 Условная вероятность. Независимость с.с. 25
2.3 Формула полной вероятности 29
2.4 Формула Байеса 31

3 Случайные величины **33**
3.1 Функция распределения 33
3.2 Непрерывные с.в. 36
3.3 Дискретные с.в. 39

4 Случайные векторы **45**
4.1 Случайный вектор и его функция распределения 45
4.2 Случайный вектор непрерывного типа 48
4.3 Случайный вектор дискретного типа 52

5 Независимость с.в. 56
5.1 Независимость с.в. 56
5.2 Функция непрерывного аргумента 60
5.3 Функция дискретного аргумента 62

6 Функции нескольких с.в. 65
6.1 Функция СВНТ 65
6.2 Функция СВДТ 68
6.3 Сумма случайных величин 69

7 Математическое ожидание 73

8 Дисперсия. Ковариация. 82
8.1 Дисперсия с.в. 82
8.2 Ковариация 88
8.3 Коэффициент корреляции 90
8.4 Линейная регрессия 91
8.5 Статистическая линеаризация нелинейностей 94

9 Типичные дискретные распределения 98
9.1 Биномиальное распределение 98
9.2 Полиномиальное распределение 102
9.3 Пуассоновское распределение 107
9.4 Многомерное пуассоновское распределение 108
9.5 Геометрическое распределение 111
9.6 Гипергеометрическое распределение 112

10 Типичные непрерывные распределения 114
10.1 Равномерное распределение 114
10.2 Нормальное распределение 116
10.3 Многомерное нормальное распределение 121
10.4 Двумерное нормальное распределение 128
10.5 Показательное распределение 130
10.6 Распределение Эрланга k-го порядка 133
10.7 Нормированное распределение Эрланга 136

10.8 t-распределение Стьюдента 138
10.9 χ^2-распределение Пирсона 141

11 Предельные теоремы **144**
11.1 Предельные теоремы для схемы Бернулли 144
11.2 Закон больших чисел и теорема Бернулли 146
11.3 Центральная предельная теорема 152

II Случайные процессы **155**

12 Некоторые типы с.п. **156**
12.1 Основные определения 156
12.2 Процессы с независимыми приращениями 158
12.3 Марковские цепи 160

13 Поведение на бесконечности **170**
13.1 Процесс размножения и гибели 170
13.2 Стационарные вероятности 174
13.3 Предельные вероятности 176

14 Корреляционная теория **181**
14.1 Основные определения 181
14.2 Дифференцирование с.п. 182
14.3 Интегрирование с.п. 184
14.4 Стационарный с.п. 187

15 Спектральная теория **190**
15.1 Спектральная плотность 190
15.2 С.п. в линейной динамической системе 192
15.3 Выделение сигнала из шума 194

III Теория массового обслуживания **198**

16 Аналитические методы изучения СМО **199**

16.1 Основные понятия. Классификация систем массового обслуживания . . . 199
16.2 Простейший поток заявок . . . 200
16.3 Марковские СМО . . . 204
16.4 Показатели эффективности СМО . . . 206
16.5 Некоторые виды СМО . . . 207
16.5.1 Пуас/Показ/N/0 . . . 207
16.5.2 Пуас/Показ/N/m . . . 210
16.5.3 Пуас/Показ/N/∞ . . . 214
16.5.4 Пуас/Произв/1/∞ . . . 216

17 Статистическое моделирование СМО 218
17.1 Основные понятия . . . 218
17.2 Генерация случайных чисел . . . 220
17.3 Приближенное вычисление показателей эффективности СМО . . . 221
17.4 Потоки Эрланга . . . 225

IV Математическая статистика 227

18 Выборочные распределения 228
18.1 Предварительная обработка наблюдений . . . 228
18.2 Аппроксимация законов распределения . . . 235
18.3 Гистограмма . . . 239

19 Точечные оценки 243
19.1 Выборочное среднее . . . 244
19.2 Выборочная дисперсия . . . 246
19.3 Упрощение ручных расчетов . . . 248
19.4 Метод наибольшего правдоподобия . . . 251

20 Интервальные оценки 255
20.1 Оценка $\mathbf{M}X$ при известной дисперсии . . . 256
20.2 Оценка $\mathbf{M}X$ при неизвестной дисперсии . . . 260

20.3 Оценка $\mathbf{D}X$ 262

21 Статистическая проверка гипотез 265
21.1 Основные определения 265
21.2 Критерий согласия Пирсона 268
21.3 Критерий согласия Романовского 275
21.4 Критерий согласия Колмогорова 277

22 Элементы корреляционного анализа 279
22.1 Выборочные уравнения регрессии 279
22.2 Корреляционное отношение 283
22.3 Понятие о криволинейной регрессии 286

V Приложения 288

A Некоторые комбинаторные тождества 289

B Полиномиальные коэффициенты 292

C Интеграл Пуассона 294

D Положительно определенные матрицы 296

E Матричное дифференцирование 299

F Гамма- и бета-функции 301

G Математические таблицы 305
G.1 Плотность нормального распределения 305
G.2 Функция нормального распределения 308
G.3 t-распределение Стьюдента 311
G.4 χ^2-распределение Пирсона 313
G.5 Функция $P(\lambda)$ 315

Предметный указатель 316

Литература **323**

Часть I

Теория вероятностей

Лекция 1
Случайные события и их вероятности

1.1 Предмет теории вероятностей

Теория вероятностей занимается изучением ***случайных событий*** (с.с.). Случайное событие — такое событие, которое при одних и тех же условиях может произойти, а может и не произойти. Например, бросание игрального кубика иногда приводит к выпадению шестерки, а иногда — нет; играя в лотерею, можно выиграть, а можно и проиграть; нельзя заранее предсказать, опоздает или придет вовремя поезд, сколько вызовов поступит на телефонную станцию в течение суток и т.д.

Дело в том, что на реализацию с.с. влияет очень большое число различных факторов, так что учесть их все невозможно. Теория вероятностей и не ставит перед собой задачу предсказания, произойдет или не произойдет единичное с.с.

Другое дело, если условия, в которых может реализоваться с.с., воспроизводятся многократно. Тогда обнаруживается, что некоторые характеристики с.с. подчиняются определенным закономерностям, которые можно использовать в практической деятельности. Так, рост отдельно взятого человека мало о чем говорит, но средний рост группы населения позволяет судить о стране, народности, расе и т.п. Некоторые закономерности проявляют такую устойчивость, что их нарушение может свидетельствовать об изменении условий, в которых они проявляются. Например, очень трудно предсказать, родится мальчик или девочка. Однако отношение числа рождений

мальчиков к числу рождений девочек чрезвычайно стабильно и $\approx 21/20$. Поэтому, когда французский ученый Лаплас в начале XIX века обнаружил, что для Парижа оно равно $25/24$, он заподозрил, что в Париже что-то не так. Действительно, оказалось, что крестьяне из окрестных деревень подбрасывали в парижские приюты своих дочерей как невыгодную в крестьянском хозяйстве часть потомства.

Так вот, ***предметом теории вероятностей*** является изучение закономерностей, которым подчиняются характеристики массовых однородных (происходящих при одних и тех же условиях) с.с.

Поскольку трудно назвать событие, которое бы не было случайным, то можно сказать, что все, что вы изучали до сих пор — бледная копия действительности. Конечно, полученные вами знания нельзя назвать ложными, но нужно понимать, что они представляют собой сильно упрощенные, идеализированные модели. Так, камень, брошенный под углом к поверхности Земли, никогда не полетит по параболе, так как сопротивление воздуха, его движение, форма камня и т.п. отклонят траекторию от параболической непредсказуемым образом. Если влияние случайных факторов незначительно, то пользоваться формулами классической механики можно и нужно, если же нет — следует обратиться к методам теории вероятностей. И хотя в теории вероятностей имеются свои идеализации и допущения, все же она гораздо ближе к реальности, чем многие другие разделы математики.

Теория вероятностей как наука зародилась в XVI–XVII вв. Ее рождение связано с именами Кардано, Гюйгенса, Паскаля, Ферма. Следующий период развития теории вероятностей связан с именем Якова Бернулли (1654-1705). Доказанная им теорема, получившая название «закона больших чисел», была первым теоретическим обоснованием накопленных ранее фактов. Дальнейшими успехами теория вероятностей обязана Муавру, Лапласу, Гауссу, Пуассону и др. Новый, наиболее плодотворный период, связан с именами П.Л. Чебышева (1821-1894) и его учеников М.А. Ляпунова (1857-1918) и А.А. Маркова (1856-1922), сделавших теорию вероятностей стройной математической наукой. Ее последующее развитие обязано в

первую очередь русским и советским математикам (С.Н. Бернштейн, В.И. Романовский, А.Н. Колмогоров, А.Я. Хинчин, В.В. Гнеденко, Н.В. Смирнов и др.). В Украине большой вклад в теорию вероятностей внесли И.И. Гихман, А.В. Скороход, В.Л. Гирко и др.

Не сразу теория вероятностей заняла достойное место среди других математических наук. Не все ученые соглашались с тем, что миром правит случай. Если бы иметь все данные о том или ином явлении в текущий момент времени, то его дальнейшее поведение можно было бы предсказать абсолютно точно, считали они. Например, если знать начальные положения и скорости молекул воздуха, то их дальнейшее движение можно рассчитать как угодно точно. Только отсутствие подобной информации заставляет прибегать к методам теории вероятностей.

Конец спорам положило возникновение в конце XIX века квантовой механики. Принцип неопределенности Гейзенберга установил, что нельзя точно измерить положение элементарной частицы, не изменив ее скорость, или измерить скорость, не изменив ее траектории. Да и частица оказалась не упругим шариком классической механики, а частицей-волной, сугубо вероятностным образом «размазанной» в пространстве.

Таким образом выяснилось, что ***случайность лежит в основе мироздания***. Правда, А. Эйнштейн до конца своих дней так и не согласился с этим фактом, заявив: «Я никогда не поверю, что прежде, чем принять решение, господь бог бросает кости».

С конца XIX в. теория вероятностей бурно развивается. Публикаций по ней каждый год выходит столько же, сколько по другим разделам математики вместе взятым. Достаточно перечислить науки, которые возникли на ее основе: теория массового обслуживания, теория надежности, корреляционный и дисперсионный анализы, математическая генетика, теория распознавания образов, анализ временных рядов, стохастическая теория управления, теория принятия решений и др.

В рамках вашей специальности теория вероятностей применяется для расчета эксплуатационных характеристик систем автоматики и

связи (теория массового обслуживания), для анализа, фильтрации и экстраполяции случайных сигналов (теория случайных процессов, статистическая радиотехника), управления в условиях случайных помех (стохастическая теория управления), определения неизвестных параметров объектов управления (теория идентификации систем) и др.

1.2 Случайные события

Итак, теория вероятностей занимается только массовыми с.с., которые происходят при одних и тех же условиях.

Воспроизведение условий, при которых возможна реализация с.с., называется, как в физике или в химии, ***опытом***.

Через Ω обозначим все возможные, исключающие друг друга результаты опыта, и назовем такое множество ***пространством элементарных случайных событий***, а его элементы — ***элементарными с.с.*** (э.с.с.). Далее мы увидим, что из э.с.с., как из кирпичиков, можно строить другие с.с.

Пример 1.1. *Иванов решил купить три лотерейных билета. Требуется построить пространство э.с.с. возможных вариантов выигрыша.*

Решение. Обозначим через В с.с. «Лотерейный билет Иванова выиграл», а через Н — «Лотерейный билет Иванова не выиграл». Тогда

$$\Omega = \{HHH, BHH, HBH, HHB, BBH, BHB, HBB, BBB\}.$$

В данном случае пространство э.с.с. конечно и описывается указанием всех своих элементов. □

Пример 1.2. *Иванов решил играть в лотерею до тех пор, пока не выиграет. Необходимо построить пространство э.с.с. для этого случая.*

Решение. Получаем бесконечное множество

$$\Omega = \{\overset{0}{B}, \overset{1}{H}B, \overset{2}{H}\overset{}{H}B, \overset{3}{H}\overset{}{H}\overset{}{H}B, \ldots\},$$

элементы которого могут быть занумерованы, например, числом букв «Н», содержащихся в элементе. Такие множества называют ***счетными***. □

Пример 1.3. *Поезд прибывает на станцию в 13.15 по расписанию, но может опоздать. Найти пространство э.с.с. всевозможных моментов прибытия поезда*

Решение. Моменты прибытия поезда образуют интервал

$$\Omega = [13.15, \infty).$$

Э.с.с. — точка из указанного интервала. Количество э.с.с. бесконечно и их нельзя занумеровать. Такие множества называют ***несчетными*** или ***континуальными***. □

Случайными событиями называют подмножества множества Ω, удовлетворяющие определенным условиям. Об этих условиях поговорим позже, а пока рассмотрим некоторые с.с. из примеров 1.1-1.3.

В примере 1.1 из э.с.с. можно образовать с.с. $A_1 = \{BHB, HBB, BBH\}$ = «Иванов выиграл в лотерею ровно два раза», $B_1 = \{BBH, BHH, BBB, BHB\}$ = «Иванов с первого раза выиграл в лотерею», $C_1 = \{BHH, HBH, HHB\}$ = «Иванов ровно один раз выиграл в лотерею». В примере 1.2 имеем с.с. $A_2 = \{B, HHB, HHHHB, \ldots\} =$ = «Иванов в нечетный раз выиграл в лотерею». В примере 1.3: $A_3 =$ $= [13.30; 13.45]$ = «Поезд прибыл на станцию между 13.30 и 13.45».

Известно, что над множествами можно выполнять определенные операции. Этим операциям в теории вероятностей соответствуют ***операции над с.с.***, приводящие к новым с.с.

$\overline{A}$ — с.с., ***противоположное*** с.с. A, состоит в ненаступлении A. Например, $\overline{B}_1$ = «Иванов не выиграл в лотерею с первого раза» = $= \{HHH, HBH, HHB, HBB\}$.

$A \cap B = A \cdot B$ — ***произведение с.с.***, состоящее в одновременном наступлении с.с. A и B. Например $A_1 \cdot B_1 = \{BHB, BBH\}$ = «У Иванова выиграл первый лотерейный билет и то ли второй, то ли третий».

$A \cup B$ — ***объединение*** с.с., происходящее, если реализуется хотя бы одно из с.с.: A или B. Например, $B_1 \cup C_1$ = «Иванов то ли с первого раза выиграл в лотерею, то ли выиграл всего один раз».

$A - B = A \cdot \overline{B}$ — ***разность*** с.с. A и B.

При пересечении множеств может получиться пустое множество $\emptyset$. В теории вероятностей ему соответствует с.с., которое называется ***невозможным***, так как никогда не наблюдается в опыте. Например, $A_1 \cdot C_1 = \emptyset$.

С.с. A и B, произведение которых является невозможным событием, называются ***несовместными***, в противном случае они называются ***совместными***. Следовательно, A_1 и C_1 несовместны.

Объединение несовместных с.с. A и B называется их ***суммой*** и обозначается $A + B$. Так, $A_1 \cup C_1 = A_1 + C_1$ = «Иванов выиграл в лотерею один или два раза». Э.с.с. все попарно несовместны и в сумме дают все пространство Ω. В примере 1.1

$$HHH + BHH + HBH + HHB + BBH + BHB + HBB + BBB = \Omega.$$

Что означает это с.с.? Только то, что один из исходов опыта обязательно реализуется. Поэтому Ω называют ***достоверным*** событием, оно обязательно происходит в опыте.

Итак, если мы хотим оперировать со с.с., мы должны считать результаты операций тоже с.с. Это условие обеспечивает система аксиом А.Н. Колмогорова, которая дает строгое определение с.с. Совокупность подмножеств из Ω называется ***алгеброй с.с.*** и обозначается $\mathcal{A}$, если

α_1) $\Omega \in \mathcal{A}$;

α_2) $A \in \mathcal{A} \Longrightarrow \overline{A} \in \mathcal{A}$;

α_3) $A_1 \in \mathcal{A}, \ldots, A_n \in \mathcal{A}, \ldots \Longrightarrow A_1 \cup \ldots \cup A_n \cup \ldots \in \mathcal{A}$.

Строго говоря, множества из $\mathcal{A}$ и называются ***случайными событиями***. Из приведенных примеров видно, что $\alpha_1) - \alpha_3)$ — вполне ра-

зумные требования, например, из них следует, что $\emptyset \in \mathcal{A}$ и $A \cdot B \in \mathcal{A}$, если A и B принадлежат $\mathcal{A}$. Действительно, ведь $\emptyset = \overline{\Omega} \in \mathcal{A}$ и $A \cdot B = \overline{\overline{A} \cup \overline{B}} \in \mathcal{A}$.

Для множеств определены отношения включения: $A \subset B$, $B \supset A$. В теории вероятностей в случае $A \subset B$ говорят, что с.с. A ***благоприятствует*** с.с. B. Если же $A \subset B$ и $B \subset A$, то говорят, что с.с. A и B ***эквивалентны*** и пишут $A = B$. Например, событие $A =$ «Второй стрелок попадет в мишень» благоприятствует событию $B =$ «Мишень будет поражена» и поэтому $A \subset B$. Примером эквивалентных событий служат события «Случайно выбранное число разделится на шесть» и «Случайно выбранное число разделится и на два, и на три». Отметим, что все свойства операций и отношений для множеств сохраняются и для с.с.

1.3 Вероятности случайных событий

Понятие вероятности связано с ***относительной частотой*** $h(A)$ с.с. A, под которой понимается отношение числа опытов n_A (***частоты***), в которых наблюдалась реализация с.с A к общему числу проведенных опытов n. Ясно, что относительная частота может принимать значения только между 0 и 1.

По аналогии ***вероятности*** определяются как числа $P(A)$, $P(B)$, ..., которые приписываются с.с. A, B, ... и обладают свойствами [22, с. 80]:

$$\beta_1)\ P(A) \geq 0,$$

$$\beta_2)\ \sum_{i=1}^{\infty} A_i = \Omega \Longrightarrow \sum_{i=1}^{\infty} P(A_i) = 1.$$

Эти аксиомы отличаются от тех, которые предложил А.Н. Колмогоров, но равносильны им и более наглядны. Первая аксиома означает, что вероятность — неотрицательная величина, а вторая — что сумма вероятностей попарно несовместных с.с., сумма которых образует достоверное событие, равна единице. Это вполне понятные требования и их очень мало — всего два. Но из них и аксиом $\alpha_1)-\alpha_3)$ следуют все

СВОЙСТВА ВЕРОЯТНОСТЕЙ

1° *Вероятность невозможного события равна нулю, вероятность достоверного события равна единице:*

$$P(\emptyset) = 0, \quad P(\Omega) = 1.$$

Так как $\emptyset + \Omega = \Omega$, то из аксиомы β_2) получаем, что $P(\emptyset) + P(\Omega) = P(\Omega) = 1$, откуда и следует требуемое.

2° *Вероятность с.с., противоположного с.с. A, вычисляется по формуле:*

$$P(\overline{A}) = 1 - P(A).$$

Снова используя аксиому β_2) и формулу $A + \overline{A} = \Omega$, получаем равенство $P(A) + P(\overline{A}) = 1$, равносильное доказываемому свойству.

3° *Вероятность с.с. заключена между 0 и 1:*

$$0 \le P(A) \le 1.$$

Свойство 2 и аксиома β_1) приводят к соотношению $0 \le P(A) = 1 - P(\overline{A})$, из которого следует, что $P(A) \le 1$.

4° ***Теорема сложения для несовместных событий.*** *Если с.с. A и B несовместны, то вероятность их суммы равна сумме их вероятностей:*

$$\boxed{P(A + B) = P(A) + P(B).}$$

Очевидно тождество: $(A + B) + (\overline{A + B}) = A + B + (\overline{A + B}) = \Omega$, применяя к которому дважды аксиому β_2), с одной стороны получаем, что $P(A + B) + P(\overline{A + B}) = 1$, а с другой стороны $P(A) + P(B) + P(\overline{A + B}) = 1$. Значит, $P(A + B) + P(\overline{A + B}) = P(A) + P(B) + P(\overline{A + B})$, из чего и следует заключение теоремы.

5° ***Теорема сложения для совместных событий.*** *Если с.с. A и B совместны, то вероятность их объединения вычисляется по формуле*

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cdot B).}$$

Так как $B = B\cdot\Omega = B\cdot(A+\overline{A}) = A\cdot B + B\cdot\overline{A}$, то по предыдущему свойству $P(B) = P(A\cdot B) + P(B\cdot\overline{A})$, значит,

$$P(B\cdot\overline{A}) = P(B) - P(A\cdot B). \tag{1.1}$$

Кроме того,

$$\begin{aligned} A\cup B &= (A\cup B)\Omega = (A\cup B)(A+\overline{A}) = \\ &= AA\cup AB + A\overline{A} + B\overline{A} = A\cup AB + B\overline{A} = \\ &= A(\Omega\cup B) + B\overline{A} = A + B\overline{A} \end{aligned} \tag{1.2}$$

и снова-таки по свойству 4 $P(A\cup B) = P(A) + P(B\cdot\overline{A})$. Подставляя в полученное равенство вместо $P(B\cdot\overline{A})$ его выражение из (1.1), получаем доказываемую формулу.

6°. *Вероятность объединения двух с.с. не превосходит суммы их вероятностей:*

$$P(A\cup B) \le P(A) + P(B).$$

Простое следствие теоремы сложения для совместных с.с.

7°. *Если событие A благоприятствует событию B, то вероятность их разности равна разности их вероятностей:*

$$P(B-A) = P(B) - P(A).$$

Так как в данном случае $A\subset B$, то из (1.2) имеем $B = B\cup A = A + B\cdot\overline{A}$. Поэтому

$$P(B) = P(A) + P(B\cdot\overline{A}),$$

что равносильно доказываемому равенству.

8°. *Если с.с. A благоприятствует с.с. B, то вероятность первого не больше вероятности второго:*

$$A\subset B \Longrightarrow P(A)\le P(B).$$

Сразу следует из предыдущего свойства, если учесть, что $P(B - A) \geq 0$.

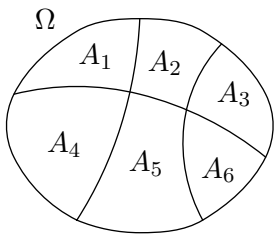
Пример 1.4. *Вероятность того, что новый компьютер поступит в отдел № 1, равна 0,7, а вероятность того, что он поступит в отдел № 2, равна 0,1. Какова вероятность того, что новый компьютер поступит в один из этих отделов?*

Решение. Обозначим события: A_1 = «Новый компьютер поступит в отдел № 1», A_2 = «Новый компьютер поступит в отдел № 2». Так как эти события исключают друг друга, то они несовместны. Следовательно, можно применить теорему сложения для несовместных с.с.:

$$P(A_1 + A_2) = P(A_1) + P(A_2) = 0{,}7 + 0{,}1 = 0{,}8.$$

□

1.4 Классическое определение вероятности



Система возможных, попарно несовместных с.с. $A_1, \ldots, A_n$, сумма которых равна Ω, называется ***разбиением***. Т.о.

$A_i \cdot A_j = \emptyset \ (i \neq j), \ A_i \neq \emptyset,$

$A_1 + \ldots + A_n = \Omega.$

С.с. $A_1, \ldots, A_n$ называются ***равновероятными***, если $P(A_1) = \ldots = P(A_n) = p$.

Теорема 1.1. *Если с.с. $A_1, \ldots, A_n$ равновероятны, образуют разбиение и $A_1 + \ldots + A_m = A$, то*

$$\boxed{P(A) = \frac{m}{n}.} \tag{1.3}$$

Доказательство. Так как $A_1, \ldots, A_n$ попарно несовместны, то по теореме сложения для несовместных с.с. имеем

$$P(A_1) + \ldots + P(A_n) = P(A_1 + \ldots + A_n) = P(\Omega) = 1,$$

значит, $np = 1$, а $p = 1/n$. Аналогично, в силу $A_1 + \ldots + A_m = A$ получаем

$$P(A) = P(A_1 + \ldots + A_m) = P(A_1) + \ldots + P(A_m) = p \cdot m = m/n.$$

□

Так как $A_1 \subset A, \ldots, A_n \subset A$, то события $A_1, \ldots, A_n$ благоприятствуют с.с. A. Это дает основание прочитать формулу так: *вероятность с.с. A равна отношению числа с.с., благоприятствующих с.с. A, к общему числу равновероятных с.с., образующих разбиение.* В таком виде до аксиоматизации теории вероятностей формулировалось определение вероятности, в силу чего выражение (1.3) называют также ***классическим определением вероятности***. Пользоваться им можно далеко не всегда, а лишь при выполнении перечисленных в теореме 1.1 условий.

В теории вероятностей для моделирования различных практических ситуаций служит так называемая ***урновая модель***. Имеется урна, в которой находятся разноцветные шарики и выбираются из нее наугад. Требуется определить вероятность того, что в выборке обнаружится определенная комбинация шариков.

Пример 1.5. *Пусть, например, в урне имеется 10 синих, 5 красных и 2 желтых шара. Какова вероятность вынуть из урны синий шар?*

Решение. Очевидно, вынуть любой шар из урны можно с одной и той же вероятностью. С.с., благоприятствующих интересующему нас с.с., десять — по количеству синих шаров, а всего возможных исходов — 17. Значит, по формуле (1.3) искомая вероятность равна $10/17$. □

10 синих
5 красных
2 желтых

Эффективное применение классического определения вероятности для решения более сложных задач может потребовать привлечения комбинаторных формул. Этот материал можно изучить по [23] и [7].

Пример 1.6 (Задача выборочного контроля). *В партии из L изделий имеется K бракованных. Случайным образом отбираются l изделий. Какова вероятность того, что среди отобранных изделий ровно k бракованных?*

Решение. Всего возможных способов выбрать l изделий из L изделий имеется $n = C_L^l$. Так как k бракованных изделий выбираются из K бракованных, а оставшиеся $l-k$ небракованных изделий выбираются из $L-K$ небракованных, то число способов получить партию из l изделий, среди которых k бракованных, равно $m = C_K^k C_{L-K}^{l-k}$. Следовательно, искомая вероятность равна

$$p = \frac{m}{n} = \frac{C_K^k C_{L-K}^{l-k}}{C_L^l}.$$

□

Пример 1.7. *В бригаде работают 6 мужчин и 4 женщины. По табельным номерам наудачу отобраны 7 человек. Найти вероятность того, что среди них оказались 3 женщины.*

Решение. Воспользуемся формулой, полученной при решении предыдущей задачи, с такими значениями входящих в нее величин: $L = 10$, $K = 4$, $l = 7$, $k = 3$. Получим

$$p = \frac{C_4^3 C_6^4}{C_{10}^7} = \frac{4!6!7!3!}{3!1!4!2!10!} = 0{,}5.$$

□

Пример 1.8. *Шифр автоматической камеры хранения состоит из четырех символов. Первый символ — одна из букв А, Б, В, Г, Д, Е, Ж, И, К, Л. Остальные три символа — одна из цифр от 0 до 9. Найти вероятность случайным набором шифра открыть камеру хранения.*

Решение. Число возможных шифровых комбинаций равно $A_{10}^4 = 10^4 = 10000$. Так как благоприятствует открыванию камеры одна-единственная комбинация символов, то искомая вероятность равна $1/10000$. □

Пример 1.9. *Имеются автопогрузчики четырех равноценных марок, которыми удовлетворяются заявки предприятий. Железнодорожный узел послал заявку на три автопогрузчика. Найти вероятность того, что он получит автопогрузчики одной марки.*

Решение. Число вариантов, которыми может быть удовлетворена заявка предприятия, равно $\overline{C}_4^3 = C_6^3 = 20$. Среди них имеется четыре благоприятствующих варианта. Поэтому искомая вероятность равна $4/20 = 0{,}2$. □

Лекция 2

Геометрическая и условная вероятности. Независимость случайных событий

2.1 Геометрическая вероятность

Рассмотрим множество Ω из n-мерного пространства, имеющее n-мерный объем, и систему $\mathcal{A}$ его подмножеств, также имеющих n-мерный объем:

$$\mathcal{A} = \{A: \ A \subset \Omega \subset R^n, \text{ A имеет объем } \mu(A)\}.$$

Для любого $A \in \mathcal{A}$ положим

$$\boxed{P(A) = \frac{\mu(A)}{\mu(\Omega)}} \tag{2.1}$$

Можно показать, что при правильном определении n-мерного объема величина $P(A)$ равна вероятности с.с. A, которое состоит в попадании точки $(x_1, \ldots, x_n)$ в область A и с этой областью отождествляется. Выражение (2.1) называется ***геометрическим определением вероятности*** и часто используется при решении практических задач.

Пример 2.1. *В точке K, положение которой на телефонной линии BD длины L равновозможно, произошел разрыв. Определить вероятность того, что точка K удалена от точки B на расстояние, не меньшее l.*

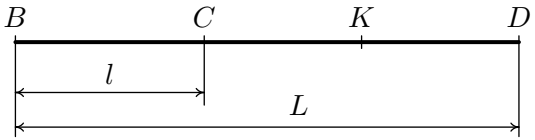


Рис. 2.1: Разрыв телефонной линии

Решение. Сделаем чертеж (рис. 2.1). Выберем в качестве множества Ω отрезок BD, тогда искомая вероятность равна вероятности попадания случайно выбранной точки на отрезок CD. Но мера отрезка — его длина и, так как $\mu(CD) = L - l$, а $\mu(BD) = L$, то по формуле (2.1) получим $P(CD) = (L - l)/L = 1 - l/L$. $\square$

Пример 2.2. *В любые моменты времени из промежутка* $[0, T]$ *равновозможны поступления в приемник двух сигналов. Приемник будет забит, если разность между моментами поступления сигналов будет меньше* τ. *Определить вероятность того, что приемник будет забит.*

Решение. В данном случае удобно считать множеством Ω квадрат $OTMT$, изображенный на рисунке.

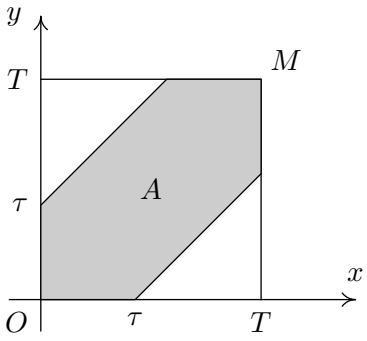


Рис. 2.2: Геометрия поступления сигналов

Пусть x и y — моменты поступления двух сигналов в приемник. По условию задачи последний будет забит, если выполняется неравенство:

$$|y - x| < \tau,$$

или

$$x - \tau < y < x + \tau.$$

На рисунке этому неравенству соответствует закрашенная область A. Следовательно, искомая вероятность будет равна отношению площадей закрашенной области к площади квадрата $OTMT$, причем, площадь фигуры A равна площади данного квадрата минус площади треугольников, которые вырезаются из квадрата закрашенной областью. Поэтому

$$P(A) = \frac{P(\Omega) - (T-\tau)^2}{P(\Omega)} = \frac{T^2 - (T-\tau)^2}{T^2} = 1 - \left(1 - \frac{\tau}{T}\right)^2.$$

□

2.2 Условная вероятность. Независимость с.с.

Вероятность с.с. существенно зависит от реализации других с.с. Например, вероятность того, что поезд доедет без происшествий, зависит от состояния локомотива, вагонов, квалификации машиниста и т.д.

Пусть осуществляется серия из n опытов, в которых могут произойти с.с. A и B. ***Относительной частотой*** осуществления с.с. A при условии реализации с.с. B называется величина $h_{A/B} = k/n_B$, где k — число осуществлений с.с. A в опытах, в которых произошло с.с. B. Очевидно, что $k = n_{AB}$ и поэтому

$$h_{A/B} = \frac{k/n}{n_B/n} = \frac{n_{AB}/n}{n_B/n} = \frac{h_{AB}}{h_B}.$$

Эта формула экспериментально обосновывает понятие условной вероятности.

Условной вероятностью $P(A/B)$ случайного события A при условии, что произошло случайное событие B, называется величина

$$\boxed{P(A/B) = \frac{P(AB)}{P(B)},\ P(B) \neq 0.}$$

Теорема умножения вероятностей.

$$\boxed{P(AB) = P(A/B)P(B).}$$

Эта формула сразу следует из определения условной вероятности, причем, для числа с.с., большего двух, принимает вид:

$$P(A_1 \dots A_n) = P(A_1)P(A_2/A_1) \cdot \ldots \cdot P(A_n/A_1 \dots A_{n-1}).$$

Два с.с. A и B называются ***независимыми***, если вероятность их произведения равна произведению их вероятностей:

$$\boxed{P(AB) = P(A)P(B).}$$

Более, чем два, с.с. $A_1, \dots, A_n$ называются ***независимыми***, если

$$P(A_i A_k \dots A_m) = P(A_i)P(A_k) \dots P(A_m)$$

при любых натуральных $i, k, \dots, m$, заключенных между 1 и n.

Если A и B независимы, то

$$P(A/B) = P(AB)/P(B) = P(A)P(B)/P(B) = P(A),$$

то есть условная вероятность с.с. A равна его безусловной вероятности. Это согласуется с интуитивным представлением о независимости с.с.

Теорема умножения вероятностей для независимых с.с. *Если* $A_1, \dots, A_n$ *независимы, то*

$$P(A_1 \dots A_n) = P(A_1)P(A_2) \cdot \ldots \cdot P(A_n).$$

Пример 2.3. *В зале ожидания станции Кондракино установлены три игральных автомата трех различных фирм. Вероятность того, что в течение года первый, второй или третий автоматы, соответственно, не выйдут из строя равны* $0,6;\ 0,7;\ 0,8$. *Найти вероятности того, что в течение года не выйдут из строя: а) только один автомат, б) ровно два автомата, в) все три автомата.*

Решение. Обозначим с.с.: A_i = «i-й автомат не выйдет из строя в течение года», $i = 1, 2, 3$; $A_{\text{а}},\ A_{\text{б}},\ A_{\text{в}}$ — с.с., указанные в пп. а), б) и в) условия задачи. С.с. $A_{\text{а}}$ можно представить следующим образом:

$$A_{\text{а}} = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3.$$

Слагаемые в этой формуле являются несовместными с.с., а множители каждого слагаемого представляют собой независимые с.с. Поэтому, применяя теорему сложения для несовместных с.с. и теорему умножения вероятностей для независимых с.с., получаем

$$\begin{aligned} P(A_a) = P(A_1)P\left(\overline{A}_2\right)P\left(\overline{A}_3\right) + P\left(\overline{A}_1\right)P\left(A_2\right)P\left(\overline{A}_3\right) + \\ + P\left(\overline{A}_1\right)P\left(\overline{A}_2\right)P(A_3) = \\ = 0{,}6\cdot 0{,}3\cdot 0{,}2 + 0{,}4\cdot 0{,}7\cdot 0{,}2 + 0{,}4\cdot 0{,}3\cdot 0{,}8 = 0{,}188; \end{aligned}$$

причем, вероятности с.с., противоположных с.с. A_1, A_2, A_3 находим по формулам:

$$P\left(\overline{A}_1\right) = 1 - P(A_1) = 1 - 0,6 = 0,4;$$

$$P\left(\overline{A}_2\right) = 1 - P(A_2) = 0,3; \qquad P\left(\overline{A}_3\right) = 1 - P(A_3) = 0,2.$$

Аналогично вычисляем остальные вероятности:

$$P(A_{\text{б}}) = 0,452,\ \ P(A_{\text{в}}) = 0,336.$$

□

Пример 2.4 (Задача о надежности)**.** *Механизм выходит из строя, если выходит из строя хотя бы одна из его* n *деталей. Детали выходят из строя независимо друг от друга. Найти вероятность исправной работы механизма в течение заданного промежутка времени* T, *если вероятность выхода из строя* i*-й детали равна* p_i.

Решение. Вероятность исправной работы i-й детали в течение заданного промежутка времени равна $q_i = 1 - p_i$. Обозначим A_i событие, состоящее в том, что в течение времени T деталь проработает исправно. Так как с.с. $A_1, \ldots, A_n$ независимы, то по теореме умножения вероятностей для независимых событий

$$P(A_1 \cdot \ldots \cdot A_n) = P(A_1) \cdot \ldots \cdot P(A_n) = q_1 \cdot \ldots \cdot q_n.$$

Следовательно, искомая вероятность равна

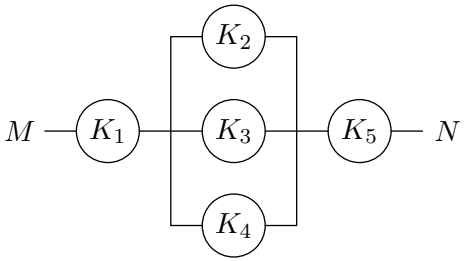
$$\widetilde{p}_n = P(A_1 \cdot \ldots \cdot A_n) = q_1 \cdot \ldots \cdot q_n.$$

В частности, если все p_i равны: $p_1 = \cdots = p_n = p$, то

$$\widetilde{p}_n = q^n = (1-p)^n,$$

где $q = 1 - p$. Заметим, что $\lim_{n\to\infty} \widetilde{p}_n = 0$. □

Пример 2.5. *Электрическая цепь между точками M и N составлена по схеме, приведенной на рис.*



Вероятности выхода из строя элементов K_i (разрыв цепи) в течение времени T равны $P(A_1) = 0{,}6$; $P(A_2) = 0{,}4$; $P(A_3) = 0{,}7$; $P(A_4) = 0{,}9$; $P(A_5) = 0{,}5$; где A_i — событие, состоящее в выходе из строя элемента K_i. Определить вероятность разрыва цепи за время T.

Решение. Обозначим A событие, состоящее в разрыве цепи за время T. Очевидно, справедлива формула

$$A = A_1 \cup A_2 A_3 A_4 \cup A_5.$$

Используя теорему сложения для совместных событий и теорему умножения для независимых событий, получаем

$$\begin{aligned}
P(A) &= P(A_1 \cup A_2A_3A_4 \cup A_5) = P(A_1 \cup (A_2A_3A_4 \cup A_5)) = \\
&= P(A_1) + P(A_2A_3A_4 \cup A_5) - P(A_1A_2A_3A_4 \cup A_1A_5) = \\
&= P(A_1) + P(A_2A_3A_4) + P(A_5) - P(A_2A_3A_4A_5) - \\
&\quad - P(A_1A_2A_3A_4) - P(A_1A_5) + P(A_1A_2A_3A_4A_5) = \\
&= P(A_1) + P(A_2)P(A_3)P(A_4) + P(A_5) - \\
&\quad - P(A_2)P(A_3)P(A_4)P(A_5) - P(A_1)P(A_2)P(A_3)P(A_4) - \\
&\quad - P(A_1)P(A_5) + P(A_1)P(A_2)P(A_3)P(A_4)P(A_5).
\end{aligned}$$

Остается подставить значения заданных вероятностей:

$$\begin{aligned}
P(A) &= 0{,}6 + 0{,}4 \cdot 0{,}7 \cdot 0{,}9 + 0{,}5 - 0{,}4 \cdot 0{,}7 \cdot 0{,}9 \cdot 0{,}5 - \\
&\quad - 0{,}6 \cdot 0{,}4 \cdot 0{,}7 \cdot 0{,}9 - 0{,}6 \cdot 0{,}5 + 0{,}6 \cdot 0{,}4 \cdot 0{,}7 \cdot 0{,}9 \cdot 0{,}5 = \\
&= 0{,}6 + 0{,}252 + 0{,}5 - 0{,}126 - 0{,}1512 - 0{,}3 + 0{,}0756 = \\
&= 0{,}8504.
\end{aligned}$$

□

2.3 Формула полной вероятности

Пусть с.с. $A_1, \ldots, A_n$ образуют разбиение. Тогда они несовместны и вместе с ними несовместны с.с. AA_i, AA_j, $i \neq j$, где A - любое с.с. Следовательно,

$$A = A\Omega = A(A_1 + \ldots + A_n) = AA_1 + \ldots + AA_n = \sum_{i=1}^{n} AA_i.$$

Поэтому, применяя теорему сложения для несовместных с.с. и теорему умножения вероятностей для зависимых с.с., получаем

$$P(A) = \sum_{i=1}^{n} P(AA_i) = \sum_{i=1}^{n} P(A/A_i)P(A_i).$$

Формула

$$\boxed{P(A) = \sum_{i=1}^{n} P(A/A_i)P(A_i).}$$

называется ***формулой полной вероятности***.

Пример 2.6. *В каждом из двух учреждений Офис1 и Офис2 работают по 3 ваших знакомых и 5 незнакомых вам людей. Известно, что один сотрудник из Офис1 перешел работать в Офис2, а потом кто-то из Офис2 организовал частное предприятие. Какова вероятность, что частное предприятие организовал ваш знакомый?*

Решение. Обозначим с.с.:

$A_{\text{зн}}$ = «Из Офис1 в Офис2 перешел знакомый»,

$A_{\text{нзн}}$ = «Из Офис1 в Офис2 перешел незнакомый»,

B = «Знакомый организовал частное предприятие».

По условию задачи $P(A_{\text{зн}}) = \frac{3}{8}$, $P(A_{\text{нзн}}) = \frac{5}{8}$. Так как с.с. $A_{\text{зн}}$ и $A_{\text{нзн}}$ образуют разбиение, то можно применить формулу полной вероятности:

$$P(B) = P(B/A_{\text{зн}})P(A_{\text{зн}}) + P(B/A_{\text{нзн}})P(A_{\text{нзн}}).$$

Условные вероятности тоже находим из условия задачи:

$$P(B/A_{\text{зн}}) = \frac{4}{9},\ P(B/A_{\text{нзн}}) = \frac{3}{9}.$$

Следовательно,

$$P(B) = \frac{4}{9} \cdot \frac{3}{8} + \frac{3}{9} \cdot \frac{5}{8} = \frac{9}{24} = \frac{3}{8} = 0,375.$$

□

2.4 Формула Байеса

При условии, что с.с. $A_1, \ldots, A_n$ образуют разбиение и $P(A) \neq 0$, справедлива ***формула Байеса***:

$$\boxed{P(A_i/A) = \frac{P(A/A_i)P(A_i)}{\sum\limits_{j=1}^{n} P(A/A_j)P(A_j)}, \ i = \overline{1,n},}$$

Действительно, по формуле условной вероятности имеем

$$P(A_i/A) = \frac{P(AA_i)}{P(A)} = \frac{P(A/A_i)P(A_i)}{P(A)}.$$

Подставляя в знаменатель вместо вероятности $P(A)$ ее выражение из формулы полной вероятности, получаем формулу Байеса. Последняя часто используется при решении практических задач.

Пример 2.7. *Запоминающие устройства, выполненные на интегральных схемах, могут иметь три вида дефектов: A_1 — дефекты схем обрамления; A_2 — дефекты, вызванные паразитными связями между ячейками; A_3 — дефекты адресных шин. Вероятности этих дефектов равны, соответственно, $P(A_1) = 0{,}1$; $P(A_2) = 0{,}6$; $P(A_3) = 0{,}3$. После диагностики запоминающего устройства с помощью набора тестов наблюдался некоторый результат (произошло событие A), условные вероятности которого до диагностики были равны $P(A/A_1) = 0{,}4$; $P(A/A_2) = 0{,}2$; $P(A/A_3) = 0{,}3$. Установить, какой дефект наиболее вероятен после диагностики.*

Решение. Так как

$$P(A) = P(A/A_1)P(A_1) + P(A/A_2)P(A_2) + P(A/A_3)P(A_3) =$$
$$= 0{,}4 \cdot 0{,}1 + 0{,}2 \cdot 0{,}6 + 0{,}3 \cdot 0{,}3 = 0{,}25,$$

то по формуле Байеса получаем

$$P(A_1/A) = \frac{P(A/A_1)P(A_1)}{P(A)} = \frac{0{,}4 \cdot 0{,}1}{P(A)} = \frac{0{,}04}{0{,}25} = 0{,}16;$$
$$P(A_2/A) = \frac{P(A/A_2)P(A_2)}{P(A)} = \frac{0{,}2 \cdot 0{,}6}{P(A)} = \frac{0{,}12}{0{,}25} = 0{,}48;$$
$$P(A_3/A) = \frac{P(A/A_3)P(A_3)}{P(A)} = \frac{0{,}3 \cdot 0{,}3}{P(A)} = \frac{0{,}09}{0{,}25} = 0{,}36.$$

Итак, в результате диагностики установлено, что наиболее вероятен дефект, вызванный паразитными связями между ячейками. □

Формулу Байеса еще называют формулой гипотез, так как с.с. A_1, $\ldots, A_n$ можно интерпретировать как различные гипотезы, высказанные о чем-либо. Следовательно, формула Байеса показывает, как изменяются вероятности гипотез в результате осуществления некоторого с.с. A: до опыта эти вероятности равнялись $P(A_i)$, а после опыта — $P(A_i/A)$.

Лекция 3
Случайные величины

3.1 Случайная величина и ее функция распределения

В теории вероятностей изучаются переменные величины, значения которых трудно предсказать заранее. Например, время расформирования состава или количество проданных за сутки железнодорожных билетов. Такие величины естественно считать случайными.

Механизм принятия случайной величиной (с.в.) своего случайного значения можно представить себе как выбор э.с.с. ω из Ω и вычисления значения функции $X(\omega)$. Таким образом, с.в. — функция на пространстве э.с.с. и то, какое ее значение реализуется в опыте, определяется тем, какое именно э.с.с. ω из Ω в данном опыте произойдет. Конечно, с.с., связанные с X, должны принадлежать алгебре $\mathcal{A}$ с.с., причем, оказывается, что этого достаточно потребовать только от с.с. вида

$$A_x = \{\omega : X(\omega) < x\},$$

где x — действительное число, а остальные с.с. уже будут принадлежать алгебре $\mathcal{A}$ автоматически.

Строгое определение с.в. таково. ***Случайной величиной*** X называется функция $X = X(\omega)$, $\omega \in \Omega$, для которой с.с. $A_x \in \Omega$.

Пример 3.1. *Построить функции, определяющие с.в., для примеров 1.1– 1.3.*

Решение. В примере 1.1 зададим с.в. X как функцию э.с.с. из Ω

следующим образом:

$$X = X(\omega) = \begin{cases} 3S, & \text{если} \quad \omega = BBB, \\ 2S, & \text{если} \quad \omega \in \{BBH, BHB, HBB\}, \\ S, & \text{если} \quad \omega \in \{BHH, HBH, HHB\}, \\ 0, & \text{если} \quad \omega = HHH, \end{cases}$$

где S — величина выигрыша по лотерейному билету. Т.о. $X(\omega)$ — сумма выигрышей Иванова по трем лотерейным билетам.

В примере 1.2 с.в. X — количество лотерейных билетов, которые придется купить Иванову, чтобы выиграть, определяется формулой:

$$X = X(\omega) = k + 1, \text{если} \quad \omega = \underbrace{H \ldots H}_{k} B.$$

В примере 1.3 можно задать функцию $X = X(\omega) = \omega - 13.15$ представляющую собой с.в. опоздания поезда. □

Однако задавать подобным образом с.в. не всегда удобно, тем более, что существует характеристика с.в., которая носит название ***функции распределения (вероятностей)*** с.в. X и определяется вероятностью с.с. A_x:

$$\boxed{F(x) = P(X < x)} = P(A_x) = P(X \in (-\infty, x)).$$

Эта характеристика полностью описывает поведение с.в., что демонстрируют

СВОЙСТВА ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

1°. $0 \leq F(x) \leq 1$.

Свойство объясняется тем, что $F(x)$ представляет собой вероятность.

2°. *Вероятность попадания с.в. в интервал* $[x_1, x_2)$ *равна разности значений функции распределения, вычисленных на концах интервала:*

$$\boxed{P(x_1 \leq X < x_2) = F(x_2) - F(x_1).}$$

Так как

$$\{\omega : X(\omega) < x_2\} = \{\omega : X(\omega) < x_1\} + \{\omega : x_1 \le X(\omega) < x_2\},$$

то

$$F(x_2) = P(X < x_2) = \\ = P(X < x_1) + P(x_1 \le X < x_2) = F(x_1) + P(x_1 \le X < x_2),$$

из чего и следует требуемое свойство.

3°. *Функция распределения не убывает:*

$$x_1 < x_2 \Longrightarrow F(x_1) \le F(x_2).$$

Пусть $x_1 < x_2$. Из предыдущего свойства следует, что

$$F(x_2) - F(x_1) = P(x_1 \le X < x_2) \ge 0,$$

так как вероятность неотрицательна.

4°. $\lim\limits_{x\to-\infty} F(x) = F(-\infty) = 0$, $\lim\limits_{x\to\infty} F(x) = F(\infty) = 1$.

Действительно, при $x \to -\infty$ интервал $(-\infty, x) \to (-\infty, -\infty)$ и, следовательно, вероятность попасть в такой интервал должна стремиться к нулю. При $x \to \infty$ интервал $(-\infty, x) \to (-\infty, \infty)$ и, значит, $F(x)$ должна стремиться к вероятности того, что с.в. примет какое-нибудь значение. Ясно, что вероятность такого с.с. равна единице.

5°. *Функция распределения непрерывна слева:*

$$F(x - 0) = F(x).$$

Так как $F(x)$ ограничена, то она существует в каждой точке оси Ox. Однако она не обязательно непрерывна и поэтому может иметь точки разрыва первого рода. Обычно принимают, что в точках разрыва функция распределения непрерывна слева.

6°. *Вероятность того, что с.в. примет значение x, равна величине скачка функции распределения в данной точке:*

$$P(X = x) = F(x + 0) - F(x).$$

Произвольная функция представляет собой функцию распределения, если для нее выполняются свойства 3, 4 и 5.

3.2 Непрерывные с.в. Плотность вероятностей.

Если функция $F(x)$ непрерывна и недифференцируема лишь в конечном числе точек, то существует другая всеобъемлющая характеристика с.в., которая называется ее ***плотностью вероятностей***, обозначается $f(x)$ и в точках дифференцируемости функции распределения связана с ней равенством

$$\boxed{f(x) = F'(x).}$$

Случайная величина, имеющие плотность вероятностей, называется ***непрерывной***. Типичный график функции распределения непрерывной с.в. имеет вид, показанный на рис. 3.1:

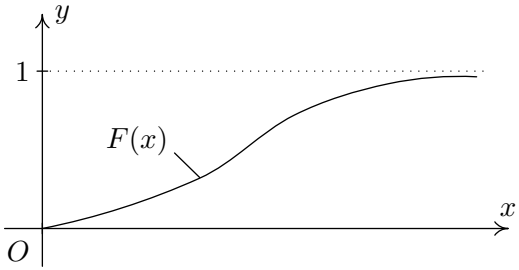


Рис. 3.1: График функции распределения непрерывной с.в.

Функцию распределения или плотность вероятностей непрерывной с.в. будем называть ее ***законом распределения***.

Примерами непрерывных с.в. могут служить времена погрузки или разгрузки ж.д. вагона, время его ремонта, объем топлива, израсходованного локомотивом, и др.

СВОЙСТВА ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ И ФУНКЦИИ РАСПРЕДЕЛЕНИЯ НЕПРЕРЫВНОЙ С.В.

1°. *Плотность вероятностей неотрицательна:* $f(x) \geq 0$.

Так как $f(x) = F'(x)$, а $F(x)$ не убывает, то $F'(x) \geq 0$ и, значит, $f(x) \geq 0$ тоже.

2°. $P(a \leq X < b) = \int\limits_a^b f(x)dx.$

Так как $F(x)$ является для $f(x)$ первообразной, то из свойства 2 функции распределения и формулы Ньютона-Лейбница сразу получаем требуемое. Таким образом, вероятность попадания с.в. в интервал $[a, b)$ численно равна площади криволинейной трапеции, ограниченной сверху графиком плотности вероятностей, а снизу — интервалом $[a, b)$:

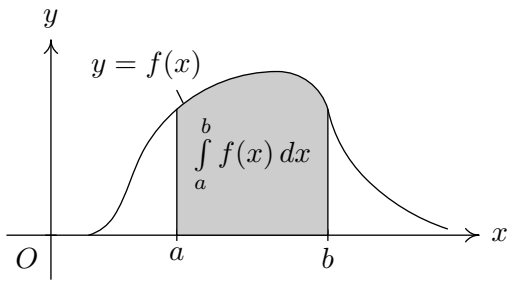


Рис. 3.2: Вероятность попадания в промежуток $[a, b)$.

3°. *Условие нормировки:* $\int\limits_{-\infty}^{\infty} f(x)dx = 1.$

Это свойство означает, что площадь криволинейной трапеции, сосредоточенной под всем графиком плотности вероятностей, равна единице. Доказательство получается из предыдущего свойства и свойства 4 функции распределения.

4°. *Функция распределения через плотность вероятностей выражается формулой*

$$F(x) = \int\limits_{-\infty}^{x} f(x)dx.$$

Это непосредственно следует из определения плотности вероятностей.

5°. *Вероятность того, что непрерывная с.в. примет конкретное значение, равна нулю:* $P(X = x) = 0$.

Это следует из непрерывности функции распределения и ее свойства 6.

Произвольная функция представляет собой плотность распределения вероятностей, если для нее выполняются свойства 1 и 3.

Пример 3.2. *Дана функция распределения непрерывной с.в.* X: $F(x) = a(b + \operatorname{arctg} x)$. *Определить значения параметров распределения* a *и* b, *найти плотность вероятностей с.в.* X, *а также вероятность ее попадания в интервал* $(-1; 1)$.

Решение. Воспользуемся свойствами функции распределения:

$$F(+\infty) = 1 = a\left(b + \frac{\pi}{2}\right), \quad F(-\infty) = 0 = a\left(b - \frac{\pi}{2}\right).$$

Из этой системы уравнений находим, что $a = 1/\pi$, $b = \pi/2$. Следовательно, функция распределения имеет вид

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\operatorname{arctg} x.$$

Найдем плотность вероятностей заданной с.в.:

$$f(x) = F'(x) = \frac{1}{\pi(1 + x^2)}.$$

Вероятность попадания с.в. X в интервал $(-1; 1)$ равна

$$P(-1 < X < 1) = F(1) - F(-1) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2} = 0{,}5.$$

□

Пример 3.3. *С.в. X имеет плотность вероятностей, равную $f(x) = C(1 - x^2)$ на интервале $(-1; 1)$, и равную нулю вне этого интервала. Найти значение параметра C и функцию распределения этой с.в.*

Решение. Так как $\int_{-\infty}^{\infty} f(x)\,dx = 1$, то

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = C\int\limits_{-1}^{1}(1-x^2)\,dx = C\left(x - \frac{x^3}{3}\right)\Bigg|_{-1}^{1} = C\cdot\frac{4}{3} = 1.$$

Следовательно, $C = 3/4$.

Найдем функцию распределения:

$$F(x) = \int\limits_{-\infty}^{x} f(t)\,dt = \begin{cases} 0, & x \le -1; \\ \frac{3}{4}\int\limits_{-1}^{x}(1-t^2)\,dt, & -1 < x < 1; \\ 1, & x \ge 1; \end{cases} =$$

$$= \begin{cases} 0, & x \le -1; \\ \frac{3}{4}\left(t - \frac{t^3}{3}\right)\Big|_{-1}^{x}, & -1 < x < 1; \\ 1, & x \ge 1; \end{cases} =$$

$$= \begin{cases} 0, & x \le -1; \\ \frac{3}{4}\left(x - \frac{x^3}{3}\right) + \frac{1}{2}, & -1 < x < 1; \\ 1, & x \ge 1. \end{cases}$$

Графики функций $f(x)$ и $F(x)$ показаны на рис. 3.3.

□

3.3 Дискретные с.в.

Вторым важным классом с.в. являются ***дискретные*** с.в. К ним относят с.в., принимающие лишь конечное или счетное число значений. Эти значения обозначают $x_1, x_2, \ldots, x_k, \ldots$, а их вероятности, соответственно, $p_1, p_2, \ldots, p_k, \ldots$, так что $p_k = P(X = x_k)$, $k = 1, 2, \ldots$. Будем считать, что $x_1 < x_2 < \ldots$.

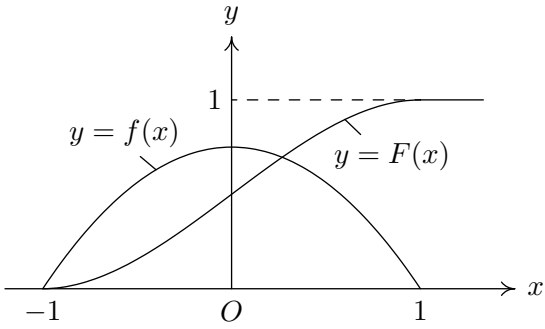


Рис. 3.3: К примеру 3.3

Примерами дискретных с.в. могут служить количество отказов аппаратуры за месяц; количество служащих, не вышедших на работу по болезни; число вызовов, поступивших на телефонную станцию за сутки и др.

Из определения дискретной с.в. видно, что ее удобно задавать парами чисел (x_k, p_k), которые можно представить либо в виде таблицы чисел, называемой ***рядом распределения***:

x_k	x_1	x_2	$\dots$	x_i	$\dots$
p_k	p_1	p_2	$\dots$	p_i	$\dots$

,

либо в виде ломаной, звенья которой соединяют точки (x_k, p_k) на плоскости xOp, и называемой ***многоугольником распределения*** (см. рис. 3.4).

Функция распределения дискретной с.в. является кусочно-постоянной, ступенчатой функцией, причем, величина «ступеньки» в точке x_k равна p_k (см. рис. 3.5).

Действительно, пусть задано число x и $x_1, x_2, \dots, x_k$ меньше x, а другие значения с.в. X $(x_{k+1}, x_{k+2}, \dots)$ больше x. Тогда в силу того, что с.с. $\{X = x_i\}$, $\{X = x_j\}$ при $i \neq j$ несовместны, на основании

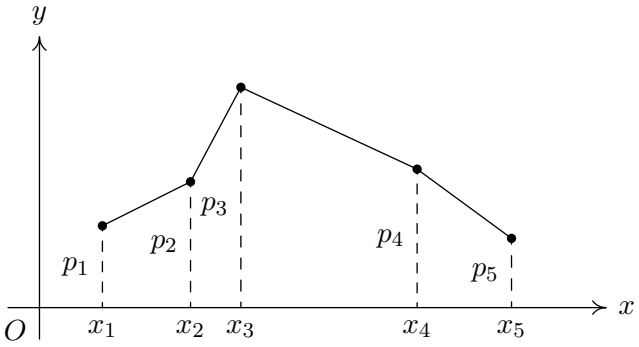


Рис. 3.4: Многоугольник распределения

определения функции распределения получаем

$$\begin{aligned} F(x) = P(X < x) = P[(X = x_1) + (X = x_2) + \ldots + (X = x_k)] = \\ = P(X = x_1) + P(X = x_2) + \ldots + P(X = x_k) = \\ = p_1 + p_2 + \ldots + p_k, \end{aligned}$$

или

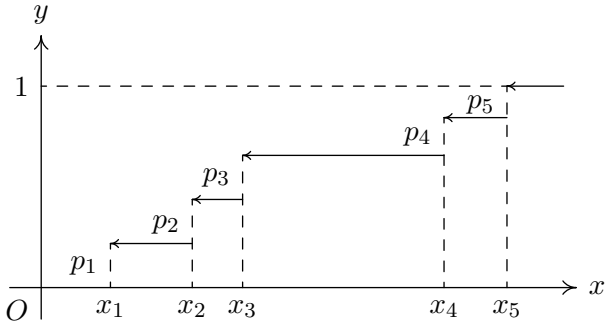


Рис. 3.5: Функция распределения дискретной с.в.

$$F(x) = \sum_{k:\, x_k < x} p_k. \tag{3.1}$$

Эта формула и доказывает ступенчатость функции распределения дискретной с.в.

Вероятности p_k обладают свойством (условие нормировки)

$$\boxed{\sum_{k=1}^{n} p_k = 1,}$$

где n — число или бесконечность.

Действительно,

$$F(\infty) = 1 = \sum_{x_k < \infty} p_k = \sum_{k=1}^{n} p_k = 1.$$

Функцию, ряд или многоугольник распределения дискретной с.в. будем называть ее ***законом распределения***.

Пример 3.4. *Пусть в примере 1.7 X — с.в., равная числу женщин среди отобранных семи человек. Построить для этой с.в. ряд, многоугольник и функцию распределения.*

Решение. Данная с.в. может принимать с ненулевыми вероятностями лишь значения 1, 2, 3 и 4. Вероятности этих значений в соответствии с формулой выборочного контроля равны

$$p_1 = P(X = 1) = \frac{C_4^1 C_6^6}{C_{10}^7} = \frac{1}{30} \approx 0{,}33;$$

$$p_2 = P(X = 2) = \frac{C_4^2 C_6^5}{C_{10}^7} = \frac{3}{10} = 0{,}3;$$

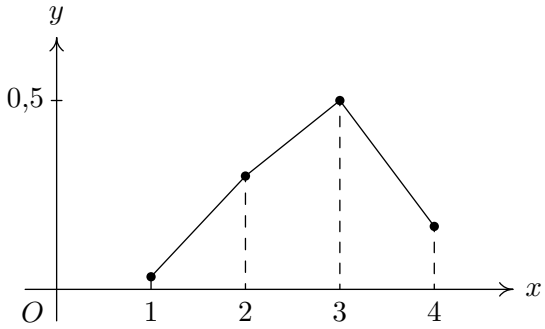
$$p_3 = P(X = 3) = \frac{C_4^3 C_6^4}{C_{10}^7} = \frac{1}{2} = 0{,}5;$$

$$p_4 = P(X = 4) = \frac{C_4^4 C_6^3}{C_{10}^7} = \frac{1}{6} \approx 0{,}17.$$

Ряд распределения показан в таблице:

x_k	1	2	3	4
p_k	1/30	3/10	1/2	1/6

а многоугольник распределения — на рис.:



Функцию распределения получим по формуле (3.1):

$$F(x) = \begin{cases} 0, & x \le 1; \\ 1/30, & 1 < x \le 2; \\ 1/30 + 3/10, & 2 < x \le 3; \\ 1/30 + 3/10 + 1/2, & 3 < x \le 4; \\ 1/30 + 3/10 + 1/2 + 1/6, & x > 4; \end{cases} =$$

$$= \begin{cases} 0, & x \le 1; \\ 1/30, & 1 < x \le 2; \\ 1/3, & 2 < x \le 3; \\ 5/6, & 3 < x \le 4; \\ 1, & x > 4. \end{cases}$$

Ее график приведен на рис. 3.6. □

Пример 3.5. *Станок непрерывно производит однотипные детали, которые подвергаются выборочному контролю. При обнаружении дефектной детали станок останавливается для переналадки. С.в. X — номер первой обнаруженной дефектной детали. Построить*

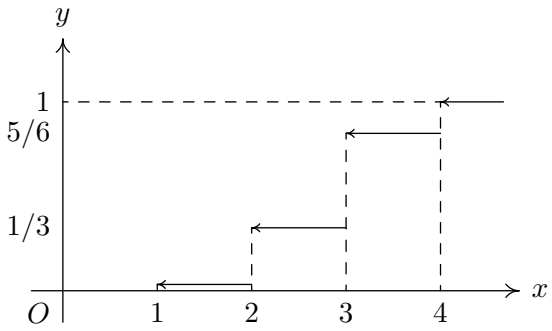


Рис. 3.6: К примеру 3.4

ряд распределения для этой с.в., считая, что вероятность появления детали с дефектом постоянна и равна p.

Решение. Обозначим вероятность того, что деталь не имеет дефекта, $q = 1 - p$. Вероятность того, что первая обнаруженная деталь с дефектом будет k-й по счету, равна вероятности того, что 1-я,...,$(k-1)$-я детали не имеют дефектов, а k-я имеет. Поэтому по теореме умножения для независимых с.с.

$$P(X = k) = p_k = pq^{k-1},$$

$k = 1, 2, \ldots$. Следовательно, ряд распределения имеет вид:

x_k	1	2	3	...	k	...
p_k	p	pq	pq^2	...	pq^{k-1}	...

□

Если привлечь формулу суммы бесконечной геометрической прогрессии, можно увидеть, что условие равенства суммы вероятностей единице выполняется:

$$\sum_{k=1}^{\infty} pq^{k-1} = p\frac{1}{1-q} = \frac{p}{p} = 1.$$

Лекция 4
Случайные векторы

4.1 Случайный вектор и его функция распределения

На практике случайные явления часто описываются не одной с.в., а несколькими, образующими систему случайных величин, или случайный вектор. Эти системы с.в. требуют специального изучения, так как их свойства не исчерпываются свойствами составляющих их случайных величин.

Случайным вектором $\mathbf{X} = (X_1, \ldots, X_n)$ называется вектор, компоненты которого являются случайными величинами.

Функцией распределения случайного вектора $\mathbf{X}$ (совместной функцией распределения с.в. $X_1, \ldots, X_n$) называется следующая вероятность:

$$F(\mathbf{x}) = F(x_1, \ldots, x_n) = P(X_1 < x_1, \ldots, X_n < x_n), \tag{4.1}$$

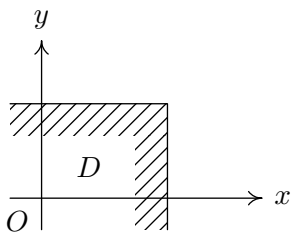
где $\mathbf{x} = (x_1, \ldots, x_n)$ — обычный вектор с действительными компонентами.

Понятие функции распределения случайного вектора и ее свойства проще всего изучить на примере двумерного случайного вектора $\mathbf{X} = (X, Y)$. Его функция распределения $F(x, y)$ в соответствии с формулой (4.1) определяется как вероятность попадания случайной точки (X, Y) в область D, изображенную на следующем рис.

Для наглядности эту область будем обозначать $\overline{x, y\big|}$. Таким образом,

$$P(x, y) = P(X < x, Y < y) = P\left((X, Y) \in \overline{x, y\big|}\right).$$

Чтобы отличить функцию распределения случайного вектора от функций распределения с.в. X и Y, будем последние обозначать $F_X(x)$, $F_Y(y)$ и называть ***частными функциями распределения***.



СВОЙСТВА ФУНКЦИИ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОГО ВЕКТОРА

1°. $0 \le F(x, y) \le 1$.

Свойство объясняется тем, что $F(x, y)$ представляет собой вероятность.

2°. *Вероятность попадания случайной точки* $\mathbf{X} = (X, Y)$ *в прямоугольник* $R = \{(x, y) : a \le x \le b,\ c \le y \le d\}$

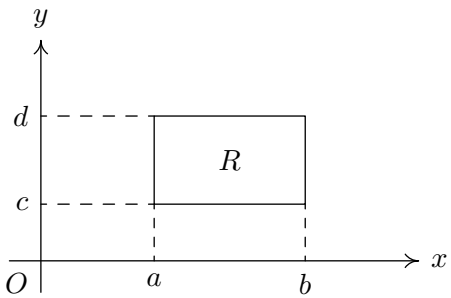


Рис. 4.1: К вероятности попадания в прямоугольник

вычисляется по формуле:

$$P(\mathbf{X} \in R) = F(b, d) + F(a, c) - F(a, d) - F(b, c).$$

Действительно, используя определение функции распределения и теорему сложения вероятностей для совместных с.с., получаем, что

$$P = \Big(\mathbf{X} \in \overline{b,d}\big|\Big) = P(\mathbf{X} \in R) + P\Big(\mathbf{X} \in \overline{a,d}\big| \cup \overline{b,c}\big|\Big) =$$

$$= P(\mathbf{X} \in R) + P\Big(\mathbf{X} \in \overline{a,d}\big|\Big) + P\Big(\mathbf{X} \in \overline{b,c}\big|\Big) - P\Big(\mathbf{X} \in \overline{a,c}\big|\Big),$$

а это равносильно доказываемому свойству.

3°. *Функция распределения не убывает как по x, так и по y.*

Пусть $x_1 < x_2$. Тогда

$$F(x_2, y) - F(x_1, y) = P(X < x_2, Y < y) - P(X < x_1, Y < y) =$$

$$= P[(X < x_1, Y < y) + (x_1 \le X < x_2, Y < y)] - P(X < x_1, Y < y) =$$

$$= P(x_1 \le X < x_2, Y < y) \ge 0,$$

откуда следует убывание $F(x, y)$ по x. Аналогично доказывается ее убывание по y.

4°. $F(x, -\infty) = F(-\infty, y) = 0$.

В самом деле, $F(x, -\infty) = P(X < x, Y < -\infty)$. Но ни одно число не может быть меньше $-\infty$ и поэтому такая вероятность должна равняться нулю. Аналогично обстоит дело со второй частью формулы.

5°. $F(x, \infty) = F_X(x),\ F(\infty, y) = F_Y(y)$.

Например, $F(x, \infty) = P(X < x, Y < \infty) = P(X < x) = F_X(x)$.

6°. $F(\infty, \infty) = 1$.

Это следует из очевидного равенства:

$$F(\infty, \infty) = P(X < \infty, Y < \infty) = 1.$$

7°. *Функция распределения непрерывна слева по каждому аргументу.*

4.2 Случайный вектор непрерывного типа

Как среди случайных величин выделяются два класса непрерывных и дискретных с.в., так и среди случайных векторов можно выделить два класса векторов, наиболее часто используемых на практике.

Если функция распределения случайного вектора $F(x, y)$ непрерывна по каждому аргументу и не имеет второй смешанной производной $F''_{xy}(x, y)$ лишь на конечном числе кривых в некоторой области на плоскости xOy, то такой случайный вектор называется ***случайным вектором непрерывного типа (СВНТ)***.

Примерами СВНТ могут служить рост и вес человека; напряжение и ток в электрической цепи; влажность, температура и давление воздуха; продолжительность эксплуатации оборудования и время его наработки на отказ и др.

Для СВНТ вместо его функции распределения $F(x, y)$ в расчетах можно использовать в некоторых отношениях более удобную характеристику , которая называется ***плотностью вероятностей***, обозначается $f(x, y)$ и определяется равенством

$$f(x, y) = F''_{xy}(x, y).$$

Зная плотность вероятностей, функцию распределения находят по формуле

$$F(x, y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(x, y)dxdy. \tag{4.2}$$

Функцию распределения или плотность вероятностей СВНТ называют его ***законом распределения***.

СВОЙСТВА ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ СЛУЧАЙНОГО ВЕКТОРА

1°. *Плотность вероятностей неотрицательна:* $f(x, y) \geq 0$.

На основании определения второй смешанной производной имеем

$$f(x,y) = F''_{xy}(x,y) = (F'_x(x,y))'_y =$$

$$= \lim_{\Delta y \to 0} \frac{F'_x(x, y+\Delta y) - F'_x(x,y)}{\Delta y} = \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \frac{H(x,y,x+\Delta x, y+\Delta y)}{\Delta x \cdot \Delta y},$$

где $H(x,y,x+\Delta x, y+\Delta y) = F(x+\Delta x, y+\Delta y) - F(x, y+\Delta y) - -F(x+\Delta x, y) + F(x,y)$. При $\Delta x \cdot \Delta y > 0$ числитель дроби представляет собой вероятность попадания СВНТ в прямоугольник с вершинами (x,y), $(x, y+\Delta y)$, $(x+\Delta x, y)$, $(x+\Delta x, y+\Delta y)$, которая, как и всякая вероятность, неотрицательна; неотрицательна, следовательно, и вся дробь. При $\Delta x \cdot \Delta y < 0$ числитель равен вероятности попадания СВНТ в указанный прямоугольник, взятой со знаком минус; но и знаменатель дроби отрицателен. Поэтому и в этом случае дробь неотрицательна. Так как предельный переход не изменяет знак результата, то $f(x,y)$ всегда неотрицательна.

2°.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)dxdy = 1.$$

На основании формулы (4.2) и свойства 6 функции распределения СВНТ получаем:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)dxdy = F(\infty,\infty) = 1.$$

3°. *Плотности вероятностей (частные) $f_X(x)$, $f_Y(y)$ с.в. X и Y, соответственно, вычисляются по формулам:*

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy, \;\; f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

В силу формулы (4.2) и свойства 5 функции распределения СВНТ справедливо равенство

$$F_X(x) = F(x, \infty) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{\infty} f(x, y) dx dy.$$

Дифференцируя его по x, приходим к первому из доказываемых выражений:

$$F_X'(x) = f_X(x) = \int\limits_{-\infty}^{\infty} f(x, y) dy.$$

Аналогично выводится вторая часть утверждения.

4°. *Вероятность попадания случайной точки (X, Y) в область D на плоскости xOy равна двойному интегралу по этой области от плотности вероятностей СВНТ $f(x, y)$:*

$$P((X, Y) \in D) = \iint\limits_{D} f(x, y) dx dy.$$

Пример 4.1. *Плотность вероятностей системы с.в. (X, Y) имеет вид*

$$f(x, y) = \begin{cases} c(x + y), & 0 \le x \le 1,\ 0 \le y \le 1; \\ 0, & \text{в остальных случаях.} \end{cases}$$

Требуется определить константу c, функцию распределения данной системы с.в., плотности вероятностей с.в. X и Y и вероятность попадания случайной точки в прямоугольник R с вершинами $(-0,5; -0,5)$, $(-0,5; 0,5)$, $(0,5; -0,5)$, $(0,5; 0,5)$.

Решение. Область D ненулевых значений плотности вероятностей заданной системы с.в., показана на рис. 4.2

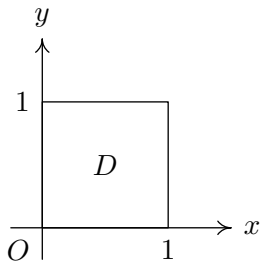


Рис. 4.2: Область ненулевых значений плотности вероятностей

Чтобы найти константу c, воспользуемся свойством 2 плотности вероятностей СВНТ:

$$1 = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y)dxdy = c\int\limits_{0}^{1} dx \int\limits_{0}^{1}(x+y)dy = c\int\limits_{0}^{1}\left(xy+\frac{y^2}{2}\right)\Bigg|_0^1 dx =$$

$$= c\int\limits_{0}^{1}\left(x+\frac{1}{2}\right)dx = c\left(\frac{x^2}{2}+\frac{x}{2}\right)\Bigg|_0^1 = c.$$

Следовательно, $c = 1$.

Функцию распределения найдем по формуле (4.2). При $x < 0$ или $y < 0$, очевидно, $F(x,y) = 0$. При $x \in [0;1]$ и $y \in [0;1]$

$$F(x,y) = \int\limits_{0}^{x}\int\limits_{0}^{y}(x+y)dxdy = \int\limits_{0}^{x}\left(xy+\frac{y^2}{2}\right)dx = \frac{1}{2}\left(x^2y+xy^2\right).$$

Если $x > 1$ а $y \in [0;1]$, то

$$F(x,y) = \int\limits_{0}^{1}\int\limits_{0}^{y}(x+y)dydx = \frac{1}{2}\left(x^2y+xy^2\right)\Bigg|_0^1 = \frac{1}{2}\left(y+y^2\right).$$

Аналогично, при $y > 1$, $x \in [0;1]$ получаем $F(x,y) = \frac{1}{2}\left(x+x^2\right)$.

Если же $x > 1$, $y > 1$, то $F(x, y) = 1$. В итоге приходим к формуле:

$$F(x,y) = \begin{cases} 0 & x \le 0 \text{ или } y \le 0; \\ xy(x+y)/2, & 0 \le x \le 1,\ 0 \le y \le 1; \\ y(y+1)/2, & x > 1,\ 0 \le y \le 1; \\ x(x+1)/2, & 0 \le x \le 1,\ y > 1; \\ 1, & x > 1, y > 1. \end{cases} \tag{4.3}$$

Вероятность попадания случайной точки (X, Y) в заданный прямоугольник найдем, воспользовавшись свойством 2 функции распределения:

$$P((X,Y) \in R) = F(0,5; 0,5) + F(-0,5; -0,5) - F(-0,5; 0,5) - \\ - F(0,5; -0,5) = \frac{1}{8} + 0 - 0 - 0 = \frac{1}{8}.$$

Функции распределения с.в. X в соответствии со свойством 5 функции распределения $F(x, y)$ равна

$$F_X(x) = F(x, \infty) = \begin{cases} 0, & x < 0; \\ (x + x^2)/2, & 0 \le x \le 1; \\ 1 & x > 1, \end{cases}$$

и поэтому

$$f_X(x) = F_X'(x) = \begin{cases} 0, & x < 0 \text{ или } x > 1; \\ x + 1/2, & 0 < x < 1. \end{cases}$$

Очевидно, что $f_Y(y) = f_X(y)$. □

4.3 Случайный вектор дискретного типа

Двумерный случайный вектор $\mathbf{X} = (X, Y)$ называется ***случайным вектором дискретного типа (СВДТ)***, если множество его возможных значений конечно или счетно. Множество возможных значений СВДТ будем представлять парами чисел (x_i, y_j), а вероятности этих значений обозначим

$$p_{ij} = P(X = x_i, Y = y_j),\ i = \overline{1,n};\ j = \overline{1,m}.$$

Совокупность всех троек чисел (x_i, y_j, p_{ij}) называется ***законом распределения*** СВДТ и оформляется в виде специальной таблицы, примеры которой будут рассмотрены дальше.

Частные законы распределения с.в. X и Y будем теперь обозначать $(x_i, p_{i\bullet})$ и $(y_j, p_{\bullet j})$ соответственно. Закон распределения СВДТ не сводится к совокупности частных законов распределения с.в. X и Y, но обратная задача всегда разрешима: зная закон распределения СВДТ, можно найти частные законы распределения его компонент по формулам:

$$p_{i\bullet} = \sum_{j=1}^{m} p_{ij}, \quad p_{\bullet j} = \sum_{i=1}^{n} p_{ij}, \tag{4.4}$$

где m, n — бесконечности или количества возможных значений с.в. X и Y соответственно.

Докажем, например, первую формулу в (4.4):

$$\begin{aligned} p_{i\bullet} &= P(X = x_i) = P[(X = x_i)\Omega] = \\ &= P[(X = x_i)(Y = y_1) + \ldots + (X = x_i)(Y = y_m)] = \\ &= P[(X = x_i)(Y = y_1)] + \ldots + P[(X = x_i)(Y = y_m)] = \sum_{j=1}^{m} p_{ij}. \end{aligned}$$

Аналогичным образом доказывается и вторая формула в (4.4).

Так как

$$\sum_{i=1}^{n} p_{i\bullet} = 1,$$

то, суммируя в первом равенстве (4.4) по i, получаем следующее свойство закона распределения СВДТ:

$$\boxed{\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1.}$$

Закон распределения СВДТ определяет его функцию распреде-

ления равенством

$$F(x,y) = \sum_{i:\, x_i<x} \sum_{j:\, y_j<y} p_{ij}.$$

Действительно, если предположить, что

$$x_1 < \ldots < x_i < x < x_{i+1} < \ldots,$$
$$y_1 < \ldots < y_j < y < y_{j+1} < \ldots,$$

то на основании определения функции распределения случайного вектора и теоремы сложения для несовместных с.с. имеем

$$\begin{aligned} F(x,y) &= P(X<x, Y<y) = \\ &= P\{[(X=x_1)+\ldots+(X=x_i)][(Y=y_1)+\ldots+(Y=y_j)]\} = \\ &= P[(X=x_1)(Y=y_1)]+\ldots+P[(X=x_1)(Y=y_j)]+\ldots+ \\ &\quad + P[(X=x_i)(Y=y_1)]+\ldots+P[(X=x_i)(Y=y_j)] = \\ &= \sum_{i:\, x_i<x} \sum_{j:\, y_j<y} p_{ij}. \end{aligned}$$

Пример 4.2. *Совместное распределение с.в. X и Y задано таблицей*

$X\backslash Y$	0	1	2
-2	$0,1$	$0,2$	$0,1$
2	$0,3$	$0,1$	$0,2$

Найти функцию распределения случайного вектора (X,Y) и безусловные законы распределения с.в. X и Y.

Решение. Так как функция распределения СВДТ определяется формулой

$$F(x,y) = \sum_{i\, x_i<x} \sum_{j:\, y_j<y} p_{ij},$$

то в данном случае она может быть представлена таблицей

$X \backslash Y$	$y \le 0$	$0 < y \le 1$	$1 < y \le 2$	$y > 2$
$x \le -2$	0	0	0	0
$-2 < x \le 2$	0	0,1	0,3	0,4
$x > 2$	0	0,4	0,7	1

Значения функции распределения стоят во внутренних клетках таблицы.

Безусловные законы распределения с.в. X и Y находим из равенств

$$p_{i\bullet} = \sum_{j=1}^{3} p_{ij}, \qquad p_{\bullet j} = \sum_{i=1}^{2} p_{ij},$$

и поэтому они представляются следующими рядами распределений:

x_i	-2	2
$p_{i\bullet}$	0,4	0,6

y_j	0	1	2
$p_{\bullet j}$	0,4	0,3	0,3

□

Лекция 5
Независимость с.в.
Функция одного случайного аргумента

5.1 Независимость с.в.

При изучении с.с. мы видели, что их независимость упрощала некоторые формулы и решение задач в целом. Независимость с.в. играет ту же роль — она не только упрощает рассуждения и выкладки, но и позволяет многие результаты получить в аналитической форме, что́ в общем случае (при отсутствии независимости с.в.) было бы невозможно.

С.в. $X_1, \ldots, X_n$ называются ***независимыми***, если их попадание в любые подмножества $B_1, \ldots, B_n$ числовой оси являются независимыми с.с.:

$$P(X_1 \in B_1, \ldots, X_n \in B_n) = P(X_1 \in B_1) \cdot \ldots \cdot P(X_n \in B_n); \qquad (5.1)$$

$B_1, \ldots, B_n \subset R$. Если это равенство нарушается для каких-либо множеств $B_1, \ldots, B_n$, то с.в. X и Y называются ***зависимыми***.

Теорема 5.1. *Пусть* $\mathbf{X} = (X, Y)$ *— СВДТ, заданная рядом распределения* (x_i, y_j, p_{ij}). *Тогда* X *и* Y *независимы, если и только если*

$$\forall (i,j) p_{ij} = P(X = x_i, Y = y_j) = P(X = x_i)(Y = y_j) = p_{i\bullet} p_{\bullet j}.$$

Доказательство. Необходимость получается из формулы (5.1) при $n = 2$, $B_1 = \{x_i\}, B_2 = \{y_j\}$. Достаточность оставим без доказательства. □

Теорема 5.2. *Пусть* $\mathbf{X} = (X, Y)$ *— СВНТ. С.в.* X *и* Y *независимы тогда и только тогда, когда*

$$f(x, y) = f_X(x) f_Y(y). \tag{5.2}$$

Доказательство. Докажем лишь необходимость (см. рис. 5.1).

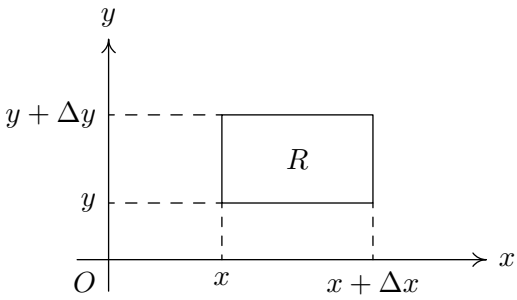


Рис. 5.1: К теореме о независимости непрерывных с.в.

Вероятность попадания случайной точки (X, Y) в прямоугольник R с одной стороны равна

$$P((X, Y) \in R) = \iint\limits_R f(x, y) dx dy$$

и по теореме о среднем для двойного интеграла

$$P((X, Y) \in R) = f(c, d) dx dy \approx f(x, y) dx dy, \ (c, d) \in R, \tag{5.3}$$

а с другой стороны по теореме умножения для зависимых с.с. имеем

$$\begin{aligned} P((X, Y) \in R) &= P(x \le X \le x + \Delta x, \, y \le Y \le y + \Delta y) = \\ &= P(x \le X \le x + \Delta x) P(\, y \le Y \le y + \Delta y / x \le X \le x + \Delta x). \end{aligned}$$

В силу независимости с.в. X и Y

$$P(\, y \le Y \le y + \Delta y / x \le X \le x + \Delta x) = P(\, y \le Y \le y + \Delta y)$$

и поэтому по теореме о среднем уже для одномерного определенного интеграла

$$P((X,Y)\in R) = P(x\le X\le x+\Delta x)P(\,y\le Y\le y+\Delta y) =$$
$$= \int\limits_{x}^{x+\Delta x} f_X(x)dx \int\limits_{y}^{y+\Delta y} f_Y(y)dy \approx f_X(x)dx\cdot f_Y(y)dy.$$

Используя это выражение и равенство (5.3), получаем

$$f(x,y)dxdy \approx f_X(x)dx\cdot f_Y(y)dy.$$

Сокращая на $dxdy$ и переходя к пределу при $dx\to 0$, $dy\to 0$, приходим к точной формуле (5.2). □

Если равенство (5.2) не выполняется, то вводят функцию

$$f(y/x) = \frac{f(x,y)}{f_X(x)}, \tag{5.4}$$

называемую ***условной плотностью вероятностей*** с.в. Y по отношению к с.в. X. Тогда $f(x,y) = f(y/x)f_X(x)$. Аналогично можно получить, что $f(x,y) = f(x/y)f_Y(y)$.

Если X и Y независимы, то из формул (5.4) и (5.2) находим, что

$$f(y/x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y),$$
$$f(x/y) = f_X(x),$$

то есть условные плотности вероятностей равны ***безусловным*** $f_X(x)$, $f_Y(y)$.

Пример 5.1. *Плотность вероятностей двух с.в. X и Y равна*

$$f(x,y) = \begin{cases} \lambda\mu e^{-(\lambda x+\mu y)}, & x\ge 0,\ y\ge 0; \\ 0, & x\le 0 \text{ или } y\le 0, \end{cases}$$

λ, $\mu > 0$. *Доказать, что X и Y независимы.*

Решение. Найдем плотность вероятностей с.в. X:

$$f_X(x) = \int\limits_0^\infty \lambda\mu e^{-(\lambda x+\mu y)}dy = \lambda\mu\frac{1}{-\mu}e^{-(\lambda x+\mu y)}\Big|_0^\infty =$$
$$= -\lambda e^{-\lambda x}(e^{-\mu\infty} - 1) = \lambda e^{-\lambda x},$$

если $x \geq 0$, и $f_X(x) = 0$, если $x < 0$. Аналогичным образом получаем плотность вероятностей с.в. Y:

$$f_Y(y) = \begin{cases} \mu e^{-\mu y}, & y \geq 0; \\ 0 & y < 0. \end{cases}$$

Очевидно выполнение равенства

$$f(x, y) = f_X(x)f_Y(y),$$

и поэтому с.в. X и Y независимы. □

Если $\mathbf{X} = (X, Y)$ — СВДТ с законом распределения (x_i, y_j, p_{ij}) и с.в. X и Y зависимы, то ***условным законом распределения компоненты*** $\mathbf{X}$ при условии, что компонента Y приняла значение y_j, называется ряд распределения $(x_i, p_{i/j\bullet})$, где

$$p_{i/j\bullet} = P(X = x_i/Y = y_j) = \frac{p_{ij}}{p_{\bullet j}}.$$

Аналогично определяется ***условный закон распределения компоненты*** $\mathbf{Y}$ при условии, что компонента X приняла значение x_i:

$$p_{\bullet j/i} = P(Y = y_j/X = x_i) = \frac{p_{ij}}{p_{i\bullet}}.$$

Так же, как и для СВНТ, доказывается, что для независимых с.в. X и Y условные законы распределения совпадают с безусловными:

$$p_{i/j\bullet} = p_{i\bullet},\ p_{\bullet j/i} = p_{\bullet j}.$$

5.2 Функция одного случайного непрерывного аргумента

Пусть $y = \varphi(x)$ — обычная функция одного аргумента x. Если вместо x в эту формулу подставить с.в. X, получим ***функцию случайной величины***, которая и сама становится случайной величиной Y:

$$Y = \varphi(X).$$

Ее еще называют ***случайной функцией*** (с.ф.).

Предположим, что с.в. X непрерывна и все ее значения принадлежат интервалу (a, b), а функция $\varphi(x)$ дифференцируема, строго возрастает на (a, b) и имеет в этом интервале обратную функцию $x = \varphi^{-1}(y)$. Найдем функцию распределения с.в. Y:

$$F_Y(y) = P(Y < y) = P(\varphi(X) < y) = P(X < \varphi^{-1}(y)) = F_X(\varphi^{-1}(y)).$$

Продифференцировав это равенство, получим и плотность вероятностей с.в. Y:

$$f_Y(y) = f_X\left((\varphi^{-1}(y)\right)\left[\varphi^{-1}(y)\right]'. \tag{5.5}$$

Если же $\varphi(x)$ строго убывает на интервале (a, b), то

$$F_Y(y) = P(Y < y) = P(\varphi(X) < y) = P(X > \varphi^{-1}(y)) = 1 - F_X(\varphi^{-1}(y)).$$

Значит,

$$f_Y(y) = F_Y'(y) = -f_X\left((\varphi^{-1}(y)\right)\left[\varphi^{-1}(y)\right]'. \tag{5.6}$$

Объединяя равенства (5.5) и (5.6) в одно, получаем

$$\boxed{f_Y(y) = f_X\left((\varphi^{-1}(y)\right)\left|\left[\varphi^{-1}(y)\right]'\right|.}$$

Пусть теперь функция $\varphi(x)$ кусочно-монотонна и пусть y — фиксированное число, а $J_1(y), \ldots, J_n(y)$ — интервалы вида $[\cdot, \cdot)$, на которых выполняется условие $\varphi(x) < y$ (см. рис. 5.2).

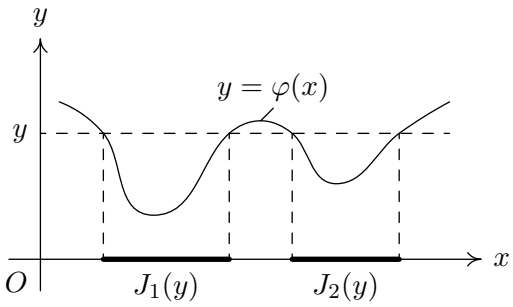


Рис. 5.2: К определению распределения функции с.в.

Тогда функция распределения с.в. Y на основании теоремы сложения для несовместных с.с. имеет вид

$$F_Y(y) = P(Y < y) = P\left(\sum_{i=1}^{n}(X \in J_i(y))\right) =$$
$$= \sum_{i=1}^{n} P(X \in J_i(y)) = \sum_{i=1}^{n} \int\limits_{J_i(y)} f_X(x)dx.$$

Из этого равенства можно найти и плотность вероятностей с.в. Y, если воспользоваться определением плотности вероятностей: $f_Y(y) = F_Y'(y)$.

Пример 5.2. *Найти закон распределения с.в.*

$$Y = \min(X, X^2),$$

если с.в. X имеет плотность вероятностей $f_X(x)$.

Решение. График функции $y = \varphi(x) = \min(x, x^2)$ имеет вид, показанный на рис. 5.3.

Функция возрастает на всей числовой оси, обратная к ней функция представляется выражением

$$\varphi^{-1}(y) = \begin{cases} y, & y \notin (0;1); \\ \sqrt{y}, & y \in (0;1); \end{cases}$$

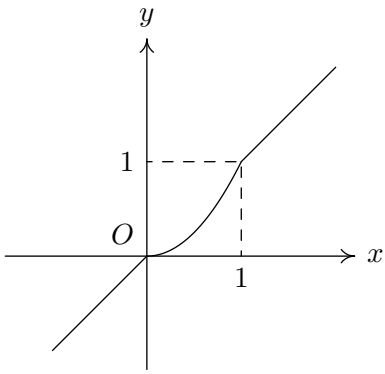


Рис. 5.3: График функции $y = \min(x, x^2)$.

и имеет производную

$$\left[\varphi^{-1}(y)\right]' = \begin{cases} 1, & y \notin (0;1); \\ \frac{1}{2\sqrt{y}}, & y \in (0;1). \end{cases}$$

Поэтому в соответствии с формулой (5.5) плотность вероятностей с.в. Y равна

$$f_Y(y) = \begin{cases} f_X(y), & y \notin (0;1); \\ \frac{f_X(\sqrt{y})}{2\sqrt{y}}, & y \in (0;1); \end{cases}$$

□

5.3 Функция одного случайного дискретного аргумента

Если с.в. X дискретна, то дискретна и с.в. $Y = \varphi(X)$. Закон ее распределения найдем, используя формулу полной вероятности и связь

между законами распределения СВДТ и его компонент:

$$P(Y = y_j) = \sum_{i=1}^{n} P(X = x_i,\, Y = y_j) =$$
$$= \sum_{i=1}^{n} P(Y = y_j / X = x_i) P(X = x_i).$$

Но

$$P(Y = y_j / X = x_i) = \begin{cases} 1, & \text{если } \varphi(x_i) = y_j; \\ 0, & \text{если } \varphi(x_i) \neq y. \end{cases}$$

Значит,

$$\boxed{P(Y = y_j) = \sum_{i:\, \varphi(x_i) = y_j} P(X = x_i).}$$

Если все y_j различны, то

$$P(Y = y_j) = P(X = x_j), \tag{5.7}$$

считая, что $y_j = \varphi(x_j)$.

Пример 5.3. *Для условий примера 4.2 проверить независимость с.в. X и Y, найти условные законы распределения $(x_i, p_{i/2\bullet})$ и $(y_j, p_{\bullet 1/j})$, найти законы распределения функций $Z = Y^2$ и $T = 2 - X$.*

С.в. X и Y не являются независимыми, так как справедливо соотношение

$$p_{11} = 0{,}1 \neq p_{1\bullet} p_{\bullet 1} = 0{,}4 \cdot 0{,}4 = 0{,}16.$$

Условные законы распределения, основываясь на формулах

$$p_{i/j\bullet} = \frac{p_{ij}}{p_{\bullet j}}, \qquad p_{\bullet j/i} = \frac{p_{ij}}{p_{i\bullet}},$$

запишем в виде рядов распределений:

x_i	-2	2
$p_{i/2\bullet}$	$2/3$	$1/3$

y_j	0	1	2
$p_{\bullet 1/j}$	$1/4$	$2/4$	$1/4$

Так как с.в. X и Y принимают только различные значения, то законы распределения с.в. $Z = Y^2$ и $T = 2 - X$ в соответствии с равенством (5.7) имеют вид:

z_j	0	2	4
p_j	$0,4$	$0,3$	$0,3$

t_i	0	4
p_i	$0,6$	$0,4$

Лекция 6
Функции нескольких с.в. Сумма случайных величин

6.1 Функция случайного вектора непрерывного типа

Рассмотрим обычную функцию нескольких аргументов

$$z = \varphi(x_1, \ldots, x_n)$$

и вместо действительных аргументов $x_1, \ldots, x_n$ подставим в нее случайные величины $X_1, \ldots, X_n$, образующие СВНТ $\mathbf{X} = (X_1, \ldots, X_n)$. В результате получится ***случайная функция нескольких случайных аргументов*** $Z = \varphi(X_1, \ldots, X_n)$, значением которой является с.в. Z.

В основном мы будем изучать случайные функции двух случайных аргументов вида $Z = \varphi(X, Y)$. Главной задачей будет отыскание закона распределения с.в. Z при известном законе распределения СВНТ $\mathbf{X} = (X, Y)$.

Пусть $f(x, y)$ — совместная плотность вероятностей с.в. X и Y, тогда функция распределения с.в. Z на основании ее определения и свойств плотности вероятностей вычисляется следующим образом:

$$\begin{aligned} F_Z(z) &= P(Z < z) = P(\varphi(x, y) < z) = P(\mathbf{X} \in D(z)) = \\ &= \iint\limits_{D(z)} f(x, y) dx dy, \end{aligned} \tag{6.1}$$

где $D(z) = \{(x, y) : \varphi(x, y) < z\}$. Далее можно найти и плотность вероятностей с.в. Z:

$$f_Z(z) = F_Z'(z).$$

Пример 6.1. *Закон распределения случайного вектора* $\mathbf{X} = (X, Y)$ *определяется плотностью вероятностей* $f(x, y)$. *Найти плотность вероятностей частного* Y/X.

Решение. Частное Y/X представляет собой функцию двух случайных переменных X и Y, то есть $Z = Y/X = \varphi(X, Y)$. Область $D(z)$, по которой проводится интегрирование в формуле (6.1), указана штриховкой на рис. 6.1 и определяется выражением

$$D(z) = \left\{ \ (x, y) : \tfrac{y}{x} < z \ \right\}.$$

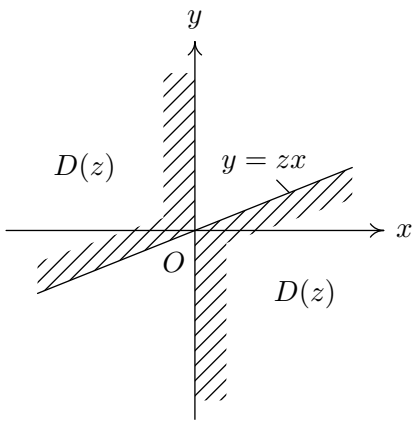


Рис. 6.1: Область $D(z)$ для частного двух с.в.

Следовательно, функция распределения с.в. Z равна

$$F_Z(z) = \iint\limits_{D(z)} f(x, y)dxdy = \int\limits_{-\infty}^{0} dx \int\limits_{zx}^{\infty} f(x, y)dy + \int\limits_{0}^{\infty} dx \int\limits_{-\infty}^{zx} f(x, y)dy.$$

Далее получаем плотность вероятностей с.в. Z:

$$f_Z(z) = F_Z'(z) = -\int\limits_{-\infty}^{0} xf(x, zx)dx + \int\limits_{0}^{\infty} xf(x, zx)dx.$$

В частности, если с.в. X и Y независимы, то

$$f_Z(z) = -\int\limits_{-\infty}^{0} xf_X(x)f_Y(zx)dx + \int\limits_{0}^{\infty} xf_X(x)f_Y(zx)dx.$$

□

Закон распределения случайной функции $Z = \varphi(X_1, \ldots, X_n)$ при $n > 2$ определяется формулой, аналогичной формуле (6.1):

$$F_Z(z) = P(Z < \varphi(x_1, \ldots, x_n)) = \int\limits_{D(z)} \cdots \int f(x_1, \ldots, x_n)dx_1 \ldots dx_n, \tag{6.2}$$

где $D(z) = \{(x_1, \ldots, x_n) : \varphi(x_1, \ldots, x_n) < z\}$, $f(x_1, \ldots, x_n)$ — плотность вероятностей СВНТ $\mathbf{X} = (X_1, \ldots, X_n)$.

Пример 6.2. *С.в. $X_1, \ldots, X_n$ независимы и одинаково распределены с функцией распределения $F_X(x)$. Найти функции распределения и плотности вероятностей с.в. $Y = \min(X_1, \ldots, X_n)$ и $Z = \max(X_1, \ldots, X_n)$.*

Решение. Используя формулу (6.2) и независимость с.в. $X_1, \ldots, X_n$, находим

$$\begin{aligned}
F_Y(y) &= P(Y < y) = 1 - P(Y \geq y) = 1 - P[\min(X_1, \ldots, X_n) \geq y] = \\
&= 1 - P(X_1 \geq y, \ldots, X_n \geq y) = \\
&= 1 - P(X_1 \geq y) \cdot \ldots \cdot P(X_n \geq y) = \\
&= 1 - [1 - P(X_1 < y)] \cdot \ldots \cdot [1 - P(X_n < y)] = \\
&= 1 - [1 - F_X(y)] \cdot \ldots \cdot [1 - F_X(y)] = 1 - [1 - F_X(y)]^n,
\end{aligned}$$

$$F_Z(z) = P(Z < z) = P[\max(X_1, \ldots, X_n) < z] =$$
$$= P(X_1 < z, \ldots, X_n < z) = P(X_1 < z) \cdot \ldots \cdot P(X_n < z) =$$
$$= F_X(z) \cdot \ldots \cdot F_X(z) = [F_X(z)]^n.$$

Дифференцируя функции распределения с.в. Y и Z, получаем их плотности вероятностей:

$$f_Y(y) = F_Y^{'}(y) = -n[1 - F_x(y)]^{n-1}(-1)F_X^{'}(y) =$$
$$= n[1 - F_X(y)]^{n-1} f_X(y),$$

$$f_Z(z) = F_Z^{'}(z) = n[F_X(z)]^{n-1} f_X(z).$$

Мы видим, что эти плотности выражаются не только через плотность вероятностей с.в. X, но и через ее функцию распределения. □

6.2 Функция случайного вектора дискретного типа

Если $\mathbf{X} = (X, Y)$ — СВДТ с законом распределения (x_i, y_j, p_{ij}), то функция $Z = \varphi(X, Y)$ продуцирует дискретную с.в. Z, функция распределения которой определяется формулой

$$F_Z(z) = P(\varphi(X, Y) < z) = \sum_{i,j:\varphi(x_i,y_j)<z} p_{ij}. \tag{6.3}$$

где суммирование ведется по всем индексам i и j, для которых выполняется неравенство $\varphi(x_i, y_j) < z$.

Пример 6.3. *Для условий примера 4.2 предыдущей лекции найти функцию и ряд распределения с.в. $Z = XY$.*

Решение. Используя выражение (6.3), находим

$$F_Z(z) = P(XY) < z) = \sum_{i,j:x_i y_j<z} p_{ij} =$$

$$= \begin{cases} 0, & z \le -4; \\ p_{13} = 0,1, & -4 < z \le -2; \\ p_{12} + p_{13} = 0,3, & -2 < z \le 0; \\ p_{11} + p_{12} + p_{13} + p_{21} = 0,7, & 0 < z \le 2; \\ p_{11} + p_{12} + p_{13} + p_{21} + p_{22} = 0,8, & 2 < z \le 4; \\ p_{11} + p_{12} + p_{13} + p_{21} + p_{22} + p_{23} = 1, & z > 4; \end{cases}.$$

Следовательно, ряд распределения имеет вид:

x_j	-4	-2	0	2	4
p_j	$0,1$	$0,2$	$0,4$	$0,1$	$0,2$

□

6.3 Сумма случайных величин. Композиция законов распределения

Очень важной функциональной зависимостью, часто встречающейся на практике, является сумма двух с.в.: $Z = \varphi(X, Y) = X + Y$. Чтобы найти ее закон распределения, прежде всего изобразим область $D(z) = \{(x, y) : x + y < z\} = \{(x, y) : y < z - x\}$ из формулы (6.1) на плоскости xOy (см. рис. 6.2.).

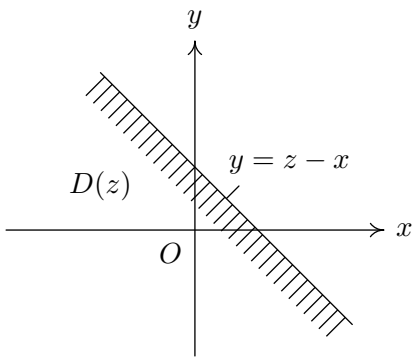


Рис. 6.2: Область $D(z)$ для суммы двух с.в.

Используя указанную формулу и рисунок для определения пределов интегрирования, получаем функцию распределения суммы двух с.в.:

$$F_Z(z) = \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{z-x} f(x,y)dy.$$

Соответствующая плотность вероятностей находится дифференцированием данного выражения:

$$f_Z(z) = F_Z'(z) = \int\limits_{-\infty}^{\infty} f_X(x, z-x)dx.$$

Если X и Y независимы, то эта формула преобразуется к виду:

$$\boxed{f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx.} \tag{6.4}$$

На основании симметричности операции сложения относительно слагаемых последнее равенство можно переписать так:

$$f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(z-x) f_Y(x)dx.$$

Таким образом, *плотность вероятностей суммы двух независимых с.в. является сверткой плотностей вероятностей слагаемых*:

$$\boxed{f_Z(z) = f_X(x) * f_Y(y),}$$

и называется также ***композицией*** законов распределения с.в. X и Y.

Если $\mathbf{X} = (X, Y)$ — СВДТ, то закон распределения суммы двух с.в. может быть представлен в виде

$$P(Z = z_k) = P(X + Y = z_k) = \sum_{i,j: x_i + y_j = z_k} p_{ij}, \tag{6.5}$$

где суммирование ведется по тем индексам i и j, для которых выполняется равенство $x_i + y_j = z_k$.

При независимости с.в. X и Y последнее равенство превращается в следующее:

$$P(Z = z_k) = \sum_{i=1}^{n} P(X = x_i)P(Y = z_k - x_i),$$

которое можно назвать ***дискретной композицией*** законов распределения с.в. X и Y.

Пример 6.4. *Пассажир ездит из города на дачу двумя электричками с пересадкой. Первую электричку ему приходится ждать не более 20 мин., а вторую — не более 40 мин. Считая времена ожидания X и Y первой и второй электричек независимыми с.в., имеющими плотности вероятностей*

$$f_X(x) = \begin{cases} 1/20, & x \in [0;20]; \\ 0, & x \notin [0;20]; \end{cases} \qquad f_Y(y) = \begin{cases} 1/40, & x \in [0;40]; \\ 0, & x \notin [0;40]; \end{cases}$$

найти плотность вероятностей суммарного времени ожидания.

Решение. Воспользуемся формулой (6.4):

$$f_{X+Y}(z) = \int\limits_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \frac{1}{20}\int\limits_{0}^{20} f_Y(z-x)dx.$$

Очевидно, что

$$\begin{aligned} f_Y(z-x) \neq 0 \wedge 0 \leq x \leq 20 &\Longleftrightarrow 0 \leq z - x \leq 40 \wedge 0 \leq x \leq 20 \Longleftrightarrow \\ &\Longleftrightarrow z - 40 \leq x \leq z \wedge 0 \leq x \leq 20 \Longleftrightarrow \\ &\Longleftrightarrow \max(z-40, 0) \leq x \leq \min(z, 20). \end{aligned}$$

Следовательно,

$$f_{X+Y}(z) = \begin{cases} \frac{1}{800}\int\limits_0^z dx = \frac{z}{800}, & 0 \le z \le 20; \\ \frac{1}{800}\int\limits_0^{20} dx = \frac{1}{40}, & 20 \le z \le 40; \\ \frac{1}{800}\int\limits_{z-40}^{20} dx = \frac{60-z}{800}, & 40 \le z \le 60; \\ 0, & z > 60. \end{cases}$$

График полученной плотности вероятностей показан на рисунке 6.3.

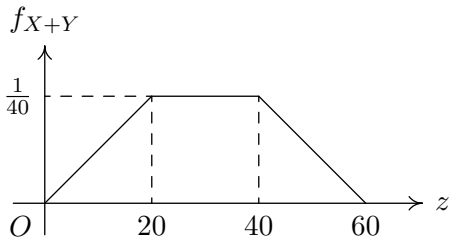


Рис. 6.3: Плотность вероятностей суммарного времени ожидания

□

Пример 6.5. *Найти сумму с.в. X и Y из примера 4.2.*

Решение. С.в. $Z = X + Y$ может принимать лишь следующие значения: $-2, -1, 0, 2, 3, 4$. В соответствии с формулой (6.5) получаем

$$P(Z=-2) = p_{11} = 0,1; \quad P(Z=-1) = p_{12} = 0,2;$$
$$P(Z=0) = p_{13} = 0,1; \quad P(Z=2) = p_{21} = 0,3;$$
$$P(Z=3) = p_{22} = 0,1; \quad P(Z=4) = p_{23} = 0,2.$$

Таким образом, ряд распределения суммы с.в. x и y имеет вид:

z_k	-2	-1	0	2	3	4
p_k	$0,1$	$0,2$	$0,1$	$0,3$	$0,1$	$0,2$

□

Лекция 7
Числовые характеристики с.в. Математическое ожидание

Хотя функция распределения, плотность вероятностей или дискретный закон распределения дают полную информацию о с.в. и случайных векторах, далеко не всегда эти характеристики удается получить на практике. Особенно это касается многомерных распределений. Если особенности решаемой задачи не помогают сделать вывод даже о типе распределения с.в., приходится прибегать к методам математической статистики, которая требует, например, для надежного определения функции распределения одномерной с.в. от 500 до 1000 наблюдений за ее реализациями. Проведение такого количества экспериментов не всегда возможно, так как опыты могут потребовать больших затрат времени (например, в сельском хозяйстве обычно на один опыт уходит один год) или значительных финансовых вложений (например, чтобы найти предел прочности какого-либо дорогостоящего устройства, его надо сломать).

В подобных обстоятельствах обычно ограничиваются определением так называемых числовых характеристик, которые дают представление лишь о некоторых особенностях поведения с.в. и случайных векторов. Впрочем, сами характеристики настолько важны, что имеет смысл найти их даже при известном законе распределения с.в. Этой последней задачей мы сейчас и займемся, а получение приближенных значений числовых характеристик с.в. по наблюдениям за ее реализациями отложим до лекций по математической статистике.

Первой числовой характеристикой, которую мы изучим, является математическое ожидание, описывающее поведение с.в. X «в среднем». Часто используются термины: средний рост, средний вес, средняя скорость молекул, средняя производительность труда и т.п. Все это — в определенном смысле примеры математических ожиданий.

Математическим ожиданием непрерывной с.в. называется несобственный интеграл (если он абсолютно сходится)

$$\boxed{\mathbf{M}X = \int\limits_{-\infty}^{\infty} x f(x) dx,} \tag{7.1}$$

где $f(x)$ — плотность вероятностей с.в. X.

Математическим ожиданием дискретной с.в. X называется сумма или ряд (в зависимости от того, чем является n — числом или символом ∞):

$$\boxed{\mathbf{M}X = \sum_{i=1}^{n} x_i p_i,} \tag{7.2}$$

где x_i — возможные значения с.в. X, а $p_i = P(X = x_i)$, — вероятности этих значений.

В формулах (7.1) и (7.2) мы прежде всего замечаем, что значения с.в. суммируются с некоторым «весом»: для непрерывной с.в. таким «весом» служит плотность вероятностей, а для дискретной с.в. — вероятности p_i. Таким образом, более вероятные значения с.в. (чаще реализующиеся в опытах) вносят больший вклад в математическое ожидание. В результате формулы (7.1) и (7.2) действительно определяют некоторое среднее значение с.в. X, называемое в теории вероятностей математическим ожиданием.

Понятно, что наличие в формулах (7.1), (7.2) несобственного интеграла и ряда может приводить к тому, что математическое ожидание не будет существовать (если интеграл или ряд расходятся).

Математическое ожидание случайного вектора $\mathbf{X} = (X_1, \ldots, X_n)$ определяется формулой

$$\mathbf{M}\mathbf{X} = (\mathbf{M}X_1, \ldots, \mathbf{M}X_n).$$

Прежде чем приступить к изучению свойств математического ожидания, условимся считать постоянную величину C случайной с плотностью вероятностей $f(x) = \delta(x - C)$, где $\delta(x)$ — известная вам дельта-функция. Тогда

$$P(X = C) = \lim_{\varepsilon \to 0} \int\limits_{C-\varepsilon}^{C+\varepsilon} \delta(x - C)dx = \langle z = x - C, dz = dx \rangle =$$

$$= \lim_{\varepsilon \to 0} \int\limits_{-\varepsilon}^{\varepsilon} \delta(x)dx = \lim_{\varepsilon \to 0} 1 = 1,$$

$$P(X \neq C) = 1 - P(X = C) = 0.$$

Следовательно, с вероятностью 1 постоянная величина принимает свое собственное значение, а другие значения — с вероятностью 0. Такие величины, в отличие от «настоящих» с.в., будем называть ***детерминированными***.

А теперь перейдем к рассмотрению свойств математического ожидания, проводя необходимые доказательства только для непрерывных с.в.

СВОЙСТВА МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

1°. *Математическое ожидание детерминированной величины равно самой этой величине:*

$$\mathbf{M}C = C.$$

На основании только что введенного определения детерминированной величины и одного из свойств дельта-функции получаем

$$\mathbf{M}C = \int\limits_{-\infty}^{\infty} x\delta(x - C)dx = C.$$

2°. *Математическое ожидание функции одного случайного аргумента находится по формуле*

$$\mathbf{M}\varphi(X) = \int\limits_{-\infty}^{\infty} \varphi(x) f(x) dx$$

для непрерывной с.в. X с плотностью вероятностей $f(x)$, а для дискретной с.в., заданной рядом распределения (x_i, p_i), — по формуле

$$\mathbf{M}\varphi(X) = \sum_{i=1}^{n} \varphi(x_i) p_i.$$

3°. *Постоянную величину можно выносить за знак математического ожидания:*

$$\mathbf{M}(CX) = C \cdot \mathbf{MX}.$$

Действительно, используя предыдущее свойство для $\varphi(x) = Cx$, получаем

$$\mathbf{M}(CX) = \int\limits_{-\infty}^{\infty} Cx f(x) dx = C \int\limits_{-\infty}^{\infty} x f(x) dx = C\mathbf{M}X.$$

4°. *Математическое ожидание функции двух случайных аргументов $\varphi(X, Y)$ для СВНТ вычисляется по формуле*

$$\mathbf{M}\varphi(X, Y) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \varphi(x, y) f(x, y) dx dy,$$

где $f(x, y)$ — совместная плотность вероятностей с.в. X и Y; а для СВДТ — по формуле

$$\mathbf{M}\varphi(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} \varphi(x_i, y_j) p_{ij},$$

где (x_i, y_i, p_{ij}) — закон распределения случайного вектора $\mathbf{X} = (X, Y)$.

5°. *Математическое ожидание суммы двух с.в. равно сумме их математических ожиданий:*

$$\boxed{\mathbf{M}(X+Y) = \mathbf{M}X + \mathbf{M}Y.}$$

Используя предыдущее свойство для функции $\varphi(x,y) = x+y$, получаем

$$\begin{aligned}
\mathbf{M}(X+Y) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x+y)f(x,y)dxdy = \\
&= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}xf(x,y)dxdy + \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}yf(x,y)dxdy = \\
&= \int\limits_{-\infty}^{\infty}x\left[\int\limits_{-\infty}^{\infty}f(x,y)dy\right]dx + \int\limits_{-\infty}^{\infty}y\left[\int\limits_{-\infty}^{\infty}f(x,y)dx\right]dy = \\
&= \int\limits_{-\infty}^{\infty}xf_X(x)dx + \int\limits_{-\infty}^{\infty}yf_Y(y)dy = \mathbf{M}X + \mathbf{M}Y.
\end{aligned}$$

6°. *Если с.в. X и Y независимы, то математическое ожидание их произведения равно произведению их математических ожиданий:*

$$\mathbf{M}XY = \mathbf{M}X \cdot \mathbf{M}Y.$$

Снова обратимся к свойству 4, принимая $\varphi(x,y) = xy$ и учитывая, что совместная плотность двух независимых с.в. раскладывается в произведение их частных плотностей: $f(x,y) = f_X(x)\cdot f_Y(y)$, и поэтому

$$\begin{aligned}
\mathbf{M}XY &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}xyf(x,y)dxdy = \int\limits_{-\infty}^{\infty}xf_X(x)dx\int\limits_{-\infty}^{\infty}yf_Y(y)dy = \\
&= \mathbf{M}X \cdot \mathbf{M}Y.
\end{aligned}$$

Пример 7.1. *Найти математическое ожидание с.в. X, имеющей плотность вероятностей*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right].$$

Решение. Воспользуемся определением математического ожидания непрерывной с.в.:

$$\begin{aligned}
\mathbf{M}X &= \int\limits_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\infty} x \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx = \\
&= \left\langle z = \frac{x-m}{\sqrt{2}\sigma}, dx = \sqrt{2}\sigma dz \right\rangle = \\
&= \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} (m + \sqrt{2}\sigma z) \exp\left(-z^2\right) dz = \\
&= \frac{m}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} \exp\left(-z^2\right) dz + \sqrt{\frac{2}{\pi}}\,\sigma \int\limits_{-\infty}^{\infty} z \exp\left(-z^2\right) dz.
\end{aligned}$$

Первый интеграл в правой части полученного выражения равен $\sqrt{\pi}$, так как является удвоенным интегралом Пуассона (см. приложение C), а второй равен нулю, потому что подынтегральная функция нечетна. Следовательно, $\mathbf{M}X = m$. □

Пример 7.2. *Имеется n идентичных блоков аппаратуры, каждый из которых с вероятностью p может иметь дефект. Блок подсоединяется к электронному устройству и включается ток; при обнаружении дефекта блок сразу же заменяется другим блоком. Найти математическое ожидание с.в. X — числа блоков, которые будут испробованы.*

Решение. Нетрудно видеть, что ряд распределения данной с.в. имеет

вид

x_k	1	2	3	$\cdots$	i	$\cdots$	n
p_k	q	pq	p^2q	$\cdots$	$p^{i-1}q$	$\cdots$	p^{n-1}

где $q = 1 - p$. Используя формулу суммы конечной геометрической прогрессии, получаем

$$\begin{aligned}
\mathbf{M}X &= \sum_{i=1}^{n-1} ip^{i-1}q + np^{n-1} = \left(\sum_{i=1}^{n-1} p^i\right)'_p q + np^{n-1} = \\
&= q\left(\frac{p-p^n}{1-p}\right)'_p + np^{n-1} = \\
&= q\,\frac{(1-np^{n-1})(1-p)+p-p^n}{(1-p)^2} + np^{n-1} = \\
&= \frac{1-np^{n-1}-p+np^n+p-p^n+np^{n-1}-np^n}{1-p} = \frac{1-p^n}{1-p}.
\end{aligned}$$

□

Если с.в. X принимает только целые неотрицательные значения, то ее математическое ожидание удобно находить с помощью ***производящей функции***

$$\boxed{G(u) = \sum_{i=0}^{\infty} p_i u^i,}$$

где $p_i = P(X = i)$, а $u, |u| \le 1$, — действительная или комплексная переменная.

Так как

$$G'(u) = \sum_{i=1}^{\infty} ip_i u^{i-1},$$

то

$$\boxed{\mathbf{M}X = G'(1).} \tag{7.3}$$

Пример 7.3 (закон Пуассона). *Дискретная с.в. X задана рядом распределения*

$$\left(i, \frac{\lambda^i}{i!}e^{-\lambda}\right), i = 0, 1, \ldots.$$

Вычислить ее математическое ожидание.

Решение. Составим производящую функцию заданной с.в. и найдем ее производную:

$$G(u) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}e^{-\lambda}u^i = e^{-\lambda}\sum_{i=0}^{\infty}\frac{(\lambda u)^i}{i!} = e^{-\lambda}e^{\lambda u} = e^{\lambda(u-1)},$$
$$G'(u) = \lambda e^{\lambda(u-1)}.$$

По формуле (7.3) находим: $\mathbf{M}X = G'(1) = \lambda$. □

Пример 7.4. *С.в. X имеет плотность вероятностей*

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0 & x < 0; \end{cases}$$

$\lambda > 0$. *Установить, при каких условиях существует и чему равно математическое ожидание с.в. $Y = e^X$.*

Решение. Используя свойство 2 математического ожидания, получаем

$$\mathbf{M}Y = \mathbf{M}\mathbf{e}^X = \int_{-\infty}^{\infty} e^x f(x)dx = \int_0^{\infty} e^x \lambda e^{-\lambda x}dx = \lambda\int_0^{\infty} e^{(1-\lambda)x}dx =$$
$$= \frac{\lambda}{1-\lambda}e^{(1-\lambda)x}\Big|_0^{\infty}.$$

При $\lambda \leq 1$ интеграл расходится и математическое ожидание не существует; при $\lambda > 1$ интеграл сходится и математическое ожидание равно

$$\mathbf{M}Y = \frac{\lambda}{\lambda - 1}.$$

□

Пример 7.5. *Совместное распределение с.в. X и Y задано таблицей*

$X \backslash Y$	-1	2
-1	$0,4$	$0,3$
1	$0,2$	$0,1$

Найти математическое ожидание с.в. $Z = X^2 - Y^2$.

Решение. Используем свойство 4 математического ожидания для СВДТ применительно к функции $\varphi(x, y) = x^2 - y^2$. Получим:

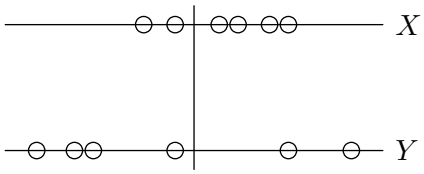
$$\mathbf{M}Z = \mathbf{M}(X^2 - Y^2) = [(-1)^2 - (-1)^2] \cdot 0{,}4 + [1^2 - (-1)^2] \cdot 0{,}2 + \\ + [(-1)^2 - 2^2] \cdot 0{,}3 + (1^2 - 2^2) \cdot 0{,}1 = -0{,}9 - 0{,}3 = -1{,}2.$$

□

Лекция 8
Дисперсия. Ковариация. Коэффициент корреляции

8.1 Дисперсия с.в.

Предположим, что имеется два типа аппаратуры, причем, X — время безотказной работы аппаратуры I типа, а Y — такое же время для аппаратуры II типа. Предположим также, что среднее время безотказной работы аппаратуры обоих видов одинаково: $\mathbf{M}X = \mathbf{M}Y$. На рисунке изображены наблюдения за этими с.в. Видно, что наблюдения за с.в. X ложатся более «кучно», хотя математические ожидания обеих с.в. равны. Это позволяет, например, плановый ремонт аппаратуры I типа проводить реже, чем аналогичное мероприятие для аппаратуры второго типа.



$\mathbf{M}X = \mathbf{M}Y$

Различие в поведении рассмотренных нами с.в. X и Y теория вероятностей фиксирует с помощью такой характеристики как ***дисперсия***, которой называется математическое ожидание квадрата от-

клонения с.в. X от ее математического ожидания:

$$\boxed{\mathbf{D}X = \mathbf{M}(X - \mathbf{M}X)^2.} \tag{8.1}$$

Возникает вопрос: почему математическое ожидание берется не от самого отклонения $|X - \mathbf{M}X|$, а от его квадрата? Только потому, что математические свойства модуля делают его «неудобным» в математических выкладках, а квадрат модуля этих недостатков не имеет. В результате разброс значений с.в. относительно ее математического ожидания имеет размерность квадрата единиц, в которых измеряется сама с.в. Впрочем, данное несоответствие устраняется введением еще одной числовой характеристики разброса с.в. X:

$$\boxed{\sigma_X = \sqrt{\mathbf{D}X},}$$

которая имеет ту же размерность, что и сама с.в. X, и называется ее ***средним квадратическим отклонением (с.к.о.)***.

Из формулы (8.1) следует, что для непрерывной с.в. X, имеющей плотность вероятностей $f(x)$, дисперсия вычисляется по формуле

$$\boxed{\mathbf{D}X = \int\limits_{-\infty}^{\infty} (x - \mathbf{M}X)^2 f(x)dx,}$$

а для дискретной с.в. X, заданной рядом распределения (x_i, p_i), — по формуле

$$\boxed{\mathbf{D}X = \sum_{i=1}^{n} (x_i - \mathbf{M}X)^2 p_i,}$$

где n — число или бесконечность.

Так как дисперсия определяется с помощью математического ожидания, она, как всякое математическое ожидание, может существовать, а может и не существовать. Чтобы лучше понять, что представляет собой дисперсия, рассмотрим ее свойства.

СВОЙСТВА ДИСПЕРСИИ

1°. *Дисперсия неотрицательна:* $\mathbf{D}X \geq 0$.

Это следует прямо из ее определения.

2°. *Дисперсия детерминированной величины равна нулю:* $\mathbf{D}C = 0$.

Так как $\mathbf{M}C = C$, то $\mathbf{D}C = \mathbf{M}(C - \mathbf{M}C)^2 = \mathbf{M}(C - C)^2 = \mathbf{M}0 = 0$.

3°. *Детерминированный множитель выносится за знак дисперсии в квадрате:*

$$\mathbf{D}(CX) = C^2\mathbf{D}X.$$

Действительно,

$$\begin{aligned}\mathbf{D}(CX) &= \mathbf{M}(CX - \mathbf{M}(CX))^2 = \mathbf{M}(CX - C\mathbf{M}X)^2 = \\ &= \mathbf{M}(C^2(X - \mathbf{M}X)^2) = C^2\mathbf{M}(X - \mathbf{M}X)^2 = C^2\mathbf{D}X.\end{aligned}$$

4°. *Дисперсия равна математическому ожиданию квадрата с.в. минус квадрат ее математического ожидания:*

$$\boxed{\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2.}$$

Вот доказательство:

$$\begin{aligned}\mathbf{D}X &= \mathbf{M}(X - \mathbf{M}X)^2 = \mathbf{M}(X^2 - 2X \cdot \mathbf{M}X + (\mathbf{M}X)^2) = \\ &= \mathbf{M}X^2 - 2 \cdot (\mathbf{M}X)^2 + (\mathbf{M}X)^2 = \mathbf{M}X^2 - (\mathbf{M}X)^2.\end{aligned}$$

5°. *Если с.в. X и Y независимы, то дисперсия суммы этих с.в. равна сумме их дисперсий:*

$$\mathbf{D}(X + Y) = \mathbf{D}X + \mathbf{D}Y.$$

На основании предыдущего свойства получаем

$$\begin{aligned}\mathbf{D}(X + Y) &= \mathbf{M}(X + Y)^2 - [\mathbf{M}(X + Y)]^2 = \\ &= \mathbf{M}(X^2 + 2XY + Y^2) - (\mathbf{M}X + \mathbf{M}Y)^2 = \\ &= \mathbf{M}X^2 + 2\mathbf{M}XY + \mathbf{M}Y^2 - (\mathbf{M}X)^2 - 2\mathbf{M}X\mathbf{M}Y - \mathbf{M}Y^2,\end{aligned}$$

то есть

$$\mathbf{D}(X+Y) = \mathbf{D}X + \mathbf{D}Y + 2\mathbf{M}XY - 2\mathbf{M}X\mathbf{M}Y. \qquad (8.2)$$

Но для независимых с.в. $\mathbf{M}XY = \mathbf{M}X\mathbf{M}Y$, поэтому $\mathbf{D}(X+Y) = \mathbf{D}X + \mathbf{D}Y$. □

Попробуйте сами выразить дисперсии

$$\mathbf{D}(-X),\ \mathbf{D}(X-Y),\ \mathbf{D}(X+C)$$

через дисперсии независимых с.в. X и Y, считая, что C — детерминированная величина.

Математическое ожидание является частным случаем (при $k=1$) ***момента k-го порядка*** с.в. X:

$$\alpha_k = \mathbf{M}X^k,$$

а дисперсия — частным случаем (при $k=2$) ее ***центрального момента k-го порядка***:

$$\mu_k = \mathbf{M}(X - \mathbf{M}X)^k.$$

Эти числовые характеристики тоже находят применение при описании некоторых особенностей поведения с.в.

Пример 8.1. *Пусть при передаче сообщения в 20 знаков может быть искажено не более 5 знаков; с.в. X — количество знаков, искаженных при передаче сообщения. Ряд распределения для нее имеет вид:*

x_k	0	1	2	3	4	5
p_k	0,1	0,1	0,2	0,4	0,1	0,1

Найти дисперсию и с.к.о. этой с.в.

Решение. Сначала вычислим ее математическое ожидание:

$$\mathbf{M}X = 0\cdot 0{,}1 + 1\cdot 0{,}1 + 2\cdot 0{,}2 + 3\cdot 0{,}4 + 4\cdot 0{,}1 + 5\cdot 0{,}1 = \\ = 0{,}1 + 0{,}4 + 1{,}2 + 0{,}4 + 0{,}5 = 2{,}6.$$

Используем свойство 4 дисперсии:

$$\begin{aligned}\mathbf{M}X^2 &= 0^2\cdot 0{,}1^2 + 1^2\cdot 0{,}1 + 2^2\cdot 0{,}2 + 3^2\cdot 0{,}4 + 4^2\cdot 0{,}1 + 5^2\cdot 0{,}1 = \\ &= 0{,}1 + 0{,}8 + 3{,}6 + 1{,}6 + 2{,}5 = 8{,}6;\\ \mathbf{D}X &= \mathbf{M}X^2 - (\mathbf{M}X)^2 = 8{,}6 - 2{,}6^2 = 1{,}84;\\ \sigma_X &\approx 1{,}36. \quad \square\end{aligned}$$

Пример 8.2. *Найти дисперсию с.в. X, заданной функцией распределения*

$$F(x) = \begin{cases} 0, & x < 1;\\ \frac{x^3-1}{7}, & 1 \le x \le 2;\\ 1 & x > 2.\end{cases}$$

Решение. Найдем плотность вероятностей заданной с.в.

$$f(x) = F'(x) = \begin{cases} 0, & x < 1 \vee x > 2;\\ \frac{3}{7}x^2, & 1 < x < 2,\end{cases}$$

и ее математическое ожидание

$$\mathbf{M}X = \int\limits_{-\infty}^{\infty} x f(x)\,dx = \int\limits_{1}^{2} \frac{3}{7}x^3\,dx = \left.\frac{3x^4}{28}\right|_1^2 = \frac{48}{28} - \frac{3}{28} = \frac{45}{28} \approx 1{,}607.$$

Далее вычислим математическое ожидание квадрата с.в.

$$\mathbf{M}X^2 = \int\limits_{-\infty}^{\infty} x^2 f(x)\,dx = \int\limits_{1}^{2} \frac{3}{7}x^4\,dx = \left.\frac{3x^5}{35}\right|_1^2 = \frac{96}{35} - \frac{3}{35} = \frac{93}{35},$$

а затем дисперсию

$$\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2 = \frac{93}{35} - \left(\frac{45}{28}\right)^2 = \frac{291}{3920} \approx 0{,}074.$$

$\square$

Иногда для дискретных с.в. (особенно имеющих бесчисленное множество значений) удобнее вычислять дисперсию с помощью производящей функции. Пусть с.в. X принимает неотрицательные значения: $0, 1, 2, \ldots$. Тогда ее дисперсию можно найти, используя производящую функцию этой с.в.:

$$G(u) = \sum_{k=0}^{n} p_k u^k.$$

Действительно, если такую функцию дважды продифференцировать, получим

$$G''(u) = \sum_{k=2}^{n} k(k-1) p_k u^{k-2},$$

а, значит,

$$G''(1) = \sum_{k=2}^{n} k(k-1) p_k = \mathbf{M}X(X-1) = \mathbf{M}X^2 - \mathbf{M}X =$$
$$= \mathbf{D}X + (\mathbf{M}X)^2 - \mathbf{M}X.$$

В результате получаем требуемую вычислительную формулу:

$$\boxed{\mathbf{D}X = G''(1) + \mathbf{M}X - (\mathbf{M}X)^2.} \tag{8.3}$$

Пример 8.3. *Найти дисперсию с.в., распределенной по закону Пуассона.*

Математическое ожидание пуассоновской с.в. было найдено в примере 7.3 и оказалось равным λ. Найдем дисперсию, используя производящую функцию и формулу (8.3). Так как в данном случае $G''(u) = \lambda^2 e^{\lambda(u-1)}$, то

$$\mathbf{D}X = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

8.2 Ковариация

Ранее нами рассматривались два типа взаимосвязей, которые могут существовать между с.в.: функциональная связь и связь, выражающаяся условными распределениями, когда распределение одной с.в. зависит от значений, принимаемых другой с.в. Последняя взаимосвязь называется также ***вероятностной***. Одной из ее характеристик является ***ковариация*** двух случайных величин X и Y, которой называется математическое ожидание произведения отклонений этих с.в. от своих математических ожиданий:

$$\boxed{\operatorname{cov}(X,Y) = \mathbf{M}(X - \mathbf{M}X)(Y - \mathbf{M}Y).} \tag{8.4}$$

В соответствии с этим определением данная характеристика для случайного вектора непрерывного типа, имеющего плотность вероятностей $f(x,y)$, вычисляется по формуле

$$\operatorname{cov}(X,Y) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x - \mathbf{M}X)(y - \mathbf{M}Y)f(x,y)dxdy,$$

а для случайного вектора дискретного типа — по формуле

$$\operatorname{cov}(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - \mathbf{M}X)(y_j - \mathbf{M}Y)p_{ij},$$

где n, m — натуральные числа или ∞, (x_i, y_j) — возможные значения случайного вектора (X,Y), $p_{ij} = P(X = x_i, Y = y_j)$ — вероятности этих значений.

Рассмотрим

СВОЙСТВА КОВАРИАЦИИ

1° *Ковариация не изменяется от перестановки с.в.:*

$$\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X).$$

2° *Ковариация с.в. с самой собой есть дисперсия этой с.в.:*

$$\operatorname{cov}(X, X) = \mathbf{D}X.$$

3° *Постоянный множитель можно выносить за знак ковариации:*

$$\operatorname{cov}(CX, Y) = \operatorname{cov}(X, CY) = C \operatorname{cov}(X, Y).$$

Эти свойства непосредственно следуют из определения ковариации.

4° *Ковариация с.в. равна математическому ожиданию произведения этих с.в. минус произведение их математических ожиданий:*

$$\operatorname{cov}(X, Y) = \mathbf{M}XY - \mathbf{M}X \cdot \mathbf{M}Y.$$

Действительно, из формулы (8.4) получаем

$$\begin{aligned}\operatorname{cov}(X, Y) &= \mathbf{M}(XY - X\mathbf{M}Y - Y\mathbf{M}X + \mathbf{M}X \cdot \mathbf{M}Y) = \\ &= \mathbf{M}XY - \mathbf{M}X \cdot \mathbf{M}Y - \mathbf{M}Y \cdot \mathbf{M}X + \mathbf{M}X \cdot \mathbf{M}Y = \\ &= \mathbf{M}XY - \mathbf{M}X \cdot \mathbf{M}Y.\end{aligned}$$

5° *Дисперсия суммы двух с.в. вычисляется по формуле*

$$\boxed{\mathbf{D}(X + Y) = \mathbf{D}X + \mathbf{D}Y + 2 \operatorname{cov}(X, Y).}$$

Это следует из равенства (8.2) и предыдущего свойства. Докажите самостоятельно, что

$$\mathbf{D}(X - Y) = \mathbf{D}X + \mathbf{D}Y - 2 \operatorname{cov}(X, Y).$$

6° *Если X и Y независимы, то их ковариация равна нулю.*

Действительно, в данном случае $\mathbf{M}XY = \mathbf{M}X \cdot \mathbf{M}Y$ и из свойства 4 сразу получаем требуемое. □

Из свойства 6 следует, что с.в. X и Y зависимы, если их ковариация не равна нулю.

Ковариационной матрицей $\operatorname{cov}(\mathbf{X})$ случайного вектора $\mathbf{X} = (X_1, \ldots, X_n)$ называется матрица

$$\operatorname{cov}(\mathbf{X}) = \begin{pmatrix} \mathbf{D}X_1 & \operatorname{cov}(X_1, X_2) & \ldots & \operatorname{cov}(X_1, X_n) \\ \operatorname{cov}(X_2, X_1) & \mathbf{D}X_2 & \ldots & \operatorname{cov}(X_2, X_n) \\ \ldots & \ldots & \ldots & \ldots \\ \operatorname{cov}(X_n, X_1) & \operatorname{cov}(X_n, X_2) & \ldots & \mathbf{D}X_n \end{pmatrix}. \quad (8.5)$$

8.3 Коэффициент корреляции

Более тонкой, чем ковариация, характеристикой взаимосвязи с.в. X и Y является ***коэффициент корреляции***, определяемый как отношение ковариации этих с.в. к произведению их с.к.о.:

$$\boxed{r = \frac{\operatorname{cov}(X, Y)}{\sigma_X \sigma_Y}.}$$

Если $r \neq 0$, то говорят, что с.в. X и Y ***коррелированы***, в противном случае их называют ***некоррелированными***. Свойство 6 ковариации позволяет утверждать, что *независимые с.в. некоррелированы*. Обратное в общем случае неверно.

Действительно, пусть, например, случайный вектор дискретного типа имеет распределение, представленное следующей таблицей

$x_i \backslash y_j$	-1	0	1
-1	0	$1/4$	0
0	$1/4$	0	$1/4$
1	0	$1/4$	0

Тогда ряды распределений с.в. X и Y одинаковы:

x_i	y_i	-1	0	1
$p_{i\bullet}$	$p_{\bullet j}$	$1/4$	$1/2$	$1/4$

Следовательно, $\mathbf{M}X = \mathbf{M}Y = (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0$, $\operatorname{cov}(X, Y) = 0$ и, значит, X и Y некоррелированы. Но они зависимы, так как

$$p_{11} = 0 \neq p_{1\bullet} \cdot p_{\bullet 1} = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}.$$

Коэффициент корреляции по модулю не превосходит единицы: $|r| \leq 1$. Чтобы доказать это, возьмем с.в. X и Y с единичной дисперсией: $\mathbf{D}X = \mathbf{D}Y = 1$. Тогда

$$\mathbf{D}(X \pm Y) = \mathbf{D}X + \mathbf{D}Y \pm 2\operatorname{cov}(X, Y) = 2 \pm 2r = 2(1 \pm r) \geq 0,$$

так как дисперсия $\mathbf{D}(X \pm Y)$ неотрицательна. Поэтому $1 \pm r \geq 0$, а $|r| \leq 1$.

Коэффициент корреляции является индикатором линейной связи между с.в. X и Y: если $Y = aX + b$, то $|r| = 1$.

Действительно, в этом случае

$$\begin{aligned}
\operatorname{cov}(X, Y) &= \operatorname{cov}(X, aX + b) = \\
&= \mathbf{M}(X - \mathbf{M}X)(aX + b - \mathbf{M}(aX + b)) = \\
&= \mathbf{M}(X - \mathbf{M}X)(aX + b - a\mathbf{M}X - b) = \\
&= a\mathbf{M}(X - \mathbf{M}X)\mathbf{M}(X - \mathbf{M}X) = \\
&= a\sigma_X^2, \\
\sigma_Y^2 &= \mathbf{D}(aX + b) = a^2\sigma_X^2,\ \sigma_Y = |a|\sigma_X, \\
r &= \frac{a\sigma_X^2}{\sigma_X|a|\sigma_X} = \frac{a}{|a|} = \pm 1.
\end{aligned}$$

Вообще говоря, *коэффициент корреляции характеризует и степень линейной связи* между X и Y , так как, чем больше модуль r, тем больше зависимость между X и Y похожа на линейную, но этот вопрос мы оставим без подробного рассмотрения.

8.4 Линейная регрессия

Рассмотрим частный случай вероятностной зависимости, которая называется ***корреляционной*** и состоит в том, что изменение одной с.в. вызывает изменение среднего значения другой с.в.

Возьмем тот случай, когда с.в. X и Y приближенно связаны между собой линейной зависимостью вида $Y \simeq aX + b$, и поставим задачу подобрать такие значения величин a и b, чтобы эта линейная связь в

каком-то смысле наилучшим образом отражала действительную вероятностную связь между данными с.в. Одним из способов решения такой задачи является ***метод наименьших квадратов***, который заключается в поиске минимума функции

$$G(a,b) = \mathbf{M}(Y - aX - b)^2 \tag{8.6}$$

по аргументам a и b. Функция $G(a,b)$ представляет собой математическое ожидание квадрата отклонения с.в. Y от с.в. $aX+b$. Понятно, что малое значение этой функции действительно свидетельствует о приближенно линейной зависимости между с.в. X и Y.

Прежде чем минимизировать функцию (8.6), преобразуем ее, обозначая $m_X = \mathbf{M}X$, $m_Y = \mathbf{M}Y$:

$$\begin{aligned} G(a,b) &= \mathbf{M}[Y - m_Y - a(X - m_X) + m_Y - am_X - b]^2 = \\ &= \mathbf{M}(Y - m_Y)^2 + a^2\mathbf{M}(X - m_X)^2 + (m_Y - am_X - b)^2 - \\ &- 2a\mathbf{M}(Y - m_Y)(X - m_X) + 2\underbrace{\mathbf{M}(Y - m_Y)}_{0}(m_Y - am_X - b) - \\ &- 2a\underbrace{\mathbf{M}(X - m_X)}_{0}(m_Y - am_X - b). \end{aligned}$$

В первых двух слагаемых возникают дисперсии с.в. X и Y, последние два равны нулю, а $\mathbf{M}(Y - m_Y)(X - m_X) = \operatorname{cov}(X,Y) = \sigma_X\sigma_Y r$, где r — коэффициент корреляции с.в. X и Y. Поэтому получаем

$$G(a,b) = \mathbf{D}Y + a^2\mathbf{D}X + (m_Y - am_X - b)^2 - 2a\sigma_X\sigma_Y r.$$

Чтобы минимизировать полученную функцию, применим теорию экстремума функции многих переменных. Сначала используем необходимые условия экстремума, для чего найдем первые частные производные функции $G(a,b)$ и приравняем их нулю:

$$\begin{aligned} \frac{\partial G}{\partial a} &= 2a\sigma_X^2 - 2m_X(m_Y - am_X - b) - 2\sigma_X\sigma_Y r = 0, \\ \frac{\partial G}{\partial b} &= -2(m_Y - am_X - b) = 0. \end{aligned}$$

Второе слагаемое в первом уравнении равно нулю в силу второго уравнения, поэтому из первого уравнения легко находится a, а затем из второго — b:

$$a = \frac{\sigma_Y}{\sigma_X} r, \quad b = m_Y - a m_X.$$

Остается показать, что такие значения аргументов функции действительно обеспечивают ее минимум. С этой целью используем достаточные условия экстремума. Найдем вторые производные функции G:

$$\frac{\partial^2 G}{\partial a^2} = 2\sigma_X^2 + 2m_X^2 > 0, \ \frac{\partial^2 G}{\partial a \partial b} = 2m_X, \ \frac{\partial^2 G}{\partial b^2} = 2,$$

и составим из них определитель:

$$H = \begin{vmatrix} 2\sigma_X^2 + 2m_X^2 & 2m_X \\ 2m_X & 2 \end{vmatrix} = 4\sigma_X^2 > 0.$$

Так как левый верхний элемент определителя и он сам положительны, то, значит, найденные значения параметров на самом деле обеспечивают минимум функции G.

Подставляя найденные значения a и b в искомую линейную зависимость, получаем ***прямую линейной среднеквадратической регрессии*** Y на X:

$$\boxed{Y - m_Y \simeq \frac{\sigma_Y}{\sigma_X} r (X - m_X),} \tag{8.7}$$

причем, величина $a = \dfrac{\sigma_Y}{\sigma_X} r$ называется ***коэффициентом регрессии*** Y на X.

Очевидно, совершенно аналогично можно определить прямую линейной среднеквадратической регрессии X на Y:

$$\boxed{X - m_X \simeq \frac{\sigma_X}{\sigma_Y} r (Y - m_Y).} \tag{8.8}$$

Подставим также найденные значения a и b в функцию G:

$$\begin{aligned} G &= \mathbf{M}\left(Y - \frac{\sigma_Y}{\sigma_X} rX - m_Y + \frac{\sigma_Y}{\sigma_X} r m_X\right)^2 = \\ &= \mathbf{M}\left[Y - m_Y - \frac{\sigma_Y}{\sigma_X} r(X - m_X)\right]^2 = \\ &= \mathbf{D}Y + \frac{\sigma_Y^2}{\sigma_X^2} r^2 \mathbf{D}X - \frac{2r^2\sigma_Y}{\sigma_X}\sigma_X\sigma_Y = \\ &= \mathbf{D}Y - r^2\sigma_Y^2 = \sigma_Y^2(1 - r^2). \end{aligned}$$

Последнее выражение называется ***остаточной дисперсией*** с.в. Y относительно с.в. X. Остаточная дисперсия показывает величину среднеквадратического отклонения с.в. Y от линейной функции $aX + b$. Мы видим, что при $r = \pm 1$ остаточная дисперсия равна нулю, то есть в этом случае с.в. X и Y связаны точной линейной зависимостью:

$$Y - m_Y = \frac{\sigma_Y}{\sigma_X}(X - m_X),$$

а прямые регрессии (8.7) и (8.8) совпадают. Попутно отметим, что при $r \neq \pm 1$, выразив, например, из (8.7) X через Y, мы не получим уравнение (8.8). Этим линии регрессии отличаются от обычных линейных зависимостей.

8.5 Статистическая линеаризация нелинейностей

Очень часто для приближенного описания зависимости вида

$$Y = \varphi(X) \tag{8.9}$$

достаточно связать математическое ожидание и дисперсию (или ковариацию) аргумента X с такими же характеристиками зависимой переменной Y. Например, можно заменить уравнение (8.9) приближенным линейным равенством

$$Y \approx U = k(X - \mathbf{M}X) + b,$$

где k и b — постоянные, которые надо определить. В методе ***статистической линеаризации нелинейностей*** для этого используют два критерия, получая два различных решения задачи.

I критерий. Постоянные k и b выбирают из условия равенства математических ожиданий и дисперсии точного и приближенного выражения функции:

$$\mathbf{M}Y = \mathbf{M}U, \ \mathbf{D}Y = \mathbf{D}U.$$

Исходя из этого критерия, сначала используем математическое ожидание:

$$\mathbf{M}Y = \mathbf{M}U = \mathbf{M}\left[k\left(X - \mathbf{M}X\right) + b\right] = b,$$

следовательно,

$$b = \mathbf{M}Y.$$

Применяя полученный результат, в качестве промежуточной величины найдем

$$\begin{aligned}\mathbf{M}U^2 &= \mathbf{M}\left[k\left(X - \mathbf{M}X\right) + b\right]^2 = \\ &= \mathbf{M}\left[k^2\left(X - \mathbf{M}X\right)^2 + 2kb\left(X - \mathbf{M}X\right) + b^2\right] = \\ &= k^2\mathbf{D}X + b^2 = k^2\mathbf{D}X + \left(\mathbf{M}Y\right)^2.\end{aligned}$$

Теперь воспользуемся условием равенства дисперсий:

$$\mathbf{D}Y = \mathbf{D}U = \mathbf{M}U^2 - \left(\mathbf{M}U\right)^2 = k^2\mathbf{D}X + \left(\mathbf{M}Y\right)^2 - \left(\mathbf{M}Y\right)^2 = k^2\mathbf{D}X.$$

Таким образом,

$$k = \pm\sqrt{\frac{\mathbf{D}Y}{\mathbf{D}X}}.$$

Знак в формуле следует выбирать так, чтобы знаки точной и аппроксимирующей функций совпадали.

II критерий. Постоянные k и b выбирают из условия минимума функционала

$$J = \mathbf{M}\left(Y - U\right)^2,$$

то есть из условия минимума средней величины квадрата отклонения точной и аппроксимирующей функции друг от друга.

Преобразуем этот функционал:

$$\begin{aligned} J &= \mathbf{M}\left[Y - k\left(X - \mathbf{M}X\right) - b\right]^2 = \\ &= \mathbf{M}\left[\left(Y - \mathbf{M}Y\right) + \mathbf{M}Y - k\left(X - \mathbf{M}X\right) - b\right]^2 = \\ &= \mathbf{D}Y + \left(\mathbf{M}Y\right)^2 + k^2\mathbf{D}X + b^2 - 2k\operatorname{cov}\left(X, Y\right) - 2b\mathbf{M}Y. \end{aligned}$$

Применим необходимые условия экстремума:

$$\frac{\partial J}{\partial k} = 2k\mathbf{D}X - 2\operatorname{cov}\left(X, Y\right) = 0, \ \frac{\partial J}{\partial b} = 2b - 2\mathbf{M}Y = 0,$$

из которых следует, что возможной точкой экстремума является пара значений

$$b = \mathbf{M}Y, \ k = \frac{\operatorname{cov}\left(X, Y\right)}{\mathbf{D}X}.$$

Определитель, составленный из вторых производных функционала, показывает, что найденная точка является точкой минимума:

$$H = \begin{vmatrix} \frac{\partial^2 J}{\partial k^2} & \frac{\partial^2 J}{\partial k \partial b} \\ \frac{\partial^2 J}{\partial b \partial k} & \frac{\partial^2 J}{\partial b^2} \end{vmatrix} = \begin{vmatrix} 2\mathbf{D}X & 0 \\ 0 & 2 \end{vmatrix} = 4\mathbf{D}X > 0.$$

Пример 8.4. *Статистически линеаризовать нелинейность вида*

$$Y = \varphi\left(x\right) = \max\left(X, 0\right),$$

если X — с.в., равномерно распределенная на отрезке $[-a, a]$.

Решение. Условие задачи позволяет сделать вывод о том, что плотность вероятностей с.в. X имеет вид

$$f_X\left(x\right) = \begin{cases} \frac{1}{2a}, & x \in [-a, a]; \\ 0, & x \notin [-a, a]. \end{cases}$$

Кроме того, числовые характеристики этой с.в. равны $\mathbf{M}X = 0$, $\mathbf{D}X = \left(2a\right)^2/12 = a^2/3$. По обоим рассмотренным критериям

$$b = \mathbf{M}Y = \int\limits_{-a}^{a} \frac{1}{2a}\max\left(x, 0\right)\, dx = \frac{1}{2a}\int\limits_{0}^{a} x\, dx = \frac{1}{2a}\cdot\frac{x^2}{2}\bigg|_0^a = \frac{a^2}{4a} = \frac{a}{4}.$$

Для применения первого критерия надо еще найти дисперсию с.в. Y:

$$\mathbf{M}Y^2 = \int\limits_{-a}^{a} \frac{1}{2a} \left[\max(x,0)\right]^2 \, dx = \frac{1}{2a} \int\limits_{0}^{a} x^2 \, dx = \frac{1}{2a} \cdot \frac{x^3}{3}\bigg|_0^a = \frac{a^3}{6a} = \frac{a^2}{6};$$

$$\mathbf{D}Y = \mathbf{M}Y^2 - (\mathbf{M}Y)^2 = \frac{a^2}{6} - \frac{a^2}{16} = \frac{5a^2}{48}.$$

Следовательно,

$$k = \pm\sqrt{\frac{\mathbf{D}Y}{\mathbf{D}X}} = \pm\sqrt{\frac{5a^2/48}{a^2/3}} = \frac{\sqrt{5}}{4}.$$

Знак плюс был выбран потому, что точная функция неотрицательна. Таким образом,

$$\max(X,0) \approx \frac{a}{4} + \frac{\sqrt{5}}{4} X.$$

Для применения второго критерия требуется найти ковариацию с.в. X и Y:

$$\begin{aligned}\operatorname{cov}(X,Y) &= \operatorname{cov}(X,\varphi(X)) = \mathbf{M}X\varphi(X) - \mathbf{M}X\mathbf{M}\varphi(X) = \\ &= \mathbf{M}X\varphi(X) = \int\limits_{-\infty}^{\infty} x\varphi(x)\,\frac{1}{2a}\,dx = \frac{1}{2a}\int\limits_{-a}^{a} x\max(x,0)\,dx = \\ &= \frac{1}{2a}\int\limits_{0}^{a} x^2\,dx = \frac{1}{2a}\cdot\frac{x^3}{3}\bigg|_0^a = \frac{a^2}{6}.\end{aligned}$$

Поэтому

$$k = \frac{\operatorname{cov}(X,Y)}{\mathbf{D}X} = \frac{5a^2/48}{a^2/3} = \frac{5}{16}$$

и в данном случае

$$\max(X,0) \approx \frac{a}{4} + \frac{5}{16} X.$$

□

Лекция 9
Типичные дискретные распределения

Некоторые законы распределения вероятностей чаще, чем другие, используются как в технических приложениях, так и в самой теории вероятностей или такой родственной ей науке, как математическая статистика. Поэтому они заслуживают специального изучения.

9.1 Схема Бернулли и биномиальное распределение

Схемой Бернулли называется совокупность n независимых опытов, в каждом из которых с одной и той же вероятностью p возможно осуществление с.с. A. В эту схему укладывается последовательность бросаний монеты ($A =$ «Выпадение герба», $p = 1/2$), кости ($A =$ $=$ «Выпадение шестерки», $p = 1/6$), испытание нескольких идентичных электромоторов ($A =$ «Электромотор проработал заданное количество часов», p определяется степенью надежности электромотора) и т.д.

Введем с.в. X — количество опытов, в которых произойдет с.с. A. Эту с.в. еще называют ***числом успехов*** в схеме Бернулли. Найдем вероятность $P(X = k)$ того, что число успехов будет равно заданному числу k ($0 \leq k \leq n$). Для этого решим сначала более простую задачу: определим вероятность того, что в первых k опытах осуществится с.с. A, а в остальных $n - k$ опытах оно не реализуется. Очевидно, искомая вероятность по теореме умножения для незави-

симых с.с. равна $p^k q^{n-k}$, где $q = 1 - p$. Точно такой же будет вероятность того, что в k опытах с заданными номерами с.с. A произойдет, а в других опытах не произойдет. Следовательно, вероятность числа успехов в схеме Бернулли равна вероятности суммы всех таких (несовместных) с.с., причем количество этих событий равно числу различных вариантов выбора k фиксированных номеров опытов из n номеров, то есть равна $C_n^k p^k q^{n-k}$. Таким образом,

$$\boxed{p_k = P(X = k) = C_n^k p^k q^{n-k}, \quad k = 0, \ldots, n,}$$

причем, $\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1$. Ряд распределения с.в. X имеет вид

x_k	0	1	$\ldots$	k	$\ldots$	n
p_k	q^n	npq^{n-1}	$\ldots$	$C_n^k p^k q^{n-k}$	$\ldots$	p^n

Такое распределение называется ***биномиальным***, а с.в. X — ***биномиальной*** с.в.

Найдем ее числовые характеристики. Для этого введем ***бернуллиевские*** с.в. X_i, которые принимают всего два значения: 1, если в i-м опыте с.с. A произошло; и 0 в противном случае. Ряд распределения для такой с.в. выглядит очень просто:

x_i	0	1
p_i	q	p

Применяя формулы расчета числовых характеристик для дискретных с.в., последовательно получаем

$$\mathbf{M}X_i = 0 \cdot q + 1 \cdot p = p, \quad \mathbf{M}X_i^2 = 0^2 \cdot q + 1^2 \cdot p = p,$$
$$\mathbf{D}X_i = \mathbf{M}X_i^2 - (\mathbf{M}X_i)^2 = p - p^2 = p(1-p) = pq.$$

Так как число успехов в схеме Бернулли представляет собой сумму всех бернуллиевских с.в., то его математическое ожидание и дисперсия (с учетом независимости X_i) вычисляются по формулам:

$$\mathbf{M}X = \sum_{i=1}^{n} \mathbf{M}X_i = np, \quad \mathbf{D}X = \sum_{i=1}^{n} \mathbf{D}X_i = npq.$$

Пример 9.1. *Пять стрелков стреляют в мишень. Вероятность попадания для каждого из них равна* 0,8. *Какова вероятность, что три из них попадут в мишень? Какое следует ожидать среднее число попаданий?*

Решение. Имеем схему Бернулли с числом опытов $n = 5$ и с.с. $A =$ =«Стрелок попадет в мишень», которое может произойти в каждом опыте с одной и той же вероятностью $P(A) = 0{,}8$.

Следовательно, для биномиальной с.в. X — количества попаданий в мишень получаем, что вероятность трех попаданий равна

$$P(X = 3) = C_5^3 0{,}8^3 0{,}2^2 = 10 \cdot 0{,}512 \cdot 0{,}04 = 0{,}2048.$$

Среднее число попаданий выражается математическим ожиданием:

$$\mathbf{M}X = np = 5 \cdot 0{,}8 = 4.$$

□

Пример 9.2. *Для условий предыдущего примера найти вероятность того, что не менее трех стрелков попадут в мишень.*

Решение. Вероятность того, что три, четыре или пять стрелков попадут в мишень, учитывая независимость перечисленных с.с., равна

$$\begin{aligned} P(X = 3) + P(X = 4) + P(X = 5) = \\ &= 0{,}2048 + C_5^4 0{,}8^4 0{,}2 + C_5^5 0{,}8^5 0{,}2^0 \\ &= 0{,}2048 + 0{,}4096 + 0{,}32768 = 0{,}94208. \end{aligned}$$

□

Представляет интерес ***наивероятнейшее число*** k_0 ***успехов*** в схеме Бернулли. Чтобы его найти, вычислим отношение

$$\begin{aligned} \frac{P(X = k+1)}{P(X = k)} &= \frac{C_n^{k+1} p^{k+1} q^{n-k-1}}{C_n^k p^k q^{n-k}} = \\ &= \frac{n!k!(n-k!)}{(k+1)!(n-k-1)!n!} \cdot \frac{p}{q} = \frac{(n-k)p}{(k+1)q}. \end{aligned} \tag{9.1}$$

Из полученного выражения следует, что $P(X = k+1) = P(X = k)$ равносильно $(n - k)p = (k + 1)q$, или $np - q = k$, т.е. равенство вероятностей для «соседних» значений с.в. эквивалентно тому, что величина $np - q$ является целым числом, равным k.

Наивероятнейшему числу k_0 соответствует наибольшее значение вероятности числа успехов в схеме Бернулли, т.е. должна выполняться система неравенств:

$$P(X = k_0) \geq P(X = k_0 - 1), \quad P(X = k_0) \geq P(X = k_0 + 1),$$

которую можно переписать так:

$$\frac{P(X = k_0 - 1)}{P(X = k_0)} \leq 1, \quad \frac{P(X = k_0 + 1)}{P(X = k_0)} \leq 1.$$

К обоим неравенствам применим формулу (9.1), причем, ко второму непосредственно, а к первому — уменьшив k на единицу и «перевернув» отношение вероятностей; последовательно получим следующие системы неравенств:

$$\begin{cases} \frac{k_0}{n-k_0+1} \cdot \frac{q}{p} \leq 1, \\ \frac{n-k_0}{k_0+1} \cdot \frac{p}{q} \leq 1; \end{cases} \quad \begin{cases} k_0 q \leq np - pk_0 + p, \\ np - k_0 p \leq k_0 q + q; \end{cases} \quad \begin{cases} k_0 \leq np + p, \\ k_0 \geq np - q. \end{cases}$$

Последнюю систему можно заменить двойным неравенством:

$$np - q \leq k_0 \leq np + p.$$

Разность между правой и левой частью составляет 1. Поэтому, если $np - q$ — дробное число, то $k_0 = \lfloor np + p \rfloor$, где $\lfloor a \rfloor$ есть наибольшее целое число, не превосходящее a. Если же $np - q$ — целое, то, как было показано ранее, вероятности двух соседних значений одинаковы и, значит, существуют два наивероятнейших значения с.в.: $k_0' = np - q$, $k_0'' = np + p$. Остается показать, что равенство друг другу трех вероятностей, соответствующих «соседним» значениям с.в., невозможно. Действительно, в этом случае

$$P(X = k - 1) = P(X = k) = P(X = k + 1)$$

и, рассуждая так же, как в случае неравенств, получим $np-q=k=$ $=np+p$. Но тогда $np+p-np+q=0$, $p+q=0$. Это возможно только при $p=0$ или $p=1$. Отбросив эти тривиальные случаи, окончательный результат сформулируем так:

$k_0=\lfloor np+p\rfloor$ — единственное наивероятнейшее значение с.в., если $np-q$ дробное,

$k_0'=np-q$, $k_0''=np+p$ — два наивероятнейших значения с.в., если $np-q$ целое.

Пример 9.3. *Испытывается 10 приборов. Вероятность того, что прибор выдержит испытание, равна* 0,7. *Найти наивероятнейшее число приборов, которые пройдут испытание.*

Решение. В данном случае $n=10$, $p=0{,}7$ и $np-q=10\cdot 0{,}7-0{,}3=6{,}7$ дробное. Поэтому,

$$k_0=\lfloor 10\cdot 0{,}7+0{,}7\rfloor=\lfloor 7{,}7\rfloor=7.$$

□

Пример 9.4. *Как изменится решение предыдущей задачи, если число приборов уменьшить на единицу?*

Решение. Теперь $n=9$ и $np-q=9\cdot 0{,}7-0{,}3=6$ целое. Следовательно, имеется два наивероятнейших значения

$$k_1=np-q=6 \text{ и } k_2=np+p=7.$$

□

9.2 Полиномиальное распределение

Пусть в схеме Бернулии в одном и том же опыте могут произойти несовместные с.с. $A_1,\ldots,A_m$ с вероятностями $p_1,\ldots,p_m$, соответственно, причем, $p_i>0$, $\sum_{i=1}^m p_i=1$. Обозначим X_i число опытов, в которых произойдет с.с. A_i («число успехов» для с.с. A_i). Введем

вектор $\mathbf{X} = (X_1, \ldots, X_m)$ и найдем вероятность осуществления вектора «числа успехов» $\mathbf{K} = (k_1, \ldots, k_m)$ в данной серии опытов:

$$P(\mathbf{X} = \mathbf{K}) = P(X = k_1, \ldots, X_m = k_m), \tag{9.2}$$

$k_1 + \cdots + k_m = n$. Так как осуществление с.с. $X_i = k_i$ влияет на вероятности осуществления других с.с. $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_m$ (например, при реализации с.с. $X_i = n$ другие с.с. A_j, $j \neq i$, вообще произойти не могут), то с.в. $X_1, \ldots, X_m$ не являются независимыми. Поэтому вероятность в формуле (9.2) вычислим по теореме умножения для зависимых с.с.:

$$\begin{aligned} P(\mathbf{X} = \mathbf{K}) = P(X_1 = k_1) \cdot P(X_2 = k_2/X_1 = k_1) \cdot \ldots \cdot \\ \cdot P(X_{m-1} = k_{m-1}/X_1 = k_1, \ldots, X_{m-2} = k_{m-2}) \cdot \\ \cdot P(X_m = k_m/X_1 = k_1, \ldots, X_{m-1} = k_{m-1}), \end{aligned} \tag{9.3}$$

причем, последний множитель, очевидно, равен 1 в силу условия $k_1 + \cdots + k_m = n$.

Вероятность $P(X_1 = k_1)$ равна произведению вероятности того, что в k_1 фиксированных опытах произойдет с.с. A_1, на число различных способов выбрать k_1 опытов из n:

$$P(X_1 = k_1) = C_n^{k_1} p_1^{k_1}.$$

Аналогично получаются и остальные вероятности:

$$P(X_2 = k_2/X_1 = k_1) = C_{n-k_1}^{k_2} p_2^{k_2},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$P(X_{m-1} = k_{m-1}/X_1 = k_1, \ldots, X_{m-2} = k_{m-2}) = C_{n-k_1-\cdots-k_{m-2}}^{k_{m-1}} p_{m-1}^{k_{m-1}},$$

$$P(X_m = k_m/X_1 = k_1, \ldots, X_{m-1} = k_{m-1}) = C_{n-k_1-\cdots-k_{m-1}}^{k_m} p_m^{k_m}.$$

Подставляя эти выражения в равенство (9.3) и используя свой-

ства биномиальных коэффициентов, получим

$$P(\mathbf{X}=\mathbf{K}) = C_n^{k_1} p_1^{k_1} \cdot C_{n-k_1}^{k_2} p_2^{k_2} \cdot \ldots \cdot C_{n-k_1-\cdots-k_{m-2}}^{k_{m-1}} p_{m-1}^{k_{m-1}} \cdot$$
$$\cdot C_{n-k_1-\cdots-k_{m-1}}^{k_m} p_m^{k_m} = \frac{n!}{k_1!(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \cdots$$
$$\cdot \frac{(n-k_1-\cdots-k_{m-2})!}{k_{m-1}!(n-k_1-\cdots-k_{m-1})!} \cdot \frac{(n-k_1-\cdots-k_{m-1})!}{k_m!(n-k_1-\cdots-k_m)!} \cdot$$
$$\cdot p_1^{k_1} p_2^{k_2} \cdot \ldots \cdot p_m^{k_m}.$$

Сокращая одинаковые множители в числителях и знаменателях дробей, приходим к равенству:

$$\boxed{P(\mathbf{X}=\mathbf{K}) = \frac{n!}{k_1!k_2!\ldots k_m!} p_1^{k_1} p_2^{k_2} \cdot \ldots \cdot p_m^{k_m},} \tag{9.4}$$

причем, из формулы (B.1) следует, что

$$\sum_{k_1+\cdots+k_m=n} \frac{n!}{k_1!\ldots k_m!} p_1^{k_1} \ldots p_m^{k_m} = (p_1+\cdots+p_m)^n = 1.$$

Распределение случайного вектора (9.4) называется ***полиномиальным***. Название связано с тем, что числа

$$\frac{n!}{k_1!k_2!\ldots k_m!} p_1^{k_1} p_2^{k_2} \cdot \ldots \cdot p_m^{k_m}$$

называются ***полиномиальными коэффициентами***, так как являются коэффициентами многочлена (полинома), получаемого из выражения $(p_1+\cdots+p_m)^n$ раскрытием скобок (см. приложение B).

Так как $X_1, \ldots, X_m$ — биномиальные с.в., то их числовые характеристики равны $\mathbf{M}X_i = np_i$, $\mathbf{D}X_i = np_iq_i$, где $q_i = 1-p_i$. Следовательно, математическое ожидание случайного полиномиального вектора $\mathbf{X}$ определяется формулой

$$\mathbf{MX} = (np_1, \ldots, np_m). \tag{9.5}$$

Найдем ковариацию с.в. X_i и X_j при $i \neq j$. Как было показано при рассмотрении схемы Бернулли, имеет место представление

$$\mathbf{X}_i = \sum_{s=1}^{n} X_{is},$$

где X_{is} — бернуллиевская с.в., которая принимает значение 1 с вероятностью p_i, если в s-м опыте произошло с.с. A_i, и значение 0 в противном случае.

Воспользуемся формулой

$$\operatorname{cov}(X_i, X_j) = \mathbf{M}X_iX_j - \mathbf{M}X_i \cdot \mathbf{M}X_j, \tag{9.6}$$

для чего вычислим

$$\mathbf{M}X_iX_j = \mathbf{M}\sum_{s=1}^{n} X_{is} \sum_{t=1}^{n} X_{jt} = \sum_{s=1}^{n}\sum_{t=1}^{n} \mathbf{M}X_{is}X_{jt}.$$

Если $s = t$, то с.в. $X_{is}X_{jt}$ принимает значение 0, так как в одном и том же опыте с.с. A_i и A_j произойти не могут. Если же $s \neq t$, то она принимает значение 1 с вероятностью p_ip_j в том случае, если в s-м опыте произойдет событие A_i, а в t-м опыте произойдет событие A_j, и принимает значение 0 в остальных случаях. Таким образом,

$$\mathbf{M}X_iX_j = \begin{cases} 0, & s = t; \\ 1 \cdot p_ip_j + 0, & s \neq t. \end{cases}$$

Следовательно,

$$\mathbf{M}X_iX_j = \sum_{\substack{s,t=1 \\ s\neq t}}^{n} p_ip_j = n\,(n-1)\,p_ip_j$$

и тогда из равенства (9.6) получим

$$\operatorname{cov}(X_i, X_j) = n\,(n-1)\,p_ip_j - np_inp_j = n\,(n-1-n)\,p_ip_j = -np_ip_j.$$

Согласно (8.5), ковариационная матрица имеет вид

$$\operatorname{cov}(\mathbf{X}) = n \begin{pmatrix} p_1q_1 & -p_1p_2 & \dots & -p_1p_m \\ -p_2p_1 & p_2q_2 & \dots & -p_2p_m \\ \dots & \dots & \dots & \dots \\ -p_mp_1 & -p_mp_2 & \dots & p_mq_m \end{pmatrix}. \tag{9.7}$$

Пример 9.5. *По статистическим данным областного центра из каждых 100 новых семей (оба супруга первый раз вступают в брак) 10% разводятся и оба супруга в новый брак не вступают, 50% разводятся и, по крайней мере, один из супругов вступает в новый брак, 40% доживают вместе до старости. Вчера в городе образовалось 7 семей. Какова вероятность, что три из них разведутся и супруги не вступят в новый брак, а две сохранят свой брак до старости? Каковы числовые характеристики будущей новой семьи?*

Решение. Введем с.с.: A_1 = «Семья разведется и супруги в новый брак не вступят»; A_2 = «Семья разведется и, по крайней мере, один из супругов вступит в новый брак»; A_3 = «Семья не разведется».

Имеем схему Бернулли с 7 опытами и тремя с.с. A_1, A_2, A_3 и соответствующими им вероятностями $p_1 = 0{,}1$; $q_1 = 0{,}9$; $p_2 = q_2 = 0{,}5$; $p_3 = 0{,}4$; $q_3 = 0{,}6$. Пусть X_i — «число успехов» в этой схеме для с.с. A_i и пусть вектор $\mathbf{X} = (X_1, X_2, X_3)$. Тогда по формуле (9.4) имеем

$$P(\mathbf{X} = (3,3,2)) = \frac{7!}{3!2!2!}\, 0{,}1^3 0{,}5^2 0{,}4^2 = 0{,}0084.$$

Для определения вектора математических ожиданий используем формулу (9.5):

$$\mathbf{MX} = (np_1, np_2, np_3) = (0{,}7; 3{,}5; 2{,}8).$$

Этот вектор показывает среднее число семей, которые после заключения брака реализуют то или иное продолжение жизненного пути.

Матрицу ковариаций вычислим на основании (9.7):

$$\operatorname{cov}(\mathbf{X}) = 7 \begin{pmatrix} 0{,}09 & -0{,}05 & -0{,}04 \\ -0{,}05 & 0{,}25 & -0{,}2 \\ -0{,}04 & -0{,}2 & 0{,}24 \end{pmatrix} = \begin{pmatrix} 0{,}63 & -0{,}35 & -0{,}28 \\ -0{,}35 & 1{,}75 & -1{,}4 \\ -0{,}28 & -1{,}4 & 1{,}68 \end{pmatrix}.$$

9.3 Пуассоновское распределение

Предположим, что число опытов в схеме Бернулли неограниченно увеличивается: $n \to \infty$, а вероятность реализации с.с. A в одном опыте неограниченно уменьшается: $p \to 0$, причем, произведение этих величин стремится к некоторой ненулевой конечной величине $\lambda : np \to \lambda, \lambda \neq 0, \lambda < \infty$. Покажем, что в этом предельном случае закон распределения числа успехов в схеме Бернулли перестает быть биномиальным и трансформируется в новый закон распределения.

Обозначим $\lambda_n = np$ и преобразуем выражение для вероятности числа успехов в схеме Бернулли:

$$P(X = k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k} =$$
$$= \frac{(n-k+1) \cdot \ldots \cdot (n-1)n}{k!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k}.$$

Поменяем местами числители первых двух дробей, а после этого каждый сомножитель числителя первой дроби разделим на n; последний множитель всего выражения представим в виде произведения двух степеней с одинаковыми основаниями:

$$P(X = k) = \frac{\lambda_n^k}{k!} \cdot \left(1 - \frac{k-1}{n}\right) \cdot \ldots \cdot \left(1 - \frac{1}{n}\right) \cdot 1 \cdot \left(1 - \frac{\lambda_n}{n}\right)^n \cdot$$
$$\cdot \left(1 - \frac{\lambda_n}{n}\right)^{-k}.$$

При $n \to \infty$ величина $\lambda_n^k \to \lambda^k$, выражения в круглых скобках, за исключением предпоследнего, стремятся к 1, а предпоследний множитель в соответствии со вторым замечательным пределом стремится к $e^{-\lambda}$. Поэтому вероятность числа успехов в данном предельном случае выражается формулой

$$\boxed{p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},}$$

причем, $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$.

Следовательно, с.в. X имеет ряд распределения вида

x_k	0	1	...	k	...
p_k	$e^{-\lambda}$	$\lambda e^{-\lambda}$	...	$\frac{\lambda^k}{k!} e^{-\lambda}$	...

Такой закон распределения называется ***пуассоновским***, как и с.в., имеющая это распределение.

Математическое ожидание пуассоновской с.в. было найдено в примере 7.3, а дисперсия — в примере 8.3. Обе величины оказались равными λ:

$$\mathbf{M}X = \mathbf{D}X = \lambda.$$

Пример 9.6. *Вероятность опечаток на одной странице равна* 0,01. *Найти вероятность того, что в книге с 200 страницами имеется 5 страниц с опечатками. Сколько в среднем имеется в книге страниц с опечатками?*

Решение. Имеем схему Бернулли с $n = 200$, $p = P(A) = 0{,}01$, где $A =$ «На странице имеются опечатки». Так как n велико, а p мало, то для определения вероятности числа успехов можно приближенно считать распределение с.в. X — количества страниц с опечатками — пуассоновским с параметром $\lambda \approx 200 \cdot 0{,}01 = 2$.

Тогда

$$P(X = 5) \approx \frac{2^5}{5!} e^{-2} = \frac{32}{120} \cdot 0{,}13534 \approx 0{,}036.$$

Среднее число страниц с опечатками выражается математическим ожиданием с.в. X, которое равно параметру λ, т.е. двум. □

9.4 Многомерное пуассоновское распределение

Возьмем в схеме Бернулли, рассмотренной в п. 9.2, не m с.с. и соответствующих им с.в., а на одно больше: $A_1, \ldots, A_{m+1}$. Пусть при этом число опытов $n \to \infty$, а вероятности $p_i \to 0$, $i = \overline{1, m}$, причем, так, что

$np_i \to \lambda_i \neq 0$, $i = \overline{1, m}$. Обозначим $\lambda_{i,n} = np_i$ и рассмотрим, как изменится при таком предельном переходе вероятность вектора «числа успехов». С учетом равенств $p_1 + \dots + p_{m+1} = 1$, $k_1 + \dots + k_{m+1} = n$ получим

$$P\left(\mathbf{X} = (k_1, \dots, k_{m+1})\right) = \frac{(k_1 + \dots + k_{m+1})!}{k_1! \dots k_{m+1}!} p_1^{k_1} \dots p_m^{k_m} p_{m+1}^{k_{m+1}} =$$

$$= \frac{n!}{k_1! \dots k_{m+1}!} p_1^{k_1} \dots p_m^{k_m} \left(1 - p_1 - \dots - p_m\right)^{k_{m+1}} =$$

$$= \frac{n!}{k_1! \dots k_{m+1}!} \cdot \frac{\lambda_{1,n}^{k_1}}{n^{k_1}} \cdot \dots \cdot \frac{\lambda_{m,n}^{k_m}}{n^{k_m}} \left(1 - \frac{\lambda_{1,n}}{n} - \dots - \frac{\lambda_{m,n}}{n}\right)^{k_{m+1}} =$$

$$= \frac{\lambda_{1,n}^{k_1} \dots \lambda_{m,n}^{k_m}}{k_1! \dots k_{m+1}!} \cdot \frac{n!}{n^{k_1} \dots n^{k_m}} \cdot$$

$$\cdot \left(1 - \frac{\lambda_{1,n} + \dots + \lambda_{m,n}}{n}\right)^{n - k_1 - \dots - k_m} =$$

$$= \frac{\lambda_{1,n}^{k_1} \dots \lambda_{m,n}^{k_m}}{k_1! \dots k_m!} \cdot \frac{n!}{k_{m+1}! n^{k_1 + \dots + k_m}} \cdot$$

$$\cdot \left(1 - \frac{\lambda_{1,n} + \dots + \lambda_{m,n}}{n}\right)^{n} \left(1 - \frac{\lambda_{1,n} + \dots + \lambda_{m,n}}{n}\right)^{-k_1 - \dots - k_m}. \quad (9.8)$$

Представим вторую дробь в виде

$$\frac{n!}{k_{m+1}! n^{k_1 + \dots + k_m}} = \frac{(n - k_{m+1} + 1)(n - k_{m+1} + 2) \dots n}{n^{n - k_{m+1}}} =$$

$$= \left(1 - \frac{k_{m+1} - 1}{n}\right) \left(1 - \frac{k_{m+1} - 2}{n}\right) \dots 1$$

и перейдем к пределу при $n \to \infty$. Очевидно, последнее выражение стремится к 1, а $\lambda_{i,n} \to \lambda_i$, следовательно, равенство (9.8) преобразуется в формулу

$$P\left(\mathbf{X} = (k_1, \dots, k_m)\right) = \frac{\lambda_1^{k_1} \dots \lambda_m^{k_m}}{k_1! \dots k_m!} e^{-(\lambda_1 + \dots + \lambda_m)}. \quad (9.9)$$

Равенство $\mathbf{X} = (k_1, \ldots, k_{m+1})$ заменено на $\mathbf{X} = (k_1, \ldots, k_m)$, так как правая часть полученной формулы не зависит от k_{m+1}.

Ометим, что

$$\sum_{k_1=1}^{\infty} \cdots \sum_{k_m=1}^{\infty} \frac{\lambda_1^{k_1} \ldots \lambda_m^{k_m}}{k_1! \ldots k_m!} e^{-(\lambda_1 + \cdots + \lambda_m)} =$$

$$= e^{-(\lambda_1 + \cdots + \lambda_m)} \sum_{k_1=1}^{\infty} \frac{\lambda_1^{k_1}}{k_1!} \cdots \sum_{k_m=1}^{\infty} \frac{\lambda_1^{k_m}}{k_m!} = e^{-(\lambda_1 + \cdots + \lambda_m)} e^{-\lambda_1} \cdots e^{-\lambda_m} = 1.$$

Распределение случайного вектора (9.9) называется ***многомерным пуассоновским*** распределением. Поскольку равенство (9.9) можно записать в виде

$$P\left(\mathbf{X} = (k_1, \ldots, k_m)\right) = \frac{\lambda_1^{k_1}}{k_1!} e^{-\lambda_1} \ldots \frac{\lambda_m^{k_m}}{k_m!} e^{-\lambda_m},$$

то многомерное пуассоновское распределение представляет собой произведение одномерных пуассоновских распределений независимых с.в. $Y_1, \ldots, Y_m$, имеющих распределения вида $P(Y_i = k_i) = \frac{\lambda_i^{k_i}}{k_i!} e^{-\lambda_i}$.

Поэтому вектор математических ожиданий и ковариационная матрица многомерного пуассоновского вектора имеют вид

$$\mathbf{M}X = (\lambda_1, \ldots, \lambda_m), \quad \operatorname{cov}(\mathbf{X}) = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \lambda_m \end{pmatrix}.$$

Пример 9.7. *Вероятность выиграть в лотерею составляет* 0,001, *получить премию* — 0,01, *попасть в магазине под акцию снижения цен* — 0,02. *Некто каждый день надеется выиграть в лотерею, получить премию и попасть под акцию снижения цен. Какова вероятность того, что ему в течение года 5 раз удастся выиграть в лотерею, трижды получить премию и 2 раза подпасть под акцию снижения цен?*

Решение. Пусть в году 365 дней. Имеем схему Бернулли из $n = 365$ опытов, в каждом из которых могут произойти с.с.: A_1 = «Некто

выиграет в лотерею», A_2 = «Некто получит премию», A_3 = «Некто подпадет под акцию снижения цен» с указанными в условии задачи вероятностями, соответственно, $p_1 = 0{,}001$, $p_2 = 0{,}01$, $p_3 = 0{,}02$. Так как эти вероятности малы, то для приближенного вычисления искомой вероятности можно использовать многомерное пуассоновское распределение с параметрами $\lambda_1 \approx np_1 = 0{,}365$, $\lambda_2 \approx np_2 = 3{,}65$, $\lambda_3 \approx np_3 = 7{,}3$. Остается получить результат по формуле (9.9):

$$P\left(\mathbf{X} = (5,3,2)\right) \approx \frac{0{,}365^5 \cdot 3{,}65^3 \cdot 7{,}3^2}{5!3!2!}\, e^{-(0{,}365+3{,}65+7{,}3)} = \\ = \frac{16{,}7876}{1440} \cdot 0{,}000012 = 1{,}42 \cdot 10^{-7}.$$

Вектор математических ожиданий и ковариационная матрица в данном случае имеют вид

$$\mathbf{M}X \approx (0{,}365; 3{,}65; 7{,}3)\,, \quad \operatorname{cov}(\mathbf{X}) \approx \begin{pmatrix} 0{,}365 & 0 & 0 \\ 0 & 3{,}65 & 0 \\ 0 & 0 & 7{,}3 \end{pmatrix}.$$

□

9.5 Геометрическое распределение

Говорят, что дискретная с.в. X имеет ***геометрическое распределение***, если ее ряд распределения имеет вид

x_k	1	2	3	...	k	...
p_k	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

Такая с.в. была рассмотрена в примере 3.5. Чтобы найти числовые характеристики с.в., имеющей геометрическое распределение, рассмотрим ее производящую функцию

$$G(u) = \sum_{i=1}^{\infty} p(1-p)^{i-1} u^i = \frac{pu}{1-(1-p)u}$$

и производные этой функции

$$G'(u) = p\frac{1-(1-p)u+(1-p)u}{[1-(1-p)u]^2} = \frac{p}{[1-(1-p)u]^2},$$
$$G''(u) = \frac{2p[1-(1-p)u](1-p)}{[1-(1-p)u]^4} = \frac{2p(1-p)}{[1-(1-p)u]^3}.$$

Тогда

$$\mathbf{M}X = G'(1) = \frac{p}{p^2} = \frac{1}{p},$$
$$\mathbf{D}X = G''(1) + \mathbf{M}X - (\mathbf{M}X)^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} =$$
$$= \frac{2-2p+p-1}{p^2} = \frac{1-p}{p^2}.$$

9.6 Гипергеометрическое распределение

Это распределение возникает в задаче выборочного контроля (пример 1.6). С.в. X имеет ***гипергеометрическое распределение***, если ее ряд распределения образуют пары

$$(x_k, p_k) = \left(k, \frac{C_K^k C_{L-K}^{l-k}}{C_L^l}\right), \quad k = \overline{0, \min(l, K)},\ l \le L,\ K \le L.$$

Это — действительно ряд распределения, так как используя следствие A.1 и договоренность о том, что $C_n^m = 0$, если $m > n$, получим

$$\sum_{k=0}^{K} \frac{C_K^k C_{L-K}^{l-k}}{C_L^l} = \frac{1}{C_L^l}\sum_{k=0}^{K} C_K^k C_{L-K}^{l-k} = \frac{1}{C_L^l}\, C_L^l = 1.$$

Математическое ожидание такой с.в. получим, применив формулу (A.3) и равенство $lC_L^l = LC_{L-1}^{l-1}$, которое следует из тождества (A.2) при $r = 1$, $n = L$ и $k = l$:

$$\mathbf{M}X = \frac{\sum_{k=1}^{K} kC_K^k C_{L-K}^{l-k}}{C_L^l} = \frac{KC_{L-1}^{l-1}}{C_L^l} = \frac{K}{L}\, l.$$

Математическое ожидание квадрата с.в. с использованием формулы (A.5) равно

$$\begin{aligned}
\mathbf{M}X^2 &= \frac{\sum_{k=0}^{K} k^2 C_K^k C_{L-K}^{l-k}}{C_L^l} = \frac{K\left[(K-1)C_{L-2}^{l-2} + C_{L-1}^{l-1}\right]}{C_L^l} = \\
&= K\frac{\left[(K-1)\frac{(L-2)!}{(l-2)!(L-l)!} + \frac{(L-1)!}{(l-1)!(L-l)!}\right] l!(L-l)!}{L!} = \\
&= K\left[\frac{(K-1)(L-2)!l!}{(l-2)!L!} + \frac{(L-1)!l!}{(l-1)!L!}\right] = \\
&= K\left[\frac{(K-1)(l-1)l}{(L-1)L} + \frac{l}{L}\right] = \frac{Kl}{L}\cdot\frac{Kl-l-K+1+L-1}{L-1} = \\
&= \frac{Kl}{L}\cdot\frac{Kl-l-K+L}{L-1}.
\end{aligned}$$

Найдем дисперсию гипергеометрической с.в.:

$$\begin{aligned}
\mathbf{D}X &= \mathbf{M}X^2 - (\mathbf{M}X)^2 = \frac{Kl}{L}\cdot\frac{Kl-l-K+L}{L-1} - \frac{K^2l^2}{L^2} = \\
&= \frac{Kl}{L}\left(\frac{Kl-l-K+L}{L-1} - \frac{Kl}{L}\right) = \\
&= \frac{Kl}{L}\cdot\frac{KlL-lL-KL+L^2-KlL+Kl}{L\,(L-1)} = \\
&= \frac{Kl}{L}\cdot\frac{-lL-KL+L^2+Kl}{L\,(L-1)} = \frac{Kl}{L}\cdot\frac{L\,(L-K)-l\,(L-K)}{L\,(L-1)} = \\
&= \frac{Kl}{L}\cdot\frac{(L-K)\,(L-l)}{L\,(L-1)}.
\end{aligned}$$

$$\mathbf{D}X = K\frac{l}{L}\left(1-\frac{K}{L}\right)\frac{L-l}{L-1}.$$

Лекция 10
Типичные непрерывные распределения

10.1 Равномерное распределение

С.в. X называется распределенной ***равномерно*** на отрезке $[a, b]$, если ее плотность вероятностей имеет вид

$$\boxed{f(x) = \begin{cases} 0, & x \notin [a, b]; \\ \dfrac{1}{b-a}, & x \in [a, b]. \end{cases}}$$

Неотрицательность этой функции очевидна, а условие нормировки легко проверяется:

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = \int\limits_{a}^{b} \frac{1}{b-a}\,dx = \frac{1}{b-a}\int\limits_{a}^{b} dx = \frac{1}{b-a}\,x\Big|_{a}^{b} = \frac{1}{b-a}\,(b-a) = 1.$$

Функция распределения определяется по известной вам формуле:

$$F(x) = \int\limits_{-\infty}^{x} f(x)dx$$

и равна

$$F(x) = \begin{cases} 0, & x < a; \\ \int\limits_a^x \frac{dx}{b-a}, & x \in [a,b]; \\ \int\limits_a^b \frac{dx}{b-a}, & x > b; \end{cases} = \begin{cases} 0, & x < a; \\ \frac{x}{b-a}\Big|_a^x, & x \in [a,b]; \\ \frac{x}{b-a}\Big|_a^b, & x > b; \end{cases} = \tag{10.1}$$

$$= \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & x \in [a,b]; \\ 1, & x > b. \end{cases}$$

Графики плотности вероятностей и функции распределения равномерной с.в. показаны на рис. 10.1.

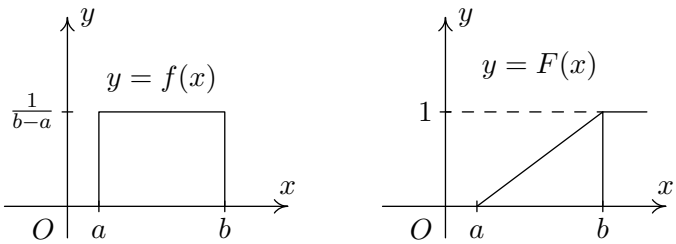


Рис. 10.1: Закон распределения равномерной с.в.

Найдем числовые характеристики с.в. X:

$$\mathbf{M}X = \int\limits_{-\infty}^{\infty} xf(x)dx = \int\limits_a^b \frac{x}{b-a}dx = \frac{1}{b-a} \cdot \frac{x^2}{2}\Big|_a^b = \frac{a+b}{2},$$

$$\mathbf{M}X^2 = \int\limits_{-\infty}^{\infty} xf(x)dx = \int\limits_a^b \frac{x^2}{b-a}dx = \frac{1}{b-a} \cdot \frac{x^3}{3}\Big|_a^b = \frac{a^2+ab+b^2}{3},$$

$$\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2 = \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} =$$

$$= \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12}.$$

Пример 10.1. *Время ожидания X (мин.) пассажиром электрички распределено равномерно на отрезке $[0; 40]$. Найти среднее время ожидания пассажиром электрички и вероятность того, что это время будет меньше 10 мин.*

Решение. Функция распределения равномерной с.в. X выражается формулой (10.1), причем, в данном случае $a = 0$, $b = 40$. Значит, вероятность того, что пассажир будет ожидать электричку меньше 10 мин., равна

$$P(X < 10) = P(-\infty < X < 10) = F(10) - F(-\infty) = \frac{10-0}{40-0} = \frac{1}{4}.$$

Среднее время ожидания пассажиром электрички является математическим ожиданием с.в. X, распределенной равномерно, и потому

$$\mathbf{M}X = \frac{a+b}{2} = \frac{0+40}{2} = 20 \text{ (мин.)}.$$

□

10.2 Нормальное распределение

Говорят, что с.в. X распределена по ***нормальному*** закону, или имеет распределение $N(m, \sigma)$, если она представлена следующей плотностью вероятностей:

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right].}$$

Функция эта действительно является плотностью вероятностей, так как она положительна и выполняется условие нормировки:

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\infty} e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx = \left\langle u = \frac{x-m}{\sqrt{2}\sigma},\ du = \frac{dx}{\sqrt{2}\sigma} \right\rangle =$$

$$= \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} e^{-u^2}\,du = \frac{2}{\sqrt{\pi}} \int\limits_{0}^{\infty} e^{-u^2}\,du = \frac{2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = 1,$$

так как интеграл Пуассона равен (см. приложение C)

$$\int\limits_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

По плотности вероятностей попытаемся найти функцию распределения этой с.в.:

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt =$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int\limits_{-\infty}^{x} \exp\left[-\frac{(t-m)^2}{2\sigma^2}\right]dt = \left\langle u = \frac{t-m}{\sigma}, du = \frac{dt}{\sigma}\right\rangle =$$

$$= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{(x-m)/\sigma} e^{-u^2/2}du.$$

Как известно, интеграл $\int \exp(-u^2)du$ не выражается в элементарных функциях и поэтому функция распределения нормальной с.в. тоже не может быть представлена в виде конечной комбинации элементарных функций. Тем не менее, применяя теорию рядов или численные методы интегрирования, можно построить график и таблицу значений функции распределения для с.в., распределенной $N(0;1)$, а по этим данным найти и значения функции распределения для общего случая с.в., имеющей распределение $N(m,\sigma)$.

Действительно, обозначим плотность вероятностей и функцию распределения с.в. с распределением $N(0;1)$ соответственно

$$\boxed{\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), \quad \Phi^*(x) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{u^2}{2}\right)du.}$$

Тогда плотность вероятностей и функция распределения произвольной нормальной с.в. с распределением $N(m,\sigma)$ выразятся через эти

функции, соответственно, следующим образом:

$$f(x) = \frac{1}{\sigma}\,\varphi\left(\frac{x-m}{\sigma}\right), \quad F(x) = \Phi^*\left(\frac{x-m}{\sigma}\right).$$

Таблицы функций $\varphi(x)$ и $\Phi^*(x)$, называемой также ***интегралом вероятностей***, представлены в приложениях G.1 и G.2, соответственно.

График функции распределения нормальной с.в. типичен для графиков функций распределения непрерывных с.в., а график плотности вероятностей показан на рис. 10.2. Обратите внимание на его симметричность относительно вертикальной прямой $x = m$ и тем большее «распластывание» вдоль оси Ox, чем больше параметр σ (мы увидим, что он равен с.к.о нормальной с.в.).

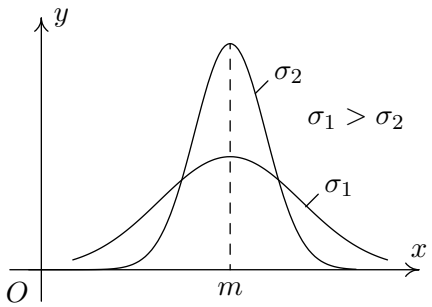


Рис. 10.2: Плотности вероятностей нормальных с.в.

Математическое ожидание нормальной с.в. было найдено в примере 7.1 и оказалось равным параметру m. Чтобы вычислить дисперсию, снова используем интеграл Пуассона.

$$\mathbf{D}X = \int\limits_{-\infty}^{\infty} (x - \mathbf{M}X)^2 f(x)dx =$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int\limits_{-\infty}^{\infty} (x-m)^2 \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx =$$

$$= \left\langle u = \frac{x-m}{\sqrt{2}\sigma}, du = \frac{dx}{\sqrt{2}\sigma} \right\rangle = \frac{2\sigma^2}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} u^2 e^{-u^2} du =$$

$$= \left\langle \alpha = u, d\beta = ue^{-u^2}, d\alpha = du, \beta = -\frac{1}{2}e^{-u^2} \right\rangle =$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \left(-\frac{1}{2} ue^{-u^2} \Big|_{-\infty}^{\infty} + \frac{1}{2} \int\limits_{-\infty}^{\infty} e^{-u^2} du \right) =$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \cdot 2 \cdot \frac{\sqrt{\pi}}{2} = \sigma^2.$$

Таким образом, параметры m и σ нормального распределения с.в. — это ее математическое ожидание и с.к.о.

Очень часто приходится вычислять вероятность попадания нормальной с.в. в заданный отрезок числовой оси $[a, b]$. Применяя известное вам свойство функции распределения с.в. и связь между функциями распределения с.в., распределенными $N(m, \sigma)$ и $N(0; 1)$, получаем

$$\begin{aligned} P(a \le X \le b) = F(b) - F(a) = \Phi^*\left(\frac{b-m}{\sigma}\right) - \Phi^*\left(\frac{a-m}{\sigma}\right) = \\ = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right), \end{aligned} \tag{10.1}$$

где $\Phi(x)$ — часто используемый вместо интеграла вероятностей $\Phi^*(x)$ интеграл $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x \exp\left(-\frac{u^2}{2}\right) du$. Применяя формулу (10.1), необходимо иметь в виду следующие свойства функций $\Phi^*(x)$ и $\Phi(x)$:

$$\Phi^*(-x) = 1 - \Phi^*(x), \quad \Phi(-x) = -\Phi(x), \quad \Phi^*(x) = \frac{1}{2} + \Phi(x).$$

Действительно,

$$\Phi^*(-x)+\Phi^*(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{-x}\exp\left(-\frac{u^2}{2}\right)du+\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{u^2}{2}\right)du=$$

$$=\langle\text{в первом интеграле: } u=-t, du=-dt\rangle=$$

$$=\frac{1}{\sqrt{2\pi}}\left(-\int\limits_{\infty}^{x}+\int\limits_{-\infty}^{x}\right)=\frac{1}{\sqrt{2\pi}}\left(\int\limits_{x}^{\infty}+\int\limits_{-\infty}^{x}\right)=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\exp\left(-\frac{u^2}{2}\right)du=1;$$

$$\Phi(-x)=\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{-x}\exp\left(-\frac{u^2}{2}\right)du=\langle u=-t, du=-dt\rangle=$$

$$=-\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{x}\exp\left(-\frac{t^2}{2}\right)dt=-\Phi(x);$$

$$\Phi^*(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{u^2}{2}\right)du=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}\exp\left(-\frac{u^2}{2}\right)du+\frac{1}{\sqrt{2\pi}}\int\limits_{0}^{x}\exp\left(-\frac{u^2}{2}\right)du=$$

$$=\frac{1}{2}+\Phi(x);$$

так как $\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{0}\exp\left(-\frac{u^2}{2}\right)du=\frac{1}{2}$.

Еще одна полезная формула показывает, что основная доля нормального распределения сосредоточена на отрезке $[m-3\sigma, m+3\sigma]$:

$$\boxed{P(|X-\mathbf{M}X|\le 3\sigma)\approx 0,9973.} \tag{10.2}$$

По-другому: вероятность отклонения нормальной с.в. от своего математического ожидания больше, чем на 3σ, практически равна нулю. Формула (10.2) называется также «***правилом трех сигм***».

Пример 10.2. *Если шарик для подшипника не проходит через отверстие диаметром d_1, но проходит через отверстие диаметром $d_2 > d_1$, то его размер считается приемлемым. В противном случае шарик бракуется. Известно, что диаметр шарика d — распределенная $N(m, \sigma)$ с.в., причем, $m = (d_1 + d_2)/2$, $\sigma = (d_2 - d_1)/4$. Определить вероятность того, что шарик будет забракован.*

Решение. Искомая вероятность p равна вероятности того, что с.в. d не попадет в интервал (d_1, d_2). Поэтому

$$p = 1 - P(d_1 < d < d_2) = 1 - \left[\Phi^*\left(\frac{d_2 - m}{\sigma}\right) - \Phi^*\left(\frac{d_1 - m}{\sigma}\right)\right] =$$
$$= 1 - \left[\Phi^*\left(\frac{d_2 - d_1}{2\sigma}\right) - \Phi^*\left(\frac{d_1 - d_2}{2\sigma}\right)\right] = 2 - 2\Phi^*\left(\frac{d_2 - d_1}{2\sigma}\right) =$$
$$= 2 - 2\Phi^*(2) \approx 0{,}0456.$$

□

10.3 Многомерное нормальное распределение

Случайный вектор $\mathbf{X}^T = (X_1, \ldots, X_n)$ имеет ***нормальное распределение***, если его многомерная плотность вероятностей может быть записана в виде

$$\boxed{f(x_1, \ldots, x_n) = \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{Q}(\mathbf{x}-\mathbf{m})},} \qquad (10.3)$$

где $\mathbf{x}^T = (x_1, \ldots, x_n)$ — вектор переменных, $\mathbf{m}^T = (m_1, \ldots, m_n)$ — числовой вектор, $\mathbf{Q} = \{\sigma_{ij}\}_n^n$ — симметричная положительно определенная (см. приложение D) матрица n-го порядка, $|\mathbf{Q}|$ — ее определитель.

Прежде чем мы убедимся в том, что функция (10.3) действительно является плотностью вероятностей и найдем числовые характеристики нормального случайного вектора, получим некоторые общие представления о преобразованиях случайных векторов. Пусть $f(x_1, \ldots, x_n)$ — плотность вероятностей случайного вектора $\mathbf{X}^T = (X_1, \ldots, X_n)$ и пусть $\mathbf{y} = \mathbf{y}(x_1, \ldots, x_n)$ — отображение, переводящее вектор $\mathbf{x}$ в вектор $\mathbf{y}^T = (y_1, \ldots, y_n)$. Будем предполагать, что данное отображение взаимно однозначно отображает все пространство переменных $(x_1, \ldots, x_n)$ на все пространство переменных $(y_1, \ldots, y_n)$ и обратное отображение имеет вид

$$\mathbf{x} = \mathbf{x}(y_1, \ldots, y_n). \tag{10.4}$$

Пусть случайный вектор $\mathbf{Y}^T = (Y_1, \ldots, Y_n)$ является функцией случайного вектора $\mathbf{X}$: $\mathbf{Y} = \mathbf{y}(X_1, \ldots, X_n)$. Рассмотрим вероятность попадания случайной точки $(X_1, \ldots, X_n)$ во множество S:

$$P((X_1, \ldots, X_n) \in S) = \int\limits_S f(x_1, \ldots, x_n)\, d\mathbf{x}^T.$$

Произведем замену переменных интегрирования в этом интеграле по формуле (10.4) с якобианом перехода $J = |\partial x_i / \partial y_j|$:

$$P((X_1, \ldots, X_n) \in S) = \int\limits_S f(x_1, \ldots, x_n)\, d\mathbf{x}^T =$$

$$= \int\limits_{S'} f[x_1(y_1, \ldots, y_n), \ldots, x_n(y_1, \ldots, y_n)]\, |J|\, d\mathbf{y}^T =$$

$$= \int\limits_{S'} g(y_1, \ldots, y_n)\, d\mathbf{y}^T,$$

где S — образ множества S' при отображении (10.4). Функция

$$g(y_1, \ldots, y_n) = f[x_1(y_1, \ldots, y_n), \ldots, x_n(y_1, \ldots, y_n)]\, |J| \tag{10.5}$$

неотрицательна, так как неотрицательна функция $f(x_1, \ldots, x_n)$. Если в качестве множества S взять все пространство, то вероятность в левой части полученного равенства будет равна 1, а тогда и интеграл справа будет равен 1. Но в силу того, что отображение (10.4) отображает пространство на пространство, множество S' тоже будет всем пространством и поэтому получающееся при этом равенство

$$\int\limits_{-\infty}^{\infty} g(y_1, \ldots, y_n)\, d\mathbf{y}^T = 1$$

окончательно характеризует функцию $g(y_1, \ldots, y_n)$ как плотность вероятностей случайного вектора $\mathbf{Y}$. Следовательно, можно написать, что

$$P((X_1, \ldots, X_n) \in S) = \int\limits_{S} f(x_1, \ldots, x_n)\, d\mathbf{x}^T =$$
$$= \int\limits_{S'} g(y_1, \ldots, y_n)\, d\mathbf{y}^T = P\left((Y_1, \ldots, Y_n) \in S'\right).$$

Вернемся теперь к функции (10.3) и убедимся в том, что она действительно является плотностью вероятностей. Неотрицательность этой функции следует из положительной определенности матрицы $\mathbf{Q}$ и положительности ее определителя (см. теорему D.1). Докажем, что

$$\int_{-\infty}^{\infty} f(x_1, \ldots, x_n)\, d\mathbf{x} = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = 1. \tag{10.6}$$

Произведем преобразование случайного вектора $\mathbf{X}$ вида

$$\mathbf{X} = \mathbf{K}\mathbf{Y} + \mathbf{m}, \tag{10.7}$$

где $\mathbf{K}$ — невырожденная матрица, приводящая матрицу $\mathbf{Q}$ к единичной матрице $\mathbf{E}$ (теорема D.2). Такому преобразованию соответствует

замена переменных $\mathbf{x}-\mathbf{m}=\mathbf{K}\mathbf{y}$ в интеграле

$$\int\limits_{-\infty}^{\infty} \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T\mathbf{Q}(\mathbf{x}-\mathbf{m})}\, d\mathbf{x}^T$$

с якобианом перехода $|\mathbf{J}|=|\mathbf{K}|=1/\sqrt{p_1\cdots p_n}$. После замены получим

$$\int\limits_{-\infty}^{\infty} \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T\mathbf{Q}(\mathbf{x}-\mathbf{m})}\, d\mathbf{x}^T =$$

$$= \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}} \frac{1}{\sqrt{p_1\cdots p_n}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T\mathbf{K}^T\mathbf{Q}\mathbf{K}\mathbf{y}} d\mathbf{y}^T =$$

$$= \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}} \frac{1}{\sqrt{p_1\cdots p_n}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T\mathbf{E}\mathbf{y}} d\mathbf{y}^T =$$

$$= \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}} \frac{1}{\sqrt{p_1\cdots p_n}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\left(y_1^2+\cdots+y_n^2\right)}\, d\mathbf{y}^T =$$

$$= \frac{\sqrt{|\mathbf{Q}|}}{\sqrt{p_1\cdots p_n}} \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}y_1^2}\, dy_1 \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}y_n^2}\, dy_n.$$

Так как $\int_{-\infty}^{\infty} e^{-x^2}\, dx=\sqrt{\pi}$, то

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\, dy = 1. \tag{10.8}$$

Действительно,

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\, dy = \left\langle y=\sqrt{2}x,\ dy=\sqrt{2}\,dx \right\rangle =$$

$$= \frac{\sqrt{2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{\sqrt{\pi}} = 1.$$

Кроме того, $|\mathbf{Q}| = p_1 \cdots p_n$ (теорема D.1). Следовательно,

$$\int\limits_{-\infty}^{\infty} \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{Q}(\mathbf{x}-\mathbf{m})}\, d\mathbf{x}^T = 1.$$

Найдем числовые характеристики нормального случайного вектора. Очевидно, с.в. Y_i распределена $N(0,1)$, потому что ее плотность вероятностей имеет вид $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y_i^2}\, dy_i$ (см. (10.8)). Тогда $\mathbf{M}(\mathbf{Y}) = \mathbf{0}$ и, применяя операцию математического ожидания к преобразованию (10.7), получим

$$\mathbf{MX} = \mathbf{M}(\mathbf{KY}) + \mathbf{m} = \mathbf{KM}(\mathbf{Y}) + \mathbf{m} = \mathbf{m}.$$

Далее найдем ковариацию с.в. Y_i и Y_j. Если $i = j$, то $\mathbf{M}Y_iY_j = \mathbf{D}Y_i = 1$; если же $i \neq j$, то

$$\mathbf{M}Y_iY_j = \frac{1}{\left(\sqrt{2\pi}\right)^2} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} y_iy_je^{-\frac{1}{2}\left(y_i^2+y_j^2\right)}\, dy_i\, dy_j =$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} y_ie^{-\frac{1}{2}y_i^2}\, dy_i \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} y_je^{-\frac{1}{2}y_j^2}\, dy_j = \mathbf{M}Y_i\mathbf{M}Y_j = 0,$$

и, значит, $\mathbf{MYY}^T = \mathbf{E}$. Но тогда

$$\operatorname{cov}(\mathbf{X}) = \mathbf{M}(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^T = \mathbf{M}\left(\mathbf{KYY}^T\mathbf{K}^T\right) =$$
$$= \mathbf{KM}\left(\mathbf{YY}^T\right)\mathbf{K}^T = \mathbf{KEK}^T = \mathbf{KK}^T.$$

С другой стороны, по определению матрицы $\mathbf{K}$ имеем $\mathbf{K}^T\mathbf{QK} = \mathbf{E}$. Следовательно, $\mathbf{Q} = \left(\mathbf{K}^T\right)^{-1}\mathbf{K}^{-1}$, или $\mathbf{Q}^{-1} = \mathbf{KK}^T$. Значит,

$$\operatorname{cov}(\mathbf{X}) = \mathbf{Q}^{-1}.$$

Тем самым доказано, что нормальный случайный вектор имеет числовые характеристики

$$\mathbf{MX} = \mathbf{m},$$
$$\operatorname{cov}(\mathbf{X}) = \mathbf{Q}^{-1}.$$

Для такого вектора будем обозначать $\mathbf{\Sigma} = \operatorname{cov}(\mathbf{X})$ и говорить, что вектор $\mathbf{X}$ имеет распределение $N(\mathbf{m}, \mathbf{\Sigma})$. Если нормальный случайный вектор $\mathbf{X}$ задается вектором $\mathbf{m}$, своими с.к.о. $\sigma_i = \sqrt{\mathbf{D}X_i}$ и коэффициентами корреляции r_{ij} с.в. X_i и X_j, $i \neq j$, то $\mathbf{\Sigma} = \{\sigma_i \sigma_j r_{ij}\}_n^n$, причем, $r_{ii} = 1$, и матрицу $\mathbf{Q}$ тогда можно получить по формуле

$$\mathbf{Q} = \mathbf{\Sigma}^{-1} = \left[\{\sigma_i \sigma_j r_{ij}\}_n^n\right]^{-1}.$$

Если с.в. $X_1, \ldots, X_n$ некоррелированы, то $r_{ij} = 0$ $(i \neq j)$ и ковариационная матрица оказывается диагональной:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \sigma_n^2 \end{pmatrix}.$$

Матрица $\mathbf{Q}$ принимает вид

$$\mathbf{Q} = \mathbf{\Sigma}^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \frac{1}{\sigma_n^2} \end{pmatrix},$$

с определителем $|\mathbf{Q}| = (\sigma_1 \ldots \sigma_n)^{-2}$, а квадратичная форма становится такой:

$$(\mathbf{x} - \mathbf{m})^T \mathbf{Q} (\mathbf{x} - \mathbf{m}) = \sum_{i=1}^n \frac{(x_i - m_i)^2}{\sigma_i^2}.$$

Следовательно, плотность вероятностей (10.3) нормального случайного вектора может быть записана в виде

$$\begin{aligned} f(x_1, \ldots, x_n) &= \frac{1}{(2\pi)^{n/2} \sigma_1 \ldots \sigma_n} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - m_i)^2}{\sigma_i^2}\right) = \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x_1 - m_1)^2}{2\sigma_1^2}} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x_n - m_n)^2}{2\sigma_n^2}} = \\ &= f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n). \end{aligned}$$

Таким образом, доказана следующая

Теорема 10.1. *Если с.в. $X_1, \ldots, X_n$ имеют совместное нормальное распределение и некоррелированны, то они независимы.*

Рассмотрим как влияет невырожденное линейное преобразование на нормальный случайный вектор.

Теорема 10.2. *Если случайный вектор $\mathbf{X}^T = (X_1, \ldots, X_n)$ имеет распределение $N(\mathbf{m}, \mathbf{\Sigma})$, то вектор $\mathbf{Y} = \mathbf{R}\mathbf{X}$, где $\mathbf{R}$ — невырожденная матрица, имеет распределение $N\left(\mathbf{R}\mathbf{m}, \mathbf{R}\mathbf{\Sigma}\mathbf{R}^T\right)$.*

Доказательство. В соответствии с рассмотренными выше элементами теории преобразований случайных векторов, в частности, с формулой (10.5), плотностью вероятностей случайного вектора $\mathbf{Y}$ будет функция

$$g(y_1, \ldots, y_n) = \frac{\sqrt{|\mathbf{Q}|}}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{Q}(\mathbf{x}-\mathbf{m})} \left|\mathbf{R}^{-1}\right|\Big|_{\mathbf{x}=\mathbf{R}^{-1}\mathbf{y}}. \tag{10.9}$$

Сначала преобразуем определитель $\left|\mathbf{R}^{-1}\right|$:

$$\left|\mathbf{R}^{-1}\right| = \frac{1}{|\mathbf{R}|} = \sqrt{\frac{1}{|\mathbf{R}|^2}} = \sqrt{\frac{|\mathbf{\Sigma}|}{|\mathbf{R}|\,|\mathbf{\Sigma}|\,|\mathbf{R}^T|}} = \frac{\sqrt{|\mathbf{\Sigma}|}}{\sqrt{\left|\mathbf{R}\mathbf{\Sigma}\mathbf{R}^T\right|}}.$$

Теперь преобразуем квадратичную форму:

$$\begin{aligned}
(\mathbf{x}-\mathbf{m})^T \mathbf{Q}\,(\mathbf{x}-\mathbf{m}) &= (\mathbf{x}-\mathbf{m})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\mathbf{m}) = \\
&= \left(\mathbf{R}^{-1}\mathbf{y}-\mathbf{m}\right)^T \mathbf{\Sigma}^{-1}\left(\mathbf{R}^{-1}\mathbf{y}-\mathbf{m}\right) = \\
&= \left(\mathbf{R}^{-1}\mathbf{y}-\mathbf{R}^{-1}\mathbf{R}\mathbf{m}\right)^T \mathbf{\Sigma}^{-1}\left(\mathbf{R}^{-1}\mathbf{y}-\mathbf{R}^{-1}\mathbf{R}\mathbf{m}\right) = \\
&= \left[\mathbf{R}^{-1}(\mathbf{y}-\mathbf{R}\mathbf{m})\right]^T \mathbf{\Sigma}^{-1}\left[\mathbf{R}^{-1}(\mathbf{y}-\mathbf{R}\mathbf{m})\right] = \\
&= (\mathbf{y}-\mathbf{R}\mathbf{m})^T \left(\mathbf{R}^{-1}\right)^T \mathbf{\Sigma}^{-1}\mathbf{R}^{-1}(\mathbf{y}-\mathbf{R}\mathbf{m}) = \\
&= (\mathbf{y}-\mathbf{R}\mathbf{m})^T \left(\mathbf{R}\mathbf{\Sigma}\mathbf{R}^T\right)^{-1}(\mathbf{y}-\mathbf{R}\mathbf{m}).
\end{aligned}$$

Используя эти формулы, из равенства (10.9) получаем плотность вероятностей случайного вектора $\mathbf{Y}$:

$$g\left(y_1,\ldots,y_n\right)=\frac{\sqrt{\left|\mathbf{\Sigma}^{-1}\right|}}{(2\pi)^{n/2}}\cdot\frac{\sqrt{\left|\mathbf{\Sigma}\right|}}{\sqrt{\left|\mathbf{R\Sigma R}^T\right|}}\,e^{-\frac{1}{2}(\mathbf{y}-\mathbf{Rm})^T\left(\mathbf{R}^T\mathbf{\Sigma R}\right)^{-1}(\mathbf{y}-\mathbf{Rm})}=$$

$$=\frac{\sqrt{\left|\mathbf{R\Sigma R}^T\right|^{-1}}}{(2\pi)^{n/2}}\,e^{-\frac{1}{2}(\mathbf{y}-\mathbf{Rm})^T\left(\mathbf{R}^T\mathbf{\Sigma R}\right)^{-1}(\mathbf{y}-\mathbf{Rm})}.$$

Это выражение и показывает, что случайный вектор $\mathbf{Y}=\mathbf{RX}$ имеет распределение $N\left(\mathbf{Rm},\mathbf{R\Sigma R}^T\right)$. □

Таким образом невырожденное линейное преобразование не нарушает нормальности случайного вектора.

10.4 Двумерное нормальное распределение

Частным случаем многомерного нормального распределения является ***двумерное нормальное*** распределение СВНТ $\mathbf{X}=(X,Y)$. Чтобы его получить, примем, что число n с.в., образующих вектор $\mathbf{X}$, равно 2, и запишем матрицу $\mathbf{\Sigma}=\{\sigma_i\sigma_j r_{ij}\}_n^n$, полученную в предыдущем п.:

$$\mathbf{\Sigma}=\begin{pmatrix}\sigma_X^2 & r\sigma_X\sigma_Y\\ r\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix},$$

где $\sigma_X=\sigma_1$, $\sigma_Y=\sigma_2$, $r=r_{12}$. Далее найдем матрицу

$$\mathbf{Q}=\mathbf{\Sigma}^{-1}=\frac{1}{1-r^2}\begin{pmatrix}\frac{1}{\sigma_X^2} & -\frac{r}{\sigma_X\sigma_Y}\\ -\frac{r}{\sigma_X\sigma_Y} & \frac{1}{\sigma_Y^2}\end{pmatrix}$$

и получим квадратичную форму

$$(\mathbf{x}-\mathbf{m})^T\mathbf{Q}(\mathbf{x}-\mathbf{m})=$$

$$= (x - m_X, y - m_Y)^T \frac{1}{1-r^2} \begin{pmatrix} \frac{1}{\sigma_X^2} & -\frac{r}{\sigma_X \sigma_Y} \\ -\frac{r}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{pmatrix} \begin{pmatrix} x - m_X \\ y - m_Y \end{pmatrix} =$$

$$= \frac{1}{1-r^2} \cdot$$

$$\cdot \left(\frac{x - m_X}{\sigma_X^2} - \frac{r(y - m_Y)}{\sigma_X \sigma_Y}, -\frac{r(x - m_X)}{\sigma_X \sigma_Y} + \frac{y - m_Y}{\sigma_Y^2} \right)^T \begin{pmatrix} x - m_X \\ y - m_Y \end{pmatrix} =$$

$$= \frac{1}{1-r^2} \left[\frac{(x - m_X)^2}{\sigma_X^2} - 2r \frac{(x - m_X)(y - m_Y)}{\sigma_X \sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2} \right].$$

Кроме того, так как для любой матрицы $\mathbf{A}$ порядка n и числа α выполняется свойство $|\alpha \mathbf{Q}| = \alpha^n |\mathbf{Q}|$, то

$$|\mathbf{Q}| = \frac{1}{(1-r^2)^2} \begin{vmatrix} \frac{1}{\sigma_X^2} & -\frac{r}{\sigma_X \sigma_Y} \\ -\frac{r}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{vmatrix} = \frac{1}{(1-r^2)^2} \left(\frac{1}{\sigma_X^2 \sigma_Y^2} - \frac{r^2}{\sigma_X^2 \sigma_Y^2} \right) =$$

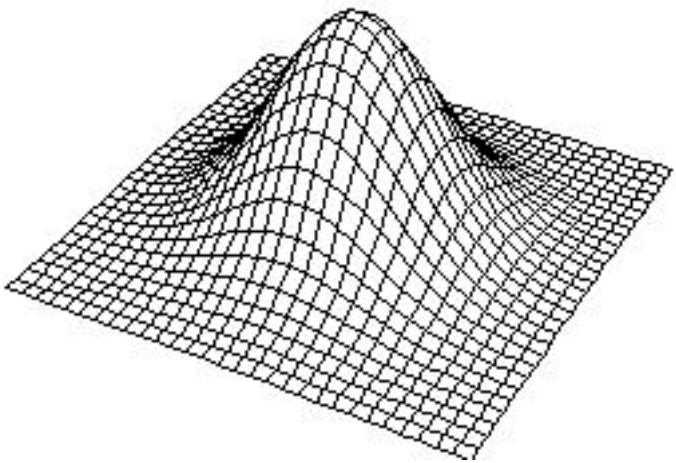
$$= \frac{1-r^2}{\sigma_X^2 \sigma_Y^2 (1-r^2)^2} = \frac{1}{\sigma_X^2 \sigma_Y^2 (1-r^2)},$$

$$\sqrt{|\mathbf{Q}|} = \frac{1}{\sigma_X \sigma_Y \sqrt{1-r^2}}.$$

Поэтому в двумерном случае плотность вероятностей (10.3) принимает вид

$$\boxed{\begin{aligned} f(x,y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-r^2}} \exp\left\{ -\frac{1}{2(1-r^2)} \left[\frac{(x - m_X)^2}{\sigma_X^2} - \right.\right. \\ \left.\left. -2r \frac{(x - m_X)(y - m_Y)}{\sigma_X \sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2} \right] \right\}; \\ \sigma_X, \sigma_Y > 0; |r| < 1. \end{aligned}}$$

На следующем рисунке показан график этой функции с максимумом в точке (m_X, m_Y), имеющий вид поверхности, распростертой над плоскостью xOy.



Как было показано, частные плотности вероятностей компонент случайного вектора имеют вид

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right],$$
$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left[-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right],$$

то есть X и Y имеют нормальное одномерное распределение.

10.5 Показательное распределение

Распределение с.в. X называется ***показательным***, если плотность ее вероятностей равна

$$\boxed{f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0, \end{cases}}$$

$\lambda > 0$.

Эта функция действительно является плотностью вероятностей, так как

$$\int\limits_{-\infty}^{\infty} f(x)\,dx = \int\limits_{0}^{\infty} \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\Big|_0^{\infty} = e^0 = 1.$$

Функция распределения выражается формулой

$$F(x) = \int\limits_{-\infty}^{x} f(x)dx = \begin{cases} \int\limits_0^x \lambda e^{-\lambda x}dx, & x \geq 0 \\ \int\limits_{-\infty}^{\infty} 0\,dx, & x < 0; \end{cases} =$$

$$= \begin{cases} 1 - e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Графики функции распределения и плотности вероятностей с.в., имеющей показательное распределение, представлены на рис. 10.3.

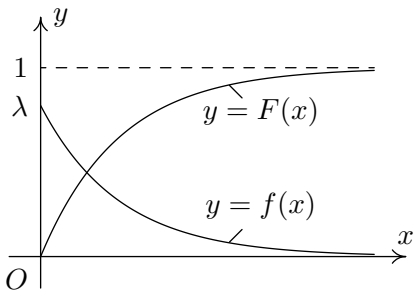


Рис. 10.3: Показательный закон распределения

Найдем числовые характеристики с.в. X. Математическое ожидание:

$$\mathbf{M}X = \int\limits_{-\infty}^{\infty} xf(x)dx = \int\limits_{0}^{\infty} x\lambda e^{-\lambda x}dx =$$

$$= \langle u = x, dv = \lambda e^{-\lambda x} dx, du = dx, v = -e^{-\lambda x} \rangle =$$

$$= -xe^{-\lambda x}\Big|_0^\infty + \int\limits_0^\infty e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x}\Big|_0^\infty = \frac{1}{\lambda};$$

дисперсия:

$$\mathbf{M}X^2 = \int\limits_{-\infty}^{\infty} x^2 f(x) dx = \int\limits_0^\infty x^2 \lambda e^{-\lambda x} dx =$$

$$= \langle u = x^2, dv = \lambda e^{-\lambda x} dx, du = 2x dx, v = -e^{-\lambda x} \rangle =$$

$$= -x^2 e^{-\lambda x}\Big|_0^\infty + 2\int\limits_0^\infty x e^{-\lambda x} dx = \frac{2\mathbf{M}X}{\lambda} = \frac{2}{\lambda^2};$$

$$\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Пример 10.3. *Время X безотказной работы радиоаппаратуры является с.в., имеющей показательное распределение с параметром $\lambda = 1/2$. Вычислить математическое ожидание, дисперсию данной с.в. и вероятность того, что радиоаппаратура не выйдет из строя в течение времени $t = \mathbf{M}X$.*

Решение. В соответствии с формулами, полученными для числовых характеристик показательной с.в., найдем, что

$$\mathbf{M}X = 1/\lambda = 2,\ \mathbf{D}X = 1/\lambda^2 = 4.$$

Затем вычислим требуемую вероятность:

$$P\left(X \geq \mathbf{M}X\right) = 1 - P\left(X < \mathbf{M}X\right) = 1 - F(\mathbf{M}X) =$$

$$= 1 - \left(1 - e^{-\lambda \cdot \mathbf{M}X}\right) = e^{-\lambda/\lambda} = e^{-1} \approx 0{,}368.$$

□

10.6 Распределение Эрланга k-го порядка

Плотность вероятностей св.в. X, распределенной по закону ***Эрланга k-го порядка***, имеет вид

$$\boxed{f_k(x) = \begin{cases} \frac{\lambda(\lambda x)^{k-1}}{(k-1)!}\, e^{-\lambda x}, & x > 0; \\ 0, & x \le 0; \end{cases}}$$

$k = 1, 2, \ldots$. При $k = 1$ имеем показательное распределение.

Чтобы доказать, что эта функция действительно является плотностью вероятностей, рассмотрим интеграл

$$L_k = \int\limits_0^\infty \frac{\lambda\,(\lambda x)^{k-1}}{(k-1)!}\, e^{-\lambda x}\, dx = \frac{\lambda^k}{(k-1)!} \int\limits_0^\infty x^{k-1}\, e^{-\lambda x}\, dx =$$

$$= \left\langle u = x^{k-1}, dv = e^{-\lambda x}\, dx, du = (k-1)\, x^{k-2}\, dx, v = -\frac{1}{\lambda} e^{-\lambda x} \right\rangle =$$

$$= \frac{\lambda^k}{(k-1)!} \left(-\frac{1}{\lambda} x^{k-1} e^{-\lambda x} \Big|_0^\infty + \frac{k-1}{\lambda} \int\limits_0^\infty x^{k-2}\, e^{-\lambda x}\, dx \right) =$$

$$= \int\limits_0^\infty \frac{\lambda\,(\lambda x)^{k-2}}{(k-2)!}\, e^{-\lambda x}\, dx = L_{k-1}.$$

Изучение показательного распределения выявило, что $L_1 = 1$. Следовательно, $L_k = 1$ для любого k. Тогда

$$\int\limits_{-\infty}^{\infty} f_k\,(x)\; dx = \int\limits_0^\infty \frac{\lambda\,(\lambda x)^{k-1}}{(k-1)!}\, e^{-\lambda x}\, dx = 1.$$

Графики плотностей распределения Эрланга k-го порядка при различных k показаны на рис 10.4.

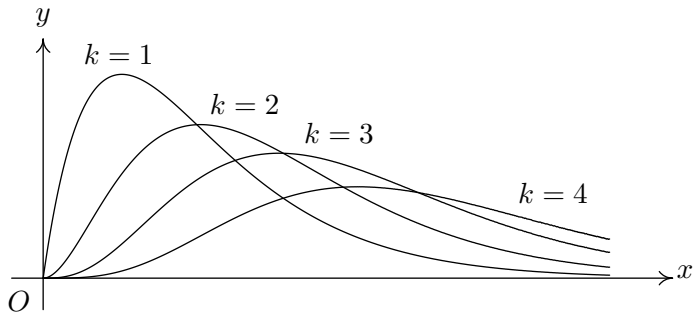


Рис. 10.4: Плотности вероятностей для распределения Эрланга

Найдем функцию распределения. Пусть $x \geq 0$. В этом случае при $k > 1$ (см. предыдущие выкладки)

$$F_k(x) = \int\limits_{-\infty}^{x} f_k(x)\, dx = \int\limits_{0}^{x} \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}\, dx =$$

$$= \frac{\lambda^k}{(k-1)!} \left(-\frac{1}{\lambda} x^{k-1} e^{-\lambda x} \Big|_0^x + \frac{k-1}{\lambda} \int\limits_0^x x^{k-2} e^{-\lambda x}\, dx \right) =$$

$$= -\frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} + F_{k-1}(x).$$

Следовательно, продолжая и т.д., получим

$$F_k(x) = -\frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} - \frac{(\lambda x)^{k-2}}{(k-2)!} e^{-\lambda x} - \cdots - \lambda x e^{-\lambda x} + F_1(x).$$

Но $F_1(x) = 1 - e^{-\lambda x}$ и поэтому

$$F_k(x) = -\frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} - \frac{(\lambda x)^{k-2}}{(k-2)!} e^{-\lambda x} - \cdots - \lambda x e^{-\lambda x} + 1 - e^{-\lambda x},$$

или

$$F_k(x) = \begin{cases} 1 - \sum\limits_{m=0}^{k-1} \frac{(\lambda x)^m}{m!} e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}.$$

Определим числовые характеристики. Математическое ожидание:

$$\mathbf{M}X = \int\limits_{-\infty}^{\infty} x f_k(x)\, dx = \int\limits_{0}^{\infty} x \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}\, dx =$$

$$= \frac{k}{\lambda} \int\limits_{0}^{\infty} \frac{\lambda(\lambda x)^k}{k!} e^{-\lambda x}\, dx = \frac{k}{\lambda} \int\limits_{-\infty}^{\infty} f_{k+1}(x)\, dx = \frac{k}{\lambda}.$$

Дисперсия:

$$\mathbf{M}X^2 = \int\limits_{-\infty}^{\infty} x^2 f_k(x)\, dx = \int\limits_{0}^{\infty} x^2 \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}\, dx =$$

$$= \frac{k(k+1)}{\lambda^2} \int\limits_{0}^{\infty} \frac{\lambda(\lambda x)^{k+1}}{(k+1)!} e^{-\lambda x}\, dx =$$

$$= \frac{k(k+1)}{\lambda^2} \int\limits_{-\infty}^{\infty} f_{k+2}(x)\, dx = \frac{k(k+1)}{\lambda^2},$$

$$\mathbf{D}X = \mathbf{M}X^2 - (\mathbf{M}X)^2 = \frac{k(k+1)}{\lambda^2} - \frac{k^2}{\lambda^2} = \frac{k}{\lambda^2}.$$

Пример 10.4. *Показать, что закон Эрланга возникает при сложении независимых с.в., подчиненных показательному закону с одним и тем же параметром λ.*

Решение. Вначале докажем вспомогательное утверждение. Пусть с.в. Z является суммой двух независимых с.в.: $Z = X + Y$, причем, с.в.

X распределена по закону Эрланга k-го порядка, а с.в. Y — по показательному закону с тем же параметром λ. Тогда в соответствии с формулой (6.4) сумма будет иметь плотность вероятностей

$$f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx = \int\limits_0^z \frac{\lambda(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x} \lambda e^{-\lambda(z-x)}\, dx =$$

$$= \frac{\lambda^{k+1}}{(k-1)!} e^{-\lambda z} \int\limits_0^z x^{k-1}\, dx = \frac{\lambda^{k+1}}{(k-1)!} \cdot \frac{x^k}{k} e^{-\lambda z}\Big|_0^z = \frac{\lambda(\lambda z)^k}{k!} e^{-\lambda z},$$

$z \geq 0$.

Таким образом, сумма имеет распределение Эрланга $(k+1)$-го порядка. Пусть с.в. X является следующей суммой: $X = X_1 + \cdots + X_n$, причем, все с.в. в этой сумме независимы и распределены по показательному закону с одним и тем же параметром λ. Из доказанного получаем, что сумма первых двух слагаемых распределена по закону Эрланга 2-го порядка, первых трех слагаемых — по закону Эрланга 3-го порядка и т.д., наконец, сумма n слагаемых распределена по закону Эрланга n-го порядка. □

10.7 Нормированное распределение Эрланга

Пусть с.в. Z имеет распределение Эрланга k-го порядка. Нормированным распределением Эрланга k-го порядка называется распределение с.в. $X = Z/k$. Найдем ее плотность вероятностей, функцию распределения и числовые характеристики. Функция распределения будет равна

$$F_{X,k}(x) = P(X < x) = P(Z/k < x) = P(Z < kx) = F_{Z,k}(kx),$$

где $F_{X,k}$ — функция распредления с.в. X, а $F_{Z,k}(z)$ — функция распределения с.в. Z, полученная в предыдущем п. Используя ее, нахо-

дим, что

$$F_{X,k}(x) = \begin{cases} 1 - \sum\limits_{m-0}^{k-1} \frac{(\lambda k x)^m}{m!} e^{-\lambda k x}, & x > 0; \\ 0, & x \le 0. \end{cases} =$$

$$= \begin{cases} 1 - \sum\limits_{m-0}^{k-1} \frac{\left(\overline{\lambda} x\right)^m}{m!} e^{-\overline{\lambda} x}, & x > 0; \\ 0, & x \le 0; \end{cases}$$

где $\overline{\lambda} = \lambda k$. Таким образом, нормированное распределение Эрланга k-го порядка совпадает с обычным распределением Эрланга k-го порядка, если параметр λ заменить параметром $\overline{\lambda}$. Поэтому остальные результаты прямо следуют из результатов предыдущей главы. Так, плотность вероятностей $f_{X,k}$ с.в. X определяется формулой

$$\boxed{f_{X,k} = \begin{cases} \frac{\overline{\lambda}\left(\overline{\lambda} x\right)^{k-1}}{(k-1)!} e^{-\overline{\lambda} x}, & x > 0; \\ 0, & x \le 0. \end{cases}}$$

Числовые характеристики получаются такими:

$$\mathbf{M}X = \frac{k}{\overline{\lambda}} = \frac{1}{\lambda},$$

$$\mathbf{D}X = \frac{k}{\overline{\lambda}^2} = \frac{1}{k\lambda^2}.$$

Отметим, что математическое ожидание не зависит от k, т.е. оно остается одним и тем же для всего семейства нормированных распределений Эрланга k-го порядка. При $k \to \infty$ дисперсия $\mathbf{D}X \to 0$. Это означает, что полученная таким образом предельная с.в. на самом деле случайной величиной уже не является и фактически отождествляется со своим математическим ожиданием $1/\lambda$.

Обратно, по числовым характеристикам можно найти параметры λ и k:

$$\lambda = \frac{1}{\mathbf{M}X}, \; k = \frac{(\mathbf{M}X)^2}{\mathbf{D}X}.$$

10.8 t-распределение Стьюдента

Говорят, что с.в. T имеет ***t-распределение***, или ***распределение Стьюдента*** с r степенями свободы, если ее плотность вероятностей может быть записана в виде

$$\boxed{f(t) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\left(1+\frac{t^2}{r}\right)^{-\frac{r+1}{2}},}$$

где $\Gamma(x)$ — гамма-функция (приложение F). Используя свойства гамма- и бета-функций, покажем, что условие нормировки выполнено:

$$\int\limits_{-\infty}^{\infty} f(t)\,dx = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\int\limits_{-\infty}^{\infty}\left(1+\frac{t^2}{r}\right)^{-\frac{r+1}{2}}dt =$$

$$= \frac{2\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\int\limits_{0}^{\infty}\left(1+\frac{t^2}{r}\right)^{-\frac{r+1}{2}}dt =$$

$$= \left\langle t = \sqrt{\frac{ru}{1-u}},\ dt = \frac{r}{2}\sqrt{\frac{1-u}{ru}}\cdot\frac{du}{(1-u)^2}\right\rangle =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\int\limits_{0}^{1}\left(1+\frac{u}{1-u}\right)^{-\frac{r+1}{2}}\sqrt{\frac{1-u}{ru}}\cdot\frac{du}{(1-u)^2} =$$

$$= \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)}\int\limits_{0}^{1}\left(\frac{1}{1-u}\right)^{-\frac{r+1}{2}}\sqrt{\frac{1-u}{u}}\cdot\frac{du}{(1-u)^2} =$$

$$= \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)}\int\limits_{0}^{1}u^{-1/2}(1-u)^{r/2-1}\,du =$$

$$= \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)}\int\limits_{0}^{1}u^{1/2-1}(1-u)^{r/2-1}\,du = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)}B(1/2, r/2) =$$

$$= \frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)\sqrt{\pi}} \cdot \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{r}{2}\right)}{\Gamma\left(\frac{1}{2}+\frac{r}{2}\right)} = \frac{\Gamma\left(\frac{r+1}{2}\right)\sqrt{\pi}}{\sqrt{\pi}\Gamma\left(\frac{r+1}{2}\right)} = 1.$$

Графики плотностей вероятностей для различных r показаны на рис 10.5.

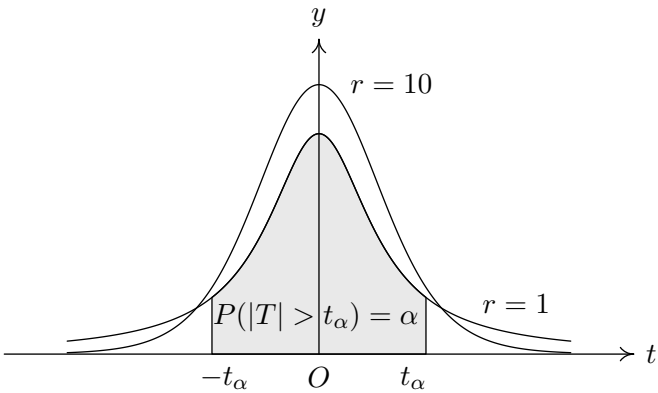


Рис. 10.5: Плотности вероятностей для t-распределения

Функция распределения данной с.в. определяется формулой

$$F(t) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{-\infty}^{t} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} dt.$$

Математическое ожидание с.в. T равно нулю:

$$\mathbf{M}T = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{-\infty}^{\infty} t\left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} dt = 0,$$

так как подынтегральная функция нечетна, а пределы интегрирования симметричны относительно нуля.

Определим дисперсию с.в. T:

$$\mathbf{D}T = \mathbf{M}T^2 = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{-\infty}^{\infty} t^2 \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} dt =$$

$$= \frac{2\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{\infty} t^2 \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}} dt =$$

$$= \left\langle t = \sqrt{\frac{ru}{1-u}},\ dt = \frac{r}{2}\sqrt{\frac{1-u}{ru}} \cdot \frac{du}{(1-u)^2} \right\rangle =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{1} \frac{ru}{1-u} \left(1 + \frac{u}{1-u}\right)^{-\frac{r+1}{2}} \sqrt{\frac{1-u}{ru}} \cdot \frac{du}{(1-u)^2} =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{1} \frac{u}{1-u} \cdot \left(\frac{1}{1-u}\right)^{-\frac{r+1}{2}} \sqrt{\frac{1-u}{u}} \cdot \frac{du}{(1-u)^2} =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{1} u^{1/2} (1-u)^{r/2-2}\, du =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{1} u^{3/2-1} (1-u)^{(r/2-1)-1}\, du =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} B\left(\frac{3}{2}\right) B\left(\frac{r}{2} - 1\right) = \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} \cdot \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{r}{2}-1\right)}{\Gamma\left(\frac{3}{2}+\frac{r}{2}-1\right)} =$$

$$= \frac{r\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} \cdot \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{r}{2}-1\right)}{\Gamma\left(\frac{r+1}{2}\right)} = \frac{r\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{r}{2}-1\right)}{\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)}.$$

Но

$$\Gamma\left(\frac{3}{2}\right) = \Gamma\left(1 + \frac{1}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2},$$

$$\Gamma\left(\frac{r}{2}\right) = \Gamma\left(\left(\frac{r}{2} - 1\right) + 1\right) = \left(\frac{r}{2} - 1\right)\Gamma\left(\frac{r}{2} - 1\right).$$

Поэтому

$$\mathbf{D}T = \frac{r\frac{\sqrt{\pi}}{2}\Gamma\left(\frac{r}{2}-1\right)}{\sqrt{\pi}\left(\frac{r}{2}-1\right)\Gamma\left(\frac{r}{2}-1\right)} = \frac{r}{r-2}.$$

Таблицы t-распределения приведены в приложении G.3 в следующем виде: для заданной вероятности α и заданного числа степеней свободы r в таблице указаны значения аргумента t_α плотности вероятностей, для которых выполняется равенство

$$P(|T| > t_\alpha) = 2\int\limits_{t_\alpha}^{\infty} f(t)\,dt = \alpha.$$

Фактически значение аргумента t_α обеспечивает попадание с.в. T с вероятностью α в интервал $(-t_\alpha, t_\alpha)$ (см. рис. 10.5).

10.9 χ^2-распределение Пирсона

С.в. U имеет ***χ^2-распределение Пирсона*** с r степенями свободы, если ее плотность вероятностей имеет вид

$$\boxed{f(u) = \begin{cases} \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} u^{r/2-1} e^{-u/2}, & u > 0; \\ 0, & u \le 0; \end{cases}}$$

где $\Gamma(x)$ — гамма-функция (см. приложение F). Условие нормировки выполняется:

$$\int\limits_{-\infty}^{\infty} f(u)\,du = \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{\infty} u^{r/2-1} e^{-u/2}\,du = \langle u = 2z,\ du = 2\,dz\rangle =$$

$$= \frac{2^{r/2}}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_{0}^{\infty} z^{r/2-1} e^{-z}\,dz = \frac{1}{\Gamma\left(\frac{r}{2}\right)}\Gamma\left(\frac{r}{2}\right) = 1.$$

Графики плотностей вероятностей для различных r показаны на рис 10.6.

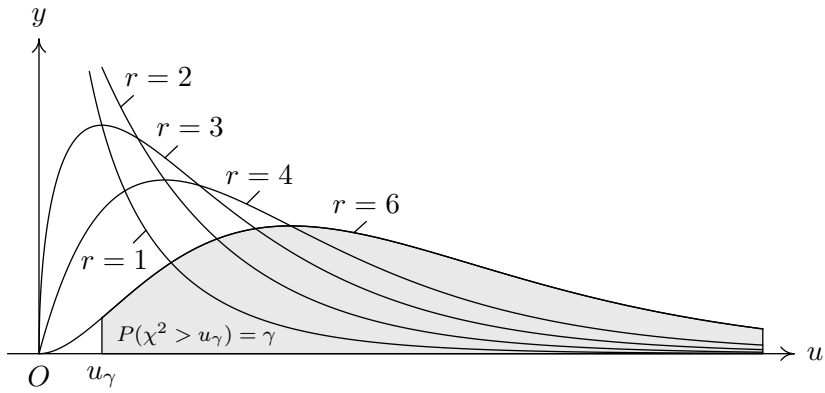


Рис. 10.6: Плотности вероятностей для χ^2-распределения

Функция распределения данной с.в. определяется формулой

$$F(u) = \begin{cases} \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^u u^{r/2-1} e^{-u/2}\, du, & u > 0; \\ 0, & u \le 0. \end{cases}$$

Найдем математическое ожидание с.в. U:

$$\mathbf{M}U = \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty u \cdot u^{r/2-1} e^{-u/2}\, du = \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty u^{r/2} e^{-u/2}\, du =$$

$$= \langle u = 2z,\ du = 2\, dz \rangle = \frac{2 \cdot 2^{r/2}}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty z^{r/2} e^{-z}\, dz =$$

$$= \frac{2}{\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty z^{(r/2+1)-1} e^{-z}\, dz = \frac{2\Gamma\left(\frac{r}{2}+1\right)}{\Gamma\left(\frac{r}{2}\right)} = \frac{2\frac{r}{2}\Gamma\left(\frac{r}{2}\right)}{\Gamma\left(\frac{r}{2}\right)} = r.$$

Вычислим математическое ожидание квадрата с.в.:

$$\begin{aligned}
\mathbf{M}U^2 &= \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty u^2 \cdot u^{r/2-1} e^{-u/2}\, du = \\
&= \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty u^{r/2+1} e^{-u/2}\, du = \langle u = 2z,\ du = 2\, dz \rangle = \\
&= \frac{4 \cdot 2^{r/2}}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty z^{r/2+1} e^{-z}\, dz = \frac{4}{\Gamma\left(\frac{r}{2}\right)} \int\limits_0^\infty z^{(r/2+2)-1} e^{-z}\, du = \\
&= \frac{4\Gamma\left(\frac{r}{2}+2\right)}{\Gamma\left(\frac{r}{2}\right)} = \frac{4\Gamma\left(\left(\frac{r}{2}+1\right)+1\right)}{\Gamma\left(\frac{r}{2}\right)} = \frac{4\left(\frac{r}{2}+1\right)\Gamma\left(\frac{r}{2}+1\right)}{\Gamma\left(\frac{r}{2}\right)} = \\
&= \frac{4\left(\frac{r}{2}+1\right)\frac{r}{2}\Gamma\left(\frac{r}{2}\right)}{\Gamma\left(\frac{r}{2}\right)} = 4\left(\frac{r}{2}+1\right)\frac{r}{2} = r\,(r+2)\,.
\end{aligned}$$

Тогда дисперсия равна

$$\mathbf{D}U = \mathbf{M}U^2 - (\mathbf{M}U)^2 = r\,(r+2) - r^2 = 2r.$$

В приложении G.4 приведены таблицы χ^2-распределения. Входами в таблицу являются число степеней свободы r и вероятность γ того, что с.в., имеющая распределение Пирсона, превзойдет число u_γ, которое и является результатом использования таблицы. Таким образом, число u_γ обеспечивает выполнение неравенства

$$P(\chi^2 > u_\gamma) = \int\limits_{u_\gamma}^\infty f(u)\, du = \gamma,$$

или, иными словами, попадание с.в. U с вероятностью γ в интервал (u_γ, ∞) (см. рис. 10.6).

Лекция 11
Предельные теоремы

11.1 Предельные теоремы для схемы Бернулли

Использование предельных теорем позволяет упростить некоторые допредельные распределения, заменив их (приближенно) более простыми законами распределений. Поэтому, помимо важного теоретического значения, они имеют большой практический смысл как эффективный способ решения прикладных задач.

Мы уже видели, что при неограниченном увеличении числа опытов в схеме Бернулли, когда вероятность p успеха в одном опыте неограниченно уменьшается, причем выполняется условие $np \to \lambda > 0$, биномиальный закон распределения трансформируется в пуассоновский. Если условие $p \to 0$ при $n \to \infty$ не выполняется, распределением Пуассона для приближенного подсчета вероятности числа успехов X в схеме Бернулли пользоваться нельзя. Однако эта вероятность может быть с достаточной точностью вычислена по формуле

$$\boxed{P(X = k) \approx \frac{1}{\sqrt{npq}}\,\varphi\left(\frac{k - np}{\sqrt{npq}}\right),} \tag{11.1}$$

где $q = 1 - p, \varphi(x)$ — плотность вероятностей с.в., распределенной $N(0, 1)$.

Приближенное равенство (11.1) составляет заключение так называемой ***локальной теоремы Муавра-Лапласа*** и может использоваться в практических расчетах при $n > 100$, $npq > 20$ и условии, что

k и n не очень сильно отличаются друг от друга. Формула (11.1) показывает, что при больших n дискретный биномиальный закон распределения может приближенно выражаться с помощью непрерывного нормального закона распределения. Это еще раз демонстрирует важность нормального закона.

Пример 11.1. *Среди 480 телефонов, имеющихся на складе, повторитель номера имеют 120 аппаратов. Найти вероятность того, что случайно отобрав из имеющихся на складе телефонов 243 телефона, мы обнаружим среди них 70 телефонов с повторителем номера.*

Решение. Случайный выбор 243 телефонов укладывается в схему Бернулли с числом опытов $n = 243$; $p = 120/480 = 0{,}25$; $q = 1 - p = 0{,}75$; $k = 70$. По формуле (11.1), учитывая, что все ограничения на ее применение выполнены, и используя таблицы функции $\varphi(x)$, находим

$$P(X = 70) \approx \frac{1}{\sqrt{243 \cdot 0,25 \cdot 0,75}} \varphi\left(\frac{70 - 243 \cdot 0,25}{\sqrt{243 \cdot 0,25 \cdot 0,75}}\right) = \\ = 0,148 \cdot \varphi(1,37) = 0,023.$$

□

Для вычисления вероятности того, что число успехов в схеме Бернулли лежит в заданном диапазоне $[k_1, k_2]$, при больших n используют ***интегральную формулу Муавра-Лапласа***:

$$\boxed{P(k_1 \le X \le k_2) \approx \Phi^*\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi^*\left(\frac{k_1 - np}{\sqrt{npq}}\right),}$$

где $\Phi^*(x)$ — функция распределения с.в., распределенной $N(0,1)$. Вместо $\Phi^*(x)$ в этом равенстве можно использовать функцию $\Phi(x)$, рассмотренную на прошлой лекции.

Пример 11.2. *Для условий предыдущего примера определить вероятность того, что среди 243 случайно отобранных телефонов от 60 до 80 из них имеют повторитель номера.*

Решение. Применим интегральную формулу Муавра-Лапласа:

$$P(k_1 \le X \le k_2) \approx \Phi^*\left(\frac{80-60{,}75}{6{,}75}\right) - \Phi^*\left(\frac{60-60{,}75}{6{,}75}\right) =$$
$$= \Phi^*(2{,}85) - \Phi^*(-0{,}11) = 0{,}99781 - (1-0{,}54380) \approx 0{,}54. \quad \square$$

11.2 Закон больших чисел и теорема Бернулли

На первой лекции говорилось о том, что непредсказуемый характер случайных событий не мешает некоторым их характеристикам проявлять удивительную устойчивость. Это значит, что в случайном мире поведение большого количества случайных величин становится не таким уж случайным, приобретая черты закономерности, а, значит, и предсказуемости. Условия, при соблюдении которых выполняется вышесказанное, отражены в теоремах, носящих название закона больших чисел.

Вначале введем в рассмотрение функцию, которая называется ***индикатором*** с.с. A и определяется следующим образом:

$$\chi(\omega, A) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Будем обозначать $\{\omega : P(\omega)\}$ множество элементов $\omega \in \Omega$, обладающих свойством $P(\omega)$. Тогда для любого действительного числа $\varepsilon > 0$ и произвольной с.в. X имеет место равенство

$$\mathbf{M}\chi(\omega, \{\omega : X(\omega) \ge \varepsilon\}) =$$
$$= 1 \cdot P(X(\omega) \ge \varepsilon) + 0 \cdot P(X(\omega) < \varepsilon) = P(X(\omega) \ge \varepsilon), \quad (11.2)$$

так как значение 1 индикатор принимает при $X \ge \varepsilon$ и, следовательно, вероятность этого значения равна $P(X \ge \varepsilon)$.

С другой стороны, для неотрицательной с.в. X справедливо соотношение

$$X(\omega) = X(\omega) \cdot \chi(\omega, \{\omega : X(\omega) \ge \varepsilon\}) + X(\omega) \cdot \chi(\{\omega, \{\omega : X(\omega) < \varepsilon\}) \ge$$
$$\ge X(\omega) \cdot \chi(\omega, \{\omega : X(\omega) \ge \varepsilon\}) \ge \varepsilon \cdot \chi(\omega, \{\omega : X(\omega) \ge \varepsilon\}).$$

Применяя к обеим частям неравенства операцию математического ожидания и используя выражение (11.2), получаем

$$\mathbf{M}X \geq \varepsilon \cdot \mathbf{M}\chi(\omega, \{\omega : X(\omega) \geq \varepsilon\}) = \varepsilon \cdot P(X \geq \varepsilon),$$

или

$$P(X \geq \varepsilon) \leq \frac{\mathbf{M}X}{\varepsilon}. \tag{11.3}$$

Это неравенство называется ***неравенством Чебышева***.

Из него следуют важные теоретические результаты. Формула (11.3) превращается в ***первое неравенство Маркова***, если вместо с.в. X взять ее модуль:

$$P(|X| \geq \varepsilon) \leq \frac{\mathbf{M}|X|}{\varepsilon}.$$

Заменяя в (11.3) X на X^2, а ε — на ε^2 и учитывая, что $P(|X| \geq \varepsilon) = P(X^2 \geq \varepsilon^2)$, получаем ***второе неравенство Маркова***:

$$P(|X| \geq \varepsilon) \leq \frac{\mathbf{M}X^2}{\varepsilon^2}.$$

Наконец, взяв в последнем соотношении разность $|X - \mathbf{M}X|$ вместо X, приходим ко ***второму неравенству Чебышева***:

$$\boxed{P(|X - \mathbf{M}X| \geq \varepsilon) \leq \frac{\mathbf{D}X}{\varepsilon^2}.}$$

Пример 11.3. *При изучении нормального распределения был сформулирован закон «трех сигм». Получить такой же закон для любого распределения.*

Решение. Используем второе неравенство Чебышева:

$$P(|X - \mathbf{M}X| < 3\sigma_X) = 1 - P(|X - \mathbf{M}X| \geq 3\sigma_X) \geq$$
$$\geq 1 - \frac{\mathbf{D}X}{9\sigma_X^2} = 1 - \frac{1}{9} = \frac{8}{9} \approx 0{,}89.$$

Напомню, что для нормального закона эта вероятность оценивается числом 0,9973. □

Пример 11.4. *Как изменить количество сигм в законе «трех сигм», чтобы оценка вероятности для любого распределения была такой же, как для нормального закона распределения:* 0,9973*?*

Решение. Нужно, чтобы величина $1 - \dfrac{\mathbf{D}X}{n^2\sigma_X^2}$ равнялась 0,9973. Из этого условия находим n:

$$1 - \frac{1}{n^2} = 0{,}9973, \quad n^2 = \frac{1}{1 - 0{,}9973} = 370{,}37;$$
$$n \approx 19{,}25, \quad n \approx 20.$$

В результате получаем закон «двадцати сигм»:

$$P(|X - \mathbf{M}X| < 20\sigma) \geq 0{,}9975.$$

Большая величина n ($n = 20$) указывает на то, что оценка вероятности с помощью неравенства Чебышева является довольно грубой (но зато пригодной для любого распределения!). □

Говорят, что последовательность с.в. $X_1, X_2, \ldots$ ***сходится по вероятности*** к с.в. X при $n \to \infty$, если

$$\forall(\varepsilon > 0) \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0,$$

и пишут $\lim_{n\to\infty} X_n \overset{P}{=} X$. Следующие далее теоремы используют это понятие и составляют содержание закона больших чисел, который устанавливает условия сходимости среднего арифметического случайных величин к среднему арифметическому их математических ожиданий.

Теорема 11.1 (Маркова)**. Закон больших чисел для зависимых с.в.**

Если $X_1, X_2, \ldots$ — последовательность с.в. и

$$\lim_{n\to\infty} \frac{1}{n^2} \mathbf{D}\left(\sum_{k=1}^{n} X_k\right) = 0, \tag{11.4}$$

то

$$\boxed{\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k \overset{P}{=} \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbf{M}X_k.} \tag{11.5}$$

Доказательство. Пусть $X = \frac{1}{n}\sum\limits_{k=1}^{n}X_k$; подставляя эту с.в. во второе неравенство Чебышева, получаем

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k - \frac{1}{n}\sum_{k=1}^{n}\mathbf{M}X_k\right| \geq \varepsilon\right) \leq \frac{\mathbf{D}\left(\frac{1}{n}\sum\limits_{k=1}^{n}X_k\right)}{\varepsilon^2} = \frac{\frac{1}{n^2}\mathbf{D}\left(\sum\limits_{k=1}^{n}X_k\right)}{\varepsilon^2}. \tag{11.6}$$

По условию теоремы правая часть неравенства при $n \to \infty$ неограниченно уменьшается, следовательно, справедлива формула (11.5). □

Таким образом, данная теорема утверждает, что практически достоверным можно считать событие, состоящее в том, что модуль отклонения среднего арифметического случайных величин от среднего арифметического их математических ожиданий будет сколь угодно малым. Следующие утверждения раскрывают смысл этого факта для все более простых систем случайных величин. Но сначала расширим понятие независимости на бесконечное множество с.в. Именно, случайные величины бесконечной последовательности X_1, X_2, …, X_n, … называются ***независимыми***, если при любом m независимы с.в. X_1, X_2, …, X_m.

Следствие 11.1. Закон больших чисел в форме Чебышева для неодинаково распределенных независимых с.в.

Если $X_1, X_2, \ldots$ — последовательность независимых с.в. и

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\mathbf{D}X_k = 0, \tag{11.7}$$

то выполняется равенство (11.5).

Доказательство. Действительно, в этом случае числитель правой части (11.6) с учетом независимости членов последовательности равен выражению под знаком предела в формуле (11.7), из чего и следует справедливость рассматриваемого следствия. □

Следствие 11.2. Закон больших чисел в форме Хинчина для одинаково распределенных независимых с.в. *Если* $X_1, X_2, \ldots$ — *последовательность независимых, одинаково распределенных с.в., причем,* $\mathbf{M}X_k = a$ $(k = 1, 2, \ldots)$, *то*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \overset{P}{=} a. \tag{11.8}$$

Доказательство. Так как члены последовательности одинаково распределены, то $\mathbf{D}X_k = d = \text{const}$ $(k = 1, 2, \ldots)$ и условие (11.7) всегда выполняется:

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \mathbf{D}X_k = \lim_{n\to\infty} \frac{1}{n^2} dn = \lim_{n\to\infty} \frac{d}{n} = 0.$$

В силу предыдущего следствия справедливо равенство (11.5), в котором правая часть принимает вид:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{M}X_k = \lim_{n\to\infty} \frac{1}{n} an = a.$$

Это и доказывает формулу (11.8). □

Пример 11.5. *Производится исследование срока эксплуатации коротковолнового передатчика до его первого отказа. Известно, что с.к.о. этого срока не превышает 10 месяцев. За каким числом передатчиков следует провести наблюдения, чтобы средняя величина срока службы передатчика до первого отказа отличалась от истинного значения не более чем на 2 месяца с вероятностью, не меньшей* 0,95*?*

Решение. Очевидно, наблюдения за эксплуатацией передатчиков можно считать независимыми с.в. Пусть X_k — фактически наблюдавшийся срок службы k-го передатчика до его первого отказа. По условию $\sigma_{X_k} \leq 10$ и, следовательно, $\mathbf{D}X_k \leq 100$. Воспользуемся законом больших чисел (11.8), который в допредельной форме имеет вид (11.6):

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a\right| \geq \varepsilon\right) \leq \frac{\frac{1}{n^2}\sum\limits_{k=1}^{n}\mathbf{D}X_k}{\varepsilon^2} \leq \frac{100}{n\varepsilon^2},$$

где $a = \mathbf{M}X_k$. Следовательно,

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a\right| < \varepsilon\right) > 1 - \frac{100}{n\varepsilon^2}.$$

Подставив в это соотношение $\varepsilon = 2$, получим

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a\right| < 2\right) > 1 - \frac{25}{n} \geq 0,95;$$

значит, неравенство выполнится, если n будет таким, что

$$\frac{25}{n} \leq 1 - 0,95 = 0,05 \Longleftrightarrow \frac{n}{25} \geq 20 \Longleftrightarrow n \geq 500.$$

Следовательно, достаточно провести 500 наблюдений. □

Неравенство (11.6) позволяет также доказать основополагающее для теории вероятностей утверждение.

Теорема 11.2 (Бернулли). *Пусть X_n — число успехов в схеме Бернулли с n опытами и p — вероятность успеха в одном опыте. Тогда относительная частота успехов X_n/n сходится по вероятности к p:*

$$\boxed{\lim_{n\to\infty}\frac{X_n}{n} \overset{P}{=} p.} \tag{11.9}$$

Доказательство. Пусть Y_k — с.в., имеющая распределение Бернулли, тогда $\mathbf{M}Y_k = p, \mathbf{D}Y_k = pq, q = 1 - p, X_n = \sum_{k=1}^{n} Y_k$. Подставляя в формулу (11.6) вместо X_k с.в. Y_k, получим

$$P\left(\left|\frac{1}{n}X_n - \frac{1}{n}pn\right| \geq \varepsilon\right) \leq \frac{\sum\limits_{k=1}^{n} \mathbf{D}Y_k}{n^2\varepsilon^2} = \frac{npq}{n^2\varepsilon^2} = \frac{pq}{n\varepsilon^2},$$

или

$$P\left(\left|\frac{X_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{pq}{n\varepsilon^2}.$$

При $n \to \infty$ правая часть последнего неравенства стремится к нулю, а это означает справедливость равенства (11.9). □

Выполнение утверждения теоремы Бернулли подтверждено экспериментально и позволяет при больших n из наблюдений за относительными частотами с.с. делать выводы об их вероятностях. В частности, такие выводы послужили основой для построения системы аксиом теории вероятностей, рассмотренной на первой лекции по этому разделу математики.

11.3 Центральная предельная теорема

Данная теорема дает возможность считать с.в. X распределенной приближенно по нормальному закону, если она является суммой независимых с.в., имеющих равномерно малое влияние на сумму. Так как такие с.в. (ошибка измерения, отклонение от цели и т.п.) часто встречаются на практике, эта теорема объясняет широкое распространение нормального закона в различных приложениях.

Теорема 11.3. (центральная предельная теорема для одинаково распределенных слагаемых)

Если с.в. $X_1, X_2, \ldots$ независимы, одинаково распределены и имеют конечную дисперсию, то при $n \to \infty$ закон распределения с.в. $X = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$ неограниченно приближается к нормальному закону $N(n\mathbf{M}X, n\mathbf{D}X)$.

Теорема 11.4. (центральная предельная теорема в формулировке Ляпунова) *Пусть с.в. $X_1, X_2, \ldots$ независимы и имеют конечный третий абсолютный момент:*

$$b_k = \mathbf{M}(|X_k - \mathbf{M}X_k|^3) < \infty.$$

Если

$$\lim_{n\to\infty} \frac{\sum\limits_{k=1}^{n} b_k}{\left(\sum\limits_{k=1}^{n} \mathbf{D}X_k\right)^{3/2}} = 0, \tag{11.10}$$

то при $n \to \infty$ закон распределения с.в. $X = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$ неограниченно приближается к нормальному.

□

Условие (11.10) означает, что ни одно слагаемое в сумме X_1+X_2+ $+X_3+\ldots$ не доминирует над остальными, то есть они все имеют равномерно малое влияние на сумму.

Пример 11.6. *За смену производится 100 микропроцессоров одного типа. Вероятность обнаружения дефекта при проверке одной микросхемы равна 0,1. Определить границы, в которых с вероятностью 0,9 будет заключено число микросхем с дефектами, обнаруженными при контроле всей сменной продукции.*

Решение. Обозначим X_i — с.в., принимающую значение 1, если в микросхеме обнаружен дефект; и значение 0 в противном случае. Ее ряд распределения имеет вид

x_i	0	1
p_i	0,9	0,1

и, значит,

$$\mathbf{M}X_i = 0\cdot 0{,}9 + 1\cdot 0{,}1 = 0{,}1; \quad \mathbf{M}X_i^2 = 0^2\cdot 0{,}9 + 1^2\cdot 0{,}1 = 0{,}1;$$
$$\mathbf{D}X_i = \mathbf{M}X_i^2 - (\mathbf{M}X_i)^2 = 0{,}1 - 0{,}1^2 = 0{,}1 - 0{,}01 = 0{,}09.$$

Тогда $X = \sum\limits_{i=1}^{100} X_i$ — число микросхем с дефектами, обнаруженными при контроле, причем,

$$\mathbf{M}X = \sum_{i=1}^{100} \mathbf{M}X_i = 0{,}1 \cdot 100 = 10;$$

и так как с.в. X_i независимы, то

$$\mathbf{D}X = \sum_{i=1}^{100} \mathbf{D}X_i = 0{,}09 \cdot 100 = 9; \quad \sigma_X = 3.$$

В силу центральной предельной теоремы с.в. X имеет приближенно нормальное распределение и поэтому

$$P(|X - \mathbf{M}X| < \alpha) = P(\mathbf{M}X - \alpha < X < \mathbf{M}X + \alpha) \approx$$
$$\approx \Phi^*\left(\frac{\alpha}{\sigma_X}\right) - \Phi^*\left(-\frac{\alpha}{\sigma_X}\right) = 2\Phi^*\left(\frac{\alpha}{\sigma_X}\right) - 1 = 0{,}9;$$

или

$$\Phi^*\left(\frac{\alpha}{\sigma_X}\right) = \frac{1{,}9}{2} = 0{,}95.$$

По таблицам функции $\Phi^*(x)$ находим, что это равенство выполняется при $\alpha/\sigma_X = 1{,}645$; значит,

$$\alpha = 1{,}645 \cdot \sigma_X = 1{,}645 \cdot 3 = 4{,}935.$$

Таким образом, с вероятностью 0,9 число микросхем с дефектами, обнаруженными при контроле, будет заключено в диапазоне от $\mathbf{M}X - \alpha = 10 - 4{,}935 = 5{,}065 \approx 5$ до $\mathbf{M}X + \alpha = 10 + 4{,}935 = 14{,}935 \approx$ ≈ 15. $\square$

Часть II

Случайные процессы

Лекция 12
Некоторые типы с.п.

12.1 Основные определения

Так же, как поведение случайных величин носит непредсказуемый характер, так и поведение развернутых во времени процессов может быть достаточно хаотичным, плохо прогнозируемым. Типичным изображением такого процесса является график температуры воздуха за год. К таким же процессам относятся сигналы, поступающие в каналы связи и проходящие по ним, радиосигналы, шумы аппаратуры и др. В теории вероятностей разработаны специальные методы, помогающие анализировать случайные процессы, находить их характеристики, в том числе вероятностные связи между процессами. Решается и обратная задача: по характеристикам случайных процессов на входе и выходе аппаратуры определяются ее оптимальные параметры. Перейдем к точным формулировкам.

Случайным процессом (с.п.) называется функция $X(t,\omega)$, $\omega \in \Omega, t$ — параметр, который чаще всего интерпретируется как время. Осуществление случайного процесса можно представлять себе как выбор э.с.с. ω из пространства э.с.с. Ω. Например, выбирая конкретный год, мы получаем график температуры воздуха за этот год. Функция $X(t,\omega_0)$ при фиксированном значении $\omega = \omega_0$ называется ***траекторией*** или ***реализацией*** с.п. $X(t,\omega)$. В дальнейшем для простоты обозначений, как и при изучении случайных величин, зависимость с.п. от ω будем опускать и писать просто $X(t)$.

А что мы получим, если зафиксируем теперь аргумент $t = t_0$ в

определении случайного процесса? Для обычной функции это было бы ее значением в точке t_0. Случайный процесс тоже превращается в свое значение, однако поскольку оно непредсказуемо, то нужно считать, что значением с.п. в момент времени t_0 является случайная величина $X(t_0)$, которая называется ***сечением*** с.п. На рис. 12.1 показаны две траектории с.п. $X(t, \omega)$, соответствующие э.с.с. $\omega = \omega_1$ и $\omega = \omega_2$ и две реализации его сечения при $t = t_0$.

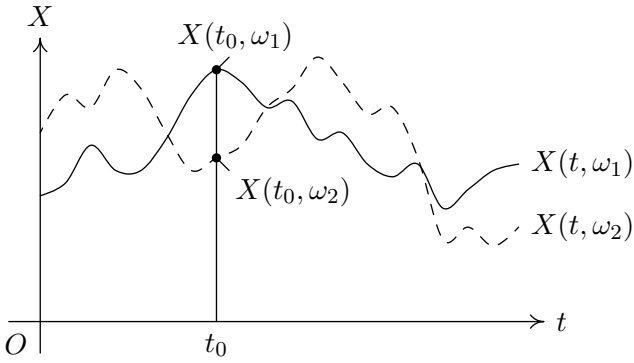


Рис. 12.1: Траектории и сечение с.п. при $t = t_0$

Далее по аналогии со случайными величинами и случайными векторами нужно, казалось бы, ввести ***функцию распределения случайного процесса*** и использовать ее для получения различных его характеристик. Действительно, такую функцию можно задать формулой

$$F_{t_1 \ldots t_n}(x_1, \ldots, x_n) = P(X(t_1) < x_1, \ldots, X(t_n) < x_n),$$

однако ее практическое применение затруднено тем, что она, как вы видите, представляет собой не одну функцию, а бесконечное их семейство, зависящее от неопределенного числа $2n$ аргументов $t_1, \ldots, t_n$, $x_1, \ldots, x_n$. Поэтому она используется в основном в теоретических исследованиях, а нам придется обойтись без нее.

Мы будем рассматривать два основных типа процессов: ***с непрерывным временем***, если $t \in R^+ = [0; \infty)$, и с ***дискретным временем***, если t принимает лишь значения $0, 1, 2, \ldots$. Важными характеристиками случайного процесса являются следующие детерминированные функции.

Математическим ожиданием с.п. $X(t)$ называется функция $m(t) = \mathbf{M}X(t)$, для каждого t равная математическому ожиданию сечения с.п. для этого момента времени.

Ковариационной функцией с.п. $X(t)$ называется функция $K(t_1, t_2) = \mathbf{M}[X(t_1) - m(t_1)][X(t_2) - m(t_2)]$, для каждой пары (t_1, t_2) равная ковариации сечений с.п. для этих моментов времени.

Нормируя ковариационную функцию, получаем ***корреляционную функцию***:

$$r_X(t_1, t_2) = \frac{K_X(t_1, t_2)}{\sigma_X(t_1)\sigma_X(t_2)}.$$

Дисперсией с.п. $X(t)$ называется функция $D(t) = K(t, t) = \mathbf{M}[X(t) - m(t)]^2$, для каждого t равная дисперсии сечения с.п. для этого момента времени.

Первыми случайными процессами, с которыми вы познакомитесь, будут процессы с независимыми приращениями и марковские цепи, играющие важную роль в теории связи, статистической радиотехнике и других родственных вашей специальности областях.

12.2 Процессы с независимыми приращениями

С.п. $X(t)$ с непрерывным временем называется процессом ***с независимыми приращениями***, если $t \in [0; \infty)$ и для любых моментов времени t_0, $t_1, \ldots$, t_n таких, что $0 \le t_0 < t_1 < \cdots < t_n$ случайные величины $X(t_0)$, $X(t_1) - X(t_0)$, $\ldots$, $X(t_n) - X(t_{n-1})$ независимы.

Процесс $X(t)$ с независимыми приращениями называется ***пуассоновским***, если 1) $X(0) = 0$, 2) с.в. $X(t) - X(s)$, $0 \le s \le t$, имеет распределение Пуассона, то есть

$$P[X(t) - X(s) = i] = \frac{[\lambda(t-s)]^i}{i!} e^{-\lambda(t-s)};\ i = 0, 1, 2, \ldots.$$

Теорема 12.1. *Пусть все траектории пуассоновского процесса непрерывны справа. Тогда почти все траектории (за исключением счетного множества) — неубывающие целочисленные функции, возрастающие только скачками величины 1 (см. рис. 12.2).*

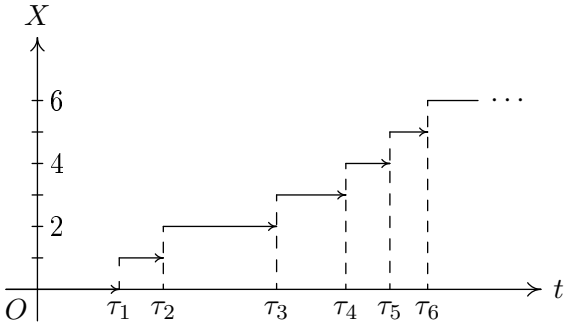


Рис. 12.2: Траектория пуассоновского процесса

Теорема 12.2. *Если $X(t)$ — пуассоновский процесс, то*

$$P[X(t)=k]=\frac{(\lambda t)^k}{k!}e^{-\lambda t};\ k=0,1,2,\ldots;$$
$$m(t)=\lambda t;\ \mathbf{D}(t)=\lambda t;\ K(t,t')=\lambda\min(t,t');$$

и при любом n независимы и имеют показательное распределение с параметром λ случайные величины длин промежутков между скачками процесса $X(t)$: τ_1, $\tau_2-\tau_1,\ldots,\tau_n-\tau_{n-1},\ldots$.

Замечание 12.1. *Величину λ трактуют как среднее число скачков процесса $X(t)$ в единицу времени.*

Пример 12.1. *Среднее число вызовов, поступающих на АТС за одну минуту, равно двум. Найти вероятность того, что за 4 минуты поступит а) 3 вызова, б) менее трех вызовов, в) не менее трех вызовов. Найти также среднее число вызовов за время t и среднее время между вызовами, предполагая, что поток вызовов — пуассоновский процесс.*

Решение. Вероятность того, что за 4 минуты поступит 3 вызова (с учетом того, что $\lambda = 2$), равна

$$P\left[X\left(4\right)=3\right]=\frac{\left(2\cdot 4\right)^{3}}{3!}\,e^{-2\cdot 4}=\frac{256}{3}\,e^{-8}\approx 0{,}029,$$

где $X\left(t\right)$ — количество вызовов, поступающих на АТС за время t. Событие «Поступит менее трех вызовов» произойдет, если наступит одно из трех несовместных событий: 1) поступит два вызова, 2) поступит один вызов, 3) не поступит ни одного вызова. Применяя теорему сложения для несовместных событий, получим

$$P\left[X\left(4\right)<3\right]=P\left[X\left(4\right)=2\right]+P\left[X\left(4\right)=1\right]+P\left[X\left(4\right)=0\right]=$$
$$=\frac{\left(2\cdot 4\right)^{2}}{2!}\,e^{-8}+\frac{2\cdot 4}{1!}\,e^{-8}+e^{-8}\approx 0{,}014.$$

Вероятность с.с. «Поступит не менее трех вызовов» равна

$$P\left[X\left(4\right)\geq 3\right]=1-P\left[X\left(4\right)<3\right]\approx 1-0{,}014=0{,}986.$$

Среднее число вызовов за время t:

$$m\left(t\right)=\lambda t=2t;$$

среднее время между вызовами равно $1/\lambda = 0{,}5$ (мин.), так как времена между вызовами имеют показательное распределение с параметром λ, для которого математическим ожиданием является величина $1/\lambda$. □

12.3 Марковские цепи. Переходные вероятности

С одной стороны, марковская цепь обобщает схему Бернулли на случай, когда опыты уже не являются независимыми, однако, зависимость между ними еще довольно слабая — каждый опыт зависит от предыдущего, но не зависит от других предшествующих ему опытов (как кольцо в цепи, отсюда и название). Поэтому говорят, что *будущее марковского процесса зависит от настоящего, но не зависит*

от прошлого. Например, частота отказов прибора, длительность его ремонта в будущем зависят в основном от его состояния в текущий момент времени, а не от того, каким образом прибор пришел в свое нынешнее состояние. Точно так же, каким будет счет футбольного мачта после следующего гола, зависит лишь от текущего счета, а не от того, как команды его достигли.

Точное определение таково. С.п. $X(t)$ называется ***марковским***, или с.п. ***без последействия***, если

$$P(X(t) < x/X(t_1) < x_1, \ldots, X(t_n) < x_n) = P(X(t) < x)/X(t_n) < x_n)$$

для всех $t_1 < \ldots < t_n < t$. Таким образом, вероятностные характеристики марковского процесса в «будущий» момент времени t зависят только от его состояния в «настоящий момент времени» t_n и не зависят от того, когда и как процесс пришел в это состояние (т.е. не зависят от его состояний в «прошлые» моменты времени $t_1, \ldots, t_{n-1}$).

Мы не станем рассматривать самый общий марковский процесс. Прежде всего предположим, что он может принимать лишь конечное число значений $A_1, \ldots, A_N$ из некоторого множества, быть может, и не числового. Эти значения называются ***состояниями*** марковского процесса $X(t)$, например, A_3 будем называть третьим состоянием, а A_N — N-м. Если при этом марковский процесс $X(t)$ является с.п. с дискретным временем, то он называется ***конечной цепью Маркова с дискретным временем***, а если $X(t)$ — с.п. с непрерывным временем, то — ***конечной цепью Маркова с непрерывным временем***.

Очень важную роль в изучении таких процессов играют вероятности

$$\boxed{P(X(t) = A_j/X(s) = A_i) = P_{ij}(s,t),}$$

называемые ***вероятностями перехода*** марковской цепи из состояния A_i в состояние A_j, или ***переходными вероятностями***. Это — условные вероятности того, что марковская цепь в момент времени t окажется в j-м состоянии, если в момент времени s она была в i-м состоянии. Сразу отметим, что «переход», в частности,

может заключаться в том, что марковская цепь в момент времени t останется в том же состоянии, в каком она пребывала в момент времени s, то есть вероятность $P(X(t) = A_i/X(s) = A_i) = P_{ii}(s,t)$ имеет смысл и не обязательно равна нулю.

Далее мы еще сузим круг марковских цепей, ограничившись ***однородными*** цепями Маркова, переходные вероятности которых зависят не от самих моментов времени t и s, а только от их разности $t - s$:

$$P_{ij}(s,t) = P_{ij}(t-s) = P_{ij}(\tau),$$

где $\tau > 0$ — интервал времени $t-s$. Для цепи Маркова с дискретным временем $P_{ij}(n)$ называется вероятностью перехода из i-го состояния в j-е ***за*** n ***шагов***. На самом деле все эти вероятности не являются совершенно независимыми; достаточно знать вероятности перехода за один шаг, чтобы получить вероятности перехода за большее число шагов. Мы убедимся в этом, доказав следующее утверждение.

Теорема 12.3. *Вероятности $P_{ij}(\tau)$ удовлетворяют уравнению Маркова-Колмогорова:*

$$\boxed{P_{ij}(\tau) = \sum_{k=1}^{N} P_{ik}(\nu) P_{kj}(\tau - \nu).} \tag{12.1}$$

Доказательство. Обозначим $A_i \xrightarrow{\tau} A_j$ с.с., заключающееся в переходе марковской цепи из состояния A_i в состояние A_j за время τ. При фиксированном i и $j = \overline{1,n}$ эти события образуют разбиение, так как из i-го состояния цепь Маркова обязательно перейдет в какое-либо и (только одно!) состояние, причем, как было отмечено, «переходом» считается и тот случай, когда цепь остается в состоянии A_i. Поэтому по формуле полной вероятности получаем

$$P(A_i \xrightarrow{\tau} A_j) = \sum_{k=1}^{N} P(A_k \xrightarrow{\tau-\nu} A_j / A_i \xrightarrow{\nu} A_k) P(A_i \xrightarrow{\nu} A_k).$$

Но в силу того, что процесс — марковский, условная вероятность, фигурирующая в этой формуле, равна безусловной. Следовательно,

$$P(A_i \xrightarrow{\tau} A_j) = \sum_{k=1}^{N} P(A_k \xrightarrow{\tau-\nu} A_j) P(A_i \xrightarrow{\nu} A_k).$$

А это и есть формула (12.1). □

Для марковских цепей с дискретным временем равенство (12.1) принимает вид

$$P_{ij}(n) = \sum_{k=1}^{N} P_{ik}(m) P_{kj}(n-m),$$

или в матричной форме:

$$\boxed{\mathbf{P}(n) = \mathbf{P}(m)\mathbf{P}(n-m),} \tag{12.2}$$

где $\mathbf{P}(n) = \{P_{ij}(n)\}$ — ***матрица вероятностей переходов за*** n ***шагов***.

Следствие 12.1. *Для марковской цепи с дискретным временем*

$$\boxed{\mathbf{P}(n) = \mathbf{P}^n,}$$

где $\mathbf{P} \triangleq \mathbf{P}(1)$.

Доказательство. Из равенства (12.2) при $n = 2, m = 1$ получаем $\mathbf{P}(2) = \mathbf{P}(1)\mathbf{P}(1) = \mathbf{P}^2$, при $n = 3, m = 2$ имеем $\mathbf{P}(3) = \mathbf{P}(2)\mathbf{P}(1) = \mathbf{P}^2\mathbf{P} = \mathbf{P}^3$. И вообще при $m < n$ выходит, что $\mathbf{P}(n) = \mathbf{P}(m)\mathbf{P}(n-m) = \mathbf{P}^m\mathbf{P}^{n-m} = \mathbf{P}^n$. □

Итак, мы видим, что вероятности переходов для марковской цепи с дискретным временем полностью определяются элементами матрицы $\mathbf{P} = \{P_{ij}\}$, причем, сумма элементов каждой строки матрицы равна единице:

$$\sum_{j=1}^{N} P_{ij} = 1.$$

Этот факт сразу следует из второй аксиомы для вероятностей с.с., если заметить, что переход цепи Маркова из i-го состояния в какое-либо другое является достоверным событием.

Возможные переходы в марковской цепи из состояния в состояние и их вероятности изображают с помощью ***графа состояний***. Он представляет собой набор ***вершин*** в виде квадратиков на плоскости и соединяющих их ***дуг*** в виде стрелок произвольной формы (рис. 12.3). Расположение вершин на плоскости тоже произвольно. Вершинам соответствуют состояния марковской цепи, а дуга, соединяющая две вершины, показывает переход из состояния, которому отвечает одна из этих вершин, в состояние, которым помечена другая вершина и на которую указывает стрелка. Над дугой проставляется вероятность такого перехода. Правда, если эта вероятность равна нулю, то дуга вообще не рисуется.

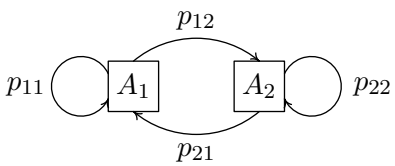


Рис. 12.3: Граф цепи Маркова с двумя состояниями

Одной из основных задач, решаемых в теории марковских цепей, является определение вероятности, с которой цепь Маркова находится в заданный момент времени в том или ином состоянии. Для решения этой задачи введем вероятность $p_i(t) = P(X(t) = A_i)$ ***пребывания марковской цепи в состоянии*** A_i. Эти вероятности образуют вектор (вектор-строку!) $\mathbf{p}(t) = (p_1(t), \ldots, p_N(t))$, который мы будем называть ***вектором вероятностей состояний***, а вектор $\mathbf{p}(0)$ — ***вектором начальных вероятностей***.

Теорема 12.4. *Для марковской цепи с дискретным временем*

$$\boxed{\mathbf{p}(n) = \mathbf{p}(0)\mathbf{P}(n) = \mathbf{p}(0)\mathbf{P}^n.} \tag{12.3}$$

Доказательство. По формуле полной вероятности

$$P(X(n) = A_j) = \sum_{i=1}^{N} P(A_i \xrightarrow{n} A_j / X(0) = A_i) P(X(0) = A_i).$$

В силу марковости процесса фигурирующая здесь условная вероятность равна безусловной, и мы получаем

$$P(X(n) = A_j) = \sum_{i=1}^{N} P(X(0) = A_i) P(A_i \xrightarrow{n} A_j),$$

или

$$P_j(n) = \sum_{i=1}^{N} p_i(0) P_{ij}(n),$$

что́ равносильно равенству (12.3). □

Достоверность с.с., заключающегося в том, что цепь Маркова находится в одном из своих состояний, приводит к равенству

$$\sum_{i=i}^{N} p_i(n) = 1.$$

Пример 12.2. *Аппаратура может находиться в одном из трех состояний: A_1 = «Исправна», A_2 = «Барахлит», A_3 = «В ремонте». Матрица вероятностей переходов за один шаг имеет вид:*

$$\mathbf{P} = \begin{pmatrix} 0{,}7 & 0{,}3 & 0 \\ 0{,}1 & 0{,}4 & 0{,}5 \\ 0{,}7 & 0{,}2 & 0{,}1 \end{pmatrix}.$$

Найти вектор вероятностей состояний через два шага, если вектор начальных вероятностей равен

$$\mathbf{p}(0) = (0{,}8; 0{,}1; 0{,}1).$$

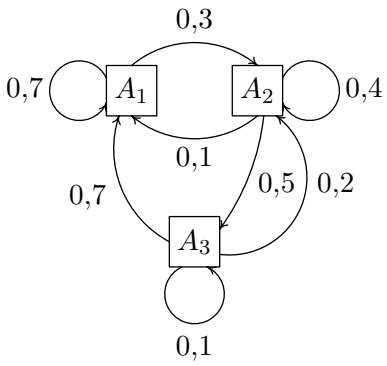


Рис. 12.4: Граф состояний для примера 12.2

Решение. По формуле (12.3) получаем

$$\mathbf{p}(2) = \mathbf{p}(0)\mathbf{P}^2.$$

Но

$$\mathbf{P}^2 = \mathbf{P} \cdot \mathbf{P} =$$
$$= \begin{pmatrix} 0{,}7 & 0{,}3 & 0 \\ 0{,}1 & 0{,}4 & 0{,}5 \\ 0{,}7 & 0{,}2 & 0{,}1 \end{pmatrix} \begin{pmatrix} 0{,}7 & 0{,}3 & 0 \\ 0{,}1 & 0{,}4 & 0{,}5 \\ 0{,}7 & 0{,}2 & 0{,}1 \end{pmatrix} = \begin{pmatrix} 0{,}52 & 0{,}33 & 0{,}15 \\ 0{,}46 & 0{,}29 & 0{,}25 \\ 0{,}58 & 0{,}31 & 0{,}11 \end{pmatrix}$$

и

$$\mathbf{p}(2) = (0{,}8; 0{,}1; 0{,}1) \begin{pmatrix} 0{,}52 & 0{,}33 & 0{,}15 \\ 0{,}46 & 0{,}29 & 0{,}25 \\ 0{,}58 & 0{,}31 & 0{,}11 \end{pmatrix} = (0{,}520; 0{,}324; 0{,}156).$$

□

Теорема 12.5. *Для марковских цепей с непрерывным временем*

$$\boxed{\mathbf{p}(t+\tau) = \mathbf{p}(t)\mathbf{P}(\tau),} \tag{12.4}$$

где $\mathbf{P}(\tau) = \{P_{ij}(\tau)\}$.

Доказательство. Это следует из равенства

$$P(X(t+\tau) = A_i) = \sum_{j=1}^{N} P(X(t+\tau) = A_i / X(t) = A_j) P(X(t) = A_j),$$

которое в силу однородности процесса приводит к формуле

$$p_i(t+\tau) = \sum_{j=1}^{N} p_j(t) P_{ji}(\tau),$$

равносильной доказываемой. □

Теорема 12.6. *Для марковских цепей с непрерывным временем вектор вероятностей состояний* $\mathbf{p}(t)$ *удовлетворяет векторно-матричному уравнению Колмогорова*

$$\boxed{\mathbf{p}'(t) = \mathbf{p}(t)\Lambda,} \tag{12.5}$$

где

$$\Lambda = \lim_{\tau \to 0+0} \frac{\mathbf{P}(\tau) - E}{\tau},$$

E *— единичная матрица.*

Доказательство. Из равенства (12.4) имеем

$$\mathbf{p}(t+\tau) - \mathbf{p}(t) = \mathbf{p}(t)[\mathbf{P}(\tau) - E].$$

Деля на τ и устремляя τ к нулю, получаем уравнение Колмогорова. □

Матрица Λ называется ***матрицей интенсивностей переходов***. Ее элемент λ_{ij} при $i \neq j$ называется ***интенсивностью перехода из состояния*** A_i ***в состояние*** A_j: $P_{ij}(\tau) = \lambda_{ij}\tau + o(\tau)$. Элемент λ_{ii} матрицы Λ называется ***интенсивностью ухода из состояния*** A_i: $1 - P_{ii}(\tau) = -\lambda_{ii}\tau + o(\tau)$.

СВОЙСТВА ИНТЕНСИВНОСТЕЙ

1°. *Интенсивности переходов неотрицательны:* $\lambda_{ij} \geq 0$ $(i \neq j)$.

Действительно, $\lambda_{ij} \approx P_{ij}(\tau)/\tau$ и $\tau > 0$.

2°. *Интенсивности уходов неположительны:* $\lambda_{ii} \leq 0$.

Это следует из того, что, $\lambda_{ii} \approx \dfrac{P_{ii}(\tau) - 1}{\tau}$ и $\tau > 0$.

3°. *Интенсивности* λ_{ij} *конечны.*

4°. *Сумма переходов и ухода из состояния* A_i *равна нулю:*

$$\boxed{\sum_{j=1}^{N} \lambda_{ij} = 0.}$$

Так как

$$\sum_{i \neq j} P_{ij}(\tau) + P_{ii}(\tau) = 1,$$

то

$$\sum_{j=1}^{N} \lambda_{ij} \approx \frac{1}{\tau} \left[\sum_{i \neq j} P_{ij}(\tau) + P_{ii}(\tau) - 1 \right] = 0.$$

□

Для наглядности марковскую цепь с непрерывным временем изображают в виде ***графа интенсивностей переходов***, который отличается от графа состояний тем, что вместо вероятностей переходов используются интенсивности переходов и граф не имеет ***петель*** (дуг, входящих в ту же вершину, из которой они выходят), то есть интенсивности ухода не обозначаются.

Для дальнейшего изложения еще сузим круг рассматриваемых цепей Маркова условием ***ординарности***, которое заключается в

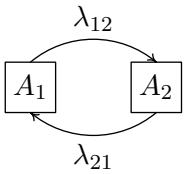


Рис. 12.5: Граф интенсивностей переходов для двух состояний

том, что марковская цепь в один и тот же момент времени не может изменить свое состояние более чем один раз.

Лекция 13
Поведение марковской цепи на бесконечности

13.1 Процесс размножения и гибели

Очень важную роль в приложениях марковских цепей играет ***процесс размножения и гибели***. И хотя этот процесс является моделью существования некоей биологической популяции, он применяется для решения широкого круга технических задач, в которых под «рождением особи», например, можно понимать возникновение заказа на ремонт какого-либо устройства, а под «гибелью» — успешное завершение выполнения заказа.

Итак, пусть имеется биологическая популяция, количество особей (объем популяции) которой никогда не превосходит числа N. Будем говорить, что популяция находится в состоянии A_i, если ее объем равен i. Предположим, что из состояния A_i популяция может перейти только либо в состояние A_{i+1} (рождение особи), либо в состояние A_{i-1} (ее гибель). Напомню, что процесс мы считаем ординарным. Обозначим через λ_i ***интенсивность размножения*** популяции (скорость, с которой происходит размножение в популяции объема i), а через μ_j — ***интенсивность гибели***. Граф интенсивностей переходов для данной марковской цепи показан на рис. 13.1,

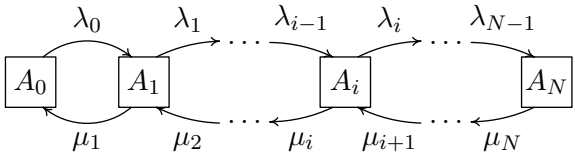


Рис. 13.1: Граф процесса размножения и гибели

а матрица интенсивностей переходов имеет вид

$$\Lambda = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots & 0 \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\mu_N \end{pmatrix}.$$

По этой матрице нетрудно составить систему дифференциальных уравнений Колмогорова (12.5).

Пример 13.1. *Осциллограф может находиться в одном из двух состояний: A_0 = «Исправен» и A_1 =«Ремонтируется», причем, интенсивность перехода из состояния A_0 в состояние A_1 равна $\lambda > 0$, а интенсивность обратного перехода равна $\mu > 0$. Считая процесс отказов осциллографа и его восстановлений процессом размножения и гибели, для вектора начальных вероятностей $\mathbf{p}(0) = (1;0)$ найти решение системы дифференциальных уравнений Колмогорова.*

Решение. Граф интенсивностей переходов данной марковской цепи с непрерывным временем и конечным числом состояний имеет вид, изображенный на рис. 13.2.

Очевидно, объем популяции в данном случае равен $N = 1$ и поэтому, так как

$$\Lambda = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix},$$

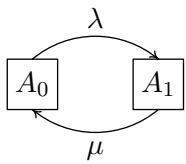


Рис. 13.2: К задаче об осциллографе

уравнение (12.5) принимает вид

$$\mathbf{p}'(t) = \mathbf{p}\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}, \tag{13.1}$$

где $\mathbf{p}(t) = (p_0(t), p_1(t))$, или в координатной форме

$$\begin{cases} p_0' = -\lambda p_0 + \mu p_1, \\ p_1' = \lambda p_0 - \mu p_1. \end{cases}$$

Это — нормальная линейная однородная система дифференциальных уравнений с постоянными коэффициентами. Чтобы получить ее решение, найдем собственные числа матрицы системы

$$\mathbf{A} = \Lambda^T = \begin{pmatrix} -\lambda & \mu \\ \lambda & -\mu \end{pmatrix},$$

решая для нее характеристическое уравнение ($\mathbf{E}$ — единичная матрица)

$$|\mathbf{A} - \kappa\mathbf{E}| = \begin{vmatrix} -\lambda - \kappa & \mu \\ \lambda & -\mu - \kappa \end{vmatrix} = \lambda\mu + (\lambda + \mu)\kappa + \kappa^2 - \lambda\mu =$$
$$= \kappa^2 + (\lambda + \mu)\kappa = \kappa(\kappa + (\lambda + \mu)) = 0.$$

Его корни: $\kappa_1 = 0,\ \kappa_2 = -(\lambda + \mu)$. Соответствующие им собственные векторы $\mathbf{s}_1$ и $\mathbf{s}_1$ матрицы системы находим, решая системы уравнений вида

$$(\mathbf{A} - \kappa_i\mathbf{E})\mathbf{s}_i, \quad i = 1, 2.$$

Пусть, например, $\mathbf{s}_1^T = (u_1, u_2)$. Тогда получаем следующую систему уравнений, соответствующую κ_1:

$$\begin{cases} -\lambda u_1 + \mu u_2 = 0, \\ \lambda u_1 - \mu u_2 = 0. \end{cases}$$

Второе уравнение можно отбросить, а первое будет выражать решение системы. Поэтому можно взять частное решение $u_1 = \mu$, $u_2 = \lambda$ и составить из него требуемый собственный вектор $\mathbf{s_1}^T = (\mu, \lambda)$.

Пусть $\mathbf{s}_2^T = (v_1, v_2)$. Система уравнений, соответствующая корню κ_2, имеет вид

$$\begin{cases} \mu v_1 + \mu v_2 = 0, \\ \lambda v_1 + \lambda v_2 = 0. \end{cases}$$

Следовательно, $v_1 = -v_2$ и можно взять $\mathbf{s}_2^T = (-1; 1)$.

Составляем общее решение уравнения (13.1):

$$\mathbf{p}^T(t) = C_1 \mathbf{s}_1 e^{\kappa_1 t} + C_2 \mathbf{s}_2 e^{\kappa_2 t} = C_1 \begin{pmatrix} \mu \\ \lambda \end{pmatrix} + C_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} e^{-(\lambda+\mu)t},$$

которое в компонентной форме имеет вид

$$p_0(t) = C_1 \mu - C_2 e^{-(\lambda+\mu)t},$$
$$p_1(t) = C_1 \lambda + C_2 e^{-(\lambda+\mu)t}.$$

Теперь учтем вектор начальных вероятностей:

$$p_0(0) = C_1 \mu - C_2 = 1,$$
$$p_1(0) = C_1 \lambda + C_2 = 0.$$

Решая эту систему уравнений, получаем

$$C_1 = \frac{1}{\lambda + \mu}, \; C_2 = -\frac{\lambda}{\lambda + \mu}.$$

Таким образом, решением задачи Коши являются функции

$$p_0(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t},$$
$$p_1(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}.$$

□

13.2 Стационарные вероятности

На практике любой динамический процесс в начале своего развития имеет фазу переходного режима, а затем — фазу устойчивой, стационарной эволюции. Это относится и к марковским процессам, но фаза их стационарности характеризуется в силу их специфической вероятностной природы стационарностью вектора вероятностей состояний. Вообще, вектор $\mathbf{p}(t)$ называется ***стационарным***, или ***стационарным распределением вероятностей***, если он не меняется с течением времени, то есть

$$\mathbf{p}(t)) \equiv \mathbf{p} = (p_1, \ldots, p_N), \ \sum_{i=1}^{N} p_i = 1.$$

Теорема 13.1. *Для цепи Маркова с дискретным временем стационарное распределение вероятностей определяется системой уравнений*

$$\boxed{\mathbf{p} = \mathbf{p}\mathbf{P}, \ \sum_{i=1}^{N} p_i = 1,} \tag{13.2}$$

а для цепи Маркова с непрерывным временем — системой

$$\boxed{\mathbf{p}\Lambda = 0, \ \sum_{i=1}^{N} p_i = 1.} \tag{13.3}$$

Доказательство. Из формулы (12.3) при $n = 1$ и $\mathbf{p}(1) = \mathbf{p}(0) \equiv \mathbf{p}$ получаем первое доказываемое равенство, а из формулы (12.5) при $\mathbf{p}'(t) \equiv 0,\ \mathbf{p}(t) \equiv \mathbf{p}$ — второе. □

Сравнивая равенства (13.2) и (13.3) с равенствами (12.3) и (12.5) соответственно, приходим к выводу, что при реализации в марковской цепи стационарного распределения вероятностей, каким бы образом цепь ни переходила из состояния в состояние, вероятности этих состояний уже измениться не могут.

Пример 13.2. *Найти стационарное распределение вероятностей для примера 12.2.*

Решение. Запишем систему уравнений (13.2):

$$\begin{cases} -0{,}3p_1 + 0{,}1p_2 + 0{,}7p_3 & = & 0, \\ 0{,}3p_1 - 0{,}6p_2 + 0{,}2p_3 & = & 0, \\ 0{,}5p_2 - 0{,}9p_3 & = & 0, \\ p_1 + p_2 + p_3 & = & 1. \end{cases}$$

Если первые три уравнения сложить, то получится тождество, значит, эти уравнения линейно зависимы. Отбросим любое из них, например, третье и получим новую систему

$$\begin{cases} -0{,}3p_1 + 0{,}1p_2 + 0{,}7p_3 & = & 0, \\ 0{,}3p_1 - 0{,}6p_2 + 0{,}2p_3 & = & 0, \\ p_1 + p_2 + p_3 & = & 1, \end{cases}$$

в которой линейно зависимых уравнений уже не будет (ее главный определитель отличен от нуля). Решая ее, находим стационарное распределение вероятностей:

$$\mathbf{p} = \left(\frac{44}{86}, \frac{27}{86}, \frac{15}{86}\right).$$

□

Пример 13.3. *Для примера 13.1 найти стационарное распределение вероятностей.*

Решение. В данном случае система уравнений (13.3) имеет вид

$$\mathbf{p}\Lambda = (p_0, p_1)\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} = 0,\ p_0 + p_1 = 1,$$

или

$$\begin{cases} -\lambda p_0 + \mu p_1 & = & 0, \\ \lambda p_0 - \mu p_1 & = & 0, \\ p_0 + p_1 & = & 1. \end{cases}$$

И эта система в силу линейной зависимости ее уравнений приводится к более простой форме:

$$\begin{cases} -\lambda p_0 + \mu p_1 & = & 0, \\ p_0 + p_1 & = & 1. \end{cases}$$

Решением системы является вектор

$$\mathbf{p} = (p_0, p_1) = \left(\frac{\mu}{\lambda + \mu}, \frac{\lambda}{\lambda + \mu}\right),$$

дающий и решение поставленной задачи. □

13.3 Предельные вероятности

Остается невыясненным вопрос, каким образом марковская цепь может придти к своему стационарному распределению вероятностей. На практике в большинстве случаев это происходит с течением времени, причем, иногда достаточно длительного. Поэтому есть смысл рассмотреть предел вектора вероятностей

$$\widetilde{\mathbf{p}} = (\widetilde{p}_1, \ldots, \widetilde{p}_N) = \lim_{t\to\infty} \mathbf{p}(t),$$

который называется ***вектором предельных вероятностей*** или ***предельным распределением вероятностей***. Из-за того, что предел может и не существовать, в случае, если он существует, цепь Маркова называют ***эргодической***. Таким образом, нужно найти условия, при которых стационарное распределение вероятностей марковской цепи совпадает с ее предельным распределением. Для этого необходимо определить некоторые понятия, связанные с графом состояний и графом интенсивностей переходов, причем, мы будем говорить, что утверждение касается графа марковской цепи, если оно будет справедливо по отношению к обоим названным графам.

Состояние A_i графа марковской цепи назовем ***несущественным***, если в этом графе найдется такое состояние A_j, что из A_i в A_j перейти по дугам можно, а из A_j в A_i — нельзя. В противном случае состояние A_i называется ***существенным***. Два состояния A_i и A_j называются ***сообщающимися***, если из A_i можно перейти в A_j, а из A_j — в A_i.

На рис. 13.3 состояния A_1 и A_2 — несущественные, так как из них можно перейти в состояние A_4, а обратно — нельзя. В то же время состояния A_1, A_2 — сообщающиеся, так же, как и существенные состояния A_3, A_4. Существенным является и состояние A_5.

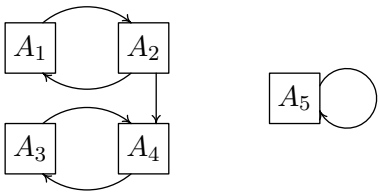


Рис. 13.3: К классификации состояний

Теорема 13.2. *Если A_i — несущественное состояние, то*

$$\lim_{t\to\infty} p_i(t) = 0.$$

Смысл этого утверждения заключается в том, что, покинув несущественное состояние, марковская цепь никогда в него уже не вернется.

Теорема 13.3. *Чтобы марковская цепь с конечным числом состояний и непрерывным временем имела единственное стационарное распределение вероятностей, совпадающее с предельным, необходимо и достаточно, чтобы все ее существенные состояния сообщались между собой.*

Поэтому полученное в примере 13.3 стационарное распределение вероятностей является и предельным, так как состояния графа интенсивностей переходов, как видно из рис. 13.2, существенны и сообщаются между собой.

Состояние марковской цепи с дискретным временем называется ***периодическим***, если в него можно вернуться только за число шагов, кратное одному и тому же числу $d > 1$, называемому ***периодом*** этого состояния. В противном случае состояние называется ***непериодическим***.

Теорема 13.4. *Чтобы марковская цепь с конечным числом состояний и дискретным временем имела единственное стационарное распределение вероятностей, совпадающее с предельным, необходимо и достаточно, чтобы все ее существенные состояния были непериодическими и сообщались между собой.*

В примере 12.2 в каждое состояние марковской цепи можно вернуться за любое число шагов (см. рис. 12.4) и, следовательно, все ее состояния непериодические. Кроме того, они существенные и сообщаются между собой. Следовательно, стационарное распределение вероятностей, найденное в примере 13.2, является также и предельным.

Пример 13.4. *Найти предельное (стационарное) распределение вероятностей для процесса размножения и гибели.*

Решение. Как показывает граф интенсивностей этого процесса (рис. 13.1), все его состояния существенны и сообщаются между собой. Следовательно, предельное распределение вероятностей совпадает со стационарным. Найдем последнее, решив систему уравнений (13.2). Для этого рассмотрим фрагмент упомянутого графа интенсивностей, показанный на рис. 13.4. Ему соответствует фрагмент матрицы

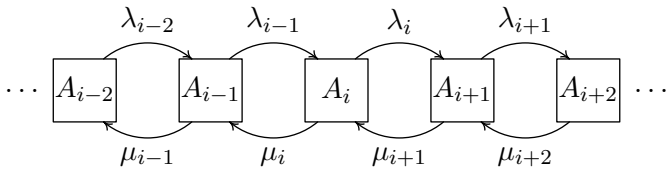


Рис. 13.4: Фрагмент графа процесса размножения и гибели

интенсивностей вида

$$\begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \mu_{i-1} & -(\lambda_{i-1}+\mu_{i-1}) & \lambda_{i-1} & 0 & 0 & \cdots \\ \cdots & 0 & \mu_i & -(\lambda_i+\mu_i) & \lambda_i & 0 & \cdots \\ \cdots & 0 & 0 & \mu_{i+1} & -(\lambda_{i+1}+\mu_{i+1}) & \lambda_{i+1} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

□

В левой части системы (13.3) умножение вектора $\mathbf{p} = (p_0, \ldots, p_N)$ на i-й столбец матрицы Λ дает i-е уравнение системы вида

$$\lambda_{i-1}p_{i-1} - (\lambda_i + \mu_i)p_i + \mu_{i+1}p_{i+1} = 0,$$

или

$$\lambda_{i-1}p_{i-1} - \mu_i p_i = \lambda_i p_i - \mu_{i+1}p_{i+1}. \tag{13.4}$$

Рассмотрим величину $g_i = \lambda_{i-1}p_{i-1} - \mu_i p_i$, причем, положим $g_0 = \lambda_{-1}p_{-1} - \mu_0 p_0 = 0$, вполне естественно считая, что $\lambda_{-1} = p_{-1} =$

$= \mu_0 = 0$. Из уравнения (13.4) получаем, что $g_i = g_{i+1}$, и, значит, $g_N = \cdots = g_1 = g_0 = 0$. Определение величины g_i тогда дает равенства $\lambda_{i-1}p_{i-1} = \mu_i p_i$, откуда

$$p_i = \frac{\lambda_{i-1}}{\mu_i}\, p_{i-1} = \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}\, p_0,$$

или

$$\boxed{p_i = \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}}\, p_0; \;\; i = \overline{1, N};} \tag{13.5}$$

где

$$\prod_{k=1}^{m} a_k = a_1 a_2 \cdots a_m.$$

Вероятность p_0 найдем из условия $\sum_{i=0}^{N} p_i = 1$, которое в данном случае принимает вид

$$\left(1 + \sum_{i=1}^{N} \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}}\right) p_0 = 1$$

Тогда

$$\boxed{p_0 = \left(1 + \sum_{i=1}^{N} \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}}\right)^{-1}.} \tag{13.6}$$

Формулы (13.6) и (13.5) и дают стационарное и предельное распределения вероятностей процесса размножения и гибели.

Лекция 14
Корреляционная теория случайных процессов

14.1 Основные определения

Корреляционная теория занимается анализом случайных процессов, основанном только на двух их характеристиках: математическом ожидании и ковариационной (корреляционной) функции.

Говорят, что с.п. $X(t)$ ***сходится в среднеквадратичном*** при $t \to t_0$ к с.в. X_0 и пишут

$$X_0 = \underset{t \to t_0}{\text{l.i.m.}}\, X(t),$$

если

$$\lim_{t \to t_0} \mathbf{M}[X(t) - X_0]^2 = 0.$$

С.п. $X(t)$ называется ***непрерывным в среднеквадратичном*** в точке t_0, если

$$\underset{t \to t_0}{\text{l.i.m.}}\, X(t) = X(t_0).$$

Теорема 14.1. *Для непрерывности в среднеквадратичном с.п. $X(t)$ в интервале (a, b) необходимо и достаточно, чтобы его математическое ожидание $m(t)$ было непрерывно в этом интервале, а ковариационная функция $K(t_1, t_2)$ была непрерывна на диагонали $t_1 = t_2 = t \in (a, b)$.*

14.2 Дифференцирование с.п.

Производной с.п. $X(t)$ называется с.п. $X'(t)$, определяемый формулой

$$X'(t) = \underset{\Delta t \to 0}{\text{l.i.m.}} \frac{X(t+\Delta t) - X(t)}{\Delta t}. \tag{14.1}$$

Теорема 14.2. *Для дифференцируемости с.п. $X(t)$ на интервале (a,b) необходимо и достаточно, чтобы его математическое ожидание $m(t)$ было дифференцируемой функцией в (a,b) и ковариационная функция имела вторую смешанную производную по t_1 и t_2 на диагонали $t_1 = t_2 = t \in (a,b)$.*

В этом случае

$$\boxed{m_{X'}(t) = m'_X(t), \quad K_{X'}(t_1,t_2) = \frac{\partial^2 K_X(t_1,t_2)}{\partial t_1 \partial t_2}.}$$

Доказательство. Возьмем математическое ожидание от обеих частей равенства (14.1):

$$\begin{aligned}
\mathbf{M}X'(t) &= \mathbf{M}\left[\underset{\Delta t \to 0}{\text{l.i.m.}} \frac{X(t+\Delta t) - X(t)}{\Delta t}\right] = \\
&= \underset{\Delta t \to 0}{\text{l.i.m.}} \frac{\mathbf{M}[X(t+\Delta t)] - \mathbf{M}X(t)}{\Delta t} = \\
&= \underset{\Delta t \to 0}{\text{l.i.m.}} \frac{m_X(t+\Delta t) - m_X(t)}{\Delta t} = m'_X(t).
\end{aligned}$$

Обоснованность перестановки операций математического ожидания и предела в среднеквадратичном примем без доказательства.

Аналогично, используя определение ковариационной функции, уже доказанную первую часть данной теоремы и обозначение $\overset{\circ}{X} = X(t) - \mathbf{M}X(t)$, получаем

$$\begin{aligned}
K_{X'}(t_1,t_2) &= \mathbf{M}[\overset{\circ}{X}{}'(t_1)\,\overset{\circ}{X}{}'(t_2)] = \mathbf{M}\frac{\partial^2[\overset{\circ}{X}(t_1)\,\overset{\circ}{X}(t_2)]}{\partial t_1 \partial t_2} = \\
&= \frac{\partial^2 \mathbf{M}[\overset{\circ}{X}(t_1)\,\overset{\circ}{X}(t_2)]}{\partial t_1 \partial t_2} = \frac{\partial^2 K_X(t_1,t_2)}{\partial t_1 \partial t_2}.
\end{aligned}$$

□

Пример 14.1. *На вход дифференцирующего устройства поступает с.п. с математическим ожиданием $m_X(t) = 3t^2 + t$ и ковариационной функцией*

$$K_X(t_1, t_2) = \sigma^2 e^{-\alpha|t_1 - t_2|}(1 + \alpha|t_1 - t_2|).$$

Дифференцируем ли данный процесс в среднеквадратичном? Найти математическое ожидание и дисперсию с.п. на выходе дифференцирующего устройства.

Решение. Используем результат предыдущей теоремы. Очевидно, производная математического ожидания с.п. на входе устройства существует и равна

$$m_{X'}(t) = (3t^2 + t)' = 6t + 1.$$

Пусть $t_1 \geq t_2$. Найдем вторую смешанную производную ковариационной функции исходного процесса:

$$\frac{\partial K_X}{\partial t_1} = -\alpha\sigma^2 e^{-\alpha|t_1 - t_2|}(1 + \alpha|t_1 - t_2|) + \alpha\sigma^2 e^{-\alpha|t_1 - t_2|};$$

$$\frac{\partial^2 K_X}{\partial t_1 \partial t_2} = -\alpha^2\sigma^2 e^{-\alpha|t_1 - t_2|}(1 + \alpha|t_1 - t_2|) + 2\alpha^2\sigma^2 e^{-\alpha|t_1 - t_2|}. \qquad (14.2)$$

Такое же выражение получится при $t_1 < t_2$. Поэтому ковариационная функция $K_{X'}(t_1, t_2)$ процесса на выходе устройства выражается формулой (14.2), а его дисперсия равна величине

$$\mathbf{D}X'(t) = K_{X'}(t_1, t_2)\big|_{t_1 = t_2 = t} = \frac{\partial^2 K_X(t_1, t_2)}{\partial t_1 \partial t_2}\bigg|_{t_1 = t_2 = t} = \alpha^2\sigma^2.$$

□

14.3 Интегрирование с.п.

Интегралом в среднеквадратичном от с.п. $X(t)$ в пределах от t_1 до t_2 называется величина

$$\int\limits_{t_1}^{t_2} X(t)dt = \underset{n\to\infty}{\text{l.i.m.}} \sum_{i=1}^{n} X(s_i)\Delta s_i,$$

если предел, который берется в предположении, что $\max\limits_i |\Delta s_i| \to 0$, существует и не зависит от способа разбиения отрезка $[t_1, t_2]$ на части точками $s_0 = t_1, s_1, \ldots, s_n = t_2$, $\Delta s_i = s_i - s_{i-1}$.

Теорема 14.3. *Если с.п. $Y(t)$ получается интегрированием в среднеквадратичном с.п. $X(t)$:*

$$Y(t) = \int\limits_0^t X(s)ds,$$

то его числовые характеристики вычисляются по формулам:

$$\boxed{m_Y(t) = \int\limits_0^t m_X(s)ds, \quad K_Y(t_1, t_2) = \int\limits_0^{t_1}\int\limits_0^{t_2} K_X(s_1, s_2)ds_1 ds_2.}$$

Доказательство. Действительно,

$$m_Y(t) = \mathbf{M}\,\underset{n\to\infty}{\text{l.i.m.}} \sum_{i=1}^{n} X(s_i)\Delta s_i = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbf{M}X(s_i)\Delta s_i =$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} m(s_i)\Delta s_i = \int\limits_0^t m_X(s)ds$$

(законность перестановки знаков $\mathbf{M}$ и l.i.m. оставляем без доказательства). Аналогично

$$K_Y(t_1,t_2) = \mathbf{M}\,\overset{\circ}{Y}(t_1)\,\overset{\circ}{Y}(t_2) = \mathbf{M}\int\limits_0^{t_1}\overset{\circ}{X}(s_1)ds_1\int\limits_0^{t_2}\overset{\circ}{X}(s_2)ds_2 =$$

$$= \int\limits_0^{t_1}\int\limits_0^{t_2}\mathbf{M}\,\overset{\circ}{X}(s_1)\,\overset{\circ}{X}(s_2)ds_1ds_2 = \int\limits_0^{t_1}\int\limits_0^{t_2}K_X(s_1,s_2)ds_1ds_2.$$

□

Пример 14.2. *На RC-цепочку (рис. 14.1), подается случайное напряжение $X(t)$ с характеристиками $m_X(t) = 2t$ и $K_X(t_1,t_2) = 4t_1t_2$. Найти математическое ожидание, ковариационную функцию и дисперсию напряжения $Y(t)$ на выходе RC-цепочки.*

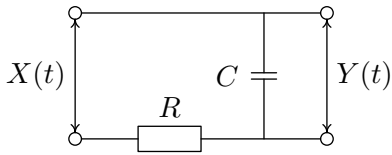


Рис. 14.1: RC-цепочка

Решение. Если бы на вход RC-цепочки подавалось детерминированное напряжение $x(t)$, то на выходе получалось бы детерминированное напряжение $y(t)$. Заряд Q конденсатора равен произведению его емкости и напряжения на его обкладках: $Q = Cy$, а ток является скоростью изменения заряда конденсатора: $i = \frac{dQ}{dt} = C\frac{dy}{dt}$. Сумма падений напряжений на резисторе и конденсаторе равна входному напряжению: $Ri + y = x$, поэтому $RC\frac{dy}{dt} + y = x$. В результате получаем уравнение:

$$\dot{y} + \beta y = \beta x,\ \ \beta = (RC)^{-1}. \tag{14.3}$$

Так как это уравнение линейно, решим его методом подстановки $y(t) = u(t)v(t)$ при начальном условии $y(0) = 0$. Подстановка приводит к уравнению

$$\dot{u}v + u\dot{v} + \beta uv = \beta x,$$

или

$$\dot{u}v + u(\dot{v} + \beta v) = \beta x.$$

Подбирая функцию v так, чтобы выражение в скобках равнялось нулю, получаем систему уравнений:

$$\begin{aligned} \dot{v} + \beta v &= 0, \\ \dot{u}v &= \beta x. \end{aligned}$$

Решаем первое из них, получаем $v(t) = e^{-\beta t}$ и подставляем эту функцию во второе уравнение вместо v:

$$\dot{u} = \beta x e^{\beta t}.$$

Тогда решение последнего уравнения, очевидно, можно выразить формулой

$$u(t) = \beta \int\limits_0^t x(\tau) e^{\beta\tau} d\tau,$$

а решение уравнения (14.3) — формулой

$$y(t) = \beta e^{-\beta t} \int\limits_0^t x(\tau) e^{\beta\tau} d\tau.$$

Мы видим, что напряжение на выходе RC-цепочки представляет собой проинтегрированное с «весом» $e^{\beta t}$входное напряжение. Следовательно, применимы результаты теоремы 14.3 и числовые характеристики с.п. на выходе устройства получаются такими:

$$m_Y(t) = \beta e^{-\beta t} \int\limits_0^t m_X(\tau) e^{\beta\tau} d\tau = 2\beta e^{-\beta t} \int\limits_0^t \tau e^{\beta\tau} d\tau =$$

$$=< u=\tau, dv=e^{\beta\tau}d\tau, du=d\tau, v=e^{\beta\tau}/\beta >=$$

$$=2\beta e^{-\beta t}\left(\frac{\tau}{\beta}e^{\beta\tau}\Big|_0^t-\frac{1}{\beta}\int\limits_0^t e^{\beta\tau}d\tau\right)=$$

$$=2\beta e^{-\beta t}\left(\frac{t}{\beta}e^{\beta t}-\frac{1}{\beta^2}e^{\beta\tau}\Big|_0^t\right)=$$

$$=2\beta e^{-\beta t}\left(\frac{t}{\beta}e^{\beta t}-\frac{1}{\beta^2}e^{\beta t}+\frac{1}{\beta^2}\right)=2\left[t-\frac{1}{\beta}\left(1-e^{-\beta t}\right)\right];$$

$$K_Y(t_1,t_2)=4\beta^2e^{-\beta(t_1+t_2)}\int\limits_0^{t_1}s_1e^{\beta s_1}ds_1\int\limits_0^{t_2}s_2e^{\beta s_2}ds_2=$$

$$=4\left[t_1-\frac{1}{\beta}\left(1-e^{-\beta t_1}\right)\right]\left[t_2-\frac{1}{\beta}\left(1-e^{-\beta t_2}\right)\right].$$

□

14.4 Стационарный с.п.

С.п. называется ***стационарным (в широком смысле)***, если его математическое ожидание постоянно, а ковариационная функция зависит лишь от разности аргументов:

$$K_X(t,t+\tau)=K_X(\tau).$$

Очевидно, что дисперсия стационарного процесса тоже постоянна: $\mathbf{D}X=K_X(t,t)=K_X(0)$.

СВОЙСТВА КОВАРИАЦИОННОЙ ФУНКЦИИ СТАЦИОНАРНОГО С.П.

1°. *Ковариационная функция стационарного с.п. четна:*

$$K_X(\tau)=K_X(-\tau).$$

2°. $|K_X(\tau)|\le K_X(0)=\mathbf{D}X.$

Теорема 14.4. *Если $X(t)$ — стационарный дифференцируемый с.п. и $Y(t) = X'(t)$, то $Y(t)$ — тоже стационарный с.п., причем,*

$$\boxed{m_Y(t) = 0, \quad K_Y(\tau) = -\frac{d^2}{d\tau^2} K_X(\tau).}$$

Для дифференцируемости с.п. $X(t)$ необходимо и достаточно, чтобы существовала дисперсия с.п. $X'(t)$.

Доказательство. Так как математическое ожидание стационарного с.п. постоянно, то в соответствии с теоремой 14.2 для математического ожидания с.п. $X'(t)$ получаем

$$m_{X'} = (m_X)' = 0.$$

Пусть $\tau = t_2 - t_1$, тогда $\frac{\partial \tau}{\partial t_1} = -1$, $\frac{\partial \tau}{\partial t_2} = 1$ и, снова используя теорему 14.2, находим, что

$$K_{X'}(\tau) = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}[K_X(t_2 - t_1)] = \frac{\partial}{\partial t_1}\left[\frac{\partial K_X(\tau)}{\partial \tau} \cdot \frac{\partial \tau}{\partial t_2}\right] =$$
$$= \frac{\partial^2 K_X(\tau)}{\partial \tau^2} \cdot \frac{\partial \tau}{\partial t_1} = -\frac{\partial^2 K_X(\tau)}{\partial \tau^2}.$$

С учетом стационарности с.п. и в силу теоремы 14.2 с.п. $X'(t)$ будет дифференцируем, если его ковариационная функция имеет производную в нуле:

$$K_Y(0) = -\left.\frac{\partial^2 K_X(\tau)}{\partial \tau^2}\right|_{\tau=0}.$$

Но $\mathbf{D}X' = K_Y(0)$, что́ доказывает последнее утверждение теоремы. □

Пример 14.3. *Найти ковариационную функцию производной стационарного с.п. $X(t)$, ковариационная функция которого равна*

$$K_X(\tau) = \sigma_X^2 e^{-\alpha^2\tau^2}, \ \alpha > 0.$$

Решение. Применяя предыдущую теорему, получаем

$$\begin{aligned} K_{X'}(\tau) &= -\sigma_X^2 \left(e^{-\alpha^2\tau^2}\right)''_\tau = 2\alpha^2\sigma_X^2 \left(\tau e^{-\alpha^2\tau^2}\right)'_\tau = \\ &= 2\alpha^2\sigma_X^2 \left(e^{-\alpha^2\tau^2} - 2\alpha^2\tau^2 e^{-\alpha^2\tau^2}\right) = \\ &= 2\alpha^2\sigma_X^2(1 - 2\alpha^2\tau^2)e^{-\alpha^2\tau^2}. \end{aligned}$$

$\square$

Лекция 15
Спектральная теория

15.1 Спектральная плотность

Стационарный с.п. $X(t)$ называется ***случайным процессом с непрерывным спектром***, если существует такая действительная непрерывная функция $S_X(\omega)$, $-\infty < \omega < \infty$, называемая ***спектральной плотностью***, что справедливы формулы Винера-Хинчина:

$$\boxed{K_X(\tau) = \frac{1}{2}\int\limits_{-\infty}^{\infty} S_X(\omega)e^{j\omega\tau}d\omega, \quad S_X(\omega) = \frac{1}{\pi}\int\limits_{-\infty}^{\infty} K_X(\tau)e^{-j\omega\tau}d\tau,}$$

где $K_X(\tau)$ — ковариационная функция с.п. $X(t)$.

Если процесс стационарен, то $S_X(\omega) \geq 0$, $S_X(-\omega) = S_X(\omega)$ (четность спектральной функции).

Очевидно, дисперсия такого с.п. может быть вычислена по формуле

$$\boxed{\mathbf{D}X = K_X(0) = \int\limits_{0}^{\infty} S_X(\omega)d\omega.}$$

Эффективной шириной спектра $\Delta\omega$ называется величина

$$\Delta\omega = \frac{1}{\max\limits_{\omega} S_X(\omega)}\int\limits_{-\infty}^{\infty} S_X(\omega)d\omega = \frac{2\sigma_X^2}{\max\limits_{\omega} S_X(\omega)},$$

а ***средним интервалом корреляции*** $\Delta\tau$ — величина

$$\Delta\tau = \frac{2}{\sigma_X^2}\int\limits_0^\infty |K_X(\tau)|d\tau.$$

Геометрически $\Delta\tau$ (соответственно $\Delta\omega$) равен основанию прямоугольника с высотой 1 (с высотой $\max\limits_\omega S_X(\omega)$), площадь которого равна площади под кривой $|r_X(\tau)|$ (площади под кривой $S_X(\omega)$).

Можно показать, что $\Delta\tau$ и $\Delta\omega$ связаны неравенством $\Delta\tau\cdot\Delta\omega \geq$ $\geq 2\pi$, которое обычно называют «соотношением неопределенности».

Его смысл заключается в том, что, чем у́же ширина спектра стационарного с. п., тем шире интервал корреляции его сечений, и наоборот. По-другому: ширину спектра и интервал корреляции нельзя уменьшить одновременно.

Пример 15.1. *Пусть* $K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|}$, $\alpha > 0$. *Найти* $S_X(\omega)$, $\Delta\tau$ *и* $\Delta\omega$.

Решение. Сначала найдем спектральную функцию

$$\begin{aligned}
S_X(\omega) &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty}\sigma_X^2 e^{-\alpha|\tau|}e^{-j\omega\tau}d\tau = \\
&= \frac{\sigma_X^2}{\pi}\left[\int\limits_0^\infty e^{-(\alpha+j\omega)\tau}d\tau + \int\limits_{-\infty}^0 e^{(\alpha-j\omega)\tau}d\tau\right] = \\
&= \frac{\sigma_X^2}{\pi}\left[-\frac{1}{\alpha+j\omega}e^{-(\alpha+j\omega)\tau}\Big|_0^\infty + \frac{1}{\alpha-j\omega}e^{(\alpha-j\omega)\tau}\Big|_{-\infty}^0\right] = \\
&= \frac{\sigma_X^2}{\pi}\left(\frac{1}{\alpha+j\omega}+\frac{1}{\alpha-j\omega}\right) = \frac{\sigma_X^2}{\pi}\frac{2\alpha}{\omega^2+\alpha^2}.
\end{aligned}$$

Эффективная ширина спектра равна

$$\Delta\omega = \frac{2\sigma_X^2}{S_X(0)} = 2\sigma_X^2\frac{\pi}{\sigma_X^2}\frac{\alpha^2}{2\alpha} = \pi\alpha,$$

а средний интервал корреляции вычисляется так:

$$\Delta\tau = \frac{2}{\sigma_X^2}\sigma_X^2 \int\limits_0^\infty e^{-\alpha\tau} d\tau = -\frac{2}{\alpha}\, e^{-\alpha\tau}\Big|_0^\infty = \frac{2}{\alpha}.$$

Нетрудно проверить, что «соотношение неопределенности» тоже выполняется. □

15.2 С.п. в линейной динамической системе

Линейной динамической системой будем называть систему, которая описывается линейным дифференциальным уравнением n-го порядка с постоянными коэффициентами:

$$a_n\dot{y}^{(n)} + a_{n-1}\dot{y}^{(n-1)} + \cdots + a_0 y = b_m\dot{x}^{(m)} + b_{m-1}\dot{x}^{(m-1)} + \cdots + b_0 x, \quad (15.1)$$

где $x(t)$ — реализация ***входного*** стационарного с.п. $X(t)$, $y(t)$ — реализация процесса $Y(t)$ на ***выходе*** системы.

Начальные условия для уравнения (15.1) примем нулевыми:

$$\dot{y}^{(i)}(0) = 0, \ i = \overline{0, n-1}.$$

Можно показать, что применение к уравнению (15.1) преобразования Фурье (при $t \gg 0$), приводит к соотношениям [26]

$$\boxed{m_Y = \frac{b_0}{a_0} \cdot m_X, \quad S_Y(\omega) = |H(j\omega)|^2 S_X(\omega),}$$

где

$$\boxed{H(p) = \frac{b_m p^m + \cdots + b_0}{a_n p^n + \cdots + a_0}}$$

— ***передаточная функция*** линейной динамической системы, а функция $|H(j\omega)|^2$ — ее ***амплитудно-частотная характеристика***.

Дисперсия на выходе системы вычисляется по формуле

$$\boxed{\mathbf{D}Y = \frac{1}{2}\int\limits_{-\infty}^{\infty} S_Y(\omega)d\omega = \frac{1}{2}\int\limits_{-\infty}^{\infty} |H(j\omega)|^2 S_X(\omega)d\omega.}$$

Пример 15.2. *На вход системы $y'' + 5y' + 4y = x' + 4x$ поступает стационарный с.п. $X(t)$ с числовыми характеристиками $m_X = m$ и ковариационной функцией*

$$K_X(\tau) = \sigma_X^2 e^{-\alpha|\tau|},\ \alpha > 0.$$

Найти числовые характеристики на выходе системы.

Решение. Сначала найдем математическое ожидание выходного сигнала:

$$m_Y = \frac{1}{1} \cdot m_X = m_X,$$

а затем — его спектральную плотность. Для этого запишем передаточную функцию:

$$H(p) = \frac{p+4}{p^2+5p+4} = \frac{p+4}{(p+4)(p+1)} = \frac{1}{p+1}.$$

и амплитудно-частотную характеристику системы:

$$|H(j\omega)|^2 = H(j\omega)H(-j\omega) = \frac{1}{1+j\omega} \cdot \frac{1}{1-j\omega} = \frac{1}{\omega^2+1}.$$

В предыдущем примере было найдено, что

$$S_X(\omega) = \sigma_X^2 \frac{2\alpha}{\pi} \frac{1}{\omega^2+\alpha^2}.$$

Поэтому

$$S_Y(\omega) = \sigma_X^2 \frac{2\alpha}{\pi} \frac{1}{(\omega^2+\alpha^2)(\omega^2+1)}.$$

Ковариационная функция на выходе системы выражается интегралом:

$$K_Y(\tau) = \frac{1}{2}\int\limits_{-\infty}^{\infty} S_Y(\omega)e^{j\omega\tau}d\omega = \sigma_X^2\frac{\alpha}{\pi}\int\limits_{-\infty}^{\infty}\frac{e^{j\omega\tau}}{(\omega^2+\alpha^2)(\omega^2+1)}d\omega.$$

Следовательно, дисперсия на выходе равна

$$\mathbf{D}Y = \sigma_X^2\frac{\alpha}{\pi}\int\limits_{-\infty}^{\infty}\frac{d\omega}{(\omega^2+\alpha^2)(\omega^2+1)}.$$

Но

$$\frac{1}{(\omega^2+\alpha^2)(\omega^2+1)} = \frac{1}{1-\alpha^2}\cdot\frac{1}{\omega^2+\alpha^2} + \frac{1}{\alpha^2-1}\cdot\frac{1}{\omega^2+1}$$

и поэтому

$$\begin{aligned}\mathbf{D}Y &= \frac{\alpha\sigma_X^2}{\pi}\int\limits_{-\infty}^{\infty}\left[\frac{1}{1-\alpha^2}\cdot\frac{1}{\omega^2+\alpha^2} + \frac{1}{\alpha^2-1}\cdot\frac{1}{\omega^2+1}\right]d\omega = \\ &= \frac{\alpha\sigma_X^2}{\pi}\cdot\frac{1}{1-\alpha^2}\left(\frac{1}{\alpha}\operatorname{arctg}\frac{\omega}{\alpha} - \operatorname{arctg}\omega\right)\Bigg|_{-\infty}^{\infty} = \\ &= \frac{\alpha\sigma_X^2}{\pi}\cdot\frac{\pi}{1-\alpha^2}\left(\frac{1}{\alpha}-1\right) = \frac{\sigma_X^2}{1+\alpha}.\end{aligned}$$

То, что дисперсия на выходе в $1+\alpha$ раз меньше дисперсии на входе, объясняется тем, что данная динамическая система по сути является интегрирующей цепью, которая имеет свойство сглаживать сигнал. Сейчас мы об этом поговорим подробнее. □

15.3 Выделение сигнала из шума

Под ***сглаживанием*** стационарного с.п. X понимается применение к нему преобразования вида

$$Z(t) = \int\limits_{-\infty}^{\infty}\varphi(s)X(t-s)ds,$$

где $\varphi(s)$ — четная функция, называемая ***весовой***, $Z(t)$ — сглаженный с.п.

Теорема 15.1. *Сглаженный с.п. $Z(t)$ стационарен и имеет следующие числовые характеристики:*

$$\boxed{\mathbf{M}Z(t) = m_X \int\limits_{-\infty}^{\infty} \varphi(s)ds, \quad S_Z(\omega) = \pi^2 S_X(\omega) S_\varphi^2(\omega).}$$

Доказательство. Найдем математическое ожидание процесса $Z(t)$:

$$\mathbf{M}Z(t) = \mathbf{M} \int\limits_{-\infty}^{\infty} \varphi(s)X(t-s)ds =$$

$$= \int\limits_{-\infty}^{\infty} \varphi(s)\mathbf{M}X(t-s)ds = m_X \int\limits_{-\infty}^{\infty} \varphi(s)ds.$$

Корреляционная функция будет равна

$$K_Z(\tau) = \mathbf{M}[\overset{\circ}{Z}(t+\tau)\,\overset{\circ}{Z}(t)] =$$

$$= \mathbf{M} \int\limits_{-\infty}^{\infty} \varphi(s_1)\,\overset{\circ}{X}(t+\tau-s_1)ds_1 \int\limits_{-\infty}^{\infty} \varphi(s_2)\,\overset{\circ}{X}(t-s_2)ds_2 =$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \varphi(s_1)\varphi(s_2)\mathbf{M}[\overset{\circ}{X}(t+\tau-s_1)\,\overset{\circ}{X}(t-s_2)]ds_1ds_2 =$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \varphi(s_1)\varphi(s_2)K_X(\tau+s_2-s_1)ds_1ds_2.$$

Теперь нетрудно найти и спектральную плотность:

$$S_Z(\omega) = \frac{1}{\pi} \int\limits_{-\infty}^{\infty} K_Z(\tau)e^{-j\omega\tau}d\tau =$$

$$= \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \varphi(s_1)\varphi(s_2) K_X(\tau + s_2 - s_1) e^{-j\omega\tau} ds_1 ds_2 d\tau =$$

$$= \langle y = \tau + s_2 - s_1, dy = d\tau \rangle =$$

$$= \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \varphi(s_1)\varphi(s_2) K_X(y) e^{-j\omega(y+s_1-s_2)} ds_1 ds_2 dy =$$

$$= \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \varphi(s_1) e^{-j\omega s_1} ds_1 \int\limits_{-\infty}^{\infty} \varphi(s_2) e^{-(-j\omega s_2)} ds_2 \int\limits_{-\infty}^{\infty} K_X(y) e^{-j\omega y} dy =$$

$$= \pi^2 S_\varphi(\omega) S_\varphi(-\omega) S_X(\omega).$$

Отсюда в силу четности функции φ получаем, что

$$S_Z(\omega) = \pi^2 S_X(\omega) S_\varphi^2(\omega).$$

□

Пусть задан стационарный с.п.

$$Y(t) = X(t) + \xi(t), \tag{15.2}$$

где $X(t)$ — полезный сигнал, $\xi(t)$ — шум, причем, оба процесса и $X(t)$ и $\xi(t)$ стационарны и некоррелированы между собой. Чтобы выделить сигнал из шума, применим к $Y(t)$ операцию сглаживания. При этом весовую функцию $\varphi(s)$ подберем так, чтобы, по возможности, не исказить спектральную характеристику сигнала S_X. Из равенства (15.2) тогда получим

$$\pi^2 S_Y S_\varphi^2 = \pi^2 (S_X + S_\xi) S_\varphi^2.$$

Приравняем правую часть к S_X и найдем

$$S_\varphi^2 = \frac{S_X}{\pi^2 (S_X + S_\xi)}. \tag{15.3}$$

Таким образом, теоретически φ можно подобрать такой, чтобы спектральные характеристики зашумленного и исходного сигнала совпали.

При этом спектральная характеристика шума станет равной

$$\overline{S}_\xi = \pi^2 S_\xi S_\varphi^2.$$

Учитывая формулу (15.3), получим

$$\boxed{\overline{S}_\xi = \frac{S_X S_\xi}{S_X + S_\xi}.}$$

Видно, что, если $S_X S_\xi \approx 0$, т.е. сигнал и шум колеблются на различных частотах (а это так обычно и бывает: сигнал — низкочастотен, а шум — высокочастотен), то $\overline{S}_\xi \approx 0$. Значит, и дисперсия сглаженного шума $\overline{\mathbf{D}}_\xi \approx 0$.

Кроме того,

$$\overline{S}_\xi = \frac{S_\xi}{S_\xi/S_X + 1} \leq S_\xi,$$

если $S_\xi/S_X \geq 0$. Это тоже говорит об уменьшении дисперсии шума.

Можно показать, что при n-кратном сглаживании процесса (15.2), спектральная плотность шума становится равной

$$\overline{S}_\xi^{(n)} = \frac{S_X S_\xi}{S_X + nS_\xi} \underset{n\to\infty}{\longrightarrow} 0,$$

следовательно, при $n \to \infty$ дисперсия шума неограниченно уменьшается.

Часть III

Теория массового обслуживания

Лекция 16
Аналитические методы изучения СМО

16.1 Основные понятия. Классификация систем массового обслуживания

Предметом теории массового обслуживания является построение, анализ и оптимизация моделей процессов, связанных с удовлетворением массового спроса на обслуживание какого-либо вида, в условиях, когда спрос и обслуживание носят случайных характер. В качестве объектов применения теории можно назвать предприятия железнодорожного, автомобильного и воздушного транспорта, телефонные станции, сети ЭВМ, предприятия массового обслуживания и многое другое. В частности, на железнодорожном транспорте методы теории массового обслуживания применяются для расчета рациональной организации подачи вагонов под погрузку и разгрузку, для расчета мощностей пунктов текущего ремонта вагонов, для выбора рационального числа кассовых аппаратов на вокзалах и станциях метрополитена. Естественно, как и в каждой науке, имеется собственная терминология. Познакомимся с ней.

Заявкой (требованием) называется необходимость в однократном обслуживании. Например, вы стали в очередь в железнодорожную кассу, значит, вы и есть заявка, которая нуждается в обслуживании (в покупке билета). Корабль подошел к причалу для погрузки — он и является заявкой.

Системой массового обслуживания (СМО) называется система обслуживающих устройств, называемых ***каналами обслу-***

живания. Различают ***одно-*** и ***многоканальные*** СМО. Например, железнодорожный кассир — канал обслуживания, причал в морском порту — канал обслуживания.

Поступление заявки в СМО или ее уход из СМО называется ***событием***. Последовательность поступающих заявок называется ***входящим***, а уходящих из СМО — ***выходящим потоком*** заявок.

Пусть поступающая заявка застает все каналы занятыми. Если она, не дожидаясь обслуживания, покидает СМО, то СМО называется ***СМО с отказами***; если же заявка становится в очередь, дожидаясь освобождения одного из каналов, то СМО называется ***СМО с очередью***.

Если число мест в очереди и время ожидания в ней ограничены, СМО относят к ***СМО смешанного типа***.

СМО с очередью называется ***открытой***, если интенсивность поступающего потока заявок не зависит от состояния самой СМО; в противном случае она называется ***замкнутой***. СМО могут отличаться по так называемой ***дисциплине обслуживания***. В одних СМО заявки обслуживаются в порядке поступления, в других — в случайном порядке, в третьих — некоторые заявки обслуживаются вне очереди (***СМО с приоритетом***) и т.д.

16.2 Простейший поток заявок

Поток заявок называется ***простейшим***, если он обладает следующими свойствами:

1) (*стационарность*) вероятность поступления одной заявки за время t равна
$$\lambda \cdot \Delta t + o(\Delta t),$$
где λ — интенсивность потока заявок;

2) (*ординарность*) вероятность поступления более одной заявки за время Δt равна
$$o\,(\Delta t)\,;$$

3) (*отсутствие последействия*) количества заявок, поступивших в непересекающиеся временные интервалы $\Delta_1, \ldots, \Delta_n$, являются независимыми с.в.

Рассмотрим с.с. $A_i =$ "Поступит i заявок"и с.в. $X(t)$ — количество заявок, поступивших за время t. Очевидно, $X(t)$ — марковская цепь с непрерывным временем, причем, как следует из свойств стационарности и ординарности, из состояния A_i она может перейти лишь в состояние A_{i+1} с интенсивностью λ. Этой цепи Маркова соответствует граф интенсивностей, показанный на рис. 16.1 и матрица интенсивностей

$$\Lambda = \begin{pmatrix} -\lambda & \lambda & 0 & \ldots & 0 & 0 & 0 & \ldots \\ 0 & -\lambda & \lambda & \ldots & 0 & 0 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -\lambda & \lambda & 0 & \ldots \\ 0 & 0 & 0 & \ldots & 0 & -\lambda & \lambda & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 0 & -\lambda & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

$$\boxed{A_0} \xrightarrow{\lambda} \boxed{A_1} \xrightarrow{\lambda} \ldots \xrightarrow{\lambda} \boxed{A_{i-1}} \xrightarrow{\lambda} \boxed{A_i} \xrightarrow{\lambda} \boxed{A_{i+1}} \xrightarrow{\lambda} \ldots$$

Рис. 16.1: Граф процесса поступления заявок

Граф интенсивностей показывает, что этот с.п. можно назвать процессом ***чистого размножения***, причем, этот процесс имеет счетное число состояний.

Такому процессу соответствует система уравнений Колмогорова (12.5):

$$\begin{cases} p_0' = -\lambda p_0, & p_0(0) = 1; \\ \dots\dots\dots\dots & \dots\dots\dots \\ p_i' = \lambda p_{i-1} - \lambda p_i, & p_i(0) = 0; \\ \dots\dots\dots\dots & \dots\dots\dots \end{cases} \tag{16.1}$$

$i \geq 1$. Решим ее, сделав замену $p_i = e^{-\lambda t}\varphi_i(t)$, $p_i' = -\lambda e^{-\lambda t}\varphi_i + e^{-\lambda t}\varphi_i'$, после которой система станет такой:

$$\begin{cases} -\lambda e^{-\lambda t}\varphi_0 + e^{-\lambda t}\varphi_0' = -\lambda e^{-\lambda t}\varphi_0, & p_0(0) = \varphi_0(0) = 1; \\ \dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots \\ -\lambda e^{-\lambda t}\varphi_i + e^{-\lambda t}\varphi_i' = \lambda e^{-\lambda t}\varphi_{i-1} - \lambda e^{-\lambda t}\varphi_i, & p_i(0) = \varphi_i(0) = 0; \\ \dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots \end{cases}$$

или

$$\begin{cases} \varphi_0' = 0, & \varphi_0(0) = 1; \\ \dots\dots\dots & \dots\dots\dots \\ \varphi_i' = \lambda\varphi_{i-1}, & \varphi_i(0) = 0; \\ \dots\dots\dots & \dots\dots\dots \end{cases} \tag{16.2}$$

Первое уравнение имеет решение

$$\varphi_0(t) = 1. \tag{16.3}$$

Покажем, что другие уравнения имеют решения вида

$$\varphi_i(t) = \frac{(\lambda t)^i}{i!},\ i = 1, 2, \dots. \tag{16.4}$$

Доказательство проведем методом математической индукции. Очевидно, в силу формулы (16.3) утверждение (16.4) выполняется при $i = 0$. Пусть оно выполняется для некоторого $i = k - 1$:

$$\varphi_{k-1}(t) = \frac{(\lambda t)^{k-1}}{(k-1)!}. \tag{16.5}$$

Покажем, что оно остается справедливым и для $i = k$. Действительно, если формула (16.5) верна, то k-е уравнение системы (16.2) примет вид

$$\varphi_k' = \lambda \frac{(\lambda t)^{k-1}}{(k-1)!},\ \varphi_k(0) = 0.$$

Интегрируя, получим

$$\varphi_k = \lambda \frac{\lambda^{k-1} t^k}{(k-1)!k} + C.$$

Используя начальное условие найдем, что $\varphi_i(0) = C = 0$. Значит решением является функция

$$\varphi_k = \frac{\lambda^k t^k}{k!},$$

которая совпадает с функцией (16.4) при $i = k$. Доказательство завершено. Таким образом, в общем случае решением системы (16.2) является функция

$$\varphi_i(t) = \frac{(\lambda t)^i}{i!},\ i = 0, 1, 2, \ldots,$$

а решением системы (16.1) — функция

$$p_i(t) = \frac{(\lambda t)^i}{i!} e^{-\lambda t},\ i = 0, 1, 2, \ldots. \tag{16.6}$$

Полученная функция $p_i(t)$ представляет собой для каждого t плотность вероятностей пуассоновского распределения с параметром λt. Поэтому простейший поток называют также ***пуассоновским***.

Из последней формулы следует, что вероятность того, что за время t не поступит ни одной заявки, равна

$$p_0(t) = P(X(t) = 0) = e^{-\lambda t}.$$

Пусть T — случайный интервал времени между двумя последовательными поступлениями заявок. На основании предыдущей формулы найдем его распределение:

$$F(t) = P(T < t) = 1 - P(T \geq t) = 1 - P(X(t) = 0) = \\ = 1 - e^{-\lambda t},\ t > 0,$$

и $F(t) = 0$ при $t \leq 0$. Видим, что интервал времени между поступлениями двух заявок распределен по показательному закону с параметром λ. Поэтому $\mathbf{M}T = \sigma_T = 1/\lambda$.

16.3 Марковские СМО

Чтобы задать СМО, надо, кроме характеристик входящего потока заявок, указать распределение с.в. $T_{\text{об}}$, времени обслуживания одной заявки. Часто бывает, что эта с.в. распределена по показательному закону:

$$F_{T_{\text{об}}}(t) = \begin{cases} 1 - e^{-\mu t}, & t > 0; \\ 0, & t \leq 0; \end{cases}$$

где μ — интенсивность потока обслуживания. ***Потоком обслуживания*** называется поток заявок, обслуживаемых одним и тем же каналом. Если $T_{\text{об}}$ имеет показательное распределение, то поток обслуживания называется ***простейшим***.

СМО, в которой входящий поток и все потоки обслуживания — простейшие, называется ***марковской СМО***, так как СМО тогда является марковской цепью с дискретными состояниями и непрерывным временем.

Пример 16.1. *Прибор состоит из $m = 3$ узлов, которые могут заменять друг друга. Для работы прибора достаточно, чтобы работал хотя бы один узел. При выходе из строя одного из узлов прибор продолжает нормально функционировать, пока не выйдут из строя все узлы. Поток отказов каждого узла — простейший. Среднее время безотказной работы одного узла $t = 20$ часов. При выходе*

из строя узел начинает сразу же ремонтироваться. Время ремонта распределено по показательному закону и в среднем составляет $s = 4$ часа. В начальный момент времени все узлы исправны. Найти среднюю производительность прибора, если с выходом из строя каждого узла прибор теряет $(100/m)\,\%$ своей производительности.

Решение. Пусть A_i — событие, состоящее в том, что i узлов прибора находятся в ремонте (i отказов). Потоки отказов и восстановления являются простейшими, поэтому их интенсивности вычисляются по формулам

$$\lambda = \frac{1}{\mathbf{M}T} = \frac{1}{20},$$

где T — время между двумя отказами (время безотказной работы);

$$\mu = \frac{1}{\mathbf{M}T_{\text{об}}} = \frac{1}{4}.$$

Переход из состояния A_i может произойти либо в состояние A_{i+1} с интенсивностью $\lambda_i = (m-i)/t = \lambda(m-i)$, либо в состояние A_{i-1} с интенсивностью $\mu_i = i/s$. Таким образом, имеем марковскую цепь с графом интенсивностей, показанным на рис. 16.2. Используя фор-

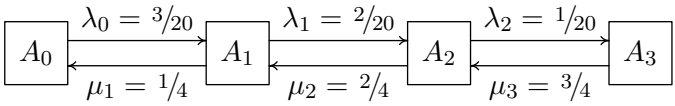


Рис. 16.2: Граф интенсивностей к примеру 16.1

мулы предельных вероятностей для процесса размножения и гибели

(13.5) и (13.6), получаем

$$p_0 = \left(1 + \frac{3}{20} \cdot 4 + \frac{6}{400} \cdot 8 + \frac{6}{8000} \cdot \frac{64}{6}\right)^{-1} =$$
$$= (1 + 0,6 + 0,12 + 0,008)^{-1} = \frac{1}{1,728} \approx 0,579;$$
$$p_1 = 0,6 \cdot \frac{1}{1,728} \approx 0,347;$$
$$p_2 = 0,12 \cdot \frac{1}{1,728} \approx 0,069;$$
$$p_3 = 0,008 \cdot \frac{1}{1,728} \approx 0,005.$$

Средняя производительность вычисляется по формуле

$$\overline{\text{Пр}} = \left(p_0 + \frac{m-1}{m}\, p_1 + \dots + \frac{m-i}{m}\, p_i + \dots + \frac{0}{m}\, p_m\right) 100\% =$$
$$= \left(0,579 + \frac{2}{3} \cdot 0,347 + \frac{1}{3} \cdot 0,069\right) 100\% = 83,3\%.$$

□

16.4 Показатели эффективности СМО

A — ***абсолютная пропускная способность***, среднее число заявок, обслуживаемых СМО в единицу времени.

$Q = A/\lambda$ — ***относительная пропускная способность***, вероятность обслуживания поступившей заявки.

$P_{\text{отк}} = 1 - Q$ — ***вероятность отказа***.

$\overline{z}$ — ***среднее число заявок в СМО*** (обслуживаемых и ожидающих в очереди).

$\overline{r}$ — ***среднее число заявок в очереди*** (если есть очередь).

$\overline{t}_{\text{с}}$ — ***среднее время пребывания заявки в СМО*** (для обслуживания и в очереди).

$\overline{t}_{\text{оч}}$ — ***среднее время пребывания заявки в очереди***.

$\overline{k}$ — ***среднее число занятых каналов***.

Для открытых СМО и стационарных входящем потоке и потоке обслуживания справедливы формулы

$$\boxed{\overline{t}_{\text{с}} = \frac{\overline{z}}{\lambda}, \quad \overline{t}_{\text{оч}} = \frac{\overline{r}}{\lambda}, \quad \overline{k} = \frac{A}{\mu},}$$

где λ — интенсивность потока заявок, μ — интенсивность потока обслуживания.

16.5 Некоторые виды СМО

В сжатом виде информацию о СМО записывают следующим образом:

Распределение входящего потока заявок	Распределение времени обслуживания	Число каналов обслуживания	Емкость накопителя

Например, **Пуас/Показ/5/0** означает пятиканальную СМО с пуассоновским входящим потоком, показательным законом распределения времени обслуживания и отсутствием накопителя.

16.5.1 Пуас/Показ/N/0

Пусть с.с. A_i означает, что i каналов занято обслуживанием заявок, а остальные каналы свободны. Обозначим λ интенсивность входящего потока заявок, а μ — интенсивность обслуживания. Функционирование СМО можно представить в виде, показанном на рис. 16.3.

Из состояния A_i с интенсивностью λ можно перейти либо в состояние A_{i+1}, если поступит еще одна заявка, либо в состояние A_{i-1} с интенсивностью $i\mu$, если от обслуживания освободится *хотя бы один канал*. Следовательно, в соответствующем процессе размножения и

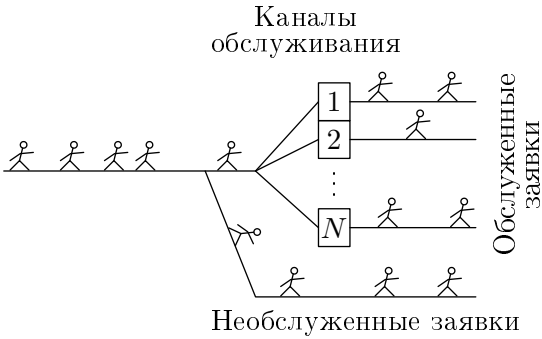


Рис. 16.3: Схема функционирования СМО

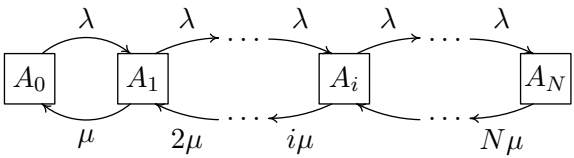


Рис. 16.4: Граф интенсивностей для СМО **Пуас/Показ/N/0**

гибели надо взять $\lambda_i = \lambda$, $\mu_i = i\mu$. Граф интенсивностей показан на рис. 16.4.

Из формул (13.5), (13.6), обозначив $\rho = \lambda/\mu$, получим

$$\frac{\lambda_0 \cdot \ldots \cdot \lambda_{i-1}}{\mu_1 \cdot \ldots \cdot \mu_i} = \frac{\lambda^i}{i!\mu^i} = \frac{\rho^i}{i!}$$

и, значит,

$$\boxed{\begin{aligned} &p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + \cdots + \frac{\rho^N}{N!}\right)^{-1}, \\ &p_i = \frac{\rho^i}{i!}\, p_0,\ i = \overline{0, N}. \end{aligned}} \qquad (16.7)$$

Эти равенства называются ***формулами Эрланга***.

Для данной СМО показатели эффективности принимают следующий вид. Абсолютная пропускная способность $A = (1 - p_N)$ равна произведению интенсивности входящего потока заявок на вероятность того, что хотя бы один канал свободен. Относительная пропускная способность $Q = 1 - p_N$ равна упомянутой вероятности. Другие показатели: $P_{\text{отк}} = p_N$, $\overline{k} = \rho\,(1 - p_N)$. Величина $\rho = \lambda/\mu$ называется ***коэффициентом загрузки системы***.

Пример 16.2. *Городская железнодорожная справочная имеет 5 различных телефонных номеров. Поток вызовов — простейший с интенсивностью $\lambda = 6$ вызовов в мин. Среднее время, затрачиваемое на справку, равно 2/3 мин. Это время распределено по показательному закону. Найти абсолютную и относительную пропускные способности справочной; вероятность отказа; среднее число занятых каналов. Определить, сколько различных телефонных номеров должна иметь справочная, чтобы вероятность отказа не превышала 0,07?*

Решение. По условию задачи среднее время обслуживания равно $\overline{T}_{\text{об}} = 2/3$. Значит, интенсивность обслуживания равна $\mu = \frac{1}{\overline{T}_{\text{об}}} = \frac{3}{2} = 1{,}5$. Кроме того, коэффициент загрузки справочной $\rho = \lambda/\mu = 6/1{,}5 = 4$, $N = 5$. Вероятности состояний справочной вычислим по формулам Эрланга (16.7):

$$p_0 = \left(1 + 4 + 8 + \frac{64}{6} + \frac{256}{24} + \frac{1024}{120}\right)^{-1} = 0{,}023;$$

$$p_1 = 0{,}093;\ p_2 = 0{,}187;\ p_3 = p_4 = 0{,}249;\ p_5 = 0{,}199.$$

Найдем показатели эффективности работы справочной: абсолютная пропускная способность $A = \lambda\,(1 - p_5) = 6 \cdot 0{,}801 = 4{,}806$ (справочная дает в среднем 4,8 справки в мин.); относительная пропускная способность $Q = 1 - p_5 = 0{,}801$ (вероятность обслуживания заявки); вероятность отказа $P_{\text{отк}} = p_5 = 0{,}199$; среднее число занятых каналов $\overline{k} = \rho\,(1 - p_5) = 3{,}204$. Так как вероятность отказа боль-

ше заданной вероятности 0,07, то попробуем увеличить число телефонных номеров на один номер. Тогда N будет равно 6 и пересчет вероятностей даст

$$p_0 = \left(1 + 4 + 8 + \frac{64}{6} + \frac{256}{24} + \frac{1024}{120} + \frac{4096}{720}\right)^{-1} = 0{,}02;$$
$$p_6 = \frac{4096}{720} \cdot 0{,}02 = 0{,}11 > 0{,}07.$$

Следовательно, шести телефонных номеров будет мало и надо попытаться увеличить количество телефонных номеров до семи. Снова выполним пересчет при $N = 7$:

$$p_0 = \left(1 + 4 + 8 + \frac{64}{6} + \frac{256}{24} + \frac{1024}{120} + \frac{4096}{720} + \frac{16384}{5040}\right)^{-1} = 0{,}019;$$
$$p_7 = \frac{16384}{5040} \cdot 0{,}019 = 0{,}063 < 0{,}07.$$

Таким образом, чтобы вероятность отказа не превышала 0,07, необходимо увеличить число телефонных номеров справочной до 7. □

16.5.2 Пуас/Показ/N/m

В отличие от предыдущей СМО рассмотрим СМО, имеющей в своем составе накопитель на m мест ожидания в очереди. Заявка, поступившая в СМО и заставшая накопитель занятым m заявками, ожидающими обслуживания, покидает СМО необслуженной. Если канал занят обслуживанием, а свободное место в накопителе имеется, заявка это место занимает и ожидает в очереди до тех пор, пока ее не обслужат. Наконец, если канал свободен от обслуживания (значит, и заявок в очереди нет), заявка поступает в него и немедленно обслуживается. Схематически функционирование такой СМО показано на рис. 16.5.

Различными состояниями процесса будем считать следующие си-

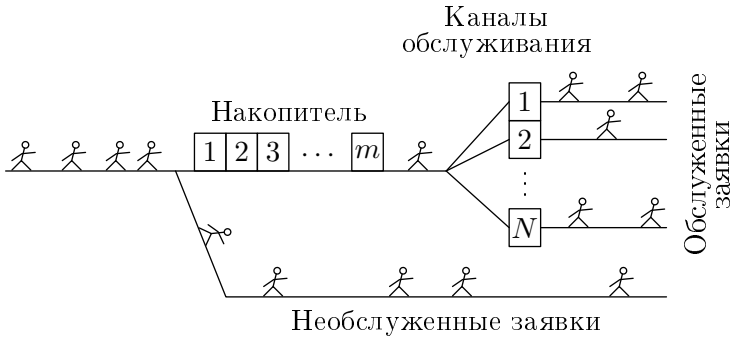


Рис. 16.5: Функционирование СМО с накопителем

туации

$$A_i = \begin{cases} i \text{ каналов заняты}, & i = \overline{0, N}; \\ \text{Все каналы заняты,} \\ \text{в очереди } i - N \text{ заявок}, & i = \overline{N+1, N+m}. \end{cases}$$

В соответствии с этим граф интенсивностей для процесса размножения и гибели приобретает вид, показанный на рис. 16.6. Подсчитаем

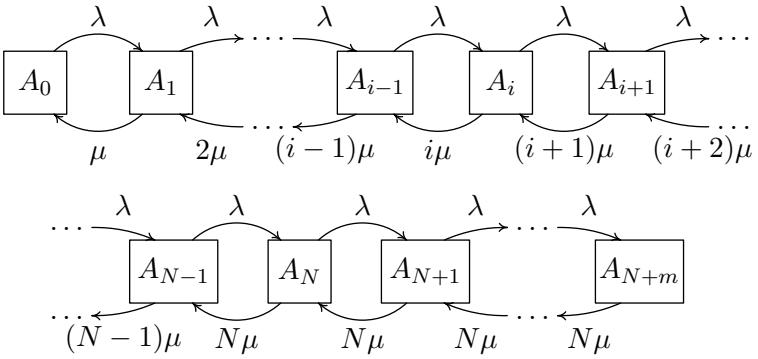


Рис. 16.6: Граф интенсивностей для СМО с накопителем

вероятности состояний СМО по формулам (13.5), (13.6), учитывая особенности графа интенсивностей:

$$p_0 = \\ = \left(1 + \frac{\lambda}{\mu} + \dots + \frac{\lambda^N}{N!\mu^N} + \frac{\lambda^{N+1}}{N \cdot N!\mu^{N+1}} + \dots + \frac{\lambda^{N+m}}{N^m N!\mu^{N+m}}\right)^{-1} = \\ = \left[1 + \rho + \dots + \frac{\rho^N}{N!} + \frac{\rho^N}{N!}\left(\frac{\rho}{N} + \dots + \frac{\rho^m}{N^m}\right)\right]^{-1},$$

или

$$\boxed{\begin{aligned} & p_0 = \left(\sum_{i=0}^{N} \frac{\rho^i}{i!} + \frac{\rho^N}{N!}\sum_{j=1}^{m} \varkappa^j\right)^{-1}; \\ & p_i = \frac{\rho^i}{i!}\, p_0,\ i = \overline{1,N};\ p_{N+j} = \frac{\rho^N}{N!}\, \varkappa^j p_0,\ j = \overline{1,m}; \end{aligned}} \tag{16.8}$$

где $\rho = \lambda/\mu$, $\varkappa = \rho/N$.

Показатели эффективности СМО вычисляют, исходя из следующих соображений. Вероятность отказа равна вероятности того, что все N каналов обслуживания будут заняты и мест в очереди не будет, то есть величине $P_{\text{отк}} = p_{N+m}$. Формулы для абсолютной и относительной пропускных способностей остаются по форме прежними: $A = \lambda(1 - P_{\text{отк}})$, $Q = 1 - P_{\text{отк}}$, но вероятность отказа приобрела, как мы видели, другой смысл. То же можно сказать о среднем числе занятых каналов: $\overline{k} = \rho(1 - P_{\text{отк}})$. На примере данной СМО познакомимся с вычислением такой характеристики как среднее число заявок в очереди:

$$\overline{r} = \sum_{j=1}^{m} j p_{N+j} = \sum_{j=1}^{m} j \frac{\rho^N}{N!}\, \varkappa^j p_0 = \frac{p_0 \rho^N}{N!}\, \varkappa \left(\sum_{j=1}^{m} \varkappa^j\right)' = \\ = \frac{p_0 \rho^N}{N!}\, \varkappa \left(\frac{\varkappa - \varkappa^{m+1}}{1 - \varkappa}\right)' = \\ = \frac{p_0 \rho^N \varkappa}{N!} \cdot \frac{[1 - (m+1)\varkappa^m](1 - \varkappa) + \varkappa - \varkappa^{m+1}}{(1 - \varkappa)^2}.$$

Следовательно,

$$\overline{r} = \frac{p_0 \rho^N \varkappa}{N!} \cdot \frac{1 - \varkappa^m (m + 1 - m\varkappa)}{(1 - \varkappa)^2}.$$

На основании этого показателя эффективности вычисляем среднее число заявок в СМО $\overline{z} = \overline{r} + \overline{k}$, среднее время пребывания заявки в СМО $\overline{t}_c = \overline{z}/\lambda$ и среднее время пребывания заявки в очереди $\overline{t}_{\text{оч}} = \overline{r}/\lambda$.

Пример 16.3 (Продолжение предыдущего примера)**.** *Предположим, что в примере с железнодорожной справочной вместо увеличения числа телефонных номеров принято решение дополнительно установить аппарат, который в случае, если телефоны заняты, ставит клиента в очередь, монотонным голосом сообщая ему: "Ждите ответа. Ждите ответа. ...". Предположим, что в таком накопителе имеется всего четыре места. Как изменятся показатели эффективности работы справочной?*

Решение. В данном случае добавляются значения следующих параметров: $m = 4$, $\varkappa = 4/5 = 0{,}8$. Прежде всего по формулам (16.8) рассчитаем вероятности состояний СМО:

$$\begin{aligned}
p_0 &= [1 + 4 + 8 + 64/6 + 256/24 + 1024/120 + \\
&\quad + 1024\left(0{,}8 + 0{,}8^2 + 0{,}8^3 + 0{,}8^4\right)/120]^{-1} = \\
&= 0{,}0159; \\
p_1 &= 0{,}063;\ p_2 = 0{,}127;\ p_3 = p_4 = 0{,}169;\ p_5 = 0{,}135; \\
p_6 &= 0{,}108;\ p_7 = 0{,}0887;\ p_8 = 0{,}069;\ p_9 = 0{,}055.
\end{aligned}$$

Показатели эффективности работы справочной станут такими:

$$\begin{aligned}
P_{\text{отк}} &= p_9 = 0{,}055 < 0{,}07;\ A = 6\,(1 - 0{,}055) = 5{,}67; \\
Q &= 1 - 0{,}055 = 0{,}945;\ \overline{k} = 4 \cdot 0{,}945 = 3{,}78; \\
\overline{r} &= \frac{0{,}159 \cdot 8{,}33 \cdot 0{,}8}{0{,}2^2} \cdot \left[1 - 0{,}8^4\,(5 - 4 \cdot 0{,}8)\right] = 0{,}72; \\
\overline{z} &= 0{,}72 + 3{,}78 = 4{,}5;\ \overline{t}_c = 4{,}5/6 = 0{,}75;\ \overline{t}_{\text{оч}} = 0{,}72/6 = 0{,}12.
\end{aligned}$$

Мы видим, что показатели работы улучшились: вероятность отказа не превосходит заданной величины 0,07; увеличилась абсолютная пропускная способность, а время ожидания в очереди незначительно. А, главное, отпала необходимость в установке дополнительных телефонных аппаратов. □

16.5.3 Пуас/Показ/N/∞

Пусть теперь количество мест в накопителе неограниченно. В принципе, такие очереди можно себе представить, так как, например, очередь к торговой точке на улице города имеет возможность удлиняться практически сколько угодно. Рассмотрим этот случай как предельный для СМО предыдущего п., когда $m \to \infty$. При этом придется предположить, что величина $\varkappa < 1$, иначе предел существовать не будет. Функционирование такой СМО иллюстрирует рис. 16.7.

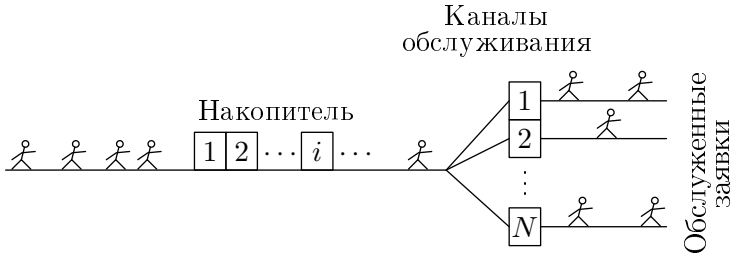


Рис. 16.7: СМО с неограниченным накопителем

Осуществление предельного перехода в равенствах (16.8) приведет к модификации лишь одного выражения: $\lim_{m\to\infty} \sum_{j=1}^{m} \varkappa^j = \varkappa/(1-\varkappa)$ по формуле суммы бесконечно убывающей геометрической прогрессии. Поэтому вероятность p_0 изменится следующим образом:

$$\boxed{p_0 = \left(\sum_{i=0}^{N} \frac{\rho^i}{i!} + \frac{\rho^{N+1}}{N!} \cdot \frac{1}{1-\varkappa} \right)^{-1},} \tag{16.9}$$

а остальные вероятности останутся прежними.

Так как $m\varkappa^m \to 0$ при $m \to \infty$ и $\varkappa < 1$, то формулы для показателей эффективности СМО приобретут вид

$$P_{\text{отк}} = 0,\ Q = 1 - P_{\text{отк}} = 1,\ A = Q\lambda = \lambda,\ \overline{k} = A/\mu = \rho,$$
$$\overline{r} = \frac{p_0\rho^{N+1}}{N!N\left(1-\varkappa\right)^2},\ \overline{z} = \overline{r} + \overline{k} = \overline{r} + \rho,$$
$$\overline{t}_c = \overline{z}/\lambda,\ \overline{t}_{\text{оч}} = \overline{r}/\lambda.$$

Пример 16.4 (Продолжение предыдущего примера)**.** *Пусть аппарат, ставящий клиента в очередь на справку, настолько мощен, что в очереди всегда есть места. Можно считать, что в этом случае очередь не ограничена. Найти показатели модифицированной СМО.*

Решение. Величина $\varkappa = 0{,}8$ удовлетворяет условию $\varkappa < 1$. Произведем перерасчет вероятностей по формулам (16.9) и (16.8):

$$p_0 = (1 + 4 + 8 + 64/6 + 256/24 + 1024/120 + 4096/600/0{,}2)^{-1} = 0{,}013;$$
$$p_1 = 0{,}052;\ p_2 = 0{,}104;\ p_3 = p_4 = 0{,}139;\ p_5 = 0{,}111;$$
$$p_6 = 0{,}089;\ p_7 = 0{,}071;\ p_8 = 0{,}057;\ p_9 = 0{,}045.$$

Вероятность того, что справочная будет занята равна

$$P_* = 1 - p_0 - p_1 - p_2 - p_3 - p_4 =$$
$$= 1 - 0{,}013 - 0{,}052 - 0{,}104 - 0{,}139 - 0{,}139 = 0{,}553.$$

Другие показатели эффективности:

$$P_{\text{отк}} = 0;\ Q = 1;\ A = \lambda = 6;\ \overline{k} = 4;$$
$$\overline{r} = \frac{0{,}013 \cdot 4^6}{5!5\left(1-0{,}8\right)^2} = 2{,}22;\ \overline{z} = 2{,}22 + 4 = 6{,}22;$$
$$\overline{t}_c = 6{,}22/6 = 1{,}04;\ \overline{t}_{\text{оч}} = 2{,}22/6 = 0{,}37.$$

По сравнению с предыдущим случаем абсолютная пропускная способность увеличилась за счет некоторого увеличения среднего времени пребывания заявки в очереди, причем последнее незначительно. Среднее число заявок в очереди также невелико. □

16.5.4 Пуас/Произв/1/∞

Продолжая считать, что входящий поток заявок — простейший, а мощность накопителя не ограничена, предположим теперь, что поток обслуживания имеет произвольное распределение и что в СМО имеется лишь один канал обслуживания. Числовые характеристики потока обслуживания следующие: $\mathbf{M}T_{\text{об}} = 1/\mu$, $\mathbf{D}T_{\text{об}} = \sigma^2$. Можно показать, что такая СМО не будет марковской, а ее характеристики вычисляются ***по формулам Полячека-Хинчина***:

$$\boxed{\begin{aligned} &\overline{r} = \frac{\rho^2\left(1+v^2\right)}{2\left(1-\rho\right)},\ \bar{t}_{\text{оч}} = \frac{\rho^2\left(1+v^2\right)}{2\lambda\left(1-\rho\right)},\\ &\overline{z} = \overline{r} + \rho,\ \bar{t}_{\text{с}} = \bar{t}_{\text{оч}} + \frac{1}{\mu}, \end{aligned}}$$

где $\rho = \lambda/\mu$ — коэффициент загрузки системы, $v = \sigma\mu$ — ***коэффициент вариации времени обслуживания***.

Пример 16.5. *В вагонное депо прибывает для ремонта простейший поток полувагонов со средним интервалом прибытия 9 час. Ремонт вагона в среднем равен 8 час. и распределен по произвольному закону, причем, с.к.о. равно 2 час. Определить среднее число вагонов в депо, из них ожидающих ремонта; среднее время ожидания начала ремонта; среднее время пребывания вагона в депо.*

Решение. Очевидно, рассматриваемый объект представляет собой изучаемый тип СМО. По условию средний интервал между прибытиями полувагонов составляет $\mathbf{M}T = 9$ (час.). Следовательно, интенсивность входящего потока заявок равна $\lambda = 1/\mathbf{M}T = 1/9 \approx 0{,}111$ (вагонов/час). Среднее время обслуживания также задано и равно $\mathbf{M}T_{\text{об}} = 8$ (час.). Значит, интенсивность потока обслуживания составляет $\mu = 1/\mathbf{M}T_{\text{об}} = 1/8 = 0{,}125$ (вагонов/час). Известна дисперсия времени обслуживания $\mathbf{D}T_{\text{об}} = \sigma^2 = 2^2 = 4$. Определим другие параметры работы СМО: $\rho = \lambda/\mu = 8/9 = 0{,}889$; $v = \sigma\mu = \frac{2}{8} = 0{,}25$. Используя формулы Полячека-Хинчина, находим, что среднее число

вагонов, ожидающих ремонта, равно

$$\overline{r} = \frac{0{,}889^2\left(1 + 0{,}25^2\right)}{2\left(1 - 0{,}889\right)} = 3{,}78;$$

среднее число вагонов в депо:

$$\overline{z} = \overline{r} + \rho = 3{,}78 + 0{,}889 = 4{,}67;$$

среднее время ожидания ремонта:

$$\overline{t}_{\text{оч}} = \frac{\overline{r}}{\lambda} = \frac{3{,}78}{0{,}111} = 34{,}02 \text{ (час.)};$$

среднее время пребывания вагона в депо:

$$\overline{t}_{\text{с}} = 34{,}02 + 8 = 42{,}02 \text{ (час.)}.$$

□

Лекция 17
Статистическое моделирование СМО

17.1 Основные понятия

К сожалению, получение аналитических результатов в теории массового обслуживания возможно только при выполнении существенных ограничений на модель СМО. Наиболее существенным ограничением является допущение о марковском характере процесса, что в реальных задачах выполняется далеко не всегда. Результатом есть то, что формулы для расчета характеристик удается получить лишь для самых простых СМО.

Во всех других случаях приходится прибегать к так называемому статистическому моделированию, которое еще называют методом Монте-Карло. Суть его заключается в том, что вручную или с помощью ЭВМ имитируется реализация случайных событий в СМО, таких, как, например, поступление заявки, интервал между поступлениями заявок, продолжительность обслуживания заявки и т.п. Эти события моделируются массово и со скоростью, намного превышающей скорость реализации таких событий в реальной действительности. В результате очень быстро получают статистический материал, который можно подвергнуть обычной статистической обработке. Особенно удобно это выполнять на ЭВМ: и статистический материал на ней получить, и тут же на ЭВМ его обработать. Например, те статистические данные, которые реальная система, например, железнодорожного транспорта может представить в течение месяца, на ЭВМ можно получить, скажем, за сутки. Более того, при мо-

делировании можно опробовать режимы работы СМО, которые на реальной системе проверить бывает или очень сложно или вообще невозможно (например, по соображениям безопасности). Зачем же мы изучали аналитические методы исследования СМО? Потому что, как и в других случаях, аналитические и численные методы имеют свои как полезные так слабые стороны. Аналитические методы дают общее представление о поведении СМО в силу того простого факта, что в формулы входят буквы, которые означают не что иное, как целое семейство значений той величины, которую буква представляет. Поэтому, зная, как может изменяться данная величина, мы можем представить себе, как будут изменяться и зависящие от нее, скажем, показатели эффективности работы СМО. Это — сильная сторона аналитического метода, а о слабых сторонах уже было сказано. В противоположность аналитическим методам метод статистического моделирования такой общностью не обладает, зато применим практически к любым типам СМО. Поэтому изучать и использовать приходится оба метода.

С чего же начинается метод Монте-Карло? Базовым в нем является имитация осуществления следующего с.с.: некая случайная величина ξ, имеющая равномерное распределение на отрезке $[0;1]$, принимает некоторое свое значение из этого отрезка. Для ручных расчетов имеются специальные таблицы случайных чисел [11, 17, 23, 27], состоящие из таких реализаций, так что при моделировании просто надо взять из таблицы очередное число. При использовании ЭВМ применяют встроенные практически во все языки программирования специальные датчики случайных чисел (это — обычные процедуры или функции), которые носят названия подобные RANDOM, после обращении к которым в качестве значения выходного параметра получают число из отрезка $[0;1]$. Оказывается, эти числа позволяют далее получить значения случайных величин, имеющих любое распределение! Рассмотрим эти возможности.

17.2 Генерация случайных чисел

Чтобы получить случайное число X из отрезка $[a, b]$, используют формулу

$$X = a + (b - a)\,\xi,$$

где ξ — упомянутая выше равномерно распределенная на отрезке $[0; 1]$ случайная величина.

Для получения с.в. X, распределенной $N\,(m, \sigma)$, надо просто сложить несколько случайных (и независимых) случайных чисел ξ_i, равномерно распределенных на отрезке $[0; 1]$:

$$X = m + \sigma\left(\sum_{i=1}^{12} \xi_i - 6\right).$$

Для ручных расчетов существуют таблицы случайных чисел, распределенных $N\,(0, 1)$ [27]. Выбрав из такой таблицы случайное число α, можно затем получить из него число X, распределенное $N\,(m, \sigma)$ по формуле

$$X = m + \sigma\alpha.$$

Самым общим является метод обратной функции [18], позволяющий вычислить случайное число, имеющее любое распределение. Пусть, например, с.в. X имеет функцию распределения $F(x)$. Тогда достаточно применить формулу

$$X = F^{-1}\,(\xi)\,.$$

Конкретизируем такую формулу для показательного распределения. В этом случае функция распределения имеет вид

$$y = F\,(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Значит, обратная функция будет иметь вид (надо решить предыдущее уравнение относительно x)

$$F^{-1}\,(y) = -\frac{1}{\lambda}\ln\,(1 - y)\,,$$

а генерировать с.в. X, имеющую показательное распределение, следует по формуле

$$X = -\frac{1}{\lambda} \ln(1 - \xi).$$

Так как $1 - \xi$ имеет точно такое же распределение, как и с.в. ξ, то предыдущую формулу можно упростить:

$$X = -\frac{1}{\lambda} \ln \xi. \tag{17.1}$$

Для с.в. X, распределенной по закону Эрланга с параметрами λ и k, в примере 10.4 было показано, что с.в. X является суммой k с.в., распределенных по показательному закону с тем же самым параметром λ. Следовательно, сначала надо получить по формуле (17.1) k случайных чисел, а затем их сложить. Соответствующая формула имеет вид

$$X = -\frac{1}{\lambda} \sum_{i=1}^{k} \ln \xi_i,$$

где $\xi_1, \ldots, \xi_k$ — случайные числа, равномерно распределенные на отрезке $[0; 1]$.

17.3 Приближенное вычисление показателей эффективности СМО

Для рассмотренных выше простейших СМО покажем, основываясь на результатах [6], как на ЭВМ можно с помощью многократных просчетов определить приближенные значения показателей эффективности СМО. Применяя методы генерирования случайных чисел, распределенных по одному из законов, упомянутых в предыдущем п., на ЭВМ получают интервалы между поступлениями заявок в СМО Δt_i, $i = \overline{1, L}$, где L — запланированное для моделирования число заявок, которые поступят в СМО. Получают также времена обслуживания заявок τ_i, $i = \overline{1, V}$, где V — число обслуженных в СМО заявок. Моменты поступления заявок в СМО t_i вычисляются по формуле $t_i = t_{i-1} + \Delta t_i$, $i = \overline{1, L}$, причем, $t_0 = T_0$, где $[T_0, T_k]$ —

интервал времени, для которого происходит моделирование работы СМО. Далее будем его называть просто интервалом моделирования.

Пусть $r = \overline{1, R}$ — шаг моделирования, который состоит в обслуживании r-й заявки или в отказе ей в обслуживании; R — число шагов моделирования. Если на текущем шаге моделирования r выполняется условие

$$t_r < t_{\text{min,оч}}(r-1) \wedge t_r < t_{\text{min,ос}}(r-1), \tag{17.2}$$

где $t_{\text{min,оч}}(r-1)$ — ближайшее время освобождения места в очереди, а $t_{\text{min,ос}}(r-1)$ — ближайшее время освобождения канала СМО, то очередная заявка не может быть обслужена, т.е. получает отказ. (Если хотя бы одно место в очереди свободно, то $t_{\text{min,оч}}(r-1) \leq t_r$). В этом случае специальный счетчик числа отказов увеличивает свое значение на единицу.

При невыполнении условия (17.2) вычисляется момент ухода заявки из системы:

$$P(r) = \begin{cases} t_r + \tau_r, & t_{\text{min,оч}}(r-1) \leq t_r \leq t_{\text{min,ос}}(r-1); \\ t_{\text{min,ос}}(r-1) + \tau_r, & t_r > t_{\text{min,ос}}(r-1). \end{cases}$$

(Если хотя бы один из каналов обслуживания свободен, то $t_r < t_{\text{min,ос}}(r-1)$).

Далее определяется момент освобождения s-го канала после обслуживания теккущей r-й заявки:

$$o(s, r) = P(r),$$

и момент освобождения m-го места в очереди от r-й заявки:

$$e(m, r) = t_{\text{min,ос}}(r-1).$$

Затем вычисляются величины

$$t_{\text{min,ос}}(r) = \min_{1 \leq s \leq M} o(s, r),$$

где s — номер канала, M — число каналов;

$$t_{\min,\text{оч}}(r) = \min_{1 \le m \le N} e(m, r),$$

где m — номер места в очереди, N — число мест в очереди. Номера канала и места в очереди, на которых достигаются эти минимумы, обозначаются, соответственно, s^* и m^*.

Моделирование прекращается при выполнении условия

$$t_r > T_k \vee r > L.$$

При этом фиксируются R ($R < L$) — число шагов моделирования (фактическое число поступивших в СМО заявок) и V — число обслуженных заявок.

Наконец, параметры работы СМО и показатели эффективности ее работы рассчитываются на основе усредненных характеристик, полученных при моделировании.

Средняя продолжительность интервала между поступлениями заявок:

$$\overline{T}_{\text{и}} = \frac{t_L - t_1}{L - 1}.$$

Интенсивность поступления заявок в СМО:

$$\lambda = \frac{1}{\overline{T}_{\text{и}}}.$$

Среднее время обслуживания заявки:

$$\overline{\tau} = \frac{1}{V} \sum_{r=1}^{V} \tau_r.$$

Интенсивность обслуживания:

$$\mu = \frac{1}{\overline{\tau}}.$$

Коэффициент загрузки системы:

$$\rho = \lambda / \mu.$$

Абсолютная пропускная способность:

$$A = \frac{V}{t_R - T_0}.$$

Вероятность отказа:

$$p_{\text{отк}} = \frac{1}{R} \sum_{r=1}^{R} E\left[t_r - t_{\text{min,оч}}\left(r-1\right)\right],$$

где $E\left(x\right)$ единичная ступенчатая функция вида

$$E\left(x\right) = \begin{cases} 0, & x \leq 0; \\ 1, & x > 0. \end{cases}$$

Относительная пропускная способность:

$$Q = 1 - p_{\text{отк}}.$$

Среднее время ожидания в очереди:

$$\bar{t}_{\text{ож}} = \frac{1}{R} \sum_{r=1}^{R} \left[t_{\text{min,ос}}\left(r-1\right) - t_r\right] \cdot$$
$$\cdot E\left[t_r - t_{\text{min,оч}}\left(r-1\right)\right] E\left[t_{\text{min,ос}}\left(r-1\right) - t_r\right].$$

Среднее число занятых каналов:

$$\overline{k} = \frac{A}{\mu} = A\overline{\tau}.$$

Среднее число заявок в очереди:

$$\overline{r} = \lambda \bar{t}_{\text{ож}}.$$

Среднее число заявок в системе:

$$\overline{z} = \overline{k} + \overline{r} = \frac{A}{\mu} + \lambda \bar{t}_{\text{ож}}.$$

Среднее время пребывания заявки в системе:

$$\bar{t}_c = \frac{\overline{z}}{\lambda}.$$

Остается добавить, что в некоторых случаях можно существенно упростить моделирование работы СМО, заменяя многократные реализации ее работы однократной, но достаточно длинной реализацией [4]. Весьма полезные и общие советы по моделированию имеются в [20].

17.4 Потоки Эрланга

Потоки Эрланга моделируют выходящие из СМО потоки необслуженных заявок, а также потоки обслуживания в многоэтапных СМО [15]. ***Потоком Эрланга k-го порядка*** называется поток, который получается из простейшего потока, если в последнем сохранить каждое k-е событие, а промежуточные события выбросить. Таким образом, промежуток времени T между событиями в потоке Эрланга k-го порядка представляет собой сумму k независимых с.в. $T_1,\ldots,T_k$, каждая из которых является аналогичным промежутком времени для простейшего потока:

$$T = \sum_{i=1}^{k} T_i.$$

Но для простейшего потока каждый промежуток времени T_i имеет одно и то же показательное распределение с одним и тем же параметром λ. Поэтому в соответствии с примером 10.4 промежуток времени T будет иметь распределение Эрланга k-го порядка с тем же параметром λ.

Рассмотрим, как будет изменяться поток Эрланга k-го порядка при $k \to \infty$ и неизменной плотности потока λ. Для этого пронормируем T:

$$\widetilde{T} = \frac{T}{k}.$$

Такой поток называется ***нормированным потоком Эрланга k-го порядка***. Его распределение было изучено в п. 10.7. Там было показано, что дисперсия с.в. $\widetilde{T}$ стремится к 0 при $k \to \infty$. Это значит, что нормированный поток Эрланга стремится к ***регулярному потоку***, который характеризуется одинаковыми интервалами между событиями, равными $1/\lambda$. Задавая различные значения параметра k можно получать любую степень последействия: от полного его отсутствия ($k = 1$) до полной регулярности ($k = \infty$).

Часть IV

Математическая статистика

Лекция 18
Выборочные распределения

Математическая статистика — это наука о том, как по наблюдениям за случайными величинами приближенно находить их распределения и числовые характеристики; делать выводы, будет ли с.в. иметь тот или иной тип распределения; можно ли считать, что две случайные величины имеют одинаковые числовые характеристики и т.д.

Математическая статистика и теория вероятностей тесно связаны. Первая берет у второй весь ее математический аппарат для получения правильных, научных выводов, а вторая у первой — ту исходную информацию, основываясь на которой, она получает свои глубокие теоретические выводы об общих законах поведения случайных величин. Вспомните, что, решая задачи по теории вероятностей, вы читали в условиях: «С.в. X имеет плотность вероятностей ...» или «Математическое ожидание с.в. равно ...». А откуда бывает известно, что с.в. имеет то или иное распределение или те или иные числовые характеристики? Их, оказывается, можно получить методами математической статистики. Да и сам фундамент теории вероятностей, ее система аксиом, покоится на наблюдениях за относительной частотой с.с.

18.1 Предварительная обработка наблюдений

Все результаты, которые получает математическая статистика, основаны на наблюдениях за случайными величинами. Что они со-

бой представляют? Прежде всего надо понять, что все возможные реализации с.в. мы наблюдать не можем — их бесконечно много. Кстати, в математической статистике имеется своя терминология, немного отличающаяся от терминологии теории вероятностей. Так, все возможные значения с.в. X называются ее ***генеральной совокупностью***. Из всего этого множества мы можем в опытах со с.в. X получить лишь конечное число ее реализаций. Чем же считать эти реализации? Конечно, после совершения опыта они становятся обычными числами, то есть детерминированными величинами. Но до того как опыт проведен, его результаты заранее мы предсказать не можем, поэтому *для изучения и прогноза результатов опыта* естественно его результаты считать случайными величинами. Однако в каждом опыте мы будем наблюдать за поведением одной и той же с.в. X. Следовательно, наши наблюдения за с.в. X нужно считать с.в. $X_1, \ldots, X_n$, имеющими то же распределение, что и с.в. X. Кроме того, выводы математической статистики проще всего получать, если опыты проводятся независимо друг от друга (на практике при проведении опытов их независимость нужно специально обеспечивать). В результате мы получаем следующее определение. ***Выборкой*** из генеральной совокупности с.в. X называется множество независимых с.в. $X_1, \ldots, X_n$, имеющих то же самое распределение, что и с.в. X. С.в. X_i называется **i-*м наблюдением*** за с.в. X, а n — ***объемом выборки***. Выборка и есть тот исходный материал, из которого строится все здание математической статистики.

Полученная в опытах выборка может представлять собой довольно обширное и хаотическое нагромождение чисел, глядя на которые вряд ли можно сказать что-либо определенное о поведении изучаемой с.в. Поэтому выборка обычно подвергается так называемой первичной обработке, после чего она становится более наглядной и информативной. Прежде всего находят наименьшее X_{min}, наибольшее X_{max} значения выборки и разность между ними $X_{max} - X_{min}$, которую называют ***размахом выборки***. Далее выборку можно упорядочить по возрастанию, после чего ее называют ***вариационным рядом***. Вариационный ряд несет в себе гораздо больше информации,

чем неупорядоченная выборка, например, по нему можно оценить скорость возрастания наблюдений, их частоту и др.

Пример 18.1. *Количество отказов усилителей низкой частоты одного типа (по месяцам) за год дает следующую выборку*

$$8, 3, 10, 4, 3, 7, 9, 6, 2, 5, 4, 6.$$

Найти объем и размах выборки, построить для нее вариационный ряд.

Решение. Объем выборки совпадает с количеством месяцев в году: $n = 12$. Минимальное наблюдение равно 2, а максимальное — 10, поэтому размах выборки равен $X_{max} - X_{min} = 10 - 2 = 8$. Располагая наблюдения в порядке возрастания, получаем вариационный ряд:

$$2, 3, 3, 4, 4, 5, 6, 6, 7, 8, 9, 10.$$

□

При большом объеме выборки часто можно заметить, что в ней имеется много одинаковых наблюдений (это бывает, когда выборку получают из генеральной совокупности дискретной с.в. или точность измерений значений непрерывной с.в. невелика). В таком случае информацию, содержащуюся в выборке, можно значительно сжать, представив ее в виде ***статистического ряда дискретного типа***: $(X^{(i)}, n_i)$, $i = \overline{1, k}$, где $\{X^{(i)}\}$ — все различные наблюдения выборки, упорядоченные по возрастанию:

$$X_{min} = X^{(1)} < X^{(2)} < \ldots < X^{(k)} = X_{max},$$

а n_i — ***частота*** i-го наблюдения, то есть количество раз, которое наблюдение $X^{(i)}$ встречается в выборке. Очевидно, что

$$\sum_{i=1}^{k} n_i = n. \tag{18.1}$$

Пример 18.2. *Имеются наблюдения о числе звонков в справочное ж.д. бюро, зарегистрированных в разные дни с 7.00 до 8.00:*

4	6	4	5	2	0	7	8	4	3	4	7	4	6	5
$\boxed{12}$	2	4	5	5	3	1	2	3	6	6	4	6	5	9
5	7	3	4	7	2	5	5	7	7	8	5	4	4	4
4	7	8	7	11	4	10	7	4	3	5	5	6	2	2
2	6	11	5	2	5	4	7	5	7	7	2	5	4	3
3	4	6	7	4	6	4	$\boxed{0}$	3	4	1	5	6	3	5
5	5	3	10	3	5	3	4	5	3	3	9	4	3	9
5	6	8	6	5	5	2	5	1	6	3	11	3	9	5

Построить статистический ряд для данной выборки.

Решение. Из представленной таблицы, имеющей 8 строк и 15 столбцов, видно, что объем наблюдений равен $n = 8 \cdot 15 = 120$. Минимальное наблюдение $X_{min} = 0$, а максимальное $X_{max} = 12$, следовательно, размах выборки составляет $X_{max} - X_{min} = 12 - 0 = 12$. Подсчитывая частоты для различных наблюдений выборки, получаем требуемый статистический ряд:

$X^{(i)}$	0	1	2	3	4	5	6	7	8	9	10	11	12
n_i	2	3	10	17	22	26	13	13	4	4	2	3	1

□

И в случае выборки из генеральной совокупности непрерывной с.в., и в случае выборки из генеральной совокупности дискретной с.в., может оказаться так, что наблюдений в выборке будет много, а одинаковых среди них мало. Тогда выборку ***группируют***. Для этого отрезок $[X_{min}, X_{max}]$ разбивают на части (не обязательно одинаковые) точками $x_0 = X_{min}, x_1, \ldots, x_{k-1}, x_k = X_{max}$, и называют интервалы $(x_{i-1}, x_i), i = \overline{1,k}$ — ***интервалами группировки***. Затем подсчитывают количество наблюдений, попавших в каждый интервал, причем, если на границу между интервалами группировки

попало несколько наблюдений (например, 7), то «меньшую половину» (3 наблюдения) относят к левому интервалу, а «бóльшую» (4 наблюдения) — к правому. Ясно, что для четного числа наблюдений, попавших на границу между интервалами, эти «половины» равны. Наблюдения, совпавшие с левой границей первого или с правой границей последнего интервалов группировки, относят, соответственно, к этим интервалам. Количество наблюдений n_i, отнесенных к i-му интервалу группировки (x_{i-1}, x_i), называют ***частотой***.

Множество пар чисел вида

$$(\overline{x}_i, n_i),\ i = \overline{1, k},$$

где $\overline{x}_i = (x_{i-1} + x_i)/2$ — середина i-го интервала группировки, называется ***группирóванным статистическим рядом*** или ***статистическим рядом непрерывного типа***. Очевидно, и для этого ряда должно выполняться свойство (18.1).

Обычно число интервалов группировки бывает в пределах $10 \div 20$ и еще его можно вычислить по формуле

$$k = \lfloor 1 + 3{,}32 \lg n \rfloor, \tag{18.2}$$

где $\lfloor x \rfloor$ — наибольшее целое, меньшее либо равное x.

Пример 18.3. *Получена выборка (табл. 18.1) из отклонений напряжения (в мВ) от номинала*[†]*: Построить группированный статистический ряд с одинаковыми интервалами группировки.*

Решение. Объем выборки $n = 180$, $X_{min} = 7$, $X_{max} = 39$, размах выборки $= 32$. Для подсчета необходимого числа интервалов воспользуемся формулой (18.2): $k = \lfloor 1 + 3{,}32 \lg 180 \rfloor = 8$. Изобразим интервалы группировки на числовой оси (см. рис. 18.1) и распределим по ним наблюдения. Вместо чисел в качестве наблюдений будем пользоваться символами последнего столбца выборки, помещая в «ящик» над интервалом символ, соответствующий строке выборки, из которой

[†]Символы последнего столбца, конечно, не относятся к выборке, их смысл будет ясен из решения задачи

32	12	30	23	32	17	18	28	30	25	20	23	20	25	32	□
28	26	32	26	27	20	24	17	30	11	35	30	27	21	27	◇
26	19	25	32	30	35	21	21	27	30	23	30	20	33	26	△
21	15	32	$\boxed{39}$	17	21	33	27	30	23	15	28	29	20	39	♣
26	22	27	22	35	23	27	31	29	23	27	31	17	22	31	★
27	31	17	30	31	19	29	24	36	23	35	28	21	29	$\boxed{7}$	♡
22	23	22	34	10	21	23	25	16	32	30	20	27	25	27	♠
26	23	23	28	21	22	21	28	18	18	31	23	18	29	22	⊕
21	23	31	22	23	33	19	25	20	26	18	20	25	23	14	⊙
28	26	18	21	30	31	29	19	23	22	24	29	23	27	27	⊗
25	34	30	25	24	25	24	14	19	14	27	15	27	32	26	◁
18	19	22	19	22	28	22	17	25	23	25	28	17	34	23	▷

Таблица 18.1: Выборка из отклонений напряжения от номинала

берется распределяемое наблюдение. Символы, попадающие точно на границу между интервалами, будем «нанизывать» на перегородки между «ящиками». Использование для каждой строки выборки своего символа позволяет проконтролировать правильность распределения наблюдений: так как в строке выборки содержится пятнадцать наблюдений, то каждый символ на рис. 18.1 должен встречаться ровно пятнадцать раз, легко подсчитать общее количество наблюдений на рис. и т.д.

Наблюдения, попавшие на границы между интервалами, распределяют по ним описанным выше способом. Так, например, на границу между интервалами $(11; 15)$ и $(15; 19)$ попало три наблюдения, поэтому к первому интервалу отнесено одно наблюдение, а ко второму — два наблюдения. Всего в интервал $(11; 15)$ попало 6 наблюдений: одно с его левой границы, одно — с правой и четыре наблюдения

попали вовнутрь интервала. В результате для каждого интервала группировки получены частоты, показанные на рис. 18.1 числами, заключенными в квадратики.

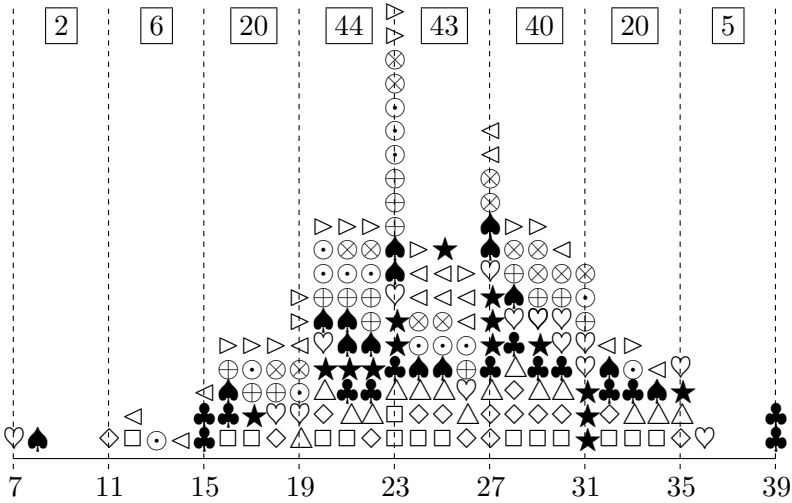


Рис. 18.1: Группировка наблюдений по интервалам

Следовательно, статистический ряд для данной группированной выборки имеет вид:

$\overline{x}_i$	9	13	17	21	25	29	33	37
n_i	2	6	20	44	43	40	20	5

□

Кроме частот, в математической статистике применяют следующие, связанные с ними величины: ***относительные частоты***

$$\frac{n_i}{n},\ i = \overline{1,k},$$

накопленные частоты

$$\sum_{j=1}^{i} n_j,\ i = \overline{1,k},$$

и ***накопленные относительные частоты***

$$\sum_{j=1}^{i} \frac{n_j}{n}, \ i = \overline{1, k}.$$

Эти величины удобно представлять в виде ***таблицы частот***. Табл. 18.2 является таблицей частот для примера 18.2, а табл. 18.3 — для примера 18.3.

18.2 Аппроксимация законов распределения

Если наблюдений достаточно много, по ним можно получить приближенное представление о распределении с.в. X, из генеральной совокупности которой получена выборка $X_1, X_2, \ldots, X_n$. Например, функция распределения (теперь мы будем называть ее ***теоретической***) $F(x)$ с.в. X аппроксимируется ***эмпирической*** функцией распределения

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

где $\mu_n(x)$ — число наблюдений выборки, меньших x.

Если выборка задана статистическим рядом дискретного типа, то эмпирическая функция распределения определяется равенством

$$\boxed{F_n(x) = \begin{cases} 0, & x \le X^{(1)}, \\ \frac{1}{n} \sum\limits_{j=1}^{i} n_j, & X^{(i)} < x \le X^{(i+1)}, \quad i = \overline{1, k-1}, \\ 1, & x > X^{(k)}. \end{cases}}$$

Действительно, число наблюдений, меньших $X^{(1)}$, равно 0, а число наблюдений, больших $X^{(i)}$, но меньших $X^{(i+1)}$, равно сумме частот, соответствующих наблюдениям от $X^{(1)}$ до $X^{(i)}$.

Для группированной выборки в качестве эмпирической функции

Таблица 18.2: Таблица частот для примера 18.2

Наблюдение, $X^{(i)}$	Частота, n_i	Относительная частота, $\frac{n_i}{n}$	Накопленная частота, $\sum\limits_{j=1}^{i} n_j$	Накопл. относит. частота, $\frac{1}{n}\sum\limits_{j=1}^{i} n_j$
0	2	0,0167	2	0,0167
1	3	0,0250	5	0,0417
2	10	0,0833	15	0,1250
3	17	0,1417	32	0,2667
4	22	0,1833	54	0,4500
5	26	0,2167	80	0,6667
6	13	0,1083	93	0,7750
7	13	0,1083	106	0,8833
8	4	0,0333	110	0,9166
9	4	0,0333	114	0,9499
10	2	0,0167	116	0,9666
11	3	0,0250	119	0,9916
12	1	0,0083	120	0,9999

Таблица 18.3: Таблица частот для примера 18.3

Номер интервала, i	Границы интервала (x_{i-1}, x_i)	Середина интервала, $\bar{x}_i$	Частота, n_i	Относительная частота, $\frac{n_i}{n}$	Накопленная частота $\sum\limits_{j=1}^{i} n_j$	Накопл. относит. частота $\frac{1}{n}\sum\limits_{j=1}^{i} n_j$
1	$7 \div 11$	9	2	0,0111	2	0,0111
2	$11 \div 15$	13	6	0,0333	8	0,0444
3	$15 \div 19$	17	20	0,1111	28	0,1556
4	$19 \div 23$	21	44	0,2444	72	0,4000
5	$23 \div 27$	25	43	0,2389	115	0,6389
6	$27 \div 31$	29	40	0,2222	155	0,8611
7	$31 \div 35$	33	20	0,1111	175	0,9722
8	$35 \div 39$	37	5	0,0278	180	1,0000

распределения принимают следующую функцию:

$$\boxed{F_n(x) = \begin{cases} 0, & x \le \overline{x}_1, \\ \sum\limits_{j=1}^{i} \frac{n_j}{n}, & \overline{x}_i < x \le \overline{x}_{i+1}, \quad i = \overline{1, k-1}, \\ 1, & x > \overline{x}_k. \end{cases}}$$

Из представленных формул видно, что эмпирические функции распределения являются кусочно-постоянными функциями (даже для непрерывных с.в., имеющих, как известно, непрерывные теоретические функции распределения). И тем не менее справедлива

Теорема 18.1. *Для любого x при $n \to \infty$ эмпирическая функция распределения с.в. X сходится по вероятности к ее теоретической функции распределения: $F_n(x) \xrightarrow{P} F(x)$.*

Аналогичную роль играет ***полигон относительных накопленных частот***, который является ломаной, соединяющей точки

$$\boxed{\left(X^{(i)}, \sum_{j=1}^{i} \frac{n_j}{n} \right), \ i = \overline{1, k},}$$

для статистического ряда дискретного типа или точки

$$\boxed{\left(\overline{x}_i, \sum_{j=1}^{i} \frac{n_j}{n} \right), \ i = \overline{1, k},}$$

для группированной выборки и в случае непрерывной функции распределения более напоминающий теоретическую функцию распределения, чем эмпирическая функция распределения.

Графики эмпирических функций распределения для выборок из примеров 18.2 и 18.3 представлены на рис. 18.2 и рис. 18.3, соответственно. Отметим, что значения этих функций брались из последних

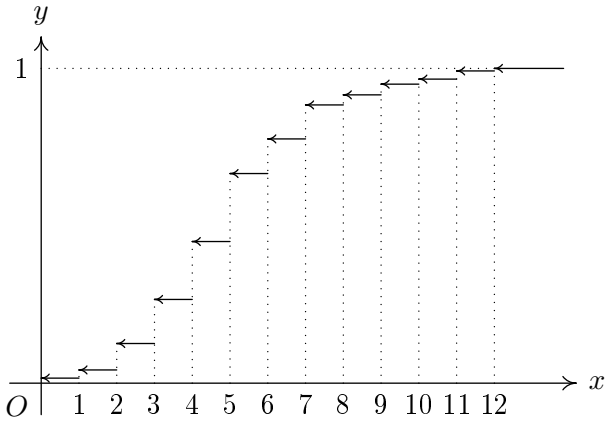


Рис. 18.2: Эмпирическая функция распределения (пример 18.2)

столбцов таблиц 18.2 и 18.3, соответственно. На рис. 18.3 для сравнения показана и теоретическая функция распределения $F(x)$. Полигон относительных накопленных частот для примера 18.3 построен на рис. 18.4 также вместе с теоретической функцией распределения. Видно хорошее согласование выборочного и теоретического распределений.

18.3 Гистограмма

Плотность вероятностей непрерывной с.в. может быть приближенно представлена с помощью ступенчатой фигуры, составленной из прямоугольников и называемой ***гистограммой*** выборки. Основанием i-го прямоугольника служит i-й интервал группировки, а высотой — частное, получаемое от деления относительной частоты, соответствующей этому интервалу, на его длину:

$$\boxed{h_i = \frac{n_i}{n(x_i - x_{i-1})}.}$$

Нетрудно видеть, что площадь гистограммы равна 1.

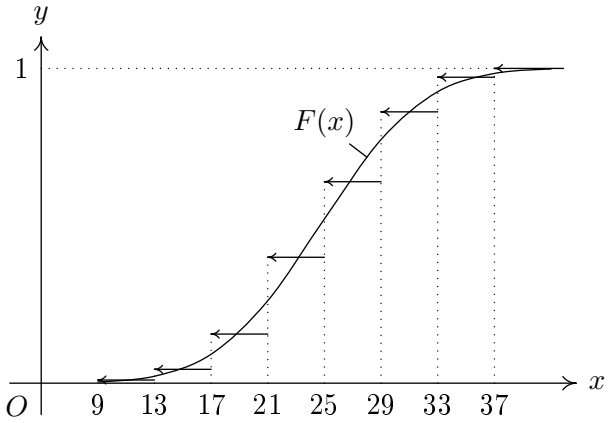


Рис. 18.3: Эмпирическая функция распределения (пример 18.3)

При больших n и удачном выборе интервалов группировки верхняя часть гистограммы имеет форму, напоминающую график плотности вероятностей с.в. X. Часто вместо гистограммы выборки строят ломаную с вершинами в точках $(X^{(i)}, n_i/n)$, $i = \overline{1,k}$, для статистического ряда дискретного типа и в точках $(\overline{x}_i, n_i/n)$, $i = \overline{1,k}$, если выборка группированная. Эту ломаную называют ***полигоном относительных частот***. Гистограмма выборки для примера 18.3 показана на рис. 18.5, а полигон относительных частот для примера 18.2 — на рис. 18.6.

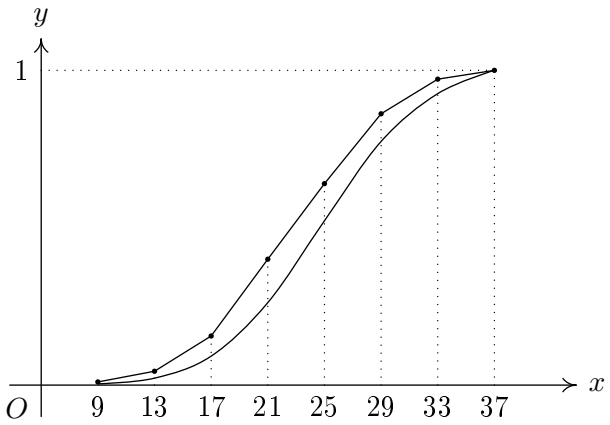


Рис. 18.4: Полигон относительных накопленных частот (пример 18.3)

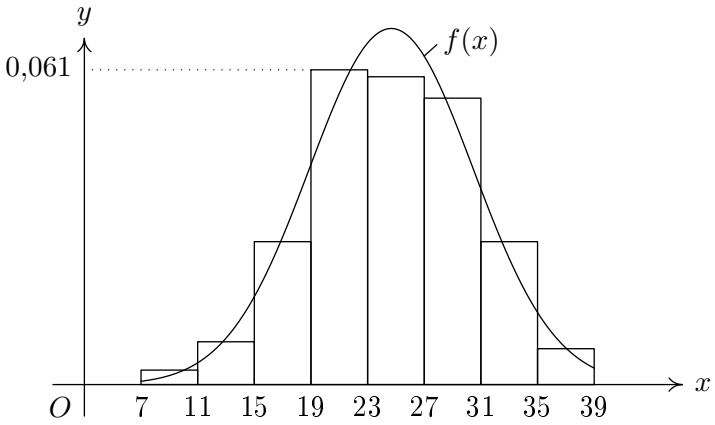


Рис. 18.5: Гистограмма выборки (пример 18.3)

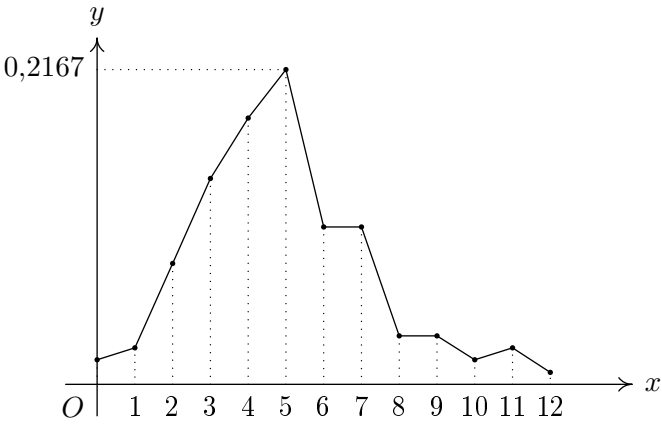


Рис. 18.6: Полигон относительных частот (пример 18.2)

Лекция 19
Точечные оценки

Записывая формулы для законов распределений с.в., мы видели, что в них фигурируют величины, которые называют ***параметрами*** распределений. Например, m и σ — параметры нормального распределения, а λ — параметр пуассоновского распределения. Числовые характеристики с.в. (математическое ожидание, дисперсию и др.) тоже можно отнести к параметрам распределения. Величины, полученные на основе выборочных данных, значения которых приближенно равны параметрам распределения, называют ***оценками*** параметров.

Пусть $X_1, \ldots, X_n$ — выборка из генеральной совокупности с.в. X и a — неизвестный параметр ее распределения. ***Точечной оценкой*** параметра a называется с.в. $a_n = g(X_1, \ldots, X_n)$, где g — некоторая функция. Чтобы оценка была «хорошей», нужно, чтобы она удовлетворяла некоторым критериям. Таких критериев можно насчитать не меньше десятка, но мы выберем два самых важных и распространенных.

Во-первых, оценка должна быть ***несмещенной*** , то есть ее математическое ожидание должно совпадать с истинным значением параметра: $\mathbf{M}a_n = a$.

Во-вторых, она должна быть ***состоятельной***. Это значит, что при неограниченном увеличении выборки она должна по вероятности сходиться к параметру a: $a_n \xrightarrow{P} a$.

Отметим, что для состоятельности оценки достаточно, чтобы вы-

полнялись равенства:

$$\lim_{n\to\infty} \mathbf{M}a_n = a, \quad \lim_{n\to\infty} \mathbf{D}a_n = 0.$$

19.1 Выборочное среднее

Для простой выборки ***выборочным средним*** называется с.в.

$$\boxed{\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i.} \tag{19.1}$$

Для статистического ряда дискретного типа эта формула эквивалентна следующей:

$$\boxed{\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i X^{(i)},} \tag{19.2}$$

а для группированной выборки в качестве выборочного среднего принимают

$$\boxed{\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i \overline{x}_i.} \tag{19.3}$$

Теорема 19.1. *Выборочное среднее является несмещенной и состоятельной оценкой математического ожидания с.в. X, причем,*

$$\mathbf{M}\overline{x} = \mathbf{M}X, \quad \mathbf{D}\overline{x} = \frac{1}{n}\mathbf{D}X.$$

Доказательство. Докажем это утверждение для выборочного среднего, определяемого формулой (19.1), для чего возьмем математическое ожидание от обеих ее частей (так как наблюдения X_i распределены так же, как и с.в. X, то их числовые характеристики совпадают с числовыми характеристиками этой с.в.):

$$\mathbf{M}\overline{x} = \mathbf{M}\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n}\mathbf{M}X_i = \mathbf{M}X\frac{1}{n}n = \mathbf{M}X.$$

Следовательно, величина $\overline{x}$ является несмещенной оценкой математического ожидания с.в. X.

Теперь вычислим дисперсию выборочного среднего, памятуя о том, что наблюдения являются независимыми с.в. и поэтому дисперсия их суммы равна сумме их дисперсий:

$$\mathbf{D}\overline{x} = \mathbf{D}\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{D}X_i = \mathbf{D}X\frac{1}{n^2}n = \frac{1}{n}\mathbf{D}X.$$

Следовательно, $\mathbf{D}\overline{x} \underset{n\to\infty}{\longrightarrow} 0$. Этот факт в совокупности с несмещенностью выборочного среднего означает его состоятельность. □

Пример 19.1. *Найти выборочное среднее для с.в. отказов усилителей из примера 18.1.*

Решение. Так как в данном случае мы имеем дело с простой выборкой, то для вычисления выборочного среднего применим формулу (19.1):

$$\overline{x} = \frac{1}{12}(8+3+10+4+3+7+9+6+2+5+4+6) = \frac{67}{12} = 5{,}58.$$

Значит, в среднем за месяц отказывает $5 \div 6$ усилителей. □

Пример 19.2. *Найти выборочное среднее для с.в. количества звонков, поступивших в ж.д. справочную, из примера 18.2.*

Решение. Применим формулу (19.2) для статистического ряда дискретного типа, полученного в примере 18.2:

$$\begin{aligned}\overline{x} = \frac{1}{120}(&2\cdot 0 + 3\cdot 1 + 10\cdot 2 + 17\cdot 3 + 22\cdot 4 + 26\cdot 5 + 13\cdot 6 + \\ &+ 13\cdot 7 + 4\cdot 8 + 4\cdot 9 + 2\cdot 10 + 3\cdot 11 + 1\cdot 12) = \\ = \frac{594}{120} &= 4{,}95.\end{aligned}$$

Таким образом, в среднем с 7.00 до 8.00 в справочную поступает около пяти звонков. □

Пример 19.3. *Найти выборочное среднее для с.в. отклонения напряжения от номинала из примера 18.3.*

Решение. Здесь нужно использовать формулу (19.3) для группированного статистического ряда, полученного в примере 18.3:

$$\overline{x} = \frac{1}{180}(2 \cdot 9 + 6 \cdot 13 + 20 \cdot 17 + 44 \cdot 21 + 43 \cdot 25 + 40 \cdot 29 + 20 \cdot 33 + \\ + 5 \cdot 37) = \frac{4440}{180} = 24{,}667.$$

На такую величину напряжение в среднем отклоняется от номинала. □

19.2 Выборочная дисперсия

Выборочная дисперсия является точечной оценкой дисперсии с.в. X. Для простой выборки она вычисляется по формуле

$$\boxed{s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \overline{x}^2,} \tag{19.4}$$

для выборки дискретного типа это эквивалентно выражению

$$\boxed{s^2 = \frac{1}{n}\sum_{i=1}^{k}n_i\left(X^{(i)} - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{k}n_i\left(X^{(i)}\right)^2 - \overline{x}^2,} \tag{19.5}$$

а для группированной выборки — определяется формулой

$$\boxed{s^2 = \frac{1}{n}\sum_{i=1}^{k}n_i(\overline{x}_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{k}n_i\overline{x}_i^2 - \overline{x}^2.} \tag{19.6}$$

Можно показать, что выборочная дисперсия является состоятельной, но смещенной оценкой дисперсии с.в. X. При больших объемах выборок смещение невелико и им можно пренебречь, а для малых

выборок лучше пользоваться так называемой ***исправленной выборочной дисперсией***, которая отличается от выборочной дисперсии только множителем:

$$\boxed{\widetilde{s}^2 = \frac{n}{n-1}s^2,} \tag{19.7}$$

но он делает исправленную выборочную дисперсию и несмещенной, и состоятельной оценкой.

Извлекая квадратные корни из выборочных дисперсий, получаем, соответственно, ***выборочное с.к.о.*** s и ***исправленное выборочное с.к.о.*** $\widetilde{s}$.

Пример 19.4. *Найти выборочную дисперсию, исправленную выборочную дисперсию, выборочное с.к.о. и исправленное выборочное с.к.о. для с.в. отказов усилителей из примера 18.1.*

Решение. Вычисление выборочной дисперсии для простой выборки состоит в применении формулы (19.4), причем, используется выборочное среднее, полученное в примере 19.1:

$$s^2 = \frac{1}{12}(8^2+3^2+10^2+4^2+3^2+7^2+9^2+6^2+2^2+5^2+4^2+6^2)-$$
$$-\left(\frac{67}{12}\right)^2 = \frac{445}{12} - \left(\frac{67}{12}\right)^2 = 5{,}91.$$

Теперь нетрудно получить и остальные оценки:

$$\widetilde{s}^2 = \frac{12}{11}s^2 = 6{,}45; \quad s = \sqrt{s^2} = 2{,}43; \quad \widetilde{s} = \sqrt{\widetilde{s}^2} = 2{,}54.$$

Видно, что выборочная и исправленная выборочные дисперсии заметно отличаются друг от друга. □

Пример 19.5. *Решить предыдущую задачу для с.в. количества звонков, поступивших в ж.д. справочную, из примера 18.2.*

Решение. Применим формулу (19.5) для статистического ряда дискретного типа, полученного в примере 18.2:

$$s^2 = \frac{1}{120}(2\cdot 0^2 + 3\cdot 1^2 + 10\cdot 2^2 + 17\cdot 3^2 + 22\cdot 4^2 + 26\cdot 5^2 + 13\cdot 6^2 + \\ + 13\cdot 7^2 + 4\cdot 8^2 + 4\cdot 9^2 + 2\cdot 10^2 + 3\cdot 11^2 + 1\cdot 12^2) - \\ - \left(\frac{594}{120}\right)^2 = \frac{3590}{120} - \left(\frac{99}{20}\right)^2 = 5{,}41;$$

$$\widetilde{s}^{\,2} = \frac{120}{119}s^2 = 5{,}46; \quad s = \sqrt{s^2} = 2{,}33; \quad \widetilde{s} = \sqrt{\widetilde{s}^{\,2}} = 2{,}34.$$

По сравнению с предыдущим примером объем выборки значительно больше и дисперсии (выборочная и исправленная выборочная) отличаются меньше. □

Пример 19.6. *Решить такую же задачу для с.в. отклонения напряжения от номинала из примера 18.3.*

Решение. На этот раз используем формулу (19.6) для группированного статистического ряда, полученного в примере 18.3:

$$s^2 = \frac{1}{180}(2\cdot 9^2 + 6\cdot 13^2 + 20\cdot 17^2 + 44\cdot 21^2 + 43\cdot 25^2 + 40\cdot 29^2 + \\ + 20\cdot 33^2 + 5\cdot 37^2) - \left(\frac{4440}{180}\right)^2 = \\ = \frac{115500}{180} - \left(\frac{74}{3}\right)^2 = 33{,}222;$$

$$\widetilde{s}^{\,2} = \frac{180}{179}s^2 = 33{,}408; \quad s = \sqrt{s^2} = 5{,}764; \quad \widetilde{s} = \sqrt{\widetilde{s}^{\,2}} = 5{,}780.$$

□

19.3 Упрощение ручных расчетов

Представьте себе, что вам нужно найти выборочное среднее для роста студентов вашей группы, то есть, попросту говоря, средний рост

студента группы АТС. Вам придется складывать около тридцати трехзначных чисел, а затем полученную сумму делить на число студентов. Вручную делать это довольно сложно. В то же время ясно, что рост может колебаться примерно от 160 до 180 см и, таким образом, возможно, средний рост приблизительно равен 170 см. Эти 170 см будем называть ***ложным средним***, так как выборочное среднее может с ним не совпасть. Но зато ложный средний рост можно использовать для того, чтобы вычислить отклонения от него (а они будут всего лишь двузначными числами да еще и разных знаков, так что при сложении вполне возможно будут взаимно уничтожаться) и найти их выборочное среднее. Теперь остается лишь догадаться, что выборочное среднее отклонений достаточно прибавить к ложному среднему, чтобы получить средний рост студента в группе. Все эти рассуждения основаны на формуле

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - x^*) + x^*,$$

правильность которой очевидна и в которой x^* — ложное среднее.

Аналогичные рассуждения можно применить и к вычислениям выборочных числовых характеристик по статистическим рядам. Общая формула преобразования исходной выборки $X_1, \ldots, X_n$ в выборку $U_1, \ldots, U_n$ для упрощенных расчетов имеет вид

$$U_i = \frac{\widetilde{X}_i - x^*}{\Delta},\ i = \overline{1, k}.$$

Для простой выборки $k = n$, $\Delta = 1$, $\widetilde{X}_i = X_i$, а x^* выбирается равным какому-нибудь наблюдению в «середине» выборки, то есть близким к числу $(X_{max} + X_{min})/2$.

Для статистического ряда дискретного типа k равно числу различных наблюдений в выборке, $\Delta = 1$, $\widetilde{X}_i = X^{(i)}$, а x^* выбирается, как в предыдущем случае.

Для группированного статистического ряда, у которого все интервалы имеют одинаковую длину Δ, а их количество равно k, ве-

личины $\widetilde{X}_i$ заменяют серединами интервалов $\overline{x}_i$, а x^* обычно принимают равным середине интервала, которому соответствует максимальная частота n_i (как правило, середина такого интервала близка к выборочному среднему).

Далее находят выборочное среднее $\overline{u}$ и дисперсию s_U^2 для выборки $U_1, \ldots, U_n$, а затем аналогичные характеристики для исходной выборки по формулам

$$\overline{x} = \Delta\overline{u} + x^*, \quad s_X^2 = \Delta^2 s_U^2.$$

При этом вычисления удобно располагать в специальную таблицу, как это показано в следующем примере.

Пример 19.7. *Найти выборочные числовые характеристики для выборки из примера 18.3.*

Решение. Статистический ряд, полученный в примере 18.3, показывает, что максимальная частота соответствует интервалу $(19; 23)$, так что его середину 21 примем в качестве ложного среднего, а длину интервала 4 — в качестве величины Δ. Дальнейшие преобразования по формуле

$$U_i = \frac{\overline{x}_i - 21}{4}, \ i = \overline{1, 8},$$

и вычисление выборочных числовых характеристик для преобразованной выборки представим в виде таблицы 19.1. Используя сумму элементов последнего столбца, деленную на объем выборки, получим выборочную дисперсию с.в. U:

$$s_U^2 = \frac{1}{n}\sum_{i=1}^{k} n_i U_i^2 - \overline{u}^2 = 2{,}91(6) - 0{,}91(6)^2 = 2{,}0763(8).$$

Далее находим выборочные характеристики заданной с.в.:

$$\overline{x} = 4 \cdot 0{,}91(6) + 21 \approx 24{,}667; \ s_X^2 = 4^2 \cdot 2{,}763(8) \approx 33{,}222.$$

□

Таблица 19.1: Таблица для ручных расчетов

n_i	U_i	n_iU_i	U_i^2	$n_iU_i^2$
2	-3	-6	9	18
6	-2	-12	4	24
20	-1	20	1	20
44	0	0	0	0
43	1	43	1	43
40	2	80	4	160
20	3	60	9	180
5	4	20	16	80
Σ		165		525
$\frac{1}{n}\Sigma$		$\overline{u} = 0{,}91(6)$		$2{,}91(6)$

19.4 Метод наибольшего правдоподобия

Данный метод тоже дает точечные оценки параметров распределения с.в., если известен закон ее распределения.

Пусть $X_1, \ldots, X_n$ — выборка из генеральной совокупности с.в. X. Если эта с.в. непрерывна и имеет плотность вероятностей $f(x, \mathbf{a})$, где $\mathbf{a} = (a_1, \ldots, a_s)$ — вектор неизвестных параметров распределения, то функция

$$\boxed{L(X_1, \ldots, X_n, \mathbf{a}) = f(X_1, \mathbf{a}) \cdot \ldots \cdot f(X_n, \mathbf{a})} \tag{19.8}$$

называется ***функцией правдоподобия***.

Если X — дискретная с.в., то ***функцией правдоподобия*** называется функция

$$\boxed{L(X_1, \ldots, X_n, \mathbf{a}) = p(X_1, \mathbf{a}) \cdot \ldots \cdot p(X_n, \mathbf{a}),} \tag{19.9}$$

где

$$p(x, \mathbf{a}) = P(X = x).$$

Идея получения точечных оценок методом наибольшего правдоподобия проще всего иллюстрируется для дискретных с.в. Формула (19.9) с учетом независимости наблюдений выборки представляет собой вероятность того, что в n опытах реализуются именно те значения с.в. X, которые наблюдались в выборке. Очевидно, что маловероятные значения с.в. вряд ли могли составить выборку, поэтому вероятность появления наблюдавшихся значений следует считать достаточно высокой, а, значит, значения неизвестных параметров $a_1, \ldots, a_s$ должны быть такими, чтобы функция правдоподобия (19.9) имела достаточно большое значение. Чтобы не гадать, насколько большой она должна быть, принимают, что она должна быть *максимально большой.* Несколько более сложные рассуждения приводят к тому же результату и в случае непрерывной с.в.

В итоге ***оценкой*** $\mathbf{a_n}$ вектора неизвестных параметров $\mathbf{a}$ с.в. X ***по методу наибольшего правдоподобия*** называют решение системы уравнений

$$\frac{\partial \ln L}{\partial a_1} = 0, \cdots, \frac{\partial \ln L}{\partial a_n} = 0 \qquad (19.10)$$

относительно неизвестных $a_1, \ldots, a_s$. Действительно, если отбросить в этой системе логарифмы, то она будет представлять собой необходимые условия экстремума для функции нескольких переменных. Зачем же тогда добавлены еще и логарифмы? Да просто потому, что введение логарифмов не влияет на результат решения системы (19.10), а вот сам процесс ее решения при наличии логарифмов значительно упрощается.

Пример 19.8. *По выборке $X_1, \ldots, X_n$ методом наибольшего правдоподобия найти оценки параметров m, σ с.в. X, имеющей нормальную плотность распределения:*

$$f(x, m, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(x-m)^2}{2\sigma^2}\right].$$

Решение. Составим функцию правдоподобия в соответствии с формулой (19.8):

$$L(X_1, \ldots, X_n, m, \sigma) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\sum_{k=1}^{n} \frac{(X_k - m)^2}{2\sigma^2}\right].$$

Возьмем логарифмы от обеих частей этого равенства:

$$\ln L = -n\left[\frac{1}{2}\ln(2\pi) + \ln\sigma\right] - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(X_k - m)^2,$$

и составим систему уравнений (19.10):

$$\frac{\partial \ln L}{\partial m} = \frac{1}{\sigma^2}\sum_{k=1}^{n}(X_k - m) = 0,$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{k=1}^{n}(X_k - m)^2 = 0.$$

Из первого уравнения этой системы нетрудно найти оценку параметра m:

$$m = \frac{1}{n}\sum_{k=1}^{n} X_k = \overline{x},$$

а из второго — оценку параметра σ:

$$\sigma^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{x})^2 = s^2.$$

Как видите, оценки оказались уже знакомыми нам выборочным средним и выборочной дисперсией, соответственно. □

Пример 19.9. *По выборке $X_1, \ldots, X_n$ методом наибольшего правдоподобия найти оценку параметра p с.в. X, имеющей геометрическое распределение:*

$$p_k = P(X = k) = (1-p)^{k-1}p, \quad k = 1, 2, \ldots\ .$$

Решение. Геометрическое распределение возникает в задаче о стрелке, который стреляет в цель до тех пор, пока не попадет; с.в. X — число выстрелов, которые придется сделать стрелку, p — вероятность попадания при одном выстреле.

Составим функцию правдоподобия вида (19.9):

$$\begin{aligned} L(X_1, \ldots, X_n, p) &= (1-p)^{X_1-1}p \cdot \ldots \cdot (1-p)^{X_n-1}p = \\ &= (1-p)^{X_1+\ldots+X_n-n}p^n, \end{aligned}$$

а затем возьмем от нее логарифм:

$$\ln L = (X_1 + \ldots + X_n - n)\ln(1-p) + n\ln p.$$

Продифференцируем полученную функцию и приравняем производную к нулю:

$$\frac{\partial \ln L}{\partial p} = -\frac{X_1 + \ldots + X_n - n}{1-p} + \frac{n}{p} = 0.$$

Приводим к общему знаменателю, раскрываем скобки и получаем требуемую оценку:

$$p = \frac{n}{X_1 + \ldots + X_n} = \frac{1}{\overline{x}},$$

которая оказывается величиной, обратной выборочному среднему. □

Лекция 20
Интервальные оценки

При малом объеме выборки желательно иметь не только точечную оценку параметра, но и меру погрешности такой оценки (при большом объеме выборки точность обычно бывает удовлетворительной). Эту задачу решает интервальное оценивание параметров, которое дает оценку в виде интервала, в котором с заданной (достаточно высокой) вероятностью содержится истинное значение неизвестного параметра. В качестве точечной оценки параметра в этом случае обычно берут середину интервала.

Более строго, пусть у нас имеется выборка $X_1, \ldots, X_n$ из генеральной совокупности с.в. X, a — неизвестный параметр ее распределения. Интервал (a_1, a_2) с концами, определяемыми наблюдениями выборки $a_1 = g_1(X_1, \ldots, X_n)$, $\quad a_2 = g_2(X_1, \ldots, X_n)$, где g_1, g_2 — некоторые функции, называется ***доверительным интервалом***, соответствующим заданной ***доверительной вероятности (надежности)*** ε, если выполняется соотношение

$$\boxed{P(a_1 < a < a_2) = \varepsilon.} \tag{20.1}$$

Так как доверительная вероятность ε берется достаточно большой (обычно 0,9, 0,95 или 0,99), то в силу формулы (20.1) доверительный интервал (a_1, a_2) дает хорошее представление о том диапазоне значений, в котором может лежать истинное значение неизвестного параметра. Однако за все надо платить: так как интервальное оценивание дает и точечную оценку параметра, и интервал, в ко-

тором с большой достоверностью находится параметр, за это приходится расплачиваться как довольно сложными выкладками при получении оценок, так и тем, что интервальные оценки в отличие от точечных не являются не зависящими от типа распределения: для каждого типа распределения приходится получать свои интервальные оценки. Мы познакомимся с интервальным оцениванием на примере нормального распределения, поэтому с.в. X у нас будет иметь распределение $N(m, \sigma)$ и в дальнейшем это специально не оговаривается.

20.1 Оценка $\mathbf{M}X$ при известной дисперсии

Пусть известна дисперсия $\mathbf{D}X = \sigma^2$ с.в. X. Введем вспомогательную с.в. (такую с.в. называют ***статистикой***)

$$Y = \frac{\mathbf{M}X - \overline{x}}{\sigma/\sqrt{n}}, \tag{20.2}$$

которая в силу нормальной распределенности с.в. X тоже оказывается распределенной по нормальному закону. Так как на предыдущей лекции было доказано, что $\mathbf{M}\overline{x} = \mathbf{M}X$, то математическое ожидание с.в. Y равно

$$\mathbf{M}Y = \mathbf{M}\frac{\mathbf{M}X - \overline{x}}{\sigma/\sqrt{n}} = \frac{\sqrt{n}}{\sigma}(\mathbf{M}X - \mathbf{M}X) = 0.$$

Используя свойства дисперсии, находим

$$\mathbf{D}Y = \mathbf{D}\frac{\mathbf{M}X - \overline{x}}{\sigma/\sqrt{n}} = \frac{n}{\sigma^2}\mathbf{D}(-\overline{x}) = \frac{\mathbf{D}X}{\sigma^2} = 1,$$

потому что $\mathbf{D}(-\overline{x}) = \mathbf{D}\overline{x} = \frac{1}{n}\mathbf{D}X$ (см. предыдущую лекцию).

Таким образом, мы получили, что с.в. Y распределена $N(0; 1)$. Этот факт дает нам право записать вероятность ее попадания в ин-

тервал (t_1, t_2) следующим образом:

$$P(t_1 < Y < t_2) = P\left(t_1 < \frac{\mathbf{M}X - \overline{x}}{\sigma/\sqrt{n}} < t_2\right) =$$
$$= P\left(t_1\frac{\sigma}{\sqrt{n}} < \mathbf{M}X - \overline{x} < t_2\frac{\sigma}{\sqrt{n}}\right) = \tag{20.3}$$
$$= P\left(\overline{x} + t_1\frac{\sigma}{\sqrt{n}} < \mathbf{M}X < \overline{x} + t_2\frac{\sigma}{\sqrt{n}}\right) = \Phi^*(t_2) - \Phi^*(t_1).$$

Приравнивая полученное соотношение к заданной надежности ε, получим:

$$P\left(\overline{x} + t_1\frac{\sigma}{\sqrt{n}} < \mathbf{M}X < \overline{x} + t_2\frac{\sigma}{\sqrt{n}}\right) = \Phi^*(t_2) - \Phi^*(t_1) = \varepsilon.$$

Из этой формулы видно, что доверительный интервал для математического ожидания можно взять в виде

$$(m_1, m_2) = \left(\overline{x} + t_1\frac{\sigma}{\sqrt{n}}, \overline{x} + t_2\frac{\sigma}{\sqrt{n}}\right).$$

Однако таких интервалов существует бесконечно много (при различных значениях t_1 и t_2), поэтому используя симметрию плотности распределения с.в. Y (распределенной $N(0;1)$) относительно оси ординат для определенности берут интервал (t_1, t_2), симметричный относительно нуля, то есть полагают $(t_1, t_2) = (-t, t)$. Тогда формула (20.1) превращается в соотношение

$$P(-t < Y < t) = P\left(\overline{x} - t\frac{\sigma}{\sqrt{n}} < \mathbf{M}X < \overline{x} + t\frac{\sigma}{\sqrt{n}}\right) = \Phi^*(t) - \Phi^*(-t),$$

из которого, во-первых, следует, что окончательно доверительный интервал будет иметь вид

$$\boxed{(m_1, m_2) = \left(\overline{x} - t\frac{\sigma}{\sqrt{n}}, \overline{x} + t\frac{\sigma}{\sqrt{n}}\right),} \tag{20.4}$$

а, во-вторых, что неизвестное значение параметра t можно найти из равенства:

$$\Phi^*(t) - \Phi^*(-t) = \varepsilon. \tag{20.5}$$

Действительно, учитывая свойство функции $\Phi^*(t)$: $\Phi^*(-t) = 1 - -\Phi^*(t)$, получаем

$$\Phi^*(t) - (1 - \Phi^*(t)) = \varepsilon$$

или

$$\Phi^*(t) = \frac{1+\varepsilon}{2}. \tag{20.6}$$

Из этого уравнения по таблицам функции $\Phi^*(t)$ (приложение G.2) находим нужное значение t, пользуясь таблицами «наоборот»: сначала находим значение функции, а затем соответствующее значение аргумента. После этого интервал (20.4) становится полностью определенным.

Вместо таблиц функции $\Phi^*(t)$ можно пользоваться таблицами функции $\Phi(t)$ [8, 9], при этом равенство (20.5) принимает вид:

$$\Phi(t) - \Phi(-t) = \varepsilon.$$

Так как функция $\Phi(t)$ нечетна, то из него получаем, что

$$\Phi(t) = \frac{\varepsilon}{2}.$$

Остается применить таблицы функции $\Phi(t)$, чтобы найти значение t.

Отметим, что доверительный интервал (20.4) симметричен относительно выборочного среднего $\overline{x}$, так что оно остается точечной оценкой математического ожидания как середина доверительного интервала.

Пример 20.1. *Было измерено время восстановления X диодов из одной партии (в наносекундах):*

69 73 70 68 61 73 70 72 67 70

Считая X нормально распределенной с.в. с дисперсией $\mathbf{D}X = \sigma^2 = 12$, найти доверительные интервалы, соответствующие доверительным вероятностям 0,95 и 0,99 для математического ожидания с.в. X.

Решение. Выборочное среднее данной с.в. вычислим по формуле (19.1):

$$\overline{x} = \frac{1}{10}(69 + 73 + 70 + 68 + 61 + 73 + 70 + 72 + 67 + 70) = \frac{693}{10} = 69{,}3.$$

Найдем величину t из равенства (20.6) для $\varepsilon = 0{,}95$:

$$\Phi^*(t) = \frac{1 + 0{,}95}{2} = 0{,}975.$$

Обращаясь к таблице функции $\Phi^*(t)$ (приложение G.2), определяем значение $t = 1{,}96$. Далее формула (20.4) дает искомый доверительный интервал

$$(m_1, m_2) = \left(69{,}3 - 1{,}96\sqrt{\frac{12}{10}}; 69{,}3 + 1{,}96\sqrt{\frac{12}{10}}\right) \approx (67{,}15; 71{,}45).$$

Аналогично решается задача для доверительной вероятности $\varepsilon = 0{,}99$. В этом случае получается $t = 2{,}57$, а доверительный интервал становится равным

$$(m_1, m_2) = \left(69{,}3 - 2{,}57\sqrt{\frac{6}{5}}; 69{,}3 + 2{,}57\sqrt{\frac{6}{5}}\right) \approx (66{,}48; 72{,}12).$$

Видно, что большей доверительной вероятности соответствует и больший доверительный интервал. Это значит, что бóльшие гарантии наличия неизвестного параметра в некотором интервале может дать только более широкий интервал. Ну, а длины интервалов показывают, что точность определения параметра по десяти наблюдениям не очень высокая. □

20.2 Оценка $\mathbf{M}X$ при неизвестной дисперсии

Пусть теперь дисперсия $\mathbf{D}X$ с.в. X неизвестна. В этом случае рассматривают статистику

$$T = \frac{\mathbf{M}X - \overline{x}}{\widetilde{s}/\sqrt{n}}, \tag{20.7}$$

аналогичную с.в. (20.2). Просто с.к.о. σ (теперь неизвестное) заменено выборочным с.к.о. Но это кардинально меняет тип с.в. дроби. Если в формуле (20.2) случайность наличествовала только в числителе, то в выражении (20.7) она присутствует и в числителе, и в знаменателе. Поэтому с.в. Y была распределена по нормальному закону, а с.в. T имеет t-распределение Стьюдента, рассмотренное в п. 10.8, с $(n-1)$-й степенью свободы. Приложение G.3 содержит таблицу этого распределения, из которой можно, задав число степеней свободы и величину α, получить число t_α, удовлетворяющее соотношению (см. рис. 10.5):

$$P(|T| > t_\alpha) = \alpha. \tag{20.8}$$

Чтобы получить интервальную оценку математического ожидания потребуем, как и в предыдущем случае, учитывая симметрию плотности вероятностей распределения Стьюдента относительно оси ординат, выполнения условия:

$$P(|T| < t) = \varepsilon. \tag{20.9}$$

Это условие отличается от условия (20.8) и поэтому, чтобы воспользоваться таблицей t-распределения, его сначала нужно привести к такому виду:

$$P(|T| > t) = 1 - \varepsilon.$$

Теперь из таблицы t-распределения (приложение G.3) по числу степеней свободы $n-1$ и величине $\alpha = 1 - \varepsilon$ можно найти значение величины t. Зная его и раскрывая формулу (20.9), получаем выра-

жение

$$P(|T| < t) = P\left(-t < \frac{\mathbf{M}X - \overline{x}}{\widetilde{s}/\sqrt{n}} < t\right) =$$

$$= P\left(\overline{x} - t\frac{\widetilde{s}}{\sqrt{n}} < \mathbf{M}X < \overline{x} + t\frac{\widetilde{s}}{\sqrt{n}}\right) = \varepsilon,$$

из которого видно, что доверительным интервалом для математического ожидания с.в. X при неизвестной дисперсии является интервал

$$\boxed{(m_1, m_2) = \left(\overline{x} - t\frac{\widetilde{s}}{\sqrt{n}}, \overline{x} + t\frac{\widetilde{s}}{\sqrt{n}}\right).} \tag{20.10}$$

Пример 20.2. *Решить предыдущий пример в предположении, что дисперсия времени восстановления диода неизвестна, а доверительная вероятность* $\varepsilon = 0{,}95$.

Решение. Сначала найдем значение t. Число степеней свободы в данном примере равно $n - 1 = 9$, а величина $\alpha = 1 - \varepsilon = 1 - 0{,}95 = 0{,}05$. Используя эти числа как входы в таблицу t-распределения (приложение G.3), получаем что t равно 2,26.

Определим еще исправленную выборочную дисперсию по формулам (19.4) и (19.7):

$$\widetilde{s}^2 = \frac{1}{n-1}\sum_{k=1}^{n} X_i^2 - \frac{n}{n-1}\overline{x}^2 =$$

$$= \frac{1}{9}(69^2 + 73^2 + 70^2 + 68^2 + 61^2 + 73^2 + 70^2 + 72^2 + 67^2 + 70^2) -$$

$$-\frac{10}{9}69{,}3^2 = \frac{48137}{9} - \frac{10}{9}69{,}3^2 = 12{,}456.$$

Следовательно, $\widetilde{s} = \sqrt{12{,}456} = 3{,}53$ и из равенства (20.10) получаем искомый доверительный интервал:

$$(m_1, m_2) = \left(69{,}3 - 2{,}26 \cdot \frac{3{,}53}{\sqrt{10}}; 69{,}3 + 2{,}26 \cdot \frac{3{,}53}{\sqrt{10}}\right) = (66{,}78; 71{,}82).$$

Видно, что доверительный интервал для математического ожидания при неизвестной дисперсии шире, чем в случае, когда дисперсия известна (при одной и той же доверительной вероятности). □

20.3 Оценка DX

Для получения интервальной оценки дисперсии тоже используют подходящую статистику, которая, однако, не похожа на две предыдущие:

$$\chi^2 = \frac{(n-1)\widetilde{s}^2}{\sigma^2}.$$

Она имеет χ^2-распределение Пирсона, рассмотренное в п. 10.9, с $(n-1)$-й степенью свободы. В приложении G.4 приведена таблица этого распределения с двумя входами, одним из которых является число степеней свободы, а другим — величина γ. На выходе таблицы получают число u, определяемое соотношением (см. рис. 10.6):

$$P(\chi^2 > u) = \gamma. \tag{20.11}$$

Подберем два числа u_1 и u_2 таких, чтобы вероятность попадания статистики χ^2 в интервал (u_1, u_2) равнялась заданной доверительной вероятности ε:

$$P\left(u_1 < \frac{(n-1)\widetilde{s}^2}{\sigma^2} < u_2\right) = \varepsilon. \tag{20.12}$$

В отличие от предыдущих случаев плотность вероятностей с.в., имеющей χ^2-распределение, не симметрична относительно какой-либо вертикальной прямой (см. рис. 10.6) и однозначность выбора чисел u_1 и u_2 в формуле (20.12) надо обеспечивать, исходя из других соображений. Поэтому вводят «симметричность по вероятности», то есть накладывают два дополнительных условия, заключающихся в том, что вероятности попадания статистики χ^2 в интервалы $(-\infty, u_1)$ и (u_2, ∞) должны быть равны и в сумме равняться $1-\varepsilon$. То есть должны выполняться соотношения:

$$P(\chi^2 < u_1) = \boxed{P(\chi^2 > u_2) = \frac{1-\varepsilon}{2}}. \tag{20.13}$$

Из этой формулы по таблице χ^2-распределения (приложение G.4) мы сразу находим u_2, используя в качестве входов число степеней

свободы $n-1$ и $\gamma = (1-\varepsilon)/2$ (ср. с (20.11)). Чтобы найти u_1, сначала преобразуем соответствующее равенство в (20.13):

$$\boxed{P(\chi^2 > u_1) = 1 - \frac{1-\varepsilon}{2} = \frac{1+\varepsilon}{2}.}$$

Снова обращаясь к таблице χ^2-распределения с тем же числом степеней свободы и $\gamma = (1+\varepsilon)/2$, определяем значение u_1.

Теперь, уже при конкретных значениях u_1 и u_2 из (20.12) получаем

$$P\left(\frac{1}{u_2} < \frac{\sigma^2}{(n-1)\widetilde{s}^2} < \frac{1}{u_1}\right) = P\left(\frac{(n-1)\widetilde{s}^2}{u_2} < \sigma^2 < \frac{(n-1)\widetilde{s}^2}{u_1}\right) = \varepsilon.$$

Из этого равенства видно, что в качестве доверительного интервала для дисперсии с.в. X можно взять интервал

$$\boxed{(s_1, s_2) = \left(\frac{(n-1)\widetilde{s}^2}{u_2}, \frac{(n-1)\widetilde{s}^2}{u_1}\right).} \tag{20.14}$$

Зная доверительный интервал для дисперсии, совсем несложно получить доверительный интервал для с.к.о. (при той же самой доверительной вероятности): достаточно извлечь квадратные корни из концов интервала. То есть доверительный интервал для с.к.о. имеет вид:

$$\boxed{(\sqrt{s_1}, \sqrt{s_2}).}$$

Пример 20.3. *Найти доверительный интервал для дисперсии и с.к.о. времени восстановления диода из примера 20.1 с доверительной вероятностью* $\varepsilon = 0{,}95$.

Решение. Число степеней свободы равно $n-1 = 9$. По таблице χ^2-распределения (приложение G.4) при $\gamma = (1-\varepsilon)/2 = (1-0{,}95)/2 = 0{,}025$ определяем $u_2 = 19{,}2$. По той же таблице и том же числе степеней свободы при $\gamma = (1+\varepsilon)/2 = (1+0{,}95)/2 = 0{,}975$ находим $u_1 = 2{,}664$. Наконец, по формуле (20.14) получаем требуемый

доверительный интервал для дисперсии:

$$(s_1, s_2) = \left(\frac{9 \cdot 12{,}456}{19{,}2}; \frac{9 \cdot 12{,}456}{2{,}664}\right) = (5{,}84; 42{,}08).$$

Извлекая квадратные корни из концов интервала, находим доверительный интервал для с.к.о.:

$$(\sqrt{s_1}; \sqrt{s_2}) = (2{,}42; 6{,}49).$$

$\square$

Лекция 21
Статистическая проверка гипотез

21.1 Основные определения

В практических задачах часто возникают определенного рода предположения относительно с.в., причем, для успешного решения задачи эти предположения надо либо подтвердить, либо опровергнуть. К таким часто встречающимся предположениям относятся предположения о типе распределения с.в., например, пуассоновская данная с.в. или нет; равно ли ее математическое ожидание заданному числу (например, нулю) или нет; равны ли дисперсии двух рассматриваемых с.в. или нет и др. Эти предположения в математической статистике называются ***статистическими гипотезами***, а процесс выяснения их истинности — ***статистической проверкой гипотез***. Так как все выводы базируются на выборке, которая является случайным материалом, то характерно, что результатом успешной проверки является не безоговорочное утверждение вроде «гипотеза верна», а лишь вероятность, с которой можно считать, что гипотеза не противоречит наблюдениям.

Запишем несколько определений. Любую выдвинутую гипотезу можно считать ***основной (нулевой)*** и обозначать H_0. ***Конкурирующей (альтернативной)*** называется гипотеза H_1, противоположная нулевой. Таким образом, в результате проверки можно принять только одну из гипотез H_0 или H_1, отвергая в то же время другую.

Простой называется гипотеза, если она однозначно определяет

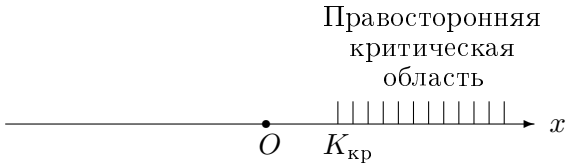
распределение с.в. В противном случае гипотеза называется ***сложной***. Например, предположение о том, что математическое ожидание показательной с.в. равно 3, является простой гипотезой, а предположение, что это же математическое ожидание больше 3, — сложной.

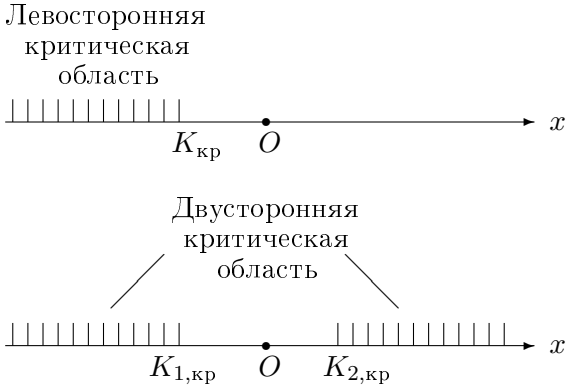
При вынесении суждения о правильности или неправильности статистической гипотезы могут быть сделаны ошибки двух типов. ***Ошибка первого рода*** заключается в том, что правильная гипотеза будет отвергнута. ***Ошибка второго рода*** состоит в том, что будет принята неправильная гипотеза. ***Уровнем значимости α*** называется вероятность совершить ошибку I рода. ***Мощностью критерия*** для проверки гипотезы называется величина $1-\beta$, где β — вероятность ошибки II рода.

Критерием проверки служит обычно некоторая с.в. $\mathbb{K}$. ***Критической областью*** называется совокупность значений критерия, при которых гипотеза H_0 отвергается. ***Областью принятия гипотезы*** называется множество значений критерия, при которых гипотеза H_0 принимается. ***Критическими точками (границами)*** $K_{\text{кр}}$ называются точки, отделяющие критическую область от области принятия гипотезы.

Правосторонняя критическая область определяется неравенством $\mathbb{K} > K_{\text{кр}} > 0$. ***Левосторонняя критическая область*** определяется неравенством $\mathbb{K} < K_{\text{кр}} < 0$. ***Двусторонняя критическая область*** определяется принадлежностью

$$\mathbb{K} \in (-\infty, K_{1,\text{кр}}) \cup (K_{2,\text{кр}}, \infty).$$





Правостороннюю критическую область находят из соотношения

$$P(\mathbb{K} > K_{\text{кр}}) = \alpha,\ \alpha = 0{,}01 \div 0{,}1,$$

где α — уровень значимости гипотезы. Если на основе выборочных данных получено значение критерия $\mathbb{K}_B$ и оказалось, что $\mathbb{K}_B > K_{\text{кр}}$, это значит, что на практике реализовалось весьма маловероятное событие, причем, его вероятность равна вероятности отвергнуть правильную гипотезу. В таком случае естественно гипотезу H_0 отвергнуть, так как считается, что в единичном опыте маловероятное событие не должно наступить.

Левостороннюю критическую область находят из соотношения

$$P(\mathbb{K} < K_{\text{кр}}) = \alpha,$$

а двустороннюю — из соотношения

$$P(\mathbb{K} < K_{1,\text{кр}}) + P(\mathbb{K} > K_{2,\text{кр}}) = \alpha.$$

Выбрав уровень значимости α, критическую область строят так, чтобы мощность критерия была наибольшей, так как при этом уменьшается до минимума (при данном α) вероятность ошибки второго рода.

При заданном объеме выборки невозможно одновременно уменьшить и α, и β. Это можно сделать, лишь увеличив объем выборки.

21.2 Критерий согласия Пирсона

Пусть имеется выборка $X_1, \ldots, X_n$ из генеральной совокупности с.в. X, имеющей неизвестную функцию распределения $F(x)$. В качестве основной гипотезы примем предположение о типе распределения с.в. X, то есть о том, будет ли данная с.в. распределенной равномерно, нормально, по биномиальному закону или какому-нибудь иному. Предполагаемую функцию распределения обозначим $F_0(x)$. Тогда основную гипотезу можно записать в виде: $H_0 : F(x) = F_0(x)$, а альтернативную — в виде $H_1 : F(x) \neq F_0(x)$. Задача проверки основной гипотезы называется ***задачей проверки согласия***, а критерий проверки гипотезы называется в данном случае ***критерием согласия***.

Идея критерия согласия Пирсона заключается в том, что основная гипотеза должна быть отвергнута, если выборочные частоты сильно отличаются от частот, которые можно получить, используя предположение, что данная с.в. имеет функцию распределения $F_0(x)$.

Сам критерий Пирсона, называемый также ***критерием*** χ^2, имеет вид

$$\boxed{\chi_B^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i},} \tag{21.1}$$

где n_i — частоты статистического ряда, np_i — так называемые ***теоретические частоты***, которые получают умножением объема выборки n на теоретические вероятности p_i. Последние получают следующим образом.

Если статистический ряд — дискретного типа, то $p_i = P(X = X^{(i)})$, $i = \overline{1, k-1}$, причем, эта вероятность определяется тем дискретным законом распределения, который соответствует функции распределения $F_0(x)$. Для первого наблюдения статистического ряда берется вероятность $p_0 = P(X \leq X^{(0)})$, если с.в. может принимать значения, меньшие, чем $X^{(0)}$, и берется вероятность $p_0 = P(X = X^{(0)})$ в противном случае. Аналогично для последнего

наблюдения берется вероятность $p_k = P(X \geq X^{(k)})$, если с.в. может принимать значения, бóльшие, чем $X^{(k)}$, и берется вероятность $p_k = P(X = X^{(k)})$ в противном случае.

Для группированного статистического ряда такие вероятности вычисляются по формуле $p_i = F_0(x_i) - F_0(x_{i-1})$, $i = \overline{1, k-1}$. Если с.в. X может принимать значения, меньшие, чем x_0, то $p_0 = F_0(x_1) - F_0(-\infty)$, в противном случае $p_0 = F_0(x_1) - F_0(x_0)$. Аналогично, если с.в. X может принимать значения, бóльшие, чем x_k, то $p_k = F_0(\infty) - F_0(x_{k-1})$, в противном случае $p_k = F_0(x_k) - F_0(x_{k-1})$.

Вид критерия Пирсона (21.1) показывает, что он действительно принимает малые значения, если выборочные частоты близки к теоретическим. (В частности, $\chi^2_B = 0$, если все $n_i = np_i$, но практически такое с.с. не наблюдается, так как слишком маловероятно. Вообще чрезмерная близость выборочных и теоретических частот может, скорее, свидетельствовать о преднамеренной подтасовке наблюдений.)

Величина χ^2_B имеет χ^2-распределение с числом степеней свободы, которое определяется как $r = k - \nu - 1$, где ν — число параметров распределения с.в., определяемых по выборке. Далее на примерах вы увидите, что величина k не является неизменной, так как, например, некоторые интервалы группировки приходится объединять и, значит, их количество может измениться. Это говорит о том, что подсчет числа степеней свободы следует производить на заключительном этапе расчетов.

Вторым параметром χ^2-распределения является уровень значимости α, равный, как вы помните, вероятности отвергнуть правильную гипотезу. Если величина χ^2_B чересчур велика, то гипотезу следует отвергнуть, поэтому естественно рассмотреть правостороннюю критическую область, определяемую соотношением

$$P(\chi^2_B > \chi^2_\alpha(r)) = \alpha, \tag{21.2}$$

где $\chi^2_\alpha(r)$ — значение статистики χ^2, найденное из таблиц (приложение G.4) по входам r и α.

Так как α мало ($\alpha = 0{,}01 \div 0{,}1$), то в случае справедливости

гипотезы H_0 вероятность (21.2) тоже мала, то есть событие $\chi^2_B > \chi^2_\alpha(r)$ маловероятно. Следовательно, если

$$\chi^2_B > \chi^2_\alpha(r),$$

то *гипотезу следует отвергнуть*, а, если

$$\boxed{\chi^2_B < \chi^2_\alpha(r),}$$

то *гипотезу можно принять как не противоречащую наблюдениям.*

Пример 21.1. *Проверить гипотезу о том, что количество звонков, поступающих утром в справочное ж.д. бюро (см. пример 18.2) распределено по закону Пуассона, взяв уровень значимости α равным* 0,2. *Для проверки использовать критерий Пирсона.*

Решение. Закон распределения Пуассона имеет вид:

$$P(X = i) = \frac{\lambda^i}{i!}e^{-\lambda};\ i = 0, 1, \ldots.$$

В этом равенстве параметр распределения λ нам неизвестен. Однако известно, что $\lambda = \mathbf{M}X$, а $\mathbf{M}X \approx \overline{x}$. Поэтому приближенно можно считать, что этот параметр равен выборочному среднему: $\lambda \approx \overline{x}$. Выборочное среднее для данной выборки было найдено в примере 19.2, откуда и получаем, что $\lambda \approx 4{,}95$. Теперь проверяемый закон Пуассона приобретает вид

$$P(X = i) = \frac{4{,}95^i}{i!}e^{-4{,}95};\ i = 0, 1, \ldots.$$

Используя эту формулу, находим теоретические частоты, соответствующие наблюдениям $X^{(i)}$, представленным в статистическом

ряде примера 18.2:

$$P(X=0)=\frac{4{,}95^0}{0!}e^{-4{,}95}=e^{-4{,}95}=0{,}0071,$$
$$P(X=1)=\frac{4{,}95^1}{1!}e^{-4{,}95}=4{,}95e^{-4{,}95}=0{,}0351,$$
$$P(X=2)=\frac{4{,}95^2}{2!}e^{-4{,}95}=0{,}0868,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$P(X=11)=\frac{4{,}95^{11}}{11!}e^{-4{,}95}=0{,}0078,$$
$$P(X\geq 12)=1-P(X=0)-\ldots-P(x=11)=0{,}0051.$$

Эти вычисления и вычисление теоретических частот представим в виде таблицы 21.1.

Из нее видно, что некоторые теоретические частоты получились довольно маленькими (см. первые две и последние четыре строки таблицы). Это значит, что соответствующие наблюдения маловероятны, и, если именно они дадут большое несовпадение выборочных и теоретических частот, то из-за маловероятных событий придется отклонить гипотезу. Поэтому на практике принято объединять строки таблицы, для которых величина np_i меньше, например, 7. Мы с вами тоже объединим строки таблицы: первые три и последние четыре. В результате получится новая таблица 21.2, из которой мы и найдем значение критерия Пирсона. Теперь определим число степеней свободы. В таблице осталось 8 строк, значит, будем считать, что теперь $k=8$. Величина $\nu=1$, так как один параметр распределения, λ, был найден по выборке. В итоге $r=8-1-1=6$. Из таблицы χ^2-распределения (приложение G.4) при $\alpha=0{,}2$ и числе степеней свободы $r=6$ находим $\chi^2_\alpha(r)=8{,}558$. Так как

$$\chi^2_B=4{,}68<8{,}558=\chi^2_\alpha(r),$$

то гипотеза о пуассоновском распределении данной с.в. не противоречит наблюдениям и ее можно принять. □

Таблица 21.1: Критерий χ^2 для примера 18.2 (1 часть)

$X^{(i)}$	n_i	p_i	np_i
0	2	0,0071	0,852
1	3	0,0351	4,212
2	10	0,0868	10,416
3	17	0,1432	17,184
4	22	0,1772	21,264
5	26	0,1754	21,048
6	13	0,1447	17,364
7	13	0,1023	12,276
8	4	0,0633	7,596
9	4	0,0348	4,176
10	2	0,0172	2,064
11	3	0,0078	0,936
12	1	0,0051	0,612

Пример 21.2. *Проверить гипотезу о том, что отклонение напряжения от номинала (см. пример 18.3) имеет нормальный закон распределения, взяв уровень значимости α равным* 0,1. *Для проверки использовать критерий Пирсона.*

Решение. В данном случае придется подсчитывать вероятности попадания в интервалы группировки (x_{i-1}, x_i) нормальной с.в. по формуле:

$$P(x_{i-1} < X < x_i) = \Phi^*\left(\frac{x_i - m}{\sigma}\right) - \Phi^*\left(\frac{x_{i-1} - m}{\sigma}\right),$$

в которой параметры m и σ неизвестны. Но $m = \mathbf{M}X \approx \overline{x}$, $\sigma =$

Таблица 21.2: Критерий χ^2 для примера 18.2 (2 часть)

$X^{(i)}$	n_i	np_i	$n_i - np_i$	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
$0 \div 2$	15	15,480	−0,480	0,2304	0,0149
3	17	17,184	−0,184	0,0339	0,0020
4	22	21,264	0,736	0,5417	0,0255
5	26	21,048	4,952	24,5220	1,1651
6	13	17,364	−4,364	19,0445	1,0968
7	13	12,276	0,724	0,5242	0,0427
8	4	7,596	−3,596	12,9312	1,7024
≥ 9	10	7,788	2,212	4,8929	0,6283
Σ	120	1			$\chi^2_B = 4{,}68$

$= \sqrt{\mathbf{D}X} \approx s$. Поэтому мы вместо них возьмем выборочное среднее и выборочное с.к.о., вычисленные в примерах 19.3 и 19.6, соответственно.

В результате предыдущая формула примет вид:

$$P(x_{i-1} < X < x_i) = \Phi^*\left(\frac{x_i - 24{,}667}{5{,}764}\right) - \Phi^*\left(\frac{x_{i-1} - 24{,}667}{5{,}764}\right).$$

По ней и с помощью таблицы функции $\Phi^*(x)$ (приложение G.2) мы

рассчитаем теоретические частоты:

$$P(-\infty < X < 11) = \Phi^*\left(\frac{11-24{,}667}{5{,}764}\right) - \Phi^*(-\infty) =$$
$$= \Phi^*(-2{,}37) = 0{,}089;$$
$$P(11 < X < 15) = \Phi^*\left(\frac{15-24{,}667}{5{,}764}\right) - \Phi^*\left(\frac{11-24{,}667}{5{,}764}\right) =$$
$$= \Phi^*(-1{,}68) - \Phi^*(-2{,}37) = 0{,}0376;$$

. . . . . .

$$P(31 < X < 35) = \Phi^*\left(\frac{35-24{,}667}{5{,}764}\right) - \Phi^*\left(\frac{31-24{,}667}{5{,}764}\right) =$$
$$= \Phi^*(1{,}79) - \Phi^*(1{,}10) = 0{,}0994;$$
$$P(35 < X < \infty) = \Phi^*(\infty) - \Phi^*\left(\frac{35-24{,}667}{5{,}764}\right) =$$
$$= 1 - \Phi^*(1{,}79) = 0{,}0367.$$

Дальнейшие вычисления поместим в таблицу 21.3.

Мы снова наблюдаем для некоторых строк малые значения np_i. Но если в последней строке эта величина равна 6,606, то есть близка к 7, то для первой строки она все же слишком мала, и, значит, первую строку нужно объединить со второй. После такого объединения результаты расчетов можно представить таблицей 21.4. Чтобы найти число степеней свободы, вычтем из количества строк $k = 7$, оставшихся в таблице, величину $\nu = 2$, так как по выборке было определено два параметра распределения: m и σ. В результате $r = 7-2-1 = 4$. По таблице χ^2-распределения (приложение G.4) при $\alpha = 0{,}1$ и числе степеней свободы $r = 4$ находим $\chi^2_\alpha(r) = 7{,}779$. Так как

$$\chi^2_B = 1{,}91 < 7{,}779 = \chi^2_\alpha(r),$$

то гипотеза о нормальном распределении данной с.в. не противоречит наблюдениям и ее можно принять. □

Таблица 21.3: Критерий χ^2 для примера 18.3 (1 часть)

(x_{i-1}, x_i)	n_i	p_i	np_i
$(-\infty; 11)$	2	0,0089	1,602
$(11; 15)$	6	0,0376	6,768
$(15; 19)$	20	0,1171	21,078
$(19; 23)$	44	0,2223	40,014
$(23; 27)$	43	0,2695	48,510
$(27; 31)$	40	0,2089	37,602
$(31; 35)$	20	0,0989	17,802
$(35; \infty)$	5	0,0367	6,606

21.3 Критерий согласия Романовского

В соответствии с этим критерием расхождение между выборочным и теоретическим распределениями считается несущественным (то есть гипотезу можно принять), если

$$\boxed{\left|\frac{\chi_B^2 - r}{\sqrt{2r}}\right| < 3.}$$

Здесь χ_B^2 — выборочное значение критерия χ^2, подсчитываемое так же, как для критерия Пирсона; r — число степеней свободы.

Данный критерий менее строг, чем критерий Пирсона, то есть некоторые гипотезы, которые отвергаются критерием Пирсона, можно принять, если следовать критерию Романовского. Вообще последний применяется, когда нет необходимости исследовать *степень близости* выборочного и теоретического распределений.

Таблица 21.4: Критерий χ^2 для примера 18.3 (2 часть)

(x_{i-1}, x_i)	n_i	np_i	$n_i - np_i$	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
$(-\infty; 15)$	8	8,370	−0,370	0,1369	0,0164
$(15; 19)$	20	21,078	−1,078	1,1621	0,0551
$(19; 23)$	44	40,014	3,986	15,8882	0,3971
$(23; 27)$	43	48,510	−5,510	30,3601	0,6259
$(27; 31)$	40	37,602	2,398	5,7504	0,1529
$(31; 35)$	20	17,802	2,198	4,8312	0,2714
$(35; \infty)$	5	6,606	−1,606	2,5792	0,3904
Σ	180	0,9999			$\chi^2_B = 1{,}91$

Для примера 21.1 критерий Романовского выполняется:

$$\left|\frac{4{,}68 - 8}{\sqrt{8}}\right| = 1{,}17 < 3.$$

То же самое можно сказать о примере 21.2:

$$\left|\frac{1{,}91 - 4}{\sqrt{4}}\right| = 1{,}05 < 3.$$

Следовательно, обе гипотезы из этих примеров, уже принятые по критерию Пирсона, можно принять и по критерию Романовского.

21.4 Критерий согласия Колмогорова

Данный критерий представляет собой с.в. $\mathbb{K} = D_n\sqrt{n}$, где

$$D_n = \max_x |\overline{F}_n(x) - F_0(x)|,$$

причем, $\overline{F}_n(x)$ — полигон относительных накопленных частот, $F_0(x)$ — предполагаемая теоретическая функция распределения.

Как показал А.Н. Колмогоров, выполняется соотношение:

$$P(D_n\sqrt{n} > \lambda) \approx P(\lambda) = 1 - \sum_{i=-\infty}^{\infty} (-1)^i e^{-2i^2\lambda^2}.$$

Выберем $\lambda = \lambda_\alpha$ так, чтобы

$$P(D_n\sqrt{n} > \lambda_\alpha) \approx P(\lambda_\alpha) = \alpha,$$

где α — уровень значимости гипотезы H_0. Следовательно, событие $D_n\sqrt{n} > \lambda_\alpha$ маловероятно. Значит, если выполнится неравенство

$$D_n\sqrt{n} > \lambda_\alpha,$$

то гипотезу следует отвергнуть, а, если выполнится противоположное неравенство:

$$\boxed{D_n\sqrt{n} < \lambda_\alpha,}$$

то гипотезу можно принять.

На применение данного критерия накладываются два существенных ограничения. Во-первых, его можно применять лишь к непрерывным с.в. Во-вторых, должны быть известны все параметры теоретического распределения (например, m и σ для нормального распределения). Если же такие параметры определяются приближенно по выборке, то критерий Колмогорова дает заведомо завышенное значение вероятности $P(\lambda)$. Поэтому в ряде случаев существует риск принять как правдоподобную гипотезу, в действительности плохо согласующуюся с наблюдениями.

Пример 21.3. *Применим критерий Колмогорова к гипотезе, сформулированной в примере 21.2, при том же уровне значимости* $\alpha = 0{,}1$.

Решение. Из таблицы значений функции $P(\lambda)$ (приложение G.5) находим, что для уровня значимости $\alpha = 0{,}1$ величина $\lambda_\alpha \approx 1{,}23$, а из рис. 18.4 видно, что $D_n = 0{,}025$. Следовательно,

$$D_n\sqrt{n} = 0{,}025\sqrt{180} = 0{,}335 < 1{,}23 = \lambda_\alpha$$

и гипотезу можно принять. □

Лекция 22
Элементы корреляционного анализа

22.1 Выборочные уравнения регрессии

При построении корреляционной связи по статистическим данным выборка является набором пар (X_i, Y_i), $i = \overline{1, n}$, причем, каждая пара представляет собой значения с.в. X и Y, соответственно, которые наблюдались в одном и том же (i-м) опыте.

Будем, как и в п. 8.4 снова искать уравнение прямой вида

$$\boxed{Y \simeq \alpha X + \beta}, \tag{22.1}$$

но в качестве критерия близости опытных данных к линейной зависимости выберем минимум суммы квадратов отклонений экспериментальных значений с.в. Y от ее значений, получаемых на основании зависимости (22.1):

$$S(a, b) = \sum_{i=1}^{n} (Y_i - \alpha X_i - \beta)^2.$$

Применяя необходимые условия экстремума, находим частные производные этой функции и приравниваем их нулю

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^{n} X_i (Y_i - \alpha X_i - \beta) = 0,$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^{n} (Y_i - \alpha X_i - \beta) = 0.$$

В результате получаем систему линейных алгебраических уравнений для определения параметров α и β уравнения (22.1):

$$\begin{cases} \alpha \sum\limits_{i=1}^{n} X_i^2 + \beta \sum\limits_{i=1}^{n} X_i = \sum\limits_{i=1}^{n} X_i Y_i, \\ \alpha \sum\limits_{i=1}^{n} X_i + \beta n = \sum\limits_{i=1}^{n} Y_i. \end{cases}$$

Разделим каждое уравнение на n и введем обозначения:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ \overline{y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,\ \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2,\ K = \frac{1}{n}\sum_{i=1}^{n} X_i Y_i.$$

В этих обозначениях полученная система уравнений примет вид:

$$\begin{cases} \alpha\overline{x^2} + \beta\overline{x} = K, \\ \alpha\overline{x} + \beta = \overline{y}. \end{cases}$$

Из нее находим, что

$$\alpha = \frac{K - \overline{x}\cdot\overline{y}}{s_X^2}, \quad \beta = \overline{y} - \alpha\overline{x}.$$

Мы не станем доказывать, что полученные значения α и β действительно обеспечивают минимум функции $S(a, b)$, но это в самом деле так.

Параметр α называется еще ***выборочным коэффициентом регрессии*** Y на X и обозначается ρ_{YX}. Он является оценкой коэффициента регрессии $\dfrac{\sigma_Y}{\sigma_X} r$. Используя эти величины нетрудно получить оценку коэффициента корреляции с.в. X и Y, которая называется ***выборочным коэффициентом корреляции***:

$$\boxed{r_B = \rho_{YX}\frac{s_X}{s_Y} = \frac{K - \overline{x}\cdot\overline{y}}{s_X s_Y}.} \tag{22.2}$$

Выборочное уравнение регрессии Y на X теперь можно записать в виде

$$\boxed{Y - \overline{y} \simeq \frac{s_Y}{s_X} r_B (X - \overline{x}).} \tag{22.3}$$

Совершенно аналогично можно получить уравнение регрессии X на Y:

$$\boxed{X - \overline{x} \simeq \frac{s_X}{s_Y} r_B (Y - \overline{y}).} \tag{22.4}$$

Пример 22.1. *Производились наблюдения за парой с.в. X — влажностью воздуха (в %) в помещении и Y — сроком службы (в месяцах) аппаратуры одного типа до первого ремонта. Наблюдения представлены таблицей 22.1. Определить выборочный коэффициент корреляции между этими с.в. и выборочное уравнение линейной регрессии Y на X.*

Таблица 22.1: Корреляционная таблица

$X \backslash Y$	3	6	9	12	15	18	21	24	n_i	$\overline{y}_j$
10		1				4	1	3	9	19,00
20			2	1	2	14	6	8	33	19,09
30	1			9	5	20	23	5	63	18,24
40		1		15	24	12	15	10	77	17,10
50		6	3	10	18	9	2	4	52	14,48
60	8	12	7	13	13	3			56	10,07
70	7	10	15	4	1		1		38	7,89
80	10	8	6	10		5			39	8,77
90	15	2		2			2		21	5,86
m_j	41	40	33	64	63	67	50	30	388	

Решение. Данные корреляционной таблицы представляют собой количества экземпляров аппаратуры, прослуживших до первого ремонта число месяцев, указанное в верхней строке таблицы, при влажности воздуха, указанной в ее первом столбце. Так, например, 15 месяцев при влажности 40% прослужило 24 экземпляра аппаратуры. Складывая числа в каждой строке таблицы, мы получаем частоты

n_i для с.в. X (они показаны в предпоследнем столбце), а складывая числа в каждом столбце, находим частоты m_j для с.в. Y (они показаны в последней строке). Складывая частоты как для с.в. X, так и для с.в. Y, мы получаем общее число наблюдений $n = 388$. О числах, стоящих в последнем столбце таблицы, поговорим позже.

Используя известные уже вам формулы для подсчета выборочных числовых характеристик с.в., последовательно находим:

$$\overline{x} = \frac{1}{n}\sum n_i X_i = 49{,}87;\ \ s_X^2 = 429{,}11;\ \ s_X = 20{,}71;$$
$$\overline{y} = \frac{1}{n}\sum m_j Y_j = 13,79;\ \ s_Y^2 = 39{,}10;\ \ s_Y = 6{,}25.$$

Величину K найдем по формуле, очевидно, равносильной той, которая была введена раньше:

$$K = \frac{1}{n}\sum n_{ij} X_i Y_j = 598{,}9175,$$

где n_{ij} — совместная частота появления наблюдений X_i и Y_j (эти числа стоят во внутренних клетках корреляционной таблицы). Далее по формуле (22.2) определяем выборочный коэффициент корреляции:

$$r_B = \frac{598{,}9175 - 49{,}87 \cdot 13{,}79}{20{,}71 \cdot 6{,}25} = -0{,}686.$$

Теперь, используя равенство (22.3), можно записать уравнение выборочной регрессии Y на X, выразив Y через X:

$$Y = 13{,}79 + \frac{6{,}25}{20{,}71}(-0{,}686)(X - 49{,}87),$$

или

$$Y = -0{,}207X + 24{,}113.$$

Применение формулы (22.4) дает уравнение выборочной регрессии X на Y:

$$X = -2{,}273Y + 81{,}215.$$

□

22.2 Корреляционное отношение

Как было отмечено на предыдущих лекциях, коэффициент корреляции (и его оценка — выборочный коэффициент корреляции) измеряет силу линейной связи между двумя с.в.: чем больше модуль этого коэффициента, тем больше связь между X и Y похожа на линейную. Если же зависимость между X и Y не является линейной, то для оценки силы связи между ними коэффициент корреляции уже не годится, нужна новая числовая характеристика.

Чтобы ее ввести, разобьем все наблюдения за парами с.в. X и Y на группы. Собственно говоря, группой будем считать наблюдения, соответствующие одному и тому же значению с.в. X (каждой группе отвечает одна строка корреляционной таблицы). Для каждой такой i-й группы можно найти выборочное среднее $\overline{y}_i$, которое называется ***групповым средним***, и выборочную дисперсию s_i^2, называемую ***групповой дисперсией***.

Внутригрупповой дисперсией называется взвешенное по объемам групп среднее арифметическое групповых дисперсий:

$$\boxed{s^2_{\text{внгр}} = \frac{1}{n}\sum N_i s_i^2,} \tag{22.5}$$

где $N_i = n_i$ — объем i-й группы.

Межгрупповой дисперсией называется дисперсия групповых средних относительно (общего) выборочного среднего с.в. Y:

$$\boxed{s^2_{\text{мжгр}} = \frac{1}{n}\sum N_i(\overline{y}_i - \overline{y})^2.} \tag{22.6}$$

Справедлива теорема, которую мы примем без доказательства.

Теорема 22.1. *(Общая) выборочная дисперсия с.в. Y равна сумме внутригрупповой и межгрупповой дисперсий:*

$$\boxed{s_Y = s^2_{внгр} + s^2_{мжгр}.} \tag{22.7}$$

Наконец, мы можем ввести числовую характеристику силы нелинейной корреляционной связи между с.в. X и Y. Ею называется ***(выборочное) корреляционное отношение*** η_{YX}, равное отношению межгруппового с.к.о. к (общему) с.к.о. s_Y с.в. Y:

$$\boxed{\eta_{YX} = \frac{s_{\text{мжгр}}}{s_Y}.} \tag{22.8}$$

Почему эта величина отражает силу нелинейной корреляционной связи, объясняют

СВОЙСТВА КОРРЕЛЯЦИОННОГО ОТНОШЕНИЯ

1°. *Корреляционное отношение заключено между* 0 *и* 1*:*

$$0 \le \eta_{YX} \le 1.$$

Неотрицательность корреляционного отношения следует из неотрицательности определяющих его величин. Из формулы (22.7), разделив на s_Y^2, получаем

$$1 = \frac{s^2_{\text{внгр}}}{s_Y^2} + \eta^2_{YX} \Longrightarrow \eta_{YX} \le 1. \tag{22.9}$$

2°. *Если корреляционное отношение равно 0, то с.в. Y не связана со с.в. X корреляционной зависимостью.*

Действительно, если $\eta_{YX} = 0$, то из определения (22.8) следует, что $s^2_{\text{мжгр}} = 0$, а тогда из формулы (22.6) вытекает равенство $\sum N_i(\overline{y}_i - \overline{y})^2 = 0$, которое означает, что все групповые средние равны (общей) выборочной средней. Таким образом, изменение с.в. X никак не влияет на изменение групповых средних с.в. Y.

3°. *Корреляционное отношение равно 1 тогда и только тогда, когда с.в. X и Y связаны строгой функциональной зависимостью.*

Если $\eta_{YX} = 1$, то $s_Y^2 = s_{\text{межгр}}^2$, и, так как справедливо равенство (22.7), то $s_{\text{внгр}}^2 = 0$. Из формулы (22.5) тогда заключаем, что все групповые дисперсии равны нулю. Это возможно лишь в том случае, когда внутри группы данному значению X соответствует одно и то же значение Y, то есть, вообще говоря, каждому значению X отвечает единственное значение Y.

Доказательство обратного утверждения опустим.

4°. *Корреляционное отношение не меньше модуля выборочного коэффициента корреляции:*

$$\eta_{YX} \geq |r_B|.$$

5°. *Если $\eta_{YX} = |r_B|$, то с.в. Y связана со с.в. X точной корреляционной зависимостью вида*

$$Y - \overline{y} = \frac{s_Y}{s_X} r_B (X - \overline{x}).$$

6°. *С возрастанием η_{YX} корреляционная связь становится все более сильной.*

В самом деле, из формулы (22.9) получаем, что

$$s_{\text{внгр}}^2 = s_Y^2 (1 - \eta_{YX}^2).$$

Ясно, что при возрастании корреляционного отношения внутригрупповая дисперсия уменьшается, а с ней уменьшаются и все групповые дисперсии. Это приводит к тому, что значения с.в. Y, соответствующие одному и тому же значению с.в. X, все меньше различаются между собой, а корреляционная связь стремится стать функциональной.

Пример 22.2. *Найти корреляционное отношение для данных предыдущего примера.*

Решение. Вот сейчас и заполним последний столбец корреляционной таблицы. В него помещаются групповые средние $\overline{y}_i$, которые вычисляются так:

$$\overline{y}_1 = \frac{1}{9}(1 \cdot 6 + 4 \cdot 18 + 1 \cdot 21 + 3 \cdot 24) = 19{,}00$$

и т.д. Затем по формуле (22.6) находим межгрупповую дисперсию и соответствующее ей с.к.о.:

$$s^2_{\text{мжгр}} = 19{,}815; \quad s_{\text{мжгр}} = 4{,}45.$$

Остается получить выборочное корреляционное отношение, использовав равенство (22.8), которое его определяет:

$$\eta_{YX} = \frac{4{,}45}{6{,}25} = 0{,}712.$$

Это число и показывает силу корреляционной связи между влажностью воздуха и продолжительностью работы аппаратуры до первого отказа. □

22.3 Понятие о криволинейной регрессии

Не обязательно искать линии регрессии в виде прямых. Очень часто наблюдения, изображенные в виде точек на плоскости, могут складываться в фигуры, близкие по очертаниям к параболе, гиперболе, экспоненте и т.д. Попытки представить регрессию в виде названных кривых в этих случаях дают меньшую погрешность приближения, чем представление с помощью прямой. Регрессию, если она задается кривой, так и называют ***криволинейной***. Наиболее часто рассматриваются квадратическая регрессия:

$$Y \simeq aX^2 + bX + c,$$

кубическая регрессия:

$$Y \simeq aX^3 + bX^2 + cX + d$$

и др.

Коэффициенты в этих формулах определяют методом наименьших квадратов, аналогично тому, как это делалось при определении коэффициентов линейной регрессии.

К сожалению, например, корреляционное отношение, способное определить силу корреляционной связи между с.в., ничем не может помочь в определении *типа* кривой, которая наилучшим образом описывала бы наблюдаемые пары с.в. Поэтому здесь большую роль начинает играть опыт исследователя и его интуиция, которые подсказывают, какая форма кривой даст наименьшую погрешность приближения. Часто в решении этого вопроса помогает сравнение остаточных дисперсий, соответствующих различным кривым, и другие методы математической статистики, которые мы, правда, изучать не будем.

Часть V

Приложения

Приложение A
Некоторые комбинаторные тождества

Утверждение A.1. *Для биномиальных коэффициентов справедлива формула*

$$C_{m+n}^{r} = \sum_{s=0}^{m} C_m^s C_n^{r-s}, \tag{A.1}$$

$0 \le r - s \le n$.

Доказательство. Возьмем тождество

$$(a+b)^{m+n} = (a+b)^m (a+b)^n$$

и применим ко всем степеням в нем формулу бинома Ньютона [7]:

$$\sum_{r=0}^{m+n} C_{m+n}^{r} a^r b^{m+n-r} = \left(\sum_{s=0}^{m} C_m^s a^s b^{m-s}\right)\left(\sum_{i=0}^{n} C_n^i a^i b^{n-i}\right).$$

Приравняем друг другу коэффициенты при одинаковых одночленах вида $a^j b^k$. Очевидно, одинаковые одночлены получаются, когда $r = s+i$, $m+n-r = (m-s)+(n-i)$, т.е. в обоих случаях при $i = r-s$. Из этого и следует доказываемая формула. □

Следствие A.1.

$$\sum_{k=0}^{K} C_K^k C_{L-K}^{l-k} = C_L^l,$$

$0 \le l - k \le L - K$.

Доказательство. Достаточно в равенстве (A.1) положить $m = K$, $n = L - K$, $s = k$, $r = l$. □

Замечание А.1. *Двойное неравенство* $0 \le l - k \le L - K$ *можно опустить, если считать, что* $C_n^m = 0$, *если* $m > n$.

Утверждение А.2. *Справедливо равенство*

$$n(n-1)\cdot\ldots\cdot(n-r+1)C_{n-r}^{k-r} = k(k-1)\cdot\ldots\cdot(k-r+1)C_n^k. \quad \text{(A.2)}$$

Доказательство. Действительно,

$$n(n-1)\cdot\ldots\cdot(n-r+1)C_{n-r}^{k-r} = \frac{n(n-1)\cdot\ldots\cdot(n-r+1)(n-r)!}{(k-r)!(n-k)!} = \\ = \frac{n!}{(k-r)!(n-k)!} = \frac{k(k-1)\cdot\ldots\cdot(k-r+1)n!}{k!(n-k)!} = \\ = k(k-1)\cdot\ldots\cdot(k-r+1)C_n^k.$$

□

Утверждение А.3. *Справедливы формулы:*

$$\sum_{k=1}^{K} kC_K^k C_{L-K}^{l-k} = KC_{L-1}^{l-1}, \quad \text{(A.3)}$$

$$\sum_{k=2}^{K} k(k-1)C_K^k C_{L-K}^{l-k} = K(K-1)C_{L-2}^{l-2}, \quad \text{(A.4)}$$

$$\sum_{k=1}^{K} k^2 C_K^k C_{L-K}^{l-k} = K\left[(K-1)C_{L-2}^{l-2} + C_{L-1}^{l-1}\right]. \quad \text{(A.5)}$$

Доказательство. Применим к левой части равенства (A.3) формулу (A.2) при $r = 1$:

$$\sum_{k=1}^{K} kC_K^k C_{L-K}^{l-k} = K\sum_{k=1}^{K} C_{K-1}^{k-1} C_{L-K}^{l-k} = \langle i = k-1\rangle = K\sum_{i=0}^{K-1} C_{K-1}^{i} C_{L-K}^{l-i-1}.$$

Далее применим тождество (A.1) при $m = K - 1$, $n = L - K$, $s = i$, $r = l - 1$ и получим правую часть (A.3).

Теперь возьмем равенство (A.4) и применим к его левой части формулу (A.2) при $r = 2$:

$$\sum_{k=2}^{K} k(k-1)C_K^k C_{L-K}^{l-k} = K(K-1)\sum_{k=2}^{K} C_{K-2}^{k-2} C_{L-K}^{l-k} = \langle i = k-2 \rangle =$$
$$= K(K-1)\sum_{i=0}^{K-2} C_{K-2}^{i} C_{L-K}^{l-i-2}.$$

Применим тождество (A.1) при $m = K-2$, $n = L-K$, $s = i$, $r = l-2$ и получим правую часть (A.4).

Остается убедится в справедливости формулы (A.5), использовав уже доказанные формулы:

$$\sum_{k=1}^{K} k^2 C_K^k C_{L-K}^{l-k} = \sum_{k=1}^{K} [k(k-1)+k] C_K^k C_{L-K}^{l-k} =$$
$$= \sum_{k=1}^{K} k(k-1) C_K^k C_{L-K}^{l-k} + \sum_{k=1}^{K} k C_K^k C_{L-K}^{l-k} =$$
$$= K(K-1)C_{L-2}^{l-2} + KC_{L-1}^{l-1} = K\left[(K-1)C_{L-2}^{l-2} + C_{L-1}^{l-1}\right].$$

□

Приложение B
Полиномиальные коэффициенты

Докажем формулу, обобщающую бином Ньютона на сумму произвольного числа слагаемых:

$$(p_1 + \cdots + p_m)^n = \sum_{k_1+\cdots+k_m=n} \frac{n!}{k_1!\ldots k_m!}\, p_1^{k_1} \ldots p_m^{k_m}. \tag{B.1}$$

Числа

$$\frac{n!}{k_1!\ldots k_m!}\, p_1^{k_1} \ldots p_m^{k_m}$$

называются ***полиномиальными коэффициентами***.

Доказательство проведем индукцией по m. При $m = 2$ справедливость формулы устанавливается тем, что она превращается в бином Ньютона:

$$(p_1 + p_2)^n = \sum_{k=0}^{n} C_n^k\, p_1^k p_2^{n-k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p_1^k p_2^{n-k} =$$
$$= \sum_{k_1+k_2=n} \frac{n!}{k_1!k_2!} p_1^{k_1} p_2^{k_2}.$$

Пусть формула (B.1) верна при $m = \nu \geq 2$. Докажем, что она верна и при $m = \nu + 1$. Используя предположение индукции и бином

Ньютона, получим

$$
\begin{aligned}
&(p_1 + \dots + p_{\nu+1})^n = [(p_1 + \dots + p_\nu) + p_{\nu+1}]^n = \\
&= \sum_{k=0}^{n} C_n^k \, (p_1 + \dots p_\nu)^k \, p_{\nu+1}^{n-k} = \\
&= \sum_{k=0}^{n} C_n^k \sum_{k_1+\dots+k_\nu=k} \frac{k!}{k_1! \dots k_\nu!} \, p_1^k \dots p_\nu^{k_\nu} p_{\nu+1}^{n-k} = \\
&= \sum_{k=0}^{n} \sum_{k_1+\dots+k_\nu=k} \frac{n!}{k!\,(n-k)!} \cdot \frac{k!}{k_1! \dots k_\nu!} \, p_1^k \dots p_\nu^{k_\nu} p_{\nu+1}^{n-k} = \\
&= \langle k_{\nu+1} = n-k, k = n - k_{\nu+1} \rangle = \\
&= \sum_{k_{\nu+1}=0}^{n} \sum_{k_1+\dots+k_\nu=n-k_{\nu+1}} \frac{n!}{k_1! \dots k_{\nu+1}!} \, p_1^{k_1} \dots p_{\nu+1}^{k_{\nu+1}} = \\
&= \sum_{k_1+\dots+k_{\nu+1}=n} \frac{n!}{k_1! \dots k_{\nu+1}!} \, p_1^{k_1} \dots p_{\nu+1}^{k_{\nu+1}},
\end{aligned}
$$

что и требовалось доказать.

Приложение C
Интеграл Пуассона

Докажем формулу (***интеграл Пуассона***)

$$\int\limits_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}.$$

Обозначим искомый интеграл через I и сделаем в нем замену $x = ut$, $u > 0$; $dx = u\,dt$. Тогда

$$I = I(u) = u\int\limits_0^\infty e^{-u^2t^2}\,dt.$$

Умножим обе части этого равенства на $e^{-u^2}\,du$ и проинтегрируем его по u в пределах от 0 до ∞:

$$I\int\limits_0^\infty e^{-u^2}\,du = I^2 = \int\limits_0^\infty ue^{-u^2}\,du\int\limits_0^\infty e^{-u^2t^2}\,dt = \int\limits_0^\infty du\int\limits_0^\infty ue^{-u^2(1+t^2)}\,dt.$$

Изменим в последнем выражении порядок интегрирования, опуская доказательство возможности такого действия:

$$I^2 = \int\limits_0^\infty dt\int\limits_0^\infty ue^{-u^2(1+t^2)}\,du = \left\langle z = u^2\left(1+t^2\right),\ dz = 2u\left(1+t^2\right)du\right\rangle =$$

$$= \frac{1}{2}\int\limits_0^\infty \frac{dt}{1+t^2}\int\limits_0^\infty e^{-z}\,dz = \left.\frac{1}{2}\operatorname{arctg} t\right|_0^\infty \left.\left(-e^{-z}\right)\right|_0^\infty = \frac{\pi}{4}.$$

Так как $I > 0$, то $I = \frac{\sqrt{\pi}}{2}$.

Приложение D
Положительно определенные матрицы

Квадратная матрица $\mathbf{A}$ порядка n называется ***положительно определенной***, если для любого ненулевого вектора $\mathbf{x}^T = (x_1, \ldots, x_n)$ выполняется неравенство

$$\mathbf{x}^T\mathbf{A}\mathbf{x} > 0. \tag{D.1}$$

При этом квадратичная форма $\mathbf{x}^T\mathbf{A}\mathbf{x}$ называется ***положительно определенной***.

Теорема D.1. *Если $\mathbf{A}$ — положительно определенная симметричная матрица, то все ее собственные значения $p_1, \ldots, p_n$ и определитель $|\mathbf{A}|$ положительны, причем, $|\mathbf{A}| = p_1 \cdots p_n$.*

Доказательство. Для симметричной матрицы существует ортогональный оператор с матрицей $\mathbf{L}$ такой, что преобразование переменных

$$\mathbf{x} = \mathbf{L}\mathbf{y},$$

где $\mathbf{y}^T = (y_1, \ldots, y_n)$ — вектор новых переменных, превращает квадратичную форму $\mathbf{x}^T\mathbf{A}\mathbf{x}$ в так называемую сумму квадратов

$$p_1 y_1^2 + \cdots + p_n y_n^2,$$

а матрицу $\mathbf{A}$ — в диагональную матрицу

$$\mathbf{\Lambda} = \begin{pmatrix} p_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & p_n \end{pmatrix} = \{\lambda_{ij}\}. \tag{D.2}$$

Следовательно, можно записать

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = (\mathbf{L}\mathbf{y})^T\,\mathbf{A}\mathbf{L}\mathbf{y} = \mathbf{y}^T\mathbf{L}^T\mathbf{A}\mathbf{L}\mathbf{y} = p_1y_1^2 + \cdots + p_ny_n^2.$$

В силу условия положительной определенности (D.1) левая часть полученного равенства положительна, значит, и его правая часть тоже положительна. Так как вектор $\mathbf{y}$ может быть любым, возьмем $y_1 = 1$, $y_2 = \cdots = y_n = 0$. Тогда условие $p_1y_1^2 + \cdots + p_ny_n^2 > 0$ станет таким: $p_1 > 0$. Аналогично доказывается, что $p_2 > 0, \ldots, p_n > 0$. Так как определитель матриц $\mathbf{L}$ и $\mathbf{L}^T$ равен ± 1, то, домножая определитель матрицы $\mathbf{A}$ слева и справа на ± 1 и учитывая диагональность матрицы $\mathbf{\Lambda}$, получаем

$$|\mathbf{A}| = \left|\mathbf{L}^T\right| |\mathbf{A}|\,|\mathbf{L}| = \left|\mathbf{L}^T\mathbf{A}\mathbf{L}\right| = |\mathbf{\Lambda}| = p_1 \cdots p_n > 0.$$

□

Теорема D.2. *Для симметричной положительно определенной матрицы* $\mathbf{A}$ *существует невырожденная матрица* $\mathbf{K}$*, такая, что* $\mathbf{K}^T\mathbf{A}\mathbf{K} = \mathbf{E}$*, где* $\mathbf{E}$ *— единичная матрица, причем* $|\mathbf{K}| = 1/\sqrt{p_1 \cdots p_n}$*, где* $p_1, \ldots, p_n$ *— собственные числа матрицы* $\mathbf{A}$*.*

Доказательство. Пусть ортогональная матрица $\mathbf{L}$ преобразованием $\mathbf{L}^T\mathbf{A}\mathbf{L}$ приводит матрицу $\mathbf{A}$ к диагональному виду (D.2).

Так как собственные числа $p_1, \ldots, p_n$ положительны, рассмотрим диагональную матрицу

$$\mathbf{D} = \begin{pmatrix} \frac{1}{\sqrt{p_1}} & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \frac{1}{\sqrt{p_n}} \end{pmatrix} = \{d_{ij}\}.$$

Выполняется равенство

$$\sum_{j,k=1}^{n} d_{ji}\lambda_{jk}d_{ks} = \begin{cases} 0, & i \neq s; \\ d_{ii}\lambda_{ii}d_{ii} = \frac{1}{\sqrt{p_i}}p_i\frac{1}{\sqrt{p_i}} = 1, & i = s; \end{cases}$$

или

$$\mathbf{D}^T\mathbf{\Lambda}\mathbf{D} = \mathbf{E}.$$

Следовательно, матрицу $\mathbf{K}$ можно взять в виде $\mathbf{K} = \mathbf{LD}$ при $|\mathbf{L}| = 1$ и в виде $\mathbf{K} = -\mathbf{LD}$, если $|\mathbf{L}| = -1$. Действительно, $\mathbf{D}^T\mathbf{L}^T\mathbf{ALD} = \mathbf{D}^T\mathbf{\Lambda}\mathbf{D} = \mathbf{E}$.

Определитель матрицы $\mathbf{K}$ равен

$$|\mathbf{K}| = |\pm\mathbf{LD}| = |\pm\mathbf{L}|\,|\mathbf{D}| = \pm(\pm)1 \cdot 1/\sqrt{p_1 \cdots p_n} = 1/\sqrt{p_1 \cdots p_n}.$$

□

Приложение E
Матричное дифференцирование

Производной скалярной функции вектора, $f\left(\mathbf{x}^T\right)$, где $\mathbf{x}^T = (x_1, \ldots, x_n)$ называется вектор

$$\frac{d\left(f\left(\mathbf{x}^T\right)\right)}{d\mathbf{x}} = \begin{pmatrix} \frac{\partial\left(f\left(\mathbf{x}^T\right)\right)}{\partial x_1} \\ \ldots \\ \frac{\partial\left(f\left(\mathbf{x}^T\right)\right)}{\partial x_n} \end{pmatrix}.$$

Покажем, что для линейной формы $\mathbf{x}^T\mathbf{b}$, где $\mathbf{b}^T = (b_1, \ldots, b_n)$, результат этого дифференцирования имеет вид

$$\frac{d\left(\mathbf{x}^T\mathbf{b}\right)}{d\mathbf{x}} = \mathbf{b}.$$

В самом деле,

$$\frac{d\left(\mathbf{x}^T\mathbf{b}\right)}{dx_s} = \frac{\partial}{\partial x_s}\left(\sum_{i=1}^{n} b_i x_i\right) = b_s$$

и поэтому

$$\frac{d\left(\mathbf{x}^T\mathbf{b}\right)}{d\mathbf{x}} = \begin{pmatrix} b_1 \\ \ldots \\ b_n \end{pmatrix} = \mathbf{b}.$$

Покажем, что для симметричной матрицы $\mathbf{A} = \{a_{ij}\}$ справедливо следующее правило дифференцирования:

$$\frac{d\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right)}{d\mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$

Действительно,

$$\frac{\partial\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right)}{\partial x_s}=\frac{\partial}{\partial x_s}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\right)=$$

$$=\frac{\partial}{\partial x_s}\left(\sum_{i=1}^{n}a_{ii}x_i^2+\sum_{\substack{i,j=1\\ i\neq j}}^{n}a_{ij}x_ix_j\right)=$$

$$=2a_{is}x_s+\sum_{\substack{i=1\\ i\neq s}}^{n}a_{is}x_i+\sum_{\substack{j=1\\ j\neq s}}^{n}a_{sj}x_j=2a_{is}x_s+2\sum_{\substack{j=1\\ j\neq s}}^{n}a_{sj}x_j=2\sum_{j=1}^{n}a_{sj}x_j.$$

Производной векторной функции $\mathbf{f}\left(\mathbf{x}^T\right)$, где $\mathbf{f}^T=(f_1,\ldots f_m)$ называется матрица

$$\frac{d\left(\mathbf{f}\left(\mathbf{x}^T\right)\right)}{d\mathbf{x}}=\left\{\frac{\partial f_i\left(\mathbf{x}^T\right)}{\partial x_j}\right\}_m^n.$$

Поэтому ***второй производной скалярной функции*** $f\left(\mathbf{x}^T\right)$ называют матрицу

$$\frac{d^2\left(f\left(\mathbf{x}^T\right)\right)}{d\mathbf{x}^2}=\frac{d}{d\mathbf{x}}\left(\frac{d\left(f\left(\mathbf{x}^T\right)\right)}{d\mathbf{x}}\right)=\left\{\frac{\partial^2\left(f\left(\mathbf{x}^T\right)\right)}{\partial x_i\partial x_j}\right\}_n^n.$$

Следовательно, второй производной квадратичной формы будет выражение

$$\frac{d^2\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right)}{d\mathbf{x}^2}=\frac{d\left(2\mathbf{A}\mathbf{x}\right)}{d\mathbf{x}}=2\mathbf{A},$$

так как

$$\frac{\partial\left(2\mathbf{A}\mathbf{x}\right)}{\partial x_k}=\frac{\partial}{\partial x_k}\left(2\sum_{j=1}^{n}a_{sj}x_j\right)=2a_{sk}.$$

Приложение F
Гамма- и бета-функции

Гамма-функцией (интегралом Эйлера II рода) называется функция вида

$$\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x}\, dx,\ \alpha > 0.$$

Теорема F.1. *Гамма-функция существует.*

Доказательство. Надо показать, что несобственный интеграл, входящий в определение гамма-функции, сходится для любого $\alpha > 0$.

Для этого разобьем его на два интеграла:

$$\int\limits_0^\infty x^{\alpha-1} e^{-x}\, dx = \int\limits_0^1 x^{\alpha-1} e^{-x}\, dx + \int\limits_1^\infty x^{\alpha-1} e^{-x}\, dx.$$

Для первого из них выполняется неравенство

$$0 < \int\limits_0^1 x^{\alpha-1} e^{-x}\, dx < \int\limits_0^1 x^{\alpha-1}\, dx = \left.\frac{x^\alpha}{\alpha}\right|_0^1 = \frac{1}{\alpha}$$

и, следовательно, он сходится.

Для доказательства сходимости второго интеграла сравним его с интегралом

$$\int\limits_1^\infty \frac{dx}{x^2},$$

который, как известно, сходится. Для этого воспользуемся предельным признаком сравнения сходимости несобственных интегралов и вычислим предел

$$\lim_{x\to\infty}\frac{x^{\alpha-1}e^{-x}}{1/x^2}=\lim_{x\to\infty}\frac{x^{\alpha+1}}{e^x}=0.$$

Следовательно, оба интеграла сходятся, значит, сходится и интеграл, определяющий гамма-функцию. □

Теорема F.2. *Справедлива формула*

$$\Gamma(\alpha+1)=\alpha\Gamma(\alpha).$$

Доказательство. Действительно,

$$\begin{aligned}\alpha\Gamma(\alpha)&=\alpha\int\limits_0^\infty x^{\alpha-1}e^{-x}\,dx=\\&=\left\langle u=-e^{-x},dv=x^{\alpha-1}\,dx,du=e^{-x}\,dx,v=\frac{x^\alpha}{\alpha}\right\rangle=\\&=-\alpha\frac{x^\alpha}{\alpha}e^{-x}\Big|_0^\infty+\int\limits_0^\infty x^\alpha e^{-x}\,dx=\int\limits_0^\infty x^\alpha e^{-x}\,dx=\Gamma(\alpha+1).\end{aligned}$$

□

Следствие F.1. *Выполняется равенство*

$$\Gamma(\alpha+n)=(\alpha+n-1)\cdot\ldots\cdot(\alpha+1)\,\alpha\Gamma(\alpha).$$

Доказательство. Применяем n раз только что доказанную формулу. □

Таким образом, достаточно иметь значения гамма-функции для $\alpha\in(0;1]$ или $\alpha\in(1;2]$, чтобы вычислить ее значения для любого α.

Следствие F.2. *При $\alpha = 1$ из предыдущей формулы получаем*

$$\Gamma(n+1) = n!.$$

Доказательство. Действительно, при $\alpha = 1$ из предыдущей формулы следует, что

$$\Gamma(n+1) = (n+1) \cdot \ldots \cdot 2 \cdot 1 \cdot \Gamma(1).$$

Но

$$\Gamma(1) = \int_0^\infty e^{-x}\,dx = -e^{-x}\Big|_0^\infty = 1$$

и поэтому

$$\Gamma(n+1) = (n+1) \cdot \ldots \cdot 2 \cdot 1 \cdot 1 = n!.$$

□

Это говорит о том, что гамма-функция является обобщением понятия факториала.

Еще одно важное значение гамма-функции:

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Действительно, используя интеграл Пуассона (приложение C), получаем

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty x^{-1/2} e^{-x}\,dx = \left\langle x = t^2, dx = 2t\,dt \right\rangle = 2\int_0^\infty t^{-1} e^{-t^2} t\,dt =$$

$$= 2\int_0^\infty e^{-t^2}\,dt = 2\frac{\sqrt{\pi}}{2} = \sqrt{\pi}.$$

Бета-функцией (интегралом Эйлера I рода) называется функция

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx;\ \alpha, \beta > 0.$$

Существует следующая зависимость между бета- и гамма-функцией:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Таблицы этих функций можно найти, например, в [25].

Приложение G
Математические таблицы

G.1 Плотность нормального распределения

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

x	0	1	2	3	4	5	6	7	8	9
0,0	39894	39892	39886	39876	39862	39844	39822	39797	39767	39733
0,1	39695	39654	39608	39559	39505	39448	39387	39322	39253	39181
0,2	39104	39024	38940	38853	38762	38667	38568	38466	38361	38251
0,3	38139	38023	37903	37780	37654	37524	37391	37255	37115	36973
0,4	36827	36678	36526	36371	36213	36053	35889	35723	35553	35381
0,5	35207	35029	34849	34667	34482	34294	34105	33912	33718	33521
0,6	33322	33121	32918	32713	32506	32297	32086	31874	31659	31443
0,7	31225	31006	30785	30563	30339	30114	29887	29659	29431	29200
0,8	28969	28737	28504	28269	28034	27798	27562	27324	27086	26848
0,9	26609	26369	26129	25888	25647	25406	25164	24923	24681	24439

x	0	1	2	3	4	5	6	7	8	9
1,0	24197	23955	23713	23471	23230	22988	22747	22506	22265	22025
1,1	21785	21546	21307	21069	20831	20594	20357	20121	19886	19652
1,2	19419	19186	18954	18724	18494	18265	18037	17810	17585	17360
1,3	17137	16915	16694	16474	16256	16038	15822	15608	15395	15183
1,4	14973	14764	14556	14350	14146	13943	13742	13542	13344	13147
1,5	12952	12758	12566	12376	12188	12001	11816	11632	11450	11270
1,6	11092	10915	10741	10567	10396	10226	10059	09893	09728	09566
1,7	09405	09246	09089	08933	08780	08628	08478	08329	08183	08038
1,8	07895	07754	07614	07477	07341	07206	07074	06943	06814	06687
1,9	06562	06438	06316	06195	06077	05959	05844	05730	05618	05508
2,0	05399	05292	05186	05082	04980	04879	04780	04682	04586	04491
2,1	04398	04307	04217	04128	04041	03955	03871	03788	03706	03626
2,2	03547	03470	03394	03319	03246	03174	03103	03034	02965	02898
2,3	02833	02768	02705	02643	02582	02522	02463	02406	02349	02294
2,4	02239	02186	02134	02083	02033	01984	01936	01888	01842	01797
2,5	01753	01709	01667	01625	01585	01545	01506	01468	01431	01394
2,6	01358	01323	01289	01256	01223	01191	01160	01130	01100	01071
2,7	01042	01014	00987	00961	00935	00909	00885	00861	00837	00814
2,8	00792	00770	00748	00727	00707	00687	00668	00649	00631	00613
2,9	00595	00578	00562	00545	00530	00514	00499	00485	00470	00457

x	0	1	2	3	4	5	6	7	8	9
3,0	00443	00430	00417	00405	00393	00381	00370	00358	00348	00337
3,1	00327	00317	00307	00298	00288	00279	00271	00262	00254	00246
3,2	00238	00231	00224	00216	00210	00203	00196	00190	00184	00178
3,3	00172	00167	00161	00156	00151	00146	00141	00136	00132	00127
3,4	00123	00119	00115	00111	00107	00104	00100	00097	00094	00090
3,5	00087	00084	00081	00079	00076	00073	00071	00068	00066	00063
3,6	00061	00059	00057	00055	00053	00051	00049	00047	00046	00044
3,7	00042	00041	00039	00038	00037	00035	00034	00033	00031	00030
3,8	00029	00028	00027	00026	00025	00024	00023	00022	00021	00021
3,9	00020	00019	00018	00018	00017	00016	00016	00015	00014	00014
4,0	00013	00013	00012	00012	00011	00011	00011	00010	00010	00009

G.2 Функция нормального распределения

$$\Phi^*(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt$$

x	0	1	2	3	4	5	6	7	8	9
0,0	50000	50399	50798	51197	51595	51994	52392	52790	53188	53586
0,1	53983	54380	54776	55172	55567	55962	56356	56750	57142	57535
0,2	57926	58317	58706	59095	59483	59871	60257	60642	61026	61409
0,3	61791	62172	62552	62930	63307	63683	64058	64431	64803	65173
0,4	65542	65910	66276	66640	67003	67364	67724	68082	68439	68793
0,5	69146	69497	69847	70194	70540	70884	71226	71566	71904	72240
0,6	72575	72907	73237	73565	73891	74215	74537	74857	75175	75490
0,7	75804	76115	76424	76731	77035	77337	77637	77935	78230	78524
0,8	78814	79103	79389	79673	79955	80234	80511	80785	81057	81327
0,9	81594	81859	82121	82381	82639	82894	83147	83398	83646	83891
1,0	84134	84375	84614	84850	85083	85314	85543	85769	85993	86214
1,1	86433	86650	86864	87076	87286	87493	87698	87900	88100	88298
1,2	88493	88686	88877	89065	89251	89435	89617	89796	89973	90147
1,3	90320	90490	90658	90824	90988	91149	91308	91466	91621	91774
1,4	91924	92073	92220	92364	92507	92647	92785	92922	93056	93189

x	0	1	2	3	4	5	6	7	8	9
1,5	93319	93448	93574	93699	93822	93943	94062	94179	94295	94408
1,6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1,7	95543	95637	95728	95818	95907	95994	96080	96164	96246	96327
1,8	96407	96485	96562	96638	96712	96784	96856	96926	96995	97062
1,9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97670
2,0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
2,1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2,2	98610	98645	98679	98713	98745	98778	98809	98840	98870	98899
2,3	98928	98956	98983	99010	99036	99061	99086	99111	99134	99158
2,4	99180	99202	99224	99245	99266	99286	99305	99324	99343	99361
2,5	99379	99396	99413	99430	99446	99461	99477	99492	99506	99520
2,6	99534	99547	99560	99573	99585	99598	99609	99621	99632	99643
2,7	99653	99664	99674	99683	99693	99702	99711	99720	99728	99736
2,8	99744	99752	99760	99767	99774	99781	99788	99795	99801	99807
2,9	99813	99819	99825	99831	99836	99841	99846	99851	99856	99861
3,0	99865	99869	99874	99878	99882	99886	99889	99893	99896	99900
3,1	99903	99906	99910	99913	99916	99918	99921	99924	99926	99929
3,2	99931	99934	99936	99938	99940	99942	99944	99946	99948	99950
3,3	99952	99953	99955	99957	99958	99960	99961	99962	99964	99965
3,4	99966	99968	99969	99970	99971	99972	99973	99974	99975	99976

x	0	1	2	3	4	5	6	7	8	9
3,5	99977	99978	99978	99979	99980	99981	99981	99982	99983	99983
3,6	99984	99985	99985	99986	99986	99987	99987	99988	99988	99989
3,7	99989	99990	99990	99990	99991	99991	99992	99992	99992	99992
3,8	99993	99993	99993	99994	99994	99994	99994	99995	99995	99995
3,9	99995	99995	99996	99996	99996	99996	99996	99996	99997	99997
4,0	99997	99998	99999	99999	99999	–	–	–	–	–

G.3 t-распределение Стьюдента

$$P(|T| > t_\alpha) = 2\int\limits_{t_\alpha}^{\infty} f(t)\,dt = \alpha$$

$r\backslash\alpha$	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001
1	3,08	6,31	12,7	31,82	63,66	127,32	318,30	636,31
2	1,89	2,92	4,30	6,96	9,92	14,09	22,33	31,60
3	1,64	2,35	3,18	4,54	5,84	7,45	10,20	12,92
4	1,53	2,13	2,78	3,75	4,60	5,60	7,17	8,61
5	1,48	2,02	2,57	3,36	4,03	4,77	5,89	6,87
6	1,44	1,94	2,45	3,14	3,71	4,32	5,21	5,96
7	1,41	1,89	2,36	3,00	3,50	4,03	4,79	5,41
8	1,40	1,86	2,31	2,90	3,36	3,83	4,50	5,04
9	1,38	1,83	2,26	2,82	3,25	3,69	4,30	4,78
10	1,37	1,81	2,23	2,76	3,17	3,58	4,14	4,59
11	1,36	1,80	2,20	2,72	3,11	3,50	4,02	4,44
12	1,36	1,78	2,18	2,68	3,05	3,43	3,93	4,32
13	1,35	1,77	2,16	2,65	3,01	3,37	3,85	4,22
14	1,35	1,76	2,14	2,62	2,98	3,33	3,79	4,14
15	1,34	1,75	2,13	2,60	2,95	3,29	3,73	4,07

$r\backslash\alpha$	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001
16	1,34	1,75	2,12	2,58	2,92	3,25	3,69	4,02
17	1,33	1,74	2,11	2,57	2,90	3,22	3,65	3,97
18	1,33	1,73	2,10	2,55	2,88	3,20	3,61	3,92
19	1,33	1,73	2,09	2,54	2,86	3,17	3,58	3,88
20	1,33	1,72	2,09	2,53	2,85	3,15	3,55	3,85
21	1,32	1,72	2,08	2,52	2,83	3,14	3,53	3,82
22	1,32	1,72	2,07	2,51	2,82	3,12	3,51	3,79
23	1,32	1,71	2,07	2,50	2,81	3,10	3,49	3,77
24	1,32	1,71	2,06	2,49	2,80	3,09	3,47	3,75
25	1,32	1,71	2,06	2,49	2,79	3,08	3,45	3,73
26	1,32	1,71	2,06	2,48	2,78	3,07	3,44	3,71
27	1,31	1,70	2,05	2,47	2,77	3,06	3,42	3,69
28	1,31	1,70	2,05	2,47	2,76	3,05	3,41	3,67
29	1,31	1,70	2,05	2,46	2,76	3,04	3,40	3,66
30	1,31	1,70	2,04	2,46	2,75	3,03	3,38	3,65
40	1,30	1,68	2,02	2,42	2,70	2,97	3,31	3,55
60	1,30	1,67	2,00	2,39	2,66	2,91	3,23	3,46
120	1,29	1,66	1,98	2,36	2,62	2,86	3,16	3,37
∞	1,28	1,64	1,96	2,33	2,58	2,81	3,09	3,29

G.4 χ^2-распределение Пирсона

$$P(\chi^2 > u_\gamma) = \int\limits_{u_\gamma}^{\infty} f(u)\,du = \gamma$$

$r\backslash\gamma$	0,99	0,98	0,95	0,90	0,80	0,20	0,10	0,05	0,02	0,01
1	0,00016	0,0063	0,0039	0,0158	0,0642	1,642	2,706	3,841	5,412	6,635
2	0,0201	0,0404	0,103	0,211	0,446	3,219	4,605	5,991	7,824	9,210
3	0,115	0,185	0,352	0,584	1,005	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,610	2,343	7,289	9,236	11,070	13,388	15,086
6	0,872	1,134	1,635	2,204	3,070	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	9,803	12,017	14,067	16,622	18,475
8	1,646	2,032	2,733	3,490	4,594	11,030	13,362	15,507	18,168	20,090
9	2,088	2,532	3,325	4,168	5,380	12,242	14,684	16,919	19,679	21,666
10	2,558	3,059	3,940	4,865	6,179	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	15,812	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,042	8,634	16,985	19,812	22,362	25,472	27,688
14	4,660	5,368	6,571	7,790	9,467	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	19,311	22,307	24,996	28,259	30,578

$r\backslash\gamma$	0,99	0,98	0,95	0,90	0,80	0,20	0,10	0,05	0,02	0,01
16	5,812	6,614	7,962	9,312	11,152	20,465	23,542	26,296	29,633	32,000
17	6,408	7,255	8,672	10,085	12,002	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,390	10,865	12,857	22,760	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	23,900	27,204	30,144	33,687	36,191
20	8,260	9,237	10,851	12,443	14,578	25,038	28,412	31,410	35,020	37,566
21	8,897	9,915	11,591	13,240	15,445	26,171	29,615	32,671	36,343	38,932
22	9,542	10,600	12,338	14,041	16,314	27,301	30,813	33,924	37,660	40,289
23	10,196	11,293	13,091	14,848	17,187	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	29,553	33,196	36,415	40,270	42,980
25	11,524	12,697	14,611	16,473	18,940	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,820	31,795	35,563	38,885	42,856	45,642
27	12,879	14,125	16,151	18,114	20,703	32,912	36,741	40,113	44,140	46,963
28	13,565	14,847	16,928	18,939	21,588	34,027	37,916	41,337	45,419	48,278
29	14,256	15,574	17,708	19,768	22,475	35,139	39,087	42,557	46,693	49,588
30	14,953	16,306	18,493	20,599	23,364	36,250	40,256	43,773	47,962	50,892

G.5 Функция $P(\lambda)$

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,0	1,000	0,7	0,711	1,4	0,040
0,1	1,000	0,8	0,544	1,5	0,022
0,2	1,000	0,9	0,393	1,6	0,012
0,3	1,000	1,0	0,270	1,7	0,006
0,4	0,997	1,1	0,178	1,8	0,003
0,5	0,964	1,2	0,112	1,9	0,002
0,6	0,864	1,3	0,068	2,0	0,001

Предметный указатель

аксиоматическое определение вероятности, 15
амплитудно-частотная характеристика, 192
бета-функция, 303
вариационный ряд, 229
вероятность, 16
внутригрупповая дисперсия, 283
выборка, 229
выборочная дисперсия, 246
выборочное корреляционное отношение, 284
выборочное с.к.о., 247
выборочное среднее, 244
выборочный коэффициент корреляции, 280
выборочный коэффициент регрессии, 280
гамма-функция, 301
генеральная совокупность, 229
геометрическое определение вероятности, 23
гипотеза
– альтернативная, 265
– двусторонняя критическая область, 266
– задача проверки согласия, 268
– конкурирующая, 265
– критерий χ^2, 268
– критерий проверки, 266
– критерий согласия, 268
– критическая область, 266
– критические точки, 266
– левосторонняя критическая область, 266
– мощность критерия, 266
– нулевая, 265
– область принятия, 266
– основная, 265
– ошибка
- второго рода, 266
- первого рода, 266
– правосторонняя критическая область, 266
– простая, 265
– сложная, 266
– статистическая, 265
– уровень значимости, 266
гистограмма, 239
группировка выборки, 231
групповая дисперсия, 283

групповое среднее, 283
детерминированная величина, 75
дисперсия, 82
– дискретной с.в., 83
– непрерывной с.в., 83
– остаточная, 94
– случайного процесса, 158
доверительная вероятность, 255
доверительный интервал, 255
закон распределения
– дискретной с.в., 42
– непрерывной с.в., 36
– СВДТ, 53
– СВНТ, 48
– условный, 59
индикатор с.с., 146
интеграл
– в среднеквадратичном, 184
– вероятностей, 118
– Пуассона, 294
интенсивность
– гибели, 170
– размножения, 170
интервал группировки, 231
исправленная выборочная дисперсия, 247
исправленное выборочное с.к.о., 247
классическое определение вероятности, 20
ковариационная
– матрица, 90
– функция, 158
ковариация, 88
композиция законов распределения
– дискретная, 71
– непрерывная, 70
конечная цепь Маркова
– с дискретным временем, 161
– с непрерывным временем, 161
корреляционная зависимость, 91
корреляционная функция, 158
коэффициент
– корреляции, 90
– регрессии, 93
криволинейная регрессия, 286
линейная динамическая система, 192
ложное среднее, 249
марковская цепь
– вектор вероятностей состояний, 164
– вектор начальных вероятностей, 164
– вероятности перехода, 161
– вероятность перехода за n шагов, 162
– вероятность пребывания в данном состоянии, 164
– граф интенсивностей переходов, 168
– граф состояний, 164
– интенсивность перехода, 167

– интенсивность ухода, 167
– матрица вероятностей переходов, 163
– матрица интенсивностей переходов, 167
– однородная, 162
– ординарность, 168
– период состояния, 178
– предельное распределение вероятностей, 177
– сообщающиеся состояния, 177
– состояние
- непериодическое, 178
- несущественное, 177
- периодическое, 178
- существенное, 177
– стационарное распределение вероятностей, 174
– эргодическая, 177
математическое ожидание
– дискретной с.в., 74
– непрерывной с.в., 74
– случайного вектора, 75
– случайного процесса, 158
межгрупповая дисперсия, 283
метод наименьших квадратов, 92
многоугольник распределения, 40
множество
– континуальное, 14
– несчетное, 14
– счетное, 14
момент k-го порядка, 85
наблюдение, 229
надежность, 255
неравенство Маркова
- второе, 147
- первое, 147
неравенство Чебышева
– второе, 147
– первое, 147
объем выборки, 229
опыт, 13
оценка параметра, 243
– методом наибольшего правдоподобия, 252
– несмещенная, 243
– состоятельная, 243
– точечная, 243
параметр распределения, 243
передаточная функция, 192
плотность вероятностей
– случайного вектора, 48
– случайной величины, 36
– условная, 58
полигон относительных накопленных частот, 238
полигон относительных частот, 240
полиномиальные коэффициенты, 104, 292
положительно определенная
– квадратичная форма, 296
– матрица, 296
поток
– входящий, 200

– выходящий, 200
– обслуживания, 204
- простейший, 204
– простейший, 200
– пуассоновский, 203
– Эрланга, 225
- нормированный, 226
– регулярный, 226
правило трех сигм, 121
производная
– векторной функции, 300
– скалярной функции вектора, 299
– функции вектора
– вторая, 300
пространство элементарных с.с., 13
прямая линейной среднеквадратической регрессии, 93
разбиение, 19
размах выборки, 229
распределение
– χ^2, 141
– t, 138
– биномиальное, 99
– геометрическое, 111
– гипергеометрическое, 112
– нормальное
- двумерное, 128
- многомерное, 121
- одномерное, 116
– Пирсона, 141
– показательное, 130
– полиномиальное, 104
– пуассоновское, 108
- многомерное, 110
– равномерное, 114
– Стьюдента, 138
– Эрланга, 133
- нормированное, 136
ряд распределения, 40
сглаживание стационарного с.п., 194
система массового обслуживания, СМО, 199
– абсолютная пропускная способность, 206
– вероятность отказа, 206
– дисциплина обслуживания, 200
– замкнутая, 200
– заявка, 199
– канал обслуживания, 200
– коэффициент вариации времени обслуживания, 216
– коэффициент загрузки системы, 209
– марковская, 204
– многоканальная, 200
– одноканальная, 200
– открытая, 200
– относительная пропускная способность, 206
– с отказами, 200
– с очередью, 200
– с приоритетом, 200
– смешанного типа, 200

– событие, 200
– среднее время пребывания заявки в очереди, 206
– среднее время пребывания заявки в СМО, 206
– среднее число занятых каналов, 207
– среднее число заявок в очереди, 206
– среднее число заявок в СМО, 206
– требование, 199
случайная величина, 33
– бернуллиевская, 99
– биномиальная, 99
– геометрическая, 111
– гипергеометрическая, 112
– дискретная, 39
– непрерывная, 36
– нормальная
- двумерная, 128
- многомерная, 121
- одномерная, 116
– показательная, 130
– пуассоновская, 108
- многомерная, 110
– равномерная, 114
– эрланговская, 133
- нормированная, 136
случайная функция, 60
– дискретного аргумента, 62
– непрерывного аргумента, 60
– нескольких аргументов, 65
– СВДТ, 68
– СВНТ, 65
случайное событие, 15
– достоверное, 15
– невозможное, 15
– относительная частота, 16
– частота, 16
случайные величины
– зависимые, 56
– коррелированные, 90
– независимые, 56, 149
– некоррелированные, 90
случайные события
– алгебра, 15
– благоприятствующие, 16
– вероятности, 16
– независимые, 26
– несовместные, 15
– операции
- объединение, 15
- произведение, 15
- разность, 15
- сумма, 15
– противоположные, 14
– равновероятные, 19
– совместные, 15
– эквивалентные, 16
– элементарные, 13
случайный вектор, 45
– дискретного типа, СВДТ, 52
– непрерывного типа, СВНТ, 48

случайный процесс, 156
– без последействия, 161
– интеграл в среднеквадратичном, 184
– марковский, 161
– состояние, 161
– непрерывность в среднеквадратичном, 181
– производная, 182
– пуассоновский, 158
– размножения и гибели, 170
– реализация, 156
– с дискретным временем, 158
– с независимыми приращениями, 158
– с непрерывным временем, 158
– с непрерывным спектром, 190
– сечение, 157
– стационарный, 187
– траектория, 156
– функция распределения, 157
– чистого размножения, 201
спектральная плотность, 190
среднее квадратическое отклонение (с.к.о.), 83
средний интервал корреляции, 191
статистика, 256
статистическая линеаризация, 95
статистическая проверка гипотез, 265
статистический ряд
– группиро́ванный, 232
– дискретного типа, 230
– непрерывного типа, 232
схема Бернулли, 98
– наивероятнейшее число успехов, 100
– число успехов, 98
сходимость
– в среднеквадратичном, 181
– по вероятности, 148
таблица частот, 235
теорема
– Бернулли, 151
– Муавра-Лапласа
- интегральная, 145
- локальная, 144
– сложения
- для несовместных событий, 17
- для совместных событий, 17
– умножения
- для зависимых событий, 26
- для независимых событий, 26
– центральная предельная, 152
урновая модель, 20
условная вероятность, 26
формула

– Байеса, 31
– полной вероятности, 30
формулы
– Полячека-Хинчина, 216
– Эрланга, 209
функция
– правдоподобия, 251
– производящая, 79
– распределения
- случайного вектора, 45
- случайного процесса, 157
- случайной величины, 34
- теоретическая, 235
- частная, 46
- эмпирическая, 235
центральный момент k-го порядка, 85
частота, 230, 232
– накопленная относительная, 235
– накопленная, 234
– относительная, 25, 234
– теоретическая, 268
эффективная ширина спектра, 190

Литература

[1] Башарин Г.П., Бочаров П.П., Коган Я.А. Анализ очередей в вычислительных сетях. Теория и методы расчета. — М.: Наука, 1989.

[2] Вентцель Е.С. Теория вероятностей. — М.: Физматгиз, 1962.

[3] Вентцель Е.С., Овчаров Л.А. Теория вероятностей. Задачи и упражнения. — М.: Наука, 1973.

[4] Вентцель Е.С. Исследование операций: задачи, принципы, методология. — М.: Наука, 1988.

[5] Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения. — М.: Наука, 1991.

[6] Волченко Ю.М. Структурно-логические модели систем обслуживания с приоритетами. — Электронное моделирование, № 3, 1990.

[7] Волченко Ю.М. Базові математичні поняття. — Донецьк: ДонІЗТ, 2005.

[8] Гмурман В.Е. Введение в теорию вероятностей и математическую статистику. — М.: Высшая школа, 1966.

[9] Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. — М.: Высшая школа, 1970. статистику. — М.: Высшая школа, 1966.

[10] Гнеденко Б.В. Курс теории вероятностей. — М.: Физматгиз, 1961.

[11] Данко П.Е., Попов А.Г., Кожевникова Т.Я. Высшая математика в упражнениях и задачах. — М.: Высшая школа, 1986, ч. II.

[12] Жевержеев В.Ф., Кальницкий Л.А, Сапогов Н.А. Специальный курс высшей математики для втузов. — М.: Высшая школа, 1970.

[13] Захаров В.К., Севастьянов Б.А., Чистяков В.П. Теория вероятностей. — М.: Наука, 1983.

[14] Карлин С. Основы теории случайных процессов. — М.: Мир, 1971.

[15] Клейнрок Л. Теория массового обслуживания. — М.: Машиностроение, 1979.

[16] Колде Я.К. Практикум по теории вероятностей и математической статистике. — М.: Высшая школа, 1991.

[17] Кузин Л.Т. Основы кибернетики. Математические основы кибернетики. — М.: Энергия, 1973, т. 1.

[18] Кнут Д. Искусство программирования для ЭВМ. Получисленные алгоритмы. — М., Мир, 1977, т. 2.

[19] Мантуров О.В. Курс высшей математики. — М.: Высшая школа, 1991.

[20] Матвеев В.Ф., Ушаков В.Г. Системы массового обслуживания. — М.: МГУ, 1984.

[21] Прохоров Ю.В., Розанов Ю.А. Теория вероятностей (Основные понятия. Предельные теоремы. Случайные процессы). — М.: Наука, 1973.

[22] Рао С. Р. Линейные статистические методы и их применение. — М.: Наука, 1968.

[23] Сборник задач по математике для втузов. Специальные курсы / Под ред. А. В. Ефимова, Б. П.Демидовича. — М.: Наука, 1984, т. 3.

[24] Свешников А.А. и др. Сборник задач по теории вероятностей, математической статистике и теории случайных функций. — М.: Наука, 1965.

[25] Справочник по специальным функциям с формулами, графиками и математическими таблицами //Под ред. Н.Абрамовица и И. Стиган. — М.:, 1979.

[26] Теория автоматического регулирования. / Под ред. В.В. Солодовникова. — М.: Машиностроение, 1967, кн. 2.

[27] Чистяков В.П. Курс теории вероятностей. — М.: Наука, 1982.

[28] Ширяев А.Н. Вероятность. — М.: Наука, 1980.