



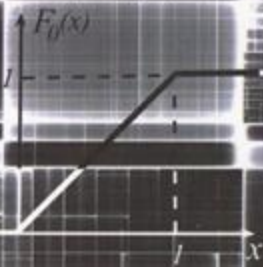
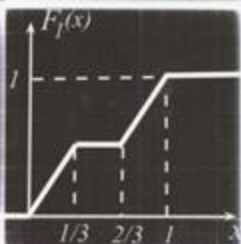
$$\frac{S_g}{S_R} = \frac{\pi \frac{R^2}{4}}{\pi R^2} = \frac{1}{4}$$

$$\frac{\mu(0, \frac{R}{2})}{R} = \frac{\frac{R}{2}}{R} = \frac{1}{2}$$

В. М. Буре, Е. М. Парилина

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА


$$F_2(x) = \begin{cases} 0, & x = 0, \\ 1/4, & x \in (1/9, 2/9), \\ 1/2, & x \in (1/3, 2/3), \\ 3/4, & x \in (7/9, 8/9), \\ 1, & x = 1, \end{cases}$$



В. М. БУРЕ,
Е. М. ПАРИЛИНА

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ДОПУЩЕНО

*УМО по классическому университетскому образованию
в качестве учебника для студентов вузов,
обучающихся по направлениям ВПО
010400 — «Прикладная математика и информатика»
и 010300 — «Фундаментальная информатика
и информационные технологии»*



САНКТ-ПЕТЕРБУРГ •
МОСКВА •
КРАСНОДАР •
2013 •

ББК 22.17я73

К 60

Буре В. М., Парилина Е. М.

К 60 Теория вероятностей и математическая статистика: Учебник. — СПб.: Издательство «Лань», 2013. — 416 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-1508-3

В книге изложены основные разделы современного курса теории вероятностей и математической статистики, включая условные распределения, основные предельные теоремы, метод характеристических функций, принципы статистического оценивания, методы построения доверительных интервалов, методы проверки статистических гипотез, регрессионный анализ и бинарная регрессия. Книга предназначена для студентов, изучающих методы теории вероятностей, математической статистики.

ББК 22.17я73

Рецензент:

А. В. ФЛЕГОНТОВ — доктор физико-математических наук, профессор, зав. кафедрой информационных систем и программного обеспечения факультета информационных технологий РГПУ им. А. И. Герцена.

Обложка

Е. А. ВЛАСОВА

*Охраняется законом РФ об авторском праве.
Воспроизведение всей книги или любой ее части запрещается без
письменного разрешения издателя.
Любые попытки нарушения закона
будут преследоваться в судебном порядке.*

© Издательство «Лань», 2013
© В. М. Буре, Е. М. Парилина, 2013
© Издательство «Лань»,
художественное оформление, 2013

ПРЕДИСЛОВИЕ

Настоящее пособие содержит подробное изложение университетского курса теории вероятностей и математической статистики.

В книге представлены основные разделы курса теории вероятностей: свойства вероятностной меры, основные формулы элементарной вероятности, классическое определение вероятности, геометрическая вероятность, различные вероятностные пространства, классические предельные теоремы для схемы Бернулли, цепи Маркова, математическое ожидание как интеграл Лебега, условные распределения и условные математические ожидания, виды сходимостей случайных величин, характеристические функции, классические предельные теоремы, некоторые специальные распределения.

При написании раздела, посвященного теории вероятностей, были использованы материалы из книг А. Н. Ширяева [35], А. А. Боровкова [4], Б. А. Севастьянова [30], К. Парасарати [27], В. В. Петрова [28], В. Феллера [33], В. Б. Невзорова [25]. Необходимо также отметить книги [2, 6, 13, 15, 19, 21, 22, 24, 31].

В раздел математической статистики также вошли основные разделы курса, включая: выборочное пространство, предельные распределения ряда статистик, критерии согласия для простых и сложных гипотез, точечные оценки и методы их построения, приближенные и точные доверительные интервалы, различные статистические критерии, таблицы сопряженности, регрессионный анализ, логит и пробит модели.

При написании раздела курса, посвященного математической статистике, были использованы материалы из книг А. А. Боровкова [5], Б. А. Севастьянова [30], С. С. Валландера [11], С. Р. Рао [29], С. М. Ермакова и А. А. Жиглявского [14],

И. А. Ибрагимова и Р. З. Хасьминского [17], Я. Ю. Никитина [26], Г. Крамера [20], Р. Е. Greenwood, М. S. Nikulin [47], также отметим книги [1, 3, 6, 10, 18, 23, 32, 34, 41, 43, 46, 49, 51].

Вопросам практического применения различных статистических методик посвящены следующие работы [3, 6, 9, 32, 34, 36, 38, 42, 44, 50, 52], которые могут быть использованы при проведении практических занятий по курсу теории вероятностей и математической статистики.

Представленный курс теории вероятностей и математической статистики читался в течение многих лет на факультете прикладной математики — процессов управления Санкт-Петербургского государственного университета.

Авторы книги стремились сделать изложение понятным и доступным самому широкому кругу студентов, изучающих основы теории вероятностей и математической статистики, сохранить все базовые теоремы и подробно изложить теоретические основы курса, включая интеграл Лебега, а также включить в курс широкий круг методик статистического анализа данных, часто применяемых в прикладных статистических исследованиях. Для проведения практических занятий по курсу теории вероятностей и математической статистики авторы рекомендуют использовать задачки Г. В. Емельянова и В. П. Скитовича [12] и А. М. Зубкова, Б. А. Севастьянова и В. П. Чистякова [16].

В классических учебниках по теории вероятностей и математической статистике редко можно встретить изложение основ регрессионного анализа. Однако в связи с большим прикладным значением этого раздела математической статистики авторы сочли необходимым включить его в настоящую книгу. При написании раздела использовался материал из книг С. С. Валландера [11] и W. H. Greene [46]. Еще одной важной прикладной методикой анализа данных является бинарная регрессия, в частности, модели логит и пробит анализов. При написании главы 19 использовались статьи и монографии [39, 46], [48]–[52].

Авторы признательны студентам Рубша Алене, Манушкиной Таисе, Хамматовой Гузель и Буре Артему, участвовавшим в подготовке рукописи.

ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО

§ 1. СЛУЧАЙНЫЕ СОБЫТИЯ

Пусть проводится некоторый опыт, в результате которого может произойти одно из *элементарных событий*. Под элементарными событиями будем понимать события (исходы), которые нельзя разделить на «составные части», также являющиеся событиями. Объединение элементарных событий образует пространство (множество) элементарных событий, которое обычно обозначается через Ω .

Пример 1.1. Пусть проводится эксперимент, в котором подкидывается 2 кубика. Тогда

$$\Omega = \{(a_1, a_2) : a_i = 1, \dots, 6\},$$

где a_i — число очков, выпавших на i -ом кубике.

В этом случае мощность множества Ω конечна и равна 36.

Пример 1.2. Проведем другой статистический эксперимент, в котором человек приходит на остановку и ждет прихода автобуса. Предположим, что автобус гарантированно приходит в промежутке времени $[0, T]$. Тогда моменты времени, когда приедет автобус, образуют множество элементарных исходов $\Omega = [0, T]$. В этом случае Ω более, чем счетно, т. е. множество Ω имеет мощность континуум.

Если в примере 1.1 нас интересует событие, которое заключается в том, что на обоих гранях выпадает четное число

очков, тогда событие A можно описать следующим образом: $\{(2k, 2l) : k, l = 1, 2, 3\}$. Если в примере 1.2 наложить условие, что время ожидания автобуса не превышает 15 минут, тогда событие B представляет собой отрезок $[0, 15]$.

События A и B не являются элементарными. Любое событие — совокупность элементарных событий, т.е. любое событие $A \subset \Omega$.

Будем говорить, что событие A произошло в том и только в том случае, если произошло элементарное событие $\omega \in A$. Событие, заключающееся в том, что происходит хотя бы одно из событий A или B , обозначим через $A \cup B = \{\omega : (\omega \in A) \text{ или } (\omega \in B)\}$ (рис. 1.1). Событие, которое происходит, когда происходят одновременно и событие A , и событие B , обозначим через $A \cap B = \{\omega : (\omega \in A) \text{ и } (\omega \in B)\}$.

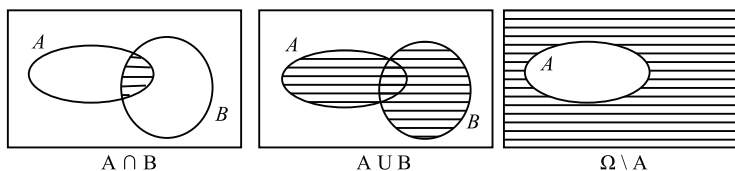


Рис. 1.1
Диаграммы Венна

Если $A \cap B = \emptyset$, то события A и B называют *несовместными*. Пустое множество \emptyset в теории вероятности принято называть *невозможным* событием, а множество Ω — *достоверным* событием.

Обозначим через $\bar{A} = \Omega \setminus A = \{\omega : \omega \notin A\}$ *дополнение* события A . События A и \bar{A} — несовместные события, т.е. $A \cap \bar{A} = \emptyset$. Также верны следующие равенства:

$$A = \bar{\bar{A}} = \Omega \setminus \bar{A},$$

$$A \cup \bar{A} = \Omega.$$

Через $A \setminus B$ обозначим множество точек из множества Ω , принадлежащих A , но не принадлежащих B , т.е. $A \setminus B = \{\omega : (\omega \in A) \text{ и } (\omega \notin B)\}$.

§ 2. АКСИОМАТИКА

Рассмотрим множества A_1, \dots, A_m такие, что выполнены следующие условия:

1. Пересечение любых двух различных множеств является пустым множеством: $A_i \cap A_j = \emptyset$ для всех $i, j, i \neq j$.
2. Объединение всех множеств $A_1 \cup A_2 \cup \dots \cup A_m$ совпадает с множеством Ω .

Определение 2.1. Система множеств $\{A_i\}_{i=1}^m$, удовлетворяющая условиям 1 и 2, называется разбиением множества Ω или полной группой событий.

Определение 2.2. Совокупность \mathcal{A} подмножеств множества Ω называется алгеброй, если выполнены следующие условия:

1. Множество Ω принадлежит совокупности \mathcal{A} .
2. Если множества A и B принадлежат совокупности \mathcal{A} , тогда их объединение $A \cup B$ также принадлежит \mathcal{A} .
3. Для любого множества A из \mathcal{A} его дополнение $\bar{A} = \Omega \setminus A$ также принадлежит \mathcal{A} .

Множества, входящие в алгебру \mathcal{A} , будем называть событиями.

Пример 2.1. Приведем несколько примеров алгебр событий:

1. $\mathcal{A}_0 = \{\emptyset, \Omega\}$.
2. $\mathcal{A}_1 = 2^\Omega = \{A : A \subset \Omega\}$ — алгебра, содержащая все подмножества Ω .
3. $\mathcal{A}_2 = \{A, \bar{A}, \emptyset, \Omega\}$, где $A \subset \Omega$.

Пусть существуют множества $\{A_\gamma\} \subset \Omega$, $\gamma \in \Gamma$. Справедливы законы двойственности:

$$\overline{\bigcup_{\gamma \in \Gamma} A_\gamma} = \bigcap_{\gamma \in \Gamma} \bar{A}_\gamma. \quad (2.1)$$

$$\overline{\bigcap_{\gamma \in \Gamma} A_\gamma} = \bigcup_{\gamma \in \Gamma} \bar{A}_\gamma. \quad (2.2)$$

Доказательство. Докажем равенство (2.1). Покажем, что множество слева от знака равенства $\overline{\bigcup_{\gamma \in \Gamma} A_\gamma}$ является подмножеством множества справа от знака равенства $\bigcap_{\gamma \in \Gamma} \bar{A}_\gamma$,

и наоборот. Рассмотрим произвольное элементарное событие ω из множества $\bigcup_{\gamma \in \Gamma} \bar{A}_\gamma$. Очевидно, что событие ω не принадлежит объединению $\bigcup_{\gamma \in \Gamma} A_\gamma$, что означает, что событие ω не принадлежит ни одному множеству A_γ , $\gamma \in \Gamma$. Следовательно, событие ω принадлежит всем дополнениям \bar{A}_γ для любого $\gamma \in \Gamma$. Это означает, что событие ω принадлежит и пересечению этих событий $\bigcap_{\gamma \in \Gamma} \bar{A}_\gamma$.

Докажем включение в обратную сторону. Рассмотрим событие ω из пересечения $\bigcap_{\gamma \in \Gamma} \bar{A}_\gamma$. Очевидно, что $\omega \in \bar{A}_\gamma$ для всех $\gamma \in \Gamma$, из чего следует, что событие ω не принадлежит ни одному из множеств A_γ , $\gamma \in \Gamma$. Тогда можно утверждать, что ω не принадлежит и объединению $\bigcup_{\gamma \in \Gamma} A_\gamma$. Следовательно, $\omega \in \overline{\bigcup_{\gamma \in \Gamma} A_\gamma}$.

Аналогично доказывается утверждение (2.2). ■

Используя законы двойственности, легко показать, что всякая алгебра замкнута относительно операции конечного пересечения своих элементов. Т.е. для любой алгебры \mathcal{A} справедливо, что, если $A \in \mathcal{A}$ и $B \in \mathcal{A}$, тогда $A \cap B \in \mathcal{A}$. Для доказательства рассмотрим выражение $\overline{A \cap B}$. Применим второе утверждение законов двойственности: $\overline{A \cap B} = \bar{A} \cup \bar{B} \in \mathcal{A}$. По определению алгебры (определение 2.2) $\overline{A \cap B} \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$.

Определение 2.3. Алгеброй, порожденной множеством $B \subset \subset \Omega$, называется система подмножеств $\alpha(B) = \{B, \bar{B}, \emptyset, \Omega\}$.

Рассмотрим некоторое событие $A \in \mathcal{A}$. Проведем статистический эксперимент N раз. В этом эксперименте событие A произошло $N(A)$ раз. Тогда отношение $N(A)/N$ — это относительная частота появления события A . При $N \rightarrow \infty$ относительная частота стабилизируется, но о сходимости говорить нельзя. Можно говорить о «статистической устойчивости» относительных частот. Вероятность появления события A , которую обозначим через p , приближенно равна $N(A)/N$.

Замечание 2.1. Теория вероятностей может применяться только к тем экспериментам, в которых наблюдается «статистическая устойчивость» относительных частот.

Определение 2.4. Пусть \mathcal{A} — алгебра подмножеств множества Ω . Числовая функция $\mu(\cdot)$, заданная на алгебре \mathcal{A} , называется конечно-аддитивной мерой, если выполнены следующие условия:

1. Для любого множества $A \in \mathcal{A}$: $\mu(A) \geq 0$.
2. Для любых множеств $A_1, A_2 \in \mathcal{A}$ таких, что $A_1 \cap A_2 = \emptyset$, имеет место равенство:

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2).$$

3. Значение функции μ для пустого множества равно нулю. Конечно-аддитивная мера μ называется *конечной*, если $\mu(\Omega)$ конечно. Если $\mu(\Omega) = 1$, то конечно-аддитивная мера μ называется *конечно-аддитивной вероятностной мерой* и обычно обозначается через P .

Замечание 2.2. Конечно-аддитивная вероятностная мера P , определенная на алгебре \mathcal{A} , обладает следующими свойствами:

1. Вероятность $P(A)$ события A неотрицательна для любого $A \in \mathcal{A}$.
2. Вероятность множества Ω равна единице, $P(\Omega) = 1$.
3. Для любых множеств $A_1, A_2 \in \mathcal{A}$ таких, что $A_1 \cap A_2 = \emptyset$, имеет место равенство $P(A_1 \cup A_2) = P(A_1) + P(A_2)$.

Определение 2.5. Пусть \mathcal{A} — алгебра подмножеств множества Ω . Конечно-аддитивная мера $\mu(\cdot)$, заданная на алгебре \mathcal{A} , называется счетно-аддитивной (σ -аддитивной) мерой, если выполнено следующее условие: если имеется счетная совокупность событий $\{A_i\}_{i=1}^{\infty}$, где $A_i \in \mathcal{A}$ для любого i , что $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, $A_i \cap A_j = \emptyset$ для всех $i \neq j$, тогда имеет место равенство:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Определение 2.6. Счетно-аддитивная мера μ называется σ -конечной, если существует такое счетное разбиение множества $\Omega = \bigcup_{i=1}^{\infty} D_i$, причем для любого i : $D_i \in \mathcal{A}$ и $D_i \cap D_j = \emptyset$ для $i \neq j$, что $\mu(D_i) < \infty$.

Определение 2.7. Счетно-аддитивная мера μ называется конечной счетно-аддитивной, если $\mu(\Omega) < \infty$. Если $\mu(\Omega) = 1$, то счетно-аддитивная мера μ называется *счетно-аддитивной вероятностной мерой* или вероятностью.

Замечание 2.3. Счетно-аддитивная вероятностная мера P , определенная на \mathcal{A} , обладает следующими свойствами:

1. Вероятность любого события неотрицательна $P(A) \geq 0$ для любого множества $A \in \mathcal{A}$.
2. Вероятность множества Ω равна единице, $P(\Omega) = 1$.
3. Для любых множеств $A_1, A_2 \in \mathcal{A}$ таких, что $A_1 \cap A_2 = \emptyset$, имеет место равенство $P(A_1 \cup A_2) = P(A_1) + P(A_2)$.
4. Если $\{A_i\}_{i=1}^{\infty}$ — последовательность событий, где $A_i \in \mathcal{A}$ для любого i , и выполняется: $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, и $A_i \cap A_j = \emptyset$ для всех $i \neq j$, тогда справедливо равенство:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

§ 3. СВОЙСТВА ВЕРОЯТНОСТЕЙ

Сформулируем и докажем свойства, которыми обладает вероятностная мера.

1. Вероятность невозможного события равна нулю, $P(\emptyset) = 0$.

Доказательство. Очевидны следующие равенства: $\Omega \cup \emptyset = \Omega$, $\Omega \cap \emptyset = \emptyset$. Так как вероятностная мера P является конечной, то из свойства аддитивности получаем следующие равенства: $P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) = 1$. Окончательно получаем: $P(\emptyset) = 0$. ■

2. Для любого события $A \in \mathcal{A}$ вероятность $P(A)$ равна $1 - P(\bar{A})$.

Доказательство. Так как $A \cup \bar{A} = \Omega$, тогда выполняется следующее равенство: $P(A) + P(\bar{A}) = P(\Omega) = 1$, следовательно, $P(A) = 1 - P(\bar{A})$. ■

3. Для событий $A, B \in \mathcal{A}$, $A \subset B$, справедливо неравенство: $P(A) \leq P(B)$.

Доказательство. Поскольку $A \subset B$, то множество B можно представить следующим образом: $B = A \cup (B \setminus A)$, где $B \setminus A = \{\omega : \omega \in B, \omega \notin A\}$. Из условия аддитивности вероятностной меры следует равенство $P(B) = P(A) + P(B \setminus A)$. При этом $P(B \setminus A) \geq 0$. Заметим, что $A \neq B \cup (A \setminus B)$, так как $A \setminus B = \emptyset$. ■

4. *Формула сложения вероятностей.* Для любых событий A и B из \mathcal{A} справедливо равенство:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Доказательство. Представим событие $A \cup B$ в следующем виде $(A \setminus B) \cup (B \setminus A) \cup (A \cap B)$. Имеет место равенство $P(A \cup B) = \{P(A \setminus B) + P(A \cap B)\} + \{P(B \setminus A) + P(A \cap B)\} - P(A \cap B)$. Так как $P(A \setminus B) + P(A \cap B) = P(A)$ и $P(B \setminus A) + P(A \cap B) = P(B)$, можно утверждать, что справедливо равенство $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. ■

5. Для любых событий $A_1, \dots, A_m \in \mathcal{A}$ выполняется неравенство:

$$P\left(\bigcup_{i=1}^m A_i\right) \leq \sum_{i=1}^m P(A_i).$$

Доказательство. Введем систему событий $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$, \dots , $B_m = A_m \setminus \bigcup_{i=1}^{m-1} A_i$. Очевидно, что $B_j \cap B_k = \emptyset$, $j \neq k$ и $\bigcup_{i=1}^m A_i = \bigcup_{i=1}^m B_i$. Тогда имеют место следующие равенства:

$$P\left(\bigcup_{i=1}^m A_i\right) = P\left(\bigcup_{i=1}^m B_i\right) = \sum_{i=1}^m P(B_i).$$

Так как событие B_i является подмножеством события A_i для любого $i = 1, \dots, m$, то по третьему свойству вероятностной меры $\sum_{i=1}^m P(B_i) \leq \sum_{i=1}^m P(A_i)$, что позволяет утверждать, что $P(\bigcup_{i=1}^m A_i) \leq \sum_{i=1}^m P(A_i)$. ■

6. Для любых событий $A_1, \dots, A_m \in \mathcal{A}$ справедливо равенство:

$$P\left(\bigcup_{i=1}^m A_i\right) = \sum_{i=1}^m P(A_i) - \sum_{1 \leq i_1 < i_2 \leq m} P(A_{i_1} \cap A_{i_2}) + \\ + \sum_{1 \leq i_1 < i_2 < i_3 \leq m} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots + (-1)^{m-1} P(A_1 \cap \dots \cap A_m).$$

Это утверждение можно доказать, используя метод математической индукции с базой $m = 2$.

7. Пусть $\{A_i\}_{i=1}^{\infty}$ — последовательность событий такая, что $A_i \in \mathcal{A}$ для любого i , и объединение этих событий $\bigcup_{i=1}^{\infty} A_i$ принадлежит \mathcal{A} , тогда имеет место следующее неравенство:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i). \quad (3.1)$$

Доказательство. Введем систему событий $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus (A_1 \cup A_2)$, \dots , $B_m = A_m \setminus \bigcup_{i=1}^{m-1} A_i$, \dots . Очевидно, что события B_1, \dots, B_m, \dots непересекающиеся, и $\bigcup_{i=1}^{\infty} A_i = \bigcup_{l=1}^{\infty} B_l$. Справедливы следующие равенства:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{l=1}^{\infty} B_l\right) = \sum_{l=1}^{\infty} P(B_l).$$

Поскольку $B_l \subset A_l$, то $\sum_{l=1}^{\infty} P(B_l) \leq \sum_{l=1}^{\infty} P(A_l)$, из чего следует справедливость неравенства (3.1). ■

Замечание 3.1. Свойства 1–6 справедливы для любой конечно-аддитивной вероятностной меры. Свойство 7 предполагает счетную аддитивность вероятностной меры.

Теорема 3.1 (о непрерывности вероятностной меры). Пусть \mathcal{A} — алгебра подмножеств множества Ω , P — конечно-аддитивная вероятностная мера, заданная на алгебре \mathcal{A} . Следующие утверждения эквивалентны:

1. Вероятностная мера P — счетно-аддитивная.
2. Конечно-аддитивная вероятностная мера P непрерывна «сверху», т. е. для любой последовательности множеств $\{A_n\}_{n=1}^{\infty}$ такой, что для любого n множество $A_n \in \mathcal{A}$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ и $A_n \subset A_{n+1}$, выполняется:

$$P(A_n) \xrightarrow{n \rightarrow \infty} P\left(\bigcup_{m=1}^{\infty} A_m\right).$$

3. Конечно-аддитивная вероятностная мера P непрерывна «снизу», т. е. для любой последовательности событий $\{A_n\}_{n=1}^{\infty}$ такой, что для любого n событие $A_n \in \mathcal{A}$, $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$ и $A_n \supset A_{n+1}$, выполняется:

$$P(A_n) \xrightarrow{n \rightarrow \infty} P\left(\bigcap_{m=1}^{\infty} A_m\right).$$

4. Конечно-аддитивная вероятностная мера P «непрерывна в нуле», т. е. для любой последовательности событий $\{A_n\}_{n=1}^{\infty}$ такой, что для любого n событие $A_n \in \mathcal{A}$, $\bigcap_{i=1}^{\infty} A_i = \emptyset$ и $A_n \supset A_{n+1}$, выполняется:

$$P(A_n) \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Покажем, что из утверждения 1 следует утверждение 2, далее из 2 следует 3, из 3 следует 4, и, наконец, из 4 следует 1.

1. Покажем, что из 1 следует 2. Считаем, что вполне первое утверждение теоремы. Рассмотрим произвольную последовательность событий $\{A_n\}_{n=1}^{\infty}$ такую, что выполнены все требования п.2 теоремы. Определим систему событий $D_1 = A_1$, $D_2 = A_2 \setminus A_1$, $D_3 = A_3 \setminus A_2$, ..., $D_k = A_k \setminus A_{k-1}$, ... Очевидно, что $D_l \cap D_m = \emptyset$ для всех $l \neq m$, и имеет место равенство $\bigcup_{k=1}^{\infty} A_k = \bigcup_{l=1}^{\infty} D_l$. Вычислим вероятность события $\bigcup_{k=1}^{\infty} A_k$. Очевидны следующие равенства:

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{l=1}^{\infty} P(D_l) = \lim_{n \rightarrow \infty} P(A_n).$$

2. Покажем, что из 2 следует 3. Рассмотрим последовательность событий $\{A_n\}$, для которой выполнены все требования п. 3, тогда для последовательности $\{\bar{A}_n\}_{n=1}^{\infty}$ выполнено: $\bar{A}_n \subset \bar{A}_{n+1}$ и $\bigcup_{n=1}^{\infty} \bar{A}_n = \overline{\bigcap_{n=1}^{\infty} A_n}$ принадлежит алгебре \mathcal{A} . Согласно утверждению 2 теоремы имеет место сходимость: $P(\bar{A}_n) \xrightarrow{n \rightarrow \infty} P\left(\overline{\bigcap_{n=1}^{\infty} A_n}\right)$, что эквивалентно следующей сходимости: $1 - P(A_n) \xrightarrow{n \rightarrow \infty} 1 - P\left(\bigcap_{n=1}^{\infty} A_n\right)$, из чего следует утверждение 3 теоремы: $P(A_n) \xrightarrow{n \rightarrow \infty} P\left(\bigcap_{n=1}^{\infty} A_n\right)$.

3. Поскольку четвертое утверждение теоремы является частным случаем третьего утверждения, то очевидно, что из утверждения 3 следует утверждение 4.

4. Покажем, что из 4 следует 1. Рассмотрим произвольную последовательность $\{B_n\}_{n=1}^{\infty}$ такую, что для любого n событие $B_n \in \mathcal{A}$ и $B_k \cap B_m = \emptyset$ для всех номеров $k \neq m$, $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$. Очевидно, что $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^k B_n + \bigcup_{l=k+1}^{\infty} B_l$. Так как $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$ и $\bigcup_{n=1}^k B_n \in \mathcal{A}$, то и $\bigcup_{l=k+1}^{\infty} B_l \in \mathcal{A}$. Из того, что P — конечно-аддитивная мера, следует равенство:

$$P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^k P(B_n) + P\left(\bigcup_{l=k+1}^{\infty} B_l\right).$$

Покажем, что

$$\lim_{k \rightarrow \infty} P\left(\bigcup_{l=k+1}^{\infty} B_l\right) = 0,$$

т. е. $P\left(\bigcup_{l=k+1}^{\infty} B_l\right) \xrightarrow{k \rightarrow \infty} 0$. Пусть $\bigcup_{l=k+1}^{\infty} B_l = A_{k+1}$. Очевидно, что $A_n \supset A_{n+1}$. Если $\bigcap_{k=1}^{\infty} A_{k+1} = \emptyset$, то получаем, что из 4 следует утверждение 1 теоремы. Предположим, что $\bigcap_{k=1}^{\infty} A_{k+1} \neq \emptyset$, т. е. существует $\tilde{\omega} \in \bigcap_{k=1}^{\infty} A_{k+1}$. Следовательно, $\tilde{\omega} \in A_n$ для всех n , $\tilde{\omega} \in \bigcup_{l=n}^{\infty} B_l$ означает, что событие $\tilde{\omega}$ должно принадлежать хотя бы одному множеству из объединения, но так как эти множества попарно несовместны ($B_k \cap B_m = \emptyset$, $k \neq m$), то $\tilde{\omega}$ должно принадлежать какому-нибудь одному множеству: $\tilde{\omega} \in B_i$ и $\tilde{\omega} \notin B_j$, $j > i$, из чего следует, что $\tilde{\omega} \notin A_j$. Пришли к противоречию, так как событие $\tilde{\omega}$ должно принадлежать A_n для любого n . Значит, предположение о том, что $\bigcap_{k=1}^{\infty} A_{k+1} \neq \emptyset$ неверно. ■

Определение 3.1. Алгебра \mathcal{F} подмножеств множества Ω называется σ -алгеброй, если она замкнута относительно объединения счетной совокупности множеств, т. е. если $\{A_n\}_{n=1}^{\infty}$ — последовательность событий, где для любого n событие $A_n \in \mathcal{F}$, тогда $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Пусть на σ -алгебре \mathcal{F} задана счетно-аддитивная вероятностная мера P , тогда считается заданным *вероятностное пространство* (Ω, \mathcal{F}, P) . Сформулированные выше определения составляют систему аксиом Колмогорова.

Замечание 3.2. В дальнейшем, в качестве вероятностного пространства будем рассматривать систему (Ω, \mathcal{F}, P) , где Ω — множество элементарных событий, \mathcal{F} — σ -алгебра, P — счетно-аддитивная вероятностная мера.

§ 4. КЛАССИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ

Рассмотрим множество $\Omega = \{\omega_1, \dots, \omega_m\}$. Пусть все элементарные события равновероятны, т. е. $P(\omega_i) = P(\omega_j) = p$ для любых индексов i и j . По свойству вероятности справедливо равенство $P(\Omega) = 1$. В силу аддитивности вероятностной меры можно записать следующее равенство: $P(\Omega) = \sum_{i=1}^m P(\omega_i) = mp$. Следовательно, вероятность любого элементарного события ω_i может быть вычислена по формуле: $p = 1/m = 1/|\Omega|$.

В качестве алгебры \mathcal{A} рассмотрим множество всех подмножеств множества Ω , которое обозначим через 2^Ω . Рассмотрим некоторое событие $A \subset \Omega$, $A = \{\omega_{i_1}, \dots, \omega_{i_k}\}$. Вероятность наступления этого события в силу свойства аддитивности вероятностной меры может быть вычислена по формуле: $P(A) = \sum_{j=1}^k P(\omega_{i_j}) = k/|\Omega| = |A|/|\Omega|$.

Определение 4.1. Формулу

$$P(A) = \frac{|A|}{|\Omega|} \quad (4.1)$$

называют классическим определением вероятности.

Теорема 4.1 (правило умножения в комбинаторике). Пусть имеется r групп различных объектов. В первую группу входит n_1 объектов, во вторую — n_2 объектов, ..., в последнюю — n_r объектов. Будем составлять различные комбинации этих объектов следующим образом: последовательно из групп будем выбирать по одному объекту и располагать их в порядке появления (т.е. сначала элемент из первой группы, потом из второй, ..., из r -ой группы). Тогда всего возможно $n_1 n_2 \cdot \dots \cdot n_r$ различных комбинаций.

Доказательство. Обозначим группы объектов следующим образом:

$$X_1 = \{x_1^{(1)}, x_2^{(1)}, \dots, x_{n_1}^{(1)}\}, \quad |X_1| = n_1;$$

$$X_2 = \{x_1^{(2)}, x_2^{(2)}, \dots, x_{n_2}^{(2)}\}, \quad |X_2| = n_2;$$

...

$$X_r = \{x_1^{(r)}, x_2^{(r)}, \dots, x_{n_r}^{(r)}\}, \quad |X_r| = n_r.$$

Рассмотрим случай, когда $r = 2$. Изобразим всевозможные комбинации объектов этих двух групп в виде таблицы, в которых столбцам соответствуют объекты группы X_2 , а строкам — объекты группы X_1 (табл. 1.1).

	$x_1^{(2)}$	$x_2^{(2)}$...	$x_{n_2}^{(2)}$
$x_1^{(1)}$	$x_1^{(1)} x_1^{(2)}$	$x_1^{(1)} x_2^{(2)}$...	$x_1^{(1)} x_{n_2}^{(2)}$
$x_2^{(1)}$	$x_2^{(1)} x_1^{(2)}$	$x_2^{(1)} x_2^{(2)}$...	$x_2^{(1)} x_{n_2}^{(2)}$
\vdots	\vdots	\vdots		\vdots
$x_{n_1}^{(1)}$	$x_{n_1}^{(1)} x_1^{(2)}$	$x_{n_1}^{(1)} x_2^{(2)}$...	$x_{n_1}^{(1)} x_{n_2}^{(2)}$

Таблица 1.1

Всевозможные комбинации объектов

Таким образом, каждая клетка таблицы соответствует одной возможной комбинации объектов. Количество всевозможных комбинаций объектов совпадает с количеством клеток

таблицы и равняется $n_1 n_2$. Для случая $r = 2$ теорема верна.

Рассмотрим случай $r = 3$. Любую комбинацию объектов из групп X_1, X_2, X_3 можем записать в виде:

$$(x_{i_1}^{(1)}, x_{i_2}^{(2)}, x_{i_3}^{(3)}) = ((x_{i_1}^{(1)}, x_{i_2}^{(2)}), x_{i_3}^{(3)}),$$

применив доказанное утверждение для случая $r = 2$ два раза. Количество всевозможных комбинаций объектов $(n_1 n_2) n_3 = n_1 n_2 n_3$, что доказывает теорему для случая $r = 3$.

Для доказательства теоремы для случая $r = m$ можно применять предложенную схему рассуждения $m - 1$ раз. ■

Рассмотрим несколько вариантов определения вероятностных пространств, когда осуществляется случайный выбор объектов.

I. Упорядоченный выбор с возвращением

Пусть имеется урна с n шарами. Каждый шар имеет свой номер от 1 до n . Произведем из них последовательный случайный выбор r шаров, при этом возвращая каждый вынутый шар обратно. Тогда элементарным событием будет вектор (a_1, \dots, a_r) , $a_i = 1, \dots, n$, шары могут совпадать. Мощность пространства элементарных событий $\Omega = \{(a_1, \dots, a_r) \mid a_i = 1, \dots, n\}$ в данном случае равна n^r , это следует из теоремы 4.1. В качестве вероятностной меры (вероятности элементарного события) возьмем отношение $1/|\Omega|$ или $P(\omega) = 1/n^r$.

II. Упорядоченный выбор без возвращения

В этом случае вынутые на каждом шаге шары не возвращаются в урну. Пространство элементарных событий может быть определено следующим образом: $\Omega = \{(a_1, \dots, a_r) \mid a_i = 1, \dots, n, a_i \neq a_j, i \neq j\}$. Мощность этого множества равна $n(n-1)\dots(n-r+1) = A_n^r$. Это число называют числом размещений из n по r . Вероятность определяется следующим образом: $P(\omega) = 1/A_n^r$.

III. Неупорядоченный выбор без возвращения

Пространство элементарных событий может быть определено следующим образом: $\Omega = \{[a_1, \dots, a_r] \mid a_i = 1, \dots, n, a_i \neq a_j, i \neq j\}$. Квадратные скобки означают, что порядок появления шаров неизвестен. Мощность этого множества

называют числом сочетаний из n по r и обозначают через C_n^r . Число сочетаний и число размещений связаны равенством: $A_n^r = C_n^r \cdot r!$ Следовательно,

$$C_n^r = \frac{n(n-1) \cdot \dots \cdot (n-r+1)(n-r)!}{r!(n-r)!} = \frac{n!}{r!(n-r)!}.$$

В этом случае вероятность определяется равенством: $P(\omega) = 1/C_n^r$.

Определение 4.2. Любую числовую функцию, заданную на конечном множестве элементарных событий Ω , будем называть случайной величиной.

Рассмотрим случайную величину ξ , равную числу белых шаров в выборке без возвращения объема n из урны, содержащей M белых и $N - M$ черных шаров. Нетрудно заметить, что

$$P\{\xi = m\} = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n},$$

где $m = 0, 1, \dots, \min(n, M)$, $n \leq N - M$.

Говорят, что случайная величина ξ подчиняется гипергеометрическому распределению.

IV. Неупорядоченный выбор с возвращением

Пространство элементарных событий может быть определено следующим образом: $\Omega = \{[a_1, \dots, a_r] : a_i = 1, \dots, n\}$. Порядок появления шаров неизвестен. Так как $|\Omega| < \infty$, и все элементарные события равновероятны, $P(\omega) = 1/|\Omega|$. Найдем мощность множества Ω .

Пусть $r_1 \geq 0$ — число появлений в выборке шара a_1 , ..., $r_n \geq 0$ — число появлений в выборке шара a_n , при этом $\sum_{i=1}^n r_i = r$ и $r_i \in \{0, \dots, r\}$. Посчитаем количество наборов (r_1, \dots, r_n) , т.е. найдем количество решений уравнения $\sum_{i=1}^n r_i = r$. Прибавим к обеим частям уравнения по n , получим:

$$(r_1 + 1) + (r_2 + 1) + \dots + (r_n + 1) = r + n.$$

Число решений этого уравнения совпадает с числом разбиений отрезка $[0, r+n]$ на n частей. Необходимо выбрать $n-1$ точку

дробления, такие точки можно выбрать C_{r+n-1}^{n-1} способами. Так как $C_k^l = C_k^{k-l} = \frac{k!}{l!(k-l)!}$, получаем $C_{r+n-1}^{n-1} = C_{r+n-1}^r = |\Omega|$. Вероятность $P(\omega) = 1/C_{r+n-1}^{n-1}$.

§ 5. ГЕОМЕТРИЧЕСКИЕ ВЕРОЯТНОСТИ.

ЗАДАЧА О ВСТРЕЧЕ. ПАРАДОКС БЕРТРАНА.

ЗАДАЧА БЮФФОНА

Рассмотрим числовую прямую и отрезок $[a, b] = \Omega$. Пусть случайным образом выбирается точка из этого отрезка. Это означает, что любая точка отрезка $[a, b]$ может появиться в результате эксперимента с равными шансами.

Зададим вероятностную меру для любого множества $A \in \mathcal{F}_{[a,b]}$, где $\mathcal{F}_{[a,b]}$ — сигма-алгебра подмножеств отрезка $[a, b]$, измеримых по Лебегу. Вероятностная мера, называемая геометрической вероятностью, может быть определена следующим образом:

$$P(A) = \frac{\mu(A)}{\mu[a, b]} = \frac{\mu(A)}{b - a},$$

где через $\mu(A)$ обозначена мера Лебега множества A .

Замечание 5.1. В качестве σ -алгебры событий нельзя рассматривать множество всех подмножеств множества Ω , так как не все подмножества множества Ω измеримы по Лебегу.

В более общем случае геометрическая вероятность определяется аналогично. В качестве множества элементарных событий рассмотрим некоторую область измеримую по Лебегу $\Omega \subset \mathbb{R}^k$. Через $\mu(\Omega)$ обозначим меру Лебега в \mathbb{R}^k множества Ω . Используя принцип геометрической вероятности по отношению к мере Лебега μ , можно определить вероятностную меру для любого измеримого по Лебегу множества $A \subset \Omega \subset \mathbb{R}^k$ следующим образом:

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}. \quad (5.1)$$

Равенство (5.1) принято называть определением геометрической вероятности.

Пример 5.1. Задача о встрече. Два лица A и B условились встретиться в определенном месте между 12 и 13 часами дня. Пришедший первым ждет другого в течение 20 минут, после чего уходит. Требуется вычислить вероятность встречи лиц A и B , если приход каждого из них в течение указанного часа происходит наудачу случайным образом, и моменты прихода независимы.

Обозначим моменты прихода лица A через x и лица B через y . Для того чтобы встреча произошла, необходимо и достаточно, чтобы

$$|x - y| \leq 20.$$

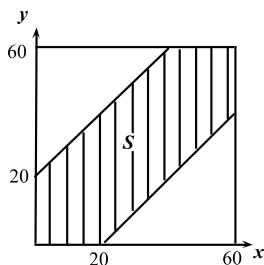


Рис. 1.2
Задача о встрече

Изобразим x и y на декартовой плоскости (рис. 1.2), в качестве единицы масштаба выберем минуту. Возможные исходы могут быть изображены квадратом со сторонами 60, благоприятствующие встрече исходы — в области S . Вероятность встречи лиц A и B равна отношению площади фигуры S к площади всего квадрата:

$$p = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

Пример 5.2. Парадокс Бертрана. Рассматривается окружность радиуса R , в которую вписан равносторонний треугольник. Случайным образом в данной окружности выбирается хорда. Какова вероятность того, что длина хорды окажется

больше стороны равностороннего треугольника, вписанного в эту окружность (рис. 1.3)?

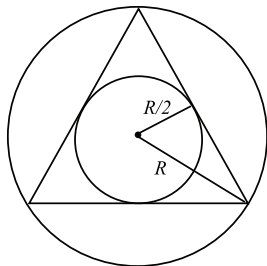


Рис. 1.3
Парадокс Бертрانا

Данную задачу можно решить тремя способами.

Решение 1. Из геометрических соображений: для того, чтобы длина хорды окружности радиуса R была больше стороны равностороннего треугольника, вписанного в эту окружность, необходимо, чтобы центр хорды оказался внутри окружности, вписанной в равносторонний треугольник. Последняя окружность будет иметь радиус $R/2$ (рис. 1.4).

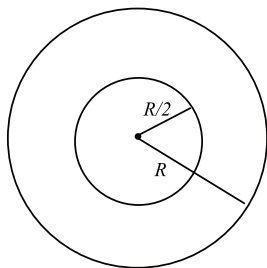


Рис. 1.4
Парадокс Бертрана. Решение 1

Таким образом, искомая вероятность определяется как отношение площадей окружностей радиуса $R/2$ и радиуса R

соответственно:

$$\frac{S_{\frac{R}{2}}}{S_R} = \frac{\pi \frac{R^2}{4}}{\pi R^2} = \frac{1}{4}.$$

Решение 2. Из соображений симметрии: пусть один конец хорды закреплен в одной из вершин равностороннего треугольника.

Вершины треугольника разбивают дугу окружности на три равные части. Длина хорды будет больше стороны равностороннего треугольника, если другой конец хорды будет лежать на дуге между двумя другими вершинами треугольника. Тогда искомая вероятность определяется как отношение длин соответствующих дуг и равна $1/3$ (рис. 1.5).

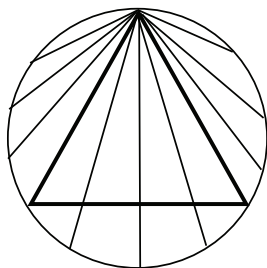


Рис. 1.5

Парадокс Бертрана. Решение 2

Решение 3. В окружности радиуса R произвольным образом проведем диаметр. Отметим точку, в которой рассматриваемая хорда пересекает диаметр. Рассматриваем хорды, перпендикулярные диаметру (рис. 1.6).

Следовательно, по расстоянию ρ от центра окружности до точки пересечения хорды и диаметра можем судить о длине хорды: длина хорды будет больше стороны равностороннего треугольника, если расстояние $\rho < R/2$. Искомая вероятность определяется как отношение соответствующих расстояний:

$$\frac{\mu(0, \frac{R}{2})}{R} = \frac{\frac{R}{2}}{R} = \frac{1}{2}.$$

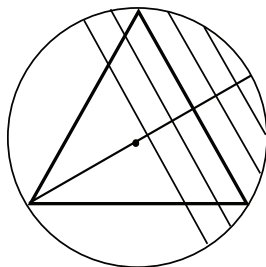


Рис. 1.6

Парадокс Бертрانا. Решение 3

Причина возникновения парадокса связана с неоднозначностью математической формализации поставленной задачи.

Пример 5.3. Задача Бюффона. На плоскость, разлинеенную параллельными прямыми, отстоящими друг от друга на расстоянии $2a$, наудачу бросается игла длиной $2l < 2a$. Требуется найти вероятность того, что игла пересечет одну из параллельных прямых (рис. 1.7).

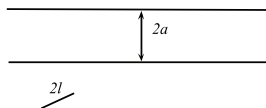


Рис. 1.7

Задача Бюффона

Решение. Пусть y — расстояние от середины иглы до ближайшей из параллельных прямых, x — острый угол между иглой и перпендикуляром, проведенным к параллельным прямым. Координаты (x, y) определяют положение иглы относительно параллельных прямых на плоскости, $0 \leq x \leq \pi/2$, $0 \leq y \leq a$. Предполагается, что точка (x, y) распределена равномерно в соответствующем прямоугольнике.

Событие, заключающееся в том, что игла пересечет какую-либо параллельную прямую, можно записать следующим

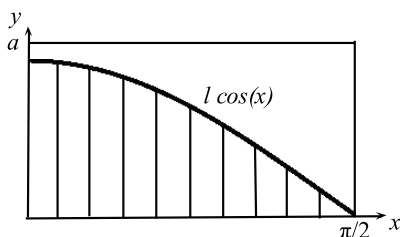


Рис. 1.8
Задача Бюффона

образом:

$$A = \{(x, y) : 0 \leq y \leq l \cos x\}.$$

Тогда искомая вероятность определяется как отношение площадей (рис. 1.8), соответствующих благоприятствующим и всем возможным исходам, и равна

$$\frac{\mu(A)}{\mu(\Omega)} = \frac{\int_0^{\pi/2} l \cos x dx}{a \cdot \frac{\pi}{2}} = \frac{2l}{a\pi}.$$

§ 6. УСЛОВНЫЕ ВЕРОЯТНОСТИ

Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) . Выберем некоторое событие $B \in \mathcal{F}$ такое, что $P(B) > 0$. Определим вероятность события $A \in \mathcal{F}$ при условии, что произошло событие B .

Проведем эксперимент N раз. Обозначим через $N(B)$ количество экспериментов, в которых произошло событие B , $N(A \cap B)$ — количество экспериментов, в которых произошли оба события A и B . Относительная частота события A , если известно, что произошло событие B , будет определяться отношением $N(A \cap B)/N(B)$. Поделим числитель и знаменатель этого отношения на N , получаем:

$$\frac{N(A \cap B)/N}{N(B)/N} \approx \frac{P(A \cap B)}{P(B)}.$$

Приближенное равенство следует из свойства «статистической устойчивости» относительных частот. Пусть множество Ω конечно, и вероятность любого элементарного события $\omega \in \Omega$ равна $p(\omega) = p = \text{const.}$ Пусть произошло событие B , что означает, что событие B стало достоверным. Воспользуемся классическим определением вероятности. Посчитаем количество элементарных событий в множествах $A \cap B$ и B . Получаем отношение:

$$\frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{P(A \cap B)}{P(B)}.$$

Полученные равенства приводят к определению.

Определение 6.1. Если вероятность события B отлична от нуля, то условная вероятность события A при условии, что событие B произошло, определяется формулой

$$P(A/B) = \frac{P(A \cap B)}{P(B)}.$$

Выберем событие $B \in \mathcal{F}$, $P(B) > 0$. Рассмотрим вероятность $P(A/B)$, где $A \in \mathcal{F}$. Определим функцию $P(\cdot/B): \mathcal{F} \rightarrow [0, 1]$. Функция $P(\cdot/B)$ обладает следующими свойствами, т. е. она является вероятностной мерой на (Ω, \mathcal{F}) :

1. Если A совпадает с множеством Ω , то $P(\Omega/B) = 1$.
2. Для любого события $A \in \mathcal{F}$ вероятность $P(A/B)$ неотрицательна.
3. Для любых несовместных событий A_1 и A_2 вероятность $P(A_1 \cup A_2/B)$ равна сумме вероятностей $P(A_1/B) + P(A_2/B)$.

Наличие первых трех свойств означает, что условная вероятность — конечно-аддитивная вероятностная мера. Если P — счетно-аддитивная вероятность, то и условная вероятность $P(\cdot/B)$ обладает свойством счетной аддитивности.

4. Для последовательности попарно несовместных событий $\{A_n\}_{n=1}^{\infty}$ справедливо равенство:

$$P\left(\bigcup_{i=1}^{\infty} A_i/B\right) = \sum_{i=1}^{\infty} P(A_i/B).$$

Из определения условной вероятности следует формула умножения вероятностей:

$$P(A \cap B) = P(B)P(A/B).$$

Обобщим формулу умножения вероятностей на случай n событий. Для событий $A_1, \dots, A_n \in \mathcal{F}$, для которых $P(A_1 \cap \dots \cap A_{n-1}) > 0$, имеет место равенство:

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

§ 7. НЕЗАВИСИМОСТЬ СОБЫТИЙ

Выберем некоторое событие $B \in \mathcal{F}$, вероятность которого положительна. Если событие $A \in \mathcal{F}$ не зависит от появления события B , то $P(A/B) = P(A)$. Используя определение условной вероятности, получаем равенство: $P(A \cap B)/P(B) = P(A)$. Следовательно, вероятность совместного появления событий $P(A \cap B)$ равна произведению вероятностей $P(A)P(B)$. Пусть $P(A) > 0$, тогда $P(B/A) = P(B \cap A)/P(A) = P(B)$. Следовательно, если событие A не зависит от события B , то и событие B не зависит от события A . Сформулируем определение независимости событий.

Определение 7.1. События A и B независимы, если справедливо следующее равенство:

$$P(A \cap B) = P(A)P(B). \quad (7.1)$$

Замечание 7.1. Определение 7.1 не нуждается в предположении о положительности $P(A)$ и $P(B)$.

Замечание 7.2. Если предположить, что события A и B несовместны, и $P(A) > 0$, $P(B) > 0$, тогда A и B — зависимые события.

Определение 7.2. События A_1, \dots, A_n называются независимыми в совокупности, если для любых m событий A_{i_1}, \dots, A_{i_m} ,

$m = 2, \dots, n$, выполнено:

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) = \prod_{k=1}^m P(A_{i_k}). \quad (7.2)$$

Если события независимы в совокупности, то они независимы попарно. Обратное не верно: из попарной независимости не следует независимость в совокупности.

Пример 7.1. Рассмотрим пространство элементарных событий $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Пусть $p(\omega_i) = 1/4$, $i = 1, \dots, 4$. Определим следующие события: $A = \{\omega_1, \omega_2\}$, $B = \{\omega_1, \omega_3\}$, $C = \{\omega_1, \omega_4\}$. Покажем, что события A , B , C попарно независимы, но в совокупности зависимы. Очевидно, что справедливы равенства:

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$

Покажем попарную независимость на примере событий A и B : $P(A \cap B) = p(\omega_1) = 1/4 = P(A)P(B)$. Аналогично попарную независимость можно показать и для пар A и C , B и C . Посчитаем вероятность $P(A \cap B \cap C) = p(\omega_1) = 1/4 \neq P(A)P(B)P(C) = 1/8$, что означает зависимость в совокупности.

Определение 7.3. Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) , пусть алгебры $\mathcal{A}_1, \dots, \mathcal{A}_n$ являются подалгебрами σ -алгебры \mathcal{F} . Говорят, что алгебры $\mathcal{A}_1, \dots, \mathcal{A}_n$ независимы или независимы в совокупности, если для любых событий $A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$ имеет место равенство:

$$P(A_1 \cap \dots \cap A_n) = \prod_{i=1}^n P(A_i). \quad (7.3)$$

Замечание 7.3. Определение 7.3 относится и к σ -алгебрам, так как σ -алгебра — частный случай алгебры.

§ 8. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ. ФОРМУЛЫ БАЙЕСА

Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) . Пусть задано разбиение множества Ω на события B_1, \dots, B_m ; $B_1 \in \mathcal{F}, \dots, B_m \in \mathcal{F}$ такие, что $P(B_i) > 0$ для всех i , $B_i \cap B_j = \emptyset$ для любых $i \neq j$, $\Omega = \bigcup_{i=1}^m B_i$.

Рассмотрим некоторое событие $A \in \mathcal{F}$. Представим его в виде: $A = \bigcup_{i=1}^m (A \cap B_i)$. По свойству аддитивности получаем:

$$P(A) = \sum_{i=1}^m P(A \cap B_i) = \sum_{i=1}^m P(B_i)P(A/B_i).$$

Формула

$$P(A) = \sum_{i=1}^m P(B_i)P(A/B_i) \quad (8.1)$$

называется *формулой полной вероятности*.

Дополнительно предположим, что $P(A) > 0$. Найдем вероятность события B_j при условии, что произошло событие A . Имеет место равенство:

$$P(B_j/A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(B_j)P(A/B_j)}{\sum_{i=1}^m P(B_i)P(A/B_i)}.$$

Формулы

$$P(B_j/A) = \frac{P(B_j)P(A/B_j)}{\sum_{i=1}^m P(B_i)P(A/B_i)}, \quad j = 1, 2, \dots, m \quad (8.2)$$

называются *формулами Байеса*. Вероятности $P(B_i)$, $i = 1, \dots, m$ называются априорными вероятностями. По формуле (8.2) можно вычислить вероятность наступления события B_i при условии, что произошло событие A . Вероятности $P(B_i/A)$, $i = 1, \dots, m$ называются апостериорными вероятностями.

СХЕМА БЕРНУЛЛИ

§ 1. СХЕМА БЕРНУЛЛИ. БИНОМИАЛЬНОЕ И ПОЛИНОМИАЛЬНОЕ РАСПРЕДЕЛЕНИЯ

Пусть проводится n независимых испытаний, причем, в каждом испытании возможны два исхода: «1» — успех и «0» — неудача. Обозначим через p вероятность успеха, через $q = 1 - p$ — вероятность неудачи.

Для определения вероятностного пространства рассмотрим элементарное событие $\omega = (a_1, \dots, a_n)$, где $a_i = 1$ или 0. Тогда $|\Omega| = 2^n$, в качестве σ -алгебры возьмем $\mathcal{F} = 2^\Omega$. Вероятность элементарного события ω вычислим по формуле

$$p(\omega) = p^{\sum_{i=1}^n a_i} q^{n - \sum_{i=1}^n a_i},$$

где $\sum_{i=1}^n a_i$ — количество единиц в векторе (a_1, \dots, a_n) .

Напомним, что если множество элементарных событий конечно, то случайной величиной называется любая числовая функция $\xi : \Omega \rightarrow \mathbb{R}$.

Пусть $\mu(a_1, \dots, a_n) = \mu(\omega) = \sum_{i=1}^n a_i$ — количество успехов в серии из n независимых испытаний. Величина $\mu(\omega)$ является случайной величиной. Нас интересует вероятность следующего события: $\mu = t$, т.е. вероятность события, что в серии из n независимых испытаний произойдет ровно t успехов, $P\{\mu = t\}$.

Случайная величина μ принимает в данном случае значение из конечной совокупности $\{0, 1, \dots, n\}$. Вероятность

события $\mu = m$ можно вычислить по формуле:

$$\begin{aligned} P\{\mu = m\} &= \sum_{\omega: \mu(\omega)=m} P(\omega) = \sum_{\omega: \sum_{i=1}^n a_i = m} p^{\sum_{i=1}^n a_i} q^{n - \sum_{i=1}^n a_i} = \\ &= p^m q^{n-m} \sum_{\omega: \sum_{i=1}^n a_i = m} 1 = C_n^m p^m q^{n-m}. \quad (1.1) \end{aligned}$$

Найдем сумму вероятностей при m , пробегающем значения от 0 до n , получаем:

$$\sum_{m=0}^n P\{\mu = m\} = \sum_{m=0}^n C_n^m p^m q^{n-m} = (p + q)^n = 1. \quad (1.2)$$

Дискретное распределение — набор (конечный или бесконечный) вероятностей несовместных событий, которые в сумме дают единицу.

Определение 1.1. Распределение случайной величины μ , равной количеству успехов в серии из n независимых испытаний, называется биномиальным распределением.

Рассмотрим более общую схему. Пусть производится n независимых испытаний, каждое из которых может закончиться одним из r исходов из множества $\{1, \dots, r\}$. Исходу i соответствует вероятность p_i , $i = 1, \dots, r$. При этом, $\sum_{i=1}^r p_i = 1$. Пусть набор (a_1, \dots, a_n) — упорядоченный набор чисел из множества $\{1, \dots, r\}$. Определим пространство элементарных событий Ω и σ -алгебру \mathcal{F} подмножеств множества Ω следующим образом:

$$\Omega = \{(a_1, \dots, a_n) : a_i = 1, \dots, r\},$$

$$\mathcal{F} = \{A : A \subset \Omega\} = 2^\Omega.$$

Вероятность события (a_1, \dots, a_n) определим так, чтобы последовательные испытания оказались независимыми:

$$P\{(a_1, \dots, a_n)\} = p_1^{m_1} \cdot \dots \cdot p_r^{m_r},$$

где m_1 — количество исходов «1» в (a_1, \dots, a_n) , ..., m_r — количество исходов « r » в (a_1, \dots, a_n) . Причем, $m_1 + \dots + m_r = n$. Найдем вероятность следующего события: в n независимых испытаниях m_1 раз выпал исход «1», ..., m_r раз выпал исход « r ». Обозначим вероятность этого события через $P_n(m_1, \dots, m_r)$, как легко видеть:

$$P_n(m_1, \dots, m_r) = D_{m_1, \dots, m_r} p_1^{m_1} \cdot \dots \cdot p_r^{m_r},$$

где

$$\begin{aligned} D_{m_1, \dots, m_r} &= C_n^{m_1} C_{n-m_1}^{m_2} \dots C_{n-m_1-\dots-m_{r-1}}^{m_r} = \\ &= \frac{n!}{m_1!(n-m_1)!} \frac{(n-m_1)!}{m_2!(n-m_1-m_2)!} \cdot \dots \cdot 1 = \frac{n!}{m_1!m_2! \dots m_r!}. \end{aligned}$$

Получаем выражение для вероятности P_n :

$$P_n(m_1, \dots, m_r) = \frac{n!}{m_1!m_2! \dots m_r!} p_1^{m_1} \dots p_r^{m_r}. \quad (1.3)$$

Определение 1.2. Распределение, определяемое формулой (1.3), называется полиномиальным распределением.

§ 2. ТЕОРЕМА ПУАССОНА

Рассмотрим схему Бернулли с n независимыми испытаниями, в каждом из которых возможны два исхода: «1» — успех и «0» — неудача. Вероятность успеха обозначим через p , а вероятность неудачи через $q = 1 - p$. Эти вероятности не зависят от номера испытания. Введем случайную величину μ , равную количеству успехов в n испытаниях.

Теорема 2.1. (Теорема Пуассона) Пусть в схеме Бернулли число испытаний $n \rightarrow \infty$ и при этом $np \xrightarrow{n \rightarrow \infty} \lambda > 0$. Тогда для любого $m = 0, 1, 2, \dots$ выполнено:

$$P\{\mu = m\} = C_n^m p^m q^{n-m} \xrightarrow{n \rightarrow \infty} \frac{\lambda^m}{m!} e^{-\lambda}. \quad (2.1)$$

Доказательство. Справедливы следующие равенства:

$$\begin{aligned} C_n^m p^m q^{n-m} &= \frac{n!}{m!(n-m)!} p^m \frac{(1-p)^n}{(1-p)^m} = \\ &= \frac{(np)^m}{m!} e^{n \ln(1-p)} \frac{1(1-\frac{1}{n}) \dots (1-\frac{m-1}{n})}{(1-p)^m} \xrightarrow{n \rightarrow \infty} \frac{\lambda^m}{m!} e^{-\lambda}, \end{aligned}$$

Утверждение справедливо, поскольку $(np)^m/m! \xrightarrow{n \rightarrow \infty} \lambda^m/m!$ и $\frac{1 \cdot (1-\frac{1}{n}) \dots (1-\frac{m-1}{n})}{(1-p)^m} \xrightarrow{n \rightarrow \infty} 1$, а также в силу того, что $\ln(1-p) = -p + o(p)$ и $e^{n \ln(1-p)} = e^{-np + np \frac{o(p)}{p}} \xrightarrow{n \rightarrow \infty} e^{-\lambda}$. ■

Замечание 2.1. Можно доказать более сильное утверждение:

$$\left| \sum_{m \in B} C_n^m p^m q^{n-m} - \sum_{m \in B} \frac{(np)^m}{m!} e^{-np} \right| \leq np^2,$$

где B — любое множество на положительной части числовой прямой.

Если вероятность p близка к 0, то есть $p \ll 1$, а $n \gg 1$, то теорема дает «хорошую» аппроксимацию для соответствующей вероятности. Эта аппроксимация называется *аппроксимацией Пуассона*.

Замечание 2.2. Используя формулу разложения экспоненты в ряд Тейлора, получаем равенство: $\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} e^{\lambda} = 1$. Таким образом, правая часть (2.1) определяет некоторое распределение, которое носит название *распределение Пуассона*. Это распределение зависит от положительного параметра λ .

§ 3. ЛОКАЛЬНАЯ И ИНТЕГРАЛЬНАЯ ТЕОРЕМЫ МУАВРА–ЛАПЛАСА

Рассмотрим схему Бернулли с n независимыми испытаниями, в каждом из которых возможны два исхода: «1» — успех и «0» — неудача. Вероятность успеха обозначим через p , а вероятность неудачи через $q = 1 - p$, причем, вероятности не зависят от номера испытания. Обозначим через μ случайную величину, равную количеству успехов в n испытаниях.

Теорема 3.1 (локальная предельная теорема Муавра–Лапласа). Пусть $\sigma = \sqrt{npq} \xrightarrow{n \rightarrow \infty} \infty$, тогда для любой константы $c > 0$ равномерно по $x = (m - np)/\sigma$ таким, что $|x| \leq c$, выполняется

$$P\{\mu = m\} = C_n^m p^m q^{n-m} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2}} (1 + o(1)), \quad (3.1)$$

где $o(1) \xrightarrow{n \rightarrow \infty} 0$.

Замечание 3.1. Утверждение теоремы можно записать в следующем виде:

$$\frac{P\{\mu = m\}}{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2}}} \xrightarrow{n \rightarrow \infty} 1.$$

Доказательство. Учитывая, что $x = (m - np)/\sigma$, $|x| \leq c$ и $\sigma = \sqrt{npq}$, получаем:

$$m = np + \sigma x = np(1 + \frac{xq}{\sigma}).$$

Пусть $k = n - m$. Очевидно, что $k = nq - \sigma x = nq(1 - xp/\sigma)$. При $n \rightarrow \infty$, $\sigma \rightarrow \infty$, получаем, что $m \rightarrow \infty$ и $k \rightarrow \infty$. Воспользуемся формулой Стирлинга:

$$n! = \sqrt{2\pi n} n^n e^{-n + \Theta_n},$$

где $\Theta_n = O(1/n)$ при $n \rightarrow \infty$. Вычислим логарифм вероятности $P\{\mu = m\}$, получаем следующее равенство:

$$\ln C_n^m p^m q^k = \ln n! - \ln m! - \ln k! + m \ln p + k \ln q. \quad (3.2)$$

Найдем логарифмы факториалов из правой части равенства (3.2), используя формулу Стирлинга:

$$\ln n! = \ln \sqrt{2\pi} + \frac{\ln n}{2} + n \ln n - n + \Theta_n. \quad (3.3)$$

Аналогичным образом получаем равенства:

$$\ln k! = \ln \sqrt{2\pi} + \frac{\ln k}{2} + k \ln k - k + \Theta_k, \quad (3.4)$$

$$\ln m! = \ln \sqrt{2\pi} + \frac{\ln m}{2} + m \ln m - m + \Theta_m. \quad (3.5)$$

Вычислим разность $\ln n! - \ln m! - \ln k!$, используя равенства (3.3), (3.4) и (3.5), получаем:

$$\begin{aligned} \ln n! - \ln k! - \ln m! &= \\ &= \ln \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \ln \frac{mk}{n} - m \ln \frac{m}{n} - k \ln \frac{k}{n} + O\left(\frac{1}{\sigma^2}\right). \end{aligned}$$

Подставив это выражение в (3.2), получаем:

$$\begin{aligned} \ln C_n^m p^m q^k &= \ln \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \ln \left(\sigma^2 \left(1 + \frac{xq}{\sigma}\right) \left(1 - \frac{xp}{\sigma}\right) \right) - \\ &\quad - m \ln \frac{m}{np} - k \ln \frac{k}{nq} + O\left(\frac{1}{\sigma^2}\right) = \\ &= \ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2} \ln \left(1 + \frac{xq}{\sigma}\right) - \frac{1}{2} \ln \left(1 - \frac{xp}{\sigma}\right) - \\ &\quad - \left[(np + \sigma x) \ln \left(1 + \frac{xq}{\sigma}\right) + (nq - \sigma x) \ln \left(1 - \frac{xp}{\sigma}\right) \right] + O\left(\frac{1}{\sigma^2}\right). \end{aligned}$$

Преобразуя выражение в квадратных скобках, получаем равенство:

$$\ln C_n^m p^m q^k = \ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{x^2}{2} + O\left(\frac{1}{\sigma}\right),$$

откуда следует утверждение теоремы:

$$C_n^m p^m q^k = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2}} e^{O(1/\sigma)},$$

где $e^{O(1/\sigma)} = 1 + O(1/\sigma)$. ■

Замечание 3.2. Рассмотрим функцию $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Вычислим интеграл $\int_{-\infty}^{\infty} \varphi(x) dx$. Для этого вычислим

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right)^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy.$$

Используя теорему о повторном интегрировании, получаем $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx dy$. Переходя к полярным координатам, получаем интеграл $\int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} r dr d\psi = 2\pi$. Следовательно, $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$, откуда следует, что $\int_{-\infty}^{\infty} \varphi(x) dx = 1$.

Функция $\varphi(x)$ обладает следующими свойствами:

1. $\varphi(x) \geq 0$ для всех $x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} \varphi(x) dx = 1$.

Любая функция, удовлетворяющая этим двум свойствам, называется *плотностью распределения*.

Определение 3.1. Функция $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ называется плотностью стандартного нормального распределения (распределения Гаусса).

Теорема 3.2 (интегральная предельная теорема Муавра–Лапласа). Пусть $\sigma = \sqrt{npq} \xrightarrow[n \rightarrow \infty]{} \infty$, тогда при $n \rightarrow \infty$ равномерно по парам (a, b) имеет место сходимость:

$$P \left\{ a \leq \frac{\mu - np}{\sqrt{npq}} \leq b \right\} \xrightarrow[n \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx = \int_a^b \varphi(x) dx. \quad (3.6)$$

Доказательство. Сначала докажем теорему для некоторого фиксированного $c > 0$, для которого $|a| \leq c$, $|b| \leq c$. Потом распространим результат на всю числовую ось.

Запишем двойное неравенство $a \leq \frac{\mu - np}{\sqrt{npq}} \leq b$ в виде:

$$\sigma a + np \leq \mu \leq \sigma b + np.$$

Выберем $m_1 \geq \sigma a + np$, $m_1 \in \mathbb{Z}^+$ и $m_2 \leq \sigma b + np$, и $m_2 \in \mathbb{Z}^+$. Число m_1 наиболее близко к $(\sigma a + np)$ справа, m_2 наиболее близко к $(\sigma b + np)$ слева. Справедливы следующие равенства:

$$\begin{aligned}
 P\left\{a \leq \frac{\mu - np}{\sqrt{npq}} \leq b\right\} &= \sum_{m=m_1}^{m_2} P\{\mu = m\} = \\
 &= \sum_{m=m_1}^{m_2} C_n^m p^m q^{n-m} = \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_m^2}{2}} \frac{1}{\sigma} \left(1 + O\left(\frac{1}{\sigma}\right)\right),
 \end{aligned}$$

где $x_m = (m - np)/\sigma$. Причем, $a \leq x_m \leq b$ для любого $m = m_1, \dots, m_2$, $1/\sigma = x_{m+1} - x_m = \Delta x_m$.

Заметим, что полученная сумма — интегральная сумма для интеграла $\int_a^b (1/\sqrt{2\pi}) e^{-\frac{x^2}{2}} dx$, и она должна сходиться к нему, причем сходимость обязана быть равномерной для всех a, b : $|a| \leq c$, $|b| \leq c$.

Теперь распространим теорему на всю числовую ось. Для любого $c > 0$ справедливо равенство:

$$\int_{|x| > c} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 - \int_{|x| \leq c} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \quad (3.7)$$

Пусть $\xi_n = (\mu - np)/\sigma$. Очевидно, что для любого c справедливо равенство:

$$P\{|\xi_n| > c\} = 1 - P\{|\xi_n| \leq c\}. \quad (3.8)$$

Вычтем (3.8) из (3.7) и рассмотрим модуль этой разности:

$$\begin{aligned}
 \left| \int_{|x| > c} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - P\{|\xi_n| > c\} \right| &= \\
 &= \left| - \int_{|x| \leq c} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + P\{|\xi_n| \leq c\} \right|. \quad (3.9)
 \end{aligned}$$

Поскольку $\int_{|x| > c} (1/\sqrt{2\pi}) e^{-\frac{x^2}{2}} dx$ убывает при $c \rightarrow \infty$, то можно гарантировать, что существует c такое, что этот интеграл будет не больше $\varepsilon/4$.

Рассмотрим отрезок $[-c, c]$. По доказанному существует номер n_1 такой, что для любых $n \geq n_1$, правая часть

(3.9) будет не больше $\varepsilon/4$. Тогда справедливо неравенство:
 $P\{|\xi_n| > c\} \leq \varepsilon/2$.

Для любых $a, b \in \mathbb{R}$ обозначим $[A, B] = [a, b] \cap [-c, c]$.
 Несложно заметить, что

$$\begin{aligned} \left| \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - P\{a \leq \xi_n \leq b\} \right| &\leq \\ &\leq \int_{|x|>c} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + P\{|\xi_n| > c\} + \\ &+ \left| \int_A^B \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - P\{A \leq \xi_n \leq B\} \right| \leq \varepsilon, \end{aligned}$$

как бы ни были выбраны точки a и b , всегда существует номер n_1 такой, что для любого номера $n \geq n_1$ модуль рассматриваемой разности будет не больше ε , то есть, имеет место равномерная сходимость для всех $a \leq b$. ■

§ 4. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Теорема 4.1 (закон больших чисел для схемы Бернулли).
 Пусть число испытаний в схеме Бернулли $n \rightarrow \infty$, тогда для любого $\varepsilon > 0$ имеет место следующая сходимость

$$P\left\{\left|\frac{\mu}{n} - p\right| > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Очевидно, что $P\{|\mu/n - p| > \varepsilon\} = 1 - P\{|\mu/n - p| \leq \varepsilon\}$. Проведем преобразования:

$$\begin{aligned} P\left\{\left|\frac{\mu}{n} - p\right| \leq \varepsilon\right\} &= P\left\{-\varepsilon \leq \frac{\mu}{n} - p \leq \varepsilon\right\} = \\ &= P\left\{-\varepsilon \sqrt{\frac{n}{pq}} \leq \frac{\mu - np}{\sqrt{npq}} \leq \varepsilon \sqrt{\frac{n}{pq}}\right\}. \end{aligned}$$

Получаем равенство:

$$P\left\{\left|\frac{\mu}{n} - p\right| > \varepsilon\right\} = \left(1 - \int_{-\varepsilon\sqrt{\frac{n}{pq}}}^{\varepsilon\sqrt{\frac{n}{pq}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx\right) + \\ + \left(\int_{-\varepsilon\sqrt{\frac{n}{pq}}}^{\varepsilon\sqrt{\frac{n}{pq}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - P\left\{-\varepsilon\sqrt{\frac{n}{pq}} \leq \frac{\mu - np}{\sqrt{npq}} \leq \varepsilon\sqrt{\frac{n}{pq}}\right\}\right).$$

Выражение в первых скобках правой части стремится к нулю, выражение во вторых скобках по теореме 3.2 тоже стремится к нулю, из чего следует утверждение теоремы. ■

§ 5. МЕТОД МОНТЕ-КАРЛО. ВЫЧИСЛЕНИЕ ИНТЕГРАЛОВ

Пусть требуется вычислить интеграл $\int_a^b g(x)dx$. При этом функция $g(x)$ ограничена на отрезке $[a, b]$, т. е. $g(x) \in [c, d]$. Преобразуем функцию $g(x)$ следующим образом:

$$\varphi(x) = \frac{g(x) - c}{d - c},$$

тогда функцию $g(x)$ можно представить в виде:

$$g(x) = c + (d - c)\varphi(x).$$

При этом $\varphi(x) \in [0, 1]$ для всех точек $x \in [a, b]$.

Запишем интеграл $\int_a^b g(x)dx$ через интеграл от функции $\varphi(x)$:

$$\int_a^b g(x)dx = c(b - a) + (d - c) \int_a^b \varphi(x)dx.$$

Проведем замену переменных: $u = (x - a)/(b - a)$, получим:

$$\int_a^b g(x)dx = c(b - a) + (d - c)(b - a) \int_0^1 h(u)du,$$

где $h(u) = \varphi(a + (b-a)u)$. Таким образом, для вычисления интеграла $\int_a^b g(x)dx$ достаточно вычислить интеграл $\int_0^1 h(u)du$. Область, ограниченную функцией $h(u)$, обозначим через A (рис. 2.1).

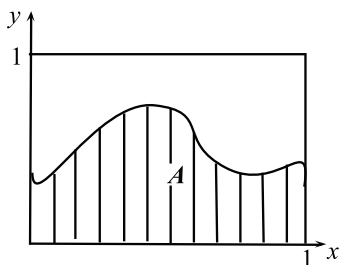


Рис. 2.1

*Вычисление интеграла с помощью
метода Монте-Карло*

Проведем следующий эксперимент. Будем многократно случайным образом выбирать точку из квадрата со стороной 1. Если точка принадлежит области A , то будем считать, что произошел успех. В противном случае будем считать, что произошла неудача. Вероятность успеха $p = \int_0^1 h(u)du$.

Пусть n_A — число успехов в серии из n независимых испытаний, тогда можно считать, что приближенно выполняется равенство:

$$\frac{n_A}{n} \cong \int_0^1 h(u)du.$$

Метод Монте-Карло особенно эффективен при вычислении многократных интегралов. В этом случае он имеет преимущества над численными методами вычисления интегралов (метод трапеций, метод Симпсона). В случае вычисления k -кратного интеграла методом Монте-Карло требуется случайным образом выбирать $k + 1$ координату и проверять, не принадлежит ли полученная точка рассматриваемой области A . Также к

преимуществам метода можно отнести его безразличие к виду подынтегральной функции.

Для нахождения случайных точек можно пользоваться таблицами случайных чисел. В этих таблицах даны цифры от 0 до 9. Эти данные можно рассматривать как реализации взаимно независимых и одинаково распределенных случайных величин, принимающих значения 0, 1, 2, ..., 9 с одной и той же вероятностью, равной 0,1. Табулированные цифры сгруппированы по десять. Каждые десять цифр представляют собой реализацию случайной величины, которая может принимать значения от 0000000000 до 9999999999 с одинаковыми вероятностями, равными 0,0000000001. Если каждую группу из k цифр, рассматриваемую как целое число, умножить на 10^{-k} , то получим реализации случайных величин ξ , принимающих k -разрядные значения от 0 до $(1 - 10^{-k})$ с одинаковыми вероятностями, равными 10^{-k} . Такое распределение вероятностей близко к равномерному на отрезке $[0, 1]$, причем, разность соответствующих функций распределения не превосходит 10^{-k} . Следовательно, реализации ξ можно рассматривать как реализации случайных величин, равномерно распределенных на отрезке $[0, 1]$. Эти реализации обычно называют равномерно распределенными случайными числами. Для программной реализации случайных чисел используют датчики случайных чисел.

Произведем оценку полученного с помощью метода Монте-Карло значения интеграла. Хотелось бы, чтобы ошибка была не больше ε :

$$\left| \frac{\mu}{n} - p \right| \leq \varepsilon,$$

где $p = \int_0^1 h(u) du$ — вероятность успеха (попадания в A). Пусть вероятность выполнения этого неравенства будет больше некоторого наперед заданного числа $1 - \alpha$:

$$P \left\{ \left| \frac{\mu}{n} - p \right| \leq \varepsilon \right\} \geq 1 - \alpha,$$

или в эквивалентном виде:

$$P \left\{ -\varepsilon \sqrt{\frac{n}{pq}} \leq \frac{\mu - np}{\sqrt{npq}} \leq \varepsilon \sqrt{\frac{n}{pq}} \right\} \geq 1 - \alpha.$$

По интегральной теореме Муавра–Лапласа вероятность в левой части близка для больших n к интегралу:

$$\frac{1}{\sqrt{2\pi}} \int_{-\varepsilon\sqrt{\frac{n}{pq}}}^{\varepsilon\sqrt{\frac{n}{pq}}} e^{-\frac{x^2}{2}} dx,$$

откуда следует, что должно выполняться неравенство: $\varepsilon\sqrt{\frac{n}{pq}} > u_{1-\alpha/2}$, где $u_{1-\alpha/2}$ — квантиль уровня $1-\alpha/2$ стандартного нормального распределения. Квантиль $u_{1-\alpha/2}$ определяется из условия:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_{1-\alpha/2}} e^{-\frac{x^2}{2}} dx = 1 - \frac{\alpha}{2}.$$

Если α задано, то можно найти нижнюю границу для количества экспериментов n :

$$n > \frac{u_{1-\alpha/2}^2}{\varepsilon^2} pq.$$

Очевидно, что $\max_{p \in [0,1]} p(1-p) = 1/4$, тогда можно получить следующее ограничение на количество испытаний:

$$n > \frac{u_{1-\alpha/2}^2}{4\varepsilon^2}. \quad (5.1)$$

Замечание 5.1. Оценка (5.1) имеет место и для многомерного случая, поскольку число испытаний не зависит от размерности пространства.

§ 6. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН

Случайные величины обычно моделируют с помощью преобразований одного или нескольких независимых значений случайной величины ξ , равномерно распределенной на отрезке $[0, 1]$.

*Моделирование случайных величин с дискретным
распределением*

Пусть требуется моделировать случайную величину η с дискретным законом распределения с конечным числом значений:

$$P\{\eta = x_i\} = p_i, \quad i = 1, \dots, k$$

Разобьем отрезок $[0, 1]$ на полуинтервалы $\delta_1 = [0, p_1)$, $\delta_2 = [p_1, p_1 + p_2)$, \dots , $\delta_k = [p_1 + \dots + p_{k-1}, 1]$.

Алгоритм моделирования случайной величины следующий:

1. Получим реализацию случайной величины, равномерно распределенной на отрезке $[0, 1]$. Пусть реализовалось число α .
2. Находим интервал среди $\delta_1, \dots, \delta_k$, которому принадлежит число α . Пусть для определенности это будет интервал δ_i .
3. Тогда присваиваем значение x_i для реализации случайной величины η .

Если требуется получить реализацию целочисленной случайной величины η с распределением $p_k = P\{\eta = k\}$ ($k = 0, 1, \dots$) с конечным или счетным числом значений, то можно использовать рекуррентное соотношение:

$$p_{i+1} = p_i r(i). \quad (6.1)$$

Пусть α — реализация равномерно распределенной на отрезке $[0, 1]$ случайной величины. Последовательно вычисляем разности $\alpha - \sum_{i=0}^k p_i$, пока не получится отрицательное число. Пусть первый раз отрицательной оказалась разность $\alpha - \sum_{i=0}^k p_i$, тогда полагаем $\eta = k$. Рекуррентное соотношение (6.1) можно использовать для последовательного вычисления вероятностей и накопленных сумм.

Пример 6.1. Для биномиального распределения с параметрами (p, n) вероятность того, что случайная величина η принимает значение i равна

$$p_i = C_n^i p^i (1-p)^{n-i}.$$

Найдем вид функции $r(i)$:

$$\begin{aligned} r(i) = \frac{p_{i+1}}{p_i} &= \frac{n!}{(i+1)!(n-i-1)!} \frac{i!(n-i)!}{n!} \frac{p}{1-p} = \\ &= \frac{n-i}{i+1} \frac{p}{1-p}, \quad i = 0, \dots, n. \end{aligned}$$

Пример 6.2. Для случайной величины, подчиняющейся распределению Пуассона с параметром λ вероятность того, что случайная величина η принимает значение i равна

$$p_i = \frac{\lambda^i}{i!} e^{-\lambda}.$$

Функция $r(i)$ имеет следующий вид:

$$r(i) = \frac{\lambda}{i+1}, \quad i = 0, 1, \dots$$

Пример 6.3. Для геометрического распределения с параметром p вероятность того, что случайная величина η принимает значение i , равна

$$p_i = p(1-p)^i.$$

Функция $r(i)$ имеет следующий вид:

$$r(i) = (1-p), \quad i = 0, 1, \dots$$

Моделирование случайных величин с непрерывными функциями распределения

Пусть требуется произвести моделирование случайной величины η с непрерывной функцией распределения $F_\eta(x) = P\{\eta \leq x\}$. Также сделаем предположение, что функция $F_\eta(x)$ монотонно возрастающая на некотором интервале (x_1, x_2) и постоянна вне этого интервала, $F_\eta(x_1) = 0$, $F_\eta(x_2) = 1$, в частном случае $x_1 = -\infty$, $x_2 = \infty$. Рассмотрим обратную к $F_\eta(x)$ функцию $F_\eta^{-1}(y)$. Нетрудно показать, что случайная величина $F_\eta^{-1}(\xi)$, где ξ — случайная величина,

равномерно распределенная на отрезке $[0, 1]$, имеет функцию распределения $F_\eta(x)$, где $x \in (x_1, x_2)$:

$$P\{F_\eta^{-1}(\xi) \leq x\} = P\{\xi \leq F_\eta(x)\} = F_\eta(x).$$

Таким образом, в случае монотонного возрастания функции $F_\eta(x)$ справедлива формула:

$$\eta = F_\eta^{-1}(\xi), \quad (6.2)$$

представляющая способ моделирования случайной величины η с непрерывной функцией распределения $F_\eta(x)$.

Пример 6.4. Пусть требуется получить значения случайной величины η подчиняющейся равномерному на отрезке $[a, b]$ распределению с функцией распределения:

$$F_\eta(x) = \begin{cases} 0, & \text{если } x < a; \\ \frac{x-a}{b-a}, & \text{если } a \leq x < b; \\ 1, & \text{если } x \geq b. \end{cases}$$

Решая уравнение $(x - a)/(b - a) = y$, найдем обратную функцию. Формула (6.2) будет иметь вид:

$$\eta = (b - a)\xi + a,$$

где ξ — случайная величина, равномерно распределенная на отрезке $[0, 1]$.

Пример 6.5. Пусть требуется получить значения непрерывной случайной величины η подчиняющейся экспоненциальному закону распределения с параметром λ с функцией распределения:

$$F_\eta(x) = \begin{cases} 0, & \text{если } x < 0; \\ 1 - e^{-\lambda x}, & \text{если } x \geq 0. \end{cases}$$

Решая уравнение $1 - e^{-\lambda x} = y$, найдем обратную функцию $F^{-1}(y) = -1/\lambda \ln(1 - y)$. Тогда формула (6.2) будет иметь вид:

$$\eta = -\frac{1}{\lambda} \ln(1 - \xi),$$

или

$$\eta = -\frac{1}{\lambda} \ln \xi,$$

поскольку случайная величина $1 - \xi$, также как и ξ , подчиняется равномерному на $[0, 1]$ распределению.

ЦЕПИ МАРКОВА

§ 1. ПОСЛЕДОВАТЕЛЬНОСТИ ЗАВИСИМЫХ
ИСПЫТАНИЙ. ЦЕПИ МАРКОВА

Рассмотрим последовательность испытаний, в каждом из которых возможно r исходов из множества $\{1, \dots, r\}$. Пусть испытание с номером 0 закончилось исходом a_0 , испытание с номером 1 — исходом a_1 , ..., испытание с номером T — исходом a_T . Упорядоченная последовательность исходов $\omega = (a_0, a_1, \dots, a_T)$, где $a_i \in \{1, \dots, r\}$, образует элементарное событие. Множество элементарных событий обозначим через Ω . Возможно другое описание рассматриваемого вероятностного пространства.

Пусть имеется система, которая может находиться в некотором фазовом пространстве в состояниях $1, 2, \dots, r$. Время предполагается дискретным и может принимать значения $1, 2, \dots, t, \dots, T$. Элементарное событие $\omega = (a_0, a_1, \dots, a_T)$ — траектория движения, где a_k — положение системы в момент времени k . Пространство элементарных событий $\Omega = \{\omega = (a_0, a_1, \dots, a_t, \dots, a_T) : a_t = 1, \dots, r\}$ конечно и $|\Omega| = r^{T+1}$.

Рассмотрим совокупность подмножеств $\mathcal{M} = \{M_\alpha\}$.

Определение 1.1. Будем говорить, что алгебра \mathcal{A} порождена совокупностью \mathcal{M} и записывать $\mathcal{A} = \alpha(\mathcal{M})$, если выполнены следующие условия:

1. Совокупность $\mathcal{M} \subset \mathcal{A}$,
2. Если D — некоторая алгебра, относительно которой известно, что $\mathcal{M} \subset D$, тогда $\mathcal{A} \subset D$.

Замечание 1.1. Термин «порожденная» равносильно выражению «минимальная, содержащая \mathcal{M} »: $\mathcal{A} = \cap_{\gamma} D_{\gamma} : D_{\gamma} \supset \mathcal{M}$, где D_{γ} — всевозможные алгебры, содержащие \mathcal{M} .

Замечание 1.2. Если в определении 1.1 заменить термин «алгебра» на термин « σ -алгебра», то получим определение σ -алгебры, порожденной семейством \mathcal{M} .

Замечание 1.3. Каким бы ни было семейство \mathcal{M} , порожденные им алгебра и σ -алгебра существуют.

Рассмотрим конечное разбиение множества Ω :

$$\Omega = \bigcup_{i=1}^m D_i : D_i \cap D_j = \emptyset.$$

Рассмотрим семейство $\mathcal{M} = \{D_1, \dots, D_m\}$. Очевидно, что

$$\alpha(\mathcal{M}) = \left\{ D_1, \dots, D_m, \emptyset, \left\{ \bigcup_{j=1}^k D_{l_j} \right\}_{k=1, \dots, m} \right\}$$

— совокупность всевозможных конечных объединений множеств D_i .

Пример 1.1. Частица переходит из одного состояния в другое. Если предположить, что время дискретно, и наблюдения за частицей происходят в промежутке времени $[0, T]$, тогда траектория движения частицы соответствует элементарному событию $\omega = (a_0, a_1, \dots, a_T)$, где a_0 — начальное положение частицы, a_1 — положение частицы в момент времени 1, ..., a_T — конечное положение частицы.

Выделим следующие события, которые являются подмножествами множества Ω :

$$A_0^{(1)} = \{\omega : a_0 = 1\},$$

$$A_0^{(2)} = \{\omega : a_0 = 2\},$$

...

$$A_0^{(r)} = \{\omega : a_0 = r\}.$$

Множества $\{A_0^{(1)}, A_0^{(2)}, \dots, A_0^{(r)}\}$ образуют конечное разбиение множества Ω , поскольку $\Omega = \bigcup_{i=1}^r A_0^{(i)}$ и $A_0^{(i)} \cap A_0^{(j)} = \emptyset$ для любых $i \neq j$. Определим алгебру $\mathcal{A}_0 = \alpha(A_0^{(1)}, A_0^{(2)}, \dots, A_0^{(r)})$. В эту алгебру входят все события, характеризующие начальное положение частицы.

Аналогичным образом определим события $\{A_1^{(1)}, A_1^{(2)}, \dots, A_1^{(r)}\}$, характеризующие положения частицы в момент времени 1, где $A_1^{(1)} = \{\omega : a_1 = 1\}$, $A_1^{(2)} = \{\omega : a_1 = 2\}$, \dots , $A_1^{(r)} = \{\omega : a_1 = r\}$. Определим алгебру $\mathcal{A}_1 = \alpha(A_1^{(1)}, A_1^{(2)}, \dots, A_1^{(r)})$. Для любого момента времени t можно аналогично построить алгебру \mathcal{A}_t . Таким образом, можно построить последовательность алгебр $\mathcal{A}_0, \mathcal{A}_1, \dots, \mathcal{A}_T$. Очевидно, что $\mathcal{A}_t \subset 2^\Omega$. Предположим, что существует вероятностная мера $P(\cdot)$, заданная на 2^Ω . Введем ограничение: $P(\omega) > 0$ для любого элементарного события $\omega \in \Omega$. Введем алгебры $\mathcal{A}_0^{t-1} = \alpha(\mathcal{A}_0, \dots, \mathcal{A}_{t-1})$, $\mathcal{A}_{t+1}^T = \alpha(\mathcal{A}_{t+1}, \dots, \mathcal{A}_T)$.

Определение 1.2. Последовательность испытаний $\mathcal{A}_0, \dots, \mathcal{A}_T$ образует цепь Маркова, если для любого целочисленного момента времени $t \in [1, T-1]$, любого события $A \in \mathcal{A}_0^{t-1}$, любого события $B \in \mathcal{A}_{t+1}^T$ и для любого исхода k выполняется условие:

$$P(A \cap B / \{a_t = k\}) = P(A / \{a_t = k\})P(B / \{a_t = k\}). \quad (1.1)$$

Замечание 1.4. Алгебра $\alpha(\mathcal{A}_0, \dots, \mathcal{A}_{t-1})$ содержит события, относящиеся к поведению частицы до момента $t-1$ включительно. Алгебра $\alpha(\mathcal{A}_{t+1}, \dots, \mathcal{A}_T)$ содержит события, относящиеся к поведению частицы после момента $t+1$, включая его. Момент времени t — текущий момент времени. Получили процесс, распадающийся на прошлое ($\alpha(\mathcal{A}_0, \dots, \mathcal{A}_{t-1}) = \mathcal{A}_0^{t-1}$), настоящее \mathcal{A}_t и будущее ($\alpha(\mathcal{A}_{t+1}, \dots, \mathcal{A}_T) = \mathcal{A}_{t+1}^T$).

Цепь Маркова — последовательность испытаний, в которой прошлое и будущее условно независимы при фиксированном настоящем.

Лемма 1.1. Последовательность испытаний образует цепь Маркова тогда и только тогда, когда для любого момента времени $t = 1, \dots, T-1$, любого исхода $k \in \{1, \dots, r\}$, любого

события $A \in \mathcal{A}_0^{t-1}$ и любого события $B \in \mathcal{A}_{t+1}^T$ выполняется условие:

$$P(B/A \cap \{a_t = k\}) = P(B/\{a_t = k\}). \quad (1.2)$$

Доказательство. Достаточность. Пусть выполнено (1.2), покажем справедливость (1.1):

$$\begin{aligned} P(A \cap B/\{a_t = k\}) &= \\ &= \frac{P(\{a_t = k\})P(A/\{a_t = k\})P(B/A \cap \{a_t = k\})}{P\{a_t = k\}} = \\ &= P(A/\{a_t = k\})P(B/\{a_t = k\}). \end{aligned}$$

Необходимость. Пусть выполнено (1.1), покажем справедливость (1.2):

$$\begin{aligned} P(B/A \cap \{a_t = k\}) &= \frac{P(\{a_t = k\})P(A \cap B/\{a_t = k\})}{P(\{a_t = k\})P(A/\{a_t = k\})} = \\ &= \frac{P(A/\{a_t = k\})P(B/\{a_t = k\})}{P(A/\{a_t = k\})} = P(B/\{a_t = k\}). \end{aligned}$$

■

Лемма 1.1 дает эквивалентное определение цепи Маркова.

Рассмотрим некоторую траекторию $\omega = (i_0, \dots, i_T)$ движения частицы по состояниям. Найдём вероятность осуществления траектории ω , используя формулу умножения вероятностей. Получаем выражение:

$$\begin{aligned} P(\omega) &= P(\{i_0, \dots, i_T\}) = P(A_0^{(i_0)} \cap \dots \cap A_T^{(i_T)}) = \\ &= P(A_0^{(i_0)})P(A_1^{(i_1)}/A_0^{(i_0)})P(A_2^{(i_2)}/A_1^{(i_1)}) \dots P(A_T^{(i_T)}/A_{T-1}^{(i_{T-1})}). \end{aligned}$$

Вероятность любого элементарного события полностью определяется начальными вероятностями $P(A_0^i)$, $i = 1, \dots, r$, и условными вероятностями $P(A_{t+1}^j/A_t^i)$. Вероятности $P(A_{t+1}^j/A_t^i)$ называются *переходными вероятностями за один шаг*. Зафиксируем момент времени t и сформируем матрицу переходных вероятностей $P_t = \{P(A_{t+1}^{(j)}/A_t^{(i)})\}$, где

$i = 1, \dots, r$ — номер строки, $j = 1, \dots, r$ — номер столбца. Зафиксируем строку i и сложим элементы этой строки:

$$P(A_{t+1}^{(1)}/A_t^{(i)}) + P(A_{t+1}^{(2)}/A_t^{(i)}) + \dots + P(A_{t+1}^{(r)}/A_t^{(i)}) = 1.$$

События $A_{t+1}^{(1)}, \dots, A_{t+1}^{(r)}$ попарно несовместны, а их объединение дает достоверное событие, вероятность которого равна единице.

Определение 1.3. Квадратная матрица с неотрицательными элементами, у которой сумма элементов в любой строке равна единице, называется стохастической.

Определение 1.4. Цепь Маркова называется однородной, если переходные вероятности за один шаг не зависят от номера испытания (момента времени t), $P(A_{t+1}^{(j)}/A_t^{(i)}) = p_{ij}$.

Обозначим через $P = \{p_{ij}\}_{i,j=1,\dots,r}$ матрицу переходных вероятностей. Найдем вероятность перехода из состояния i в состояние j за два шага:

$$\begin{aligned} P(A_{t+2}^{(j)}/A_t^{(i)}) &= P\left\{\bigcup_{k=1}^r (A_{t+2}^{(j)} \cap A_{t+1}^{(k)})/A_t^{(i)}\right\} = \\ &= \sum_{k=1}^r P(A_{t+2}^{(j)} \cap A_{t+1}^{(k)}/A_t^{(i)}) = \\ &= \sum_{k=1}^r \frac{P(A_t^{(i)})P(A_{t+1}^{(k)}/A_t^{(i)})P(A_{t+2}^{(j)}/A_{t+1}^{(k)})}{P(A_t^{(i)})} = \sum_{k=1}^r p_{ik}p_{kj}. \end{aligned}$$

Полученная формула доказывает, что для однородной цепи Маркова переходные вероятности за два шага не зависят от момента времени t . Аналогично можно показать, что переходные вероятности за любое число шагов не зависят от момента времени t .

Обозначим через $p_{ij}(k) = P(A_{t+k}^{(j)}/A_t^{(i)})$ вероятность перехода частицы из состояния i в состояние j за k шагов. Очевидно, что $P(A_{t+1}^{(j)}/A_t^{(i)}) = p_{ij} = p_{ij}(1)$.

Лемма 1.2 (уравнение Чепмена–Колмогорова). Для однородной цепи Маркова для любого момента времени $t \geq 1$ и любого $s \geq 1$ переходную вероятность из состояния i в состояние j за $t + s$ шагов можно найти по формуле

$$p_{ij}(t + s) = \sum_{k=1}^r p_{ik}(t) p_{kj}(s). \quad (1.3)$$

Доказательство. Покажем, что уравнение (1.3) справедливо для однородной цепи Маркова:

$$\begin{aligned} p_{ij}(t + s) &= P(A_{t+s}^{(j)} / A_0^{(i)}) = \sum_{k=1}^r P(A_t^{(k)} \cap A_{t+s}^{(j)} / A_0^{(i)}) = \\ &= \sum_{k=1}^r \frac{P(A_0^{(i)}) P(A_t^{(k)} / A_0^{(i)}) P(A_{t+s}^{(j)} / A_t^{(k)})}{P(A_0^{(i)})} = \\ &= \sum_{k=1}^r p_{ik}(t) p_{kj}(s). \end{aligned}$$

■

Запишем уравнение (1.3) в матричном виде:

$$P(t + s) = P(t)P(s), \quad (1.4)$$

где $P(t)$ — матрица переходных вероятностей за t шагов, $P(1) = P$. Если взять $s = 1$, тогда получим формулу $P(t+1) = P(t)P(1)$ верную для любого t . Следовательно, для любого $t \geq 1$ верно равенство:

$$P(t) = P^t. \quad (1.5)$$

По определению можно положить, что $P^0 = \mathbb{E}$, где \mathbb{E} — единичная матрица порядка r .

§ 2. ТЕОРЕМА О ПРЕДЕЛЬНЫХ ВЕРОЯТНОСТЯХ ДЛЯ ЦЕПЕЙ МАРКОВА

Теорема 2.1. Пусть задана однородная цепь Маркова, и существует момент времени t_0 такой, что $p_{ij}(t_0) > 0$ для любых $i, j = 1, \dots, r$, тогда существуют $\lim_{t \rightarrow \infty} p_{ij}(t) = p_j$ для любого $j = 1, \dots, r$. Причем пределы не зависят от начального состояния. Предельные вероятности p_1, \dots, p_r являются единственным решением системы линейных уравнений:

$$\sum_{j=1}^r x_j = 1, \quad x_j = \sum_{k=1}^r x_k p_{kj}, \quad j = 1, \dots, r.$$

Доказательство. Рассмотрим матрицу переходных вероятностей $P(t)$ и выберем произвольным образом столбец $(p_{1j}(t), p_{2j}(t), \dots, p_{rj}(t))'$. Выберем максимальный и минимальный элемент в столбце:

$$M_j(t) = \max_i p_{ij}(t),$$

$$m_j(t) = \min_i p_{ij}(t).$$

Очевидно, что справедливо неравенство:

$$m_j(t) \leq p_{ij}(t) \leq M_j(t). \quad (2.1)$$

Рассмотрим матрицу переходных вероятностей $P(t+1)$. Воспользовавшись уравнением (1.3), получим:

$$p_{ij}(t+1) = \sum_{k=1}^r p_{ik} p_{kj}(t).$$

Тогда $m_j(t) \leq p_{ij}(t+1) \leq M_j(t)$ для любого i . Для любого t справедливы неравенства:

$$m_j(t) \leq m_j(t+1) \leq p_{ij}(t+1) \leq M_j(t+1) \leq M_j(t),$$

следовательно, $m_j(t)$ — неубывающая последовательность, ограниченная сверху, а $M_j(t)$ — невозрастающая последовательность, ограниченная снизу: $m_j(t) \leq 1$, $M_j(t) \geq 0$.

Значит, для любого j существуют пределы $\lim_{t \rightarrow \infty} M_j(t)$, $\lim_{t \rightarrow \infty} m_j(t)$, кроме того, выполнено неравенство:

$$0 \leq M_j(t+1) - m_j(t+1) \leq M_j(t) - m_j(t).$$

Причем последовательность разностей монотонна и ограничена снизу, следовательно, для любого j существует предел $\lim_{t \rightarrow \infty} (M_j(t) - m_j(t))$.

Предположим, что верно следующее утверждение:

$$\lim_{t \rightarrow \infty} (M_j(t) - m_j(t)) = 0. \quad (2.2)$$

Тогда из (2.1) видно, что если существует предел

$$\lim_{t \rightarrow \infty} m_j(t) = \lim_{t \rightarrow \infty} M_j(t) = p_j,$$

то для любого i выполняется $\lim_{t \rightarrow \infty} p_{ij}(t) = p_j$. Докажем справедливость утверждения (2.2). Из уравнения (1.3) следуют соотношения:

$$M_j(t_0 + t) = p_{uj}(t_0 + t) = \sum_{k=1}^r p_{uk}(t_0) p_{kj}(t),$$

$$m_j(t_0 + t) = p_{vj}(t_0 + t) = \sum_{k=1}^r p_{vk}(t_0) p_{kj}(t).$$

Найдем разность:

$$\begin{aligned} M_j(t_0 + t) - m_j(t_0 + t) &= \sum_{k=1}^r (p_{uk}(t_0) - p_{vk}(t_0)) p_{kj}(t) = \\ &= \sum_{k=1}^+ (p_{uk}(t_0) - p_{vk}(t_0)) p_{kj}(t) - \sum_{k=1}^- |p_{uk}(t_0) - p_{vk}(t_0)| p_{kj}(t), \end{aligned}$$

все слагаемые в обеих суммах неотрицательны, т. е. во вторую сумму входят только слагаемые $(p_{uk}(t_0) - p_{vk}(t_0)) p_{kj}(t) < 0$. Заметим, что

$$\begin{aligned} \sum_{k=1}^+ (p_{uk}(t_0) - p_{vk}(t_0)) - \sum_{k=1}^- |p_{uk}(t_0) - p_{vk}(t_0)| &= \\ &= \sum_{k=1}^r (p_{uk}(t_0) - p_{vk}(t_0)) = 0. \end{aligned}$$

Введем обозначение $d_{u,v} = \sum^+ (p_{uk}(t_0) - p_{vk}(t_0))$, причем $0 \leq d_{u,v} < 1$. Если рассмотреть произвольные u, v , то ничего не изменится. Введем

$$\max_{(u,v)}(d_{u,v}) = d.$$

Очевидно, что $0 \leq d < 1$. Таким образом, для любого $t \geq 1$ оценка исходной разности примет следующий вид:

$$M_j(t_0 + t) - m_j(t_0 + t) \leq d(M_j(t) - m_j(t)).$$

При $t = t_0$ получаем:

$$M_j(2t_0) - m_j(2t_0) \leq d(M_j(t_0) - m_j(t_0)),$$

при $t = 2t_0$ получаем:

$$M_j(2t_0 + t_0) - m_j(2t_0 + t_0) \leq d^2(M_j(t_0) - m_j(t_0)).$$

Получаем, что для любого l справедливо неравенство:

$$M_j((l+1)t_0) - m_j((l+1)t_0) \leq d^l(M_j(t_0) - m_j(t_0)),$$

где $0 \leq d < 1$. Заметим, что при $l \rightarrow \infty$ имеет место сходимость: $d^l(M_j(t_0) - m_j(t_0)) \rightarrow 0$. Следовательно, предел последовательности существует и равен нулю.

Докажем вторую часть теоремы. Рассмотрим систему уравнений:

$$\sum_{j=1}^r x_j = 1, \quad x_j = \sum_{k=1}^r x_k p_{kj}, \quad j = 1, \dots, r.$$

Запишем систему в матричном виде $x^T = x^T P$. Перейдем к пределу в равенстве:

$$\sum_{j=1}^r p_{ij}(t) = 1$$

при $t \rightarrow \infty$, получим $\sum_{j=1}^r p_j = 1$.

Из уравнения Чепмена–Колмогорова имеем: $p_{ij}(t+1) = \sum_{k=1}^r p_{ik}(t)p_{kj}$. При $t \rightarrow \infty$ получаем равенство: $p_j = \sum_{k=1}^r p_k p_{kj}$.

Докажем единственность. Домножим на P уравнение $x^T = x^T P$, получим уравнение $x^T P = x^T P^2$. Домножим полученное уравнение на P , получим уравнение $x^T P^2 = x^T P^3$ и так далее. Если рассмотреть полученные уравнения в виде системы уравнений, то окажется, что если x — решение, то $x^T = x^T P^t$ для любого $t \in \mathbb{Z}$ или в скалярной форме: $x_j = \sum_{k=1}^r x_k p_{kj}(t)$. Пусть $t \rightarrow \infty$, тогда $x_j = \sum_{k=1}^r x_k p_j = p_j$. Таким образом, любое решение совпадает с тем, которое мы получили. ■

ОСНОВНЫЕ ВЕРОЯТНОСТНЫЕ ПРОСТРАНСТВА

§ 1. ПОЛУАЛГЕБРЫ. ТЕОРЕМЫ О ПРОДОЛЖЕНИИ МЕРЫ

Определение 1.1. Систему подмножеств \mathcal{S} множества Ω будем называть полуалгеброй, если выполнены следующие условия:

1. $\Omega \in \mathcal{S}$.
2. Если $A \in \mathcal{S}, B \in \mathcal{S}$, то $A \cap B \in \mathcal{S}$.
3. Если $A \in \mathcal{S}$, то множество \bar{A} можно представить в виде $\bar{A} = \sum_{k=1}^m B_k$, где $B_k \in \mathcal{S}$, $B_i \cap B_j = \emptyset$ для всех $i \neq j$.

Определение 1.2. Неотрицательная числовая функция $\mu: \mathcal{S} \rightarrow [0, \infty]$ называется конечно-аддитивной мерой на полуалгебре \mathcal{S} , если выполняются следующие условия:

1. $\mu(\emptyset) = 0$.
2. Для всех множеств $B_1, \dots, B_k \in \mathcal{S}$ таких, что $B_i \cap B_j = \emptyset$ для всех $i \neq j$, $\sum_{i=1}^k B_i \in \mathcal{S}$ выполнено:

$$\mu\left(\sum_{i=1}^k B_i\right) = \sum_{i=1}^k \mu(B_i).$$

Объединение множеств $\bigcup_{i=1}^k B_i$ будем обозначать через $\sum_{i=1}^k B_i$, если $B_i \cap B_j = \emptyset$ для всех $i \neq j$.

Если $\mu(\Omega) = 1$, то конечно-аддитивная мера μ называется конечно-аддитивной вероятностной мерой и обозначается P .

Определение 1.3. Пусть μ — конечно-аддитивная мера на \mathcal{S} . Если для всех множеств $B_k \in \mathcal{S}$ таких, что $B_i \cap B_j = \emptyset$ для всех $i \neq j$, $\sum_{i=1}^{\infty} B_i \in \mathcal{S}$ выполнено: $\mu(\sum_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i)$, то μ — счетно-аддитивная мера на полуалгебре \mathcal{S} .

Если существует последовательность множеств $\{A_i\}$, $A_i \in \mathcal{S}$ для любого i , при этом $A_i \subset A_{i+1}$ и $\bigcup_{i=1}^{\infty} A_i = \Omega$, $\mu(A_i) < +\infty$ для любого i , то счетно-аддитивная мера μ называется σ -конечной мерой. Если $\mu(\Omega) = 1$, то μ — счетно-аддитивная вероятностная мера.

Как отмечалось в главе 3, алгебра \mathcal{A} порождена системой множеств \mathcal{M} , и обозначать такую алгебру будем через $\alpha(\mathcal{M})$, если $\mathcal{A} \supset \mathcal{M}$ и, кроме того, любая другая алгебра \mathcal{D} , содержащая \mathcal{M} , содержит алгебру \mathcal{A} , $\mathcal{A} \subset \mathcal{D}$.

Теорема 1.1. Пусть \mathcal{S} — полуалгебра подмножеств множества Ω . Семейство \mathcal{A} , состоящее из всевозможных конечных объединений непересекающихся элементов полуалгебры \mathcal{S} представляет собой алгебру, порожденную полуалгеброй \mathcal{S} , $\mathcal{A} = \alpha(\mathcal{S})$.

Доказательство. Для доказательства теоремы достаточно заметить, что система множеств, о которой говорится в условии теоремы, является алгеброй, содержит полуалгебру \mathcal{S} и, кроме того, она должна содержаться в любой алгебре, содержащей полуалгебру \mathcal{S} . ■

Теорема 1.2 (теорема о продолжении конечно-аддитивной меры с полуалгебры на алгебру). Пусть \mathcal{S} — полуалгебра подмножеств множества Ω . Пусть на \mathcal{S} задана конечно-аддитивная мера μ , тогда существует и единственная конечно-аддитивная мера ν , определенная на $\alpha(\mathcal{S})$ и такая, что для любого множества $E \in \mathcal{S}$: $\nu(E) = \mu(E)$.

Доказательство. В силу теоремы 1.1 любое множество $B \in \alpha(\mathcal{S})$ может быть представлено следующим образом: $B = \sum_{i=1}^m E_i$, где $E_i \cap E_j = \emptyset$, $i \neq j$, $E_i \in \mathcal{S}$. Следовательно, по определению можно записать представление:

$$\nu(B) = \sum_{i=1}^m \mu(E_i). \quad (1.1)$$

Необходимо показать, что функция ν задана корректно, и что она является конечно-аддитивной. Представим, что множество B можно записать в виде конечной суммы: $B = \sum_{j=1}^n F_j$, $F_i \cap F_j = \emptyset$, $i \neq j$, $F_j \in \mathcal{S}$, из чего следует, что должно выполняться равенство:

$$\nu(B) = \sum_{j=1}^n \mu(F_j). \quad (1.2)$$

Покажем, что из (1.1) следует (1.2). Поскольку $\sum_{j=1}^n F_j = \sum_{i=1}^m E_i$, следовательно, F_j представимо в виде $\sum_{i=1}^m F_j \cap E_i$. Тогда $\mu(F_j) = \sum_{i=1}^m \mu(F_j \cap E_i)$. Рассуждая аналогично, получаем $\mu(E_i) = \sum_{j=1}^n \mu(F_j \cap E_i)$. Берем от первого выражения сумму по j , от второго — по i и получаем (1.2). Нетрудно проверить, что функция $\nu(\cdot)$ конечно-аддитивная. ■

Теорема 1.3. Пусть \mathcal{S} — полуалгебра подмножеств множества Ω . Пусть на \mathcal{S} задана счетно-аддитивная σ -конечная мера μ . Тогда существует и единственна счетно-аддитивная σ -конечная мера ν , определенная на $\alpha(\mathcal{S})$, такая, что для любого множества $E \in \mathcal{S}$ имеет место равенство: $\nu(E) = \mu(E)$.

Доказательство. По предположению μ — счетно-аддитивная, следовательно, μ — конечно-аддитивная мера. По теореме 1.2 существует единственная конечно-аддитивная мера ν , для которой выполнено условие теоремы. Нетрудно показать, что мера ν , определенная равенством (1.1), счетно-аддитивна. ■

Замечание 1.1. Теоремы 1.2 и 1.3 относятся к конечно-аддитивной и счетно-аддитивной мерам, заданным на полуалгебре.

Теорема 1.4 (о продолжении σ -конечной счетно-аддитивной меры с алгебры на σ -алгебру). Пусть \mathcal{A} — алгебра подмножеств множества Ω . Пусть на \mathcal{A} задана σ -конечная счетно-аддитивная мера μ . Тогда существует и единственна счетно-аддитивная σ -конечная мера ν , определенная на σ -алгебре, порожденной алгеброй \mathcal{A} такая, что

для любого множества $E \in \mathcal{A}$ имеет место равенство $\nu(E) = \mu(E)$.

Доказательство. Доказательство теоремы можно найти в [27]. ■

Замечание 1.2. 1. Из теорем 1.2, 1.3, 1.4 следует, что для того, чтобы задать меру на σ -алгебре, достаточно задать ее на полуалгебре, если σ -алгебра порождена полуалгеброй.

2. Все теоремы справедливы для вероятностных мер.

Пара (Ω, \mathcal{F}) называется измеримым пространством. Будем предполагать, что σ -алгебра \mathcal{F} порождается полуалгеброй \mathcal{S} , то есть

$$\mathcal{F} = \sigma(\mathcal{S})$$

или можно рассматривать два этапа: $\mathcal{F} = \sigma(\alpha(\mathcal{S}))$.

Замечание 1.3. Рассмотрим измеримое пространство и некоторую меру на нем. Может оказаться так, что σ -алгебра \mathcal{F} неполна относительно меры μ . Это означает, что существует множество E нулевой меры, $\mu(E) = 0$, что подмножества множества E могут не принадлежать σ -алгебре \mathcal{F} . В этом случае всегда можно провести процедуру пополнения σ -алгебры, добавив в нее все подмножества множества нулевой меры. Строгое изложение процедуры пополнения содержится, например, в [27].

§ 2. ПРИМЕРЫ ИЗМЕРИМЫХ ПРОСТРАНСТВ

Пример 2.1. Пусть $\Omega = \mathbb{R}$ — вся числовая ось. Рассмотрим семейство множеств:

$$\mathcal{I}_1 = \{(-\infty, +\infty), (-\infty, a], (a, b], (b, +\infty)\},$$

где $a \in \mathbb{R}, b \in \mathbb{R}$. Нетрудно проверить, что \mathcal{I}_1 — полуалгебра. Назовем $\sigma(\alpha(\mathcal{I}_1)) = \sigma(\mathcal{I}_1) = \mathcal{B}(\mathbb{R})$ борелевской σ -алгеброй на числовой прямой. Справедливо равенство:

$$(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n} \right] \in \mathcal{B}(\mathbb{R}).$$

Как известно [19], любое открытое множество на числовой прямой представляет собой не более чем счетное объединение интервалов. Так как операция счетного объединения не выводит из σ -алгебры, то все открытые множества будут входить в борелевскую σ -алгебру, следовательно, все замкнутые множества также будут входить в борелевскую σ -алгебру. Так же будут входить в нее и все одноточечные множества:

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a \right].$$

Пример 2.2. Рассмотрим декартово n -кратное произведение числовой прямой $\Omega = \mathbb{R}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}\}$. Введем множество

$$I = I_1 \times I_2 \times \dots \times I_n,$$

где $I_1 \in \mathcal{I}_1(\mathbb{R}), \dots, I_n \in \mathcal{I}_1(\mathbb{R})$.

Рассмотрим систему подмножеств $\mathcal{I}_2 = \{I = I_1 \times \dots \times I_n\}$. Нетрудно показать, что \mathcal{I}_2 — полуалгебра. Введем множество $B = B_1 \times \dots \times B_n$, где $B_1 \in \mathcal{B}(\mathbb{R}), \dots, B_n \in \mathcal{B}(\mathbb{R})$. Рассмотрим всевозможные множества $\{B = B_1 \times \dots \times B_n\} = \mathcal{I}_3$. Нетрудно показать, что \mathcal{I}_3 — полуалгебра.

Понятно, что $\mathcal{I}_3 \supset \mathcal{I}_2$. Из этого следует, что $\alpha(\mathcal{I}_3) \supset \alpha(\mathcal{I}_2)$, но можно показать [35], что $\sigma(\alpha(\mathcal{I}_3)) = \sigma(\alpha(\mathcal{I}_2)) = \mathcal{B}(\mathbb{R}^n)$. Назовем борелевской σ -алгеброй σ -алгебру $\mathcal{B}(\mathbb{R}^n)$.

Пример 2.3. Рассмотрим $\Omega = \mathbb{R}^\infty = \{(x_1, \dots, x_n, \dots) : x_k \in \mathbb{R}\}$. Введем цилиндрические множества $C_n(I_1 \times \dots \times I_n) = \{(x_1, \dots, x_n, \dots) \in \mathbb{R}^\infty : x_1 \in I_1, \dots, x_n \in I_n\}$, где $I_1 \in \mathcal{I}_1, \dots, I_n \in \mathcal{I}_1$. Рассмотрим семейство всех таких цилиндрических множеств:

$$\{C_n(I_1 \times \dots \times I_n)\} = \mathcal{I}_4.$$

Нетрудно показать, что \mathcal{I}_4 — полуалгебра. Можно рассмотреть цилиндры другого вида $C_n(B_1 \times \dots \times B_n) = \{(x_1, \dots, x_n, \dots) : x_1 \in B_1, \dots, x_n \in B_n\}$, где $B_1 \in \mathcal{B}(\mathbb{R}), \dots, B_n \in \mathcal{B}(\mathbb{R})$. Рассмотрим семейство таких цилиндрических множеств:

$$\{C_n(B_1 \times \dots \times B_n)\} = \mathcal{I}_5 \supset \mathcal{I}_4.$$

Система множеств \mathcal{I}_5 также является полуалгеброй. Возьмем в основании цилиндра множество B , а именно

$$C_n(B) = \{(x_1, \dots, x_n, \dots) : (x_1, \dots, x_n) \in B, \quad B \in \mathcal{B}(\mathbb{R}^n)\}$$

и рассмотрим совокупность множеств $\{C_n(B)\} = \mathcal{I}_6 \supset \mathcal{I}_5$. Система множеств \mathcal{I}_6 образует алгебру. Можно показать, что σ -алгебры, порожденные введенными системами множеств совпадают: $\sigma(\mathcal{I}_6) = \sigma(\mathcal{I}_5) = \sigma(\mathcal{I}_4) = \mathcal{B}(\mathbb{R}^\infty)$. Далее будем называть σ -алгебру $\mathcal{B}(\mathbb{R}^\infty)$ σ -алгеброй борелевских множеств бесконечномерного пространства.

Замечание 2.1. В примере 2.1 можем рассматривать расширенную числовую прямую $\Omega = [-\infty, +\infty]$, т.е. добавляем ко множеству \mathbb{R} еще два элемента: $-\infty$ и $+\infty$. Нетрудно построить полуалгебру и σ -алгебру на расширенной числовой прямой, включив в нее два одноточечных множества $\{-\infty\}$ и $\{\infty\}$.

§ 3. ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$

Пусть вероятностная мера P задана на борелевской σ -алгебре $\mathcal{B}(\mathbb{R})$. Для любого $A \in \mathcal{B}(\mathbb{R})$ задана вероятность $P(A)$. Будем рассматривать событие $B = (-\infty, x]$, $x \in \mathbb{R}$. Определим функцию распределения $F(x)$ следующим образом:

$$F(x) = P(B) = P(-\infty, x].$$

Свойства функции распределения

1. Пусть $x_1 < x_2$, тогда $F(x_1) \leq F(x_2)$, функция $F(x)$ монотонно неубывающая.

Доказательство. Как легко заметить, $P(-\infty, x_2] = P(-\infty, x_1] + P(x_1, x_2]$. Следовательно, $F(x_2) \geq F(x_1)$. ■

2. В любой точке $y \in \mathbb{R}$ выполняется равенство: $F(y+0) = F(y)$, функция непрерывна справа, $\lim_{x \rightarrow y+0} F(x) = F(y)$, и существует предел слева $\lim_{x \rightarrow y-0} F(x)$, который может не совпадать с $F(y)$.

Доказательство. Пусть $x < y$, следовательно по первому свойству $F(x) \leq F(y)$, но тогда существование предела $\lim_{x \rightarrow y-0} F(x)$ очевидно вследствие монотонности функции. Возьмем любую последовательность $\{x_n\} \xrightarrow[n \rightarrow \infty]{} y + 0$, $x_{n+1} < x_n$. Рассмотрим $F(x_n) - F(y) = P(y, x_n]$. Введем событие $A_n = (y, x_n]$, тогда $A_{n+1} \subset A_n$. Очевидно, что $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Счетно-аддитивная вероятностная мера непрерывна. Следовательно, можно гарантировать, что $P(y, x_n] \xrightarrow[n \rightarrow \infty]{} 0$. ■

3. Справедливы следующие равенства: $\lim_{x \rightarrow \infty} F(x) = F(\infty) = 1$, $\lim_{x \rightarrow -\infty} F(x) = F(-\infty) = 0$.

Доказательство. В силу монотонности $F(x)$ достаточно рассмотреть любую возрастающую последовательность, поэтому выберем последовательность натуральных чисел $\{n\}$, она возрастающая, и последовательность $\{-n\}$ — убывающая. Рассмотрим разность $F(n) - F(-n) = P(-\infty, n] - P(-\infty, -n] = P(-n, n]$. Обозначим через A_n полуинтервал $(-n, n]$, тогда $A_n \subset A_{n+1}$, $\bigcup_{n=1}^{\infty} A_n = \mathbb{R}$.

По свойству непрерывности:

$$P(-n, n] \xrightarrow[n \rightarrow \infty]{} P(\mathbb{R}) = 1,$$

но тогда $\lim_{n \rightarrow \infty} (F(n) - F(-n)) = 1$, что равносильно, в виду монотонности, $F(n) \xrightarrow[n \rightarrow \infty]{} 1$, $F(-n) \xrightarrow[n \rightarrow \infty]{} 0$. ■

Определение 3.1. Рассмотрим на числовой прямой класс функций, удовлетворяющий свойствам 1, 2, 3. Будем называть представителей этого класса функциями распределения.

Теорема 3.1. Пусть $F(x)$ — некоторая произвольно выбранная функция распределения, тогда существует и единственная вероятностная мера P , заданная на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ такая, что для любых a, b , $a < b$ имеет место равенство $P(a, b] = F(b) - F(a)$.

Доказательство. Составим полуалгебру $\mathcal{I}_1 = \{(-\infty, \infty), (-\infty, a], (a, b], (b, +\infty)\}$, где $a, b \in \mathbb{R}$. Далее рассмотрим σ -алгебру $\sigma(\alpha(\mathcal{I}_1)) = \sigma(\mathcal{I}_1)$. Введем функцию P_0 на полуалгебре и убедимся, что она конечно-аддитивна, затем, что счетно-аддитивна: $P_0(-\infty, \infty) = 1 = F(\infty) - F(-\infty)$, $P_0(-\infty, a] = F(a)$, $P_0(a, b] = F(b) - F(a)$, $P_0(b, +\infty) = 1 - F(b)$.

Видим, что функция P_0 неотрицательна. Проверим, что она конечно-аддитивна. Для этого разобьем промежуток $(a, b]$ точкой $c: (a, b] = (a, c] \cup (c, b]$, где $c \in (a, b]$. $P_0(a, b] = F(b) - F(a) = F(c) - F(a) + F(b) - F(c) = P_0(a, c] + P_0(c, b]$. Аналогично можно проверить конечную аддитивность для любых элементов полуалгебры. По теореме § 1. о продолжении меры с полуалгебры на алгебру считаем, что P_0 — конечно-аддитивная вероятностная мера на $\alpha(\mathcal{I}_1)$.

Докажем, что P_0 — счетно-аддитивная мера на полуалгебре. Пусть $(a, b] = \sum_{i=1}^{\infty} (a_i, b_i]$, $(a_i, b_i] \cap (a_j, b_j] = \emptyset$, $i \neq j$. Теперь для любых $\varepsilon_i > 0$ и для любого $\delta > 0$ имеет место включение $[a + \delta, b] \subset \bigcup_{i=1}^{\infty} (a_i, b_i + \varepsilon_i)$, $\varepsilon_i > 0$. Для компактного множества из любого открытого покрытия можно выделить конечное подпокрытие, поэтому, существует номер m , что $[a + \delta, b] \subset \bigcup_{i=1}^m (a_i, b_i + \varepsilon_i)$, тогда $(a + \delta, b] \subset \bigcup_{i=1}^m \{(a_i, b_i] \cup (b_i, b_i + \varepsilon_i)\}$, но тогда

$$\begin{aligned} P_0(a + \delta, b] &\leq P_0 \left\{ \bigcup_{i=1}^m \{(a_i, b_i] \cup (b_i, b_i + \varepsilon_i)\} \right\} \leq \\ &\leq \sum_{i=1}^m (P_0(a_i, b_i] + P_0(b_i, b_i + \varepsilon_i]) \leq \\ &\leq \sum_{i=1}^{\infty} P_0(a_i, b_i] + \sum_{i=1}^{\infty} (F(b_i + \varepsilon_i) - F(b_i)). \quad (3.1) \end{aligned}$$

Функция $F(x)$ непрерывна справа в каждой точке, следовательно, можно выбрать ε_i так, чтобы разность $F(b_i + \varepsilon_i) - F(b_i)$ была меньше $\varepsilon/2^i$. Следовательно для любого $\varepsilon > 0$ и для любого $\delta > 0$ справедливо неравенство:

$$P_0(a + \delta, b] = F(b) - F(a + \delta) \leq \sum_{i=1}^{\infty} P_0(a_i, b_i] + \varepsilon.$$

Можем устремить ε и δ к 0, следовательно $P_0(a, b] \leq \leq \sum_{i=1}^{\infty} P_0(a_i, b_i]$. С другой стороны, очевидно, что $(a, b] \supset \supset \sum_{i=1}^m (a_i, b_i]$ при любом m . Тогда $P_0(a, b] \geq \sum_{i=1}^m P_0(a_i, b_i]$. Переходя к пределу при $m \rightarrow \infty$, получаем противоположное неравенство. Следовательно, $P_0(a, b] = \sum_{i=1}^{\infty} P_0(a_i, b_i]$.

Возьмем любое подмножество $I \subset \mathcal{I}_1$: $I = \bigcup_{n=-\infty}^{+\infty} I \cap (n, n+1]$. Нетрудно показать, что

$$P_0(I) = \sum_{n=-\infty}^{+\infty} P_0(I \cap (n, n+1]).$$

Например, рассмотрим следующее множество: $I = (b, +\infty)$, тогда $\bigcup_{n=-\infty}^{+\infty} \{(b, +\infty) \cap (n, n+1]\} = (b, l] + \bigcup_{n=l}^{+\infty} (n, n+1]$, $l-1 \leq b < l$. Очевидно, что

$$\begin{aligned} \sum_{n=-\infty}^{+\infty} P_0(I \cap (n, n+1]) &= \\ &= F(l) - F(b) + \lim_{k \rightarrow \infty} \sum_{n=l}^k (F(n+1) - F(n)) = \\ &= \lim_{k \rightarrow \infty} (F(l) - F(b) + F(l+1) - F(l) + \dots + F(k+1)) = \\ &= \lim_{k \rightarrow \infty} (F(k+1) - F(b)) = F(\infty) - F(b) = \\ &= 1 - F(b) = P_0(I). \end{aligned}$$

Аналогично рассматриваются все другие элементы полуалгебры. Теперь докажем счетную аддитивность меры P_0 на полуалгебре. Пусть $I = \sum_{k=1}^{\infty} I_k$, $I \in \mathcal{I}_1$, $I_k \in \mathcal{I}_1$, $I_k \cap I_n = \emptyset$, $k \neq n$. По доказанному ранее получаем:

$$\begin{aligned} P_0(I) &= \sum_{n=-\infty}^{+\infty} P_0(I \cap (n, n+1]) = \\ &= \sum_{n=-\infty}^{+\infty} P_0\left(\bigcup_{k=1}^{+\infty} (I_k \cap (n, n+1])\right) = \\ &= \sum_{n=-\infty}^{+\infty} \sum_{k=1}^{\infty} P_0(I_k \cap (n, n+1]) = \end{aligned}$$

$$= \sum_{k=1}^{\infty} \sum_{n=-\infty}^{+\infty} P_0(I_k \cap (n, n+1]) = \sum_{k=1}^{\infty} P_0(I_k).$$

Отсюда следует, что P_0 — счетно-аддитивная мера на полуалгебре, поэтому, по теореме 1.4 она единственным образом продолжается на σ -алгебру. ■

Замечание 3.1. Каждой вероятностной мере P на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ соответствует единственная функция распределения F и наоборот.

Отметим ряд важных свойств:

$$\begin{aligned} P(\{a\}) &= P\left\{\bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a\right]\right\} = \\ &= \lim_{n \rightarrow \infty} P\left\{\left(a - \frac{1}{n}, a\right]\right\} = \lim_{n \rightarrow \infty} (F(a) - F(a - \frac{1}{n})) = \\ &= F(a) - F(a - 0). \end{aligned}$$

Если функция $F(x)$ непрерывна в точке a , то $P(\{a\}) = 0$. Также справедливы равенства:

$$\begin{aligned} P(a, b) &= P(a, b] - P(\{b\}) = \\ &= F(b) - F(a) - F(b) + F(b - 0) = F(b - 0) - F(a), \end{aligned}$$

$$\begin{aligned} P[a, b] &= P(\{a\}) + P(a, b] = \\ &= F(a) - F(a - 0) + F(b) - F(a) = F(b) - F(a - 0). \end{aligned}$$

§ 4. КЛАССИФИКАЦИЯ ВЕРОЯТНОСТНЫХ МЕР

Дискретная вероятностная мера

Определение 4.1. Будем говорить, что мера P , заданная на числовой оси, дискретна, если существует не более, чем счетная совокупность точек $\{a_i\}$: $P\{a_i\} > 0$, причем $\sum_i P\{a_i\} = 1$, точки $\{a_i\}$ называются носителями меры.

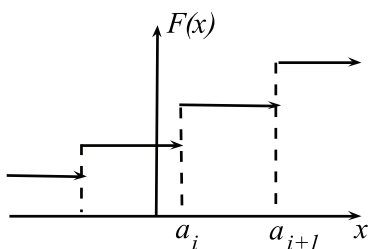


Рис. 4.1

Функция распределения, соответствующая дискретной мере

Так как $P\{a_i\} = F(a_i) - F(a_i - 0) > 0$, то $\sum_i (F(a_i) - F(a_i - 0)) = 1$, тогда функция $F(x)$ может быть только кусочно-постоянной (рис. 4.1).

Абсолютно непрерывная вероятностная мера

Определение 4.2. Будем говорить, что вероятностная мера P , заданная на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, абсолютно непрерывна относительно меры Лебега (для краткости, абсолютно непрерывна), если существует $f(x) \geq 0$ такая, что для любого борелевского множества $B \in \mathcal{B}(\mathbb{R})$ выполнено:

$$P(B) = \int_B f(x) dx. \quad (4.1)$$

В равенстве (4.1) рассматривается интеграл Лебега по мере Лебега, неотрицательную функцию $f(x)$ называют плотностью вероятностной меры P относительно меры Лебега.

Если $B = (-\infty, x]$, то $P(B) = F(x)$, из (4.1) следует, что

$$F(x) = \int_{-\infty}^x f(t) dt. \quad (4.2)$$

Таким образом, формула (4.2) — частный случай формулы (4.1). Нетрудно показать, что (4.1) и (4.2) эквивалентны, то

есть неотрицательная функция $f(x)$ — плотность, если для любого $x \in \mathbb{R}$ выполнено (4.2).

Замечание 4.1. Важное свойство интеграла Лебега заключается в том, что если мы изменим под знаком интеграла подынтегральную функцию на множестве меры 0, то интеграл от этого никак не изменится.

Замечание 4.2. Монотонные функции почти всюду дифференцируемы и почти всюду имеет место равенство (см. [19]):

$$F'(x) \stackrel{\text{п.в.}}{=} f(x). \quad (4.3)$$

Если для любого $x \in \mathbb{R}$ функция $f(x)$ неотрицательна и $\int_{-\infty}^{+\infty} f(x)dx = 1$, то $\int_{-\infty}^x f(t)dt$ — функция распределения на \mathbb{R} , которой взаимно однозначно соответствует мера P такая, что $P(B) = \int_B f(x)dx$.

Функцию

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

называют плотностью стандартного нормального распределения.

Сингулярная мера. Сингулярная функция распределения

Определение 4.3. Пусть $F(x)$ — некоторая произвольная функция распределения на \mathbb{R} . Будем говорить, что точка z — точка роста функции F , если для любого $\varepsilon > 0$ справедливо:

$$F(z + \varepsilon) - F(z - \varepsilon) > 0.$$

Определение 4.4. Будем говорить, что непрерывная функция распределения $F(x)$ сингулярна, если множество точек роста имеет меру Лебега равную нулю. Вероятностная мера, взаимно однозначно соответствующая сингулярной функции распределения, называется сингулярной.

Для сингулярной функции выполняется соотношение: $F'(x) \stackrel{\text{п.н.}}{=} 0$, но тогда $F(x) \neq \int_{-\infty}^x F'(x)dx$, при этом $F(x)$ непрерывна. Следовательно, функция распределения не восстанавливается по своей производной.

Пример 4.1. *Пример Кантора сингулярной функции распределения.* Кусочно-линейная непрерывная функция $F_0(x)$ представлена графиком (рис. 4.2).

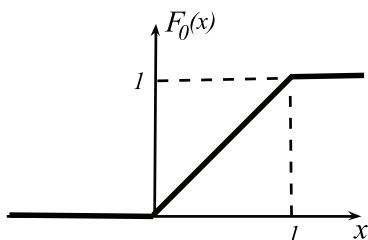


Рис. 4.2
Функция $F_0(x)$

Непрерывная функция распределения $F_1(x)$ равна нулю левее точки $x = 0$ и равна 1 правее точки $x = 1$. Отрезок $[0, 1]$ делим на 3 равные части. Значение функции внутри среднего интервала постоянно и равно среднему арифметическому ближайших слева и справа уже заданных значений (рис. 4.3).

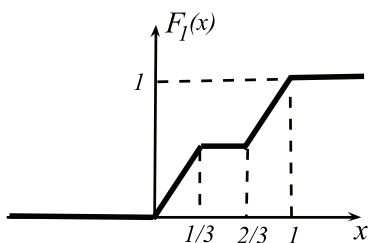


Рис. 4.3
Функция $F_1(x)$

Значение функции на оставшихся интервалах определим с помощью линейной интерполяции, т. о. определим функцию

$F_1(x)$ на отрезке $[0, 1]$ следующим образом:

$$F_1(x) = \begin{cases} 0, & x = 0, \\ 1/2, & x \in (1/3, 2/3), \\ 1, & x = 1, \end{cases} \quad (4.4)$$

доопределяя ее в остальных точках отрезка $[0, 1]$ с помощью линейной интерполяции.

Далее каждый из отрезков $[0, 1/3]$ и $[2/3, 1]$ также разбиваем на три части по вышеописанному правилу, определяем функцию:

$$F_2(x) = \begin{cases} 0, & x = 0, \\ 1/4, & x \in (1/9, 2/9), \\ 1/2, & x \in (1/3, 2/3), \\ 3/4, & x \in (7/9, 8/9), \\ 1, & x = 1, \end{cases} \quad (4.5)$$

доопределяя ее в остальных точках отрезка $[0, 1]$ с помощью линейной интерполяции (рис. 4.4).

Таким образом строится последовательность функций распределения $F_n(x)$. Все функции $F_n(x)$ непрерывны. Для любого x последовательность $F_n(x)$ сходится в себе, то есть для

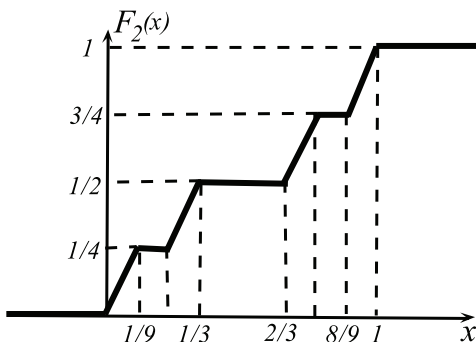


Рис. 4.4
Функция $F_2(x)$

любого $\varepsilon > 0$ существует номер n_0 такой, что для любых $n > n_0, m > n_0$ справедливо неравенство: $|F_n(x) - F_m(x)| < \varepsilon$. То есть существует предел $\lim_{n \rightarrow \infty} F_n(x) = F(x)$. В данном случае сходимость не только поточечная, но и равномерная:

$$\sup_x |F(x) - F_n(x)| \xrightarrow{n \rightarrow \infty} 0.$$

Следовательно, предельная функция $F(x)$ — непрерывная функция распределения на числовой оси. Функция $F(x)$ является сингулярной функцией распределения. Проанализируем участки функции внутри отрезка $[0, 1]$, где она является постоянной:

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots = \frac{1}{3} \left(1 + \frac{2}{3} + \left(\frac{2}{3} \right)^2 + \dots \right) = \frac{1}{3} \frac{1}{(1 - \frac{2}{3})} = 1.$$

Таким образом, суммарная длина участков, содержащихся в $[0, 1]$, на которых $F(x)$ является постоянной, равна 1. Производная функции $F(x)$ почти всюду существует и почти всюду равна 0.

Теорема 4.1 (теорема Лебега). Любая функция распределения $F(x)$ на числовой прямой представима в следующем виде:

$$F(x) = p_1 F_1(x) + p_2 F_2(x) + p_3 F_3(x),$$

где $p_1, p_2, p_3 \geq 0, p_1 + p_2 + p_3 = 1$, $F_1(x)$ — кусочно-постоянная функция распределения, $F_2(x)$ — абсолютно-непрерывная функция распределения, $F_3(x)$ — сингулярная функция распределения.

§ 5. КОНЕЧНОМЕРНОЕ ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

Рассмотрим вероятностное пространство $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P)$. Для любого $B \in \mathcal{B}(\mathbb{R}^n)$ задана вероятность $P(B)$. Возьмем в качестве B множество специального вида:

$$B = (-\infty, x_1] \times \dots \times (-\infty, x_n].$$

Таким образом, определена функция $F(x_1, \dots, x_n): P(B) = F(x_1, \dots, x_n)$. Функция $F(x_1, \dots, x_n)$ называется функцией распределения, соответствующей вероятностной мере P в пространстве \mathbb{R}^n . Выясним ее свойства, введя для этого разностный оператор $\Delta_{a_i, b_i}^{(i)} h(x_1, \dots, x_n) = h(x_1, \dots, x_{i-1}, b_i, x_{i+1}, \dots, x_n) - h(x_1, \dots, x_{i-1}, a_i, x_{i+1}, \dots, x_n)$.

1. Последовательно применим разностный оператор (каждый по своей оси), тогда

$$\Delta_{a_1, b_1}^{(1)} \Delta_{a_2, b_2}^{(2)} \dots \Delta_{a_n, b_n}^{(n)} F(x_1, \dots, x_n) \geq 0,$$

здесь $a_i < b_i$, $i = 1, \dots, n$.

Доказательство. Рассмотрим случай $n = 2$, т.е.

$$\Delta_{a_1, b_1}^{(1)} \Delta_{a_2, b_2}^{(2)} F(x_1, x_2).$$

Имеют место равенства:

$$P\{(-\infty, x_1] \times (-\infty, b_2]\} = F(x_1, b_2),$$

$$P\{(-\infty, x_1] \times (-\infty, a_2]\} = F(x_1, a_2),$$

$$\Delta_{a_2, b_2}^{(2)} F(x_1, x_2) = P\{(-\infty, x_1] \times (a_2, b_2]\},$$

$$\Delta_{a_1, b_1}^{(1)} P\{(-\infty, x_1] \times (a_2, b_2]\} = P\{(a_1, b_1] \times (a_2, b_2]\} \geq 0.$$

■

2. Функция $F(x)$ непрерывна справа в любой точке по всем аргументам. Для любого $y = (y_1, \dots, y_n)$ имеет место равенство:

$$\lim_{\forall i: x_i \rightarrow y_i + 0} F(x_1, \dots, x_n) = F(y_1, \dots, y_n).$$

Доказательство. Рассмотрим множество $A_m = (-\infty, x_1(m)] \times \dots \times (-\infty, x_n(m)]$, где последовательности выбраны следующим образом: $x_1(m) \rightarrow y_1 + 0, \dots, x_n(m) \rightarrow y_n + 0$. Пусть $A = (-\infty, y_1] \times \dots \times (-\infty, y_n]$. Из построения очевидно, что $A_{m+1} \subset A_m$, и $A = \bigcap_{m=1}^{\infty} A_m$. По свойству непрерывности $P(A_m) \xrightarrow{m \rightarrow \infty} P(A)$, то есть $F(x_1(m), \dots, x_n(m)) \xrightarrow{m \rightarrow \infty} F(y)$. ■

3. Справедливы утверждения: $F(x_1, \dots, x_n) \xrightarrow[\forall i: x_i \rightarrow +\infty]{} 1$ и $F(x_1, \dots, x_n) \xrightarrow[\exists i: x_i \rightarrow -\infty]{} 0$, если существует $x_i \rightarrow -\infty$.

Доказательство. Пусть выбраны следующие последовательности: $x_1(m) \rightarrow +\infty, \dots, x_n(m) \rightarrow +\infty$. Пусть $A_m = (-\infty, x_1(m)] \times \dots \times (-\infty, x_n(m)]$. Очевидно, что $A_m \subset A_{m+1}$ и $\bigcup_m A_m = \mathbb{R}^n$, $P(A_m) \xrightarrow{m \rightarrow \infty} 1$.

Пусть мы зафиксировали кроме одного все остальные аргументы. Пусть $x_i(m) \xrightarrow{m \rightarrow \infty} -\infty$, $A_m = (-\infty, x_1] \times \dots \times (-\infty, x_{i-1}] \times (-\infty, x_i(m)] \times (-\infty, x_{i+1}] \times \dots \times (-\infty, x_n]$. Очевидно, что $A_{m+1} \subset A_m$, $\bigcap_{m=1}^{\infty} A_m = \emptyset$, тогда $P(A_m) \xrightarrow{m \rightarrow \infty} 0$. ■

Определение 5.1. Любая функция от n аргументов, удовлетворяющая свойствам 1, 2, 3, называется функцией распределения в \mathbb{R}^n .

Теорема 5.1. Пусть $F(x_1, \dots, x_n)$ — функция распределения в \mathbb{R}^n , тогда существует и единственная вероятностная мера P на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ такая, что для любых $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$ таких, что $a_i < b_i$ для любого i имеет место равенство:

$$P\{(a_1, b_1] \times \dots \times (a_n, b_n]\} = \Delta_{a_1, b_1}^{(1)} \dots \Delta_{a_n, b_n}^{(n)} F(x),$$

где $x = (x_1, \dots, x_n)$.

Доказательство аналогично доказательству соответствующей теоремы на прямой.

Замечание 5.1. Пусть $F_1(x), \dots, F_n(x)$ — произвольные функции распределения на числовой прямой \mathbb{R} , тогда $F(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n)$ является функцией распределения в пространстве \mathbb{R}^n .

Выделим 2 класса мер:

1. Дискретная вероятностная мера

Определение 5.2. Будем говорить, что мера P на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ дискретна, если существует не более чем

счетное множество $\{z_i = (z_{1i}, \dots, z_{ni})\}$ такое, что $P\{z_i\} > 0$ и $\sum_i P(z_i) = 1$.

2. Абсолютно непрерывная мера относительно меры Лебега

Определение 5.3. Будем говорить, что вероятностная мера P абсолютно непрерывна относительно меры Лебега в \mathbb{R}^n , если существует плотность, т.е. существует $f(x) \geq 0$ такая, что для любого $B \in \mathcal{B}(\mathbb{R}^n)$: $P(B) = \int_B f(x)dx$, где $x = (x_1, \dots, x_n)$.

Возьмем множество: $B = (-\infty, x_1] \times \dots \times (-\infty, x_n]$, следовательно, для любого $x \in \mathbb{R}^n$ выполнено равенство:

$$\int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_1 \dots dt_n = F(x_1, \dots, x_n). \quad (5.1)$$

Нетрудно показать, что можно дать эквивалентное определение плотности, опираясь на выражение (5.1).

Определение 5.4. Если для некоторой функции $f(x) \geq 0$ справедливо равенство (5.1), где $F(x)$ — функция распределения в \mathbb{R}^n , соответствующая мере P , то $f(x)$ — плотность меры P относительно меры Лебега.

Замечание 5.2. Рассмотрим выражение (5.1). Производная по верхнему пределу почти всюду существует и почти всюду совпадает со значениями f в точке:

$$\frac{d^n F(x)}{dx_1 \dots dx_n} \stackrel{\text{п.в.}}{=} f(x_1, \dots, x_n). \quad (5.2)$$

§ 6. ВЕРОЯТНОСТНОЕ ПРОСТРАНСТВО $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty), P)$

Пусть на пространстве $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ задана мера P . Сигма-алгебра в бесконечномерном пространстве определяется следующим образом: $\mathcal{B}(\mathbb{R}^\infty) = \sigma\{C_n(B)\}$, где $C_n(B)$ — цилиндрические множества, $C_n(B) = \{x = (x_1, \dots, x_n, \dots) : (x_1, \dots, x_n) \in B\}$, $B \in \mathcal{B}(\mathbb{R}^n)$. Пусть $P(C_n(B))$ определена.

Зафиксируем n , тогда можно определить вероятностную меру в конечномерном пространстве $P(C_n(B)) = P_n(B)$ для любого $B \in \mathcal{B}(\mathbb{R}^n)$, где $P_n(B)$ — вероятностная мера на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, n — любое, но фиксированное.

Мера P порождает бесконечную последовательность вероятностных мер: P_1, \dots, P_n, \dots , каждая мера задана на пространстве своей размерности, то есть P_1 задана на \mathbb{R} , P_2 задана на \mathbb{R}^2 и так далее. Очевидно, что $C_n(B) = C_{n+1}(B \times \mathbb{R})$, следовательно,

$$P_n(B) = P_{n+1}(B \times \mathbb{R}). \quad (6.1)$$

Условие (6.1) — условие согласованности мер, оно должно выполняться для всех n и любого множества $B \in \mathcal{B}(\mathbb{R}^n)$.

Верно и обратное: если имеется бесконечная последовательность вероятностных мер и выполнены условия согласованности (6.1), то существует и единственная вероятностная мера P , заданная на бесконечномерном пространстве.

Теорема 6.1. Пусть существует последовательность вероятностных пространств

$$\{(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_1), \dots, (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_n), \dots\}$$

таких, что для всех n и любого множества $B \in \mathcal{B}(\mathbb{R}^n)$ выполнено условие (6.1), тогда существует и единственная вероятностная мера P на $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ такая, что для всех n и любого множества $B \in \mathcal{B}(\mathbb{R}^n)$: $P(C_n(B)) = P_n(B)$.

Доказательство. Доказательство можно найти в [35]. ■

Замечание 6.1. Пусть $F(x)$ — любая функция распределения на числовой прямой \mathbb{R} . Нетрудно проверить, что $F_2(x_1, x_2) = F(x_1)F(x_2)$ — функция распределения на плоскости \mathbb{R}^2 , ..., $F_n(x_1, \dots, x_n) = F(x_1) \dots F(x_n)$ — функция распределения в \mathbb{R}^n и так далее. Каждой из таких функций распределения взаимно однозначно соответствует мера P_n . Нетрудно проверить, что последовательность P_1, \dots, P_n, \dots удовлетворяет условиям согласованности (6.1), следовательно, существует мера P на $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$, согласованная со всеми введенными в замечании мерами.

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 1. РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Определение 1.1. Случайной величиной, заданной на вероятностном пространстве (Ω, \mathcal{F}, P) или \mathcal{F} -измеримой числовой функцией называется функция $\xi : \Omega \rightarrow \mathbb{R}$ такая, что для любого $B \in \mathcal{B}(\mathbb{R})$: $\xi^{-1}(B) = \{\omega \mid \xi(\omega) \in B\} \in \mathcal{F}$ (полный прообраз любого борелевского множества содержится в σ -алгебре \mathcal{F}).

Пусть $B \in \mathcal{B}(\mathbb{R})$ — некоторое борелевское множество. Рассмотрим $\xi^{-1}(B) = \{\omega \mid \xi(\omega) \in B\} = \{\xi \in B\}$. Можно говорить о $P(\xi^{-1}(B)) = P(\xi \in B)$, так как полный прообраз $\xi^{-1}(B)$ принадлежит \mathcal{F} .

Определим функцию $P_\xi(B)$ как функцию аргумента $B \in \mathcal{B}(\mathbb{R})$ следующим образом:

$$P_\xi(B) = P(\xi^{-1}(B)) = P\{\xi \in B\}. \quad (1.1)$$

Определение 1.2. Функция P_ξ , определенная равенством (1.1) на σ -алгебре борелевских множеств, называется вероятностным распределением случайной величины ξ .

Замечание 1.1. $P_\xi(\cdot)$ представляет собой вероятностную меру на борелевской прямой $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Доказательство.

1. $P_\xi(\mathbb{R}) = P(\xi^{-1}(\mathbb{R})) = P(\Omega) = 1$.
2. Неравенство $P_\xi(B) \geq 0$ следует из определения.

3. Пусть $B_1, B_2 \in \mathcal{B}(\mathbb{R})$: $B_1 \cap B_2 = \emptyset$, тогда $\xi^{-1}(B_1 \cup B_2) = \xi^{-1}(B_1) \cup \xi^{-1}(B_2)$, причем $\xi^{-1}(B_1) \cap \xi^{-1}(B_2) = \emptyset$. Следовательно, справедливы равенства: $P_\xi(B_1 \cup B_2) = P\{\xi^{-1}(B_1 \cup B_2)\} = P\{\xi^{-1}(B_1) \cup \xi^{-1}(B_2)\} = P\{\xi^{-1}(B_1)\} + P\{\xi^{-1}(B_2)\} = P_\xi(B_1) + P_\xi(B_2)$. Таким образом, распределение ξ — конечно-аддитивная вероятностная мера.
4. Рассмотрим последовательность $\{B_k\}$: $B_k \in \mathcal{B}(\mathbb{R})$ для всех k , $B_k \cap B_m = \emptyset$, $k \neq m$. Повторяя рассуждения из предыдущего пункта, получаем, что

$$P_\xi \left(\bigcup_{k=1}^{\infty} B_k \right) = \sum_{k=1}^{\infty} P_\xi(B_k).$$

Таким образом, P_ξ — счетно-аддитивная вероятностная мера. ■

Замечание 1.2. Из определения 1.1 и замечания 1.1 следует, что функции $P_\xi(B)$ взаимно однозначно соответствует функция распределения $F_\xi(x)$. Если взять в качестве $B = (-\infty, x]$, получаем $P_\xi(-\infty, x] = P\{\xi \in (-\infty, x]\} = P\{\xi \leq x\} = F_\xi(x)$. Таким образом, можно перейти от вероятностного распределения случайной величины ξ к ее функции распределения.

Определение 1.3. Функцией распределения случайной величины ξ будем называть функцию

$$F_\xi(x) = P\{\xi \leq x\}. \quad (1.2)$$

Функция распределения $F_\xi(x) = P\{\xi \in (-\infty, x]\}$ обладает следующими свойствами:

1. Свойство монотонности. Пусть $x_1 < x_2$, тогда $F_\xi(x_1) \leq F_\xi(x_2)$.
2. Функция распределения непрерывна справа в любой точке числовой прямой:

$$F_\xi(y) \xrightarrow{y \rightarrow x, y > x} F_\xi(x)$$

для любой точки $x \in \mathbb{R}$.

3. Имеют место сходимости:

$$F_{\xi}(x) \xrightarrow{x \rightarrow +\infty} 1, \quad F_{\xi}(x) \xrightarrow{x \rightarrow -\infty} 0.$$

Замечание 1.3. Пусть $F_{\xi}(x)$ — функция распределения случайной величины ξ , тогда для любых точек $a < b$ имеет место равенство:

$$F_{\xi}(b) - F_{\xi}(a) = P\{a < \xi \leq b\}.$$

Любая точка на числовой прямой является борелевским множеством. Имеет место равенство:

$$P\{\xi = a\} = F_{\xi}(a) - F_{\xi}(a - 0).$$

Справедливы следующие равенства:

$$P\{a \leq \xi \leq b\} = F_{\xi}(b) - F_{\xi}(a - 0),$$

$$P\{a < \xi < b\} = F_{\xi}(b - 0) - F_{\xi}(a).$$

Лемма 1.1. Пусть $F(x)$ — произвольно выбранная функция распределения на \mathbb{R} , тогда существует случайная величина ξ такая, что для любого $x : F_{\xi}(x) \equiv F(x)$.

Доказательство. Возьмем в качестве $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, $\xi(\omega) \equiv \omega$. Каждой функции распределения $F(x)$ взаимно однозначно соответствует вероятностная мера P . Итак, определили $(\Omega, \mathcal{F}, P) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$. Таким образом, для любого x справедливы равенства: $F_{\xi}(x) = P\{\xi \leq x\} = P\{\xi^{-1}(-\infty, x]\} = P(-\infty, x] = F(x)$. ■

Замечание 1.4. В большинстве прикладных задач требуется знать только закон распределения случайной величины без задания самой функции $\xi(\omega)$.

Определение 1.4. Будем говорить, что случайная величина ξ дискретна, если ее распределение представляет собой дискретную вероятностную меру на \mathbb{R} , т. е. существует не более, чем счетная совокупность точек $\{z_i\}$ такая, что $P_{\xi}(z_i) = P\{\xi = z_i\} > 0$ и $\sum_i P\{\xi = z_i\} = 1$.

Если число слагаемых в сумме конечно (то есть набор z_i конечен), то случайная величина называется простой. Из определения следует, что $F_\xi(x)$ — кусочно-постоянная функция: $P_\xi(z_i) = F_\xi(z_i) - F_\xi(z_i - 0) > 0$, $\sum_i (F_\xi(z_i) - F_\xi(z_i - 0)) = 1$. Простую случайную величину можно записать следующим образом:

$$\xi(\omega) = \sum_{i=1}^m z_i I\{\omega \in A_i\},$$

где $A_i \in \mathcal{F}$, $A_i \cap A_j = \emptyset$ для $i \neq j$, $\bigcup_{i=1}^m A_i = \Omega$, $I\{\omega \in A_i\}$ — индикатор множества A_i :

$$I\{\omega \in A_i\} = \begin{cases} 1, & \omega \in A_i \\ 0, & \omega \notin A_i. \end{cases}$$

Если $A_i = \xi^{-1}(z_i)$, то $z_i \neq z_j$ для всех $i \neq j$. Может оказаться так, что A_i будет разбито на подмножества, например, $A_1 = D_1 + D_2 + D_3$, где $D_1, D_2, D_3 \in \mathcal{F}$, $D_i \cap D_j = \emptyset$ для всех $i \neq j$. Тогда $I(D_1 + D_2 + D_3) = I(D_1) + I(D_2) + I(D_3)$, поэтому, вообще говоря, некоторые коэффициенты z_i могут совпадать.

Определение 1.5. Будем говорить, что распределение $P_\xi(\cdot)$ случайной величины ξ абсолютно непрерывно относительно меры Лебега, если существует функция $f_\xi(x)$, заданная на всей числовой прямой, такая, что $f_\xi(x) \geq 0$ для любого $x \in \mathbb{R}$ и для любого $B \in \mathcal{B}(\mathbb{R})$ выполнено:

$$P_\xi(B) = \int_B f_\xi(x) dx. \quad (1.3)$$

Определение 1.6. Функцию $f_\xi(x)$ из определения 1.5 называют плотностью распределения случайной величины ξ относительно меры Лебега.

Из определения 1.5 следует, что если $B = (-\infty, x]$, где $x \in \mathbb{R}$, то справедливо равенство:

$$F_\xi(x) = \int_{-\infty}^x f_\xi(y) dy. \quad (1.4)$$

Как было отмечено ранее, равенства (1.3) и (1.4) представляют собой эквивалентные определения функции плотности. Плотность распределения определяется единственным образом с точностью до множеств Лебега меры нуль по формуле:

$$\frac{dF_{\xi}(x)}{dx} \stackrel{\text{п.в.}}{=} f_{\xi}(x). \quad (1.5)$$

Функция распределения $F_{\xi}(x)$ называется абсолютно непрерывной функцией, если у нее существует плотность распределения относительно меры Лебега.

Как было отмечено ранее, на числовой прямой могут быть заданы сингулярные функции распределения и соответствующие им вероятностные меры. Соответствующие случайные величины можно было бы назвать сингулярными.

Если случайная величина дискретная, то для задания распределения достаточно задать множество точек $\{z_i\}$ и «весов», $P_{\xi}(z_i) > 0$, $\sum_i P_{\xi}(z_i) = 1$. Для того, чтобы задать абсолютно-непрерывное распределение случайной величины, достаточно задать плотность распределения $f_{\xi}(x) \geq 0$ такую, что

$$\int_{-\infty}^{\infty} f_{\xi}(x) dx = 1.$$

Пример 1.1. *Распределение Бернулли.* Случайная величина принимает значения 0 или 1:

$$\xi = 0 \cdot I(\bar{A}) + 1 \cdot I(A) = I(A),$$

где вероятность $P_{\xi}(1) = p$, $P_{\xi}(0) = 1 - p = q$.

Пример 1.2. *Биномиальное распределение.* Случайная величина принимает значение, равное числу успехов в серии из n независимых испытаний Бернулли:

$$P\{\mu = m\} = C_n^m p^m q^{n-m}, \quad m = 0, 1, \dots, n.$$

Пример 1.3. *Распределение Пуассона.* Случайная величина ξ распределена по закону Пуассона:

$$P\{\xi = m\} = \frac{\lambda^m}{m!} e^{-\lambda}, \quad \lambda > 0, \quad m = 0, 1, \dots$$

Пример 1.4. Стандартное нормальное распределение $N(0, 1)$ с плотностью распределения:

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Пример 1.5. Нормальное распределение $N(a, \sigma^2)$ с параметрами $a \in \mathbb{R}$ и σ^2 ($\sigma > 0$). Рассмотрим случайную величину $\eta = a + \sigma\xi$, $a \in \mathbb{R}$, $\sigma > 0$, где случайная величина ξ подчиняется стандартному нормальному распределению. Тогда функция распределения случайной величины η может быть определена следующим образом:

$$\begin{aligned} F_{\eta}(x) &= P\{\eta \leq x\} = P\{a + \sigma\xi \leq x\} = \\ &= P\left\{\xi \leq \frac{x-a}{\sigma}\right\} = F_{\xi}\left(\frac{x-a}{\sigma}\right). \end{aligned}$$

Функция плотности распределения случайной величины η легко находится:

$$\begin{aligned} f_{\eta}(x) &= F'_{\eta}(x) = \frac{1}{\sigma} F'_{\xi}\left(\frac{x-a}{\sigma}\right) = \\ &= \frac{1}{\sigma} f_{\xi}\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \end{aligned}$$

где $x \in \mathbb{R}$.

Пример 1.6. Гамма-распределение $G(\lambda, p)$.

Определение 1.7. Будем говорить, что случайная величина ξ подчиняется гамма-распределению с параметрами λ, p , если плотность имеет следующий вид:

$$f_{\xi}(x) = \begin{cases} \frac{\lambda^p x^{p-1}}{\Gamma(p)} e^{-\lambda x}, & \text{если } x > 0, \\ 0, & \text{если } x \leq 0, \end{cases}$$

где λ — параметр масштаба, p — параметр формы, $p, \lambda > 0$, $\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx$.

Пример 1.7. Экспоненциальное распределение является частным случаем гамма-распределения при $p = 1$. Плотность распределения имеет вид:

$$f_{\xi}(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{если } x > 0, \\ 0, & \text{если } x \leq 0. \end{cases}$$

Пример 1.8. Бета-распределение с параметрами p_1 и p_2 . Плотность распределения имеет вид:

$$f_{\xi}(x) = \begin{cases} \frac{x^{p_1-1}(1-x)^{p_2-1}}{B(p_1, p_2)}, & \text{если } x \in (0, 1), \\ 0, & \text{если } x \notin (0, 1), \end{cases}$$

где β -функция определяется следующим образом: $B(p_1, p_2) = \int_0^1 x^{p_1-1}(1-x)^{p_2-1}dx$, $p_1 > 0$, $p_2 > 0$.

Пример 1.9. Равномерное распределение является частным случаем бета-распределения при $p_1 = 1$ и $p_2 = 1$. Плотность равномерного распределения имеет вид:

$$f_{\xi}(x) = \begin{cases} 1, & \text{если } x \in [0, 1], \\ 0, & \text{если } x \notin [0, 1]. \end{cases}$$

§ 2. СВОЙСТВО ИЗМЕРИМОСТИ

Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) и случайную величину $\xi : \Omega \rightarrow \mathbb{R}$. Функция $\xi(\omega)$ является \mathcal{F} -измеримым отображением, т. е. для любого борелевского множества $B \in \mathcal{B}(\mathbb{R})$: $\xi^{-1}(B) \in \mathcal{F}$.

Определение 2.1. Будем называть функцию $h : \mathbb{R}^k \rightarrow \mathbb{R}^l$ борелевской, если $h^{-1}(B) \in \mathcal{B}(\mathbb{R}^k)$ для любого $B \in \mathcal{B}(\mathbb{R}^l)$.

Лемма 2.1. Пусть \mathcal{M} — семейство множеств на числовой прямой \mathbb{R} такое, что $\sigma(\mathcal{M}) = \mathcal{B}(\mathbb{R})$. Для того, чтобы функция $\xi : \Omega \rightarrow \mathbb{R}$ была \mathcal{F} -измеримой, необходимо и достаточно, чтобы для любого $C \in \mathcal{M}$ выполнялось равенство $\xi^{-1}(C) \in \mathcal{F}$.

Доказательство. *Необходимость.* Очевидно, что если $\sigma(\mathcal{M}) = \mathcal{B}(\mathbb{R})$, то $\mathcal{M} \subset \mathcal{B}(\mathbb{R})$, тогда для любого $C \in \mathcal{M}$ имеет место включение: $\xi^{-1}(C) \in \mathcal{F}$.

Достаточность. Рассмотрим семейство множеств

$$\mathcal{D} = \{A \in \mathcal{B}(\mathbb{R}) : \xi^{-1}(A) \in \mathcal{F}\}, \quad \mathcal{M} \subset \mathcal{D} \subset \mathcal{B}(\mathbb{R}).$$

Нетрудно доказать следующие равенства:

1. $\xi^{-1}(\bigcup_i T_i) = \bigcup_i \xi^{-1}(T_i),$
2. $\xi^{-1}(\bigcap_i T_i) = \bigcap_i \xi^{-1}(T_i),$
3. $\xi^{-1}(\bar{T}) = \overline{\xi^{-1}(T)}.$

Из равенств следует, что система множеств \mathcal{D} представляет собой σ -алгебру, и, следовательно, $\mathcal{D} = \mathcal{B}(\mathbb{R})$. ■

Замечание 2.1. Справедливо следующее равенство: $\sigma\{(-\infty, a], a \in \mathbb{R}\} = \mathcal{B}(\mathbb{R})$.

Лемма 2.2. Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) задана случайная величина $\xi(\omega)$ и задана борелевская функция $h : \mathbb{R} \rightarrow \mathbb{R}$. Тогда $\eta(\omega) = h(\xi(\omega))$ — \mathcal{F} -измеримая числовая функция, т. е. случайная величина.

Доказательство. Очевидно, что $\eta : \Omega \rightarrow \mathbb{R}$. Для произвольного борелевского множества $B \in \mathcal{B}(\mathbb{R})$ выполнено: $\eta^{-1}(B) = \xi^{-1}(h^{-1}(B)) \in \mathcal{F}$. ■

Из леммы следует, что элементарные функции от случайных величин являются случайными величинами. Если для любого $\omega \in \Omega$ существует предел:

$$\lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega),$$

причем $\xi_n(\omega)$ — \mathcal{F} -измеримые функции, то $\xi(\omega)$ также является \mathcal{F} -измеримой функцией [19, 35].

Лемма 2.3. Семейство полных прообразов $\xi^{-1}(B)$, $B \in \mathcal{B}(\mathbb{R})$ является σ -алгеброй.

Доказательство.

1. $\xi^{-1}(\mathbb{R}) = \Omega$.
2. Рассмотрим счетные или конечные совокупности множеств $\{B_i\}$ таких, что $B_i \in \mathcal{B}(\mathbb{R})$ для любого i . Тогда справедливо равенство: $\xi^{-1}(\bigcup_i B_i) = \bigcup_i \xi^{-1}(B_i)$, при этом $\bigcup_i B_i \in \mathcal{B}(\mathbb{R})$. Следовательно, $\xi^{-1}(\bigcup_i B_i) \in \mathcal{F}$.

$$3. \overline{\xi^{-1}(B)} = \xi^{-1}(\bar{B}).$$

Таким образом, семейство замкнуто относительно операции объединения и дополнения, следовательно, оно является σ -алгеброй, которая содержится в \mathcal{F} . ■

Определение 2.2. Сигма-алгебру $\sigma_\xi = \{\xi^{-1}(B), B \in \mathcal{B}(\mathbb{R})\}$ будем называть сигма-алгеброй, порожденной случайной величиной ξ .

§ 3. СЛУЧАЙНЫЕ ЭЛЕМЕНТЫ СО ЗНАЧЕНИЯМИ В КОНЕЧНОМЕРНОМ ПРОСТРАНСТВЕ \mathbb{R}^k

В дальнейшем будем говорить о σ -алгебре борелевских множеств $\mathcal{B}(\mathbb{R}^n)$ в пространстве \mathbb{R}^k , т. е.

$$\sigma\{(a_1, b_1] \times (a_2, b_2] \times \dots \times (a_k, b_k]\} = \mathcal{B}(\mathbb{R}^k).$$

К этой σ -алгебре можно прийти другим способом: пусть $B_1 \in \mathcal{B}(\mathbb{R}_1)$, $B_2 \in \mathcal{B}(\mathbb{R}_2)$, ..., $B_n \in \mathcal{B}(\mathbb{R}_k)$, тогда

$$\sigma\{B_1 \times B_2 \times \dots \times B_n\} = \mathcal{B}(\mathbb{R}^k).$$

Построенную σ -алгебру называют прямым произведением n сигма-алгебр и обозначают:

$$\mathcal{B}(\mathbb{R}^k) = \mathcal{B}(\mathbb{R}_1) \otimes \mathcal{B}(\mathbb{R}_2) \otimes \dots \otimes \mathcal{B}(\mathbb{R}_k).$$

Определение 3.1. Отображение $\xi : \Omega \rightarrow \mathbb{R}^k$ будем называть \mathcal{F} -измеримым, или, что то же самое, ξ является случайным элементом со значениями в \mathbb{R}^k или случайным вектором, если для любого множества $B \in \mathcal{B}(\mathbb{R}^k)$: $\xi^{-1}(B) \in \mathcal{F}$.

Тогда, по аналогии со скалярным случаем, определим вероятностное распределение случайного вектора ξ или совместное распределение случайных величин ξ_1, \dots, ξ_n :

$$P\{\xi^{-1}(B)\} = P\{\xi \in B\} = P_\xi(B).$$

Совместное распределение случайных величин ξ_1, \dots, ξ_k определено на σ -алгебре $\mathcal{B}(\mathbb{R}^k)$. Как и в скалярном случае,

совместное распределение является вероятностной мерой $P_\xi(\cdot)$ на $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. Леммы 2.1, 2.2, 2.3 переносятся на рассматриваемые отображения $\xi : \Omega \rightarrow \mathbb{R}^k$. При этом, конечно, $h : \mathbb{R}^k \rightarrow \mathbb{R}^m$ и $\eta(\omega) = h(\xi(\omega))$ — случайный элемент со значениями в \mathbb{R}^m .

Лемма 3.1. Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) задано отображение $\xi : \Omega \rightarrow \mathbb{R}^k$. Для того, чтобы отображение ξ было \mathcal{F} -измеримым, необходимо и достаточно, чтобы для любого $c \in \mathcal{M}$, $\sigma(\mathcal{M}) = \mathcal{B}(\mathbb{R}^k)$, выполнялось: $\xi^{-1}(c) \in \mathcal{F}$, где \mathcal{M} — система множеств.

Лемма 3.2. Пусть φ — борелевская функция, такая, что $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^m$, тогда если $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))^T$ — случайный вектор, то $\varphi(\xi(\omega))$ — случайный вектор.

Лемма 3.3. Пусть $F(x_1, \dots, x_k)$ — произвольная функция распределения в \mathbb{R}^k . Тогда существует случайный вектор $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))^T$, заданный на некотором вероятностном пространстве (Ω, \mathcal{F}, P) , такой, что

$$F_{\xi_1, \dots, \xi_k}(x_1, \dots, x_k) \equiv F(x_1, \dots, x_k).$$

Лемма 3.4. Семейство полных прообразов $\xi^{-1}(B)$, $B \in \mathcal{B}(\mathbb{R}^k)$, является σ -алгеброй.

Определение 3.2. Сигма-алгебру $\sigma_\xi = \{\xi^{-1}(B), B \in \mathcal{B}(\mathbb{R}^k)\}$ будем называть сигма-алгеброй, порожденной случайным вектором ξ .

Рассмотрим $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))^T$. Отображение $\xi_l(\omega) : \Omega \rightarrow \mathbb{R} = (-\infty, \infty)$. Для произвольного $B \in \mathcal{B}(\mathbb{R})$ справедливо следующее равенство:

$$\begin{aligned} \xi_l^{-1}(B) &= \xi_1^{-1}(\mathbb{R}) \cap \dots \cap \xi_{l-1}^{-1}(\mathbb{R}) \cap \xi_l^{-1}(B) \cap \xi_{l+1}^{-1}(\mathbb{R}) \cap \dots \\ &\dots \cap \xi_k^{-1}(\mathbb{R}) = \xi^{-1}(\mathbb{R} \times \dots \times \mathbb{R} \times B \times \mathbb{R} \times \dots \times \mathbb{R}) \in \mathcal{F}. \end{aligned}$$

Таким образом, любая проекция случайного вектора является случайной величиной. Верно и обратное. Пусть $\eta_1(\omega), \dots, \eta_k(\omega)$ — случайные величины, заданные на вероятностном пространстве (Ω, \mathcal{F}, P) , тогда $\eta(\omega) = (\eta_1(\omega), \dots, \eta_k(\omega))^T$ является случайным элементом со значениями в \mathbb{R}^k . Обратное доказательство можно провести, применив лемму 3.1.

§ 4. ФУНКЦИИ РАСПРЕДЕЛЕНИЯ В КОНЕЧНОМЕРНЫХ ПРОСТРАНСТВАХ

Рассмотрим конечномерное вероятностное пространство $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P_\xi)$.

Определение 4.1. Совместной функцией распределения случайных величин ξ_1, \dots, ξ_k (функцией распределения вектора $\xi = (\xi_1, \dots, \xi_k)$) называется функция

$$\begin{aligned} F_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k) &= F_\xi(x) = \\ &= P_\xi \{(-\infty, x_1] \times \dots \times (-\infty, x_k]\} = \\ &= P \{\xi_1 \leq x_1, \dots, \xi_k \leq x_k\}. \end{aligned} \quad (4.1)$$

Определение 4.2. Будем говорить, что распределение $P_\xi(\cdot)$ случайного вектора ξ в пространстве $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ абсолютно непрерывно относительно меры Лебега, если существует функция $f_\xi(x)$ такая, что $f_\xi(x) \geq 0$ для любого $x \in \mathbb{R}^k$, и для любого $B \in \mathcal{B}(\mathbb{R}^k)$ выполнено:

$$P_\xi(B) = \int_B f_\xi(x) dx. \quad (4.2)$$

Определение 4.3. Функцию $f_\xi(x) = f_{\xi_1 \dots \xi_k}(x_1, \dots, x_k)$ будем называть совместной плотностью распределения случайных величин ξ_1, \dots, ξ_k .

Выберем множество $B = (-\infty, x_1] \times \dots \times (-\infty, x_k]$, тогда справедливо равенство:

$$\begin{aligned} P\{(-\infty, x_1] \times \dots \times (-\infty, x_k]\} &= F_{\xi_1 \dots \xi_k}(x_1, \dots, x_k) = \\ &= \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_k} f(y_1, \dots, y_k) dy_1 \dots dy_k. \end{aligned} \quad (4.3)$$

Аналогично скалярному случаю нетрудно показать, что равенства (4.2) и (4.3) эквивалентны. Переходя к повторному

интегрированию, используя свойства интеграла Лебега с переменным верхним пределом, получаем:

$$\frac{\partial^k F_{\xi}(x_1, \dots, x_k)}{\partial x_1 \dots \partial x_k} \stackrel{\text{п.в.}}{=} f_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k). \quad (4.4)$$

Свойства многомерной функции распределения изучались в § 5 главы 4. Там же было показано, что имеется взаимно однозначное соответствие между вероятностными мерами и функциями распределения в конечномерных пространствах.

Установим еще одно свойство, верное для всех многомерных функций распределения. Пусть задана функция распределения $F_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k)$. Выберем совокупность аргументов x_{l+1}, \dots, x_k и устремим каждый из них к бесконечности, тогда имеет место сходимостъ:

$$F_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k) \xrightarrow{x_{l+1} \rightarrow \infty, \dots, x_k \rightarrow \infty} F_{(\xi_1, \dots, \xi_l)}(x_1, \dots, x_l),$$

где

$$\begin{aligned} F_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k) &= P\{\xi_1 \leq x_1, \xi_2 \leq x_2, \dots, \xi_k \leq x_k\} = \\ &= P\{\xi_1^{-1}(-\infty, x_1] \cap \xi_2^{-1}(-\infty, x_2] \cap \dots \cap \xi_k^{-1}(-\infty, x_k]\}. \end{aligned}$$

Доказательство. Выберем произвольные числовые последовательности: $x_{l+1}(n) \rightarrow \infty, \dots, x_k(n) \rightarrow \infty$ и подставим их вместо x_{l+1}, \dots, x_k в выражение для функции распределения, тогда под знаком вероятности получим возрастающие последовательности: $P_{\xi}\{(-\infty, x_1] \times \dots \times (-\infty, x_{l+1}(n)] \times \dots \times (-\infty, x_k(n)]\}$.

Обозначим через $A_n = (-\infty, x_1] \times \dots \times (-\infty, x_l] \times (-\infty, x_{l+1}(n)] \times \dots \times (-\infty, x_k(n)]$. Очевидно, что $A_n \subset A_{n+1}$. Тогда в силу непрерывности вероятностной меры: $P_{\xi}(A_n) \rightarrow P_{\xi}((-\infty, x_1] \times \dots \times (-\infty, x_l] \times \mathbb{R}^{n-l}) = F_{\xi_1 \dots \xi_l}(x_1, \dots, x_l)$. ■

Замечание 4.1. Утверждение справедливо для любого набора аргументов $x_{i_1}, x_{i_2}, \dots, x_{i_s}$. Устремив выбранные аргументы к бесконечности, получим функцию распределения подвектора, зависящую от оставшихся аргументов.

Следствие 4.1. *Имеет место сходимость:*

$$F_{(\xi_1, \dots, \xi_k)}(x_1, \dots, x_k) \xrightarrow{x_1 \rightarrow \infty, \dots, x_{i-1} \rightarrow \infty, x_{i+1} \rightarrow \infty, \dots, x_k \rightarrow \infty} F_{\xi_i}(x_i),$$

где $F_{\xi_i}(x_i)$ — маргинальная функция распределения (одномерная функция распределения случайной величины ξ_i , $i = 1, \dots, k$).

Справедливо еще одно свойство многомерной плотности распределения. Пусть $f_{\xi_1, \dots, \xi_k}(x_1, \dots, x_k)$ — совместная плотность распределения случайных величин ξ_1, \dots, ξ_k . Тогда имеет место равенство:

$$\begin{aligned} f_{\xi_1, \dots, \xi_l}(x_1, \dots, x_l) &= \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\xi_1, \dots, \xi_l, \xi_{l+1}, \dots, \xi_k}(x_1, \dots, x_l, x_{l+1}, \dots, x_k) dx_{l+1} \dots dx_k. \end{aligned}$$

Доказательство. Рассмотрим произвольное борелевское множество $B \in \mathcal{B}(\mathbb{R}^l)$. Перейдя к повторному интегрированию, найдем вероятность:

$$\begin{aligned} P\{(\xi_1, \dots, \xi_l) \in B\} &= P\{(\xi_1, \dots, \xi_l, \xi_{l+1}, \dots, \xi_k) \in B \times \mathbb{R}^{k-l}\} = \\ &= \int_B \dots \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f_{\xi_1, \dots, \xi_k}(x_1, \dots, x_l, x_{l+1}, \dots, x_k) dx_1 \dots dx_k = \\ &= \int_B \dots \int \left[\int_{\mathbb{R}} \dots \int_{\mathbb{R}} f_{\xi}(x_1, \dots, x_{l+1}, \dots, x_k) dx_{l+1} \dots dx_k \right] \times \\ &\quad \times dx_1 \dots dx_l, \end{aligned}$$

функция в квадратных скобках неотрицательна и удовлетворяет определению плотности. Так как равенство верно для любого $B \in \mathcal{B}(\mathbb{R}^l)$, то в соответствии с определением $\int_{\mathbb{R}} \dots \int_{\mathbb{R}} f_{\xi}(x_1, \dots, x_l, x_{l+1}, \dots, x_k) dx_{l+1} \dots dx_k$ представляет собой плотность. ■

Замечание 4.2. Утверждение сохраняется для любого набора аргументов $x_{i_1}, x_{i_2}, \dots, x_{i_s}$. Проинтегрировав по выбранным аргументам, получим плотность распределения подвектора.

Замечание 4.3. Указанным способом можно получить плотность распределения любой компоненты вектора ξ , интегрируя по всем «ненужным» компонентам.

Пример 4.1. Плотность стандартного нормального k -мерного распределения $N(0, \mathbb{E}_k)$ имеет вид:

$$f_{\xi}(x_1, \dots, x_k) = \frac{1}{(2\pi)^{\frac{k}{2}}} e^{-\frac{1}{2} \sum_{i=1}^k x_i^2}.$$

Пример 4.2. Плотность нормального k -мерного распределения $N(a, \Sigma)$ имеет вид:

$$f_{\xi}(x_1, \dots, x_k) = \frac{1}{(2\pi)^{\frac{k}{2}} \sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-a)^T \Sigma^{-1}(x-a)},$$

где матрица $\Sigma > 0$ положительно определена и вектор $a = (a_1, \dots, a_k)^T \in \mathbb{R}^k$.

§ 5. НЕЗАВИСИМЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) задан случайный вектор $\xi(\omega) = (\xi_1(\omega), \dots, \xi_k(\omega))$.

Определение 5.1. Случайные величины ξ_1, \dots, ξ_k взаимно независимы, если для любых борелевских множеств $B_1, \dots, B_k \in \mathcal{B}(\mathbb{R})$ выполнено следующее равенство:

$$P\{\xi_1 \in B_1, \dots, \xi_k \in B_k\} = P\{\xi_1 \in B_1\} \dots P\{\xi_k \in B_k\} \quad (5.1)$$

или

$$P_{(\xi_1, \dots, \xi_k)}(B_1 \times \dots \times B_k) = P_{\xi_1}(B_1) \cdot \dots \cdot P_{\xi_k}(B_k). \quad (5.2)$$

Замечание 5.1. Данное определение эквивалентно следующему: случайные величины ξ_1, \dots, ξ_k взаимно независимы, если взаимно независимы порожденные ими σ -алгебры.

Теорема 5.1. *Случайные величины ξ_1, \dots, ξ_k взаимно независимы тогда и только тогда, когда совместная функция распределения удовлетворяет условию:*

$$F_{\xi_1, \dots, \xi_k}(x_1, \dots, x_k) = F_{\xi_1}(x_1) \cdot \dots \cdot F_{\xi_k}(x_k),$$

при любом $x = (x_1, \dots, x_k)$.

Доказательство. *Необходимость* очевидна.

Достаточность. К обеим частям равенства применим разностный оператор $\Delta_{a_1 b_1}^{(1)} \Delta_{a_2 b_2}^{(2)} \dots \Delta_{a_k b_k}^{(k)}$:

$$\begin{aligned} P\{\xi_1 \in (a_1, b_1], \xi_2 \in (a_2, b_2], \dots, \xi_k \in (a_k, b_k]\} = \\ = P\{\xi_1 \in (a_1, b_1]\} P\{\xi_2 \in (a_2, b_2]\} \dots P\{\xi_k \in (a_k, b_k]\}. \end{aligned}$$

В силу непрерывности вероятностных мер получаем:

$$P\{\xi_1 \in I_1, \dots, \xi_k \in I_k\} = P\{\xi_1 \in I_1\} \cdot \dots \cdot P\{\xi_k \in I_k\},$$

где I_1, \dots, I_k — произвольные множества из полуалгебры $\mathcal{I}_1 = \{(-\infty, +\infty), (-\infty, a], (b, +\infty), (a, b]\}$.

Зафиксируем множества I_2, \dots, I_k , получим две конечные меры, заданные на пространстве $(\mathbb{R}_1, \mathcal{B}(\mathbb{R}_1))$. Они совпадают на полуалгебре, а следовательно, они совпадают на σ -алгебре, порожденной этой полуалгеброй.

$$\begin{aligned} P\{\xi_1 \in B_1, \xi_2 \in I_2, \dots, \xi_k \in I_k\} = \\ = P\{\xi_1 \in B_1\} P\{\xi_2 \in I_2\} \dots P\{\xi_k \in I_k\}, \end{aligned}$$

где B_1 — произвольное борелевское множество. Далее повторяем рассуждение и приходим к формуле (5.1). ■

Предположим, что распределение случайного вектора ξ абсолютно непрерывно относительно меры Лебега, т.е. существует совместная плотность распределения $f_\xi(x_1, \dots, x_k)$ случайного вектора ξ , тогда существует плотность $f_{\xi_i}(x_i)$ у каждой случайной величины ξ_i , $i = 1, \dots, k$, причем

$$f_{\xi_i}(x_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_\xi(x_1, \dots, x_k) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_k.$$

Следствие 5.1. Если существует совместная плотность распределения, случайные величины ξ_1, \dots, ξ_k взаимно независимы тогда и только тогда, когда совместная плотность представляется в виде:

$$f_{\xi_1, \dots, \xi_k}(x_1, \dots, x_k) \stackrel{n.б.}{=} f_{\xi_1}(x_1) \cdot \dots \cdot f_{\xi_k}(x_k).$$

Следствие 5.2. Пусть все случайные величины дискретны: $\xi_1 \in \{z_1^{(i_1)}\}, \dots, \xi_k \in \{z_k^{(i_k)}\}$. Тогда $\xi = (\xi_1, \dots, \xi_k) \in \{(z_1^{(i_1)}, \dots, z_k^{(i_k)})\}$. Необходимым и достаточным условием независимости случайных величин ξ_1, \dots, ξ_k является равенство:

$$P\{\xi_1 = z_1^{(i_1)}, \dots, \xi_k = z_k^{(i_k)}\} = P\{\xi_1 = z_1^{(i_1)}\} \dots \{ \xi_k = z_k^{(i_k)} \}.$$

Рассмотрим суммирование двух взаимно независимых случайных величин с абсолютно непрерывными распределениями. Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) заданы случайные величины $\xi(\omega)$ с плотностью распределения $f_\xi(x)$ и $\eta(\omega)$ с плотностью распределения $f_\eta(y)$. Пусть случайные величины ξ и η взаимно независимы, тогда совместная плотность распределения имеет вид $f_{\xi\eta}(x, y) = f_\xi(x)f_\eta(y)$.

Рассмотрим случайную величину $\zeta = \xi + \eta$ и найдем ее плотность распределения. Для этого достаточно найти

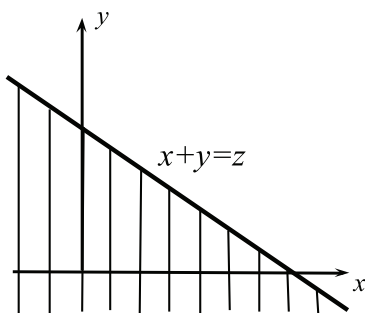


Рис. 5.1

Область интегрирования

$F_{\zeta}(z) = P\{\zeta \leq z\}$ (рис. 5.1):

$$\begin{aligned}
 F_{\zeta}(z) &= P\{\zeta \leq z\} = \\
 &= P\{\xi + \eta \leq z\} = P\{(\xi, \eta) \in \{(x, y) : x + y \leq z\}\} = \\
 &= \iint_{\{(x, y) : x + y \leq z\}} f_{\xi}(x) f_{\eta}(y) dx dy = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{z-x} f_{\xi}(x) f_{\eta}(y) dy \right) dx = \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^z f_{\xi}(x) f_{\eta}(u - x) du dx = \\
 &= \int_{-\infty}^z \left(\int_{-\infty}^{+\infty} f_{\xi}(x) f_{\eta}(u - x) dx \right) du = \int_{-\infty}^z f_{\zeta}(u) du.
 \end{aligned}$$

Таким образом, получили формулу свертки. В силу симметричности справедливо равенство:

$$f_{\zeta}(u) = \int_{-\infty}^{+\infty} f_{\xi}(x) f_{\eta}(u - x) dx = \int_{-\infty}^{+\infty} f_{\eta}(y) f_{\xi}(u - y) dy.$$

МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ КАК ИНТЕГРАЛ ЛЕБЕГА

§ 1. ОПРЕДЕЛЕНИЕ И СВОЙСТВА МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

Пусть задано вероятностное пространство (Ω, \mathcal{F}, P) и числовая функция $\xi(\omega)$. Могут существовать такие элементарные события ω , что $\xi(\omega) = +\infty$ или $\xi(\omega) = -\infty$. Поэтому имеет смысл рассматривать расширенную числовую прямую $\mathbb{R} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

Пусть для любого борелевского множества $B \in \mathcal{B}(\mathbb{R})$ выполняется $\xi^{-1}(B) \in \mathcal{F}$, $\xi^{-1}(-\infty) \in \mathcal{F}$, $\xi^{-1}(+\infty) \in \mathcal{F}$. Такие \mathcal{F} -измеримые функции будем называть расширенными случайными величинами. Ниже перечислим некоторые соотношения, которые будем предполагать справедливыми:

- $\frac{a}{\pm\infty} = 0$, $a \in \mathbb{R}$;
- $a \cdot \infty = \infty$, $a > 0$;
- $a \cdot \infty = -\infty$, $a < 0$;
- $0 \cdot \infty = 0$.

Определение 1.1. Математическое ожидание расширенной случайной величины ξ представляет собой интеграл Лебега:

$$E\xi = \int_{\Omega} \xi(\omega) dP, \quad (1.1)$$

если этот интеграл существует. Последовательно рассмотрим следующие типы случайных величин:

1. Рассмотрим простую неотрицательную случайную величину

$$0 \leq \xi(\omega) = \sum_{i=1}^m a_i I\{\omega \in A_i\},$$

где $a_i \in \mathbb{R}$, $a_i \geq 0$, $A_i \in \mathcal{F}$, $\sum_{i=1}^m A_i = \Omega$, $A_i \cap A_j = \emptyset$ для всех $i \neq j$.

Полагаем

$$E\xi = \int_{\Omega} \xi(\omega) dP = \sum_{i=1}^m a_i P(A_i).$$

2. Рассмотрим неотрицательную расширенную случайную величину $\xi(\omega) \geq 0$. В этом случае существует последовательность простых случайных величин $\xi_n(\omega) \geq 0$ таких, что в каждой точке множества Ω случайные величины $\xi_n(\omega)$ монотонно сходятся к случайной величине $\xi(\omega)$, то есть $\lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)$, $\xi_n(\omega) \leq \xi_{n+1}(\omega)$ для любого ω . Тогда

$$E\xi = \int_{\Omega} \xi(\omega) dP = \lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP.$$

3. Рассмотрим любую \mathcal{F} -измеримую функцию $\xi(\omega)$ (расширенную случайную величину). Ее можно представить единственным образом в виде разности двух неотрицательных функций:

$$\xi(\omega) = \xi^+(\omega) - \xi^-(\omega),$$

где

$$\xi^+(\omega) = \xi(\omega) I\{\omega : \xi(\omega) \geq 0\} = \begin{cases} \xi(\omega), & \xi(\omega) \geq 0 \\ 0, & \xi(\omega) < 0 \end{cases},$$

$$\xi^-(\omega) = -\xi(\omega) I\{\omega : \xi(\omega) < 0\} = \begin{cases} -\xi(\omega), & \xi(\omega) < 0 \\ 0, & \xi(\omega) \geq 0 \end{cases}.$$

Обе функции ξ^+ и ξ^- являются неотрицательными. В каждой точке ω выполняется: $\xi^+ \xi^- = 0$. Положим

$$E\xi = \int_{\Omega} \xi(\omega) dP = \int_{\Omega} \xi^+(\omega) dP - \int_{\Omega} \xi^-(\omega) dP.$$

Из определения 1.1 следует, что интеграл Лебега может не существовать только в одном случае, когда в части 3 определения 1.1 получаем $\int_{\Omega} \xi^+(\omega) dP = \int_{\Omega} \xi^-(\omega) dP = \infty$, то есть, возникает неопределенность вида $\infty - \infty$.

Замечание 1.1. Интеграл Лебега может принимать значения ∞ и $-\infty$. В этом случае считается, что интеграл определен (существует), но не конечен.

Свойства математического ожидания

I. Свойство линейности

1. Пусть существует $\int_{\Omega} \xi(\omega) dP$ и $c \in \mathbb{R}$, тогда существует

$$\int_{\Omega} c\xi(\omega) dP = c \int_{\Omega} \xi(\omega) dP.$$

2. Пусть существуют $\int_{\Omega} \xi(\omega) dP$, $\int_{\Omega} \eta(\omega) dP$, $\int_{\Omega} \xi(\omega) dP + \int_{\Omega} \eta(\omega) dP$, тогда существует

$$\int_{\Omega} (\xi(\omega) + \eta(\omega)) dP = \int_{\Omega} \xi(\omega) dP + \int_{\Omega} \eta(\omega) dP.$$

II. Свойство положительности

1. Пусть $\xi(\omega) \geq 0$, тогда $\int_{\Omega} \xi(\omega) dP \geq 0$.
2. Пусть $\xi(\omega) \leq \eta(\omega)$, и существует $\int_{\Omega} \xi(\omega) dP > -\infty$, тогда существует $\int_{\Omega} \eta(\omega) dP \geq \int_{\Omega} \xi(\omega) dP$.
3. Пусть $\xi(\omega) \leq \eta(\omega)$, и существует $\int_{\Omega} \eta(\omega) dP < \infty$, тогда существует $\int_{\Omega} \xi(\omega) dP \leq \int_{\Omega} \eta(\omega) dP$.

Докажем, что введенное ранее определение 1.1 математического ожидания корректно. Рассмотрим сначала первую часть определения. В этом случае могут быть разные формы записи $\xi(\omega)$:

$$\xi(\omega) = \sum_{i=1}^m a_i I\{\omega \in A_i\} = \sum_{j=1}^n b_j I\{\omega \in B_j\},$$

где $B_j \in \mathcal{F}$, $B_j \cap B_k = \emptyset$, $j \neq k$, $\bigcup_{j=1}^n B_j = \Omega$. Необходимо показать, что $\sum_{i=1}^m a_i P(A_i) = \sum_{j=1}^n b_j P(B_j)$. Всякое A_i представимо в виде $A_i = \bigcup_{j=1}^n (A_i \cap B_j)$, и аналогично,

$B_j = \bigcup_{i=1}^m (A_i \cap B_j)$. Тогда имеет место равенство:

$$\begin{aligned} \sum_{i=1}^m a_i P(A_i) &= \sum_{i=1}^m \sum_{j=1}^n a_i P(A_i \cap B_j) = \\ &= \sum_{j=1}^n \sum_{i=1}^m b_j P(A_i \cap B_j) = \sum_{j=1}^n b_j P(B_j). \end{aligned}$$

Если $A_i \cap B_j = \emptyset$, то $P(A_i \cap B_j) = 0$. Если $A_i \cap B_j \neq \emptyset$, то обязательно $a_i = b_j$. Таким образом, первая часть определения 1.1 математического ожидания корректна.

Проверим выполнение свойств линейности и положительности для первой части определения 1.1 математического ожидания. Докажем свойство линейности:

1. Если $\xi(\omega) = \sum_{i=1}^m a_i I\{\omega \in A_i\}$, то имеет место равенство:

$$c\xi(\omega) = \sum_{i=1}^m ca_i I\{\omega \in A_i\}.$$

Следовательно, получаем:

$$\int_{\Omega} c\xi(\omega) dP = \sum_{i=1}^m ca_i P(A_i) = c \sum_{i=1}^m a_i P(A_i) = c \int_{\Omega} \xi(\omega) dP.$$

2. Введем случайную величину $\eta(\omega) = \sum_{j=1}^n d_j I\{\omega \in D_j\}$. Можем записать представления:

$$A_i = \bigcup_{j=1}^n (A_i \cap D_j), \quad D_j = \bigcup_{i=1}^m (A_i \cap D_j).$$

Тогда

$$\xi(\omega) = \sum_{i=1}^m \sum_{j=1}^n a_i I\{\omega \in A_i \cap D_j\},$$

$$\eta(\omega) = \sum_{j=1}^n \sum_{i=1}^m d_j I\{\omega \in A_i \cap D_j\}.$$

Рассмотрим сумму

$$\xi(\omega) + \eta(\omega) = \sum_{i=1}^m \sum_{j=1}^n (a_i + d_j) I\{\omega \in A_i \cap D_j\} \quad (1.2)$$

и найдем ее математическое ожидание:

$$\int_{\Omega} (\xi(\omega) + \eta(\omega)) dP = \sum_{i=1}^m \sum_{j=1}^n (a_i + d_j) P(A_i \cap D_j).$$

Сумма простых функций также является простой функцией. Учитывая представления (1.2), можем продолжить равенство:

$$\begin{aligned} & \sum_{i=1}^m \sum_{j=1}^n (a_i + d_j) P(A_i \cap D_j) = \\ &= \sum_{i=1}^m a_i P(A_i) + \sum_{j=1}^n d_j P(D_j) = \int_{\Omega} \xi(\omega) dP + \int_{\Omega} \eta(\omega) dP. \end{aligned}$$

Свойство линейности выполняется для первой части определения 1.1 математического ожидания.

Докажем свойство положительности:

1. Так как $\xi(\omega)$ — простая случайная величина, $\xi(\omega) = \sum_{i=1}^m a_i I\{\omega \in A_i\} \geq 0$, то все коэффициенты $a_i \geq 0$. Следовательно,

$$\int_{\Omega} \xi(\omega) dP = \sum_{i=1}^m a_i P(A_i) \geq 0,$$

так как все слагаемые неотрицательные.

2. Так как $\xi(\omega) \leq \eta(\omega)$, то $\eta(\omega) - \xi(\omega) \geq 0$. Рассмотрим выражение $(\eta(\omega) - \xi(\omega)) + \xi(\omega) = \eta(\omega)$, где $\eta(\omega) - \xi(\omega) \geq 0$, $\xi(\omega) \geq 0$, $\eta(\omega) \geq 0$. Из доказательства линейности следует, что

$$\int_{\Omega} \eta(\omega) dP = \int_{\Omega} (\eta(\omega) - \xi(\omega)) dP + \int_{\Omega} \xi(\omega) dP. \quad (1.3)$$

Опираясь на доказательство предыдущего пункта, делаем вывод, что $\int_{\Omega} (\eta(\omega) - \xi(\omega)) dP \geq 0$. Если отбросить это неотрицательное слагаемое из равенства (1.3), то правая часть равенства может только уменьшиться. Свойство положительности выполняется для первой части определения 1.1 математического ожидания.

Теорема 1.1. *Справедливы свойства:*

1. Пусть $\xi(\omega)$ — произвольная \mathcal{F} -измеримая функция, принимающая значения на расширенной числовой прямой. Тогда существует последовательность простых функций $\xi_n(\omega)$ такая, что $\lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)$ и $|\xi_n(\omega)| \leq |\xi(\omega)|$ для всех n .
2. Пусть $\xi(\omega) \geq 0$ — \mathcal{F} -измеримая функция со значениями в расширенной числовой прямой. Тогда существует последовательность простых функций, что имеет место сходимости:

$$0 \leq \xi_n(\omega) \leq \xi_{n+1}(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega).$$

Доказательство. Сначала докажем второе утверждение теоремы. Запишем последовательность $\xi_n(\omega)$ явным образом. Для этого разделим каждый интервал на 2^n подинтервалов:

$$\xi_n(\omega) = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \cdot I \left\{ \frac{k}{2^n} \leq \xi(\omega) < \frac{k+1}{2^n} \right\} + nI\{\xi(\omega) \geq n\}.$$

Проверим, что для любых ω и n справедливо неравенство:

$$0 \leq \xi_n(\omega) \leq \xi_{n+1}(\omega) \leq \xi(\omega).$$

Выбирая $n > \xi(\omega)$, получаем: $0 \leq \xi(\omega) - \xi_n(\omega) \leq \frac{1}{2^n}$. Если $\xi(\omega_0) = \infty$, то по построению $\xi_n(\omega_0) = n \rightarrow \infty$.

Докажем первое утверждение теоремы. Представим $\xi(\omega)$ единственным образом в виде разности двух положительных функций:

$$\xi(\omega) = \xi^+(\omega) - \xi^-(\omega),$$

где $\xi^+(\omega) = \xi(\omega)I\{\xi(\omega) \geq 0\}$, $\xi^-(\omega) = -\xi(\omega)I\{\xi(\omega) < 0\}$.

Таким образом, $|\xi(\omega)| = \xi^+(\omega) + \xi^-(\omega)$. Применим второе утверждение теоремы для каждой из частей функции $\xi(\omega)$, получаем:

$$0 \leq \xi_n^+(\omega) \xrightarrow{n \rightarrow \infty} \xi^+(\omega),$$

$$0 \leq \xi_n^-(\omega) \xrightarrow{n \rightarrow \infty} \xi^-(\omega).$$

Таким образом, имеет место сходимость:

$$\xi_n(\omega) = \xi_n^+(\omega) - \xi_n^-(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega).$$

При этом $|\xi_n(\omega)| \leq \xi_n^+(\omega) + \xi_n^-(\omega) \xrightarrow{n \rightarrow \infty} \xi^+(\omega) + \xi^-(\omega)$, и имеют место неравенства: $\xi_n^+(\omega) < \xi^+(\omega)$, $\xi_n^-(\omega) < \xi^-(\omega)$, тогда справедливо неравенство:

$$|\xi_n(\omega)| \leq \xi_n^+(\omega) + \xi_n^-(\omega) \leq \xi^+(\omega) + \xi^-(\omega) = |\xi(\omega)|.$$

Теорема доказана. ■

Замечание 1.2. Данная теорема доказывает существование последовательности, которая рассматривается во второй части определения математического ожидания. Однако может существовать много последовательностей, обладающих такими свойствами. Надо доказать, что предел математических ожиданий у всех последовательностей одинаков.

Лемма 1.1. Пусть последовательность неотрицательных простых функций $\xi_n(\omega)$ такова, что $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, $0 \leq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$ и $0 \leq \eta(\omega) \leq \xi(\omega)$, где $\eta(\omega)$ — простая функция. Тогда справедливо следующее неравенство:

$$\lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP \geq \int_{\Omega} \eta(\omega) dP.$$

Доказательство. Рассмотрим некоторое $\varepsilon > 0$, введем событие:

$$A_n^\varepsilon = \{\omega : \xi_n(\omega) \geq \eta(\omega) - \varepsilon\}.$$

Последовательность $A_n^\varepsilon \subset A_{n+1}^\varepsilon$ — возрастающая последовательность событий, и $\bigcup_{n=1}^\infty A_n^\varepsilon = \Omega$. Следовательно,

$$\bar{A}_n^\varepsilon \supset \bar{A}_{n+1}^\varepsilon, \quad \bigcap_{n=1}^\infty \bar{A}_n^\varepsilon = \emptyset. \quad (1.4)$$

Очевидно, что выполнено неравенство:

$$\begin{aligned} \xi_n(\omega) &\geq \xi_n(\omega) I\{\omega \in A_n^\varepsilon\} \geq (\eta(\omega) - \varepsilon) I\{\omega \in A_n^\varepsilon\} = \\ &= \eta(\omega) - \eta(\omega) I\{\omega \in \bar{A}_n^\varepsilon\} - \varepsilon I\{\omega \in A_n^\varepsilon\} \geq \\ &\geq \eta(\omega) - C I\{\omega \in \bar{A}_n^\varepsilon\} - \varepsilon, \end{aligned} \quad (1.5)$$

где $C = \max_{\omega \in \Omega} \eta(\omega)$.

Слева и справа в неравенстве (1.5) — простые функции, для них свойство положительности уже установлено. Таким образом, получаем неравенство:

$$\int_{\Omega} \xi_n(\omega) dP \geq \int_{\Omega} \eta(\omega) dP - CP(\bar{A}_n^\varepsilon) - \varepsilon.$$

В силу (1.4) получаем сходимость: $CP(\bar{A}_n^\varepsilon) \xrightarrow{n \rightarrow \infty} 0$. Перейдем к пределу:

$$\lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP \geq \int_{\Omega} \eta(\omega) dP - \varepsilon.$$

Неравенство справедливо при любом ε , а, значит, справедливо и утверждение леммы. ■

Рассмотрим следующие последовательности простых функций: $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, $0 \leq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$, $\eta_m(\omega) \leq \eta_{m+1}(\omega)$, $0 \leq \eta_m(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$.

Возьмем конкретное m и применим лемму 1.1:

$$\lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP \geq \int_{\Omega} \eta_m(\omega) dP.$$

Также справедливо неравенство:

$$\lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP \geq \lim_{m \rightarrow \infty} \int_{\Omega} \eta_m(\omega) dP.$$

Данные рассуждения применимы для функций $\eta_n(\omega)$, $\xi_m(\omega)$, рассматриваемых в обратном порядке. Таким образом,

$$\lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP = \lim_{m \rightarrow \infty} \int_{\Omega} \eta_m(\omega) dP.$$

Следовательно, вторая часть определения 1.1 математического ожидания корректна.

Проверим выполнение свойств линейности и положительности для второй части определения 1.1 математического ожидания. Докажем сначала свойство линейности:

1. Рассмотрим $c\xi(\omega)$, где $c \geq 0$, $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, $0 \leq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$. Тогда $c\xi_n(\omega) \leq c\xi_{n+1}(\omega)$, $0 \leq c\xi_n(\omega) \xrightarrow{n \rightarrow \infty} c\xi(\omega)$ и

$$\begin{aligned} \int_{\Omega} c\xi(\omega) dP &= \lim_{n \rightarrow \infty} \int_{\Omega} c\xi_n(\omega) dP = \\ &= c \lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP = c \int_{\Omega} \xi(\omega) dP. \end{aligned}$$

2. Рассмотрим последовательности простых функций $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, $0 \leq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$, $\eta_n(\omega) \leq \eta_{n+1}(\omega)$, $0 \leq \eta_n(\omega) \xrightarrow{n \rightarrow \infty} \eta(\omega)$. Для них справедливо неравенство:

$$\xi_n(\omega) + \eta_n(\omega) \leq \xi_{n+1}(\omega) + \eta_{n+1}(\omega),$$

$$0 \leq \xi_n(\omega) + \eta_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega) + \eta(\omega).$$

Тогда имеют место равенства:

$$\begin{aligned} \int_{\Omega} (\xi(\omega) + \eta(\omega)) dP &= \lim_{n \rightarrow \infty} \int_{\Omega} (\xi_n(\omega) + \eta_n(\omega)) dP = \\ &= \lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP + \lim_{n \rightarrow \infty} \int_{\Omega} \eta_n(\omega) dP = \\ &= \int_{\Omega} \xi(\omega) dP + \int_{\Omega} \eta(\omega) dP. \end{aligned}$$

Докажем свойство положительности:

1. Пункт 1 свойства положительности очевидно выполняется.
2. Рассмотрим п. 2, 3 свойства положительности. Если $\xi(\omega) \leq \eta(\omega)$, то $\eta(\omega) = (\eta(\omega) - \xi(\omega)) + \xi(\omega)$. В данном равенстве $\eta(\omega) - \xi(\omega) \geq 0$, $\xi(\omega) \geq 0$. Воспользуемся свойством аддитивности:

$$\int_{\Omega} \eta(\omega) dP = \int_{\Omega} (\eta(\omega) - \xi(\omega)) dP + \int_{\Omega} \xi(\omega) dP.$$

Здесь $\int_{\Omega} (\eta(\omega) - \xi(\omega)) dP \geq 0$. Если отбросить это слагаемое, то значение правой части равенства может только уменьшиться.

Рассмотрим третью часть определения 1.1 математического ожидания. Корректность третьей части не вызывает сомнений. Необходимым и достаточным условием существования интеграла Лебега является:

$$\min \left(\int_{\Omega} \xi^+(\omega) dP, \int_{\Omega} \xi^-(\omega) dP \right) \in \mathbb{R}.$$

Проверка справедливости свойств линейности и положительности для третьей части определения 1.1 математического ожидания выполняется аналогично приведенным выше доказательствам.

III. Свойство конечности

Неравенство $|\int_{\Omega} \xi(\omega) dP| < \infty$ имеет место тогда и только тогда, когда $\int_{\Omega} |\xi(\omega)| dP < \infty$.

Доказательство. Согласно третьей части определения 1.1 математического ожидания:

$$\int_{\Omega} \xi(\omega) dP = \int_{\Omega} \xi^+(\omega) dP - \int_{\Omega} \xi^-(\omega) dP. \quad (1.6)$$

В силу аддитивности имеет место равенство:

$$\int_{\Omega} |\xi(\omega)| dP = \int_{\Omega} \xi^+(\omega) dP + \int_{\Omega} \xi^-(\omega) dP. \quad (1.7)$$

Интегралы (1.6) и (1.7) конечны тогда и только тогда, когда $\int_{\Omega} \xi^+(\omega) dP < \infty$ и $\int_{\Omega} \xi^-(\omega) dP < \infty$. ■

IV. Справедливо неравенство:

$$\left| \int_{\Omega} \xi(\omega) dP \right| \leq \int_{\Omega} |\xi(\omega)| dP. \quad (1.8)$$

Доказательство. Очевидно, что имеет место неравенство:

$$-|\xi(\omega)| \leq \xi(\omega) \leq |\xi(\omega)|.$$

Если $\int_{\Omega} \xi(\omega) dP = \infty$ или $\int_{\Omega} \xi(\omega) dP = -\infty$, то $\int_{\Omega} |\xi(\omega)| dP = \infty$, и неравенство (1.8) выполняется как равенство. Если $|\int_{\Omega} \xi(\omega) dP| < \infty$, то согласно свойству II математического ожидания справедливы неравенство:

$$\int_{\Omega} \xi(\omega) dP \leq \int_{\Omega} |\xi(\omega)| dP, \quad (1.9)$$

$$-\int_{\Omega} |\xi(\omega)| dP \leq \int_{\Omega} \xi(\omega) dP. \quad (1.10)$$

Умножим неравенство (1.10) на -1 , получим:

$$\int_{\Omega} |\xi(\omega)| dP \geq -\int_{\Omega} \xi(\omega) dP. \quad (1.11)$$

Из (1.9) и (1.11) следует:

$$\left| \int_{\Omega} \xi(\omega) dP \right| \leq \int_{\Omega} |\xi(\omega)| dP.$$

■

V. Свойство мультипликативности

1. Пусть случайные величины $\xi(\omega) \geq 0$ и $\eta(\omega) \geq 0$ независимы, тогда имеет место следующее равенство:

$$E(\xi\eta) = E\xi E\eta.$$

2. Пусть случайные величины $\xi(\omega)$ и $\eta(\omega)$ независимы. Пусть у них существуют и конечны математические ожидания $E\xi \in \mathbb{R}$, $E\eta \in \mathbb{R}$, тогда имеет место равенство:

$$E(\xi\eta) = E\xi E\eta.$$

Доказательство. Оба свойства докажем одновременно. Пусть $\xi(\omega) \geq 0$, $\eta(\omega) \geq 0$ — взаимно независимые простые функции, тогда их можно представить в виде:

$$\begin{aligned} \xi(\omega) &= \sum_{i=1}^n a_i I\{\omega \in A_i\}, \\ \eta(\omega) &= \sum_{j=1}^m b_j I\{\omega \in B_j\}, \end{aligned}$$

причем

$$\begin{aligned} A_k \cap A_l &= \emptyset, \quad k \neq l, \quad \bigcup_{k=1}^n A_k = \Omega; \\ B_r \cap B_s &= \emptyset, \quad r \neq s, \quad \bigcup_{r=1}^m B_r = \Omega. \end{aligned}$$

Без ограничения общности можно считать, что $a_k \neq a_l$ для любых k, l и $b_r \neq b_s$ для любых r, s . Очевидно, что $\xi(\omega)\eta(\omega) = \sum_{i=1}^n \sum_{j=1}^m b_j a_i I\{\omega \in A_i \cap B_j\}$ — простая случайная величина.

Тогда

$$E(\xi\eta) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j P(A_i \cap B_j),$$

но так как $a_k \neq a_l$ для всех k, l и $b_r \neq b_s$ для всех r, s , то $A_i = \xi^{-1}(a_i)$, $B_j = \eta^{-1}(b_j)$. Тогда

$$P(\xi^{-1}(a_i) \cap \eta^{-1}(b_j)) = P(A_i)P(B_j).$$

Таким образом, получаем равенство:

$$\begin{aligned} E(\xi\eta) &= \sum_{i=1}^n \sum_{j=1}^m a_i P(A_i) b_j P(B_j) = \\ &= \sum_{i=1}^n a_i P(A_i) \sum_{j=1}^m b_j P(B_j) = E\xi E\eta. \end{aligned}$$

Покажем теперь справедливость пункта 1 свойства V. Пусть $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, $0 \leq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$, $\eta_n(\omega) \leq \eta_{n+1}(\omega)$, $0 \leq \eta_n(\omega) \xrightarrow{n \rightarrow \infty} \eta(\omega)$ — последовательности простых функций из теоремы 1.1. Следовательно, $\xi_n(\omega)\eta_n(\omega) \leq \xi_{n+1}(\omega)\eta_{n+1}(\omega)$, $0 \leq \xi_n(\omega)\eta_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)\eta(\omega)$, где $\xi_n(\omega)\eta_n(\omega)$ — простая измеримая функция. Тогда по определению интеграла Лебега:

$$E(\xi\eta) = \lim_{n \rightarrow \infty} E(\xi_n\eta_n).$$

Очевидно, $\sigma_{\xi_n} \subset \sigma_{\xi}$, $\sigma_{\eta_n} \subset \sigma_{\eta}$. Так как сигма-алгебры σ_{ξ} и σ_{η} независимы, то σ_{ξ_n} и σ_{η_n} независимы. Тогда, используя полученное выше равенство для простых случайных величин, получаем:

$$E(\xi\eta) = \lim_{n \rightarrow \infty} (E\xi_n E\eta_n) = \lim_{n \rightarrow \infty} E\xi_n \lim_{n \rightarrow \infty} E\eta_n = E\xi E\eta.$$

Докажем пункт 2 свойства V. Очевидно, что $E\xi \in \mathbb{R}$ тогда и только тогда, когда $E\xi^+$, $E\xi^- \in \mathbb{R}$, и $E\eta \in \mathbb{R}$ тогда и только тогда, когда $E\eta^+$, $E\eta^- \in \mathbb{R}$.

Рассмотрим $E(\xi\eta)$. Так как $\xi = \xi^+ - \xi^-$, $\eta = \eta^+ - \eta^-$, то справедливо равенство:

$$E(\xi\eta) = E(\xi^+\eta^+ + \xi^-\eta^- - \xi^-\eta^+ - \xi^+\eta^-).$$

Выполняется $\sigma_{\xi^+} \subset \sigma_{\xi}$, $\sigma_{\xi^-} \subset \sigma_{\xi}$, $\sigma_{\eta^+} \subset \sigma_{\eta}$, $\sigma_{\eta^-} \subset \sigma_{\eta}$. Таким образом, все сомножители в рассматриваемом равенстве взаимно независимы. Справедлив пункт 1 свойства V, а для выполнения свойства аддитивности нужно, чтобы интегралы были конечны:

$$\begin{aligned} E(\xi\eta) &= E(\xi^+\eta^+ + \xi^-\eta^- - \xi^-\eta^+ - \xi^+\eta^-) = \\ &= E\xi^+E\eta^+ + E\xi^-E\eta^- - E\xi^-E\eta^+ - E\xi^+E\eta^- = \\ &= (E\xi^+ - E\xi^-)(E\eta^+ - E\eta^-) = E\xi E\eta. \end{aligned}$$

■

VI. Пусть существует $\int_{\Omega} \xi(\omega) dP$. Тогда для любого $A \in \mathcal{F}$ существует $\int_{\Omega} \xi(\omega) I\{\omega \in A\} dP$.

Доказательство. Справедливы неравенства:

$$0 \leq \xi^+(\omega) I\{\omega \in A\} \leq \xi^+(\omega),$$

$$0 \leq \xi^-(\omega) I\{\omega \in A\} \leq \xi^-(\omega).$$

Из этих неравенств следует свойство VI. ■

Определение 1.2. По определению будем считать, что $\int_A \xi(\omega) dP = \int_{\Omega} \xi(\omega) I\{\omega \in A\} dP$.

Замечание 1.3. Ранее говорилось, что $0 \cdot (\pm\infty) = 0$. Следовательно, даже если $|\xi(\omega_0)| = \infty$ и $\omega_0 \notin A$, то значение подинтегральной функции окажется равным нулю во всех точках $\omega \notin A$.

VII. Пусть $\xi(\omega) \stackrel{n.н.}{=} 0$, тогда существует $\int_{\Omega} \xi(\omega) dP = 0$.

Доказательство. Проверим выполнение свойства VII для первой части определения 1.1 математического ожидания. Пусть $\xi(\omega) \geq 0$ — простая функция, т. е. $\xi(\omega) = \sum_{i=1}^n a_i I\{\omega \in A_i\}$, $A_j \cap A_i = \emptyset$, $A_i \in \mathcal{F}$, $\bigcup_{i=1}^n A_i = \Omega$, тогда $\int_{\Omega} \xi(\omega) dP = \sum_{i=1}^n a_i P(A_i)$. Если $a_i \neq 0$, то $P(A_i) = 0$. Если $P(A_i) \neq 0$, то $a_i = 0$. Следовательно, $\int_{\Omega} \xi(\omega) dP = 0$.

Проверим выполнение свойства VII для второй части определения 1.1 математического ожидания. Так как $\xi(\omega) \geq 0$, то

существует последовательность простых функций $\xi_n(\omega)$ такая, что $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, $0 \leq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$. Тогда для любого n : $0 \leq \xi_n(\omega) \leq \xi(\omega)$, для любого n и $\xi_n(\omega) \stackrel{\text{п.н.}}{=} 0$. Таким образом, $\xi_n(\omega) \stackrel{\text{п.н.}}{=} 0$ и $\int_{\Omega} \xi_n(\omega) dP = 0$. По определению 1.1 математического ожидания справедливо равенство:

$$\int_{\Omega} \xi(\omega) dP = \lim_{n \rightarrow \infty} \int_{\Omega} \xi_n(\omega) dP = 0.$$

Проверим выполнение свойства VII для третьей части определения 1.1 математического ожидания. Выполняется $\xi(\omega) = \xi^+(\omega) - \xi^-(\omega)$. Следовательно,

$$0 \leq \xi^+(\omega) \leq \xi^+(\omega) + \xi^-(\omega) = |\xi(\omega)| \stackrel{\text{п.н.}}{=} 0,$$

$$0 \leq \xi^-(\omega) \leq \xi^+(\omega) + \xi^-(\omega) = |\xi(\omega)| \stackrel{\text{п.н.}}{=} 0.$$

Тогда $\xi^+(\omega) \stackrel{\text{п.н.}}{=} 0$, $\xi^-(\omega) \stackrel{\text{п.н.}}{=} 0$. Таким образом, справедливо равенство:

$$\int_{\Omega} \xi(\omega) dP = \int_{\Omega} \xi^+(\omega) dP - \int_{\Omega} \xi^-(\omega) dP = 0.$$

■

VIII. Пусть существует конечный интеграл $\int_{\Omega} \xi(\omega) dP \in \mathbb{R}$. Тогда $P\{\omega : |\xi(\omega)| = \infty\} = 0$.

Доказательство. Рассмотрим неравенство:

$$|\xi(\omega)| \geq |\xi(\omega)| I\{|\xi(\omega)| = \infty\}.$$

Проинтегрируя неравенство, получаем:

$$\begin{aligned} \int_{\Omega} |\xi(\omega)| dP &\geq \int_{\Omega} |\xi(\omega)| I\{|\xi(\omega)| = \infty\} dP = \\ &= \lim_{n \rightarrow \infty} n P\{|\xi(\omega)| = \infty\}. \end{aligned}$$

Если бы вероятность $P\{\omega : |\xi(\omega)| = \infty\}$ была бы отлична от нуля, то предел был бы равен бесконечности, чего быть не может по условию. Таким образом, $P\{\omega : |\xi(\omega)| = \infty\} = 0$. ■

IX. Пусть $\xi(\omega) \geq 0$ и $\int_{\Omega} \xi(\omega) dP = 0$. Тогда $\xi(\omega) \stackrel{n.н.}{=} 0$.

Доказательство. Рассмотрим множества $B = \{\omega : \xi(\omega) > 0\}$, $B_n = \{\omega : \xi(\omega) \geq 1/n\}$. Очевидно, что $B_n \subset B_{n+1}$ для любого n , $B = \bigcup_{n=1}^{\infty} B_n$. Следовательно, по свойству непрерывности вероятностной меры $P(B) = \lim_{n \rightarrow \infty} P(B_n)$. Рассмотрим неравенство:

$$\xi(\omega) \geq \xi(\omega) I\{\omega \in B_n\} \geq \frac{1}{n} I\{\omega \in B_n\}.$$

Согласно свойству положительности:

$$0 = \int_{\Omega} \xi(\omega) dP \geq \frac{1}{n} P(B_n).$$

Тогда $P(B_n) = 0$, что значит, что $P(B) = 0$. ■

X. Пусть $\xi(\omega) \stackrel{n.н.}{=} \eta(\omega)$, при этом существует $\int_{\Omega} \xi(\omega) dP$, тогда существует $\int_{\Omega} \eta(\omega) dP = \int_{\Omega} \xi(\omega) dP$.

Доказательство. Рассмотрим множество

$$N = \{\omega : \xi(\omega) \neq \eta(\omega)\},$$

по условию $P(N) = 0$. Тогда имеют место равенства:

$$\begin{aligned} \int_{\Omega} \xi(\omega) dP &= \int_{\Omega} \xi(\omega) I\{\omega \in N\} dP + \int_{\Omega} \xi(\omega) I\{\omega \in \bar{N}\} dP = \\ &= \int_{\Omega} \eta(\omega) I\{\omega \in \bar{N}\} dP + \int_{\Omega} \eta(\omega) I\{\omega \in N\} dP = \int_{\Omega} \eta(\omega) dP, \end{aligned}$$

так как $\int_{\Omega} \xi(\omega) I\{\omega \in N\} dP = 0$, $\int_{\Omega} \eta(\omega) I\{\omega \in N\} dP = 0$. ■

§ 2. ПРЕДЕЛЬНЫЙ ПЕРЕХОД ПОД ЗНАКОМ ИНТЕГРАЛА ЛЕБЕГА

Теорема 2.1 (о монотонном предельном переходе под знаком интеграла Лебега). Пусть $\xi_1, \dots, \xi_n, \dots$ — последовательность неотрицательных случайных величин, заданных

на вероятностном пространстве (Ω, \mathcal{F}, P) . Пусть имеет место сходимост $0 \leq \xi_n(\omega) \xrightarrow[n \rightarrow \infty]{} \xi(\omega)$, $\xi_n(\omega) \leq \xi_{n+1}(\omega)$, для любого $\omega \in \Omega$, тогда имеет место сходимост математических ожиданий:

$$E\xi_n \xrightarrow[n \rightarrow \infty]{} E\xi.$$

Доказательство. Так как $\xi_n(\omega) \leq \xi(\omega)$ для любого ω , то в силу свойства положительности математического ожидания справедливы неравенства: $E\xi_n \leq E\xi$ и $E\xi_n \leq E\xi_{n+1}$. Тогда выполняется неравенство:

$$\lim_{n \rightarrow \infty} E\xi_n \leq E\xi.$$

Докажем обратное неравенство. Рассмотрим последовательность неотрицательных случайных величин $\xi_1(\omega)$, $\xi_2(\omega)$, \dots , $\xi_k(\omega)$, \dots и следующие последовательности простых функций:

$$\xi_{1n}(\omega) \leq \xi_{1,n+1}(\omega), \quad 0 \leq \xi_{1n}(\omega) \xrightarrow[n \rightarrow \infty]{} \xi_1(\omega),$$

$$\xi_{2n}(\omega) \leq \xi_{2,n+1}(\omega), \quad 0 \leq \xi_{2n}(\omega) \xrightarrow[n \rightarrow \infty]{} \xi_2(\omega),$$

...

$$\xi_{kn}(\omega) \leq \xi_{k,n+1}(\omega), \quad 0 \leq \xi_{kn}(\omega) \xrightarrow[n \rightarrow \infty]{} \xi_k(\omega).$$

Тогда $\zeta_n(\omega) = \max_{k=1, \dots, n} \{\xi_{kn}(\omega)\}$ — простая неотрицательная функция, $0 \leq \zeta_n(\omega) \leq \zeta_{n+1}(\omega)$. Справедливо неравенство:

$$\xi_{kn}(\omega) \leq \zeta_n(\omega) \leq \xi_n(\omega), \quad k \leq n.$$

При $n \rightarrow \infty$ имеет место неравенство:

$$\xi_k(\omega) \leq \zeta(\omega) \leq \xi(\omega),$$

где предел $\lim_{n \rightarrow \infty} \zeta_n(\omega)$ обозначен через $\zeta(\omega)$.

При $k \rightarrow \infty$ справедливо неравенство:

$$\xi(\omega) \leq \zeta(\omega) \leq \xi(\omega),$$

тогда $\zeta(\omega) = \xi(\omega)$ для любого ω .

По определению 1.1 математического ожидания $E\zeta = E\xi = \lim_{n \rightarrow \infty} E\zeta_n$. Из неравенств следует, что $E\zeta_n \leq E\xi_n$. Тогда получаем неравенство:

$$E\zeta = E\xi = \lim_{n \rightarrow \infty} E\zeta_n \leq \lim_{n \rightarrow \infty} E\xi_n.$$

Получили противоположное неравенство, следовательно, $E\xi = \lim_{n \rightarrow \infty} E\xi_n$. ■

Следствие 2.1. Пусть $P\{\omega : \xi_n(\omega) \leq \xi_{n+1}(\omega), 0 \leq \xi_n(\omega) \xrightarrow[n \rightarrow \infty]{} \xi(\omega)\} = 1$. Тогда

$$\lim_{n \rightarrow \infty} E\xi_n = E\xi.$$

Доказательство. Пусть N — множество тех точек, в которых не выполнено условие теоремы, $P(N) = 0$. Введем случайные величины

$$\tilde{\xi}_n(\omega) = \begin{cases} \xi_n(\omega), & \omega \notin N, \\ 0, & \omega \in N; \end{cases}$$

$$\tilde{\xi}(\omega) = \begin{cases} \xi(\omega), & \omega \notin N, \\ 0, & \omega \in N. \end{cases}$$

Для $\tilde{\xi}_n(\omega)$ и $\tilde{\xi}(\omega)$ выполнено условие теоремы для любого ω . Следовательно,

$$E\tilde{\xi} = \lim_{n \rightarrow \infty} E\tilde{\xi}_n.$$

Так как мера множества N равна нулю, то $E\xi = E\tilde{\xi} = \lim_{n \rightarrow \infty} E\tilde{\xi}_n = \lim_{n \rightarrow \infty} E\xi_n$. ■

Следствие 2.2. Пусть для любого $\omega \in \Omega$ имеет место сходямость $\eta(\omega) \leq \xi_n(\omega) \xrightarrow[n \rightarrow \infty]{} \xi(\omega)$ ($\xi_n(\omega) \leq \xi_{n+1}(\omega)$ для любого $\omega \in \Omega$). Пусть $E\eta > -\infty$, тогда

$$\lim_{n \rightarrow \infty} E\xi_n = E\xi.$$

Доказательство. Рассмотрим $\tilde{\xi}_n = \xi_n - \eta \geq 0$ и $\tilde{\xi} = \xi - \eta \geq 0$. Для таких $\tilde{\xi}_n, \tilde{\xi}$ выполнены условия теоремы, следовательно,

$$E\tilde{\xi}_n \xrightarrow{n \rightarrow \infty} E\tilde{\xi}.$$

Тогда $E\tilde{\xi}_n + E\eta \xrightarrow{n \rightarrow \infty} E\tilde{\xi} + E\eta$. В силу свойства аддитивности получаем равенства:

$$E\tilde{\xi}_n + E\eta = E\xi_n,$$

$$E\tilde{\xi} + E\eta = E\xi.$$

Таким образом, следствие доказано. ■

Следствие 2.3. Пусть для любого $\omega \in \Omega$ имеет место сходимость $\eta(\omega) \geq \xi_n(\omega) \xrightarrow{n \rightarrow \infty} \xi(\omega)$ ($\xi_n(\omega) \geq \xi_{n+1}(\omega)$ для любого $\omega \in \Omega$). Пусть $E\eta < \infty$, тогда

$$\lim_{n \rightarrow \infty} E\xi_n = E\xi.$$

Доказательство. Рассмотрим $\tilde{\eta}(\omega) = -\eta(\omega)$, $\tilde{\xi}(\omega) = -\xi(\omega)$, $\tilde{\xi}_n(\omega) = -\xi_n(\omega)$. Таким образом, мы оказываемся в условиях следствия 2.2. ■

Следствие 2.4. Пусть при всех n : $\xi_n(\omega) \geq 0$. Тогда

$$\int_{\Omega} \sum_{n=1}^{\infty} \xi_n(\omega) dP = \sum_{n=1}^{\infty} \int_{\Omega} \xi_n(\omega) dP.$$

Доказательство. Рассмотрим частичную сумму ряда:

$$0 \leq S_n(\omega) = \sum_{k=1}^n \xi_k(\omega) \xrightarrow{n \rightarrow \infty} \sum_{k=1}^{\infty} \xi_k(\omega).$$

Таким образом, условия теоремы 2.1 выполняются:

$$\begin{aligned} \int_{\Omega} \sum_{k=1}^{\infty} \xi_k(\omega) dP &= \lim_{n \rightarrow \infty} \int_{\Omega} \sum_{k=1}^n \xi_k(\omega) dP = \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \int_{\Omega} \xi_k(\omega) dP = \sum_{k=1}^{\infty} \int_{\Omega} \xi_k(\omega) dP. \end{aligned}$$

■

§ 3. ТЕОРЕМА ЛЕБЕГА О МАЖОРИРУЕМОЙ СХОДИМОСТИ

Лемма 3.1 (лемма Фату).

1. Пусть заданы случайные величины $\eta, \xi_1, \dots, \xi_n, \dots$ на вероятностном пространстве (Ω, \mathcal{F}, P) , при этом $\eta(\omega) \leq \leq \xi_n(\omega)$ для любого n , $E\eta > -\infty$, тогда $\varliminf_{n \rightarrow \infty} E\xi_n \geq \geq E \varliminf_{n \rightarrow \infty} \xi_n$.
2. Пусть $\xi_n(\omega) \leq \eta(\omega)$ для любого n , $E\eta < \infty$, тогда $\varliminf_{n \rightarrow \infty} E\xi_n \leq E \varliminf_{n \rightarrow \infty} \xi_n$.
3. Пусть заданы случайные величины $\eta, \xi_1, \dots, \xi_n, \dots$ на вероятностном пространстве (Ω, \mathcal{F}, P) такие, что $|\xi_n(\omega)| \leq \eta(\omega)$ для любого n , и $E\eta < \infty$, тогда

$$E \varliminf_{n \rightarrow \infty} \xi_n \leq \varliminf_{n \rightarrow \infty} E\xi_n \leq \overline{\lim}_{n \rightarrow \infty} E\xi_n \leq E \overline{\lim}_{n \rightarrow \infty} \xi_n.$$

Доказательство. Докажем первое утверждение леммы. В равенстве $\varliminf_{n \rightarrow \infty} \xi_n(\omega) = \lim_{n \rightarrow \infty} \inf_{m \geq n} \xi_m(\omega)$ обозначим $\inf_{m \geq n} \xi_m(\omega)$ через $\zeta_n(\omega)$. Тогда по построению $\eta(\omega) \leq \zeta_n(\omega) \leq \zeta_{n+1}(\omega)$ и $\zeta_n(\omega) \xrightarrow[n \rightarrow \infty]{} \varliminf_{n \rightarrow \infty} \xi_n(\omega)$. Согласно следствию 2.2 справедливо:

$$E \varliminf_{n \rightarrow \infty} \xi_n = \lim_{n \rightarrow \infty} E\zeta_n = \varliminf_{n \rightarrow \infty} E\zeta_n \leq \varliminf_{n \rightarrow \infty} E\xi_n.$$

Так как $\zeta_n(\omega) \leq \xi_n(\omega)$, то $E\zeta_n(\omega) \leq E\xi_n(\omega)$.

Для доказательства второго утверждения леммы нужно все случайные величины умножить на -1 . Третье утверждение леммы следует из первого и второго утверждений. ■

Теорема 3.1 (о мажорируемой сходимости). Пусть $\eta(\omega), \xi(\omega), \xi_1(\omega), \dots, \xi_n(\omega), \dots$ — случайные величины, заданные на вероятностном пространстве (Ω, \mathcal{F}, P) . Выполнены следующие условия: $E\eta < \infty$, $|\xi_n(\omega)| \leq \eta(\omega)$ для любого n , $\lim_{n \rightarrow \infty} \xi_n(\omega) \stackrel{n.н.}{=} \xi(\omega)$, тогда

1. Существует конечное математическое ожидание $E\xi$.
2. Имеет место сходимость: $E\xi_n \xrightarrow[n \rightarrow \infty]{} E\xi$.
3. Имеет место сходимость: $E|\xi(\omega) - \xi_n(\omega)| \xrightarrow[n \rightarrow \infty]{} 0$.

Доказательство. Докажем первое утверждение теоремы. Так как $|\xi_n(\omega)| \leq \eta(\omega)$ для любых n , можно сделать предельный переход: $|\xi(\omega)| \leq \eta(\omega)$ п.н. Согласно свойству положительности интеграла Лебега справедливо неравенство: $E|\xi| \leq E\eta < \infty$. Согласно свойству конечности интеграла Лебега $E\xi \in \mathbb{R}$.

Докажем второе утверждение теоремы. Условия леммы Фату 3.1 выполнены, тогда имеют место неравенства:

$$E \lim_{n \rightarrow \infty} \xi_n \leq \lim_{n \rightarrow \infty} E\xi_n \leq \overline{\lim}_{n \rightarrow \infty} E\xi_n \leq E \overline{\lim}_{n \rightarrow \infty} \xi_n.$$

Так как

$$\begin{aligned} \lim_{n \rightarrow \infty} \xi_n &\stackrel{\text{п.н.}}{=} \lim_{n \rightarrow \infty} \xi_n(\omega) \stackrel{\text{п.н.}}{=} \xi(\omega), \\ \overline{\lim}_{n \rightarrow \infty} \xi_n &\stackrel{\text{п.н.}}{=} \lim_{n \rightarrow \infty} \xi_n(\omega) \stackrel{\text{п.н.}}{=} \xi(\omega), \end{aligned}$$

то получаем искомое равенство.

Третье утверждение доказывается аналогично. ■

§ 4. ТЕОРЕМА О ЗАМЕНЕ ПЕРЕМЕННОЙ ПОД ЗНАКОМ ИНТЕГРАЛА ЛЕБЕГА

Теорема 4.1 (о замене переменной). Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) задан случайный вектор $\xi(\omega) = (\xi_1(\omega), \dots, \xi_n(\omega))$, функция $g : \mathbb{R}^n \rightarrow \mathbb{R}$ измерима относительно борелевских σ -алгебр. Тогда если существует один из двух интегралов: $\int_{\Omega} g(\xi_1(\omega), \dots, \xi_n(\omega)) dP$, $\int_{\mathbb{R}^n} g(x_1, \dots, x_n) dP_{\xi}$, то существует и второй, и они равны:

$$\int_{\Omega} g(\xi_1(\omega), \dots, \xi_n(\omega)) dP = \int_{\mathbb{R}^n} g(x_1, \dots, x_n) dP_{\xi}, \quad (4.1)$$

где P_{ξ} — распределение случайного вектора ξ .

Доказательство. Вспомним определение 1.1, рассмотрим каждый из его пунктов. Покажем, что равенство (4.1) выполняется для каждого из них. Заметим также, что борелевская функция $g(x)$ на пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_{\xi})$ является случайной величиной.

1. Пусть $g(x) \geq 0$ — простая числовая функция такая, что $g(x) = \sum_{i=1}^m a_i I\{x \in A_i\}$, $A_k \cap A_l = \emptyset$, $k \neq l$, $\bigcup_{k=1}^m A_k = \mathbb{R}^n$, $A_k \in \mathcal{B}(\mathbb{R}^n)$.

Рассмотрим следующую функцию этого класса:

$$g(x) = 1 \cdot I\{x \in A\} + 0 \cdot I\{x \in \bar{A}\}.$$

Подставим это выражение в левую часть равенства (4.1) и рассмотрим интеграл:

$$\int_{\Omega} I\{\xi(\omega) \in A\} dP = P\{\xi \in A\} = P\{\xi^{-1}(A)\}.$$

Подставим данное представление в правую часть равенства (4.1), получим:

$$\int_{\mathbb{R}^n} I\{x \in A\} dP_{\xi} = P_{\xi}(A).$$

Таким образом, для данного вида функции $g(x)$ равенство (4.1) доказано. Следовательно, можно утверждать справедливость формулы:

$$\int_{\Omega} I\{\xi(\omega) \in A_i\} dP = \int_{\mathbb{R}^n} I\{x \in A_i\} dP_{\xi}.$$

Домножим обе части равенства на a_i и просуммируем, воспользовавшись свойством аддитивности:

$$\int_{\Omega} g(\xi(\omega)) dP = \int_{\mathbb{R}^n} g(x) dP_{\xi}.$$

2. Пусть $g(x) \geq 0$ — расширенная числовая функция. Существует последовательность простых функций $g_k(x)$ такая, что $g_k(x) \leq g_{k+1}(x)$, $0 \leq g_k(x) \xrightarrow{k \rightarrow \infty} g(x)$. По определению 1.1 получаем:

$$\int_{\Omega} g(\xi(\omega)) dP = \lim_{k \rightarrow \infty} \int_{\Omega} g_k(\xi(\omega)) dP.$$

Также можно написать: $\int_{\mathbb{R}^n} g(x) dP_\xi = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} g_k(x) dP_\xi$. Следовательно,

$$\int_{\Omega} g(\xi(\omega)) dP = \int_{\mathbb{R}^n} g(x) dP_\xi.$$

3. Любую расширенную числовую функцию $g(x)$ можно представить в виде $g(x) = g^+(x) - g^-(x)$. По отдельности для $g^+(x)$ и $g^-(x)$ равенство (4.1) доказано:

$$\int_{\Omega} g^+(\xi(\omega)) dP = \int_{\mathbb{R}^n} g^+(x) dP_\xi,$$

$$\int_{\Omega} g^-(\xi(\omega)) dP = \int_{\mathbb{R}^n} g^-(x) dP_\xi.$$

Но тогда можем утверждать, что справедливо и равенство:

$$\int_{\Omega} g(\xi(\omega)) dP = \int_{\mathbb{R}^n} g(x) dP_\xi,$$

так как если существует левая часть, то существует и правая часть равенства, и наоборот. ■

Следствие 4.1. Пусть выполнены все условия теоремы 4.1, $n = 1$, $g(x) = x$, тогда

$$E\xi = \int_{\Omega} \xi(\omega) dP = \int_{\mathbb{R}} x dP_\xi. \quad (4.2)$$

Замечание 4.1. В равенствах (4.1) и (4.2) интегралы в правых частях зависят от случайных величин только через их распределения, поэтому интегралы в (4.1) и (4.2) могут быть записаны специальным образом:

$$\int_{\mathbb{R}^n} g(x_1, \dots, x_n) dP_\xi \equiv \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(x_1, \dots, x_n) dF_\xi(x),$$

$$\int_{\mathbb{R}} x dP_{\xi} \equiv \int_{-\infty}^{+\infty} x dF_{\xi}(x).$$

Эти интегралы называются интегралами Лебега–Стилтьеса.

Следствие 4.2. Пусть $n = 1$, тогда

$$\int_{\Omega} g(\xi(\omega)) dP = \int_{\mathbb{R}} g(x) dP_{\xi} = \int_{-\infty}^{+\infty} g(x) dF_{\xi}(x).$$

Здесь важно, что суперпозиция функций также является случайной величиной $\eta(\omega) = g(\xi_1(\omega))$. Согласно следствию 4.1 получаем:

$$E(g(\xi)) = \int_{\Omega} g(\xi(\omega)) dP = \int_{-\infty}^{\infty} g(x) dF_{\xi}(x) = \int_{-\infty}^{+\infty} y dF_{\eta}(y).$$

Следствие 4.3. Пусть выполнены все условия теоремы 4.1 и $B \in \mathcal{B}(\mathbb{R}^n)$, тогда если существует один из интегралов, то существует и второй, и они равны:

$$\int_{\xi^{-1}(B)} g(\xi_1(\omega), \dots, \xi_n(\omega)) dP = \int_B g(x_1, \dots, x_n) dP_{\xi}.$$

Для доказательства достаточно рассмотреть функцию $\tilde{g}(x_1, \dots, x_n) = I\{x \in B\}g(x_1, \dots, x_n)$.

Теорема 4.2. Пусть выполнены условия теоремы 4.1, и у случайного вектора ξ существует плотность распределения $f_{\xi}(x_1, \dots, x_n)$. Тогда если существует один из интегралов, то существует и второй, и они равны:

$$\begin{aligned} \int_{\mathbb{R}^n} g(\xi_1, \dots, \xi_n) dP_{\xi} &= \\ &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(x_1, \dots, x_n) f_{\xi}(x_1, \dots, x_n) dx_1 \dots dx_n. \end{aligned}$$

Доказательство. Доказательство теоремы 4.2 аналогично доказательству теоремы 4.1, поэтому сделаем только первый шаг в доказательстве теоремы и укажем дальнейшие шаги.

1. Рассмотрим функцию $g(x) = I\{x \in A\}$, $A \in \mathcal{B}(\mathbb{R}^n)$. Тогда

$$\int_{\Omega} I\{\xi(\omega) \in A\} dP = P\{\xi^{-1}(A)\}.$$

Справедливы равенства:

$$\begin{aligned} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} I\{x \in A\} f_{\xi}(x) dx &= \int_{\mathbb{R}^n} I\{x \in A\} f_{\xi}(x) dx = \\ &= \int_A f_{\xi}(x) dx = P_{\xi}(A) = P\{\xi^{-1}(A)\}. \end{aligned}$$

Далее рассуждение повторяется (см. теорему 4.1), имеется только одно отличие. В части 2 доказательства теоремы, применив теорему о монотонной сходимости под знаком интеграла Лебега, получим:

$$\begin{aligned} \int_{\Omega} g(\xi(\omega)) dP &= \lim_{k \rightarrow \infty} \int_{\Omega} g_k(\xi(\omega)) dP, \\ \int_{\mathbb{R}^n} g(x) f_{\xi}(x) dx &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} g_k(x) f_{\xi}(x) dx, \end{aligned}$$

откуда следует доказываемое утверждение. ■

Замечание 4.2 (о связи интеграла Лебега и интеграла Римана).

1. Пусть $g(x)$ задана на отрезке $[a, b]$, и существует интеграл Римана, который обозначим через $(R) \int_a^b g(x) dx$, тогда существует и интеграл Лебега, который обозначим через $(L) \int_a^b g(x) dx$, и они равны:

$$(R) \int_a^b g(x) dx = (L) \int_a^b g(x) dx.$$

2. Пусть $g(x) \geq 0$ на $[a, +\infty)$ и на любом отрезке $[a, b]$ существует интеграл Римана $(R) \int_a^b g(x) dx$, тогда существуют и равны следующие интегралы:

$$(R) \int_a^{+\infty} g(x) dx = (L) \int_a^{+\infty} g(x) dx.$$

Несобственный интеграл Римана совпадает с интегралом Лебега. Отметим, что оба интеграла могут быть равны $+\infty$.

§ 5. ФОРМУЛЫ ДЛЯ ВЫЧИСЛЕНИЯ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

1. Пусть $\xi(\omega) \geq 0$ — простая случайная величина, т. е.

$$\xi(\omega) = \sum_{i=1}^m a_i I\{\omega \in A_i\}.$$

Математическое ожидание простой случайной величины можно вычислить по формуле:

$$E\xi = \sum_{i=1}^m a_i P(A_i).$$

Всегда можно добиться того, чтобы все значения a_i были бы различными, тогда $A_i = \xi^{-1}(a_i)$, получаем формулу для вычисления математического ожидания неотрицательной простой случайной величины:

$$E\xi = \sum_{i=1}^m a_i P\{\xi = a_i\}. \quad (5.1)$$

2. Пусть $\xi(\omega)$ — простая случайная величина, т. е.

$$\xi(\omega) = \sum_{i=1}^m b_i I\{\omega \in B_i\}.$$

Также $\xi(\omega)$ можно представить в виде:

$$\xi(\omega) = \xi^+(\omega) - \xi^-(\omega),$$

где

$$\xi^+(\omega) = \sum_{i=1}^m b_i^+ I\{\omega \in B_i\},$$

$$\xi^-(\omega) = \sum_{i=1}^m b_i^- I\{\omega \in B_i\}.$$

Математическое ожидание простой случайной величины можно вычислить по формуле:

$$E\xi = E\xi^+ - E\xi^- = \sum_{i=1}^m b_i^+ P(B_i) - \sum_{i=1}^m b_i^- P(B_i) = \sum_{i=1}^m b_i P(B_i),$$

где $b_i = b_i^+ - b_i^-$. Всегда можно добиться того, чтобы все значения b_i были бы различными, тогда $B_i = \xi^{-1}(b_i)$, получаем формулу для вычисления математического ожидания простой случайной величины:

$$E\xi = \sum_{i=1}^m b_i P\{\xi = b_i\}. \quad (5.2)$$

3. Пусть $\xi(\omega)$ — дискретная случайная величина, $\xi \in \{b_i, i \in N\}$, другими словами, множество значений случайной величины счетно:

$$\xi(\omega) = \sum_{i=1}^{\infty} b_i I\{\omega \in B_i\},$$

где $B_i \cap B_j = \emptyset$ для всех $i \neq j$, $\bigcup_{i=1}^{\infty} B_i = \Omega$. Всегда можно добиться того, чтобы $b_i \neq b_j$, тогда $B_i = \xi^{-1}(b_i)$. Выделим положительную и отрицательную части:

$$\xi(\omega) = \xi^+(\omega) - \xi^-(\omega),$$

где

$$\xi^+(\omega) = \sum_{i=1}^{\infty} b_i^+ I\{\omega \in B_i\},$$

$$\xi^-(\omega) = \sum_{i=1}^{\infty} b_i^- I\{\omega \in B_i\}.$$

Тогда

$$\xi_n^+(\omega) \leq \xi_{n+1}^+(\omega), \quad 0 \leq \xi_n^+(\omega) = \sum_{i=1}^n b_i^+ I\{\omega \in B_i\} \xrightarrow{k \rightarrow \infty} \xi^+(\omega),$$

$$\xi_n^-(\omega) \leq \xi_{n+1}^-(\omega), \quad 0 \leq \xi_n^-(\omega) = \sum_{i=1}^n b_i^- I\{\omega \in B_i\} \xrightarrow{k \rightarrow \infty} \xi^-(\omega).$$

Следовательно,

$$E\xi^+ = \lim_{n \rightarrow \infty} E\xi_n^+ = \lim_{n \rightarrow \infty} \sum_{i=1}^n b_i^+ P(B_i) = \sum_{i=1}^{\infty} b_i^+ P(B_i),$$

$$E\xi^- = \lim_{n \rightarrow \infty} E\xi_n^- = \lim_{n \rightarrow \infty} \sum_{i=1}^n b_i^- P(B_i) = \sum_{i=1}^{\infty} b_i^- P(B_i).$$

Тогда математическое ожидание случайной величины $\xi(\omega)$ можно вычислить следующим образом:

$$E\xi = E\xi^+ - E\xi^- = \sum_{i=1}^{\infty} b_i^+ P(B_i) - \sum_{i=1}^{\infty} b_i^- P(B_i). \quad (5.3)$$

Справедливы утверждения:

- Если $E\xi^+ = \infty$, $E\xi^- \in \mathbb{R}$, то $E\xi = \infty$.
- Если $E\xi^+ \in \mathbb{R}$, $E\xi^- = \infty$, то $E\xi = -\infty$.
- Если $E\xi^+ \in \mathbb{R}$, $E\xi^- \in \mathbb{R}$, то $E\xi = \sum_{i=1}^{\infty} b_i P(B_i)$, причем, ряд сходится в абсолютном смысле.
- Если $E\xi^+ = \infty$, $E\xi^- = \infty$, то математическое ожидание ξ не существует.

4. Рассмотрим случайную величину ξ , у которой существует плотность распределения $f_\xi(x)$. Математическое ожидание такой случайной величины можно вычислить по формуле:

$$E\xi = \int_{-\infty}^{\infty} x f_\xi(x) dx = \int_0^{\infty} x f_\xi(x) dx - \int_{-\infty}^0 |x| f_\xi(x) dx, \quad (5.4)$$

если интеграл Лебега существует. В общем случае верна формула:

$$E\xi = \int_{-\infty}^{\infty} x dF_{\xi}(x). \quad (5.5)$$

Если существует плотность распределения, то

$$Eg(\xi) = \int_{-\infty}^{\infty} g(x) dF_{\xi}(x) = \int_{-\infty}^{\infty} g(x) f_{\xi}(x) dx. \quad (5.6)$$

Пример 5.1. Пусть случайная величина ξ подчиняется нормальному распределению с параметрами (a, σ^2) . Плотность ее распределения имеет вид: $f_{\xi}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$. Вычислим математическое ожидание этой случайной величины:

$$\begin{aligned} E\xi &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-a) e^{-\frac{(x-a)^2}{2\sigma^2}} dx + \\ &\quad + a \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = a. \end{aligned}$$

Пример 5.2. Рассмотрим распределение Коши с плотностью распределения $f_{\xi}(x) = \frac{1}{\pi} \frac{1}{1+x^2}$. Согласно формуле (5.4), получаем

$$E\xi = \frac{1}{\pi} \int_0^{\infty} x \frac{1}{1+x^2} dx - \frac{1}{\pi} \int_{-\infty}^0 |x| \frac{1}{1+x^2} dx.$$

Оба интеграла в правой части равны бесконечности, следовательно, математического ожидания у случайной величины, подчиняющейся распределению Коши, не существует.

Пример 5.3. Пусть μ — случайная величина, равная числу успехов в n независимых испытаниях Бернулли с вероятностью успеха в каждом испытании p и вероятностью неудачи

$q = 1 - p$. Математическое ожидание случайной величины μ можно вычислить по формуле

$$E\mu = \sum_{m=0}^n m C_n^m p^m q^{n-m}.$$

Однако можно получить более удобную формулу. Определим случайные величины $\xi_i = I\{\text{успех в } i\text{-ом испытании}\}$, тогда $\mu = \sum_{i=1}^n \xi_i$. Заметим, что $E\xi_i = p$, тогда

$$E\mu = \sum_{i=1}^n E\xi_i = np.$$

§ 6. МОМЕНТЫ СЛУЧАЙНЫХ ВЕЛИЧИН. ДИСПЕРСИЯ

Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) .

Определение 6.1. Начальным моментом порядка k случайной величины ξ называется интеграл Лебега–Стилтьеса:

$$a_k = E(\xi^k) = \int_{-\infty}^{\infty} x^k dF_{\xi}(x), \quad (6.1)$$

если математическое ожидание $E(\xi^k)$ существует и конечно.

Определение 6.2. Центральным моментом порядка k случайной величины ξ называется интеграл Лебега–Стилтьеса:

$$a_k^0 = E(\xi - E\xi)^k = \int_{-\infty}^{\infty} (\xi - E\xi)^k dF_{\xi}(x), \quad (6.2)$$

если интеграл существует и конечен.

Среди центральных моментов особое место занимает момент второго порядка $a_2^0 = E(\xi - E\xi)^2$, который называется *дисперсией* случайной величины ξ и обозначается $D\xi$. Среднеквадратическим отклонением случайной величины ξ называется число $\sigma_{\xi} = \sqrt{D\xi}$.

Свойства дисперсии

Будем предполагать, что выполнено условие $E(\xi^2) < \infty$.

I. Дисперсия случайной величины ξ удовлетворяет условию:

$$D\xi \geq 0.$$

При этом $D\xi = 0$ тогда и только тогда, когда $\xi \stackrel{\text{п.н.}}{=} \text{const}$, т. е. является вырожденной случайной величиной.

Доказательство. Очевидно, что $D\xi = \sigma_\xi^2 \geq 0$. По свойству интеграла Лебега $D\xi = E(\xi - E\xi)^2 = 0$ тогда и только тогда, когда $\xi(\omega) \stackrel{\text{п.н.}}{=} E\xi \in \mathbb{R}$. ■

II. Пусть случайная величина η связана со случайной величиной ξ следующим равенством: $\eta \stackrel{\text{п.н.}}{=} a\xi + b$, тогда

$$D\eta = a^2 D\xi.$$

Доказательство. Согласно определению дисперсии:

$$D\eta = D(a\xi + b) = E(a\xi + b - E(a\xi + b))^2.$$

Так как $E(a\xi + b) = aE\xi + b$, то

$$D\eta = E(a\xi - aE\xi)^2 = E(a(\xi - E\xi))^2 = a^2 D\xi.$$

■

III. Пусть случайные величины ξ и η независимы, тогда

$$D(\xi \pm \eta) = D\xi + D\eta.$$

Доказательство. Согласно определению дисперсии:

$$\begin{aligned} D(\xi \pm \eta) &= E((\xi - E\xi) \pm (\eta - E\eta))^2 = \\ &= D\xi + D\eta \pm 2E(\xi - E\xi)(\eta - E\eta) = D\xi + D\eta, \end{aligned}$$

так как в силу независимости случайных величин ξ и η слагаемое $E(\xi - E\xi)(\eta - E\eta)$ можно представить в виде $E(\xi - E\xi)E(\eta - E\eta) = 0$. ■

IV. Имеет место формула для вычисления дисперсии:

$$D\xi = E(\xi^2) - (E\xi)^2.$$

Доказательство. По определению дисперсии:

$$D\xi = E(\xi - E\xi)^2 = E\xi^2 - 2E\xi E\xi + (E\xi)^2 = E\xi^2 - (E\xi)^2.$$

■

V. Справедливо неравенство:

$$D\xi \leq E(\xi - a)^2,$$

где a — константа, при этом $D\xi = E(\xi - a)^2$ тогда и только тогда, когда $a = E\xi$.

Доказательство. Рассмотрим $E(\xi - a)^2$. К выражению в скобках прибавим и вычтем величину $E\xi$:

$$\begin{aligned} E(\xi - a)^2 &= E((\xi - E\xi) + (E\xi - a))^2 = E(\xi - E\xi)^2 + \\ &+ 2(E\xi - a)E(\xi - E\xi) + (E\xi - a)^2 = D\xi + (E\xi - a)^2, \end{aligned}$$

где $E(\xi - E\xi)^2 = D\xi$ по определению дисперсии, и $E(\xi - E\xi) = 0$.

■

Пример 6.1. Пусть случайная величина ξ подчиняется нормальному распределению с параметрами (a, σ^2) . Математическое ожидание случайной величины ξ , как было показано, равно a . Вычислим дисперсию:

$$\begin{aligned} D\xi &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \sigma^2 \int_{-\infty}^{\infty} \left(\frac{x-a}{\sigma}\right)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} d\left(\frac{x-a}{\sigma}\right) = \\ &= \frac{1}{\sqrt{2\pi}} \sigma^2 \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy = \\ &= -\frac{1}{\sqrt{2\pi}} \sigma^2 y e^{-\frac{y^2}{2}} \Big|_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \sigma^2. \end{aligned}$$

Пример 6.2. Пусть μ — случайная величина, равная числу успехов в n независимых испытаниях Бернулли с вероятностью успеха в каждом испытании p и вероятностью неудачи $q = 1 - p$. Математическое ожидание случайной величины μ равно np . Вычислим дисперсию μ :

$$D\mu = E(\mu - E\mu)^2 = \sum_{m=0}^n (m - np)^2 C_n^m p^m q^{n-m}.$$

Воспользуемся другим способом. Определим случайные величины $\xi_i = I\{\text{успех в } i\text{-ом испытании}\}$, тогда $\mu = \sum_{i=1}^n \xi_i$. Используя то, что случайные величины ξ_i , $i = 1, \dots, n$ независимы, свойство III дисперсии, и, вычислив дисперсию $D\xi_i = pq$, получаем:

$$D\mu = \sum_{i=1}^n D\xi_i = npq.$$

Если из случайной величины ξ вычесть ее математическое ожидание $E\xi$, то получим центрированную случайную величину, а если центрированную случайную величину разделить на $\sqrt{D\xi}$, то получим нормированную случайную величину. Вычислим дисперсию нормированной случайной величины:

$$E\left(\frac{\xi - E\xi}{\sqrt{D\xi}}\right)^2 = \frac{E(\xi - E\xi)^2}{(D\xi)} = 1.$$

§ 7. НЕРАВЕНСТВО ЧЕБЫШЕВА. ДРУГИЕ НЕРАВЕНСТВА

Лемма 7.1 (неравенство Чебышева). Пусть $\xi \geq 0$. Тогда для любого $\varepsilon > 0$ справедливо неравенство:

$$P\{\xi \geq \varepsilon\} \leq \frac{E\xi}{\varepsilon}.$$

Доказательство. Справедливо неравенство:

$$\xi(\omega) \geq \xi(\omega)I\{\xi(\omega) \geq \varepsilon\} \geq \varepsilon I\{\xi(\omega) \geq \varepsilon\}.$$

Проинтегрируем неравенство и применим свойство положительности интеграла Лебега:

$$E\xi \geq \varepsilon P\{\xi(\omega) \geq \varepsilon\}.$$

■

Следствие 7.1. *Справедливо неравенство:*

$$P\{|\xi - E\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}.$$

Доказательство. Преобразуем выражение:

$$P\{|\xi - E\xi| \geq \varepsilon\} = P\{(\xi - E\xi)^2 \geq \varepsilon^2\}.$$

Воспользуемся леммой 7.1, получаем неравенство:

$$P\{(\xi - E\xi)^2 \geq \varepsilon^2\} \leq \frac{E(\xi - E\xi)^2}{\varepsilon^2} = \frac{D\xi}{\varepsilon^2}.$$

■

Следствие 7.2 (закон больших чисел для взаимно независимых случайных величин с равномерно ограниченными дисперсиями). Пусть последовательность взаимно независимых случайных величин $\xi_n(\omega)$ задана на вероятностном пространстве (Ω, \mathcal{F}, P) , $E\xi_n(\omega) = a \in \mathbb{R}$ для любого n , дисперсии равномерно ограничены $D\xi_n \leq c$ для любого n . Тогда для любого $\varepsilon > 0$ имеет место сходимость:

$$P\left\{\left|\frac{\sum_{k=1}^n \xi_k}{n} - a\right| \geq \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Рассмотрим последовательность случайных величин $\eta_n = (\sum_{k=1}^n \xi_k)/n$. По свойству линейности математического ожидания справедливо равенство: $E\eta_n = a$. Применим следствие 7.1:

$$P\left\{\left|\frac{\sum_{k=1}^n \xi_k}{n} - a\right| \geq \varepsilon\right\} \leq \frac{nc}{n^2\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0,$$

так как

$$D\eta_n = E(\eta_n - a)^2 = E\left(\frac{1}{n} \sum_{k=1}^n (\xi_k - a)\right)^2 = \frac{1}{n^2} \sum_{k=1}^n D\xi_k.$$

■

Следствие 7.3 (закон больших чисел Бернулли). Пусть в схеме Бернулли с n независимыми испытаниями число успехов обозначается μ_n . Тогда для любого $\varepsilon \geq 0$ имеет место сходимость:

$$P\left\{\left|\frac{\mu_n}{n} - p\right| \geq \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Доказываемое утверждение представляет собой частный случай следствия 7.2. Пусть

$$\mu_n = \sum_{k=1}^n \xi_k,$$

где ξ_k : $P(\xi_k = 1) = p$, $P(\xi_k = 0) = q$. Тогда

$$E\xi_k = 1 \cdot p + 0 \cdot q = p,$$

$$D\xi_k = E(\xi_k - p)^2 = (1 - p)^2 p + (0 - p)^2 q = qp(q + p) = pq \leq \frac{1}{4}.$$

■

Лемма 7.2 (неравенство Йенсена). Пусть ξ — случайная величина, математическое ожидание которой конечно. Пусть $g(x)$ — выпуклая функция, заданная на числовой прямой, тогда

$$E(g(\xi)) \geq g(E(\xi)).$$

Доказательство. Для того, чтобы функция была выпуклой, необходимо и достаточно, чтобы для любого x_0 существовало число $\lambda(x_0)$ такое, что

$$g(x) \geq \lambda(x_0)(x - x_0) + g(x_0).$$

Пусть $x_0 = E\xi$, $x = \xi$, тогда имеет место неравенство:

$$g(\xi) \geq \lambda(E\xi)(\xi - E\xi) + g(E\xi).$$

При интегрировании этого неравенства получаем, что $E(\xi - E\xi) = 0$ и $E(g(E\xi)) = g(E\xi)$. Следовательно,

$$E(g(\xi)) \geq g(E(\xi)).$$

■

Замечание 7.1. Если функция $g(x)$ — линейная, то неравенство Йенсена выполняется как равенство.

Лемма 7.3 (неравенство Коши–Шварца–Буняковского). Пусть $E(\xi^2) < \infty$, $E(\eta^2) < \infty$, тогда существует $E(\xi\eta)$, причем справедливо неравенство:

$$|E(\xi\eta)| \leq \sqrt{E\xi^2} \sqrt{E\eta^2}. \quad (7.1)$$

Доказательство. Пусть $E\xi^2 = 0$. Следовательно, $\xi \stackrel{\text{п.н.}}{=} 0$, $\xi\eta \stackrel{\text{п.н.}}{=} 0$ и $E(\xi\eta) = 0$. Тогда неравенство Коши–Шварца–Буняковского выполнено как равенство. Аналогично для $E\eta^2 = 0$.

Пусть $0 < E\xi^2 < \infty$, $0 < E\eta^2 < \infty$. Введем случайные величины:

$$\tilde{\xi} = \frac{\xi}{\sqrt{E\xi^2}} \quad \text{и} \quad \tilde{\eta} = \frac{\eta}{\sqrt{E\eta^2}}.$$

Рассмотрим выражение:

$$(\tilde{\xi} \pm \tilde{\eta})^2 \geq 0.$$

Раскроем скобки и преобразуем:

$$\tilde{\xi}^2 + \tilde{\eta}^2 \geq \pm 2\tilde{\xi}\tilde{\eta}.$$

Проинтегрировав последнее неравенство, получаем:

$$2 \geq \pm 2E(\tilde{\xi}\tilde{\eta}),$$

или эквивалентное неравенство:

$$\pm E(\xi\eta) \leq \sqrt{E\xi^2} \sqrt{E\eta^2}.$$

Последнее неравенство доказывает лемму. ■

Следствие 7.4. Если $E\xi^2 \neq 0$, $E\eta^2 \neq 0$, то неравенство (7.1) выполняется как равенство тогда и только тогда, когда случайные величины ξ и η почти наверное пропорциональны, т. е. существует константа $c \in \mathbb{R}$ такая, что

$$\eta(\omega) \stackrel{\text{п.н.}}{=} c\xi(\omega).$$

Доказательство. Достаточность очевидна:

$$|c|E\xi^2 = \sqrt{c^2} \sqrt{E\xi^2} \sqrt{E\xi^2}.$$

Докажем необходимость. Пусть неравенство (7.1) выполнено как равенство. Тогда возможны два случая:

1. $E(\xi\eta) = \sqrt{E\xi^2} \sqrt{E\eta^2}$.
2. $-E(\xi\eta) = \sqrt{E\xi^2} \sqrt{E\eta^2}$.

Рассмотрим первый случай:

$$E(\tilde{\xi}\tilde{\eta}) = 1,$$

где $\tilde{\xi} = \xi/\sqrt{E\xi^2}$, $\tilde{\eta} = \eta/\sqrt{E\eta^2}$. Домножим обе части равенства на два и проделаем преобразования из доказательства леммы 7.3, получим:

$$2E(\tilde{\xi}\tilde{\eta}) = 2 = E\tilde{\xi}^2 + E\tilde{\eta}^2.$$

Следовательно, $E(\tilde{\xi} - \tilde{\eta})^2 = 0$, $\tilde{\xi} - \tilde{\eta} \stackrel{\text{п.н.}}{=} 0$ или $\eta \stackrel{\text{п.н.}}{=} \sqrt{\frac{E\xi^2}{E\eta^2}} \xi$.

Аналогично рассматривается второй случай. ■

§ 8. КОВАРИАЦИЯ, КОРРЕЛЯЦИЯ

Пусть заданы случайные величины ξ и η на вероятностном пространстве (Ω, \mathcal{F}, P) , и существуют конечные математические ожидания $E\xi \in \mathbb{R}$, $E\eta \in \mathbb{R}$, а также $D\xi > 0$ и $D\eta > 0$. Кроме того, полагаем, что $E\xi^2 < \infty$ и $E\eta^2 < \infty$.

Определение 8.1. Ковариация $cov(\xi, \eta)$ случайных величин ξ и η определяется формулой:

$$cov(\xi, \eta) = E\{(\xi - E\xi)(\eta - E\eta)\}.$$

Определение 8.2. Коэффициент корреляции $\varrho(\xi, \eta)$ случайных величин ξ и η определяется формулой:

$$\varrho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}}.$$

Свойства

I. Справедлива формула:

$$\text{cov}(\xi, \eta) = E(\xi\eta) - E\xi E\eta.$$

Доказательство. Очевидно, что

$$\begin{aligned} \text{cov}(\xi, \eta) &= E\{(\xi - E\xi)(\eta - E\eta)\} = \\ &= E\{\xi\eta - \xi E\eta - \eta E\xi + E\xi E\eta\} = \\ &= E(\xi\eta) - E\xi E\eta - E\eta E\xi + E\xi E\eta = E(\xi\eta) - E\xi E\eta. \end{aligned}$$

■

II. Пусть ξ и η взаимно независимы, и выполнены все условия, содержащиеся в начале параграфа, тогда

$$\text{cov}(\xi, \eta) = 0 \quad \text{и} \quad \varrho(\xi, \eta) = 0.$$

Доказательство. Воспользуемся свойством мультипликативности:

$$\text{cov}(\xi, \eta) = E\{(\xi - E\xi)(\eta - E\eta)\} = E(\xi - E\xi)E(\eta - E\eta),$$

так как случайные величины ξ и η взаимно независимы. Можно также записать:

$$E(\xi - E\xi) = E\xi - E\xi = 0,$$

$$E(\eta - E\eta) = E\eta - E\eta = 0.$$

Таким образом, $\text{cov}(\xi, \eta) = 0$. Так как по определению 8.2:

$$\varrho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}},$$

то $\varrho(\xi, \eta) = 0$.

■

III. Имеет место неравенство $|\varrho(\xi, \eta)| \leq 1$. Если $|\varrho(\xi, \eta)| = 1$, тогда существует константа c такая, что $\eta(\omega) \stackrel{\text{п.н.}}{=} c\xi(\omega) + b$.

Для доказательства этого свойства следует применить неравенство Коши–Шварца–Буняковского (см. лемму 7.3) к числителю выражения из определения 8.2.

Случайные величины ξ и η не коррелируют, если $\text{cov}(\xi, \eta) = 0$ и $\varrho(\xi, \eta) = 0$. Из независимости случайных величин следует их некоррелированность. Обратное верно лишь в случае, когда случайные величины ξ и η имеют совместное нормальное распределение.

IV. Пусть $\xi^T = (\xi_1, \dots, \xi_n)$ — случайный вектор. Определим матрицу

$$D\xi = E\{(\xi - E\xi)(\xi - E\xi)^T\} = \\ = \begin{pmatrix} D\xi_1 & \text{Cov}(\xi_1, \xi_2) & \dots & \text{Cov}(\xi_1, \xi_n) \\ \text{Cov}(\xi_2, \xi_1) & D\xi_2 & \dots & \text{Cov}(\xi_2, \xi_n) \\ \dots & \dots & \dots & \dots \\ \text{Cov}(\xi_n, \xi_1) & \text{Cov}(\xi_n, \xi_2) & \dots & D\xi_n \end{pmatrix}.$$

Эта матрица называется ковариационной или дисперсионной матрицей случайного вектора ξ и обозначается через $D\xi$. Нетрудно показать, что $D\xi \geq 0$ — неотрицательно определенная матрица.

§ 9. ЗАДАЧА О НАИЛУЧШЕМ ЛИНЕЙНОМ ПРОГНОЗЕ

Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) . Пусть на нем заданы случайные величины ξ и η . Будем считать, что осуществляются наблюдения случайной величины ξ и по этим наблюдениям делается прогноз случайной величины η .

Рассмотрим функцию $\varphi(\xi)$, которая будет являться прогнозом значений случайной величины η . Пусть эта функция будет линейна:

$$\varphi(\xi) = a\xi + b.$$

В качестве критерия точности прогноза выберем функцию $Q(a, b) = E(\eta - \varphi(\xi))^2$. Тогда задача принимает следующий

вид: надо выбрать параметры a и b , доставляющие минимум $Q(a, b)$, т. е.

$$\min_{a, b} Q(a, b) = \min_{a, b} E(\eta - a\xi - b)^2.$$

Введем новую случайную величину $\zeta = \eta - a\xi$ и сначала будем решать задачу:

$$\min_b E(\zeta - b)^2.$$

Как было показано при доказательстве свойства V для дисперсий минимум математического ожидания $E(\zeta - b)^2$ достигается в точке $\hat{b} = E\zeta$. Найдем математическое ожидание случайной величины ζ :

$$\hat{b} = E\zeta = E(\eta - a\xi) = E\eta - aE\xi,$$

и подставим его в функцию $Q(a, b)$. Будем решать задачу: $\min_a Q(a, \hat{b})$.

$$\min_a Q(a, \hat{b}) = \min_a E((\eta - E\eta) - a(\xi - E\xi))^2.$$

Преобразуем функцию $Q(a, \hat{b})$:

$$\begin{aligned} Q(a, \hat{b}) &= D\eta - 2acov(\xi, \eta) + a^2 D\xi = \\ &= D\eta - 2a\sqrt{D\xi}\sqrt{D\eta}\varrho(\xi, \eta) + a^2 D\xi = D\eta - D\eta\varrho^2(\xi, \eta) + \\ &+ (D\eta\varrho^2(\xi, \eta) - 2a\sqrt{D\xi}\sqrt{D\eta}\varrho(\xi, \eta) + a^2 D\xi) = \\ &= D\eta(1 - \varrho^2(\xi, \eta)) + \left[\sqrt{D\eta}\varrho(\xi, \eta) - a\sqrt{D\xi}\right]^2. \end{aligned}$$

Минимум функции $Q(a, \hat{b})$ по параметру a достигается в точке

$$\hat{a} = \sqrt{\frac{D\eta}{D\xi}}\varrho(\xi, \eta).$$

Оптимальное значение функции: $Q(\hat{a}, \hat{b}) = D\eta(1 - \varrho^2(\xi, \eta))$.

Получаем оптимальное значение прогноза:

$$\hat{\varphi}(\xi) = \hat{a}\xi + \hat{b} = E\eta + \sqrt{\frac{D\eta}{D\xi}}\varrho(\xi, \eta)(\xi - E\xi).$$

Если $|\varrho(\xi, \eta)| = 1$, а это возможно при линейной связи ξ и η , тогда $Q(\hat{a}, \hat{b}) = 0$.

Пусть случайные величины ξ и η не коррелируют (к примеру, ξ и η независимы), т. е. $\varrho(\xi, \eta) = 0$, тогда величина ошибки максимальна: $Q(\hat{a}, \hat{b}) = D\eta$. Чем больше $|\varrho(\xi, \eta)|$, тем точнее линейный прогноз, коэффициент корреляции $\varrho(\xi, \eta)$ является величиной, характеризующей степень линейной зависимости.

УСЛОВНЫЕ РАСПРЕДЕЛЕНИЯ

§ 1. УСЛОВНЫЕ МАТЕМАТИЧЕСКИЕ ОЖИДАНИЯ ПРОСТЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) заданы две простые случайные величины:

$$\xi(\omega) = \sum_{i=1}^n a_i I\{\omega \in A_i\},$$

где $a_k \neq a_l$, $A_k \cap A_l = \emptyset$, $k \neq l$, $\bigcup_{k=1}^n A_k = \Omega$, $A_k = \xi^{-1}(a_k)$ и

$$\eta(\omega) = \sum_{j=1}^m b_j I\{\omega \in B_j\},$$

где $b_k \neq b_l$, $B_k \cap B_l = \emptyset$, $k \neq l$, $\bigcup_{k=1}^m B_k = \Omega$, $B_k = \eta^{-1}(b_k)$.

Так как справедливы представления:

$$A_i = \bigcup_{j=1}^m (A_i \cap B_j) \quad \text{и} \quad B_j = \bigcup_{i=1}^n (A_i \cap B_j),$$

то можно записать случайные величины $\xi(\omega)$ и $\eta(\omega)$ в следующем виде:

$$\xi(\omega) = \sum_{i=1}^n \sum_{j=1}^m a_i I\{\omega \in A_i \cap B_j\},$$

$$\eta(\omega) = \sum_{i=1}^n \sum_{j=1}^m b_j I\{\omega \in A_i \cap B_j\}.$$

Совместное распределение простых случайных величин $\xi(\omega)$ и $\eta(\omega)$ представимо в виде таблицы:

$\xi \setminus \eta$	b_1	b_2	\dots	b_m	Σ
a_1	p_{11}	p_{12}	\dots	p_{1m}	$p_{1\cdot}$
a_2	p_{21}	p_{22}	\dots	p_{2m}	$p_{2\cdot}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
a_n	p_{n1}	p_{n2}	\dots	p_{nm}	$p_{n\cdot}$
Σ	$p_{\cdot 1}$	$p_{\cdot 2}$	\dots	$p_{\cdot m}$	1

где

$$p_{ij} = P\{A_i \cap B_j\} = P\{\xi^{-1}(a_i) \cap \eta^{-1}(b_j)\} = P\{\xi = a_i, \eta = b_j\},$$

$$\begin{aligned} p_{i\cdot} &= \sum_{j=1}^m p_{ij} = \sum_{j=1}^m P(A_i \cap B_j) = \\ &= P\left\{\bigcup_{j=1}^m (A_i \cap B_j)\right\} = P(A_i) = P\{\xi = a_i\}, \end{aligned}$$

$$\begin{aligned} p_{\cdot j} &= \sum_{i=1}^n p_{ij} = \sum_{i=1}^n P(A_i \cap B_j) = \\ &= P\left\{\bigcup_{i=1}^n (A_i \cap B_j)\right\} = P(B_j) = P\{\eta = b_j\}, \end{aligned}$$

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

Случайные величины $\xi(\omega)$ и $\eta(\omega)$ независимы тогда и только тогда, когда для любых i, j выполняется равенство:

$$p_{ij} = p_{i\cdot} p_{\cdot j}.$$

Далее будем предполагать, что $p_{i\cdot} > 0$ для любого i и $p_{\cdot j} > 0$ для любого j . Рассмотрим строку с номером i таблицы совместного распределения простых случайных величин $\xi(\omega)$ и

$\eta(\omega)$. Рассмотрим отношения:

$$\frac{p_{i1}}{p_{i\cdot}}, \quad \frac{p_{i2}}{p_{i\cdot}}, \quad \dots, \quad \frac{p_{im}}{p_{i\cdot}}.$$

Все эти числа являются неотрицательными, и в сумме дают единицу. Рассмотрим условную вероятность:

$$\frac{p_{ij}}{p_{i\cdot}} = \frac{P\{A_i \cap B_j\}}{P\{A_i\}} = P\{B_j/A_i\} = P\{\eta = b_j/\xi = a_i\}.$$

Набор вероятностей $P\{\eta = b_1/\xi = a_i\}$, $P\{\eta = b_2/\xi = a_i\}$, \dots , $P\{\eta = b_m/\xi = a_i\}$ — условное распределение случайной величины $\eta(\omega)$ при условии, что $\xi = a_i$.

Рассмотрим столбец с номером j таблицы совместного распределения простых случайных величин $\xi(\omega)$ и $\eta(\omega)$. Рассмотрим отношения:

$$\frac{p_{1j}}{p_{\cdot j}}, \quad \frac{p_{2j}}{p_{\cdot j}}, \quad \dots, \quad \frac{p_{nj}}{p_{\cdot j}}.$$

Все эти числа являются неотрицательными, и в сумме дают единицу. Рассмотрим условную вероятность:

$$\frac{p_{ij}}{p_{\cdot j}} = \frac{P\{A_i \cap B_j\}}{P\{B_j\}} = P\{A_i/B_j\} = P\{\xi = a_i/\eta = b_j\}.$$

Набор вероятностей $P\{\xi = a_1/\eta = b_j\}$, $P\{\xi = a_2/\eta = b_j\}$, \dots , $P\{\xi = a_n/\eta = b_j\}$ — условное распределение случайной величины $\xi(\omega)$ при условии, что $\eta = b_j$.

Определение 1.1. Условное математическое ожидание случайной величины $\eta(\omega)$ при условии, что $\xi = a_i$, определяется равенством:

$$E(\eta/\xi = a_i) = \sum_{j=1}^m b_j P\{\eta = b_j/\xi = a_i\}. \quad (1.1)$$

Аналогично можно определить условное математическое ожидание случайной величины $\xi(\omega)$ при условии, что $\eta = b_j$:

$$E(\xi/\eta = b_j) = \sum_{i=1}^n a_i P\{\xi = a_i/\eta = b_j\}. \quad (1.2)$$

Определение 1.2. Условным математическим ожиданием случайной величины $\eta(\omega)$ относительно случайной величины ξ называется случайная величина $E(\eta/\xi)$, распределение которой определяется следующим образом:

$$\frac{E(\eta/\xi)}{P} \quad \left| \quad \frac{E(\eta/\xi = a_1)}{P\{\xi = a_1\}} \quad \right| \quad \frac{E(\eta/\xi = a_2)}{P\{\xi = a_2\}} \quad \left| \quad \dots \quad \right| \quad \frac{E(\eta/\xi = a_n)}{P\{\xi = a_n\}}$$

Случайная величина $E(\eta/\xi)$ — математическое ожидание простой случайной величины $\eta(\omega)$ относительно простой случайной величины $\xi(\omega)$.

Возможен и другой способ определения случайной величины $E(\eta/\xi)$.

Определение 1.3. Пусть

$$h(a_i) = E(\eta/\xi = a_i),$$

где функция $h(x)$ задана на множестве $\{a_1, \dots, a_n\}$. Тогда случайная величина $h(\xi(\omega)) = E(\eta/\xi)$ — условное математическое ожидание случайной величины η относительно случайной величины ξ .

Аналогично определяется условное математическое ожидание $E(\xi/\eta)$ случайной величины ξ относительно случайной величины η . Введенные таким образом определения 1.2 и 1.3 эквивалентны. Случайная величина $E(\eta/\xi)$ определяется единственным образом с точностью до множества меры нуль, то есть если $\zeta \stackrel{\text{п.н.}}{=} E(\eta/\xi)$, то случайная величина ζ также может рассматриваться как условное математическое ожидание случайной величины η относительно случайной величины ξ .

Сформулируем свойства условного математического ожидания.

1. Если случайные величины ξ и η взаимно независимы, то справедливо равенство:

$$E(\eta/\xi) = E\eta.$$

Доказательство. Доказательство следует из равенства (1.1). Если случайные величины ξ и η взаимно независимы, то

$$P\{\eta = b_j/\xi = a_i\} = P\{\eta = b_j\}.$$



2. Пусть случайная величина $\eta \stackrel{\text{п.н.}}{=} c$, $c \in \mathbb{R}$, тогда

$$E(\eta/\xi) \stackrel{\text{п.н.}}{=} c.$$

Доказательство. Как отмечалось ранее,

$$E(\eta/\xi = a_i) = \sum_{j=1}^m b_j P\{\eta = b_j/\xi = a_i\}.$$

Если $b_j \neq c$, то

$$P\{\eta = b_j/\xi = a_i\} = 0.$$

Если $b_j = c$, то нетрудно заметить, что

$$P\{\eta = c/\xi = a_i\} = \frac{P\{\eta^{-1}(c) \cap \xi^{-1}(a_i)\}}{P\{\xi^{-1}(a_i)\}} = 1.$$

Действительно,

$$\xi^{-1}(a_i) = \xi^{-1}(a_i) \cap \eta^{-1}(c) + \xi^{-1}(a_i) \cap \overline{\eta^{-1}(c)}.$$

Следовательно,

$$P\{\xi^{-1}(a_i)\} = P\{\xi^{-1}(a_i) \cap \eta^{-1}(c)\} + P\{\xi^{-1}(a_i) \cap \overline{\eta^{-1}(c)}\}.$$

Очевидно, что $P\{\xi^{-1}(a_i) \cap \overline{\eta^{-1}(c)}\} = 0$. Таким образом, $P\{\eta = c/\xi = a_i\} = 1$. ■

3. Справедливо равенство:

$$EE(\eta/\xi) = E\eta.$$

Доказательство. Как легко заметить,

$$\begin{aligned} EE(\eta/\xi) &= \sum_{i=1}^n E\{\eta/\xi = a_i\} P\{\xi = a_i\} = \\ &= \sum_{i=1}^n \sum_{j=1}^m b_j P\{\eta = b_j/\xi = a_i\} P\{\xi = a_i\} = \\ &= \sum_{i=1}^n \sum_{j=1}^m b_j P\{\eta = b_j, \xi = a_i\} = \sum_{j=1}^m b_j P\{\eta = b_j\} = E\eta. \end{aligned}$$

■

4. Пусть $\varphi(x)$ — борелевская функция, тогда имеет место равенство:

$$E\{\eta\varphi(\xi)/\xi\} = \varphi(\xi)E(\eta/\xi).$$

Доказательство. Воспользуемся определением 1.1:

$$\begin{aligned} E\{\eta\varphi(\xi)/\xi = a_i\} &= \sum_{j=1}^m \varphi(a_i)b_j P\{\eta = b_j/\xi = a_i\} = \\ &= \varphi(a_i)E(\eta/\xi = a_i) = g(a_i), \end{aligned}$$

где функция $g(x)$ задана на множестве $\{a_1, a_2, \dots, a_n\}$. По определению 1.3 справедливо равенство:

$$E\{\eta\varphi(\xi)/\xi\} = g(\xi) = \varphi(\xi)E(\eta/\xi).$$

■

Замечание 1.1. В теории вероятностей все случайные величины определяются с точностью до множества меры нуль, поэтому доказываемое равенство можно записать следующим образом:

$$E\{\eta\varphi(\xi)/\xi\} \stackrel{\text{п.н.}}{=} \varphi(\xi)E(\eta/\xi).$$

5. Пусть $\varphi(x)$ — борелевская функция. Для случайных величин ξ и η справедливо неравенство:

$$E(\eta - E(\eta/\xi))^2 \leq E(\eta - \varphi(\xi))^2, \quad (1.3)$$

причем неравенство выполняется как равенство только, когда $\varphi(\xi) \stackrel{\text{п.н.}}{=} E(\eta/\xi)$.

Доказательство. Преобразуем правую часть неравенства (1.3). К выражению в скобках прибавим и вычтем величину $E(\eta/\xi)$, раскроем скобки, тогда

$$\begin{aligned} E(\eta - \varphi(\xi))^2 &= E\{(\eta - E(\eta/\xi)) + (E(\eta/\xi) - \varphi(\xi))\}^2 = \\ &= E(\eta - E(\eta/\xi))^2 + E(E(\eta/\xi) - \varphi(\xi))^2 + \\ &\quad + 2E\{(\eta - E(\eta/\xi))(E(\eta/\xi) - \varphi(\xi))\}. \end{aligned}$$

Рассмотрим последнее слагаемое, воспользуемся третьим свойством условного математического ожидания:

$$\begin{aligned} E(E\{(\eta - E(\eta/\xi))(E(\eta/\xi) - \varphi(\xi))/\xi\}) &= \\ &= E[(E(\eta/\xi) - \varphi(\xi))E\{(\eta - E(\eta/\xi))/\xi\}]. \end{aligned}$$

Нетрудно проверить, что условное математическое ожидание обладает свойством линейности, следовательно,

$$\begin{aligned} E\{(\eta - E(\eta/\xi))/\xi\} &= E(\eta/\xi) - E\{E(\eta/\xi)/\xi\} = E(\eta/\xi) - \\ &- E(\eta/\xi)E\{1/\xi\} = E(\eta/\xi) - E(\eta/\xi) \stackrel{\text{п.н.}}{=} 0. \end{aligned}$$

Таким образом, последнее слагаемое равно нулю. Итак, получили равенство:

$$E(\eta - \varphi(\xi))^2 = E(\eta - E(\eta/\xi))^2 + E(E(\eta/\xi) - \varphi(\xi))^2. \quad (1.4)$$

Так как все слагаемые положительные, то очевидно, что имеет место неравенство:

$$E(\eta - \varphi(\xi))^2 \geq E(\eta - E(\eta/\xi))^2.$$

Из (1.4) следует, что неравенство выполняется как равенство только, когда $\varphi(\xi) \stackrel{\text{п.н.}}{=} E(\eta/\xi)$. ■

Замечание 1.2. Доказанное свойство можно записать в виде:

$$E(\eta/\xi) = \arg \inf_{\varphi(\cdot)} E(\eta - \varphi(\xi))^2, \quad E(\varphi^2(\xi)) \in \mathbb{R}.$$

Следовательно, случайная величина $E(\eta/\xi)$ дает наилучший в среднеквадратическом смысле прогноз случайной величины η по случайной величине ξ . Ранее была рассмотрена задача о наилучшем линейном прогнозе.

§ 2. УСЛОВНЫЕ МАТЕМАТИЧЕСКИЕ ОЖИДАНИЯ ПРОИЗВОЛЬНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Определение 2.1. Пусть на вероятностном пространстве (Ω, \mathcal{F}, P) заданы две случайные величины $\xi(\omega)$, $\eta(\omega)$. Назовем функцию $P(B/x)$, где $B \in \mathcal{B}(\mathbb{R})$, $x \in \mathbb{R}$, условным

вероятностным распределением случайной величины η при условии $\xi = x$, если выполнены следующие условия:

1. При любом фиксированном $x \in \mathbb{R}$ функция $P(\cdot/x)$ как функция первого аргумента является вероятностной мерой на борелевской прямой $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
2. При любом фиксированном $B \in \mathcal{B}(\mathbb{R})$ функция $P(B/\cdot)$ как функция второго аргумента является борелевской функцией (т. е. $P(B/\cdot): \mathbb{R} \rightarrow \mathbb{R}$ и полный прообраз любого борелевского множества есть борелевское множество).
3. Для любых борелевских множеств $A \in \mathcal{B}(\mathbb{R})$, $B \in \mathcal{B}(\mathbb{R})$ справедливо равенство:

$$P\{\xi \in A, \eta \in B\} = \int_A P(B/x) dP_\xi. \quad (2.1)$$

Замечание 2.1. Положим в третьем пункте определения $A = \mathbb{R}$. Тогда событие $\xi \in \mathbb{R}$ является достоверным:

$$\begin{aligned} P\{\xi^{-1}(A) \cap \eta^{-1}(B)\} &= P\{\xi^{-1}(\mathbb{R}) \cap \eta^{-1}(B)\} = \\ &= P\{\Omega \cap \eta^{-1}(B)\} = P\{\eta \in B\}. \end{aligned}$$

Следовательно, имеет место равенство:

$$P\{\eta \in B\} = \int_{\mathbb{R}} P(B/x) dP_\xi.$$

Данное равенство является аналогом формулы полной вероятности.

Замечание 2.2. Определение 2.1 также справедливо в векторном случае: $\xi = (\xi_1, \dots, \xi_n)^T$, $\eta = (\eta_1, \dots, \eta_m)^T$, $x \in \mathbb{R}^n$, $B \in \mathcal{B}(\mathbb{R}^m)$.

Определение 2.2. Пусть при любом $x \in \mathbb{R}$ существует интеграл Лебега по вероятностной мере: $\int_{\mathbb{R}} y P(dy/x)$. Условным математическим ожиданием случайной величины η при условии $\xi = x$ назовем величину

$$E(\eta/\xi = x) = \int_{\mathbb{R}} y P(dy/x),$$

где $P(B/x)$ — условное распределение η при условии $\xi = x$.

Замечание 2.3. Сформулированное определение полностью согласуется с определением математического ожидания, если вспомнить теорему о замене переменной под знаком интеграла Лебега.

Замечание 2.4. Пусть $g: \mathbb{R} \rightarrow \mathbb{R}$ — борелевская функция. Тогда условным математическим ожиданием случайной величины $g(\eta)$ при условии $\xi = x$ называется

$$E(g(\eta)/\xi = x) = \int_{\mathbb{R}} g(y)P(dy/x),$$

если интеграл существует.

Если существует $E(\eta/\xi = x)$, то с учетом определения 2.1, а также с учетом определения интеграла Лебега получаем, что $E(\eta/\xi = x) = h(x)$ — борелевская функция.

Определение 2.3. Пусть для любого $x \in \mathbb{R}$ существует $E(\eta/\xi = x) = h(x)$. Назовем случайную величину $h(\xi) = E(\eta/\xi)$ условным математическим ожиданием случайной величины η относительно случайной величины ξ .

Во всех определениях требовалось выполнение условий для любых $x \in \mathbb{R}$, но достаточно требовать выполнение всех условий почти для всех $x \in \mathbb{R}$ относительно меры P_ξ .

Если условное математическое ожидание $E(\eta/\xi = x)$ существует и конечно почти при всех $x \in \mathbb{R}$ относительно меры P_ξ , и если существует и конечно математическое ожидание $E(\eta)$, то, как нетрудно показать, все свойства условного математического ожидания $E(\eta/\xi)$, рассмотренные в предыдущем параграфе, полностью переносятся на общий случай. В дальнейшем, пусть при любом фиксированном $x \in \mathbb{R}$ условное распределение $P(B/x)$ имеет плотность $f_{\eta/\xi}(y/x)$ относительно меры Лебега, т. е. справедливо равенство:

$$P(B/x) = \int_B f_{\eta/\xi}(y/x)dy$$

для любого $B \in \mathcal{B}(\mathbb{R})$, $f_{\eta/\xi}(y/x) \geq 0$ для любых $y \in \mathbb{R}$, $x \in \mathbb{R}$.

Тогда справедливо следующее представление:

$$\begin{aligned} P\{\xi \in A, \eta \in B\} &= \int_A P(B/x) P_\xi(dx) = \\ &= \int_A \int_B f_{\eta/\xi}(y/x) dy P_\xi(dx). \end{aligned} \quad (2.2)$$

Кроме того, будем предполагать, что существует плотность распределения $f_\xi(x)$ относительно меры Лебега, т. е.

$$P_\xi(A) = \int_A f_\xi(x) dx$$

для любого $A \in \mathcal{B}(\mathbb{R})$. Тогда равенство (2.2) можно переписать следующим образом:

$$P\{\xi \in A, \eta \in B\} = \int_A \int_B f_{\eta/\xi}(y/x) f_\xi(x) dy dx. \quad (2.3)$$

Будем называть плотность $f_{\eta/\xi}(y/x)$ условной плотностью распределения случайной величины η при условии $\xi = x$.

Из равенства (2.3) следует, что

$$P\{(\xi, \eta) \in A \times B\} = \int_{A \times B} f_{\eta/\xi}(y/x) f_\xi(x) dx dy$$

для любых $A \in \mathcal{B}(\mathbb{R})$, $B \in \mathcal{B}(\mathbb{R})$. Тогда для любого $D \in \mathcal{B}(\mathbb{R}^2)$ справедливо

$$P\{(\xi, \eta) \in D\} = \int_D f_{\eta/\xi}(y/x) f_\xi(x) dx dy.$$

Таким образом, доказана следующая теорема.

Теорема 2.1. Пусть $f_{\eta/\xi}(y/x)$ — условная плотность распределения случайной величины η при условии $\xi = x$. Пусть $f_\xi(x)$ — плотность распределения случайной величины ξ .

Тогда совместная плотность распределения $f_{\xi,\eta}(x,y)$ случайных величин ξ, η может быть представлена в виде:

$$f_{\xi,\eta}(x,y) = f_{\eta/\xi}(y/x)f_{\xi}(x).$$

Аналогично $f_{\xi,\eta}(x,y) = f_{\xi/\eta}(x/y)f_{\eta}(y)$.

Справедлива и обратная теорема.

Теорема 2.2. Пусть $f_{\xi,\eta}(x,y)$ — совместная плотность распределения случайных величин ξ, η . Тогда имеют место равенства:

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy, \quad (2.4)$$

$$f_{\eta}(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx, \quad (2.5)$$

$$f_{\eta/\xi}(y/x) = \frac{f_{\xi,\eta}(x,y)}{f_{\xi}(x)}, \quad f_{\xi/\eta}(x/y) = \frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}.$$

Доказательство. Равенства (2.4) и (2.5) были доказаны ранее. Справедливо равенство

$$P\{\xi \in A, \eta \in B\} = \int_A \int_B f_{\xi,\eta}(x,y)dydx.$$

В правой его части выражение под знаком интеграла умножим и разделим на величину $f_{\xi}(x)$ и перейдем к повторному интегрированию:

$$\begin{aligned} P\{\xi \in A, \eta \in B\} &= \int_A \int_B f_{\xi,\eta}(x,y) \frac{f_{\xi}(x)}{f_{\xi}(x)} dydx = \\ &= \int_A \left(\int_B \frac{f_{\xi,\eta}(x,y)}{f_{\xi}(x)} dy \right) f_{\xi}(x) dx. \end{aligned}$$

Выражение в круглых скобках — функция, зависящая от x и от B . В силу определения 2.1 эта функция — условное распределение случайной величины η при условии $\xi = x$. Таким образом, $f_{\xi,\eta}(x,y)/f_{\xi}(x)$ — плотность меры $P(B/x)$ относительно меры Лебега. ■

Так как плотность $f_\xi(x)$ может принимать нулевое значение, то всегда ли корректно делить на $f_\xi(x)$? Пусть $N = \{x : f_\xi(x) = 0\}$, тогда

$$P_\xi(N) = \int_N f_\xi(x) dx = 0.$$

Таким образом, функция $f_{\eta/\xi}(y/x)$ не определена на множестве меры нуль, но интегрирование производится по мере P_ξ , следовательно, функцию можно доопределить на этом множестве произвольным образом, что не окажет никакого влияния на правую часть рассматриваемого в теореме выражения.

§ 3. МНОГОМЕРНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Пусть $\xi_1 \sim N(0, 1)$, $\xi_2 \sim N(0, 1)$, ..., $\xi_n \sim N(0, 1)$. Предположим, что все случайные величины взаимно независимы. Составим случайный вектор $\xi = (\xi_1, \dots, \xi_n)^T$. Как было установлено, необходимым и достаточным условием независимости случайных величин является равенство:

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \prod_{i=1}^n F_{\xi_i}(x_i),$$

которое должно выполняться для любого $x \in \mathbb{R}^n$. Если есть совместная плотность, то необходимое и достаточное условие имеет вид:

$$f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \prod_{i=1}^n f_{\xi_i}(x_i).$$

В случае стандартного нормального распределения:

$$f_{\xi_i}(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}.$$

Следовательно,

$$f_{\xi_1, \dots, \xi_n}(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}x^T x}, \quad x^T = (x_1, \dots, x_n).$$

Отсюда следует, что величины ξ_1, \dots, ξ_n подчиняются многомерному нормальному распределению $\xi \sim N(0, E_n)$, где $\xi^T = (\xi_1, \dots, \xi_n)$.

Рассмотрим случайный вектор:

$$\eta = (\eta_1, \dots, \eta_n)^T = a + c\xi,$$

где $a = (a_1, \dots, a_n)^T \in \mathbb{R}^n$, c — квадратная матрица, $|c| > 0$.

Возьмем любое множество $B \in \mathcal{B}(\mathbb{R}^n)$ и рассмотрим случайный вектор $\eta \in B$. Тогда существует множество D , что следующие события равносильны $\{\eta \in B\} = \{\xi \in D\}$. Так как $y \in B$, тогда и только тогда, когда $x = c^{-1}(y - a) \in D$, то

$$P\{\eta \in B\} = P\{\xi \in D\} = \int_D f_\xi(x) dx.$$

Сделаем следующую замену переменной: $x = c^{-1}(y - a)$, тогда

$$\int_D f_\xi(x) dx = \int_B \frac{1}{|c|} f_\xi(c^{-1}(y - a)) dy, \quad B \in \mathcal{B}(\mathbb{R}^n).$$

Следовательно, плотность распределения случайного вектора η имеет вид:

$$f_\eta(y) = \frac{1}{|c|} f_\xi(c^{-1}(y - a)) = \frac{1}{(2\pi)^{\frac{n}{2}} |c|} e^{-\frac{1}{2}(y-a)^T ((c^{-1})^T c^{-1})(y-a)},$$

где $(c^{-1})^T c^{-1} = (cc^T)^{-1}$, $\Sigma = cc^T$, $|\Sigma| = |c|^2$. Тогда можно записать выражение для плотности распределения вектора η в следующем виде:

$$f_\eta(y) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\Sigma|}} e^{-\frac{1}{2}(y-a)^T \Sigma^{-1}(y-a)}.$$

Назовем эту плотность плотностью многомерного нормального распределения $N(a, \Sigma)$.

Найдем математическое ожидание вектора $\eta = a + c\xi$:

$$E\eta = \begin{pmatrix} E\eta_1 \\ \vdots \\ E\eta_n \end{pmatrix} = a + cE\xi = a.$$

Найдем ковариационную матрицу:

$$D\eta = E\{(\eta - E\eta)(\eta - E\eta)^T\} = E\{c\xi\xi^T c^T\} = cE(\xi\xi^T)c^T,$$

здесь $E\xi\xi^T$ — ковариационная матрица случайного вектора ξ . По предположению справедливо равенство:

$$E(\xi\xi^T) = E_n.$$

Тогда ковариационная матрица вектора η равна:

$$D\eta = E\{(\eta - E\eta)(\eta - E\eta)^T\} = E\{c\xi\xi^T c^T\} = cE(\xi\xi^T)c^T = \Sigma.$$

Далее будем рассматривать двумерное нормальное распределение случайного вектора (ξ, η) . В дальнейшем удобно обозначить первую компоненту через ξ , а вторую компоненту через η . Пусть случайный вектор $(\xi, \eta)^T \sim N(a, \Sigma)$, где $a^T = (a_\xi, a_\eta)$, $a_\xi = E\xi$, $a_\eta = E\eta$. Плотность распределения случайного вектора (ξ, η) имеет вид:

$$f_{\xi, \eta}(x, y) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-a_\xi, y-a_\eta)\Sigma^{-1}\begin{pmatrix} x-a_\xi \\ y-a_\eta \end{pmatrix}},$$

$$\Sigma = D(\xi, \eta) = \begin{pmatrix} \sigma_\xi^2 & cov(\xi, \eta) \\ cov(\xi, \eta) & \sigma_\eta^2 \end{pmatrix}.$$

Ковариацию случайных величин $cov(\xi, \eta)$ можно записать в виде $\sigma_\xi\sigma_\eta\rho$, где $\rho = \rho(\xi, \eta)$ — коэффициент корреляции случайных величин ξ и η . Тогда справедливо равенство:

$$\begin{pmatrix} \sigma_\xi^2 & cov(\xi, \eta) \\ cov(\xi, \eta) & \sigma_\eta^2 \end{pmatrix} = \begin{pmatrix} \sigma_\xi^2 & \sigma_\xi\sigma_\eta\rho \\ \sigma_\xi\sigma_\eta\rho & \sigma_\eta^2 \end{pmatrix}.$$

Найдем $f_\xi(x)$, $f_\eta(y)$, не предполагая независимости случайных величин ξ и η , также найдем плотности условных распределений $f_{\xi/\eta}(x/y)$, $f_{\eta/\xi}(y/x)$ и условное математическое ожидание $E(\eta/\xi)$. Нетрудно видеть, что

$$\begin{aligned}\Sigma^{-1} &= \frac{1}{\sigma_\xi^2 \sigma_\eta^2 (1 - \rho^2)} \begin{pmatrix} \sigma_\eta^2 & -\sigma_\xi \sigma_\eta \rho \\ -\sigma_\xi \sigma_\eta \rho & \sigma_\xi^2 \end{pmatrix} = \\ &= \frac{1}{\sigma_\eta^2 (1 - \rho^2)} \begin{pmatrix} \frac{\sigma_\eta^2}{\sigma_\xi^2} & -\frac{\sigma_\eta}{\sigma_\xi} \rho \\ -\frac{\sigma_\eta}{\sigma_\xi} \rho & 1 \end{pmatrix}.\end{aligned}$$

Запишем квадратичную форму:

$$\begin{aligned}(x - a_\xi, y - a_\eta) \Sigma^{-1} \begin{pmatrix} x - a_\xi \\ y - a_\eta \end{pmatrix} &= \frac{1}{\sigma_\eta^2 (1 - \rho^2)} \times \\ &\times \left(\frac{\sigma_\eta^2}{\sigma_\xi^2} (x - a_\xi)^2 - 2 \frac{\sigma_\eta}{\sigma_\xi} \rho (x - a_\xi)(y - a_\eta) + (y - a_\eta)^2 \right) = \\ &= \frac{1}{\sigma_\eta^2 (1 - \rho^2)} \left\{ (y - a_\eta - \frac{\sigma_\eta}{\sigma_\xi} \rho (x - a_\xi))^2 + \right. \\ &\quad \left. + \frac{\sigma_\eta^2}{\sigma_\xi^2} (1 - \rho^2) (x - a_\xi)^2 \right\}.\end{aligned}$$

Отсюда следует

$$\begin{aligned}f_{\xi, \eta}(x, y) &= \frac{1}{\sqrt{2\pi} \sqrt{\sigma_\eta^2 (1 - \rho^2)}} \times \\ &\times \exp \left(-\frac{(y - a_\eta - \frac{\sigma_\eta}{\sigma_\xi} \rho (x - a_\xi))^2}{2\sigma_\eta^2 (1 - \rho^2)} \right) \cdot \\ &\cdot \frac{1}{\sqrt{2\pi} \sigma_\xi} \exp \left(-\frac{(x - a_\xi)^2}{2\sigma_\xi^2} \right) = f_{\eta/\xi}(y/x) f_\xi(x),\end{aligned}$$

где $f_\xi(x) = \frac{1}{\sqrt{2\pi} \sigma_\xi} \exp \left(-\frac{(x - a_\xi)^2}{2\sigma_\xi^2} \right)$ — плотность нормального распределения $N(a_\xi, \sigma_\xi^2)$, $f_{\eta/\xi}(y/x)$ — плотность условного

распределения, которое также является нормальным распределением. Очевидно, что

$$\int_{-\infty}^{\infty} f_{\eta/\xi}(y/x) dy = 1.$$

Следовательно,

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x, y) dy.$$

Таким образом, если совместное распределение компонент случайного вектора является нормальным, тогда и любая из компонент подчиняется нормальному распределению (независимо от того, зависимы они или независимы). Аналогичные выражения можно получить для $f_{\eta}(y)$ и $f_{\xi/\eta}(x/y)$.

Так как условное распределение является также нормальным $N(a_{\eta} + \rho \frac{\sigma_{\eta}}{\sigma_{\xi}}(x - a_{\xi}), \sigma_{\eta}^2(1 - \rho^2))$, то

$$E(\eta/\xi = x) = \int_{-\infty}^{\infty} y f_{\eta/\xi}(y/x) dy = a_{\eta} + \rho \frac{\sigma_{\eta}}{\sigma_{\xi}}(x - a_{\xi}).$$

Следовательно,

$$E(\eta/\xi) = a_{\eta} + \rho \frac{\sigma_{\eta}}{\sigma_{\xi}}(\xi - a_{\xi}).$$

Из полученного выражения можно сделать вывод: если совместное распределение случайных величин ξ и η нормальное, то наилучший прогноз случайной величины η по случайной величине ξ является линейным.

Необходимое и достаточное условие независимости случайных величин ξ и η , если существует совместная плотность, имеет вид:

$$f_{\xi,\eta}(x, y) = f_{\xi}(x) f_{\eta}(y).$$

Для двумерного нормального распределения, как следует из полученных формул, тождество возможно тогда и только тогда, когда коэффициент корреляции $\rho = 0$.

НЕКОТОРЫЕ РАСПРЕДЕЛЕНИЯ

§ 1. ДИСКРЕТНЫЕ РАСПРЕДЕЛЕНИЯ

1. *Распределение Бернулли.* Случайная величина ξ подчиняется распределению Бернулли, если

$$P\{\xi = 1\} = p, \quad P\{\xi = 0\} = q = 1 - p,$$

$$E\xi = p, \quad D\xi = E\xi^2 - (E\xi)^2 = p - p^2 = pq.$$

2. *Биномиальное распределение* (рис. 8.1, 8.2).

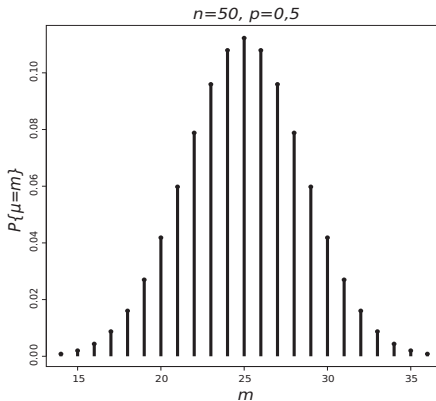


Рис. 8.1

Биномиальное распределение

Пусть μ — число успехов в серии из n испытаний Бернулли, с каждым испытанием можно связать случайную величину:

$$\xi_i = \begin{cases} 1, & \text{в испытании с номером } i \text{ произошел успех;} \\ 0, & \text{в испытании с номером } i \text{ произошла неудача.} \end{cases}$$

Тогда μ можно представить в виде:

$$\mu = \sum_{i=1}^n \xi_i.$$

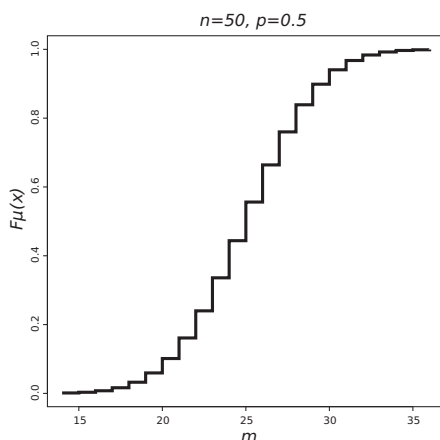


Рис. 8.2

Функция биномиального распределения

Случайная величина μ подчиняется биномиальному распределению:

$$P\{\mu = m\} = C_n^m p^m q^{n-m}, \quad m = 0, 1, \dots, n.$$

Математическое ожидание, дисперсия и среднеквадратическое отклонение случайной величины μ равны соответственно:

$$E\mu = np, \quad D\mu = npq, \quad \sigma_\mu = \sqrt{npq}.$$

3. *Распределение Пуассона* (рис. 8.3, 8.4). Случайная величина ξ подчиняется распределению Пуассона, если

$$P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda > 0, \quad k = 0, 1, \dots$$

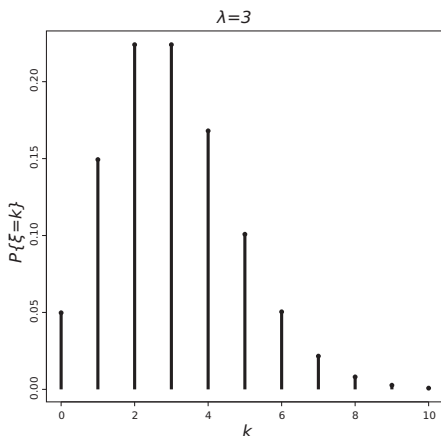


Рис. 8.3
Распределение Пуассона

Вычислим математическое ожидание случайной величины ξ :

$$E\xi = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Вычислим дисперсию $D\xi$. Воспользуемся следующим приемом:

$$\begin{aligned} E\{\xi(\xi - 1)\} &= \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=2}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \\ &= \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} e^{-\lambda} = \lambda^2. \end{aligned}$$

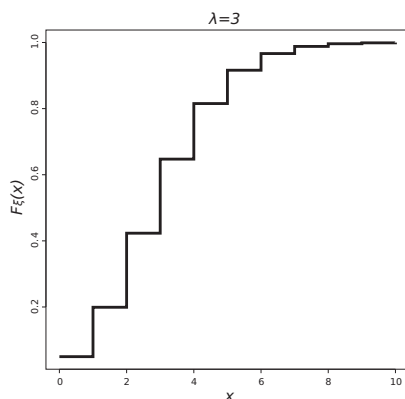


Рис. 8.4
Функция распределения Пуассона

С другой стороны, справедливы равенства:

$$E\{\xi(\xi - 1)\} = E\{\xi^2 - \xi\} = E\xi^2 - E\xi = \lambda^2.$$

Следовательно,

$$E\xi^2 = \lambda^2 + \lambda,$$

тогда дисперсия может быть вычислена следующим образом:

$$D\xi = E\xi^2 - (E\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

§ 2. НЕПРЕРЫВНЫЕ РАСПРЕДЕЛЕНИЯ

1. *Стандартное нормальное распределение* $N(0, 1)$ (рис. 8.5, 8.6). Рассмотрим плотность распределения

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Вычислим математическое ожидание случайной величины ξ :

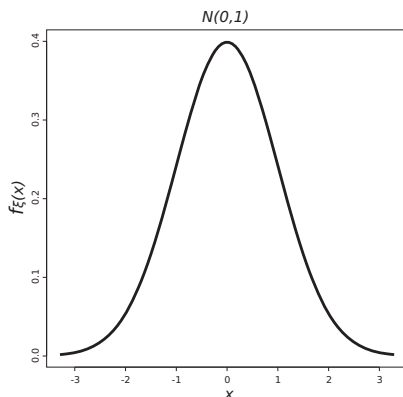


Рис. 8.5

Функция плотности стандартного
нормального распределения

$$\begin{aligned}
 E\xi &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\
 &= \int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - \int_{-\infty}^0 |x| e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} dx = \\
 &= \int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - \int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0.
 \end{aligned}$$

Вычислим дисперсию:

$$\begin{aligned}
 D\xi = E\xi^2 &= \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\
 &= -x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x^2}{2}} dx = 1,
 \end{aligned}$$

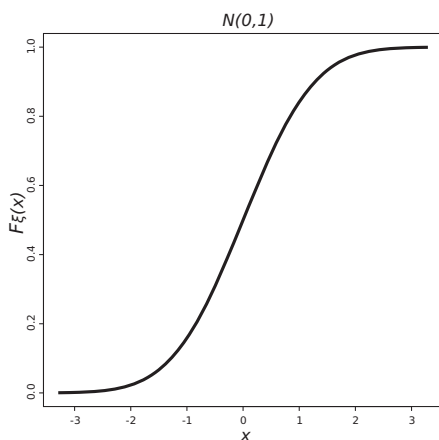


Рис. 8.6

Функция стандартного нормального
распределения

следовательно, $\sigma_\xi = 1$. Вычислим коэффициент асимметрии:

$$\begin{aligned}
 K_a = E\xi^3 &= \int_{-\infty}^{\infty} x^3 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\
 &= -x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + 2 \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0.
 \end{aligned}$$

Коэффициент асимметрии для стандартного нормального распределения равен нулю.

Найдем коэффициент эксцесса:

$$\begin{aligned} K_e &= E\xi^4 - 3 = \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - 3 = \\ &= -x^3 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + 3 \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx - 3 = 3 - 3 = 0. \end{aligned}$$

Коэффициент эксцесса K_e характеризует островершинность распределения.

Коэффициенты K_a и K_e подобраны таким образом, чтобы стандартное нормальное распределение служило шаблоном для других распределений.

Для нормально распределенной случайной величины ξ справедливо *правило трех сигм*:

$$P\{-3 < \xi < 3\} > 0,98.$$

2. *Нормальное распределение* $N(a, \sigma^2)$ (см. рис. 8.7, 8.8). Пусть ξ подчиняется стандартному нормальному распределению, $\xi \sim N(0, 1)$. Рассмотрим следующую случайную величину:

$$\eta = a + \sigma\xi, \quad a \in \mathbb{R}, \quad \sigma > 0.$$

Найдем функцию $F_\eta(x)$ и плотность распределения $f_\eta(x)$:

$$\begin{aligned} F_\eta(x) &= P\{\eta \leq x\} = P\{a + \sigma\xi \leq x\} = P\left\{\xi \leq \frac{x-a}{\sigma}\right\} = \\ &= F_\xi\left(\frac{x-a}{\sigma}\right) = \int_{-\infty}^{\frac{(x-a)}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-a)^2}{2\sigma^2}} dz. \end{aligned}$$

Следовательно,

$$f_\eta(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

здесь $f_\eta(x)$ — плотность нормального распределения $N(a, \sigma^2)$.

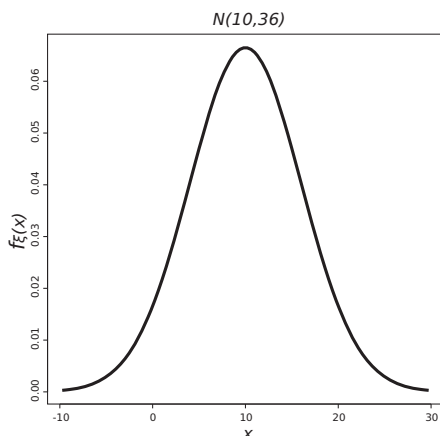


Рис. 8.7

Функция плотности нормального распределения $N(10, 36)$

Вычислим математическое ожидание и дисперсию случайной величины η :

$$E\eta = a + \sigma E\xi = a,$$

$$D\eta = E(\eta - E\eta)^2 = \sigma^2 D\xi = \sigma^2.$$

Коэффициенты в функции плотности распределения имеют следующий смысл: a — математическое ожидание, σ — среднеквадратическое отклонение, σ^2 — дисперсия. Если случайная величина подчиняется распределению $\eta \sim N(a, \sigma^2)$, то случайная величина $\xi = (\eta - a)/\sigma$ подчиняется стандартному нормальному распределению $N(0, 1)$, следовательно, можно переформулировать свойства, полученные ранее для стандартного нормального распределения.

Например, *правило трех сигм* можно записать в виде:

$$\begin{aligned} P\{|\xi| < 3\} &= P\left\{\left|\frac{\eta - a}{\sigma}\right| < 3\right\} = \\ &= P\{-3\sigma < \eta - a < 3\sigma\} > 0,98. \end{aligned}$$

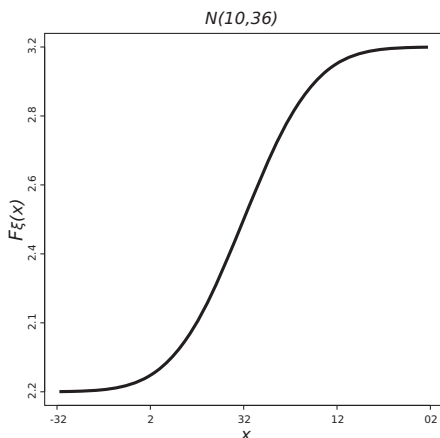


Рис. 8.8

Функция нормального распределения $N(10, 36)$

Вычислим коэффициенты асимметрии и эксцесса:

$$K_a = \frac{E(\eta - a)^3}{\sigma_\eta^3} = E\left(\frac{\eta - a}{\sigma}\right)^3 = E\xi^3 = 0,$$

$$K_e = \frac{E(\eta - E\eta)^4}{\sigma_\eta^4} - 3 = E\left(\frac{\eta - a}{\sigma}\right)^4 - 3 = E\xi^4 - 3 = 0.$$

Коэффициенты K_a и K_e равны нулю для любой нормально распределенной случайной величины.

3. *Гамма-распределение* (см. рис. 8.9, 8.10). Случайная величина ξ подчиняется гамма-распределению с параметрами $p > 0$, $\lambda > 0$, т. е. $\xi \sim G(p, \lambda)$, если плотность распределения имеет вид:

$$f_\xi(x) = \begin{cases} 0, & x \leq 0; \\ \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x}, & x > 0, \end{cases}$$

где p — параметр формы, λ — параметр масштаба, $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$ — гамма-функция.

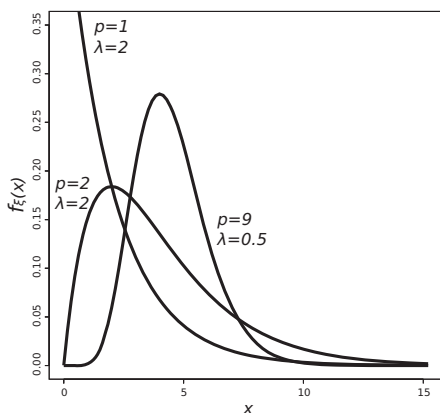


Рис. 8.9

Функции плотности гамма-распределения

Если $p = 1$, то гамма-распределение называется экспоненциальным (показательным) распределением.

Математическое ожидание случайной величины ξ , подчиняющейся гамма-распределению с параметрами p и λ равно

$$\begin{aligned}
 E\xi &= \int_0^{\infty} x f_{\xi}(x) dx = \frac{-\lambda^{p-1}}{\Gamma(p)} x^p e^{-\lambda x} \Big|_0^{\infty} + \\
 &\quad + \frac{p}{\lambda} \int_0^{\infty} \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} dx = \frac{p}{\lambda},
 \end{aligned}$$

так как $\int_0^{\infty} \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} dx = 1$. Найдём дисперсию случайной величины ξ :

$$D\xi = E\xi^2 - (E\xi)^2,$$

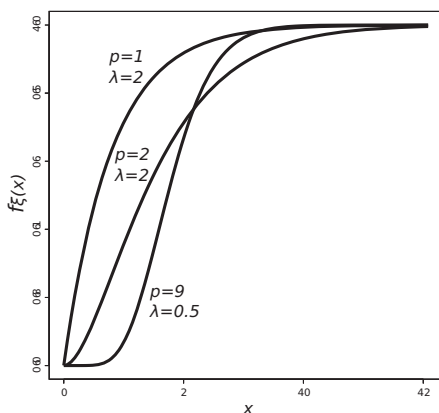


Рис. 8.10
Функции гамма-распределения

где первое слагаемое $E\xi^2$ может быть вычислено следующим образом:

$$\begin{aligned}
 E\xi^2 &= \int_0^{\infty} x^2 \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} dx = -\frac{\lambda^p}{\lambda \Gamma(p)} x^2 x^{p-1} e^{-\lambda x} \Big|_0^{\infty} + \\
 &+ (p+1) \frac{1}{\lambda} \int_0^{\infty} x \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} dx = \frac{p(p+1)}{\lambda^2}.
 \end{aligned}$$

Следовательно, получаем, что

$$D\xi = \frac{p}{\lambda^2}.$$

4. Экспоненциальное распределение (см. рис. 8.11, 8.12), то есть гамма-распределение с $p = 1$, $G(\lambda, 1)$. Плотность распределения экспоненциально распределенной случайной величины ξ имеет вид:

$$f_{\xi}(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

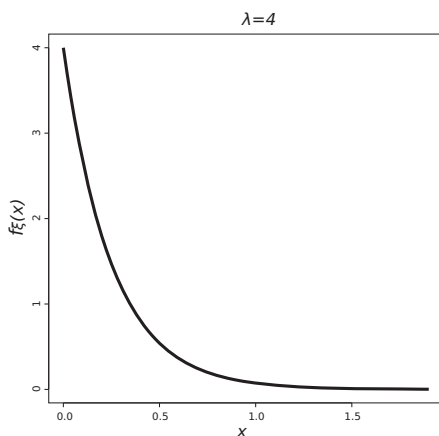


Рис. 8.11

Функция плотности экспоненциального распределения

Математическое ожидание и дисперсия соответственно равны

$$E\xi = \frac{1}{\lambda},$$

$$D\xi = \frac{1}{\lambda^2}.$$

Вычислим вероятность того, что $\xi \leq x$:

$$P\{\xi \leq x\} = \lambda \int_0^x e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x},$$

следовательно,

$$F_\xi(x) = P\{\xi \leq x\} = \begin{cases} 1 - e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Часто аргумент x представляет собой время, поэтому будем обозначать его t , тогда

$$F_\xi(t) = 1 - e^{-\lambda t}, \quad t > 0.$$

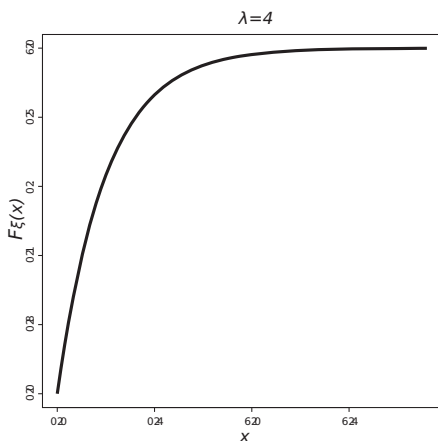


Рис. 8.12

Функция экспоненциального распределения

Найдем условную вероятность:

$$\begin{aligned}
 P\{\xi \leq (t + \tau) / \xi > \tau\} &= 1 - P\{\xi > t + \tau / \xi > \tau\} = \\
 &= 1 - \frac{P\{(\xi > t + \tau) \cap (\xi > \tau)\}}{P\{\xi > \tau\}} = 1 - \frac{P\{\xi > \tau + t\}}{P\{\xi > \tau\}} = \\
 &= 1 + \frac{1 - F_\xi(t + \tau)}{1 - F_\xi(\tau)} = 1 - \frac{e^{-\lambda(t + \tau)}}{e^{-\lambda\tau}} = 1 - e^{-\lambda t} = F_\xi(t).
 \end{aligned}$$

Таким образом, доказано следующее свойство экспоненциального распределения:

$$P\{\xi \leq t + \tau / \xi > \tau\} = P\{\xi \leq t\}.$$

Доказанное свойство называется свойством отсутствия последствия.

5. *Бета-распределение* (см. рис. 8.13, 8.14). Функцию распределения будем обозначать $B(x; p_1, p_2)$, где $p_1 > 0$, $p_2 > 0$.

Определим плотность бета-распределения:

$$f_{\xi}(x) = \begin{cases} 0, & x \notin (0, 1); \\ \frac{x^{p_1-1}(1-x)^{p_2-1}}{\beta(p_1, p_2)}, & x \in (0, 1), \end{cases}$$

где

$$\beta(p_1, p_2) = \int_0^1 x^{p_1-1}(1-x)^{p_2-1} dx$$

— бета-функция.

Свойства бета-функции:

1. Имеет место равенство: $\beta(p_1, p_2) = \beta(p_2, p_1)$. Нетрудно заметить, что

$$\begin{aligned} \int_0^1 x^{p_1-1}(1-x)^{p_2-1} dx &= - \int_1^0 (1-t)^{p_1-1} t^{p_2-1} dt = \\ &= \int_0^1 t^{p_2-1}(1-t)^{p_1-1} dt. \end{aligned}$$

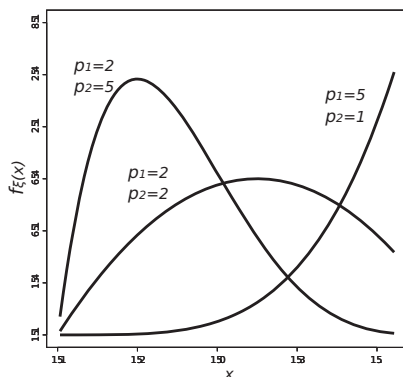


Рис. 8.13

Функция плотности бета-распределения

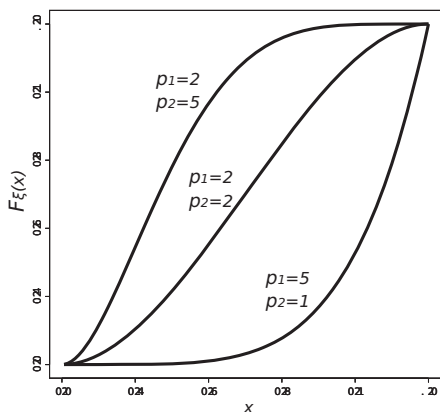


Рис. 8.14
Функция бета-распределения

2. Справедлива формула понижения порядка:

$$\beta(p_1 + 1, p_2) = \frac{p_1}{p_1 + p_2} \beta(p_1, p_2).$$

Заметим, что

$$\begin{aligned} \beta(p_1 + 1, p_2) &= \int_0^1 x^{p_1} (1-x)^{p_2-1} dx = -\frac{1}{p_2} x^{p_1} (1-x)^{p_2} \Big|_0^1 + \\ &+ \frac{p_1}{p_2} \int_0^1 x^{p_1-1} (1-x)^{p_2} dx = \frac{p_1}{p_2} (\beta(p_1, p_2) - \beta(p_1 + 1, p_2)). \end{aligned}$$

Следовательно,

$$\frac{p_2 + p_1}{p_1} \beta(p_1 + 1, p_2) = \beta(p_1, p_2).$$

Вычислим моменты:

$$E\xi = \int_0^1 x^{p_1} (1-x)^{p_2-1} \frac{1}{\beta(p_1, p_2)} dx = \frac{\beta(p_1 + 1, p_2)}{\beta(p_1, p_2)} = \frac{p_1}{p_1 + p_2}.$$

$$\begin{aligned}
 E\xi^2 &= \frac{1}{\beta(p_1, p_2)} \int_0^1 x^{p_1+1} (1-x)^{p_2-1} dx = \frac{\beta(p_1+2, p_2)}{\beta(p_1, p_2)} = \\
 &= \frac{\beta(p_1+2, p_2)}{\beta(p_1+1, p_2)} \frac{\beta(p_1+1, p_2)}{\beta(p_1, p_2)} = \frac{p_1+1}{p_1+p_2+1} \cdot \frac{p_1}{p_1+p_2}.
 \end{aligned}$$

$$\begin{aligned}
 D\xi &= E\xi^2 - (E\xi)^2 = \frac{p_1}{p_1+p_2} \left(\frac{p_1+1}{p_1+p_2+1} - \frac{p_1}{p_1+p_2} \right) = \\
 &= \frac{p_1 p_2}{(p_1+p_2)^2 (p_1+p_2+1)}.
 \end{aligned}$$

6. *Равномерное распределение на $[0, 1]$* (рис. 8.15, 8.16). Равномерное распределение — частный случай бета-распределения, когда $p_1 = 1$, $p_2 = 1$, т. е. $B(x; 1, 1)$. Равномерное распределение обозначают $U[0, 1]$. Плотность равномерного распределения имеет вид:

$$f_\xi(x) = \begin{cases} 1, & x \in [0, 1]; \\ 0, & x \notin [0, 1]. \end{cases}$$

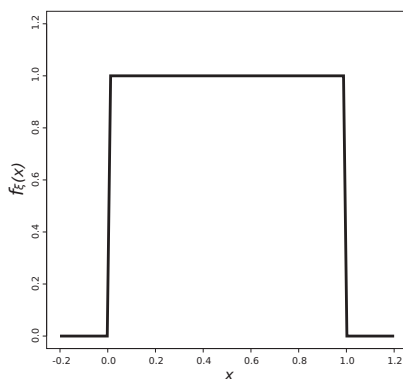


Рис. 8.15

Функция плотности равномерного распределения $U[0, 1]$

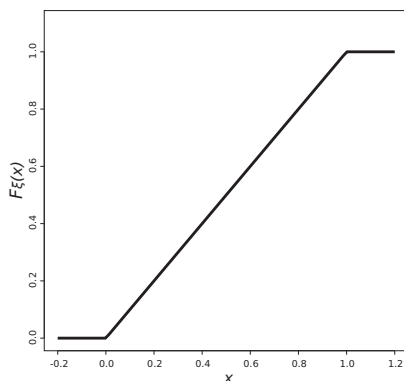


Рис. 8.16

Функция равномерного распределения $U[0, 1]$

Рассмотрим точки $\alpha, \alpha', \beta, \beta'$ такие, что $\beta - \alpha = \beta' - \alpha'$, при этом $(\alpha, \beta) \subset [0, 1]$, и $(\alpha', \beta') \subset [0, 1]$.

Вычислим вероятность:

$$P\{\xi \in (\alpha, \beta)\} = \int_{(\alpha, \beta)} 1 dx = \int_{\alpha}^{\beta} dx = \beta - \alpha,$$

аналогично получаем:

$$P\{\xi \in (\alpha', \beta')\} = \beta' - \alpha'.$$

Таким образом, вероятности событий $\xi \in (\alpha, \beta)$ и $\xi \in (\alpha', \beta')$ одинаковы.

7. *Распределение Коши* (см. рис. 8.17). Будем говорить, что случайная величина ξ подчиняется распределению Коши, если плотность распределения имеет вид:

$$f_{\xi}(x) = \frac{1}{\pi} \frac{1}{1 + x^2}, \quad x \in \mathbb{R}.$$

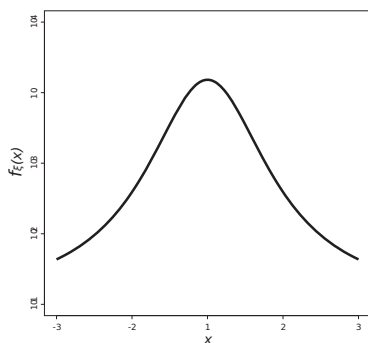


Рис. 8.17
Функция плотности распределения Коши

Распределение Коши имеет «тяжелые» хвосты, вследствие чего

$$E\xi = \int_{-\infty}^{\infty} x \frac{1}{\pi} \frac{1}{1+x^2} dx = \int_0^{\infty} x \frac{1}{\pi} \frac{1}{1+x^2} dx - \int_{-\infty}^0 |x| \frac{1}{\pi} \frac{1}{1+x^2} dx,$$

здесь каждый из интегралов равен $+\infty$, следовательно, математическое ожидание $E\xi$ не существует.

ВИДЫ СХОДИМОСТЕЙ СЛУЧАЙНЫХ ВЕЛИЧИН

§ 1. СХОДИМОСТИ ПО ВЕРОЯТНОСТИ, ПОЧТИ НАВЕРНОЕ, ПО РАСПРЕДЕЛЕНИЮ, В СРЕДНЕМ, В ОСНОВНОМ

Будем считать, что в вероятностном пространстве (Ω, \mathcal{F}, P) сигма-алгебра \mathcal{F} полна относительно меры P .

Определение 1.1. Пусть случайная величина ξ и последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ заданы на вероятностном пространстве (Ω, \mathcal{F}, P) . Будем говорить, что последовательность $\{\xi_n\}$ сходится почти наверное к случайной величине ξ и записывать:

$$\xi_n \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \xi, \quad (1.1)$$

если $P \left\{ \omega : \xi_n(\omega) \not\xrightarrow[n \rightarrow \infty]{} \xi(\omega) \right\} = 0$ или

$$P \left\{ \omega : \lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega) \right\} = 1.$$

Определение 1.2. Пусть случайная величина ξ и последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ заданы на вероятностном пространстве (Ω, \mathcal{F}, P) . Будем говорить, что последовательность $\{\xi_n\}$ сходится по вероятности к случайной величине ξ и записывать:

$$\xi_n \xrightarrow[n \rightarrow \infty]{P} \xi, \quad (1.2)$$

если для любого $\varepsilon > 0$ имеет место сходимость

$$P \{ |\xi_n - \xi| > \varepsilon \} \xrightarrow[n \rightarrow \infty]{} 0 \text{ или } P \{ |\xi_n - \xi| \leq \varepsilon \} \xrightarrow[n \rightarrow \infty]{} 1.$$

Определение 1.3. Пусть случайная величина ξ и последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ заданы на вероятностном пространстве (Ω, \mathcal{F}, P) . Будем говорить, что последовательность $\{\xi_n\}$ сходится в среднем порядка r к случайной величине ξ и записывать:

$$\xi_n \xrightarrow[n \rightarrow \infty]{L^r} \xi, \quad r > 1, \quad (1.3)$$

если $E|\xi_n - \xi|^r \xrightarrow[n \rightarrow \infty]{} 0$, $E\xi^r \in \mathbb{R}$, $E\xi_n^r \in \mathbb{R}$ при любом n .

Определение 1.4. Пусть заданы случайная величина ξ и последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$. Будем говорить, что последовательность $\{\xi_n\}$ сходится по распределению к случайной величине ξ и записывать:

$$\xi_n \xrightarrow[n \rightarrow \infty]{d} \xi, \quad (1.4)$$

если для любой непрерывной и ограниченной функции $g(x)$ выполнено $Eg(\xi_n) \xrightarrow[n \rightarrow \infty]{} Eg(\xi)$.

Замечание 1.1. В определении 1.4 не требуется, чтобы все случайные величины были заданы на одном и том же вероятностном пространстве.

Действительно, используя теорему о замене переменной под знаком интеграла, получаем

$$\int_{-\infty}^{\infty} g(x) dF_{\xi_n}(x) \xrightarrow[n \rightarrow \infty]{} \int_{-\infty}^{\infty} g(x) dF_{\xi}(x).$$

Интегралы зависят от случайных величин только через закон распределения.

Определение сходимости по распределению можно обобщить на случайные вектора $\xi_n = (\xi_1^{(n)}, \dots, \xi_m^{(n)})$. Будем говорить, что последовательность случайных векторов ξ_n сходится по распределению к случайному вектору ξ и записывать:

$$\xi_n = (\xi_1^{(n)}, \dots, \xi_m^{(n)}) \xrightarrow[n \rightarrow \infty]{d} \xi = (\xi_1, \dots, \xi_m), \quad (1.5)$$

если для любой непрерывной ограниченной функции $g(x_1, \dots, x_m)$ выполнено условие: $Eg(\xi_n) \xrightarrow{n \rightarrow \infty} Eg(\xi)$.

Определение 1.5. Пусть задана последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$. Пусть $F_\xi(x) = P\{\xi \leq x\}$ — функция распределения случайной величины ξ , $F_{\xi_n}(x) = P\{\xi_n \leq x\}$ — функция распределения случайной величины ξ_n , $n = 1, \dots, k, \dots$. Будем говорить, что последовательность функций распределения $F_{\xi_n}(x)$ сходится в основном к функции распределения $F_\xi(x)$ и записывать

$$F_{\xi_n}(x) \xrightarrow{n \rightarrow \infty} F_\xi(x), \quad x \in C(F_\xi), \quad (1.6)$$

если для любой точки $x \in C(F_\xi) = \{x \in R : F_\xi \text{ непрерывна в } x\}$ выполнено $F_{\xi_n}(x) \xrightarrow{n \rightarrow \infty} F_\xi(x)$.

Определение 1.6. Пусть задана последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$. Будем говорить, что данная последовательность фундаментальна с вероятностью 1, если $P\{\omega : \xi_n(\omega) \text{ — фундаментальна}\} = 1$.

§ 2. КРИТЕРИЙ СХОДИМОСТИ ПОЧТИ НАВЕРНОЕ. ТЕОРЕМА БОРЕЛЯ–КАНТЕЛЛИ

Теорема 2.1 (необходимое и достаточное условие сходимости с вероятностью единица).

Справедливы следующие утверждения:

1. Последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ сходится почти наверное к случайной величине ξ тогда и только тогда, когда для любого $\varepsilon > 0$ имеет место сходимость

$$P\left\{\omega : \sup_{m \geq n} |\xi_m - \xi| > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0. \quad (2.1)$$

2. Последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ фундаментальна почти наверное тогда и только тогда, когда для любого $\varepsilon > 0$:

$$P\{\omega : \sup_{m, l \geq n} |\xi_m - \xi_l| > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Докажем первое утверждение. Выясним условия, когда $P\{\xi_n(\omega) \not\rightarrow \xi(\omega)\} = 0$. Как нетрудно заметить, справедливо равенство:

$$\{\xi_n(\omega) \not\rightarrow \xi(\omega)\}_{n \rightarrow \infty} = \bigcup_{\varepsilon > 0} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{\omega : |\xi_m(\omega) - \xi(\omega)| > \varepsilon\}. \quad (2.2)$$

Очевидно, что

$$\begin{aligned} \bigcup_{\varepsilon > 0} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \{\omega : |\xi_m(\omega) - \xi(\omega)| > \varepsilon\} = \\ = \bigcup_{r=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{ \omega : |\xi_n(\omega) - \xi(\omega)| > \frac{1}{r} \right\}. \end{aligned}$$

Введем обозначение:

$$B_r = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{ \omega : |\xi_n(\omega) - \xi(\omega)| > \frac{1}{r} \right\}.$$

Таким образом, мы доказали следующее утверждение:

$$P\{\xi_n \not\rightarrow \xi\} = 0 \iff P\left\{ \bigcup_{r=1}^{\infty} B_r \right\} = 0. \quad (2.3)$$

Так как для любого r справедливо неравенство: $P(B_r) \leq P\{\bigcup_{r=1}^{\infty} B_r\} \leq \sum_{r=1}^{\infty} P(B_r)$, то условие, что $P\{\bigcup_{r=1}^{\infty} B_r\} = 0$, эквивалентно тому, что для любого r вероятность $P(B_r)$ равна нулю. Обозначим $\bigcup_{m=n}^{\infty} \left\{ \omega : |\xi_m - \xi| > \frac{1}{r} \right\}$ через A_n , $A_n \supset A_{n+1}$, тогда в силу непрерывности вероятностной меры равенство $P(B_r) = P\{\bigcap_{n=1}^{\infty} A_n\} = 0$ эквивалентно равенству $\lim_{n \rightarrow \infty} P(A_n) = 0$. Последнее равенство эквивалентно тому, что для любого r имеет место сходимость:

$$P\left\{ \bigcup_{m=n}^{\infty} \left\{ \omega : |\xi_m - \xi| > \frac{1}{r} \right\} \right\} \xrightarrow{n \rightarrow \infty} 0. \quad (2.4)$$

Справедливо:

$$\left\{ \bigcup_{m=n}^{\infty} \left\{ \omega : |\xi_m - \xi| > \frac{1}{r} \right\} \right\} = \left\{ \omega : \sup_{m \geq n} |\xi_m - \xi| > \frac{1}{r} \right\},$$

поэтому утверждение о том, что для любого r имеет место сходимость $\lim_{n \rightarrow \infty} P(A_n) = 0$, эквивалентно утверждению, что для любого r имеет место сходимость:

$$P \left\{ \sup_{m \geq n} |\xi_m - \xi| > \frac{1}{r} \right\} \xrightarrow{n \rightarrow \infty} 0. \quad (2.5)$$

Сходимость (2.5) для любого r эквивалентна тому, что для любого $\varepsilon > 0$ выполнено: $P \left\{ \sup_{m \geq n} |\xi_m - \xi| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0$.

Докажем второе утверждение. Запишем в явном виде следующее событие:

$$\begin{aligned} \{ \xi_n(\omega) \text{ не фундаментальна} \} = \\ = \bigcup_{\varepsilon > 0} \bigcap_{n=1}^{\infty} \bigcup_{\substack{m=n \\ l=n}}^{\infty} \{ \omega : |\xi_m(\omega) - \xi_l(\omega)| > \varepsilon \}. \end{aligned}$$

Далее все рассуждения повторяются аналогично доказательству пункта 1.

Замечание 2.1. Необходимое и достаточное условие фундаментальности с вероятностью единица можно записать иначе, а именно: для любого $\varepsilon > 0$ справедливо:

$$P \left\{ \omega : \sup_{k > 0} |\xi_{n+k} - \xi_n| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0.$$

Докажем, что данное условие эквивалентно условию 2 из теоремы 2.1. Достаточно заметить, что

$$\sup_{k > 0} |\xi_{n+k} - \xi_n| \leq \sup_{l, m \geq n} |\xi_l - \xi_m| \leq 2 \sup_{k > 0} |\xi_{n+k} - \xi_n|.$$

Тогда для любого $\varepsilon > 0$ справедливы неравенства:

$$\begin{aligned} P \left\{ \sup_{k > 0} |\xi_{n+k} - \xi_n| > \varepsilon \right\} &\leq P \left\{ \sup_{l, m \geq n} |\xi_l - \xi_m| > \varepsilon \right\} \leq \\ &\leq P \left\{ \sup_{k > 0} |\xi_{n+k} - \xi_n| > \frac{\varepsilon}{2} \right\}. \end{aligned}$$

Следовательно, утверждение, что

$$P \left\{ \sup_{k \geq 0} |\xi_{n+k} - \xi_n| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

для любого $\varepsilon > 0$ эквивалентно утверждению, что

$$P \left\{ \sup_{l, m \geq n} |\xi_l - \xi_m| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0 \text{ для любого } \varepsilon > 0. \quad \blacksquare$$

Лемма 2.1 (лемма Бореля–Кантелли). Пусть $\{A_n\}$ — бесконечная последовательность событий, $A_n \in \mathcal{F}$ для любого номера n . Пусть $A^* = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$, тогда справедливы утверждения:

1. Если ряд сходится $\sum_{k=1}^{\infty} P(A_k) < \infty$, тогда $P(A^*) = 0$.
2. Если ряд расходится $\sum_{k=1}^{\infty} P(A_k) = \infty$, и события A_k взаимно независимы, тогда $P(A^*) = 1$.

Доказательство. Докажем первый пункт леммы. Если ряд сходится, то при $n \rightarrow \infty$ справедливо:

$$P \left(\bigcup_{k=n}^{\infty} A_k \right) \leq \sum_{k=n}^{\infty} P(A_k) \xrightarrow{n \rightarrow \infty} 0.$$

Из следующего соотношения событий

$$\bigcup_{k=n}^{\infty} A_k \supset \bigcup_{k=n+1}^{\infty} A_k$$

и из непрерывности счетно-аддитивной меры получаем:

$$P(A^*) = \lim_{n \rightarrow \infty} P \left(\bigcup_{k=n}^{\infty} A_k \right) = 0.$$

Докажем второй пункт леммы. Для этого достаточно показать, что $P(\bar{A}^*) = 0$. Применяя законы двойственности, получим:

$$\bar{A}^* = \overline{\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m} = \bigcup_{n=1}^{\infty} \overline{\bigcup_{m=n}^{\infty} A_m} = \bigcup_{n=1}^{\infty} \left(\bigcap_{m=n}^{\infty} \bar{A}_m \right).$$

По свойству непрерывности вероятностной меры сверху, так как $\bigcap_{m=n}^{\infty} \bar{A}_m \subset \bigcap_{m=n+1}^{\infty} \bar{A}_m$, то $P(\bar{A}^*) = \lim_{n \rightarrow \infty} P \left(\bigcap_{m=n}^{\infty} \bar{A}_m \right)$.

Проведем следующие преобразования:

$$P\left(\bigcap_{m=n}^{\infty} \bar{A}_m\right) = P\left(\bigcap_{N=n}^{\infty} \bigcap_{m=n}^N \bar{A}_m\right) = \lim_{N \rightarrow \infty} P\left(\bigcap_{m=n}^N \bar{A}_m\right),$$

учитывая, что $\bigcap_{m=n}^N \bar{A}_m \supset \bigcap_{m=n}^{N+1} \bar{A}_m$. Рассмотрим вероятность:

$$\begin{aligned} P\left(\bigcap_{m=n}^N \bar{A}_m\right) &= \prod_{m=n}^N P(\bar{A}_m) = \prod_{m=n}^N (1 - P(A_m)) = \\ &= \exp\left(\sum_{m=n}^N \ln(1 - P(A_m))\right). \end{aligned} \quad (2.6)$$

Докажем, что из неравенства $1 > x \geq 0$ следует неравенство $\ln(1 - x) \leq -x$. Для этого рассмотрим функцию:

$$F(x) = \ln(1 - x) + x, \quad x \in [0, \alpha),$$

где $0 < \alpha < 1$. Очевидно, что $F(0) = 0$, $F'(x) = -1/(1 - x) + 1 = -x/(1 - x) < 0$. Следовательно, $F(x) < 0$ при $x \in (0, \alpha)$. Справедливо неравенство:

$$P\left(\bigcap_{m=n}^N \bar{A}_m\right) \leq \exp\left(-\sum_{m=n}^N P(A_m)\right) \xrightarrow{N \rightarrow \infty} 0.$$

Как видим, $\sum_{m=1}^{\infty} P(A_m) = \infty$ при условии взаимной независимости событий влечет: $\lim_{N \rightarrow \infty} P\left(\bigcap_{m=n}^N \bar{A}_m\right) = 0$, тогда $P(\bar{A}^*) = 0$. ■

Замечание 2.2. Установим содержательный смысл события A^* . Событие $A^* = \bigcap_{n=1}^{\infty} (\bigcup_{k=n}^{\infty} A_k)$ происходит, если одновременно происходят события $\{\bigcup_{k=n}^{\infty} A_k\}$. Таким образом, событие A^* происходит тогда и только тогда, когда совместно происходит бесконечно много событий, входящих в последовательность $\{A_n\}$.

Теорема 2.2 (иерархия видов сходимости). Пусть заданы последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ и случайная величина ξ , тогда справедливы утверждения:

1. Если $\xi_n \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \xi$, тогда $\xi_n \xrightarrow[n \rightarrow \infty]{P} \xi$.
2. Если $\xi_n \xrightarrow[n \rightarrow \infty]{L^r} \xi$, тогда $\xi_n \xrightarrow[n \rightarrow \infty]{P} \xi$.
3. Если $\xi_n \xrightarrow[n \rightarrow \infty]{P} \xi$, тогда $\xi_n \xrightarrow[n \rightarrow \infty]{d} \xi$.

Доказательство. Докажем первый пункт теоремы. Очевидно, что для любого $\varepsilon > 0$ имеет место неравенство:

$$P\left\{\sup_{m \geq n} |\xi_m - \xi| > \varepsilon\right\} \geq P\{|\xi_n - \xi| > \varepsilon\}.$$

Докажем второй пункт теоремы. Применим неравенство Чебышева:

$$P\{|\xi_n - \xi| > \varepsilon\} = P\{|\xi_n - \xi|^r > \varepsilon^r\} \leq \frac{E|\xi_n - \xi|^r}{\varepsilon^r} \xrightarrow[n \rightarrow \infty]{} 0.$$

Докажем третье утверждение теоремы. Пусть $g(x)$ — любая непрерывная и ограниченная функция, $|g(x)| \leq k$, тогда справедливы следующие неравенства:

$$\begin{aligned} |Eg(\xi_n) - Eg(\xi)| &= |E(g(\xi_n) - g(\xi))| \leq \\ &\leq E\{I\{|\xi_n - \xi| \leq \delta\}|g(\xi_n) - g(\xi)|\} + \\ &+ E\{I\{|\xi_n - \xi| > \delta\}|g(\xi_n) - g(\xi)|\} = \\ &= E\{I\{|\xi| > N\}I\{|\xi_n - \xi| \leq \delta\}|g(\xi_n) - g(\xi)|\} + \\ &+ E\{I\{|\xi| \leq N\}I\{|\xi_n - \xi| \leq \delta\}|g(\xi_n) - g(\xi)|\} + \\ &+ E\{I\{|\xi_n - \xi| > \delta\}|g(\xi_n) - g(\xi)|\}. \end{aligned}$$

Рассмотрим отдельно слагаемые в правой части неравенства:

$$\begin{aligned} E\{I\{|\xi| > N\}I\{|\xi_n - \xi| \leq \delta\}|g(\xi_n) - g(\xi)|\} &\leq \\ &\leq 2kP\{|\xi| > N\} \leq 2k(F_\xi(-N) + 1 - F_\xi(N)) \xrightarrow[N \rightarrow \infty]{} 0. \end{aligned}$$

Выберем δ_0 так, чтобы $|g(x) - g(y)| < \varepsilon/3$ для всех $x, y \in [-N-1, N+1]$, $|x - y| \leq \delta$, при $\delta < \delta_0$, тогда справедливы

соотношения:

$$E\{I\{|\xi| \leq N\}I\{|\xi_n - \xi| \leq \delta\}|g(\xi_n) - g(\xi)|\} < \varepsilon/3,$$

$$E\{I\{|\xi_n - \xi| > \delta\}|g(\xi_n) - g(\xi)|\} < 2kP\{|\xi_n - \xi| > \delta\} \xrightarrow{n \rightarrow \infty} 0.$$

Легко заметить, что последовательно выбирая N , δ и n для произвольного $\varepsilon > 0$, можно добиться, чтобы каждое слагаемое оказалось меньше $\varepsilon/3$, следовательно, сумма окажется меньше ε . ■

Пример 2.1. Пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}\{[0, 1]\}$. В качестве меры возьмем меру Лебега. Построим последовательность случайных величин: $\xi_1^{(1)}(\omega) = I\{\omega \in [0, 1]\}$, $\xi_2^{(1)}(\omega) = I\{\omega \in [0, 1/2]\}$, $\xi_2^{(2)}(\omega) = I\{\omega \in [1/2, 1]\}$, $\xi_3^{(1)}(\omega) = I\{\omega \in [0, 1/3]\}$, $\xi_3^{(2)}(\omega) = I\{\omega \in [1/3, 2/3]\}$, $\xi_3^{(3)}(\omega) = I\{\omega \in [2/3, 1]\}$, ..., $\xi_n^{(1)}(\omega) = I\{\omega \in [0, 1/n]\}$ и так далее. Расположим случайные величины следующим образом: $\xi_1^{(1)}$, $\xi_2^{(1)}$, $\xi_2^{(2)}$, $\xi_3^{(1)}$, $\xi_3^{(2)}$, $\xi_3^{(3)}$, Изучим сходимость этой последовательности по вероятности. Очевидно, что для любого $\varepsilon \in (0, 1)$: $P\{|\xi_n^{(m)}| > \varepsilon\} = 1/n \xrightarrow{n \rightarrow \infty} 0$, следовательно, имеет место сходимость: $\xi_n^{(m)} \xrightarrow[n \rightarrow \infty]{P} 0$.

Теперь проверим сходимость в среднем. Возьмем $r > 0$ и рассмотрим $E|\xi_n^{(m)}|^r = 1/n \xrightarrow[n \rightarrow \infty]{} 0$, следовательно, имеет место сходимость $\xi_n^{(m)} \xrightarrow[n \rightarrow \infty]{L^r} 0$. Проверим наличие сходимости почти наверное. Для этого рассмотрим точку $\omega \in [0, 1]$ и рассмотрим $\xi_n^{(m)}(\omega)$. В этой числовой последовательности есть бесконечная подпоследовательность с бесконечным числом единиц, следовательно, ни в одной точке ω сходимости нет. Этот пример показывает, что из сходимости по вероятности и сходимости в среднем любого порядка не следует сходимости почти наверное.

Пример 2.2. Пусть имеется такое же вероятностное пространство, что и в примере 2.1. Зададим последовательность

случайных величин отличным от предыдущего примера образом:

$$\xi_n(\omega) = \begin{cases} e^n, & \text{если } 0 \leq \omega \leq \frac{1}{n}, \\ 0, & \text{если } \omega > \frac{1}{n}. \end{cases}$$

Возьмем любое $\varepsilon > 0$, тогда $P\{|\xi_n| > \varepsilon\} = 1/n \xrightarrow{n \rightarrow \infty} 0$, следовательно, $\xi_n \xrightarrow[n \rightarrow \infty]{P} 0$.

Нетрудно заметить, что для любого $\omega \in (0, 1]$, если рассматривать последовательность ξ_n в этой точке ω , т. е. $\xi_n(\omega)$, начиная с некоторого номера, все члены последовательности будут равны нулю: $\xi_n(\omega) \xrightarrow[n \rightarrow \infty]{} 0$. Это верно для всех точек, за исключением точки $\omega = 0$, следовательно, $\xi_n \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \xi$.

Проверим сходимость в среднем порядка r . Очевидно, что $E|\xi_n|^r = e^{nr}/n \xrightarrow[n \rightarrow \infty]{} \infty$, следовательно, сходимости в среднем порядка r нет. Из сходимости почти наверное не следует сходимости в среднем порядка r .

§ 3. ЭКВИВАЛЕНТНОСТЬ СХОДИМОСТИ В ОСНОВНОМ И СХОДИМОСТИ ПО РАСПРЕДЕЛЕНИЮ. СВОЙСТВА СХОДИМОСТИ В ОСНОВНОМ

Теорема 3.1. Пусть заданы последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ и случайная величина ξ , $F_\xi(x)$ — функция распределения случайной величины ξ , $F_{\xi_n}(x)$ — функция распределения случайной величины ξ_n , $n = 1, \dots$. Последовательность $\xi_1, \dots, \xi_n, \dots$ сходится по распределению к ξ при $n \rightarrow \infty$ тогда и только тогда, когда функции $F_{\xi_n}(x)$ сходятся в основном при $n \rightarrow \infty$ к функции $F_\xi(x)$.

Доказательство. *Необходимость.* Возьмем любое $a \in \mathbb{C}(F)$. Пусть

$$g(x) = \begin{cases} 1, & x \in (-\infty, a], \\ \frac{-x+a+\varepsilon}{\varepsilon}, & x \in (a, a+\varepsilon], \\ 0, & x \in (a+\varepsilon, \infty). \end{cases}$$

Справедливы рассуждения:

$$\begin{aligned} F_{\xi_n}(a) = EI\{\xi_n \in (-\infty, a]\} &\leq Eg(\xi_n) \xrightarrow{n \rightarrow \infty} Eg(\xi) \leq \\ &\leq EI\{\xi \in (-\infty, a + \varepsilon]\} = F_\xi(a + \varepsilon). \end{aligned} \quad (3.1)$$

Следовательно, $\overline{\lim}_{n \rightarrow \infty} F_{\xi_n}(a) \leq F_\xi(a + \varepsilon)$.

Рассмотрим функцию $\tilde{g}(x)$ следующего вида: $\tilde{g}(x) = g(x + \varepsilon)$. Аналогично рассуждая, получаем

$$\begin{aligned} F_{\xi_n}(a) = EI\{\xi_n \in (-\infty, a]\} &\geq E\tilde{g}(\xi_n) \rightarrow E\tilde{g}(\xi) \geq \\ &\geq EI\{\xi \in (-\infty, a - \varepsilon]\} = F_\xi(a - \varepsilon). \end{aligned}$$

Следовательно,

$$\underline{\lim}_{n \rightarrow \infty} F_{\xi_n}(a) \geq F_\xi(a - \varepsilon). \quad (3.2)$$

Объединив (3.1) и (3.2), получаем неравенства:

$$F_\xi(a - \varepsilon) \leq \underline{\lim}_{n \rightarrow \infty} F_{\xi_n}(a) \leq \overline{\lim}_{n \rightarrow \infty} F_{\xi_n}(a) \leq F_\xi(a + \varepsilon). \quad (3.3)$$

Функция $F_\xi(x)$ непрерывна в точке $x = a$, следовательно, $\lim_{n \rightarrow \infty} F_{\xi_n}(a)$ существует и равен $F_\xi(a)$.

Достаточность. Возьмем любую ограниченную непрерывную функцию $g(x)$. Пусть $|g(x)| \leq k$ при любом $x \in \mathbb{R}$. Возьмем произвольный полуинтервал $(a, b]$: $a, b \in \mathbb{C}(F_\xi)$ — точки непрерывности. Покажем, что

$$\int_{(a, b]} g(x) dP_{\xi_n} \xrightarrow{n \rightarrow \infty} \int_{(a, b]} g(x) dP_\xi. \quad (3.4)$$

Рассмотрим замкнутый промежуток $[a, b]$, $g(x)$ равномерно непрерывна на нем, т.е. для любого $\varepsilon > 0$ существует $\delta > 0$, что имеет место неравенство: $|x - y| < \delta$, тогда $|g(x) - g(y)| < \varepsilon/3$. Разобьем $[a, b]$ с шагом меньше δ на части: $\Delta_1 \cup \dots \cup \Delta_m = [a, b]$. Причем, $\Delta_i \cap \Delta_j = \emptyset$, $\Delta_i = (a_{i-1}, a_i]$ для всех $i \neq 1$, $\Delta_1 = [a_0, a_1]$, где $a_0 = a$, $\Delta_m = (a_{m-1}, a_m]$, $a_m = b$, $a_{i+1} - a_i < \delta$.

Пусть $\tilde{g}(x) = \sum_{i=1}^m g(x_i)I\{x \in \Delta_i\}$, где $x_i \in \Delta_i$, следовательно $|g(x) - \tilde{g}(x)| < \varepsilon/3$ при $x \in [a, b]$. Рассмотрим выражение:

$$\begin{aligned} & \left| \int_{[a,b]} g(x) dP_{\xi_n} - \int_{[a,b]} g(x) dP_{\xi} \right| = \\ & = \left| \int_{[a,b]} g(x) dP_{\xi_n} - \int_{[a,b]} g(x) dP_{\xi} \pm \int_{[a,b]} \tilde{g}(x) dP_{\xi_n} \mp \int_{[a,b]} \tilde{g}(x) dP_{\xi} \right| \leq \\ & \leq \int_{[a,b]} |g(x) - \tilde{g}(x)| dP_{\xi_n} + \int_{[a,b]} |g(x) - \tilde{g}(x)| dP_{\xi} + \\ & \quad + \sum_{i=1}^m |g(x_i)| \cdot |P_{\xi_n}(\Delta_i) - P_{\xi}(\Delta_i)|. \quad (3.5) \end{aligned}$$

Первый и второй интегралы по отдельности не больше $\varepsilon/3$. Очевидно, что

$$\begin{aligned} P_{\xi_n}(\Delta_i) - P_{\xi}(\Delta_i) &= F_{\xi_n}(a_i) - F_{\xi_n}(a_{i-1}) - F_{\xi}(a_i) + F_{\xi}(a_{i-1}) = \\ &= (F_{\xi_n}(a_i) - F_{\xi}(a_i)) + (F_{\xi_n}(a_{i-1}) - F_{\xi}(a_{i-1})). \end{aligned}$$

В качестве точек дробления a_i были выбраны точки непрерывности F_{ξ} , следовательно, с ростом n выражения в скобках стремятся к нулю. Выберем n_0 таким, чтобы для любого $n > n_0$ и для всех i было выполнено:

$$|F_{\xi_n}(a_i) - F_{\xi}(a_i)| < \frac{\varepsilon}{3km}.$$

Тогда абсолютная величина разности двух интегралов окажется меньше ε . Соотношение (3.4) доказано.

Теперь покажем, что (3.4) будет верно и при интегрировании по всей оси \mathbb{R} . Для этого выберем $\varepsilon > 0$ и отрезок $[a, b]$ такие, что $F_{\xi}(a) < \varepsilon/(6k)$, $1 - F_{\xi}(b) < \varepsilon/(6k)$, $a, b \in \mathbb{C}(F_{\xi})$. Очевидно, что можно выбрать n_1 , чтобы при $n \geq n_1$ было выполнено неравенство: $|F_{\xi_n}(a) - F_{\xi}(a)| < \varepsilon/(6k)$. Аналогично, чтобы для $n \geq n_2$ было выполнено неравенство:

$|F_{\xi_n}(b) - F_{\xi}(b)| < \varepsilon/(6k)$. Тогда справедливы рассуждения:

$$\begin{aligned} & \left| \int_{\mathbb{R}} g(x) dP_{\xi_n} - \int_{\mathbb{R}} g(x) dP_{\xi} \right| = \\ &= \left| \int_{(-\infty, a]} g(x) dP_{\xi_n} + \int_{(b, \infty)} g(x) dP_{\xi_n} - \int_{(-\infty, a]} g(x) dP_{\xi} - \right. \\ & \quad \left. - \int_{(b, \infty)} g(x) dP_{\xi} + \left(\int_{(a, b]} g(x) dP_{\xi_n} - \int_{(a, b]} g(x) dP_{\xi} \right) \right| \leq \frac{\varepsilon}{3} + \\ & \quad + \frac{\varepsilon}{3} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \left| \left(\int_{(a, b]} g(x) dP_{\xi_n} - \int_{(a, b]} g(x) dP_{\xi} \right) \right| \leq 2\varepsilon, \end{aligned}$$

где n_3 выбрано так, чтобы при $n \geq n_3$ последнее слагаемое было меньше ε , после чего для $n > \max(n_1, n_2, n_3)$ полученная оценка справедлива. ■

Замечание 3.1. Теорема справедлива и в конечномерных пространствах, когда рассматриваются последовательности случайных векторов.

Теорема 3.2. Пусть последовательность функций распределения $F_{\xi_n}(x)$ сходится в основном к функции распределения $F_{\xi}(x)$. Пусть множество $C(F_{\xi}) = \mathbb{R}$, т. е. предельная функция $F_{\xi}(x)$ непрерывна во всех точках \mathbb{R} , тогда последовательность $F_{\xi_n}(x)$ равномерно сходится к $F_{\xi}(x)$, т. е.

$$\sup_{x \in \mathbb{R}} |F_{\xi_n}(x) - F_{\xi}(x)| \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Для любого $\varepsilon > 0$, выберем точки a и b так, чтобы выполнялись неравенства: $F_{\xi}(a) < \varepsilon$, $1 - F_{\xi}(b) < \varepsilon$. Разобьем промежутки $[a, b]$ на непересекающиеся отрезки:

$$[a, b] = [a_0, a_1] + (a_1, a_2] + \dots + (a_{m-1}, a_m] : F_{\xi}(a_i) - F_{\xi}(a_{i-1}) < \varepsilon.$$

Для любой точки a_i существует n_i , что при $n \geq n_i$ выполнено $|F_{\xi_n}(a_i) - F_{\xi}(a_i)| < \varepsilon$. Пусть $\bar{n} = \max_{i=0, m} \{n_i\}$. Будем считать, что $n \geq \bar{n}$. Возьмем точку $x \in [a_i, a_{i+1}]$, тогда справедливы неравенства:

$$\begin{aligned} F_{\xi_n}(x) - F_{\xi}(x) &\leq F_{\xi_n}(a_{i+1}) - F_{\xi}(a_i) \leq \\ &\leq F_{\xi_n}(a_{i+1}) - F_{\xi}(a_{i+1}) + \varepsilon \leq 2\varepsilon, \end{aligned}$$

$$\begin{aligned} F_{\xi_n}(x) - F_{\xi}(x) &\geq F_{\xi_n}(a_i) - F_{\xi}(a_{i+1}) \geq \\ &\geq F_{\xi_n}(a_i) - F_{\xi}(a_i) - \varepsilon \geq -2\varepsilon. \end{aligned}$$

Следовательно, неравенство $|F_{\xi_n}(x) - F_{\xi}(x)| \leq 2\varepsilon$ справедливо для всех $x \in [a, b]$.

Пусть $x \in (-\infty, a]$, тогда $F_{\xi_n}(x) - F_{\xi}(x) \leq F_{\xi_n}(a) \leq 2\varepsilon$, $F_{\xi_n}(x) - F_{\xi}(x) \geq -F_{\xi}(a) \geq -\varepsilon$.

Пусть $x \in [b, +\infty)$, тогда $F_{\xi_n}(x) - F_{\xi}(x) \leq 1 - F_{\xi}(b) < \varepsilon$, $F_{\xi_n}(x) - F_{\xi}(x) \geq F_{\xi_n}(b) - 1 \pm F_{\xi}(b) > -2\varepsilon$.

Следовательно, для всех x справедливо неравенство: $|F_{\xi_n}(x) - F_{\xi}(x)| < 2\varepsilon$, что означает равномерную сходимость.

■

Замечание 3.2. Теорема справедлива и в конечномерных пространствах.

Определение 3.1. Случайная величина называется вырожденной, если с вероятностью единица она принимает постоянное значение: $P\{\xi = a\} = 1$.

Теорема 3.3. Пусть последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ сходится по распределению при $n \rightarrow \infty$ к случайной величине $\xi \stackrel{n.h.}{=} a$, тогда $\xi_n \xrightarrow[n \rightarrow \infty]{P} a$.

Доказательство. Покажем, что для любого $\varepsilon > 0$ выполнено: $P\{|\xi_n - \xi| > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0$. Справедливы рассуждения:

$$\begin{aligned} P\{|\xi_n - \xi| > \varepsilon\} &= \\ &= P\{\xi_n - \xi < -\varepsilon\} + P\{\xi_n - \xi > \varepsilon\} = P\{\xi_n < a - \varepsilon\} + \\ &+ P\{\xi_n > a + \varepsilon\} \leq P\{\xi_n \leq a - \varepsilon\} + P\{\xi_n > a + \varepsilon\} = \\ &= F_{\xi_n}(a - \varepsilon) + (1 - F_{\xi_n}(a + \varepsilon)). \end{aligned}$$

Последние два слагаемых стремятся к нулю, следовательно, сходимость по вероятности имеетя: $P\{|\xi_n - \xi| > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0$ или $P\{|\xi_n - a| > \varepsilon\} \xrightarrow{n \rightarrow \infty} 0$, так как $P\{|\xi_n - a| > \varepsilon\} = P\{|\xi_n - \xi| > \varepsilon\}$. ■

ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

§ 1. СВОЙСТВА ХАРАКТЕРИСТИЧЕСКИХ ФУНКЦИЙ

Рассмотрим вероятностное пространство (Ω, \mathcal{F}, P) , на котором заданы случайные величины $\xi(\omega)$ и $\eta(\omega)$.

Определение 1.1. Под комплексной случайной величиной будем понимать случайную величину

$$\zeta(\omega) = \xi(\omega) + i\eta(\omega).$$

Комплексную случайную величину можно представить по-другому:

$$\zeta(\omega) = |\zeta(\omega)|e^{i\text{Arg}(\zeta(\omega))}.$$

Пусть существуют и конечны математические ожидания случайных величин ξ и η . Математическое ожидание ζ определяется следующим образом:

$$E\zeta = E\xi + iE\eta.$$

Определение 1.2. Комплексные случайные величины $\zeta_1(\omega) = \xi_1(\omega) + i\eta_1(\omega)$ и $\zeta_2(\omega) = \xi_2(\omega) + i\eta_2(\omega)$ называются независимыми, если вектора $(\xi_1(\omega), \eta_1(\omega))^T$ и $(\xi_2(\omega), \eta_2(\omega))^T$ взаимно независимы.

Свойства комплексных случайных величин

1. Мультипликативность.

Если комплексные случайные величины ζ_1 и ζ_2 независимы и существуют математические ожидания $E\zeta_1$, $E\zeta_2$, тогда имеет место равенство:

$$E(\zeta_1 \zeta_2) = E\zeta_1 E\zeta_2.$$

Доказательство. По правилу умножения комплексных чисел замечаем, что

$$\zeta_1 \zeta_2 = (\xi_1 \xi_2 - \eta_1 \eta_2) + i(\xi_1 \eta_2 + \xi_2 \eta_1).$$

Следовательно, используя определение, линейность и попарную взаимную независимость элементов, получаем равенства:

$$\begin{aligned} E(\zeta_1 \zeta_2) &= E(\xi_1 \xi_2 - \eta_1 \eta_2) + iE(\xi_1 \eta_2 + \xi_2 \eta_1) = \\ &= E\xi_1 E\xi_2 - E\eta_1 E\eta_2 + i(E\xi_1 E\eta_2 + E\xi_2 E\eta_1) = \\ &= (E\xi_1 + iE\eta_1)(E\xi_2 + iE\eta_2) = E\zeta_1 E\zeta_2. \end{aligned}$$

■

2. Пусть $\zeta = \xi + i\eta$ и существует математическое ожидание $E\zeta$, тогда $|E\zeta| \leq E|\zeta|$.

Доказательство. 1) Пусть $\xi(\omega) = \sum_{k=1}^n a_k I\{\omega \in A_k\}$ и $\eta(\omega) = \sum_{l=1}^m b_l I\{\omega \in B_l\}$ — простые случайные величины. Индикатор множества A_k определяется следующим образом:

$$I\{\omega \in A_k\} = \begin{cases} 1, & \text{если } \omega \in A_k; \\ 0, & \text{если } \omega \notin A_k, \end{cases}$$

где $A_k \cap A_{k'} = \emptyset$ для любых $k \neq k'$, $\bigcup_{k=1}^n A_k = \Omega$, $A_k \in \mathcal{F}$. Аналогичные свойства выполнены для B_l . Тогда можно записать равенства:

$$\xi(\omega) = \sum_{k=1}^n \sum_{l=1}^m a_k I\{\omega \in A_k \cap B_l\},$$

$$\eta(\omega) = \sum_{k=1}^n \sum_{l=1}^m b_l I\{\omega \in A_k \cap B_l\}.$$

Нетрудно заметить, что справедливо равенство: $\sum_{k=1}^n I(\omega \in A_k \cap B_l) = I\{B_l\}$. Следовательно, имеет место равенство:

$$\zeta(\omega) = \xi(\omega) + i\eta(\omega) = \sum_{k=1}^n \sum_{l=1}^m (a_k + ib_l) I\{\omega \in A_k \cap B_l\}.$$

Очевидно, что

$$|\zeta(\omega)| = \sum_{k=1}^n \sum_{l=1}^m |a_k + ib_l| I\{\omega \in A_k \cap B_l\}.$$

По определению интеграла Лебега для простой функции получаем равенство:

$$E|\zeta(\omega)| = \sum_{k=1}^n \sum_{l=1}^m |a_k + ib_l| P(A_k \cap B_l). \quad (1.1)$$

Учитывая, что

$$E\xi(\omega) = \sum_{k=1}^n \sum_{l=1}^m a_k P(A_k \cap B_l),$$

$$E\eta(\omega) = \sum_{k=1}^n \sum_{l=1}^m b_l P(A_k \cap B_l),$$

получаем выражение:

$$|E\zeta(\omega)| = \left| \sum_{k=1}^n \sum_{l=1}^m (a_k + ib_l) P(A_k \cap B_l) \right|. \quad (1.2)$$

Сравним (1.1) и (1.2). Очевидно, что $|E\zeta(\omega)| \leq E|\zeta(\omega)|$.
 2) Рассмотрим общий случай: $E\xi \in \mathbb{R}$, $E\eta \in \mathbb{R}$, существует последовательность $\xi_n(\omega)$ простых случайных величин такая, что $\xi_n(\omega) \xrightarrow[n \rightarrow \infty]{} \xi(\omega)$ для всех $\omega \in \Omega$, и, кроме того, выполняется неравенство: $|\xi_n(\omega)| \leq |\xi(\omega)|$.

Существует последовательность $\eta_n(\omega)$ простых случайных величин: $\eta_n(\omega) \xrightarrow{n \rightarrow \infty} \eta(\omega)$ для всех $\omega \in \Omega$, и, кроме того, $|\eta_n(\omega)| \leq |\eta(\omega)|$.

По теореме о мажорируемой сходимости получаем: $E\xi_n(\omega) \xrightarrow{n \rightarrow \infty} E\xi$, $E\eta_n(\omega) \xrightarrow{n \rightarrow \infty} E\eta$.

Введем последовательность $\zeta_n(\omega) = \xi_n(\omega) + i\eta_n(\omega) \xrightarrow{n \rightarrow \infty} \zeta(\omega) = \xi(\omega) + i\eta(\omega)$, но тогда $E\zeta_n \xrightarrow{n \rightarrow \infty} E\zeta$.

Очевидно, что

$$\begin{aligned} |\zeta_n(\omega)| &= \sqrt{\xi_n^2(\omega) + \eta_n^2(\omega)} \xrightarrow{n \rightarrow \infty} |\zeta(\omega)| = \\ &= \sqrt{\xi^2(\omega) + \eta^2(\omega)}. \end{aligned}$$

При этом выполняется неравенство: $|\zeta_n(\omega)| \leq |\zeta(\omega)|$. Кроме того, уже доказано, что для всех n справедливо неравенство: $|E\zeta_n(\omega)| \leq E|\zeta_n(\omega)|$.

Устремим n к бесконечности. Так как $E|\zeta(\omega)| \leq E|\xi(\omega)| + E|\eta(\omega)|$, $E|\zeta(\omega)| \in \mathbb{R}$, то применима теорема Лебега о мажорируемой сходимости, следовательно, мы можем перейти к пределу в неравенстве, в результате получим неравенство:

$$|E\zeta(\omega)| \leq E|\zeta(\omega)|.$$

■

Определение 1.3. Характеристической функцией случайной величины ξ назовем функцию

$$\varphi_\xi(t) = Ee^{i\xi t}, \quad t \in \mathbb{R}. \quad (1.3)$$

Замечание 1.1. По существу, $\varphi_\xi(t) = \int_{\Omega} e^{i\xi t} dP = \int_{-\infty}^{\infty} e^{itx} dF_\xi(x)$. Так как $e^{i\xi t} = \cos(\xi t) + i\sin(\xi t)$, получаем еще одну форму представления характеристической функции:

$$\begin{aligned} \varphi_\xi(t) &= E(\cos(\xi t) + i\sin(\xi t)) = E\cos(\xi t) + iE\sin(\xi t) = \\ &= \int_{-\infty}^{\infty} \cos(xt) dF_\xi(x) + i \int_{-\infty}^{\infty} \sin(xt) dF_\xi(x). \end{aligned}$$

Как видим, характеристическая функция $\varphi_\xi(t)$ порождается функцией распределения $F_\xi(x)$ и зависит только от распределения случайной величины ξ .

Определение 1.4. Характеристической функцией случайного вектора $\xi = (\xi_1, \dots, \xi_m)$ назовем функцию

$$\varphi_\xi(t) = \varphi_{(\xi_1, \dots, \xi_m)}(t_1, \dots, t_m) = Ee^{i\xi^T t}. \quad (1.4)$$

Характеристическую функцию случайного вектора $\xi = (\xi_1, \dots, \xi_m)$ можно найти по формуле:

$$\varphi_\xi(t) = \int_{\Omega} e^{i\xi^T t} dP = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{ix^T t} dF_\xi(x).$$

Очевидно, что характеристическая функция существует для любой функции распределения.

Свойства характеристических функций

1. Для любого $t \in \mathbb{R}$ имеет место неравенство:

$$|\varphi_\xi(t)| \leq 1 = \varphi_\xi(0).$$

Доказательство. Запишем определение характеристической функции:

$$\varphi_\xi(t) = Ee^{i\xi t} = E \cos(\xi t) + iE \sin(\xi t).$$

Очевидно, что

$$|\varphi_\xi(t)| = |Ee^{i\xi t}| \leq E|e^{i\xi t}| = 1.$$

При $t = 0$: $\varphi_\xi(0) = 1$. ■

2. Пусть $\eta = a + b\xi$, где $a, b \in \mathbb{R}$, тогда справедливо равенство:

$$\varphi_\eta(t) = e^{iat} \varphi_\xi(bt).$$

Доказательство. Легко заметить:

$$\varphi_\eta(t) = Ee^{it\eta} = E(e^{iat} e^{itb\xi}) = e^{iat} Ee^{itb\xi} = e^{iat} \varphi_\xi(bt). \quad \blacksquare$$

3. Пусть $S_n = \xi_1 + \dots + \xi_n$, случайные величины ξ_1, \dots, ξ_n независимы, тогда характеристическая функция случайной величины S_n имеет вид:

$$\varphi_{S_n}(t) = \prod_{k=1}^n \varphi_{\xi_k}(t). \quad (1.5)$$

Доказательство. Достаточно доказать формулу (1.5) для двух слагаемых:

$$\varphi_{S_2}(t) = E(e^{i\xi_1 t} e^{i\xi_2 t}) = \varphi_{\xi_1}(t) \varphi_{\xi_2}(t).$$

Результат переносится на случай произвольного числа n . ■

4. Характеристическая функция $\varphi_\xi(t)$ случайной величины ξ равномерно непрерывна на \mathbb{R} .

Доказательство. Рассмотрим модуль разности:

$$\begin{aligned} |\varphi_\xi(t+h) - \varphi_\xi(t)| &= |Ee^{i\xi(t+h)} - Ee^{i\xi t}| = \\ &= |E(e^{i\xi t}(e^{i\xi h} - 1))| \leq E|e^{i\xi h} - 1| = \\ &= E|(\cos(\xi h) - 1) + i \sin(\xi h)|, \end{aligned}$$

где мажоранта не зависит от t .

По теореме Лебега о мажорируемой сходимости справедлива сходимость:

$$E|(\cos(\xi h) - 1) + i \sin(\xi h)| \xrightarrow{h \rightarrow 0} 0,$$

следовательно, характеристическая функция равномерно непрерывна на \mathbb{R} . ■

5. Имеют место равенства:

$$\overline{\varphi_\xi(t)} = \overline{Ee^{i\xi t}} = E\overline{e^{i\xi t}} = Ee^{-i\xi t} = \varphi_\xi(-t) = \varphi_{-\xi}(t).$$

Определение 1.5. Распределение случайной величины ξ симметрично, если для любого $x \in \mathbb{R}$ выполняется неравенство: $P\{\xi \leq x\} = P\{\xi \geq -x\}$.

Для симметричной случайной величины ξ справедливо равенство $F_\xi(x) = F_{-\xi}(x)$. Распределение случайной величины ξ симметрично тогда и только тогда, когда функции распределения случайных величин ξ и $-\xi$ совпадают.

6. Характеристическая функция $\varphi_\xi(t)$ является вещественной, т. е. $\overline{\varphi_\xi(t)} = \varphi_\xi(t)$, тогда и только тогда, когда распределение случайной величины ξ симметрично, $F_\xi(x) = F_{-\xi}(x)$.

Доказательство. 1) *Достаточность.* Пусть распределение симметрично, то есть $F_\xi(x) \equiv F_{-\xi}(x)$. Из свойства 5 следует, что $\overline{\varphi_\xi(t)} = \varphi_{-\xi}(t) = \varphi_\xi(t)$.

2) *Необходимость.* Пусть $\varphi_\xi(t) = \overline{\varphi_\xi(t)} = \varphi_{-\xi}(t)$. Это означает, что характеристические функции случайных величин ξ и $-\xi$ совпадают. Далее будет доказано, что характеристическим функциям взаимно однозначно соответствуют функции распределения, следовательно, выполняется равенство:

$$F_\xi(x) = F_{-\xi}(x).$$

■

7. Пусть существует начальный момент порядка n , и он конечен, $E\xi^n \in \mathbb{R}$, тогда для любого номера $k \leq n$ существует $\varphi_\xi^{(k)}(t) = E\{i^k \xi^k e^{it\xi}\}$ и $\varphi_\xi^{(k)}(0) = i^k E\xi^k$. Кроме того, справедливо представление:

$$\varphi_\xi(t) = \sum_{k=0}^n \frac{(it)^k}{k!} E\xi^k + \frac{(it)^n}{n!} \varepsilon(t),$$

где $\varepsilon(t) \xrightarrow[t \rightarrow 0]{} 0$.

Доказательство. Как легко видеть,

$$\begin{aligned} \frac{\varphi_\xi(t+h) - \varphi_\xi(t)}{h} &= \\ &= \frac{E(e^{i\xi(t+h)} - e^{i\xi t})}{h} = E\left(e^{i\xi t} \frac{e^{i\xi h} - 1}{h}\right) = \end{aligned}$$

$$= E \left(e^{i\xi t} \left(\frac{\cos(\xi h) - 1}{h} + i \frac{\sin(\xi h)}{h} \right) \right). \quad (1.6)$$

Рассмотрим $h > 0$. Используем формулу конечных приращений Лагранжа для синуса и косинуса:

$$\cos(\xi h) - 1 = -\xi \sin(\Theta_1 \xi h) h,$$

$$\sin(\xi h) = \xi \cos(\Theta_2 \xi h) h,$$

где $\Theta_1 \in (0, 1)$ и $\Theta_2 \in (0, 1)$. Тогда

$$\frac{\varphi_\xi(t+h) - \varphi_\xi(t)}{h} = E(e^{i\xi t} (-\xi \sin(\Theta_1 \xi h) + i\xi \cos(\Theta_2 \xi h))).$$

Очевидно, что справедливо неравенство:

$$|e^{i\xi t} (-\xi \sin(\Theta_1 \xi h) + i\xi \cos(\Theta_2 \xi h))| \leq 2|\xi|,$$

причем, по условию $E|\xi| \in \mathbb{R}$. Таким образом, по теореме Лебега о мажорируемой сходимости имеет место сходимость:

$$\frac{\varphi_\xi(t+h) - \varphi_\xi(t)}{h} \xrightarrow{h \rightarrow 0} E(i\xi e^{i\xi t}).$$

Аналогичным образом можно рассмотреть $(\varphi'_\xi(t+h) - \varphi'_\xi(t))/h$, мажорантой будет $2\xi^2$, если $n \geq 2$, то $E|\xi|^2 \in \mathbb{R}$. Рассуждения повторяются последовательно нужное число раз. При рассмотрении $(\varphi_\xi^{(n-1)}(t+h) - \varphi_\xi^{(n-1)}(t))/h$ мажорантой будет $2|\xi|^n$, так как $E|\xi|^n \in \mathbb{R}$, можно применить теорему Лебега и получить соответствующий предел.

Далее запишем формулу Тейлора для подынтегральной функции:

$$\begin{aligned} \cos(\xi t) + i \sin(\xi t) &= \sum_{k=0}^{n-1} \frac{(it)^k}{k!} \xi^k + \\ &+ \frac{(it)^n}{n!} \xi^n (\cos(\xi t \Theta_1) + i \sin(\xi t \Theta_2)), \end{aligned}$$

где $\Theta_1 \in (0, 1)$ и $\Theta_2 \in (0, 1)$.

Прибавим и вычтем единицу в последнем слагаемом, введем обозначение: $\varepsilon_n(t) = E\{\xi^n(\cos(\Theta_1 \xi t) - 1 + i \sin(\Theta_2 \xi t))\}$. Выражение в фигурных скобках стремится к 0 при $t \rightarrow 0$, оно также ограничено сверху по модулю величиной $3|\xi|^n$. Теорема Лебега о мажорируемой сходимости в данном случае применима, поэтому можно поменять местами интегрирование и предельный переход, следовательно, $\varepsilon_n(t) \rightarrow 0$ при $t \rightarrow 0$. ■

8. Следующее свойство представляет собой теорему.

Теорема 1.1 (о взаимной однозначности функций распределения и характеристических функций). Пусть $F(x)$ и $G(x)$ — функции распределения, заданные на \mathbb{R} . Если в каждой точке $t \in \mathbb{R}$ имеет место равенство $\int_{-\infty}^{\infty} e^{ixt} dF(x) = \int_{-\infty}^{\infty} e^{ixt} dG(x)$, тогда $F(x) \equiv G(x)$.

Доказательство. Возьмем любые $a \in \mathbb{R}$ и $b \in \mathbb{R}$, $a < b$. Рассмотрим для любого положительного ε функцию $h(x, \varepsilon, a, b)$ следующего вида:

$$h(x, \varepsilon, a, b) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{\varepsilon}, & a < x \leq a + \varepsilon, \\ 1, & a + \varepsilon < x \leq b, \\ \frac{b+\varepsilon-x}{\varepsilon}, & b < x \leq b + \varepsilon, \\ 0, & x > b + \varepsilon. \end{cases}$$

Очевидно, что $h(x, \varepsilon, a, b) \rightarrow I(a, b]$ при $\varepsilon \rightarrow 0$, при этом $0 \leq h(x, \varepsilon, a, b) \leq 1$. Мы хотим доказать справедливость следующего равенства:

$$\int_{-\infty}^{+\infty} h(x, \varepsilon, a, b) dF(x) = \int_{-\infty}^{+\infty} h(x, \varepsilon, a, b) dG(x). \quad (1.7)$$

Можем утверждать, что как только будет доказана истинность (1.7), то будет доказана вся теорема. Действительно, переходя к пределу при $\varepsilon \rightarrow 0$, по теореме Лебега из (1.7) получаем, что $F(b) - F(a) = G(b) - G(a)$,

устремляя $a \rightarrow -\infty$, переходим к равенству $F(b) = G(b)$ при любом $b \in \mathbb{R}$.

Выберем $\varepsilon > 0$ и натуральное число n такое, что $(-n, n) \supset [a, b + \varepsilon]$. Рассмотрим функцию $h(x, \varepsilon, a, b)$ на сегменте $[-n, n]$, найдем преобразование, которое устанавливает взаимнооднозначное соответствие между промежутками $[-n, n]$ и $[-\pi, \pi]$.

Пусть $y = \frac{\pi}{n}x$, тогда $x = \frac{n}{\pi}y$. Рассмотрим функцию:

$$g(y) = h\left(\frac{n}{\pi}y, \varepsilon, a, b\right),$$

$g(-\pi) = g(\pi) = 0$. Воспользуемся теоремой Вейерштрасса, которая говорит о том, что непрерывная функция на отрезке $[-\pi, \pi]$, принимающая на его концах одинаковые значения в равномерной метрике с любой степенью точности, аппроксимируется тригонометрическим многочленом, а именно: для любого δ_n из интервала $(0, 1)$ существует тригонометрический многочлен

$$T_m(y) = A_0 + \sum_{k=1}^m (A_k \cos(ky) + B_k \sin(ky)),$$

такой что для любого $y \in [-\pi, \pi]$ выполняется неравенство:

$$|g(y) - T_m(y)| \leq \delta_n. \quad (1.8)$$

Так как

$$\begin{aligned} e^{iky} &= \cos(ky) + i \sin(ky), \\ e^{-iky} &= \cos(ky) - i \sin(ky), \end{aligned}$$

то

$$\begin{aligned} \cos(ky) &= \frac{e^{iky} + e^{-iky}}{2}, \\ \sin(ky) &= \frac{e^{iky} - e^{-iky}}{2i}, \end{aligned}$$

и, следовательно, можно записать тригонометрический многочлен $T_m(y)$ в виде:

$$T_m(y) = \sum_{k=-m}^m C_k e^{iky},$$

где C_k — комплексные числа. При этом $|T_m(y)| \leq 2$ для любого $y \in \mathbb{R}$, так как $-\delta_n - g(y) \leq T_m(y) \leq \delta_n + g(y)$. Перепишем (1.8), сделав замену:

$$\left| h(x, \varepsilon, a, b) - T_m\left(\frac{\pi}{n}x\right) \right| \leq \delta_n \quad (1.9)$$

для любого $x \in [-n, n]$.

$$\begin{aligned} & \left| \int_{-\infty}^{+\infty} h(x, \varepsilon, a, b) dF(x) - \int_{-\infty}^{+\infty} h(x, \varepsilon, a, b) dG(x) \right| = \\ & = \left| \int_{-n}^n h(x, \varepsilon, a, b) dF(x) - \int_{-n}^n h(x, \varepsilon, a, b) dG(x) \right| = \\ & = \left| \int_{-n}^n \left(h(x, \varepsilon, a, b) - T_m\left(\frac{\pi}{n}x\right) \right) dF(x) + \right. \\ & \quad \left. + \int_{-n}^n \left(T_m\left(\frac{\pi}{n}x\right) - h(x, \varepsilon, a, b) \right) dG(x) + \right. \\ & \quad \left. + \int_{-n}^n T_m\left(\frac{\pi}{n}x\right) dF(x) - \int_{-n}^n T_m\left(\frac{\pi}{n}x\right) dG(x) \right| \leq \\ & \leq \int_{-n}^n \left| h(x, \varepsilon, a, b) - T_m\left(\frac{\pi}{n}x\right) \right| dF(x) + \\ & + \int_{-n}^n \left| T_m\left(\frac{\pi}{n}x\right) - h(x, \varepsilon, a, b) \right| dG(x) + \left| \int_{-n}^n T_m\left(\frac{\pi}{n}x\right) dF(x) - \right. \\ & \quad \left. - \int_{-n}^n T_m\left(\frac{\pi}{n}x\right) dG(x) \right| \leq 2\delta_n + \left| \int_{-\infty}^{+\infty} T_m\left(\frac{\pi}{n}x\right) dF(x) - \right. \\ & \quad \left. - \int_{-\infty}^{+\infty} T_m\left(\frac{\pi}{n}x\right) dG(x) - \int_S T_m\left(\frac{\pi}{n}x\right) dF(x) + \right. \end{aligned}$$

$$\begin{aligned}
& \left| + \int_S T_m \left(\frac{\pi}{n} x \right) dG(x) \right| \leq 2\delta_n + \\
& + \left| \sum_{k=-m}^m C_k \left(\int_{-\infty}^{+\infty} e^{i \frac{k}{n} \pi x} dF(x) - \int_{-\infty}^{+\infty} e^{i \frac{k}{n} \pi x} dG(x) \right) \right| + \\
& + \int_S \left| T_m \left(\frac{\pi}{n} x \right) \right| dF(x) + \int_S \left| T_m \left(\frac{\pi}{n} x \right) \right| dG(x) \leq \\
& \leq 2\delta_n + 2(F(-n) + 1 - F(n)) + 2(G(-n) + 1 - G(n)) \rightarrow 0,
\end{aligned}$$

где $S = (-\infty, -n) \cup (n, +\infty)$. Правая часть стремится к нулю с ростом n , в то время как левая часть от n не зависит. Теорема доказана. ■

Замечание 1.2. Теорема остается справедливой, если речь идет о функциях распределения в пространстве \mathbb{R}^n .

9. Пусть задан случайный вектор $\xi^T = (\xi_1, \dots, \xi_n)$, случайные величины ξ_1, \dots, ξ_n взаимно независимы тогда и только тогда, когда $\varphi_\xi(t_1, \dots, t_n) = \prod_{k=1}^n \varphi_{\xi_k}(t_k)$.

Доказательство. Необходимость очевидна. Если компоненты взаимно независимы, то можно применить свойство мультипликативности.

Докажем достаточность:

$$\begin{aligned}
& \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{it^T x} dF(x) = \\
& = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{it^T x} d(F_{\xi_1}(x_1) \cdot \dots \cdot F_{\xi_n}(x_n)).
\end{aligned}$$

Следовательно, применяя многомерный аналог теоремы 1.1, получаем равенство:

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \cdot \dots \cdot F_{\xi_n}(x_n),$$

что является необходимым и достаточным условием взаимной независимости случайных величин. ■

10. Справедливо следующее равенство:

$$\varphi_{\xi_1, \dots, \xi_n}(t_1, \dots, t_n)|_{t_{k+1}=0, \dots, t_n=0} = \varphi_{\xi_1, \dots, \xi_k}(t_1, \dots, t_k).$$

Доказательство следует из определения характеристической функции.

Замечание 1.3. Свойство сохраняется, если мы выберем любой набор аргументов (они не обязательно должны идти по порядку), которые затем приравняем к нулю.

11. Свойство представляет собой теорему.

Теорема 1.2 (теорема Бохнера–Хинчина). *Непрерывная функция $\varphi(t)$, заданная на числовой прямой, $\varphi(0) = 1$, является характеристической тогда и только тогда, когда она неотрицательно определена, т. е. для любых t_1, \dots, t_n и любых комплексных чисел c_1, \dots, c_n выполнено неравенство:*

$$\sum_{r=1}^n \sum_{k=1}^n \varphi(t_r - t_k) c_r \bar{c}_k \geq 0.$$

Сформулированная теорема дает необходимое и достаточное условие того, что функция является характеристической.

§ 2. ПРИМЕРЫ ХАРАКТЕРИСТИЧЕСКИХ ФУНКЦИЙ

Пример 2.1. Пусть случайная величина ξ почти наверное равняется константе a , тогда ее характеристическая функция будет иметь вид: $\varphi_\xi(t) = e^{ita}$. Действительно, $\varphi_\xi(t) = Ee^{i\xi t} = Ee^{iat} = e^{ita}$.

Пример 2.2. Пусть случайная величина η подчиняется распределению Бернулли, т. е. она принимает значение 1 с вероятностью p и значение 0 с вероятностью $q = 1 - p$, тогда ее характеристическая функция будет иметь вид: $\varphi_\eta(t) = Ee^{i\eta t} = pe^{it} + q$.

Пример 2.3. Пусть случайная величина μ подчиняется биномиальному распределению, $P\{\mu = m\} = C_n^m p^m q^{n-m}$, $m = 0, \dots, n$. Заметим, что $\mu = \sum_{k=1}^n \xi_k$, где ξ_k — случайная величина, принимающая значение «1», если в k -ом испытании произошел успех, и значение «0», если k -ое испытание закончилось неудачей. Случайные величины ξ_1, \dots, ξ_n взаимно независимые бернуллевские величины, тогда, используя свойство 3 характеристических функций, получаем

$$\varphi_\mu(t) = \prod_{k=1}^n \varphi_{\xi_k}(t) = (pe^{it} + q)^n.$$

Пример 2.4. Рассмотрим распределение Пуассона: $P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, 2, \dots$. Найдем характеристическую функцию случайной величины ξ :

$$\varphi_\xi(t) = Ee^{i\xi t} = \sum_{k=0}^{\infty} e^{ikt} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

§ 3. ФОРМУЛЫ ОБРАЩЕНИЯ

Рассмотрим характеристическую функцию

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x). \quad (3.1)$$

Каждой функции распределения случайной величины взаимно однозначным образом соответствует характеристическая функция. Хотелось бы получить формулу преобразования, обратного преобразованию (3.1).

Пусть существует плотность распределения $f(x)$, тогда по определению:

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

В этом случае:

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx. \quad (3.2)$$

Формула (3.2) представляет собой преобразование Фурье функции плотности распределения $f(x)$. В математическом анализе есть теорема об обратном преобразовании: пусть интеграл $\int_{-\infty}^{+\infty} |\varphi(t)| dt \in \mathbb{R}$ сходится, тогда плотность распределения $f(x)$ представима следующим образом:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi(t) dt. \quad (3.3)$$

Задача — опираясь на выражение (3.3), найти общую формулу обращения.

Пусть дана случайная величина ξ и характеристическая функция $\varphi_{\xi}(t)$. Пусть случайная величина η подчиняется стандартному нормальному распределению, и случайные величины ξ и η взаимно независимы. Введем малый параметр $\sigma > 0$ и рассмотрим сумму случайных величин $\xi + \sigma\eta$. Ее характеристической функцией по ранее доказанным свойствам будет следующее выражение:

$$\varphi_{\xi+\sigma\eta}(t) = \varphi_{\xi}(t) e^{-\frac{\sigma^2}{2} t^2},$$

при этом, как будет показано в следующем параграфе, $\varphi_{\eta}(t) = e^{-\frac{1}{2} t^2}$, в силу первого свойства характеристических функций получаем, что для любого $\sigma \neq 0$ справедливо неравенство:

$$\int_{-\infty}^{+\infty} |\varphi_{\xi+\sigma\eta}(t)| dt = \int_{-\infty}^{+\infty} |\varphi_{\xi}(t)| e^{-\frac{\sigma^2}{2} t^2} dt < +\infty,$$

так как $|\varphi_{\xi}(t)| \leq 1$.

Найдем плотность распределения случайной величины $\xi + \sigma\eta$:

$$\begin{aligned} F_{\xi+\sigma\eta}(z) &= P\{\xi + \sigma\eta \leq z\} = \\ &= \iint_L dF_{\xi}(x_1) dF_{\sigma\eta}(x_2) = \iint_L dF_{\xi}(x_1) f_{\sigma\eta}(x_2) dx_2 = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x_1} f_{\sigma\eta}(x_2) dx_2 dF_{\xi}(x_1), \quad (3.4) \end{aligned}$$

где $L = \{(x_1, x_2) : x_1 + x_2 \leq z\}$.

Сделаем замену переменных $s = x_1 + x_2$, тогда $ds = dx_2$, равенство (3.4) примет следующий вид:

$$F_{\xi+\sigma\eta}(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^z f_{\sigma\eta}(s - x_1) ds dF_{\xi}(x_1) =$$

так как по теореме Тонелли–Фубини можно поменять порядок интегрирования:

$$= \int_{-\infty}^z \left(\int_{-\infty}^{+\infty} f_{\sigma\eta}(s - x_1) dF_{\xi}(x_1) \right) ds. \quad (3.5)$$

Выражение в скобках согласно определению будет плотностью распределения. Таким образом, плотность распределения случайной величины $\xi + \sigma\eta$ существует, и характеристическая функция интегрируема, следовательно, справедливо представление:

$$f_{\xi+\sigma\eta}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi_{\xi}(t) e^{-\frac{\sigma^2}{2}t^2} dt. \quad (3.6)$$

Возьмем любые точки a и b , являющиеся точками непрерывности функции $F_{\xi}(x)$, пусть $a < b$. Проинтегрируем выражение (3.6) на промежутке $(a, b]$:

$$\begin{aligned} F_{\xi+\sigma\eta}(b) - F_{\xi+\sigma\eta}(a) &= \int_a^b f_{\xi+\sigma\eta}(x) dx = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\xi}(t) e^{-\frac{\sigma^2}{2}t^2} dt. \end{aligned}$$

Следовательно, справедлива формула обращения:

$$F_{\xi}(b) - F_{\xi}(a) = \frac{1}{2\pi} \lim_{\sigma \rightarrow 0} \int_{-\infty}^{+\infty} \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\xi}(t) e^{-\frac{\sigma^2}{2}t^2} dt. \quad (3.7)$$

Получим эквивалентное представление формулы обращения. Действительно,

$$\begin{aligned} \lim_{\sigma \rightarrow 0} \lim_{A \rightarrow \infty} \int_{-A}^A \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\xi}(t) e^{-\frac{\sigma^2}{2} t^2} dt = \\ = \lim_{A \rightarrow \infty} \int_{-A}^A \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\xi}(t) dt. \end{aligned}$$

Следовательно, в точках непрерывности a и b справедливо представление:

$$F_{\xi}(b) - F_{\xi}(a) = \frac{1}{2\pi} \lim_{A \rightarrow \infty} \int_{-A}^A \frac{e^{-ita} - e^{-itb}}{it} \varphi_{\xi}(t) dt. \quad (3.8)$$

Покажем обоснованность предельного перехода в формуле (3.7). Возьмем любую последовательность $\sigma_n \xrightarrow{n \rightarrow \infty} 0$, $\sigma_n > 0$, тогда для любого положительного ε выполнено следующее:

$$\begin{aligned} P\{|\xi + \sigma_n \eta - \xi| > \varepsilon\} &= P\{|\sigma_n \eta| > \varepsilon\} = P\left\{|\eta| > \frac{\varepsilon}{\sigma_n}\right\} \leq \\ &\leq 1 - F_{\eta}\left(\frac{\varepsilon}{\sigma_n}\right) + F_{\eta}\left(\frac{-\varepsilon}{\sigma_n}\right) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Следовательно, имеет место сходимост по вероятности:

$$\xi + \sigma_n \eta \xrightarrow[n \rightarrow \infty]{P} \xi.$$

Как известно, из сходимости по вероятности следует сходимость по распределению:

$$\xi + \sigma_n \eta \xrightarrow[n \rightarrow \infty]{d} \xi,$$

что эквивалентно сходимости:

$$F_{\xi + \sigma_n \eta}(x) \xrightarrow[n \rightarrow \infty]{} F_{\xi}(x),$$

если точка x — точка непрерывности функции распределения F_{ξ} .

§ 4. ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ ДЛЯ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Пусть случайная величина ξ подчиняется стандартному нормальному распределению. Распределение симметрично, следовательно, по свойству 6 характеристическая функция принимает только вещественные значения:

$$\varphi_{\xi}(t) = Ee^{i\xi t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx.$$

Используя свойство 7 характеристических функций, возьмем производную от характеристической функции, получим

$$\begin{aligned} \varphi'_{\xi}(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ix e^{itx} e^{-\frac{x^2}{2}} dx = \\ &= -\frac{i}{\sqrt{2\pi}} e^{itx} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \frac{i^2 t}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx = -t\varphi_{\xi}(t). \end{aligned}$$

Таким образом, получается дифференциальное уравнение $\varphi'_{\xi}(t) = -t\varphi_{\xi}(t)$ с начальным условием $\varphi_{\xi}(0) = 1$, у которого существует единственное решение:

$$\varphi_{\xi}(t) = e^{-\frac{t^2}{2}}. \quad (4.1)$$

Рассмотрим теперь случайную величину $\eta = \sigma\xi + a$, $\sigma > 0$. Запишем выражение для функции распределения случайной величины η :

$$F_{\eta}(x) = P\{\eta \leq x\} = P\left\{\xi \leq \frac{x-a}{\sigma}\right\} = F_{\xi}\left(\frac{x-a}{\sigma}\right).$$

Плотность распределения случайной величины η имеет вид:

$$f_{\eta}(x) = \frac{1}{\sigma} f_{\xi}\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Говорят, что случайная величина η распределена нормально с параметрами a и σ^2 , $N(a, \sigma^2)$. Используя свойство 2 характеристических функций, найдем $\varphi_\eta(t)$:

$$\varphi_\eta(t) = e^{iat} \varphi_\xi(\sigma t) = e^{iat - \frac{\sigma^2 t^2}{2}}. \quad (4.2)$$

Перейдем к случайному вектору $\xi^T = (\xi_1, \dots, \xi_n)$, где ξ_k при любом номере k подчиняется стандартному нормальному распределению, $N(0, 1)$, и случайные величины ξ_1, \dots, ξ_n взаимно независимы. Плотность распределения вектора ξ имеет вид:

$$\begin{aligned} f_\xi(x_1, \dots, x_n) &= \prod_{k=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{x_k^2}{2}} = \\ &= \frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum_{k=1}^n x_k^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} x^T x}. \end{aligned}$$

Характеристическая функция случайного вектора ξ будет иметь вид:

$$\varphi_\xi(t_1, \dots, t_n) = \prod_{k=1}^n \varphi_{\xi_k}(t_k) = \prod_{k=1}^n e^{-\frac{t_k^2}{2}} = e^{-\frac{1}{2} t^T t}. \quad (4.3)$$

Рассмотрим теперь случайный вектор $\eta = A\xi + a$, где A — матрица размерности $n \times n$, причем $|A| > 0$, $a^T = (a_1, \dots, a_n)$. Для любого борелевского множества $B \in \mathcal{B}(\mathbb{R}^n)$:

$$P\{\eta \in B\} = P\{A\xi + a \in B\}.$$

Пусть $B = \{x\}$ — множество векторов, $Ay + a = x$, $y = A^{-1}(x - a)$. Все возможные вектора y образуют множество $\tilde{B} = \{y\}$. Очевидно, что между множествами B и \tilde{B} существует взаимно однозначное соответствие:

$$\begin{aligned} P\{\eta \in B\} &= P\{\xi \in \tilde{B}\} = \int_{\tilde{B}} \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} y^T y} dy = \\ &= \int_B \frac{1}{(2\pi)^{n/2}} \frac{1}{|A|} e^{-\frac{1}{2} (x-a)^T (A^{-1})^T A^{-1} (x-a)} dx. \end{aligned}$$

При интегрировании перешли к новым переменным и вычислили якобиан линейного преобразования, который оказался равным $1/|A|$. Нетрудно заметить, что справедливо равенство:

$$E\eta = (E\eta_1, \dots, E\eta_n)^T = AE\xi + a = a.$$

Выпишем ковариационную матрицу случайного вектора η :

$$D\eta = E\{(\eta - E\eta)(\eta - E\eta)^T\} = AD\xi A^T = AA^T = \Sigma.$$

Очевидно, что $|\Sigma| = |A|^2$. Запишем плотность распределения случайного вектора η :

$$f_\eta(x) = \frac{1}{(2\pi)^{n/2} \sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-a)^T \Sigma^{-1}(x-a)}$$

— плотность многомерного нормального распределения с параметрами (a, Σ) , где $a = E\eta$, $\Sigma = D\eta$, т.е. $N(a, \Sigma)$. Найдём вид характеристической функции случайного вектора η :

$$\begin{aligned} \varphi_\eta(t_1, \dots, t_n) &= Ee^{i\eta^T t} = Ee^{ia^T t + i\xi^T A^T t} = \\ &= e^{ia^T t} \varphi_\xi(A^T t) = e^{ia^T t - \frac{1}{2}t^T \Sigma t}. \end{aligned} \quad (4.4)$$

Нетрудно убедиться, что линейное преобразование не меняет тип нормального распределения. Пусть $\zeta = b + C\eta$, где $b \in \mathbb{R}^k$, размерность C равна $k \times n$, тогда

$$\begin{aligned} \varphi_\zeta(t) &= Ee^{i(b^T + \eta^T C^T)t} = e^{ib^T t} \varphi_\eta(C^T t) = \\ &= e^{i(b^T + a^T C^T)t - \frac{1}{2}t^T C \Sigma C^T t}, \end{aligned} \quad (4.5)$$

т.е. вектор ζ распределен нормально с вектором математических ожиданий $b + Ca$ и ковариационной матрицей $C\Sigma C^T$, другими словами, ζ подчиняется распределению $N(b + Ca, C\Sigma C^T)$.

§ 5. МЕТОД ХАРАКТЕРИСТИЧЕСКИХ ФУНКЦИЙ В ДОКАЗАТЕЛЬСТВЕ ПРЕДЕЛЬНЫХ ТЕОРЕМ

В параграфе рассматривается скалярный случай, но все утверждения можно перенести и на векторный случай.

Определение 5.1. Обобщенной функцией распределения на числовой прямой называется функция $G(x)$, обладающая свойствами:

1. Для любых точек $x_1 \leq x_2$ справедливо неравенство $G(x_1) \leq G(x_2)$.
2. Для любой точки $x \in \mathbb{R}$ имеет место сходимость $G(y) \xrightarrow{y \rightarrow x+0} G(x)$.
3. Выполняются соотношения: $\lim_{x \rightarrow \infty} G(x) = G(\infty) \leq 1$,
 $\lim_{x \rightarrow -\infty} G(x) \geq 0$.

Обозначим через J класс обобщенных функций распределения на числовой прямой.

Теорема 5.1. Пусть $\{G_n\}$ — последовательность обобщенных функций распределения, тогда существует функция $G \in J$ и подпоследовательность $\{G_{n_k}\}$ такие, что подпоследовательность $\{G_{n_k}\}$ сходится в основном к функции G , т. е. для любой точки x из множества непрерывности функции G имеет место сходимость $G_{n_k}(x) \xrightarrow{n_k \rightarrow \infty} G(x)$.

Доказательство. Рассмотрим счетное множество $S = \{x_1, x_2, \dots\}$ плотное в \mathbb{R} . Последовательность $G_n(x_1)$ ограничена, следовательно, существует сходящаяся подпоследовательность $G_{n_k^{(1)}}(x_1)$, $k = 1, 2, \dots$. Перебирая последовательно точки из S , построим «диагональную» последовательность $G_{n_k^{(k)}}(x_l)$, сходящуюся к некоторому пределу, который обозначим через $G_s(x_l)$. Здесь предполагается, что подпоследовательности вложены друг в друга, $\{n_l^{(k+1)}\} \subset \{n_l^{(k)}\}$. Далее определим функцию $G(x) = \inf\{G_s(x_k) : x_k \in S, x_k > x\}$. Нетрудно показать, что функция $G(x)$ является обобщенной функцией распределения и для всех точек x на прямой, где

$G(x)$ непрерывна, выполняется:

$$G(x) = \lim_{k \rightarrow \infty} G_{n_k}(x).$$

■

Теорема 5.2 (метод характеристических функций).

1. Пусть задана последовательность функций распределения $\{F_n\}$, причем эта последовательность сходится в основном к функции распределения F . Пусть

$$\varphi_n(t) = \int_{-\infty}^{\infty} e^{itx} dF_n(x), \quad (5.1)$$

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x). \quad (5.2)$$

Тогда для любой точки t имеет место сходимостъ $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t)$. Таким образом, из сходимости в основном функций распределения следует поточечная сходимостъ характеристических функций.

2. Пусть $\varphi_n(t)$ из (5.1) — характеристическая функция, соответствующая функции распределения $F_n(x)$. Пусть в каждой точке t существует предел $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t)$. Предположим, что функция $\varphi(t)$ непрерывна в точке 0, тогда функция $\varphi(t)$ также является характеристической, т. е. существует функция распределения $F(x)$ такая, что $\varphi(t)$ удовлетворяет (5.2), и при этом, последовательность функций распределения $F_n(x)$ сходится в основном к функции распределения $F(x)$.

Доказательство. Докажем первый пункт теоремы. Для любой функции $F_n(x)$ существует случайная величина ξ_n такая, что $F_{\xi_n}(x) = F_n(x)$, то есть, $F_n(x) = P\{\xi_n \leq x\}$. Для функции распределения $F(x)$ существует случайная величина ξ такая, что выполнено равенство $F_\xi(x) \equiv F(x)$, т. е.

$F(x) = P\{\xi \leq x\}$. Так как сходимость в основном равносильна сходимости по распределению, то имеет место следующая сходимость:

$$Eg(\xi_n) \xrightarrow{n \rightarrow \infty} Eg(\xi) \quad (5.3)$$

для любой непрерывной и ограниченной функции $g(x)$.

Рассмотрим $g(x) = \cos(tx)$ и $g(x) = \sin(tx)$. Подставим их в выражение (5.3) и получим:

$$E \cos(t\xi_n) \xrightarrow{n \rightarrow \infty} E \cos(t\xi),$$

$$E \sin(t\xi_n) \xrightarrow{n \rightarrow \infty} E \sin(t\xi).$$

Но тогда очевидно, что

$$E \cos(t\xi_n) + iE \sin(t\xi_n) \xrightarrow{n \rightarrow \infty} E \cos(t\xi) + iE \sin(t\xi),$$

или

$$\varphi_{\xi_n}(t) \xrightarrow{n \rightarrow \infty} \varphi_{\xi}(t).$$

Докажем второе утверждение теоремы. Для любой случайной величины ζ при любых положительных вещественных постоянных x и τ , удовлетворяющих условию $x\tau > 1$ справедливо неравенство:

$$P\{|\zeta| \leq x\} \geq \frac{\left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi_{\zeta}(t) dt \right| - \frac{1}{x\tau}}{1 - \frac{1}{x\tau}}. \quad (5.4)$$

Докажем неравенство (5.4). Преобразуем выражение под знаком модуля:

$$\begin{aligned} \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi_{\zeta}(t) dt \right| &= \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} E e^{i\zeta t} dt \right| = \\ &= \left| \frac{1}{2\tau} E \int_{-\tau}^{\tau} \{e^{i\zeta t} (I\{\zeta = 0\} + I\{|\zeta| > 0\})\} dt \right|. \end{aligned} \quad (5.5)$$

Используя формулу Эйлера для экспоненты и равенство

$$\begin{aligned} \frac{1}{2\tau} \int_{-\tau}^{\tau} \{e^{i\zeta t} I\{|\zeta| > 0\}\} dt &= \frac{1}{2\tau} \left\{ \frac{\sin(\zeta t)}{\zeta} I\{|\zeta| > 0\} \right\} \Big|_{-\tau}^{\tau} = \\ &= \frac{\sin(\zeta \tau)}{\zeta \tau} I\{|\zeta| > 0\}, \end{aligned}$$

запишем выражение (5.5) в следующем виде:

$$\begin{aligned} \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi_{\zeta}(t) dt \right| &= \left| EI\{\zeta = 0\} + E \left\{ \frac{\sin(\zeta \tau)}{\zeta \tau} I\{|\zeta| > 0\} \right\} \right| \leqslant \\ &\leqslant P\{\zeta = 0\} + P\{0 < |\zeta| \leqslant x\} + \frac{1}{x\tau} P\{|\zeta| > x\} = \\ &= P\{|\zeta| \leqslant x\} + \frac{1}{x\tau} (1 - P\{|\zeta| \leqslant x\}). \end{aligned}$$

Неравенство (5.4) доказано.

Положим $x\tau = 2$, тогда

$$P\left\{|\zeta| \leqslant \frac{2}{\tau}\right\} \geqslant 2 \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi_{\zeta}(t) dt \right| - 1 \quad (5.6)$$

для любого положительного τ . Неравенство (5.6) справедливо при любом $\tau > 0$ для случайных величин ξ_n , т.е. для последовательности $\{F_n(x)\}$. Справедлива теорема 5.1, поэтому существует подпоследовательность $F_{n_k}(x)$ такая, что при $k \rightarrow \infty$ имеет место сходимость в основном последовательности $F_{n_k}(x)$ к функции $G(x) \in J$. Далее будем рассматривать последовательность $F_{n_k}(x)$. Справедливо неравенство:

$$1 \leqslant \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi(t) dt \right| + \frac{1}{2\tau} \int_{-\tau}^{\tau} |1 - \varphi(t)| dt.$$

Так как для любого n выполняются соотношения:

$$\varphi_n(0) = 1, \quad \lim_{n \rightarrow \infty} \varphi_n(0) = \varphi(0) = 1,$$

то для любого положительного ε существует такое положительное τ_0 , что при $|t| \leq \tau_0$ выполнено следующее неравенство:

$$|1 - \varphi(t)| < \frac{\varepsilon}{4}.$$

Будем выбирать $\tau < \tau_0$, но тогда выполнено неравенство:

$$1 - \frac{\varepsilon}{4} \leq \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi(t) dt \right|.$$

Преобразуем интеграл:

$$\int_{-\tau}^{\tau} \varphi(t) dt = \int_{-\tau}^{\tau} \varphi_n(t) dt + \int_{-\tau}^{\tau} (\varphi(t) - \varphi_n(t)) dt,$$

тогда

$$\begin{aligned} 1 - \frac{\varepsilon}{4} &\leq \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi(t) dt \right| \leq \\ &\leq \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi_n(t) dt \right| + \frac{1}{2\tau} \int_{-\tau}^{\tau} |\varphi(t) - \varphi_n(t)| dt, \quad (5.7) \end{aligned}$$

причем $|\varphi(t) - \varphi_n(t)| \xrightarrow{n \rightarrow \infty} 0$ для любого t , кроме того, $|\varphi(t) - \varphi_n(t)| \leq 2$, что следует из свойств характеристической функции. Из данных двух условий следует, что применима теорема Лебега о мажорируемой сходимости, поэтому из неравенства (5.7) следует выполнение следующего неравенства:

$$1 - \frac{\varepsilon}{2} \leq \left| \frac{1}{2\tau} \int_{-\tau}^{\tau} \varphi_n(t) dt \right|$$

при $n \geq n_0$, где n_0 выбрано из условия, что последнее слагаемое в (5.7) не превосходит $\varepsilon/4$. Из неравенства (5.6) и, учитывая, что $\varphi_n(t) = \int_{-\infty}^{+\infty} e^{itx} dF_n(x)$, $F_n(x) \equiv F_{\xi_N}(x)$ следует,

что

$$P \left\{ |\xi_n| \leq \frac{2}{\tau} \right\} \geq 2 \left(1 - \frac{\varepsilon}{2} \right) - 1 = 1 - \varepsilon.$$

Теперь возьмем любое $n \geq n_0$, $0 < \tau \leq \tau_0$, $A < -2/\tau$, $B > 2/\tau$, где точки A и B являются точками непрерывности функции $G(x)$. Будет выполняться неравенство:

$$F_n(B) - F_n(A) \geq P \left\{ |\xi_n| \leq \frac{2}{\tau} \right\} \geq 1 - \varepsilon.$$

Рассмотрим разность $F_{n_k}(B) - F_{n_k}(A) \geq 1 - \varepsilon$. Устремив k к бесконечности, получим

$$G(B) - G(A) \geq 1 - \varepsilon,$$

тогда $G(\infty) = 1$, $G(-\infty) = 0$, иначе возникло бы противоречие. Поэтому $G(x)$ — обычная функция распределения. Обозначим $G(x) \equiv F(x)$.

Так как $F_{n_k}(x)$ сходится в основном к $F(x)$ при $k \rightarrow \infty$, то применимо уже доказанное первое утверждение теоремы, тогда для любого t имеет место сходимостъ:

$$\varphi_{n_k}(t) \rightarrow \int_{-\infty}^{+\infty} e^{itx} dF(x),$$

с другой стороны, для любого t выполняется:

$$\varphi_{n_k}(t) \rightarrow \varphi(t).$$

Из этого следует, что $\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x)$ — характеристическая функция. Теперь докажем сходимостъ в основном $F_n(x)$ к $F(x)$. Докажем методом от противного. Пусть сходимости нет, тогда существует точка x_0 , которая является точкой непрерывности функции распределения $F(x)$ такая, что нет сходимости $F_n(x_0)$ к $F(x_0)$. Тогда пусть $F_{n_k(1)}(x_0)$ сходится к $q \neq F(x_0)$ при $k \rightarrow \infty$. Вернемся к началу доказательства и будем рассматривать последовательность $\{F_{n_k(1)}(x)\}$, повторяя все рассуждения.

Существует $\{F_{n_k^{(2)}}(x)\}$ такая, что

$$\{F_{n_k^{(2)}}(x)\} \subset \{F_{n_k^{(1)}}(x)\},$$

и $F_{n_k^{(2)}}(x)$ сходится в основном к $\tilde{F}(x)$, следовательно, имеют место следующие сходимости:

$$\varphi_{n_k^{(2)}}(t) \xrightarrow[k \rightarrow \infty]{} \int_{-\infty}^{+\infty} e^{itx} d\tilde{F}(x),$$

$$\varphi_{n_k^{(2)}}(t) \xrightarrow[k \rightarrow \infty]{} \varphi(t).$$

Но тогда для любого t справедливы равенства:

$$\int_{-\infty}^{+\infty} e^{itx} d\tilde{F}(x) = \varphi(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x).$$

Пришли к противоречию, так как $\tilde{F}(x) \equiv F(x)$, и $\tilde{F}(x_0) = q \neq F(x_0)$. ■

Замечание 5.1. Утверждение теоремы переносится на многомерный случай.

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

§ 1. ЗАКОН БОЛЬШИХ ЧИСЕЛ ДЛЯ НЕЗАВИСИМЫХ ОДИНАКОВО РАСПРЕДЕЛЕННЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Определение 1.1. Пусть $\xi_1, \dots, \xi_n, \dots$ — последовательность случайных величин, определенных на (Ω, \mathcal{F}, P) . Будем говорить, что к этой последовательности применим закон больших чисел (ЗБЧ), если

$$\frac{\sum_{k=1}^n \xi_k}{n} - \frac{\sum_{k=1}^n E\xi_k}{n} \xrightarrow[n \rightarrow \infty]{P} 0. \quad (1.1)$$

Замечание 1.1. В определении 1.1 предполагается, что все математические ожидания существуют и конечны.

Если $E\xi_k = a$, $k = 1, 2, \dots$, тогда сходимость (1.1) можно записать в виде:

$$\frac{\sum_{k=1}^n \xi_k}{n} \xrightarrow[n \rightarrow \infty]{P} a. \quad (1.2)$$

Теорема 1.1. Пусть $\xi_1, \dots, \xi_n, \dots$ — последовательность взаимно независимых одинаково распределенных случайных величин. Пусть $E\xi_k = a \in \mathbb{R}$ для любого $k = 1, 2, \dots$, тогда к последовательности $\{\xi_k\}$ применим закон больших чисел, т. е.

$$\zeta_n = \frac{1}{n} \sum_{k=1}^n \xi_k \xrightarrow[n \rightarrow \infty]{P} a. \quad (1.3)$$

Доказательство. Воспользуемся методом характеристических функций и докажем, что $\zeta_n \xrightarrow[n \rightarrow \infty]{d} \zeta \stackrel{\text{п.н.}}{=} a$, тогда по теореме 3.3 из главы 9 имеет место сходимость: $\zeta_n \xrightarrow[n \rightarrow \infty]{P} a$. Рассмотрим характеристическую функцию:

$$\begin{aligned}\varphi_{\zeta_n}(t) &= Ee^{it\zeta_n} = Ee^{i\frac{t}{n} \sum_{k=1}^n \xi_k} = E \prod_{k=1}^n e^{i\frac{t}{n} \xi_k} = \\ &= \prod_{k=1}^n Ee^{i\frac{t}{n} \xi_k} = \varphi_{\xi_1}^n \left(\frac{t}{n} \right) = e^{n \ln \varphi_{\xi_1} \left(\frac{t}{n} \right)}.\end{aligned}$$

Выберем главную ветвь $\ln 1 = 0$. С ростом n аргумент будет стремиться к нулю. Тогда из свойства 7 характеристических функций следует равенство:

$$\varphi_{\xi_1} \left(\frac{t}{n} \right) = 1 + \frac{ita}{n} + o \left(\frac{1}{n} \right),$$

$$\ln(1 + \varepsilon) = \varepsilon + o(\varepsilon).$$

Таким образом, $e^{n \ln \varphi_{\xi_1} \left(\frac{t}{n} \right)} = e^{n \{ i\frac{t}{n} a + o(\frac{1}{n}) \}} = e^{ita + no(\frac{1}{n})} \xrightarrow[n \rightarrow \infty]{} e^{ita}$, где e^{ita} взаимно однозначно соответствует вырожденному в точке a распределению $P\{\zeta = a\} = 1$. Получаем, что сходимость в основной последовательности $F_{\zeta_n}(x)$ к функции $F_{\zeta}(x)$ равносильна сходимости $\zeta_n \xrightarrow[n \rightarrow \infty]{d} \zeta$. ■

Следствие 1.1. Пусть $\xi_1, \dots, \xi_n, \dots$ — последовательность случайных величин, причем $P\{\xi_k = 1\} = p$, $P\{\xi_k = 0\} = q = 1 - p$. Обозначим через μ число успехов в n испытаниях (схема Бернулли), т. е. $\mu = \sum_{k=1}^n \xi_k$. Так как $\{\xi_k\}$ — последовательность независимых и одинаково распределенных случайных величин, и $E\xi_k = 1 \cdot p + 0 \cdot q = p$, тогда

$$\frac{\mu}{n} \xrightarrow[n \rightarrow \infty]{P} p. \quad (1.4)$$

Доказательство. Очевидным образом следует из теоремы 1.1. ■

§ 2. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ДЛЯ НЕЗАВИСИМЫХ ОДИНАКОВО РАСПРЕДЕЛЕННЫХ СЛУЧАЙНЫХ ВЕЛИЧИН И СЛУЧАЙНЫХ ВЕКТОРОВ

Определение 2.1. Пусть $\eta_1, \dots, \eta_k, \dots$ — последовательность случайных величин, определенных на (Ω, \mathcal{F}, P) . Будем говорить, что к последовательности $\{\eta_k\}$ применима центральная предельная теорема (ЦПТ), если для любого $x \in \mathbb{R}$ имеет место сходимость

$$P \left\{ \frac{\eta_k - E\eta_k}{\sqrt{D\eta_k}} \leq x \right\} \xrightarrow{k \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy. \quad (2.1)$$

Замечание 2.1. В определении 2.1 предполагается, что все математические ожидания $E\eta_k$ и дисперсии $D\eta_k$ существуют и конечны. Обозначим через η случайную величину, подчиняющуюся стандартному нормальному распределению. Функция распределения случайной величины η будет иметь вид:

$$F_\eta(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy.$$

Тогда сходимость (2.1) можно записать следующим образом:

$$\frac{\eta_k - E\eta_k}{\sqrt{D\eta_k}} \xrightarrow[k \rightarrow \infty]{d} \eta. \quad (2.2)$$

Теорема 2.1 (центральная предельная теорема для независимых одинаково распределенных случайных величин). Пусть $\xi_1, \dots, \xi_n, \dots$ — последовательность взаимно независимых, одинаково распределенных случайных величин. Пусть $E\xi_k = a \in \mathbb{R}$, $D\xi_k = \sigma^2 \in (0, \infty)$ для любого $k = 1, 2, \dots$, тогда для последовательности случайных величин имеет место сходимость:

$$\zeta_n = \frac{\sum_{k=1}^n (\xi_k - a)}{\sqrt{n}\sigma} \xrightarrow[n \rightarrow \infty]{d} \zeta \sim N(0, 1),$$

т. е. функция распределения сходится равномерно по $x \in \mathbb{R}$:

$$F_{\zeta_n}(x) = P\{\zeta_n \leq x\} \xrightarrow{n \rightarrow \infty} \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Доказательство. Воспользуемся методом характеристических функций:

$$\begin{aligned} \varphi_{\zeta_n}(t) &= E e^{it\zeta_n} = E e^{i \frac{t}{\sigma\sqrt{n}} \sum_{k=1}^n (\xi_k - a)} = E \prod_{k=1}^n e^{i \frac{t}{\sigma\sqrt{n}} (\xi_k - a)} = \\ &= \prod_{k=1}^n \varphi_{(\xi_k - a)} \left(\frac{t}{\sigma\sqrt{n}} \right) = \varphi_{(\xi_1 - a)}^n \left(\frac{t}{\sigma\sqrt{n}} \right) = e^{n \ln \varphi_{(\xi_1 - a)} \left(\frac{t}{\sigma\sqrt{n}} \right)}. \end{aligned}$$

Выберем главную ветвь $\ln 1 = 0$, $\ln(1+\varepsilon) = \varepsilon + o(\varepsilon)$. Используя свойство 7 характеристических функций, получаем равенство:

$$\varphi_{\xi_1 - a} \left(\frac{t}{\sigma\sqrt{n}} \right) = 1 + \frac{\left(i \frac{t}{\sigma\sqrt{n}} \right)^2}{2!} \sigma^2 + o \left(\frac{1}{n} \right).$$

По условию теоремы $E(\xi_1 - a) = 0$, $E(\xi_1 - a)^2 = D_{\xi_1} = \sigma^2$, тогда имеет место сходимость:

$$e^{n \ln \varphi_{(\xi_1 - a)} \left(\frac{t}{\sigma\sqrt{n}} \right)} = e^{n \left\{ -\frac{t^2}{2n} + o\left(\frac{1}{n}\right) \right\}} \xrightarrow{n \rightarrow \infty} e^{-\frac{t^2}{2}}.$$

Таким образом, получили характеристическую функцию $e^{-\frac{t^2}{2}}$ взаимно однозначно соответствующую стандартному нормальному распределению, тогда $\zeta_n \xrightarrow[n \rightarrow \infty]{d} \zeta \sim N(0, 1)$. ■

Замечание 2.2. В теореме 2.1 пусть $\eta_n = \sum_{l=1}^n \xi_l$, тогда $D\eta_n = \sigma^2 n$.

Следствие 2.1. Пусть $\xi_1, \dots, \xi_n, \dots$ — последовательность случайных величин, причем $P\{\xi_k = 1\} = p$, $P\{\xi_k = 0\} =$

$= q = 1 - p$. Обозначим через μ число успехов в серии из n испытаний (схема Бернулли), т. е. $\mu = \sum_{k=1}^n \xi_k$. Для любого $k = 1, 2, \dots$: $E\xi_k = p$, $D\xi_k = pq$, тогда $D\mu = npq$. Имеет место сходимость:

$$P \left\{ a \leq \frac{\mu - np}{\sqrt{npq}} \leq b \right\} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy \quad (2.3)$$

равномерно по $a < b$.

Доказательство. Из теоремы 2.1 следует, что

$$P \left\{ a < \frac{\mu - np}{\sqrt{npq}} \leq b \right\} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy.$$

По локальной предельной теореме имеет место сходимость:

$$P \left\{ \frac{\mu - np}{\sqrt{npq}} = a \right\} \xrightarrow{n \rightarrow \infty} 0.$$

■

Следствие 2.2. Пусть справедливы условия теоремы 2.1.

Обозначим через $\tilde{\zeta}_n = \frac{\sum_{k=1}^n (\xi_k - a)}{\sqrt{n}}$. Тогда имеет место сходимость

$$\tilde{\zeta}_n \xrightarrow[n \rightarrow \infty]{d} \tilde{\zeta} \sim N(0, \sigma^2), \quad (2.4)$$

что эквивалентно равномерной по $x \in \mathbb{R}$ сходимости:

$$P \{ \tilde{\zeta}_n \leq x \} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{y^2}{2\sigma^2}} dy. \quad (2.5)$$

Доказательство. Рассмотрим $P \{ \tilde{\zeta}_n \leq x \}$ и применим теорему 2.1:

$$P \{ \tilde{\zeta}_n \leq x \} = P \left\{ \frac{\tilde{\zeta}_n}{\sigma} \leq \frac{x}{\sigma} \right\} \xrightarrow{n \rightarrow \infty} \int_{-\infty}^{\frac{x}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Сделаем замену $v = \sigma y$, $dy = dv/\sigma$, тогда функция распределения $\int_{-\infty}^{\frac{x}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{v^2}{2\sigma^2}} dv$, соответствует нормальному распределению $N(0, \sigma^2)$, что и означает сходимость в основном, а из сходимости в основном следует сходимость по распределению. ■

Следствие 2.3 (центральная предельная теорема для независимых одинаково распределенных случайных векторов). Пусть $\{\xi_n^T = (\xi_{1n}, \dots, \xi_{mn})\}$ — последовательность взаимно независимых одинаково распределенных случайных векторов (компоненты внутри векторов могут быть зависимы). Пусть $E\xi_{kn} = a_k \in \mathbb{R}$, и $E\xi_n = (a_1, \dots, a_m)^T = a$. Пусть дисперсионная матрица $D\xi_n$ существует, обозначим ее через $\Sigma = D\xi_n = E\{(\xi_n - a)(\xi_n - a)^T\}$, тогда случайный вектор $\varsigma_n = 1/\sqrt{n} \sum_{k=1}^n (\xi_k - a)$ сходится по распределению к случайному вектору ς , распределенному по многомерному нормальному закону с параметрами $(0, \Sigma)$, т. е.

$$\varsigma_n \xrightarrow[n \rightarrow \infty]{d} \varsigma \sim N(0, \Sigma). \quad (2.6)$$

Доказательство. Теорема о методе характеристических функций в доказательстве предельных теорем верна и в многомерном случае. Рассмотрим многомерную характеристическую функцию и докажем, что справедлива сходимость:

$$\varphi_{\varsigma_n}(t_1, \dots, t_m) = E e^{i(t_1, \dots, t_m) \varsigma_n} \xrightarrow[n \rightarrow \infty]{} e^{-\frac{1}{2}(t_1, \dots, t_m) \Sigma (t_1, \dots, t_m)^T}.$$

Возьмем произвольный вектор $\theta = (\theta_1, \dots, \theta_m)^T$, тогда $\{\theta^T \xi_n\}$ — последовательность независимых случайных величин с математическими ожиданиями $E(\theta^T \xi_n) = \theta^T a$ и дисперсиями $D(\theta^T \xi_n) = E\{(\theta^T \xi_n - \theta^T a)^2\} = E\{\theta^T (\xi_n - a)(\xi_n - a)^T \theta\} = \theta^T E\{(\xi_n - a)(\xi_n - a)^T\} \theta = \theta^T \Sigma \theta$.

Таким образом, можно применить теорему 2.1, тогда

$$\theta^T \varsigma_n = \frac{\sum_{k=1}^n (\theta^T \xi_k - \theta^T a)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \delta \sim N(0, \theta^T \Sigma \theta).$$

Следовательно, имеет место сходимость:

$$Ee^{it\theta^T\zeta_n} \xrightarrow{n \rightarrow \infty} e^{-\frac{1}{2}t\theta^T\Sigma\theta t},$$

где $t \in \mathbb{R}$, $t\theta^T = (t\theta_1, \dots, t\theta_m) = (t_1, \dots, t_m)$. Таким образом, получаем:

$$\varphi_{\zeta_n}(t_1, \dots, t_m) = Ee^{i(t_1, \dots, t_m)\zeta_n} \xrightarrow{n \rightarrow \infty} e^{-\frac{1}{2}(t_1, \dots, t_m)\Sigma(t_1, \dots, t_m)^T}.$$

■

§ 3. ЗАКОН БОЛЬШИХ ЧИСЕЛ ДЛЯ НЕЗАВИСИМЫХ ПРОИЗВОЛЬНО РАСПРЕДЕЛЕННЫХ СЛУЧАЙНЫХ ВЕЛИЧИН. СХЕМА СЕРИЙ

Рассмотрим последовательность $\{\xi_k\}$ взаимно независимых и произвольно распределенных случайных величин. Пусть существует $E\xi_k = a_k \in \mathbb{R}$. Определим, при каком условии имеет место сходимость по вероятности:

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n a_k \xrightarrow[n \rightarrow \infty]{P} 0. \quad (3.1)$$

Обозначим через $\varsigma_n = \frac{1}{n} \sum_{k=1}^n (\xi_k - a_k)$, тогда сходимость в (3.1) можно записать в виде:

$$\varsigma_n \xrightarrow[n \rightarrow \infty]{P} 0.$$

Обозначим через ξ_{kn} случайную величину $\xi_{kn} = (\xi_k - a_k)/n$, тогда $\varsigma_n = \frac{1}{n} \sum_{k=1}^n (\xi_k - a_k) = \sum_{k=1}^n \xi_{kn}$. Далее будем рассматривать более общую задачу, введя произвольную схему серий, которая будет удовлетворять сформулированным ниже условиям.

Назовем бесконечную последовательность конечных наборов случайных величин $\{\xi_{1n}, \xi_{2n}, \dots, \xi_{nn}\}$, $n = 1, 2, \dots$, схемой серий, будем предполагать, что для любых n и k :

$$E\xi_{kn} = 0. \quad (3.2)$$

В любом конечном наборе $\xi_{1n}, \xi_{2n}, \dots, \xi_{nn}$ случайные величины взаимно независимы.

Как видим, определенные ранее случайные величины ξ_{kn} представляют собой частный случай схемы серий. Докажем вспомогательные леммы.

Лемма 3.1. Пусть β — комплексное число такое, что $\operatorname{Re} \beta \leq 0$, тогда имеют место неравенства:

$$|e^\beta - 1| \leq |\beta|, \quad (3.3)$$

$$|e^\beta - 1 - \beta| \leq \frac{|\beta|^2}{2}, \quad (3.4)$$

$$|e^\beta - 1 - \beta - \frac{\beta^2}{2}| \leq \frac{|\beta|^3}{6}. \quad (3.5)$$

Доказательство. Применим формулу Ньютона–Лейбница на комплексной плоскости к аналитической функции:

$$e^\beta - 1 = \int_0^\beta e^z dz = \beta \int_0^1 e^{\beta v} dv,$$

где $v \in [0, 1]$, тогда справедливо неравенство:

$$|e^\beta - 1| \leq |\beta| \int_0^1 |e^{\beta v}| dv \leq |\beta|.$$

Аналогично доказывается второе неравенство:

$$e^\beta - 1 - \beta = \int_0^\beta (e^z - 1) dz = \beta \int_0^1 (e^{\beta v} - 1) dv,$$

где $v \in [0, 1]$, тогда

$$|e^\beta - 1 - \beta| \leq |\beta| \int_0^1 |e^{\beta v} - 1| dv \leq \frac{|\beta|^2}{2}.$$

Третье неравенство доказывается аналогичными рассуждениями. ■

Лемма 3.2. Пусть существуют комплексные числа $a_1, \dots, a_n, b_1, \dots, b_n$ и для любого $k = 1, \dots, n$ имеют место неравенства: $|a_k| \leq 1, |b_k| \leq 1$. Тогда, если $A_m = \prod_{k=1}^m a_k, B_m = \prod_{k=1}^m b_k$, при $m = 1, \dots, n$, то справедливо неравенство:

$$|B_n - A_n| \leq \sum_{k=1}^n |b_k - a_k|. \quad (3.6)$$

Доказательство. Сделаем преобразования:

$$\begin{aligned} B_n - A_n &= B_{n-1}b_n - A_{n-1}a_n = B_{n-1}b_n - B_{n-1}a_n + \\ &+ B_{n-1}a_n - A_{n-1}a_n = B_{n-1}(b_n - a_n) + a_n(B_{n-1} - A_{n-1}). \end{aligned}$$

Заметим, что $|B_{n-1}| = \prod_{k=1}^{n-1} |b_k| \leq 1$, тогда $|B_n - A_n| \leq |b_n - a_n| + |B_{n-1} - A_{n-1}|$. Повторяя рассуждения, получаем искомую оценку. ■

Будем рассматривать определенную выше схему серий.

Определение 3.1. Будем говорить, что выполнено условие (M_1) , если для любого $\tau > 0$ имеет место сходимостъ:

$$M_{1n}(\tau) = \sum_{k=1}^n E \{ |\xi_{kn}| I\{|\xi_{kn}| > \tau\} \} \xrightarrow{n \rightarrow \infty} 0. \quad (3.7)$$

Определение 3.2. Будем говорить, что выполнено условие (D_1) , если существует $s \in (1, 2]$ такое, что имеет место сходимостъ:

$$D_{1n}(s) = \sum_{k=1}^n E \{ \min(|\xi_{kn}|, |\xi_{kn}|^s) \} \xrightarrow{n \rightarrow \infty} 0. \quad (3.8)$$

Замечание 3.1. Если $s_1 < s_2$, то $D_{1n}(s_1) \geq D_{1n}(s_2)$ для $s_1 \in (1, 2], s_2 \in (1, 2]$, т. е. функция D_{1n} убывает.

Замечание 3.2. Если $\tau_1 < \tau_2$, то $M_{1n}(\tau_1) \geq M_{1n}(\tau_2)$ для $\tau > 0$, т. е. функция M_{1n} убывает.

Функцию $D_{1n}(s)$ из определения 3.2 можно эквивалентно записать в виде:

$$D_{1n}(s) = \sum_{k=1}^n E(|\xi_{kn}| I\{|\xi_{kn}| > 1\}) + \sum_{k=1}^n E(|\xi_{kn}|^s I\{|\xi_{kn}| \leq 1\}).$$

Лемма 3.3. Пусть существует константа $c > 0$, что для любого n имеет место неравенство

$$\sum_{k=1}^n E|\xi_{kn}| < c, \quad (3.9)$$

тогда условия (M_1) и (D_1) эквивалентны.

Доказательство. Доказательство следует из двух неравенств. Для любого $\tau \in (0, 1)$ справедливы неравенства:

$$D_{1n}(s) \leq M_{1n}(\tau) + \tau^{s-1}c,$$

$$M_{1n}(\tau) \leq \frac{1}{\tau^{s-1}} D_{1n}(s).$$

■

Теорема 3.1. (закон больших чисел для схемы серий) Рассмотрим схему серий. Пусть выполнено условие (D_1) , тогда имеет место сходимост по вероятности:

$$\zeta_n = \sum_{k=1}^n \xi_{kn} \xrightarrow[n \rightarrow \infty]{P} 0. \quad (3.10)$$

Доказательство. Сходимость по вероятности будет следовать из сходимости по распределению. Таким образом, достаточно показать, что выполнено $\zeta_n \xrightarrow[n \rightarrow \infty]{d} \zeta \stackrel{\text{п.н.}}{=} 0$. Применим метод характеристических функций:

$$\begin{aligned} |\varphi_{\zeta_n}(t) - \varphi_{\zeta}(t)| &= \left| E e^{it \sum_{k=1}^n \xi_{kn}} - 1 \right| = \\ &= \left| \prod_{k=1}^n E e^{it \xi_{kn}} - \prod_{k=1}^n 1 \right| \leq \sum_{k=1}^n |E e^{it \xi_{kn}} - 1| = \\ &= \sum_{k=1}^n |E(e^{it \xi_{kn}} - 1 - it \xi_{kn})| \leq \sum_{k=1}^n E |e^{it \xi_{kn}} - 1 - it \xi_{kn}|. \end{aligned}$$

Проанализируем отдельные слагаемые. Если $\beta = it\xi_{kn}$, то из леммы 3.1 следует справедливость неравенств:

$$|e^{it\xi_{kn}} - 1 - it\xi_{kn}| \leq \frac{|it\xi_{kn}|^2}{2},$$

$$|e^{it\xi_{kn}} - 1 - it\xi_{kn}| \leq |e^{it\xi_{kn}} - 1| + |it\xi_{kn}| \leq 2|it\xi_{kn}|.$$

Тогда имеют место неравенства:

$$\begin{aligned} \sum_{k=1}^n E |e^{it\xi_{kn}} - 1 - it\xi_{kn}| &\leq \\ &\leq \sum_{k=1}^n E \left\{ \min \left(2|t\xi_{kn}|, \frac{|t\xi_{kn}|^2}{2} \right) \right\} \leq \\ &\leq \max \left\{ 2|t|, \frac{|t|^2}{2} \right\} D_{1n}(2) \leq h(t)D_{1n}(s), \end{aligned}$$

где $h(t) = \max \left\{ 2|t|, \frac{|t|^2}{2} \right\}$. Так как выполнено условие (D_1) , то $h(t)D_{1n}(s) \xrightarrow{n \rightarrow \infty} 0$. ■

Замечание 3.3. Вернемся к последовательности случайных величин $\xi_{kn} = (\xi_k - a_k)/n$, тогда при выполнении условий теоремы 3.1 имеет место сходимость:

$$\varsigma_n = \frac{1}{n} \sum_{k=1}^n (\xi_k - a_k) = \frac{\sum_{k=1}^n \xi_k}{n} - \frac{\sum_{k=1}^n E\xi_k}{n} \xrightarrow[n \rightarrow \infty]{P} 0, \quad (3.11)$$

где $a_k = E\xi_k$.

Для $\xi_{kn} = (\xi_k - a_k)/n$ справедливы следующие представления:

$$D_{1n}(s) = \frac{1}{n} \sum_{k=1}^n E \left\{ \min \left\{ |\xi_k - a_k|, \frac{|\xi_k - a_k|^s}{n^{s-1}} \right\} \right\}, \quad (3.12)$$

$$M_{1n}(\tau) = \frac{1}{n} \sum_{k=1}^n E \{ |\xi_k - a_k| I \{ |\xi_k - a_k| \geq \tau n \} \}. \quad (3.13)$$

Теорема 3.2. Пусть $\{\xi_n\}$ — последовательность взаимно независимых случайных величин, причем $E\xi_k = a_k \in \mathbb{R}$ для любого k . Если $D_{1n}(s) \xrightarrow[n \rightarrow \infty]{} 0$ при некотором s , $1 < s \leq 2$, тогда

$$\varsigma_n = \frac{1}{n} \sum_{k=1}^n (\xi_k - a_k) \xrightarrow[n \rightarrow \infty]{P} 0, \quad (3.14)$$

т. е. к последовательности случайных величин $\{\xi_n\}$ применим закон больших чисел.

Доказательство. Доказательство следует из теоремы 3.1.

■

Замечание 3.4. ЗБЧ для одинаково распределенных независимых случайных величин является прямым следствием теоремы 3.2.

Для доказательства рассмотрим сумму:

$$\sum_{k=1}^n E \left(\frac{|\xi_k - a_k|}{n} \right) = \frac{1}{n} \sum_{k=1}^n E|\xi_k - a_k| = E|\xi_1 - a| = c,$$

считая, что все $a_k = a$. Таким образом, можно применить лемму 3.3, следовательно, условие (D_1) эквивалентно условию (M_1) . Как легко видеть,

$$\begin{aligned} M_{1n}(\tau) &= \sum_{k=1}^n \frac{E\{|\xi_k - a| I\{|\xi_k - a| > \tau n\}\}}{n} = \\ &= \frac{1}{n} \sum_{k=1}^n E\{|\xi_1 - a| I\{|\xi_1 - a| > \tau n\}\} = \\ &= E\{|\xi_1 - a| I\{|\xi_1 - a| > \tau n\}\}. \end{aligned}$$

Так как $|\xi_1(\omega) - a| I\{|\xi_1(\omega) - a| > \tau n\} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0$ и $|\xi_1(\omega) - a| I\{|\xi_1(\omega) - a| > \tau n\} \leq |\xi_1(\omega)| + |a|$, то по теореме Лебега имеет место сходимость:

$$E\{|\xi_1 - a| I\{|\xi_1 - a| > \tau n\}\} \xrightarrow[n \rightarrow \infty]{} 0,$$

следовательно, выполнены условия (M_1) и (D_1) .

■

§ 4. ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ДЛЯ НЕЗАВИСИМЫХ ПРОИЗВОЛЬНО РАСПРЕДЕЛЕННЫХ СЛУЧАЙНЫХ ВЕЛИЧИН. СХЕМА СЕРИЙ

Рассмотрим последовательность $\{\xi_n\}$ взаимно независимых и произвольно распределенных случайных величин. Пусть существуют $E\xi_k = a_k \in \mathbb{R}$ и $D\xi_k = \sigma_k^2 \in \mathbb{R}$. Введем обозначение:

$$B_n^2 = \sum_{k=1}^n \sigma_k^2 = D \left(\sum_{k=1}^n \xi_k \right).$$

Определим условия, при выполнении которых имеют место сходимости:

$$\varsigma_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n} \xrightarrow[n \rightarrow \infty]{d} \varsigma \sim N(0, 1)$$

или

$$P\{\varsigma_n \leq x\} \xrightarrow[n \rightarrow \infty]{} \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Введем обозначение: $\xi_{kn} = \frac{\xi_k - a_k}{B_n}$. Очевидно, что $E\xi_{kn} = 0$, $\sigma_{kn}^2 = D\xi_{kn}$, $\sum_{k=1}^n \sigma_{kn}^2 = 1$, тогда $\varsigma_n = \sum_{k=1}^n \xi_{kn}$. Далее будем рассматривать более общую задачу, введя произвольную схему серий, которая будет удовлетворять сформулированным ниже условиям.

Пусть для любого n задана конечная последовательность случайных величин $\{\xi_{kn}\}$, $k = 1, \dots, n$. Рассмотрим схему серий $\varsigma_n = \xi_{1n} + \dots + \xi_{nn}$, $n = 1, \dots$. Предположим, что внутри последовательности $\xi_{1n}, \xi_{2n}, \dots, \xi_{nn}$ случайные величины взаимно независимы, причем для любого k и для любого n : $E\xi_{kn} = 0$, $\sigma_{kn}^2 = D\xi_{kn}$, $\sum_{k=1}^n \sigma_{kn}^2 = 1$. Очевидно, что ранее введенные случайные величины $\xi_{kn} = (\xi_k - a_k)/B_n$ являются частным случаем рассматриваемой схемы серий.

Определение 4.1. Будем говорить, что выполнено условие (M_2) (условие Линдеберга), если для любого $\tau > 0$ имеет

МЕСТО СХОДИМОСТЬ:

$$M_{2n}(\tau) = \sum_{k=1}^n E \{ \xi_{kn}^2 I \{ |\xi_{kn}| > \tau \} \} \xrightarrow{n \rightarrow \infty} 0. \quad (4.1)$$

Замечание 4.1. Функция $M_{2n}(\tau)$ — невозрастающая, т. е. если $\tau_1 < \tau_2$, то $M_{2n}(\tau_1) \geq M_{2n}(\tau_2)$.

Доказательство. Преобразуем $M_{2n}(\tau_1)$:

$$\begin{aligned} M_{2n}(\tau_1) &= \sum_{k=1}^n E \{ \xi_{kn}^2 I \{ \tau_1 < |\xi_{kn}| \leq \tau_2 \} \} + \\ &+ \sum_{k=1}^n E \{ \xi_{kn}^2 I \{ |\xi_{kn}| > \tau_2 \} \} = \\ &= \sum_{k=1}^n E \{ \xi_{kn}^2 I \{ \tau_1 < |\xi_{kn}| \leq \tau_2 \} \} + M_{2n}(\tau_2). \end{aligned}$$

Так как $\sum_{k=1}^n E \{ \xi_{kn}^2 I \{ \tau_1 < |\xi_{kn}| \leq \tau_2 \} \} \geq 0$, то $M_{2n}(\tau_1) \geq M_{2n}(\tau_2)$. ■

Определение 4.2. Будем говорить, что выполнено условие (D_2) , если существует $s > 2$ такое, что имеет место сходимость:

$$D_{2n}(s) = \sum_{k=1}^n E \{ \min(\xi_{kn}^2, |\xi_{kn}|^s) \} \xrightarrow{n \rightarrow \infty} 0. \quad (4.2)$$

Замечание 4.2. Левую часть в условии (4.2) можно записать в виде:

$$D_{2n}(s) = \sum_{k=1}^n E \{ |\xi_{kn}|^s I \{ |\xi_{kn}| \leq 1 \} \} + \sum_{k=1}^n E \{ \xi_{kn}^2 I \{ |\xi_{kn}| > 1 \} \}.$$

Замечание 4.3. Функция $D_{2n}(s)$ — невозрастающая, т. е. если $s_1 < s_2$, то $D_{2n}(s_1) \geq D_{2n}(s_2)$.

Определение 4.3. Будем говорить, что выполнено условие (L) (условие Ляпунова), если существует $s > 2$ такое, что имеет место сходимость:

$$L_n(s) = \sum_{k=1}^n E|\xi_{kn}|^s \xrightarrow{n \rightarrow \infty} 0. \quad (4.3)$$

Определение 4.4. Будем говорить, что выполнено условие (S) (условие равномерной малости дисперсий), если имеет место сходимость:

$$\max_{k=1, n} \sigma_{kn}^2 \xrightarrow{n \rightarrow \infty} 0. \quad (4.4)$$

Замечание 4.4. Пусть выполнено условие (S) . По неравенству Чебышева: для любого $\varepsilon > 0$ и любого $k > 0$: $P\{|\xi_{kn}| \geq \varepsilon\} \leq \sigma_{kn}^2/\varepsilon^2$. Неравенство останется верным и в том случае, если взять максимум по k :

$$\max_{k=1, n} P\{|\xi_{kn}| \geq \varepsilon\} \leq \frac{\max_{k=1, n} \sigma_{kn}^2}{\varepsilon^2}.$$

По условию (S) имеет место сходимость:

$$\frac{\max_{k=1, n} \sigma_{kn}^2}{\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0,$$

тогда и $\max_{k=1, n} P\{|\xi_{kn}| \geq \varepsilon\} \xrightarrow{n \rightarrow \infty} 0$. Таким образом, все случайные величины, входящие в серию, равномерно малы по вероятности.

Лемма 4.1. Условия (M_2) и (D_2) эквивалентны.

Доказательство. Достаточность. Покажем, что из (M_2) следует (D_2) . Пусть $\tau \in (0, 1)$, тогда справедливо неравенство:

$$\begin{aligned}
D_{2n}(s) &= \sum_{k=1}^n E\{|\xi_{kn}|^s I\{|\xi_{kn}| \leq 1\}\} + \\
&\quad + \sum_{k=1}^n E\{\xi_{kn}^2 I\{|\xi_{kn}| > 1\}\} \leq \\
&\leq \tau^{s-2} \sum_{k=1}^n E\{\xi_{kn}^2 I\{|\xi_{kn}| \leq \tau\}\} + M_{2n}(\tau) \leq \tau^{s-2} + M_{2n}(\tau).
\end{aligned}$$

Величину τ можно выбрать сколь угодно малой и устремить n к бесконечности.

Необходимость. Покажем, что из (D_2) следует (M_2) . Для любого $\tau \in (0, 1)$ справедливы неравенства:

$$\begin{aligned}
M_{2n}(\tau) &= \sum_{k=1}^n E\{\xi_{kn}^2 I\{|\xi_{kn}| > 1\}\} + \\
&+ \sum_{k=1}^n E\{\xi_{kn}^2 I\{\tau < |\xi_{kn}| \leq 1\}\} \leq \sum_{k=1}^n E\{\xi_{kn}^2 I\{|\xi_{kn}| > 1\}\} + \\
&+ \frac{1}{\tau^{s-2}} \sum_{k=1}^n E\{\xi_{kn}^s I\{\tau < |\xi_{kn}| \leq 1\}\} \leq \frac{1}{\tau^{s-2}} D_{2n}(s).
\end{aligned}$$

■

Замечание 4.5. Из доказательства леммы следует, что условие (D_2) инвариантно относительно s , т. е. если условие (D_2) выполнено для произвольного $s_0 > 2$, то оно будет выполнено для любого $s > 2$.

Лемма 4.2. Справедливы следующие утверждения:

1. Из условия (L) следуют условия (M_2) и (D_2) .
2. Из условия (M_2) следует условие (S) .

Доказательство.

1) Очевидно, что $L_n(s) = \sum_{k=1}^n E|\xi_{kn}|^s \geq \sum_{k=1}^n E \min\{\xi_{kn}^2, |\xi_{kn}|^s\}$ для любого $s \geq 2$. Таким образом, из (L) следует (D_2) . Выполнение условия (M_2) следует из леммы 4.1.

2) Как легко видеть,

$$\begin{aligned}\sigma_{kn}^2 = D\xi_{kn} = E\xi_{kn}^2 &= E\{\xi_{kn}^2 I\{|\xi_{kn}| \leq \tau\}\} + \\ &+ E\{\xi_{kn}^2 I\{|\xi_{kn}| > \tau\}\} \leq \tau^2 + M_{2n}(\tau)\end{aligned}$$

для любого $\tau \in (0, 1)$. Следовательно,

$$0 \leq \max_{k=1, n} \sigma_{kn}^2 \leq \tau^2 + M_{2n}(\tau).$$

Для произвольного $\varepsilon > 0$ можно выбрать τ так, что $\tau^2 < \varepsilon/2$, по выбранному τ можно взять n_0 таким образом, что $M_{2n}(\tau) < \varepsilon/2$ для любого $n > n_0$. Следовательно,

$$\max_{k=1, n} \sigma_{kn}^2 \xrightarrow{n \rightarrow \infty} 0.$$

Таким образом, условие (S) выполнено. ■

Теорема 4.1 (центральная предельная теорема для схемы серий). Пусть задана схема серий $\varsigma_n = \sum_{k=1}^n \xi_{kn}$, $n = 1, 2, \dots$. Все случайные величины, входящие в любую серию $\xi_{1n}, \xi_{2n}, \dots, \xi_{nn}$, взаимно независимы. Для любых k и n выполнены равенства: $E\xi_{kn} = 0$, $D\xi_{kn} = \sigma_{kn}^2 < \infty$, причем $D\varsigma_n = \sum_{k=1}^n \sigma_{kn}^2 = 1$. Пусть выполнено условие (M_2) (или эквивалентное ему (D_2)), тогда имеет место сходимость по распределению:

$$\varsigma_n \xrightarrow[n \rightarrow \infty]{d} \varsigma \sim N(0, 1), \quad (4.5)$$

или, что эквивалентно, имеет место равномерная по $x \in \mathbb{R}$ сходимость:

$$P\{\varsigma_n \leq x\} \xrightarrow[n \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy. \quad (4.6)$$

Доказательство. Воспользуемся методом характеристических функций. Докажем, что для любого t выполняется:

$$\left| \varphi_{\varsigma_n}(t) - e^{-\frac{t^2}{2}} \right| \xrightarrow[n \rightarrow \infty]{} 0.$$

Заметив, что $1 = \sigma_{1n}^2 + \dots + \sigma_{nn}^2 = \sum_{k=1}^n \sigma_{kn}^2$ и применив лемму 3.2, получаем следующее неравенство:

$$\begin{aligned} & \left| \varphi_{\varsigma_n}(t) - e^{-\frac{t^2}{2}} \right| = \\ & = \left| \prod_{k=1}^n \varphi_{\xi_{kn}} - \prod_{k=1}^n e^{-\frac{t^2}{2} \sigma_{kn}^2} \right| \leq \sum_{k=1}^n \left| E e^{it\xi_{kn}} - e^{-\frac{t^2}{2} \sigma_{kn}^2} \right| = \\ & = \sum_{k=1}^n \left| E e^{it\xi_{kn}} - e^{-\frac{t^2}{2} \sigma_{kn}^2} + \left(1 - \frac{1}{2} \sigma_{kn}^2 t^2\right) - \left(1 - \frac{1}{2} \sigma_{kn}^2 t^2\right) \right| \leq I_1 + I_2, \end{aligned}$$

где

$$\begin{aligned} I_1 &= \sum_{k=1}^n \left| E \left(e^{-it\xi_{kn}} - 1 - it\xi_{kn} - \frac{(i\xi_{kn})^2}{2} t^2 \right) \right|, \\ I_2 &= \sum_{k=1}^n \left| e^{-\frac{\sigma_{kn}^2}{2} t^2} - 1 - \left(-\frac{\sigma_{kn}^2}{2} t^2 \right) \right|. \end{aligned}$$

Положив $\beta = -\frac{\sigma_{kn}^2}{2} t^2$ и применив лемму 3.1, получим:

$$I_2 \leq \sum_{k=1}^n \frac{\sigma_{kn}^4}{8} t^4 = \frac{t^4}{8} \sum_{k=1}^n \sigma_{kn}^2 \sigma_{kn}^2 \leq \frac{t^4}{8} \max_{k=1, n} \sigma_{kn}^2 \xrightarrow{n \rightarrow \infty} 0.$$

Оценим сверху I_1 , получаем:

$$I_1 \leq \sum_{k=1}^n E \left| e^{it\xi_{kn}} - 1 - it\xi_{kn} - \frac{(it\xi_{kn})^2}{2} \right|.$$

Положив $\beta = it\xi_{kn}$, из леммы 3.1 следует неравенство:

$$\left| e^{it\xi_{kn}} - 1 - it\xi_{kn} - \frac{(it\xi_{kn})^2}{2} \right| \leq \frac{|it\xi_{kn}|^3}{6}.$$

Кроме того,

$$\left| e^{it\xi_{kn}} - 1 - it\xi_{kn} - \frac{(it\xi_{kn})^2}{2} \right| \leq t^2 \xi_{kn}^2.$$

В итоге получаем:

$$I_1 \leq \sum_{k=1}^n \min \left\{ \frac{|t|^3 |\xi_{kn}|^3}{6}, t^2 \xi_{kn}^2 \right\} \leq \max \left\{ \frac{|t|^3}{6}, t^2 \right\} D_{2n}(3) \xrightarrow{n \rightarrow \infty} 0.$$

■

Замечание 4.6. В теореме 4.1 условие (M_2) (или эквивалентное ему (D_2)) является не только достаточным, но и необходимым, если выполнено условие равномерной малости дисперсий (S) .

§ 5. ТЕОРЕМА ЛИНДЕБЕРГА. ТЕОРЕМА ЛЯПУНОВА

Теорема 5.1 (теорема Линдеберга). Пусть $\{\xi_k\}$ — последовательность взаимно независимых случайных величин. Пусть для любого k существуют и конечны $E\xi_k = a_k \in \mathbb{R}$, $D\xi_k = \sigma_k^2 < \infty$. Обозначим $B_n^2 = \sum_{k=1}^n \sigma_k^2$. Пусть для любого τ имеет место сходимость:

$$M_{2n}(\tau) = \frac{1}{B_n^2} \sum_{k=1}^n E \{ (\xi_k - a_k)^2 I \{ |\xi_k - a_k| > \tau B_n \} \} \xrightarrow{n \rightarrow \infty} 0,$$

тогда

$$\varsigma_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n} \xrightarrow[n \rightarrow \infty]{d} \varsigma \sim N(0, 1), \quad (5.1)$$

или, что эквивалентно, имеет место равномерная по $x \in \mathbb{R}$ сходимость:

$$P\{\varsigma_n \leq x\} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy. \quad (5.2)$$

Замечание 5.1. Теорема 5.1 является частным случаем теоремы 4.1. Условие теоремы будет необходимым и достаточным, если выполнено условие равномерной малости дисперсий (S) .

Теорема 5.2 (теорема Ляпунова). Пусть $\{\xi_k\}$ — последовательность взаимно независимых случайных величин. Пусть для любого k существуют и конечны $E\xi_k = a_k \in \mathbb{R}$, $D\xi_k = \sigma_k^2 < \infty$. Обозначим $B_n^2 = \sum_{k=1}^n \sigma_k^2$. Пусть существует $s > 2$, что имеет место сходимость:

$$L_n(s) = \frac{\sum_{k=1}^n E|\xi_k - a_k|^s}{B_n^s} \xrightarrow{n \rightarrow \infty} 0,$$

тогда

$$\varsigma_n = \frac{\sum_{k=1}^n (\xi_k - a_k)}{B_n} \xrightarrow{n \rightarrow \infty} \varsigma \sim N(0, 1), \quad (5.3)$$

или, что эквивалентно, имеет место равномерная по $x \in \mathbb{R}$ сходимость:

$$P\{\varsigma_n \leq x\} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy. \quad (5.4)$$

Доказательство. Следует из леммы 4.2 и теоремы 4.1. ■

§ 6. УСИЛЕННЫЙ ЗАКОН БОЛЬШИХ ЧИСЕЛ ДЛЯ ПРОИЗВОЛЬНО РАСПРЕДЕЛЕННЫХ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Лемма 6.1 (неравенство Колмогорова). Пусть $\xi_1, \xi_2, \dots, \xi_n$ — совокупность взаимно независимых случайных величин, $E\xi_k = 0$ и $D\xi_k \in \mathbb{R}$ для любого k . Пусть $s_m = \sum_{k=1}^m \xi_k$, $m = 1, \dots, n$. Тогда для любого $\varepsilon > 0$ выполняется:

$$P\left\{\max_{k=1, n} |s_k| \geq \varepsilon\right\} \leq \frac{Ds_n}{\varepsilon^2}. \quad (6.1)$$

Доказательство. Рассмотрим событие:

$$A = \left\{\max_{k=1, n} |s_k| \geq \varepsilon\right\},$$

$$A = \bigcup_{k=1}^n A_k,$$

где $A_k = \{|s_1| < \varepsilon, \dots, |s_{k-1}| < \varepsilon, |s_k| \geq \varepsilon\}$. Очевидно, что $A_k \cap A_l = \emptyset$ для всех $k \neq l$. Тогда $I\{\omega \in A\} = \sum_{k=1}^n I\{\omega \in A_k\}$.

Справедливы неравенства:

$$\begin{aligned} Ds_n &= Es_n^2 \geq E\{s_n^2 I\{\omega \in A\}\} = \sum_{k=1}^n E\{s_n^2 I\{\omega \in A_k\}\} = \\ &= \sum_{k=1}^n E\{[s_k + (s_n - s_k)]^2 I\{\omega \in A_k\}\} = \\ &= \sum_{k=1}^n E\{s_k^2 I\{\omega \in A_k\}\} + \sum_{k=1}^n E\{(s_n - s_k)^2 I\{\omega \in A_k\}\} + \\ &+ 2 \sum_{k=1}^n E\{[s_k I\{\omega \in A_k\}](s_n - s_k)\} \geq \sum_{k=1}^n E\{s_k^2 I\{\omega \in A_k\}\} \geq \\ &\geq \varepsilon^2 \sum_{k=1}^n EI\{\omega \in A_k\} = \varepsilon^2 EI\{\omega \in A\} = \varepsilon^2 P(A). \end{aligned}$$

Поделив обе части неравенства на ε^2 , получим:

$$P\left\{\max_{k=1, n} |s_k| \geq \varepsilon\right\} \leq \frac{Ds_n}{\varepsilon^2}$$

для любого положительного ε . В доказательстве использовалось утверждение, что

$$E\{[s_k I\{\omega \in A_k\}](s_n - s_k)\} = E[s_k I\{\omega \in A_k\}]E(s_n - s_k) = 0.$$

■

Лемма 6.2. Пусть $\{\xi_n(\omega)\}$ — последовательность взаимно независимых случайных величин, $E\xi_n = 0$, $D\xi_n < \infty$ для любого n , $\sum_{k=1}^{\infty} D\xi_k < \infty$, тогда $\sum_{k=1}^{\infty} \xi_k(\omega) \stackrel{\text{п.н.}}{\in} \mathbb{R}$ или $P\{\omega : \sum_{k=1}^{\infty} \xi_k(\omega) \in \mathbb{R}\} = 1$.

Доказательство. Равенство $P\{\omega : \sum_{k=1}^{\infty} \xi_k(\omega) \in R\} = 1$ верно тогда и только тогда, когда почти наверное существует предел последовательности частичных сумм:

$$\lim_{n \rightarrow \infty} s_n(\omega) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \xi_k(\omega) \stackrel{\text{п.н.}}{\in} \mathbb{R}.$$

Для выполнения указанного свойства необходимо и достаточно, чтобы последовательность $\{s_n(\omega)\}$ была фундаментальна с вероятностью единица. Покажем, что для любого положительного ε имеет место сходимость:

$$P\left\{\sup_{k \geq 1} |s_{n+k} - s_n| > \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 0,$$

и, следовательно, последовательность частичных сумм почти наверное фундаментальна. Введем событие

$$\begin{aligned} A^{(n)} &= \left\{\sup_{k \geq 1} |s_{n+k} - s_n| > \varepsilon\right\} = \\ &= \bigcup_{k=1}^{\infty} \{|s_{n+k} - s_n| > \varepsilon\} = \bigcup_{N=1}^{\infty} \bigcup_{k=1}^N \{|s_{n+k} - s_n| > \varepsilon\} = \\ &= \bigcup_{N=1}^{\infty} \left\{\max_{k=1, N} |s_{n+k} - s_n| > \varepsilon\right\} = \bigcup_{N=1}^{\infty} A_N^{(n)}. \end{aligned}$$

Заметим, что $A_N^{(n)} \subset A_{N+1}^{(n)}$, следовательно, имеет место равенство:

$$P\left(A^{(n)}\right) = \lim_{N \rightarrow \infty} P\left(A_N^{(n)}\right).$$

Из леммы 6.1 следует неравенство:

$$P\left(A_N^{(n)}\right) \leq \frac{\sum_{k=1}^N D(\xi_{n+k})}{\varepsilon^2} = \frac{\sum_{l=n+1}^{n+N} D(\xi_l)}{\varepsilon^2}.$$

Устремим N к бесконечности, получаем:

$$P\left(A^{(n)}\right) \leq \frac{\sum_{l=n+1}^{\infty} D\xi_l}{\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0,$$

так как последовательность остаточных сумм стремится к нулю. Таким образом, $\sum_{k=1}^{\infty} \xi_k(\omega) \stackrel{\text{п.н.}}{\in} \mathbb{R}$. ■

Лемма 6.3 (лемма Теплица). Пусть $\{x_n\}$ — числовая последовательность, причем, $x_n \xrightarrow{n \rightarrow \infty} x \in \mathbb{R}$. Последовательность $\{a_n\}$ состоит из положительных элементов. Пусть последовательность $b_n = \sum_{k=1}^n a_k$ монотонно стремится к бесконечности при $n \rightarrow \infty$, тогда имеет место сходимость:

$$\frac{\sum_{k=1}^n a_k x_k}{b_n} \xrightarrow{n \rightarrow \infty} x. \quad (6.2)$$

Доказательство. Рассмотрим $\varepsilon > 0$ и сделаем следующую оценку:

$$\left| \frac{\sum_{k=1}^n a_k x_k}{b_n} - x \right| = \frac{\left| \sum_{k=1}^n a_k (x_k - x) \right|}{b_n} \leq \frac{\sum_{k=1}^n a_k |x_k - x|}{b_n}.$$

Из условий леммы следует, что для произвольного положительного ε найдется n_0 такое, что для любого $n > n_0$ будет выполнено неравенство: $|x_k - x| < \varepsilon/2$, $k > n_0$. Тогда для любого $n > n_0$ справедлива оценка:

$$\begin{aligned} \frac{\sum_{k=1}^n a_k |x_k - x|}{b_n} &= \frac{\sum_{k=1}^{n_0} a_k |x_k - x|}{b_n} + \frac{\sum_{k=n_0+1}^n a_k |x_k - x|}{b_n} \leq \\ &\leq \frac{\sum_{k=1}^{n_0} a_k |x_k - x|}{b_n} + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

так как $b_n \rightarrow \infty$, то существует номер n_1 , что при $n > n_1$ первое слагаемое в правой части меньше $\varepsilon/2$. ■

Следствие 6.1. Пусть выполнены условия леммы 6.3. Выберем $a_k = 1$ для любого k . Тогда имеет место сходимость:

$$\frac{\sum_{k=1}^n x_k}{n} \xrightarrow{n \rightarrow \infty} x.$$

Доказательство. Очевидным образом следует из леммы 6.3. ■

Лемма 6.4 (лемма Кронекера). Пусть задан ряд $\sum_{k=1}^{\infty} y_k = s \in \mathbb{R}$. Последовательность $b_n \xrightarrow{n \rightarrow \infty} \infty$ и для любого n : $b_n > 0$, $b_{n+1} > b_n$, тогда имеет место сходимость:

$$\frac{\sum_{k=1}^n b_k y_k}{b_n} \xrightarrow{n \rightarrow \infty} 0. \quad (6.3)$$

Доказательство. Пусть $s_n = \sum_{k=1}^n y_k$ — частичные суммы, $y_0 = 0$, $b_0 = 0$, $s_0 = 0$. Справедливы равенства:

$$\begin{aligned} \sum_{k=1}^n b_k y_k &= \sum_{k=1}^n b_k (s_k - s_{k-1}) = \sum_{k=1}^n b_k s_k - \sum_{k=1}^n b_k s_{k-1} = b_n s_n + \\ &+ \sum_{k=1}^{n-1} b_k s_k - \sum_{k=1}^n b_k s_{k-1} = b_n s_n + \sum_{k=1}^{n-1} b_{k-1} s_{k-1} - \sum_{k=1}^n b_k s_{k-1} = \\ &= b_n s_n - \sum_{k=1}^n (b_k - b_{k-1}) s_{k-1} = b_n s_n - \sum_{k=1}^n a_k x_k, \end{aligned}$$

где $x_k = s_{k-1}$, $a_k = b_k - b_{k-1}$. Поделив обе части равенства на b_n и устремив n к бесконечности, в силу леммы 6.3 получим:

$$\frac{\sum_{k=1}^n b_k y_k}{b_n} = s_n - \frac{\sum_{k=1}^n a_k x_k}{b_n} \xrightarrow{n \rightarrow \infty} 0.$$

■

Следствие 6.2. *Имеет место сходимость:*

$$\frac{\sum_{k=1}^n \frac{1}{k}}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Для доказательства достаточно положить $b_n = n$, $y_k = 1/k^2$. ■

Теорема 6.1 (усиленный закон больших чисел для независимых произвольно распределенных случайных величин). Пусть $\{\xi_n(\omega)\}$ — последовательность взаимно независимых случайных величин, заданных на (Ω, \mathcal{F}, P) . Пусть для любого k существуют и конечны $E\xi_k = a_k \in \mathbb{R}$, $D\xi_k \in \mathbb{R}$. Последовательность $b_n \xrightarrow{n \rightarrow \infty} \infty$ и для любого n : $b_n > 0$, $b_{n+1} > b_n$. Пусть также ряд $\sum_{n=1}^{\infty} D\xi_n/b_n^2$ сходится, тогда имеет место сходимость:

$$\frac{\sum_{k=1}^n (\xi_k - a_k)}{b_n} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0. \quad (6.4)$$

Доказательство. Как следует из условия теоремы:

$$\sum_{k=1}^{\infty} \frac{D\xi_k}{b_k^2} = \sum_{k=1}^{\infty} D\left(\frac{\xi_k - a_k}{b_k}\right) = \sum_{k=1}^{\infty} D\eta_k < \infty,$$

где $\eta_k = (\xi_k - a_k)/b_k$, причем η_k удовлетворяет всем условиям леммы 6.2, следовательно,

$$\sum_{k=1}^{\infty} \frac{\xi_k - a_k}{b_k} = \sum_{k=1}^{\infty} \eta_k \overset{\text{п.н.}}{\in} \mathbb{R}.$$

По лемме 6.4 имеет место сходимость:

$$\frac{\sum_{k=1}^n b_k \eta_k}{b_n} = \frac{\sum_{k=1}^n (\xi_k(\omega) - a_k)}{b_n} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0.$$

■

Следствие 6.3. Пусть $\{\xi_k\}$ — последовательность взаимно независимых случайных величин, заданных на (Ω, \mathcal{F}, P) . Пусть для любого k существуют и конечны $E\xi_k = a_k \in \mathbb{R}$, $D\xi_k < \infty$. Пусть также ряд $\sum_{k=1}^{\infty} D\xi_k/k^2$ сходится, тогда имеет место сходимост

$$\frac{\sum_{k=1}^n (\xi_k - a_k)}{n} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0.$$

Доказательство. Доказательство следует из теоремы 6.1, если положить $b_n = n$. ■

§ 7. УСИЛЕННЫЙ ЗАКОН БОЛЬШИХ ЧИСЕЛ ДЛЯ ОДИНАКОВО РАСПРЕДЕЛЕННЫХ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Лемма 7.1. Математическое ожидание произвольной случайной величины ξ конечно тогда и только тогда, когда сходится ряд $\sum_{k=0}^{\infty} P\{|\xi| > k\}$.

Доказательство. Из свойств интеграла Лебега следует, что $E\xi \in \mathbb{R}$ тогда и только тогда, когда $E|\xi| \in \mathbb{R}$. Очевидно, что

$$\begin{aligned} \sum_{k=1}^{\infty} (k-1)I\{k-1 < |\xi(\omega)| \leq k\} &\leq \\ &\leq |\xi(\omega)| \leq \sum_{k=1}^{\infty} kI\{k-1 < |\xi(\omega)| \leq k\}. \end{aligned}$$

Следовательно,

$$\begin{aligned} \sum_{k=1}^{\infty} (k-1)P\{k-1 < |\xi(\omega)| \leq k\} &\leq \\ &\leq E|\xi(\omega)| \leq \sum_{k=1}^{\infty} kP\{k-1 < |\xi(\omega)| \leq k\}. \end{aligned}$$

Таким образом, математическое ожидание конечно тогда и только тогда, когда ряд $\sum_{k=1}^{\infty} kP\{k-1 < |\xi(\omega)| \leq k\}$ сходится. Все слагаемые положительны, поэтому можно поменять местами порядок суммирования:

$$\begin{aligned} \sum_{k=1}^{\infty} kP\{k-1 < |\xi(\omega)| \leq k\} &= \\ &= \sum_{k=1}^{\infty} \sum_{n=1}^k P\{k-1 < |\xi(\omega)| \leq k\} = \\ &= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P\{k-1 < |\xi(\omega)| \leq k\} = \sum_{n=1}^{\infty} P\{|\xi(\omega)| > n-1\}. \end{aligned}$$

Ряд $\sum_{n=1}^{\infty} P\{|\xi(\omega)| > n\}$ сходится тогда и только тогда, когда $E\xi \in \mathbb{R}$. ■

Теорема 7.1 (усиленный закон больших чисел Колмогорова для независимых одинаково распределенных случайных величин). Пусть $\{\xi_n\}$ — последовательность взаимно независимых одинаково распределенных случайных величин, заданных на (Ω, \mathcal{F}, P) . Сходимость

$$\frac{1}{n} \sum_{k=1}^n \xi_k \xrightarrow[n \rightarrow \infty]{n.н.} a \in \mathbb{R} \quad (7.1)$$

имеет место тогда и только тогда, когда существует $E\xi_1$ и $E\xi_1 = a$.

Доказательство. Достаточность. Предположим, что $E\xi_1 = a \in \mathbb{R}$. Покажем, что имеет место сходимость (7.1). Определим случайную величину:

$$\tilde{\xi}_n(\omega) = \xi_n(\omega) I\{|\xi_n(\omega)| \leq n\}.$$

Тогда можно записать равенства:

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n \xi_k(\omega) - a &= \frac{\sum_{k=1}^n \xi_k(\omega) - na + \sum_{k=1}^n \tilde{\xi}_k(\omega) - \sum_{k=1}^n \tilde{\xi}_k(\omega)}{n} + \\ &+ \frac{\sum_{k=1}^n E\tilde{\xi}_k(\omega) - \sum_{k=1}^n E\tilde{\xi}_k(\omega)}{n} = \frac{\sum_{k=1}^n (\xi_k(\omega) - \tilde{\xi}_k(\omega))}{n} + \\ &+ \frac{\sum_{k=1}^n (\tilde{\xi}_k(\omega) - E\tilde{\xi}_k(\omega))}{n} + \frac{\sum_{k=1}^n (E\tilde{\xi}_k(\omega) - a)}{n}. \end{aligned}$$

Рассмотрим каждую дробь по отдельности. Преобразуем последнюю дробь:

$$\begin{aligned} \frac{\sum_{k=1}^n (E\tilde{\xi}_k(\omega) - a)}{n} &= \frac{\sum_{k=1}^n (E\tilde{\xi}_k(\omega) - E\xi_k(\omega))}{n} = \\ &= - \frac{\sum_{k=1}^n E\{\xi_k(\omega) I\{|\xi_k(\omega)| > k\}\}}{n} = \\ &= - \frac{\sum_{k=1}^n E\{\xi_1(\omega) I\{|\xi_1(\omega)| > k\}\}}{n}. \end{aligned}$$

Введем следующее обозначение: $E\{\xi_1(\omega) I\{|\xi_1(\omega)| > k\}\} = x_k$. Очевидно, что имеет место сходимость:

$$\xi_1(\omega) I\{|\xi_1(\omega)| > k\} \xrightarrow[k \rightarrow \infty]{\text{п.н.}} 0.$$

Справедливо неравенство: $|\xi_1(\omega) I\{|\xi_1(\omega)| > k\}| \leq |\xi_1(\omega)|$. Так как математическое ожидание $E\xi_1(\omega)$ конечно, то конечно и математическое ожидание $E|\xi_1(\omega)|$ (следует из свойств интеграла Лебега). Таким образом, по теореме Лебега о мажорируемой сходимости: $x_k \xrightarrow[k \rightarrow \infty]{} 0$. Из следствия к лемме Теплица получаем сходимость:

$$\frac{\sum_{k=1}^n (E\tilde{\xi}_k(\omega) - a)}{n} \xrightarrow[n \rightarrow \infty]{} 0.$$

Рассмотрим первую дробь. Введем событие:

$$A_k = \{\xi_k(\omega) = \tilde{\xi}_k(\omega)\} = \{|\xi_k(\omega)| \leq k\},$$

$$\bar{A}_k = \{|\xi_k(\omega)| > k\}.$$

Рассмотрим событие $A = \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} A_k$. Пусть $\omega_0 \in A$, тогда существует m_0 такое, что $\omega_0 \in \bigcap_{k=m_0}^{\infty} A_k$. Тогда при выбранном ω_0 выполняется сходимость:

$$\frac{\sum_{k=1}^{m_0-1} (\xi_k(\omega_0) - \tilde{\xi}_k(\omega_0))}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Применяя законы двойственности, получаем равенство:

$$\bar{A} = \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} \bar{A}_k.$$

Из первого утверждения леммы Бореля–Кантелли следует, что если сходится ряд $\sum_{k=1}^{\infty} P(\bar{A}_k)$, то $P(\bar{A}) = 0$. Так как $\sum_{k=1}^{\infty} P\{|\xi_k(\omega)| > k\} = \sum_{k=1}^{\infty} P\{|\xi_1(\omega)| > k\} < \infty$, то $P(\bar{A}) = 0$, следовательно, $P(A) = 1$. Таким образом, выполняется сходимость:

$$\frac{\sum_{k=1}^n (\xi_k(\omega) - \tilde{\xi}_k(\omega))}{n} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0.$$

Для доказательства сходимости второй дроби достаточно показать, что ряд $\sum_{n=1}^{\infty} D\tilde{\xi}_n(\omega)/n^2$ сходится. Так как $|\tilde{\xi}_n(\omega)|^2 \leq \sum_{k=1}^n k^2 I\{k-1 < |\xi_k(\omega)| \leq k\}$, то справедливо неравенство:

$$D\tilde{\xi}_n \leq E\tilde{\xi}_n^2 \leq \sum_{k=1}^n k^2 P\{k-1 < |\xi_k(\omega)| \leq k\}.$$

Рассмотрим ряд и произведем оценку сверху:

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{D\tilde{\xi}_n}{n^2} &\leq \sum_{n=1}^{\infty} \sum_{k=1}^n \frac{k^2 P\{k-1 < |\xi_k(\omega)| \leq k\}}{n^2} \leq \\ &\leq \sum_{k=1}^{\infty} k^2 P\{k-1 < |\xi_k(\omega)| \leq k\} \sum_{n=k}^{\infty} \frac{1}{n^2}. \end{aligned}$$

Так как

$$\begin{aligned} \sum_{n=k}^{\infty} \frac{1}{n^2} &= \frac{1}{k^2} + \sum_{n=k+1}^{\infty} \frac{1}{n^2} \leq \\ &\leq \int_k^{\infty} \frac{1}{x^2} dx + \frac{1}{k^2} = \frac{1}{k} + \frac{1}{k^2} = \frac{1+k}{k^2}, \end{aligned}$$

то справедливы неравенства:

$$\begin{aligned} \sum_{k=1}^{\infty} k^2 P\{k-1 < |\xi_k(\omega)| \leq k\} \sum_{n=k}^{\infty} \frac{1}{n^2} &\leq \\ &\leq \sum_{k=1}^{\infty} (1+k) P\{k-1 < |\xi_1(\omega)| \leq k\} \leq \\ &\leq \sum_{k=1}^{\infty} (k-1) P\{k-1 < |\xi_1(\omega)| \leq k\} + 2P\{|\xi_1(\omega)| > 0\} \leq \\ &\leq E|\xi_1(\omega)| + 2P\{|\xi_1(\omega)| > 0\}. \end{aligned}$$

Последнее неравенство следует из леммы 7.1. Таким образом, ряд сходится.

Необходимость. Введем обозначения: $S_n = \sum_{k=1}^n \xi_k$, $S_{n-1} = \sum_{k=1}^{n-1} \xi_k$. Пусть $S_n/n \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a$, тогда $\frac{S_{n-1}}{n-1} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a$.

Так как

$$\frac{S_n}{n} = \frac{\xi_n}{n} + \frac{n-1}{n} \frac{S_{n-1}}{n-1},$$

при этом $(n-1)/n \xrightarrow[n \rightarrow \infty]{} 1$, тогда $\xi_n/n \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0$.

Введем событие

$$B = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} B_n,$$

где $B_n = \{|\xi_n(\omega)| > n\}$, события B_n — независимые. Тогда из леммы Бореля–Кантелли следует, что если $\sum_{n=1}^{\infty} P(B_n) < \infty$, то $P(B) = 0$, и если $\sum_{n=1}^{\infty} P(B_n) = \infty$, то $P(B) = 1$.

Если выполнено событие B , то

$$B \subset \left\{ \omega : \frac{\xi_n}{n} \nrightarrow 0 \right\}.$$

Следовательно, $P(B) = 0$ и $\sum_{n=1}^{\infty} P(B_n) = \sum_{n=1}^{\infty} P\{|\xi_1(\omega)| > n\} < \infty$. Таким образом, по лемме 7.1 получаем, что $E\xi_1(\omega) \in \mathbb{R}$, но тогда из первой части теоремы следует, что $E\xi_1(\omega) = a$. ■

ВЫБОРОЧНОЕ ПРОСТРАНСТВО

§ 1. ВЫБОРКА. ВЫБОРОЧНОЕ ПРОСТРАНСТВО

Рассмотрим случайную величину

$$\xi(\omega) : \Omega \longrightarrow \mathbb{R}$$

и вероятностное пространство значений случайной величины

$$(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_\xi),$$

где $\mathcal{B}(\mathbb{R})$ — сигма-алгебра борелевских множеств числовой прямой; P_ξ — вероятностная мера такая, что $P_\xi(-\infty, x] = F_\xi(x) = P\{\xi \leq x\}$.

Если речь идет о наборе случайных величин $(\xi_1, \dots, \xi_n) : \Omega \longrightarrow \mathbb{R}^n$, то вероятностное пространство определим следующим образом:

$$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_\xi),$$

здесь P_ξ — совместное распределение случайных величин: $P_\xi(-\infty, x_1] \times \dots \times (-\infty, x_n] = F_\xi(x_1, \dots, x_n)$.

Часто в практических задачах функция распределения неизвестна. Аппроксимация неизвестных функций распределения — одна из задач математической статистики.

Определение 1.1. Совокупность взаимно независимых реализаций случайной величины ξ образует выборку $X_{[n]}$ объема n :

$$X_{[n]} = (X_1, \dots, X_n),$$

где X_i — числовая реализация случайной величины ξ в i -ом эксперименте ($i = 1, \dots, n$).

Определение 1.2. Случайная величина ξ , реализации которой мы наблюдаем, называется генеральной совокупностью.

Следует отметить, что требование взаимной независимости сужает допустимую область исследований. Однако требование взаимной независимости наблюдений необходимо для построения строгой математической теории.

Функция распределения выборки строится по функции распределения генеральной совокупности:

$$F_{X_{[n]}}(x_1, \dots, x_n) = F_{\xi}(x_1) \cdot \dots \cdot F_{\xi}(x_n),$$

где x_i — числовая переменная, соответствующая i -ой координатной оси.

Получили *выборочное пространство*

$$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_{X_{[n]}}),$$

соответствующее выборкам объема n , где вероятностная мера $P_{X_{[n]}}$ взаимно однозначно соответствует функции распределения $F_{X_{[n]}}$.

Учитывая необходимость предельного перехода, когда $n \rightarrow \infty$, рассмотрим бесконечномерное пространство:

$$(\mathbb{R}^{\infty}, \mathcal{B}(\mathbb{R}^{\infty}), P_{X_{[\infty]}}).$$

Элементарным событием в этом пространстве является бесконечная числовая последовательность (бесконечная выборка). Указанное выше конечномерное пространство размерности n для выборок объема n , является подпространством бесконечномерного пространства, соответствующим первым n координатам.

Пусть $B \in \mathcal{B}(\mathbb{R}^n)$, рассмотрим цилиндрическое множество $J_n(B) = \{x \in \mathbb{R}^{\infty} : x = (x_1, \dots, x_n, \dots), (x_1, \dots, x_n) \in B\}$, тогда $P_{X_{[\infty]}}(J_n(B)) = P_{X_{[n]}}(B)$. Мера $P_{X_{[\infty]}}$ определяется единственным образом по мерам $P_{X_{[n]}}$, $n = 1, 2, \dots$ (см. теорему Ионеску Тулчи в книге Ж. Неве [24]).

Определение 1.3. Статистикой будем называть любую борелевскую функцию, заданную на выборочном пространстве.

Если $\xi = (\xi_1, \dots, \xi_m)^T$ — случайный вектор, то при проведении экспериментов фиксируются значения всей совокупности, получаются взаимно независимые вектора

$$X_1 = \begin{pmatrix} X_{11} \\ X_{21} \\ \vdots \\ X_{m1} \end{pmatrix}, \dots, X_n = \begin{pmatrix} X_{1n} \\ X_{2n} \\ \vdots \\ X_{mn} \end{pmatrix}.$$

Аналогично скалярному случаю можно построить выборочное вероятностное пространство для выборок такого типа.

Введем обозначения $\mathcal{X} = \mathbb{R}^m$, получаем

$$(\mathcal{X}^n, \mathcal{B}(\mathcal{X}) \otimes \dots \otimes \mathcal{B}(\mathcal{X}), P_{X_{[n]}}),$$

где $P_{X_{[n]}}(B_1 \times \dots \times B_n) = P_\xi(B_1)P_\xi(B_2) \dots P_\xi(B_n)$, а $B_1 \in \mathcal{B}(\mathcal{X}), \dots, B_n \in \mathcal{B}(\mathcal{X})$. Можно считать, что все построенные конечномерные пространства являются проекциями бесконечномерного пространства

$$(\mathcal{X}^\infty, \mathcal{B}(\mathcal{X}^\infty) \otimes \dots \otimes \mathcal{B}(\mathcal{X}^\infty), P_{X_{[\infty]}}).$$

Элементарными событиями будут бесконечные выборки. Мера $P_{X_{[\infty]}}$ определяется единственным образом по мерам $P_{X_{[n]}}$, $n = 1, 2, \dots$

§ 2. ЭМПИРИЧЕСКАЯ ВЕРОЯТНОСТНАЯ МЕРА. ГИСТОГРАММА

Пусть имеется генеральная совокупность ξ , представляющая собой случайную величину, и выборка $X_{[n]} = (X_1, \dots, X_n)$.

Определение 2.1. Эмпирическим распределением назовем вероятностную меру, определенную следующим образом:

$$P_n^*(B) = \frac{\nu(B)}{n},$$

где $B \in \mathcal{B}(\mathbb{R})$, а $\nu(B)$ — количество элементов выборки, попавших в B .

Если n фиксировано, и выборка $X_{[n]}$ фиксирована, то $P_n^*(\cdot)$ является вероятностной мерой на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, следовательно, ей соответствует единственная функция распределения.

Определение 2.2. Эмпирической функцией распределения называется функция

$$F_n^*(x) = P_n^*(-\infty; x] = \frac{\nu(-\infty; x]}{n}, \quad x \in \mathbb{R}.$$

Введем порядковые статистики:

$$\begin{aligned} X_{(1)} &= \min \{X_1, \dots, X_n\} \text{ —} \\ &\text{первая порядковая статистика,} \\ X_{(2)} &= \min \{ \{X_1, \dots, X_n\} \setminus X_{(1)} \} \text{ —} \\ &\text{вторая порядковая статистика,} \\ &\dots \quad \dots \quad \dots \\ X_{(n)} &= \max \{X_1, \dots, X_n\} \text{ —} \\ &n\text{-ая порядковая статистика.} \end{aligned}$$

Очевидно, что $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$. Величины $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ образуют *вариационный ряд*. Если предположить, что все элементы вариационного ряда различны, то есть $X_{(1)} < X_{(2)} < \dots < X_{(n)}$, то можно определить эмпирическую функцию распределения следующим образом:

$$F_n^*(x) = \begin{cases} 0, & \text{если } x < X_{(1)}; \\ \frac{1}{n}, & \text{если } X_{(1)} \leq x < X_{(2)}; \\ \frac{2}{n}, & \text{если } X_{(2)} \leq x < X_{(3)}; \\ \dots & \\ \frac{k}{n}, & \text{если } X_{(k)} \leq x < X_{(k+1)}; \\ \dots & \\ 1, & \text{если } x \geq X_{(n)}. \end{cases} \quad (2.1)$$

Имея вариационный ряд, можно построить *гистограмму*. Возьмем интервал (a, b) , где $a < X_{(1)}$ и $X_{(n)} < b$, разобьем

этот интервал на конечную совокупность непересекающихся промежутков:

$$a_0 = a < a_1 < a_2 < \dots < a_m = b,$$

$$(a_{i-1}, a_i], i = 1, \dots, m.$$

Пусть n_i — количество элементов выборки, попавших в полуинтервал $(a_{i-1}, a_i]$. Тогда

$$n_1 + n_2 + \dots + n_m = n,$$

$$l_i = a_i - a_{i-1},$$

$$h_i = \frac{n_i}{l_i n}.$$

Получаем гистограмму:

$$f_n^*(x) = \begin{cases} 0, & \text{если } x \leq a_0; \\ h_1, & \text{если } a_0 < x \leq a_1; \\ \dots & \\ h_m, & \text{если } a_{m-1} < x \leq a_m; \\ 0, & \text{если } x > a_m. \end{cases}$$

Гистограмма $f_n^*(x)$ — эмпирический аналог плотности распределения. Если в знаменателе при вычислении h_i убрать l_i , получится гистограмма относительных частот, если, кроме того, в знаменателе убрать n , то получится гистограмма частот n_i . Часто при построении гистограммы полагают $l_i = l = \text{const}$.

§ 3. ТЕОРЕМА ГЛИВЕНКО–КАНТЕЛЛИ.

ТЕОРЕМА О ПРЕДЕЛЬНОМ РАСПРЕДЕЛЕНИИ ЭМПИРИЧЕСКИХ ВЕРОЯТНОСТЕЙ

Справедливы следующие теоремы.

Теорема 3.1. Для любого $B \in \mathcal{B}(\mathbb{R})$ выполняется:

$$P_n^*(B) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} P_\xi(B) \quad (3.1)$$

и для любого $x \in \mathbb{R}$ выполняется:

$$F_n^*(x) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} F_\xi(x). \quad (3.2)$$

Доказательство. Очевидно, что справедливо равенство:

$$P_n^*(B) = \frac{1}{n} \sum_{k=1}^n I\{X_k \in B\},$$

все слагаемые в этой сумме являются случайными величинами, они независимы и одинаково распределены. Каждая случайная величина принимает значение 1 с вероятностью $P_\xi(B)$ и 0 с вероятностью $1 - P_\xi(B)$. Очевидны равенства:

$$P\{X_k \in B\} = P\{\xi \in B\} = P_\xi(B),$$

тогда

$$EI\{X_k \in B\} = P_\xi(B),$$

откуда, учитывая усиленный закон больших чисел Колмогорова, следует (3.1):

$$P_n^*(B) = \frac{1}{n} \sum_{k=1}^n I\{X_k \in B\} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} P_\xi(B).$$

Для доказательства второго утверждения возьмем $B = (-\infty, x]$. ■

Теорема 3.2 (теорема Гливенко–Кантелли). Пусть заданы функция распределения $F_\xi(x)$ и эмпирическая функция распределения $F_n^*(x)$, тогда

$$\sup_{x \in \mathbb{R}} |F_n^*(x) - F_\xi(x)| \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0.$$

Доказательство. 1. Рассмотрим случай, когда $F_\xi(x)$ непрерывна на \mathbb{R} . Выберем любое $\varepsilon > 0$, разобьем ось ординат шагом меньше ε , на оси абсцисс получим точки дробления:

$$-\infty = z_0 < z_1 < \dots < z_r = \infty.$$

При этом:

$$0 \leq F_{\xi}(z_{k+1}) - F_{\xi}(z_k) < \varepsilon.$$

Введем событие $A_k = \{F_n^*(z_k) \xrightarrow{n \rightarrow \infty} F_{\xi}(z_k)\}$, тогда по теореме 3.1: $P(A_k) = 1$ для всех $k = 1, \dots, r-1$. Рассмотрим событие $A = \bigcap_{k=1}^{r-1} A_k$, тогда справедливо равенство $\bar{A} = \bigcup_{k=1}^{r-1} \bar{A}_k$. Очевидно, что $P(\bar{A}) \leq \sum_{k=1}^{r-1} P(\bar{A}_k) = 0$, следовательно, $P(\bar{A}) = 0$ или $P(A) = 1$.

Будем рассматривать только события $\omega \in A$. В каждой из точек z_1, \dots, z_{r-1} выполняется сходимост, тогда существует номер n_0 такой, что для всех номеров $n \geq n_0$ и для любого $k = 1, \dots, r-1$ выполнено:

$$|F_n^*(z_k) - F_{\xi}(z_k)| < \varepsilon.$$

Возьмем любую точку $x \in \mathbb{R}$. Она обязательно попадет в какой-нибудь промежуток $z_k \leq x < z_{k+1}$. Оценим сверху и снизу разность $F_n^*(x) - F_{\xi}(x)$:

$$\begin{aligned} F_n^*(x) - F_{\xi}(x) &\leq F_n^*(z_{k+1}) - F_{\xi}(z_k) \leq \\ &\leq F_n^*(z_{k+1}) - F_{\xi}(z_{k+1}) + \varepsilon \leq 2\varepsilon, \end{aligned}$$

$$\begin{aligned} F_n^*(x) - F_{\xi}(x) &\geq F_n^*(z_k) - F_{\xi}(z_{k+1}) \geq \\ &\geq F_n^*(z_k) - F_{\xi}(z_k) - \varepsilon \geq -2\varepsilon. \end{aligned}$$

Следовательно, для любого $x \in \mathbb{R}$ справедливо неравенство:

$$|F_n^*(x) - F_{\xi}(x)| \leq 2\varepsilon.$$

Таким образом, мы доказали утверждение о равномерной сходимости с вероятностью единица, так как $P(A) = 1$.

2. Рассмотрим случай, когда $F_{\xi}(x)$ может иметь разрывы, т. е. существует точка x , для которой: $F_{\xi}(x-0) < F_{\xi}(x)$. Выберем произвольное $\varepsilon > 0$ и такие точки разрыва, для которых имеет место неравенство:

$$F_{\xi}(y_k) - F_{\xi}(y_k - 0) > \frac{\varepsilon}{2}.$$

Таких точек y_k конечное число, включаем их в точки дробления. Обозначим через z_i , $i = 1, \dots, r - 1$ все получившиеся в итоге точки дробления:

$$-\infty = z_0 < z_1 < \dots < z_r = \infty,$$

$$F_\xi(z_{k+1} - 0) - F_\xi(z_k) < \varepsilon.$$

Рассмотрим события $A_k = \{F_n^*(z_k) \xrightarrow{n \rightarrow \infty} F_\xi(z_k)\}$ и $A_k^- = \{F_n^*(z_k - 0) \xrightarrow{n \rightarrow \infty} F_\xi(z_k - 0)\}$. Очевидно, что $P(A_k) = 1$, $P(A_k^-) = 1$, так как $F_\xi(z_k - 0) = P_\xi(-\infty, z_k)$ и $B = (-\infty, z_k)$. Рассмотрим $A = \bigcap_{k=1}^{r-1} (A_k \cap A_k^-)$. Далее полностью повторяются рассуждения предыдущего пункта доказательства. ■

Теорема 3.3. Для любого борелевского множества $B \in \mathcal{B}(\mathbb{R})$ выполняется:

$$\sqrt{n} (P_n^*(B) - P_\xi(B)) \xrightarrow[n \rightarrow \infty]{d} \sqrt{P_\xi(B)(1 - P_\xi(B))} \zeta,$$

где $\zeta \sim N(0, 1)$.

Доказательство. Справедливо равенство:

$$P_n^*(B) - P_\xi(B) = \frac{1}{n} \sum_{k=1}^n (I\{X_k \in B\} - P_\xi(B)).$$

Применим центральную предельную теорему для одинаково распределенных слагаемых:

$$\sqrt{n} (P_n^*(B) - P_\xi(B)) = \frac{\sum_{k=1}^n (I\{X_k \in B\} - P_\xi(B))}{\sqrt{n}}.$$

Из чего следует доказательство теоремы. ■

Замечание 3.1. Теоремы 3.1, 3.2, 3.3 справедливы и в многомерном случае.

§ 4. ОПИСАТЕЛЬНАЯ СТАТИСТИКА

В описательную статистику входят оценки числовых характеристик генеральной совокупности ξ , найденные по имеющейся у статистика выборке $X_{[n]} = (X_1, \dots, X_n)$ объема n , а также всевозможные функции от выборки. Если элементы одномерной выборки упорядочить по возрастанию (построить *вариационный ряд* $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$) и отметить повторяемость наблюдений (подсчитать *частоту*), то получится *статистический ряд*, построенный по одномерной выборке $X_{[n]}$.

Разность между максимальным и минимальным элементами выборки называется *размахом*, $R = X_{\max} - X_{\min}$. При большом объеме выборки ее элементы иногда объединяются в группы, представляя результаты опытов в виде *группированного статистического ряда*. Для этого интервал, содержащий все элементы выборки, разбивается на k непересекающихся интервалов. Обычно разбиение производится на интервалы одинаковой длины $b = R/k$. После чего нетрудно определить частоты — количества n_i элементов выборки, попавших в i -ый интервал. Статистический ряд часто записывают в виде таблицы. В первой строке таблицы указывают середины интервалов группировки X_i , а во второй — частоты n_i . Подсчитываются также *накопленные частоты* $\sum_{j=1}^i n_j$, *относительные частоты* n_i/n , *накопленные относительные частоты* $\sum_{j=1}^i n_j/n$.

Для наглядного представления выборки применяют *гистограмму* и *полигон частот*. Полигоном частот называется ломаная с вершинами в точках $(X_i, n_i/b)$, а полигоном относительных частот — ломаная с вершинами в точках $(X_i, n_i/(nb))$. Эмпирическая функция распределения определяется равенством (2.1).

Выборочный начальный момент r -го порядка определяется равенством

$$a_r^* = \frac{1}{n} \sum_{i=1}^n X_i^r,$$

если выборка представлена статистическим рядом, то

$$a_r^* = \frac{1}{n} \sum_{i=1}^k n_i X_i^r,$$

выборочный центральный момент r -го порядка определяется равенством

$$a_r^{0*} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^r,$$

если выборка представлена статистическим рядом, то

$$a_r^{0*} = \frac{1}{n} \sum_{i=1}^k n_i (X_i - \bar{X})^r,$$

где $\bar{X} = a_1^* = \frac{1}{n} \sum_{i=1}^n X_i$ — выборочное среднее, для статистического ряда $\bar{X} = a_1^* = \frac{1}{n} \sum_{i=1}^k n_i X_i$.

Выборочная квантиль x_p порядка p определяется как элемент вариационного ряда $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ выборки $X_{[n]}$ с номером $[np] + 1$, где $[a]$ — целая часть числа a . В описательной статистике используют ряд квантилей, имеющих специальные названия *персентили* (квантили порядков 0.01; 0.02; ...; 0.99), *децили* (квантили порядков 0.1; 0.2; ...; 0.9), *квартили* (квантили порядков 0.25; 0.5; 0.75).

Наиболее распространенными характеристиками положения являются *выборочное среднее*, *выборочная медиана* (медианой называется число, которое делит вариационный ряд на две части, содержащие равное количество элементов; если $n = 2k + 1$, то медианой выборки является элемент вариационного ряда $X_{(k+1)}$, если $n = 2k$, то медианой выборки является число $(X_{(k)} + X_{(k+1)})/2$), *выборочная мода* (модой называется элемент выборки, имеющий наибольшую частоту).

Наиболее распространенными мерами рассеяния являются *размах* (размах $R = X_{\max} - X_{\min}$), *средний межквартильный размах* (три квартили Q_1, Q_2, Q_3 делят вариационный ряд на четыре части с равным числом элементов, тогда средний межквартильный размах равен $(Q_3 - Q_1)/2$), *персентильный размах* (персентильный размах равен разности персентилей

$P_{90} - P_{10}$), дисперсия (дисперсия $s^2 = a_2^{0*}$; исправленная дисперсия $\tilde{s}^2 = ns^2/(n-1)$) и среднее квадратическое отклонение (среднее квадратическое отклонение $\tilde{s} = \sqrt{\tilde{s}^2}$).

В качестве меры относительного разброса используют коэффициент вариации $v = s/\bar{X}$, иногда коэффициент записывают в процентах $C_v = v \cdot 100\%$.

Для оценки формы распределения служат коэффициент асимметрии $S_{k1} = a_3^{0*}/s^3$ и коэффициент эксцесса $K = a_4^{0*}/s^4 - 3$, для нормального распределения теоретические коэффициенты асимметрии и эксцесса, вычисляемые по распределению генеральной совокупности, равны нулю. Еще один показатель асимметрии вычисляется на основе квантилей $S_{k2} = (Q_3 + Q_1 - 2Q_2)/(Q_3 - Q_1)$.

СТАТИСТИКИ. ПРЕДЕЛЬНЫЕ РАСПРЕДЕЛЕНИЯ

§ 1. СТАТИСТИКИ ПЕРВОГО ТИПА

Будем рассматривать функционал $G(F)$, заданный на множестве функций распределения, следуя [5]:

$$G(F) = h \left(\int_{-\infty}^{\infty} g(x) dF(x) \right),$$

где $g : \mathbb{R} \rightarrow \mathbb{R}^m$ — заданная борелевская функция, $a = \int_{-\infty}^{\infty} g(x) dF_{\xi}(x) \in \mathbb{R}^m$, некоторая борелевская функция $h : \mathbb{R}^m \rightarrow \mathbb{R}^l$ непрерывная в точке a . Тогда назовем статистику $S(X_{[n]}) = G(F_n^*)$ статистикой первого типа. Таким образом, $S(X_{[n]}) = G(F_n^*) = h \left(\int_{-\infty}^{\infty} g(x) dF_n^*(x) \right) = h \left(\frac{1}{n} \sum_{i=1}^n g(X_i) \right)$.

Замечание 1.1. Если $h(t) \equiv t$, $g(x) = x^k$ то $G(F_n^*) = a_k^*$ начальный эмпирический момент порядка k . Нетрудно заметить, что центральный эмпирический момент также является статистикой первого типа. Покажем это для дисперсий (центральный момент второго порядка):

$$s^2 = a_2^{0*} = \frac{1}{n} \sum_{i=1}^n (X_i - a_1^*)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2,$$

$$a_2^0 = \int_{-\infty}^{+\infty} (x - a_1)^2 dF_{\xi}(x) = \int_{-\infty}^{+\infty} x^2 dF_{\xi}(x) - \left(\int_{-\infty}^{+\infty} x dF_{\xi}(x) \right)^2.$$

Выберем функции h и g следующего вида:

$$h(t_1, t_2) = t_2 - t_1^2,$$

$$g(x) = (g_1(x), g_2(x)) = (x, x^2).$$

Тогда получаем равенства:

$$a_2^0 = h \left(\int_{-\infty}^{+\infty} g(x) dF_\xi(x) \right),$$

$$a_2^{0*} = h \left(\frac{1}{n} \sum_{i=1}^n g(X_i) \right).$$

Теорема 1.1. Пусть $S(X_{[n]}) = G(F_n^*)$ — статистика первого типа, тогда имеет место сходимость:

$$S(X_{[n]}) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} h(a) = G(F_\xi).$$

Доказательство. Из усиленного закона больших чисел Колмогорова следует сходимость:

$$\frac{1}{n} \sum_{i=1}^n g(X_i) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} E g(\xi) = \int_{-\infty}^{+\infty} g(x) dF_\xi(x),$$

так как все слагаемые взаимно независимые, одинаково распределенные случайные величины. Функция $h(\cdot)$ непрерывна в точке a , следовательно,

$$G(F_n^*) = h \left(\frac{1}{n} \sum_{i=1}^n g(X_i) \right) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} h \left(\int_{-\infty}^{+\infty} g(x) dF_\xi(x) \right).$$

■

Следствие 1.1. Пусть у генеральной совокупности ξ существует теоретический начальный момент порядка k , $a_k = E\xi^k \in R$, тогда

$$a_k^* \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a_k.$$

Следствие 1.2. Пусть у генеральной совокупности ξ существует теоретический центральный момент порядка k , $a_k^0 = E(\xi - E\xi)^k \in R$, тогда

$$a_k^{0*} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a_k^0.$$

Замечание 1.2. Рассмотрим случайный вектор $\xi = (\xi_1, \dots, \dots, \xi_m)^T$ с m компонентами. Пусть имеется выборка:

$$(X_1, X_2, \dots, X_n) = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \dots & \dots & \dots & \dots \\ X_{m1} & X_{m2} & \dots & X_{mn} \end{pmatrix}.$$

Рассмотрим компоненты ξ_k и ξ_l , им соответствуют элементы выборки: X_{k1}, \dots, X_{kn} и X_{l1}, \dots, X_{ln} . Вычислим выборочные моменты:

$$a_{1k}^* = \bar{X}_k = \frac{1}{n} \sum_{i=1}^n X_{ki} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} E\xi_k,$$

$$a_{1l}^* = \bar{X}_l = \frac{1}{n} \sum_{i=1}^n X_{li} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} E\xi_l.$$

Кроме того,

$$a_{2k}^* = s_k^2 = \frac{1}{n} \sum_{i=1}^n (X_{ki} - \bar{X}_k)^2 \xrightarrow[n \rightarrow \infty]{\text{п.н.}} D\xi_k,$$

$$a_{2l}^* = s_l^2 = \frac{1}{n} \sum_{i=1}^n (X_{li} - \bar{X}_l)^2 \xrightarrow[n \rightarrow \infty]{\text{п.н.}} D\xi_l.$$

Коэффициент корреляции равен

$$\varrho(\xi_k, \xi_l) = \frac{\text{cov}(\xi_k, \xi_l)}{\sqrt{D\xi_k} \sqrt{D\xi_l}},$$

где

$$\text{cov}(\xi_k, \xi_l) = E\{(\xi_k - E\xi_k)(\xi_l - E\xi_l)\} = E(\xi_k \xi_l) - E(\xi_k)E(\xi_l).$$

Подберем эмпирический аналог ковариации. Рассмотрим элементы выборки:

$$\begin{pmatrix} X_{k1} \\ X_{l1} \end{pmatrix}, \begin{pmatrix} X_{k2} \\ X_{l2} \end{pmatrix}, \dots, \begin{pmatrix} X_{kn} \\ X_{ln} \end{pmatrix}.$$

Надо центрировать величины:

$$\begin{pmatrix} X_{k1} - \bar{X}_k \\ X_{l1} - \bar{X}_l \end{pmatrix}, \begin{pmatrix} X_{k2} - \bar{X}_k \\ X_{l2} - \bar{X}_l \end{pmatrix}, \dots, \begin{pmatrix} X_{kn} - \bar{X}_k \\ X_{ln} - \bar{X}_l \end{pmatrix}.$$

Перемножив их, получаем выборочную ковариацию:

$$\begin{aligned} \widehat{cov}(\xi_k, \xi_l) &= \frac{1}{n} \sum_{i=1}^n (X_{ki} - \bar{X}_k)(X_{li} - \bar{X}_l) = \\ &= \frac{1}{n} \sum_{i=1}^n X_{ki} X_{li} - \bar{X}_k \bar{X}_l \xrightarrow[n \rightarrow \infty]{\text{п.н.}} cov(\xi_k, \xi_l). \end{aligned}$$

Так как имеют место сходимости:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n X_{ki} X_{li} &\xrightarrow[n \rightarrow \infty]{\text{п.н.}} E(\xi_k \xi_l), \\ \bar{X}_k \bar{X}_l &\xrightarrow[n \rightarrow \infty]{\text{п.н.}} E\xi_k E\xi_l, \end{aligned}$$

тогда имеет место сходимость почти наверное для выборочного коэффициента корреляции:

$$\hat{\varrho}(\xi_k \xi_l) = \frac{\widehat{cov}(\xi_k, \xi_l)}{\sqrt{s_k^2 s_l^2}} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \varrho(\xi_k \xi_l).$$

Напомним, что под одномерной выборкой понимаем взаимно независимые случайные величины, распределенные так же, как генеральная совокупность.

§ 2. ТЕОРЕМЫ НЕПРЕРЫВНОСТИ

Теорема 2.1. Пусть η — случайная величина, заданная на вероятностном пространстве (Ω, \mathcal{F}, P) , последовательность случайных величин $\{\eta_n\}$ также задана на (Ω, \mathcal{F}, P) . Пусть борелевская функция $H : \mathbb{R} \rightarrow \mathbb{R}$ непрерывна на борелевском множестве $B \in \mathcal{B}(\mathbb{R})$, $P\{\eta \in B\} = 1$, тогда справедливы утверждения:

1. Если $\eta_n \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \eta$, тогда $H(\eta_n) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} H(\eta)$.
2. Если $\eta_n \xrightarrow[n \rightarrow \infty]{P} \eta$, тогда $H(\eta_n) \xrightarrow[n \rightarrow \infty]{P} H(\eta)$.

Доказательство. Докажем первое утверждение. Определим множество A следующим образом:

$$A = \{\omega : \eta_n(\omega) \xrightarrow[n \rightarrow \infty]{} \eta(\omega)\}.$$

По условию: $P(A) = 1$. Рассмотрим множество $A \cap \eta^{-1}(B)$. Определим вероятность:

$$P(\overline{A \cap \eta^{-1}(B)}) = P(\bar{A} \cup \overline{\eta^{-1}(B)}) \leq P(\bar{A}) + P(\overline{\eta^{-1}(B)}) = 0,$$

тогда $P(A \cap \eta^{-1}(B)) = 1$.

Теперь рассмотрим событие $\omega \in A \cap \eta^{-1}(B)$. Тогда

$$\eta_n(\omega) \xrightarrow[n \rightarrow \infty]{} \eta(\omega) \in B.$$

Следовательно, имеет место сходимость:

$$H(\eta_n(\omega)) \xrightarrow[n \rightarrow \infty]{} H(\eta(\omega)).$$

Докажем второе утверждение теоремы от противного. Пусть утверждение 2 неверно. Это означает, что существует $\varepsilon > 0$, существует $\delta > 0$ и существует последовательность $\{n_k^{(1)}\}$, что справедливо неравенство

$$P\{|H(\eta_{n_k^{(1)}}) - H(\eta)| > \varepsilon\} > \delta.$$

Последовательность $\eta_{n_k^{(1)}}$ сходится по вероятности к η , $\eta_{n_k^{(1)}} \xrightarrow[k \rightarrow \infty]{P} \eta$, следовательно, существует подпоследовательность $\{n_k^{(2)}\} \subset \{n_k^{(1)}\}$, для которой имеет место сходимость почти наверно:

$$\eta_{n_k^{(2)}} \xrightarrow[k \rightarrow \infty]{\text{п.н.}} \eta.$$

Возникло противоречие, так как

$$\eta_{n_k^{(2)}} \xrightarrow[k \rightarrow \infty]{\text{п.н.}} \eta,$$

но с другой стороны

$$P\{|H(\eta_{n_k^{(2)}}) - H(\eta)| > \varepsilon\} > \delta,$$

отсюда следует, что утверждение 2 верно. ■

Теорема 2.2. Пусть для последовательности случайных векторов имеет место сходимость по распределению: $\eta_n = (\eta_n^{(1)}, \dots, \eta_n^{(m)}) \xrightarrow[n \rightarrow \infty]{d} \eta = (\eta^{(1)}, \dots, \eta^{(m)})$. Пусть задана функция $H : \mathbb{R}^m \rightarrow \mathbb{R}^l$ непрерывная на \mathbb{R}^m , тогда $H(\eta_n) \xrightarrow[n \rightarrow \infty]{d} H(\eta)$.

Доказательство. Для любой непрерывной и ограниченной функции $g(x)$ должно выполняться следующее предельное соотношение:

$$Eg(H(\eta_n)) \xrightarrow[n \rightarrow \infty]{} Eg(H(\eta)).$$

Рассмотрим $g(H(t)) \equiv h(t)$, т. е. h — суперпозиция двух непрерывных функций, следовательно, она непрерывна. Так как g ограничена, то отсюда следует, что h ограничена, то есть выполнено следующее соотношение:

$$Eh(\eta_n) \xrightarrow[n \rightarrow \infty]{} Eh(\eta).$$

■

Теорема 2.3. Пусть борелевская функция $H : \mathbb{R}^m \rightarrow \mathbb{R}^k$ непрерывна на $B \in \mathcal{B}(\mathbb{R}^m)$ и $P\{\eta \in B\} = 1$. Пусть $\eta_n^T = (\eta_{1n}, \dots, \eta_{mn}) \xrightarrow[n \rightarrow \infty]{d} \eta$, тогда $H(\eta_n) \xrightarrow[n \rightarrow \infty]{d} H(\eta)$.

Доказательство. Доказательство теоремы можно найти в [2]. ■

Теорема 2.4. Пусть последовательность случайных величин η_n сходится по распределению к случайной величине η , $\eta_n \xrightarrow[n \rightarrow \infty]{d} \eta$. Пусть функция $H : \mathbb{R} \rightarrow \mathbb{R}$ — борелевская функция. Числовая последовательность $b_n \xrightarrow[n \rightarrow \infty]{} 0$, причем $b_n \neq 0$ для любого n . Тогда справедливы утверждения:

1. Если функция H дифференцируема в точке $a \in \mathbb{R}$, то

$$\frac{H(a + b_n \eta_n) - H(a)}{b_n} \xrightarrow[n \rightarrow \infty]{d} H'(a) \eta.$$

2. Если функция H дифференцируема в некоторой окрестности точки a , $H'(a) = 0$, и существует $H''(a)$, то

$$\frac{H(a + b_n \eta_n) - H(a)}{b_n^2} \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2} H''(a) \eta^2.$$

Доказательство. Докажем утверждение 1. Введем $H_1(x)$ и $\tilde{H}_1(x, y) = H_1(x)y$, причем

$$H_1(x) = \begin{cases} \frac{H(a+x) - H(a)}{x}, & x \neq 0, \\ H'(a), & x = 0. \end{cases}$$

Из условия теоремы ясно, что $H_1(x)$ непрерывна в нуле, тогда $\tilde{H}_1(x, y)$ непрерывна на множестве $\{(0, y), y \in \mathbb{R}\}$. Покажем, что $(b_n \eta_n, \eta_n) \xrightarrow[n \rightarrow \infty]{d} (0, \eta)$. Воспользуемся методом характеристических функций. Надо показать, что

$$\varphi_{(b_n \eta_n, \eta_n)}(t_1, t_2) \xrightarrow[n \rightarrow \infty]{} \varphi_{(0, \eta)}(t_1, t_2). \quad (2.1)$$

Рассмотрим левую часть (2.1):

$$\begin{aligned} \varphi_{(b_n \eta_n, \eta_n)}(t_1, t_2) &= E e^{i(t_1 b_n \eta_n + t_2 \eta_n)} = \\ &= E e^{i(t_1 b_n + t_2) \eta_n} = \varphi_{\eta_n}(b_n t_1 + t_2) \pm \varphi_{\eta}(b_n t_1 + t_2) = \\ &= (\varphi_{\eta_n}(b_n t_1 + t_2) - \varphi_{\eta}(b_n t_1 + t_2)) + \varphi_{\eta}(b_n t_1 + t_2) \xrightarrow[n \rightarrow \infty]{} \\ &\xrightarrow[n \rightarrow \infty]{} \varphi_{\eta}(t_2) = E e^{i t_2 \eta + i t_1 0} = \varphi_{(0, \eta)}(t_1, t_2), \end{aligned}$$

так как t_1, t_2 фиксированы, то $(b_n t_1 + t_2) \in [a, b]$, то есть сходимость на данном промежутке равномерная (как показано в [22], из поточечной сходимости последовательности характеристических функций к некоторой характеристической функции на конечном промежутке $[a, b]$ следует равномерная сходимость на этом промежутке), тогда $(\varphi_{\eta_n}(b_n t_1 + t_2) - \varphi_{\eta}(b_n t_1 + t_2)) \xrightarrow[n \rightarrow \infty]{} 0$. Таким образом, показали, что

$$\varphi_{(b_n \eta_n, \eta_n)}(t_1, t_2) \xrightarrow[n \rightarrow \infty]{} \varphi_{(0, \eta)}(t_1, t_2).$$

Заметим, что $P\{(0, \eta) \in \{(0, y) : y \in R\}\} = 1$, подставим в \tilde{H}_1 случайный вектор $(b_n \eta_n, \eta_n)$, получим:

$$\tilde{H}_1(b_n \eta_n, \eta_n) = \frac{H(a + b_n \eta_n) - H(a)}{b_n},$$

$$\tilde{H}_1(b_n \eta_n, \eta_n) \xrightarrow[n \rightarrow \infty]{d} \tilde{H}_1(0, \eta) = H'(a)\eta,$$

как следует из теоремы 2.3.

Докажем второе утверждение теоремы. Запишем формулу Тейлора второго порядка:

$$H(a + x) = H(a) + \frac{1}{2}H''(a)x^2 + o(x^2).$$

Введем функцию:

$$H_2(x) = \begin{cases} \frac{H(a+x) - H(a)}{x^2}, & x \neq 0, \\ \frac{1}{2}H''(a), & x = 0. \end{cases}$$

Введем функцию $\tilde{H}_2(x, y) = H_2(x)y^2$, функция непрерывна на множестве $\{(0, y) : y \in \mathbb{R}\}$. Подставим в \tilde{H}_2 случайный вектор $(b_n \eta_n, \eta_n)$:

$$\tilde{H}_2(b_n \eta_n, \eta_n) \xrightarrow[n \rightarrow \infty]{d} \tilde{H}_2(0, \eta) = \frac{1}{2}H''(a)\eta^2,$$

$$\tilde{H}_2(b_n \eta_n, \eta_n) = \frac{H(a + b_n \eta_n) - H(a)}{b_n^2}.$$

Утверждение 2 доказано. ■

Замечание 2.1. Теорема 2.4 допускает обобщение на многомерный случай.

Теорема 2.5. Пусть задана функция $H : \mathbb{R}^m \rightarrow \mathbb{R}$ и последовательность случайных векторов $\eta_n = (\eta_n^{(1)}, \dots, \eta_n^{(m)}) \xrightarrow[n \rightarrow \infty]{d} \eta = (\eta^{(1)}, \dots, \eta^{(m)})$. Пусть числовая последовательность $b_n \xrightarrow[n \rightarrow \infty]{} 0$ такая, что $b_n \neq 0$ для любого n . Тогда справедливы утверждения:

1. Если существует $\frac{\partial H}{\partial t}|_{t=a} = \left(\frac{\partial H}{\partial t_1}, \dots, \frac{\partial H}{\partial t_m} \right) \Big|_{t=a}$, $a \in \mathbb{R}^m$, то

$$\frac{H(a + b_n \eta_n) - H(a)}{b_n} \xrightarrow[n \rightarrow \infty]{d} H'(a) \eta^T.$$

2. Если $H'(a) = 0$, и существует $H''(a) = \left(\frac{\partial^2 H}{\partial t_i \partial t_j} \right) \Big|_{t=a}$, то

$$\frac{H(a + b_n \eta_n) - H(a)}{b_n^2} \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2} \eta H''(a) \eta^T.$$

Доказательство. Доказательство этой теоремы аналогично доказательству теоремы 2.4. ■

Замечание 2.2. Можно продолжить обобщение теоремы 2.5 на случай $H : \mathbb{R}^m \rightarrow \mathbb{R}^l$.

§ 3. ПРЕДЕЛЬНОЕ РАСПРЕДЕЛЕНИЕ СТАТИСТИК ПЕРВОГО ТИПА

Пусть задана статистика первого типа:

$$S(X_{[n]}) = G(F_n^*) = h \left(\frac{1}{n} \sum_{i=1}^n g(X_i) \right),$$

где $G(F) = h \left(\int_{-\infty}^{\infty} g(x) dF(x) \right)$, $\int_{-\infty}^{\infty} g(x) dF_{\xi}(x) = a \in \mathbb{R}^m$, т. е. $G(F_{\xi}) = h(a)$.

Теорема 3.1. Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ с функцией распределения F_{ξ} и $S(X_{[n]}) = h \left(\frac{1}{n} \sum_{i=1}^n g(X_i) \right)$ — статистика I типа, борелевские функции $h : \mathbb{R} \rightarrow \mathbb{R}$, $g : \mathbb{R} \rightarrow \mathbb{R}$, тогда справедливы утверждения:

1. Если существует $h'(a)$, то

$$\sqrt{n} (S(X_{[n]}) - h(a)) \xrightarrow[n \rightarrow \infty]{d} h'(a) \zeta,$$

где ζ — случайная величина, распределенная нормально с параметрами $(0, Dg(\xi))$, $\zeta \sim N(0, Dg(\xi))$.

2. Если $h'(a) = 0$ и существует $h''(a)$, то

$$n(S(X_{[n]}) - h(a)) \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2}h''(a)\zeta^2.$$

Доказательство. Применима теорема 2.4. Преобразуем функцию $G(F^*)$:

$$\begin{aligned} G(F^*) &= h\left(\frac{1}{n}\sum_{i=1}^n g(X_i)\right) = h\left(a + \frac{1}{n}\sum_{i=1}^n (g(X_i) - a)\right) = \\ &= h\left(a + \frac{1}{\sqrt{n}\sqrt{n}}\sum_{i=1}^n (g(X_i) - a)\right), \end{aligned}$$

где $a = Eg(X_1)$. По центральной предельной теореме для одинаково распределенных слагаемых, справедливо:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (g(X_i) - a) = \xi_n \xrightarrow[n \rightarrow \infty]{d} \xi \sim N(0, \sigma^2),$$

где $\sigma^2 = Dg(\xi)$. Заметим, что $\frac{1}{\sqrt{n}} = b_n$. ■

Теорема 3.2. Пусть задана статистика I типа $S(X_{[n]}) = h\left(\frac{1}{n}\sum_{i=1}^n g(X_i)\right)$ и борелевские функции $h: \mathbb{R}^m \rightarrow \mathbb{R}$, $g: \mathbb{R} \rightarrow \mathbb{R}^m$, тогда справедливы утверждения:

1. Если существует $h'(a) = \left(\frac{\partial h}{\partial t_1}, \dots, \frac{\partial h}{\partial t_m}\right)\Big|_{t=a}$, где $a = Eg(\xi) = (Eg_1(\xi), \dots, Eg_m(\xi))$, то

$$\sqrt{n}(S(X_{[n]}) - h(a)) \xrightarrow[n \rightarrow \infty]{d} h'(a)\zeta^T,$$

где случайный вектор $\zeta = (\zeta_1, \dots, \zeta_m)$ подчиняется многомерному нормальному распределению с параметрами $(0, Dg(\xi))$, $\zeta \sim N(0, Dg(\xi))$.

2. Если $h'(a) = 0$ и существует $h''(a)$, то

$$n(S(X_{[n]}) - h(a)) \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2}\zeta h''(a)\zeta^T.$$

Доказательство. Доказательство проводится аналогично доказательству теоремы 3.1. Применяется центральная предельная теорема для одинаково распределенных случайных векторов. ■

Пример 3.1. Рассмотрим генеральную совокупность ξ , для которой $E\xi = \alpha > 0$, $D\xi = \sigma^2$. Получена выборка $X_{[n]} = (X_1, \dots, X_n)$ из ξ . Найдем асимптотическое распределение статистики $1/\bar{X}$, где $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = a_1^*$. Покажем, что $1/\bar{X}$ — статистика первого типа.

Для доказательства достаточно взять $h(t) = 1/t$, $g(x) = x$ и заметить, что

$$1/\bar{X} = h\left(\frac{1}{n} \sum_{i=1}^n g(X_i)\right).$$

Очевидно, что

$$a = \int_{-\infty}^{+\infty} g(x) dF_{\xi}(x) = Eg(\xi) = E\xi = \alpha,$$

при этом, $h(\alpha) = 1/\alpha$, $1/\bar{X} = h(\frac{1}{n} \sum_{i=1}^n g(x_i))$ является статистикой первого типа.

Из теоремы 3.1 следует, что

$$\sqrt{n} \left(\frac{1}{\bar{X}} - \frac{1}{\alpha} \right) \xrightarrow[n \rightarrow \infty]{d} \xi \left(\frac{-1}{\alpha^2} \right) = -\xi \frac{1}{\alpha^2},$$

где $\xi \sim N(0, \sigma^2)$.

§ 4. ПРЕДЕЛЬНОЕ РАСПРЕДЕЛЕНИЕ СТАТИСТИКИ ПИРСОНА

Пусть задана генеральная совокупность ξ с функцией распределения F_{ξ} и выборка $X_{[n]} = (X_1, \dots, X_n)$. Разобьем числовую ось на r непересекающихся интервалов: $-\infty = a_0 < a_1 < \dots < a_r = \infty$. Обозначим через $\Delta_1 = (-\infty, a_1]$, $\Delta_2 = (a_1, a_2]$, \dots , $\Delta_r = (a_{r-1}, \infty)$.

Пусть $p_i = F_\xi(a_i) - F_\xi(a_{i-1})$ — вероятность того, что случайная величина ξ попадет в интервал Δ_i , $\sum_{i=1}^r p_i = 1$. Пусть n_i — количество элементов выборки $X_{[n]}$, попавших в Δ_i . Определим статистику χ^2 следующим образом:

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}, \quad (4.1)$$

где n_i — частота (количество элементов выборки, попавших в Δ_i), np_i — ожидаемое количество наблюдений в интервале Δ_i .

Определение 4.1. Статистики вида (4.1) называются *статистиками χ^2* или *статистиками Пирсона*.

Покажем, что статистика Пирсона может быть преобразована к статистике первого типа:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i} = n \sum_{i=1}^r \left(\frac{1}{\sqrt{p_i}} \frac{n_i}{n} - \sqrt{p_i} \right)^2 = \\ &= n \sum_{i=1}^r \left(\frac{1}{n} \sum_{j=1}^n \frac{1}{\sqrt{p_i}} I\{X_j \in \Delta_i\} - \sqrt{p_i} \right)^2. \end{aligned} \quad (4.2)$$

Рассмотрим статистику первого типа $S(X_{[n]}) = h\left(\frac{1}{n} \sum_{j=1}^n g(X_j)\right)$,

где в качестве h возьмем функцию $h(t_1, \dots, t_r) = \sum_{i=1}^r (t_i - \sqrt{p_i})^2$, $g(x) = (g_1(x), \dots, g_r(x))$ и $g_i(x) = \frac{1}{\sqrt{p_i}} I\{x \in \Delta_i\}$, $i = 1, \dots, r$. Получаем:

$$\frac{1}{n} \sum_{j=1}^n g(X_j) = \left(\frac{1}{\sqrt{p_1}} \frac{n_1}{n}, \dots, \frac{1}{\sqrt{p_r}} \frac{n_r}{n} \right).$$

Таким образом, статистика χ^2 представляет собой произведение константы n на статистику первого типа. Следовательно, можно воспользоваться теоремой 3.2 о предельном распределении статистик первого типа.

Очевидно, что $a = (a_1, \dots, a_r) = (Eg_1(\xi), \dots, Eg_r(\xi)) = (\sqrt{p_1}, \dots, \sqrt{p_r})$, при этом $h'(a) = 0$, $\frac{1}{2}h''(a) = E_r$, где E_r — единичная матрица порядка r . Из теоремы 3.2 получаем:

$$\chi^2(X_{[n]}) \xrightarrow[n \rightarrow \infty]{d} \zeta^T \zeta = \zeta_1^2 + \dots + \zeta_r^2,$$

где $\zeta \sim N(0, Dg(\xi))$. Вычислим ковариационную матрицу $Dg(\xi)$:

$$Dg(\xi) = E \{g(\xi)g^T(\xi)\} - Eg(\xi)(Eg(\xi))^T,$$

после дополнительных преобразований получаем

$$Dg(\xi) = E_r - (\sqrt{p_1}, \dots, \sqrt{p_r})^T (\sqrt{p_1}, \dots, \sqrt{p_r}).$$

Пусть C — ортонормированная матрица $CC^T = C^TC = E_r$. Зафиксируем первую строку $c_1 = (\sqrt{p_1}, \dots, \sqrt{p_r})$ матрицы C . Остальные строки будем искать методом ортогонализации Грама–Шмидта.

Случайный вектор $\eta = C\zeta$ подчиняется многомерному нормальному распределению $N(0, CDg(\xi)C^T)$, при этом из выбора матрицы C следует, что $D\eta_1 = 0$, кроме того, $E\eta_1 = 0$, следовательно, $\eta_1 \stackrel{\text{п.н.}}{=} 0$.

$$CDg(\xi)C^T = \begin{pmatrix} 0 & 0 \dots 0 \\ 0 & \\ \dots & E_{r-1} \\ 0 & \end{pmatrix}.$$

Как легко видеть, $\eta^T \eta = \zeta^T C^T C \zeta = \zeta^T \zeta = \eta_1^2 + \eta_2^2 + \dots + \eta_r^2 = \zeta_1^2 + \zeta_2^2 + \dots + \zeta_r^2$. Следовательно, распределения сумм одинаковы, поэтому

$$\chi^2 \xrightarrow[n \rightarrow \infty]{d} \eta_1^2 + \dots + \eta_r^2,$$

но $\eta_1^2 + \dots + \eta_r^2 \stackrel{\text{п.н.}}{=} \eta_2^2 + \dots + \eta_r^2$, тогда

$$\chi^2 \xrightarrow[n \rightarrow \infty]{d} \eta_2^2 + \dots + \eta_r^2,$$

где η_2, \dots, η_r — взаимно независимые одинаково распределенные случайные величины, $\eta_i \sim N(0, 1)$, $i = 2, \dots, r$, $\eta_1 \stackrel{\text{п.н.}}{=} 0$.

Определение 4.2. Пусть $\delta_1, \dots, \delta_k$ — взаимно независимые одинаково распределенные стандартные гауссовы случайные величины, тогда распределение случайной величины $\delta_1^2 + \dots + \delta_k^2$, называется распределением χ^2 с k степенями свободы (или распределением Пирсона с k степенями свободы).

Проведенные рассуждения доказывают следующую теорему.

Теорема 4.1. Статистика χ^2 , определяемая равенством (4.1), асимптотически распределена по закону хи-квадрат с $r - 1$ степенью свободы:

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i} \xrightarrow[n \rightarrow \infty]{d} \tau,$$

где τ подчиняется распределению хи-квадрат с $r - 1$ степенью свободы.

Замечание 4.1. Нетрудно заметить, что к статистике χ^2 , определяемой формулой (4.1), можно прийти, исходя из генеральной совокупности, подчиняющейся полиномиальному распределению с r возможными исходами, где вероятности p_1, p_2, \dots, p_r представляют собой вероятности появления соответствующих исходов ($\sum_{i=1}^r p_i = 1$), n — число испытаний, n_1 — количество появлений первого исхода, n_2 — количество появлений второго исхода, \dots , n_r — количество появлений исхода с номером r , $\sum_{i=1}^r n_i = n$. В статистическом эксперименте непосредственно наблюдается выборка частот (n_1, n_2, \dots, n_r) . Утверждение теоремы 4.1 сохраняется и для рассматриваемого полиномиального распределения.

§ 5. ГАММА-РАСПРЕДЕЛЕНИЕ

Гамма-функция определяется следующим образом:

$$\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx, \quad p > 0.$$

Свойства гамма-функции:

1. Справедливы равенства:

$$\Gamma(1) = 1, \Gamma(2) = 1,$$

$$\begin{aligned} \Gamma\left(\frac{1}{2}\right) &= \int_0^{+\infty} x^{-\frac{1}{2}} e^{-x} dx = 2 \int_0^{+\infty} e^{-x} d(x^{\frac{1}{2}}) = \\ &= 2 \int_0^{+\infty} e^{-y^2} dy = \int_{-\infty}^{+\infty} e^{-y^2} dy = \left\{ \left(\int_{-\infty}^{+\infty} e^{-y^2} dy \right)^2 \right\}^{\frac{1}{2}} = \\ &= \left\{ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-y^2} e^{-x^2} dx dy \right\}^{\frac{1}{2}} = \sqrt{\pi}, \end{aligned}$$

в последнем интеграле переходим к полярным координатам.

2. При $p > 1$, интегрируя по частям, нетрудно получить равенство:

$$\Gamma(p) = (p-1)\Gamma(p-1).$$

Если $n \in \mathbb{N}$, то $\Gamma(n) = (n-1)!$

Плотность распределения случайной величины ξ , соответствующая стандартному гамма-распределению с параметром формы p , определяется формулой:

$$f_{\xi}(x) = \begin{cases} \frac{x^{p-1}}{\Gamma(p)} e^{-x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Нетрудно заметить, что $f_{\xi}(x) \geq 0$ и $\int_{-\infty}^{+\infty} f_{\xi}(x) dx = 1$.

Рассмотрим случайную величину $\eta = \xi/\lambda$, параметр $\lambda > 0$, нетрудно получить выражение для плотности распределения, соответствующей гамма-распределению $G(\lambda, p)$ с параметром формы p и параметром масштаба λ :

$$f_{\eta}(x) = \begin{cases} \frac{\lambda^p x^{p-1}}{\Gamma(p)} e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases} \quad (5.1)$$

Если в формуле (5.1) положить $\lambda = 1$, то получим плотность стандартного гамма-распределения, а если в формуле (5.1) положить $p = 1$, то получим плотность экспоненциального распределения.

Лемма 5.1. Пусть δ — случайная величина, подчиняющаяся стандартному нормальному распределению, $\delta \sim N(0, 1)$, тогда случайная величина δ^2 подчиняется гамма-распределению с параметрами $p = 1/2$, $\lambda = 1/2$.

Доказательство. Пусть $\delta \sim N(0, 1)$, то есть $f_\delta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Функцию распределения случайной величины δ обозначим через

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt.$$

Найдем функцию распределения случайной величины δ^2 :

$$P\{\delta^2 \leq y\} = P(-\sqrt{y} \leq \delta \leq \sqrt{y}) = F(\sqrt{y}) - F(-\sqrt{y}),$$

дифференцируя, найдем плотность распределения для $y > 0$:

$$\begin{aligned} f_{\delta^2}(y) &= \frac{1}{2} \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} + \frac{1}{2} \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} = \\ &= \left(\frac{1}{2}\right)^{\frac{1}{2}} y^{-\frac{1}{2}} \frac{1}{\sqrt{\pi}} e^{-\frac{y}{2}} = \begin{cases} \frac{(\frac{1}{2})^{\frac{1}{2}} y^{-\frac{1}{2}}}{\Gamma(\frac{1}{2})} e^{-\frac{y}{2}}, & y > 0; \\ 0, & y < 0. \end{cases} \end{aligned}$$

Следовательно, случайная величина δ^2 подчиняется гамма-распределению с параметрами $p = 1/2$, $\lambda = 1/2$. ■

Лемма 5.2. Пусть заданы взаимно независимые случайные величины $\xi_1 \sim G(\lambda, p_1)$, $\xi_2 \sim G(\lambda, p_2)$, тогда $\xi = \xi_1 + \xi_2 \sim G(\lambda, p_1 + p_2)$.

Доказательство. Очевидно, что

$$\begin{aligned} P\{\xi \leq y\} &= \iint_{\{(x_1, x_2) | x_1 + x_2 \leq y\}} f_{\xi_1}(x_1) f_{\xi_2}(x_2) dx_1 dx_2 = \\ &= \int_{-\infty}^{+\infty} f_{\xi_1}(x_1) \int_{-\infty}^{y-x_1} f_{\xi_2}(x_2) dx_2 dx_1. \end{aligned}$$

Сделаем замену переменных: $u = x_1 + x_2$, $du = dx_2$, тогда

$$\begin{aligned} P\{\xi \leq y\} &= \int_{-\infty}^{+\infty} f_{\xi_1}(x_1) \int_{-\infty}^y f_{\xi_2}(u - x_1) du dx_1 = \\ &= \int_{-\infty}^y \int_{-\infty}^{+\infty} f_{\xi_1}(x_1) f_{\xi_2}(u - x_1) dx_1 du. \end{aligned}$$

Получили формулу свертки для плотности суммы двух независимых случайных величин:

$$\begin{aligned} f_{\xi}(u) &= \int_{-\infty}^{+\infty} f_{\xi_1}(x_1) f_{\xi_2}(u - x_1) dx_1 = \int_0^u f_{\xi_1}(x_1) f_{\xi_2}(u - x_1) dx_1 = \\ &= \int_0^u \frac{\lambda^{p_1+p_2}}{\Gamma(p_1)\Gamma(p_2)} x_1^{p_1-1} (u - x_1)^{p_2-1} e^{-\lambda x_1} e^{-\lambda(u-x_1)} dx_1 = \\ &= \frac{\lambda^{p_1+p_2} e^{-\lambda u}}{\Gamma(p_1)\Gamma(p_2)} \int_0^u x_1^{p_1-1} (u - x_1)^{p_2-1} dx_1. \end{aligned}$$

Сделаем замену переменных под знаком интеграла: $x_1 = su$, $s \in (0, 1)$, $dx_1 = u ds$, тогда

$$\begin{aligned} f_{\xi}(u) &= \frac{\lambda^{p_1+p_2} e^{-\lambda u}}{\Gamma(p_1)\Gamma(p_2)} u^{p_1+p_2-1} \int_0^1 s^{p_1-1} (1-s)^{p_2-1} ds = \\ &= \begin{cases} cu^{p_1+p_2-1} e^{-\lambda u} \lambda^{p_1+p_2}, & u > 0; \\ 0, & u \leq 0, \end{cases} \end{aligned}$$

где $c = (\int_0^1 s^{p_1-1}(1-s)^{p_2-1}ds)/(\Gamma(p_1)\Gamma(p_2))$. Найдем c из условия нормировки:

$$c \int_0^{+\infty} \lambda^{p_1+p_2} u^{p_1+p_2-1} e^{-\lambda u} du = 1.$$

Сделаем замены переменных: $x = \lambda u$, $u = x/\lambda$, $du = dx/\lambda$, тогда

$$c \int_0^{+\infty} x^{p_1+p_2-1} e^{-x} dx = c\Gamma(p_1 + p_2) = 1$$

или

$$c = \frac{1}{\Gamma(p_1 + p_2)}.$$

Плотность $f_\xi(x)$ имеет вид:

$$f_\xi(x) = \begin{cases} \frac{\lambda^{p_1+p_2}}{\Gamma(p_1+p_2)} x^{p_1+p_2-1} e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

■

В ходе доказательства леммы 5.2 получено тождество, связывающее бета-функцию с гамма-функциями:

$$B(p_1, p_2) = \int_0^1 s^{p_1-1}(1-s)^{p_2-1} ds = \frac{\Gamma(p_1)\Gamma(p_2)}{\Gamma(p_1 + p_2)}.$$

Следствие 5.1. Пусть случайные величины δ_i , $i = 1, \dots, m$ взаимно независимы, одинаково распределены и подчиняются стандартному нормальному распределению, $\delta_i \sim N(0, 1)$. Тогда $\delta_1^2 + \dots + \delta_m^2 \sim G(\frac{1}{2}, \frac{m}{2})$.

Доказательство следует из лемм 5.1, 5.2.

Замечание 5.1. Доказано важное утверждение: распределение хи-квадрат с m степенями свободы является частным случаем гамма распределения с параметрами $\lambda = 1/2$, $p = m/2$.

Если случайная величина τ подчиняется распределению хи-квадрат с m степенями свободы, то ее плотность имеет вид:

$$f_{\tau}(x) = \begin{cases} \frac{x^{\frac{m}{2}-1} e^{-\frac{1}{2}x}}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

КРИТЕРИИ СОГЛАСИЯ ДЛЯ ПРОСТЫХ ГИПОТЕЗ

§ 1. КРИТЕРИЙ СОГЛАСИЯ ПИРСОНА

Определение 1.1. *Статистической гипотезой* называется любое предположение о законе распределения генеральной совокупности.

Определение 1.2. Гипотеза называется *простой*, если в ней единственным образом определяется закон распределения генеральной совокупности. В противном случае гипотеза называется *сложной*.

Один из типов гипотез — *гипотезы согласия*. Методы проверки этих гипотез — *критерии согласия*. Пусть задана генеральная совокупность ξ , функция распределения F_ξ которой взаимно однозначно соответствует распределению генеральной совокупности P_ξ , и выборка $X_{[n]} = (X_1, \dots, X_n)$. Гипотезы принято обозначать через H . Пусть H_0 — основная (нулевая) гипотеза.

Пусть проверяется гипотеза согласия $H_0 : F_\xi = F_0$, при этом предполагается, что $F_0(x)$ известна. Очевидно, что данная гипотеза — простая. Сформулируем альтернативную гипотезу: $H_1 : F_\xi \neq F_0$. Для нас важна гипотеза H_0 , необходимо решить: принять ее или отклонить. Решение принимается по имеющейся выборке $X_{[n]}$, т. е. проверяется «хорошо» ли выборка $X_{[n]}$ согласуется с F_0 . Мы можем принять гипотезу, но при этом она на самом деле может быть неверной.

Рассмотрим критерий Пирсона (критерий χ^2). Числовую ось разбиваем на r промежутков $-\infty = a_0 < a_1 < \dots < a_r = \infty$, $\Delta_i = (a_{i-1}, a_i]$, $i = 1, \dots, r$, и построим статистику χ^2 :

$$\chi^2(X_{[n]}) = \sum_{i=1}^r \frac{(n_i - np_i^{(0)})^2}{np_i^{(0)}},$$

где $p_i^{(0)} = F_0(a_i) - F_0(a_{i-1})$. Если H_0 верна, тогда по теореме 4.1 главы 13: $\chi^2(X_{[n]}) \xrightarrow[n \rightarrow \infty]{d} \zeta$, где ζ подчиняется распределению хи-квадрат с $r - 1$ степенью свободы. В § 4 главы 13 было показано, что $\chi^2(X_{[n]}) = nS(X_{[n]}) = nh(\frac{1}{n} \sum_{j=1}^n X_j)$, где

$$h(t_1, \dots, t_r) = \sum_{i=1}^r \left(t_i - \sqrt{p_i^{(0)}} \right)^2, \quad g(x) = (g_1(x), \dots, g_r(x)) \quad \text{и} \\ g_i(x) = \frac{1}{\sqrt{p_i^{(0)}}} I \{x \in \Delta_i\}, \quad i = 1, \dots, r.$$

Статистика $S(X_{[n]}) = h\left(\frac{1}{n} \sum_{j=1}^n g(X_j)\right)$ является статистикой первого типа. При гипотезе H_0 имеем

$$a = \left(\sqrt{p_1^{(0)}}, \dots, \sqrt{p_r^{(0)}} \right) \quad \text{и} \quad h(a) = 0.$$

Если H_0 верна, тогда $h\left(\frac{1}{n} \sum_{j=1}^n g(X_j)\right) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} h(a) = 0$.

Если верна гипотеза H_1 , то $a \neq \left(\sqrt{p_1^{(0)}}, \dots, \sqrt{p_r^{(0)}} \right)$, тогда $h(a) \neq 0$, но статистика $h\left(\frac{1}{n} \sum_{j=1}^n g(X_j)\right) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} h(a) \neq 0$, следовательно,

$$\chi^2 = nh\left(\frac{1}{n} \sum_{j=1}^n g(X_j)\right) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \infty.$$

Поведение статистики χ^2 зависит от того, верна нулевая гипотеза или нет.

Нетрудно понять, как выбрать критическую область для гипотезы H_0 . Если $\chi^2(X_{[n]}) \geq C$, то H_0 отклоняется, если $\chi^2(X_{[n]}) < C$, то нет оснований для отклонения H_0 .

Выберем вероятность $\alpha \in (0, 1)$. В качестве α часто выбирают 0,05. Константу $C(r - 1, \alpha)$ выберем из условия

$$P\{\eta \geq C\} = \alpha,$$

где случайная величина η подчиняется распределению хи-квадрат с $r - 1$ степенью свободы. Константа $C(r - 1, \alpha)$ представляет собой квантиль уровня $1 - \alpha$ распределения хи-квадрат с $r - 1$ степенью свободы. Для практического нахождения квантили можно использовать статистические таблицы. Область $(C(r - 1, \alpha), \infty)$ является критической для гипотезы H_0 . Если $\chi^2(X_{[n]}) > C(r - 1, \alpha)$, то H_0 отклоняется, а если $\chi^2(X_{[n]}) \leq C(r - 1, \alpha)$, то для отклонения нет оснований.

В результате применения критерия могут возникать следующие ошибки:

- Ошибка первого рода, если мы отбросили гипотезу H_0 , а она на самом деле верна.
- Ошибка второго рода, если мы принимаем гипотезу H_0 , а она на самом деле не верна.

Учитывая доказанную асимптотику, вероятность ошибки первого рода приближенно совпадает с заданной вероятностью α . Вероятность ошибки первого рода называют часто уровнем значимости критерия.

Возможен альтернативный подход. Найдем вероятность

$$P\{\eta > \chi^2(X_{[n]})\} = 1 - F_{\chi^2_{r-1}}(\chi^2(X_{[n]}),$$

где $F_{\chi^2_{r-1}}(\cdot)$ — функция распределения хи-квадрат с $r - 1$ степенью свободы. Вероятность вычисляется при условии справедливости H_0 . Эта вероятность называется p -значением (p -value). В программах типа SPSS, STATISTICA реализован второй подход.

Большие значения статистики χ^2 свидетельствуют против H_0 . Интервал $(\chi^2(n_1, \dots, n_r); +\infty)$ — критическая область значений статистики χ^2 , там находятся еще более «худшие»

значения статистики для гипотезы H_0 :

$$P_{H_0}(\chi^2(n_1, \dots, n_r); +\infty) = 1 - F_{\chi^2_{r-1}}(\chi^2(n_1, \dots, n_r)) = p.$$

Далее можно выбрать значение α (например, 0,05) и применить следующий критерий. Если $p \leq \alpha$, то мы отвергаем H_0 и принимаем альтернативу H_1 . Если $p > \alpha$, то гипотеза H_0 согласуется с наблюдениями. Конечно, указанный критерий полностью эквивалентен критерию в первом подходе.

Можно отметить некоторые общие особенности статистических критериев, которые обсуждались выше. Любой статистический критерий устроен следующим образом:

1. Выбираем статистику критерия, закон распределения которой известен, если справедлива H_0 .
2. Определяем критическую область для гипотезы H_0 , попадание в которую маловероятно, если гипотеза H_0 верна, но при этом вероятность попадания в эту область при условии истинности гипотезы H_1 больше, чем при условии истинности гипотезы H_0 .
3. Уровень значимости или вероятность ошибки первого рода — вероятность попадания в критическую область при условии истинности H_0 , обычно выбирается малой, например, 0,05. Существует правило p -значения: находится вероятность получения значений статистики, которые еще «хуже», чем полученное значение при условии истинности H_0 . Если p -значение оказывается меньше заданного уровня значимости α , то нулевую гипотезу отвергают.

§ 2. КРИТЕРИЙ СОГЛАСИЯ КОЛМОГОРОВА

Пусть задана генеральная совокупность ξ , функция распределения F_ξ которой взаимно однозначно соответствует распределению P_ξ , и выборка $X_{[n]} = (X_1, \dots, X_n)$. Выдвинем нулевую гипотезу $H_0 : F_\xi = F_0$, $H_1 : F_\xi \neq F_0$. Рассмотрим другой критерий согласия — критерий Колмогорова. Дополнительно наложим ограничение: функция $F_0(x)$ непрерывна

на \mathbb{R} . Рассмотрим статистику Колмогорова:

$$D_n(X_{[n]}) = \sup_{x \in R} |F_n^*(x, X_{[n]}) - F_0(x)|. \quad (2.1)$$

Если верна гипотеза H_0 , то $D_n(X_{[n]}) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0$. Если верна гипотеза H_1 , т. е. $F_\xi \equiv G \neq F_0$, тогда

$$D_n(X_{[n]}) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \sup_{x \in R} |G(x) - F_0(x)| > 0.$$

Можно показать, что при условии справедливости гипотезы H_0 распределение статистики $D_n(X_{[n]})$ не зависит от конкретного вида F_0 .

Лемма 2.1. Если гипотеза H_0 верна, и $F_0(x)$ — непрерывная функция на \mathbb{R} , тогда распределение статистики

$$D_n = \sup_{x \in R} |F_n^*(x; X_{[n]}) - F_0(x)|$$

не зависит от закона распределения генеральной совокупности.

Доказательство. Предположим дополнительно, что существует (α, β) : $\alpha \geq -\infty$, $\beta \leq +\infty$, что $F_0(x)$ строго монотонна на (α, β) , при этом $F_0(\alpha) = 0$, $F_0(\beta) = 1$. Подставим в качестве аргумента в $F_0(x)$ случайную величину ξ . Заметим, что $\xi \overset{\text{п.н.}}{\in} (\alpha, \beta)$ и $X_i \overset{\text{п.н.}}{\in} (\alpha, \beta)$ для всех i . Тогда из строгой монотонности F_0 следуют равенства:

$$P\{F_0(\xi) \leq y\} = P\{\xi \leq F_0^{-1}(y)\} = F_0(F_0^{-1}(y)) = y.$$

Таким образом, получили соотношение: $P\{F_0(\xi) \leq y\} = y$, если $y \in [0, 1]$, то есть $F_0(\xi)$ подчиняется равномерному распределению. Обозначим $F_0(X_i)$ через Y_i , $i = 1, \dots, n$, тогда выборка $Y_{[n]} = (Y_1, \dots, Y_n)$ — выборка из равномерного распределения на отрезке $[0, 1]$. Очевидно, что

$$\begin{aligned} F_n^*(x, X_{[n]}) &= \frac{1}{n} \sum_{j=1}^n I\{X_j \leq x\} = \\ &= \frac{1}{n} \sum_{j=1}^n I\{F_0(X_j) \leq F_0(x)\} = \frac{1}{n} \sum_{j=1}^n I\{Y_j \leq y\} = F_n^*(y, Y_{[n]}), \end{aligned}$$

где $F_n^*(y, Y_{[n]})$ — эмпирическая функция распределения, построенная по выборке из равномерного распределения, $y = F_0(x)$.

Тогда имеет место равенство:

$$\sup_{x \in \mathbb{R}} |F_n^*(x, X_{[n]}) - F_0(x)| = \sup_{y \in [0,1]} |F_n^*(y, Y_{[n]}) - y|.$$

Отсюда следует равенство:

$$\begin{aligned} P\{\sup_{x \in \mathbb{R}} |F_n^*(x, X_{[n]}) - F_0(x)| \leq z\} = \\ = P\{\sup_{y \in [0,1]} |F_n^*(y, Y_{[n]}) - y| \leq z\}. \end{aligned}$$

Последнюю вероятность можно вычислять для различных значений z как многомерный интеграл от плотности, тождественно равной единице внутри гиперкуба $[0, 1]^n$. Интегрирование проводится по множеству точек в $[0, 1]^n$, удовлетворяющих неравенству, содержащемуся под знаком вероятности. Таким образом, для рассматриваемого случая можно построить точные критические области с заданным уровнем значимости независимо от конкретного вида $F_0(x)$.

Можно найти квантиль $z_{1-\alpha}$ для некоторого α . Рассмотрим критическую область $(z_{1-\alpha}, 1]$, если статистика (2.1) попадает в данную область, тогда отвергаем H_0 и принимаем H_1 . Если $D_n \in [0, z_{1-\alpha})$, тогда принимаем гипотезу H_0 . ■

Замечание 2.1. Доказательство леммы 2.1 можно распространить на случай, когда от $F_0(x)$ не требуется строгой монотонности, для доказательства необходимо рассмотреть обобщенную обратную функцию:

$$F_0^{-1}(y) = \sup\{x : F_0(x) \leq y\}.$$

Таким образом, возможно построение точного критерия согласия для фиксированного n . Однако, при больших n возникают серьезные вычислительные трудности. Справедлива следующая теорема А. Н. Колмогорова.

Теорема 2.1. Если гипотеза H_0 верна и $F_0(x)$ — непрерывная функция на \mathbb{R} , тогда имеет место сходимость:

$$P\{\sqrt{n}D_n(X_{[n]}) \leq z\} \xrightarrow{n \rightarrow \infty} K(z) = 1 + 2 \sum_{m=1}^{\infty} (-1)^m e^{-2m^2 z^2}.$$

Находим константу $d_{1-\alpha}$ как решение уравнения:

$$K(d_{1-\alpha}) = 1 - \alpha.$$

Правило проверки гипотез будет следующим. Если $\sqrt{n}D_n(X_{[n]}) \in (d_{1-\alpha}, \infty)$, тогда гипотеза H_0 отвергается; если $\sqrt{n}D_n(X_{[n]}) \notin (d_{1-\alpha}, \infty)$, тогда гипотеза H_0 принимается. Для практической проверки гипотез согласия по критерию Колмогорова можно воспользоваться таблицами математической статистики.

Статистику $D_n(X_{[n]})$ можно вычислить с помощью простого вычислительного алгоритма:

$$D_n(X_{[n]}) = \max_{1 \leq i \leq n} \left[\frac{i}{n} - F_0(X_{(i)}), F_0(X_{(i)}) - \frac{i-1}{n} \right],$$

где $X_{(1)} < \dots < X_{(n)}$ — вариационный ряд, построенный по выборке $X_{[n]}$.

Замечание 2.2. Критерий $\bar{\omega}^2$.

Пусть задана генеральная совокупность ξ с функцией распределения F_ξ и выборка $X_{[n]} = (X_1, \dots, X_n)$ из этой генеральной совокупности. Выдвинем нулевую гипотезу $H_0 : F_\xi = F_0$, при конкурирующей гипотезе $H_1 : F_\xi \neq F_0$. Статистика критерия имеет вид:

$$\bar{\omega}_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left\{ F_0(X_{(i)}) - \frac{2i-1}{2n} \right\}^2,$$

где $X_{(1)} < \dots < X_{(n)}$ — вариационный ряд, построенный по выборке $X_{[n]}$.

При справедливости гипотезы H_0 и непрерывности функции F_0 распределение статистики омега-квадрат зависит только от n и не зависит от F_0 . При малых n имеются таблицы

критических точек, а для больших значений n следует использовать предельное (при $n \rightarrow \infty$) распределение статистики $\bar{\omega}_n^2$. Для него составлены подробные таблицы и вычислительные программы. Важное с теоретической точки зрения свойство критериев, основанных на D_n и $\bar{\omega}_n^2$: они состоятельны против любой альтернативной гипотезы $F_\xi \neq F_0$.

Статистический критерий для проверки гипотезы H_0 называют состоятельным против альтернативной гипотезы H_1 , если вероятность отвергнуть H_0 , когда на самом деле верна H_1 , стремится к 1 при неограниченном увеличении объема наблюдений.

ТОЧЕЧНЫЕ ОЦЕНКИ

§ 1. СВОЙСТВА ТОЧЕЧНЫХ ОЦЕНОК

Рассмотрим выборку $X_{[n]} = (X_1, \dots, X_n)$, генеральную совокупность ξ и ее функцию распределения $F_\xi(x, \theta)$, где $\theta = (\theta_1, \dots, \theta_m)$ — неизвестные параметры в распределении случайной величины ξ . По имеющейся выборке можно построить оценку для этих параметров.

Определение 1.1. Пусть $\theta \in \Theta \subset \mathbb{R}^m$. Под *оценкой* понимается статистика $\hat{\theta}(X_{[n]})$ такая, что получившееся значение можно рассматривать как точечную оценку параметра θ ($\hat{\theta}(X_{[n]}) \sim \theta$).

Невозможно найти численное значение вероятности $P \left\{ \left| \hat{\theta}(X_{[n]}) - \theta \right| > \varepsilon \right\}$ для произвольного ε , так как вероятность содержит неизвестный параметр θ . Тогда какую оценку считать «хорошей»?

Свойства точечных оценок

1. Несмещенность.

Определение 1.2. Пусть параметр $\theta \in \Theta \subset \mathbb{R}$. Говорят, что оценка $\hat{\theta}(X_{[n]})$ является несмещенной оценкой параметра θ , если

$$E\hat{\theta}(X_{[n]}) = \theta \quad (1.1)$$

для любого $\theta \in \Theta$.

Определение 1.3. Говорят, что оценка $\hat{\theta}(X_{[n]})$ является асимптотически несмещенной оценкой параметра θ , если

$$E\hat{\theta}(X_{[n]}) \xrightarrow{n \rightarrow \infty} \theta \quad (1.2)$$

для любого $\theta \in \Theta$.

Замечание 1.1. Свойство несмещенности позволяет агрегировать информацию, накопленную в различных научных центрах. Рассмотрим следующий пример. Пусть $\hat{\theta}_1$ — несмещенная оценка параметра θ , полученная в некотором научном центре, $\hat{\theta}_2$ — несмещенная оценка того же параметра, полученная в другом научном центре. Предполагая, что техническая оснащенность научных центров одинаковая, будем считать, что дисперсии оценок одинаковы:

$$D(\hat{\theta}_i) = E(\hat{\theta}_i - \theta)^2 = \sigma^2(\theta),$$

$$E(\hat{\theta}_i) = \theta, \quad i = 1, 2.$$

Рассмотрим новую оценку:

$$\hat{\theta} = \frac{\hat{\theta}_1 + \hat{\theta}_2}{2},$$

$$E\hat{\theta} = \frac{E\hat{\theta}_1 + E\hat{\theta}_2}{2} = \theta,$$

тогда имеют место равенства:

$$D\hat{\theta} = E(\hat{\theta} - \theta)^2 = \frac{1}{4}E\{(\hat{\theta}_1 - \theta) + (\hat{\theta}_2 - \theta)\}^2 = \frac{\sigma^2(\theta)}{2}.$$

Как видим, агрегированная оценка оказывается более точной, дисперсия уменьшилась в два раза.

Пример 1.1. Выборочное среднее является несмещенной оценкой для математического ожидания:

$$E\bar{X} = E\left\{\frac{1}{n}\sum_{k=1}^n X_k\right\} = \frac{1}{n}\sum_{k=1}^n EX_k = E\xi = a_1.$$

Рассмотрим выборочную дисперсию, проверим, выполнено ли свойство несмещенности:

$$\begin{aligned} Es^2 &= E \left\{ \frac{1}{n} \sum_{k=1}^n \left(X_k - a_1 - \frac{1}{n} \sum_{i=1}^n (X_i - a_1) \right) \right\}^2 = \\ &= E \left\{ \frac{1}{n} \sum_{k=1}^n (X_k - a_1)^2 - \left(\frac{1}{n} \sum_{k=1}^n (X_k - a_1) \right)^2 \right\} = \\ &= \frac{1}{n} \sum_{k=1}^n E(X_k - a_1)^2 - \frac{1}{n^2} \sum_{k=1}^n \sigma^2 = \sigma^2 - \frac{1}{n} \sigma^2 = \frac{n-1}{n} \sigma^2, \end{aligned}$$

при выводе формулы учитывалось следующее соотношение:

$$E\{(X_k - a_1)(X_l - a_1)\} = \sigma^2 \delta_{kl} = \begin{cases} \sigma^2, & k = l; \\ 0, & k \neq l. \end{cases}$$

Таким образом, получаем равенство:

$$Es^2 = \frac{n-1}{n} \sigma^2,$$

следовательно, s^2 — смещенная оценка, однако она является асимптотически несмещенной оценкой: $Es^2 \xrightarrow[n \rightarrow \infty]{} \sigma^2$.

Рассмотрим исправленную оценку дисперсии:

$$\tilde{s}^2 = \frac{n}{n-1} s^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2,$$

как легко видеть, \tilde{s}^2 — несмещенная оценка дисперсии.

2. Состоятельность.

Определение 1.4. Пусть параметр $\theta \in \Theta \subset \mathbb{R}$. Говорят, что оценка $\hat{\theta}(X_{[n]})$ состоятельна, если

$$\hat{\theta}(X_{[n]}) \xrightarrow[n \rightarrow \infty]{P} \theta \tag{1.3}$$

для любого $\theta \in \Theta$.

Определение 1.5. Оценка $\hat{\theta}(X_{[n]})$ называется сильно состоятельной оценкой параметра θ , если

$$\hat{\theta}(X_{[n]}) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \theta \quad (1.4)$$

для любого $\theta \in \Theta$.

Замечание 1.2. В случае, когда $\hat{\theta}(X_{[n]})$ — векторная оценка, свойство состоятельности и сильной состоятельности рассматриваются покомпонентно.

Замечание 1.3. В определении оценки предполагалось, что n фиксировано, но обычно под оценкой понимают некоторое правило, по которому можно построить оценку для любого n .

Замечание 1.4. Пусть существует $E\xi^k$, тогда a_k^* — статистика первого типа, где a_k^* — эмпирический момент порядка k . Тогда $a_k^* \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a_k = E\xi^k$, т. е. a_k^* является сильно состоятельной оценкой.

Пусть существует $E(\xi - E\xi)^k$, тогда

$$a_k^{0*} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$$

— сильно состоятельная оценка для теоретического момента:

$$a_k^{0*} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a_k^0 = E(\xi - E\xi)^k.$$

Пример 1.2. Выборочное среднее и выборочная дисперсия представляют собой сильно состоятельные оценки соответствующих числовых характеристик случайной величины при условии, что они существуют и конечны.

$$a_1^* = \bar{X} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a_1 = E\xi,$$

$$s^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X})^2 = a_2^{0*} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} a_2^0 = D\xi.$$

3. Эффективность.

Пусть в распределении генеральной совокупности имеет-ся неизвестный параметр $\theta \in \Theta \subset \mathbb{R}$. Рассмотрим некоторый класс оценок $K = \{\hat{\theta}(X_{[n]})\}$ параметра θ .

Определение 1.6. Говорят, что оценка $\theta^*(X_{[n]}) \in K$ является эффективной оценкой параметра θ в классе K , если для любой другой оценки $\hat{\theta} \in K$ имеет место неравенство:

$$E(\theta^* - \theta)^2 \leq E(\hat{\theta} - \theta)^2 \quad (1.5)$$

для любого $\theta \in \Theta$.

Класс несмещенных оценок обозначим через

$$K_0 = \left\{ \hat{\theta}(X_{[n]}) : E\hat{\theta} = \theta, \forall \theta \in \Theta \right\}.$$

Оценка, эффективная в классе K_0 , называется эффективной оценкой.

Рассмотрим случай, когда $m > 1$, то есть $\theta = (\theta_1, \dots, \theta_m)$. Для любого $y \in \mathbb{R}^m$ определим $\alpha_y = (\theta, y) = \theta_1 y_1 + \dots + \theta_m y_m$. Тогда $\alpha_y^* = (\theta^*, y)$ — оценка параметра α_y .

Определение 1.7. Будем говорить, что оценка $\theta^* \in K$ является эффективной оценкой параметра $\theta = (\theta_1, \dots, \theta_m)$ в классе K , если для любой другой оценки $\hat{\theta} \in K$ и любого $y \in \mathbb{R}^m$ при любом допустимом значении $\theta \in \Theta$ имеет место неравенство:

$$E(\alpha_y^* - \alpha_y)^2 \leq E(\hat{\alpha}_y - \alpha_y)^2, \quad (1.6)$$

где $\hat{\alpha}_y = (\hat{\theta}, y)$.

Теорема 1.1. Пусть несмещенные оценки $\hat{\theta}_1$ и $\hat{\theta}_2$ параметра $\theta \in \Theta \subset \mathbb{R}$ являются эффективными, тогда оценки $\hat{\theta}_1$ и $\hat{\theta}_2$ почти наверное совпадают.

Доказательство. Рассмотрим $\tilde{\theta} = (\hat{\theta}_1 + \hat{\theta}_2)/2$. Нетрудно показать несмещенность данной оценки. Справедливо тождество:

$$\begin{aligned} \frac{(\hat{\theta}_1 - \theta)^2}{2} + \frac{(\hat{\theta}_2 - \theta)^2}{2} &= \left(\frac{(\hat{\theta}_1 - \theta)}{2} - \frac{(\hat{\theta}_2 - \theta)}{2} \right)^2 + \\ &+ \left(\frac{(\hat{\theta}_1 - \theta)}{2} + \frac{(\hat{\theta}_2 - \theta)}{2} \right)^2. \end{aligned}$$

Из условия теоремы следует, что для любого $\theta \in \Theta$ имеет место равенство:

$$D\hat{\theta}_1 = D\hat{\theta}_2 = d^2(\theta).$$

Кроме того, для любого $\tilde{\theta} \in K_0$ должно выполняться неравенство: $D\tilde{\theta} \geq d^2(\theta)$. Найдем для тождества математическое ожидание:

$$d^2(\theta) = \frac{1}{4}E(\hat{\theta}_1 - \hat{\theta}_2)^2 + D\tilde{\theta}.$$

Следовательно, $E(\hat{\theta}_1 - \hat{\theta}_2)^2 = 0$ (так как если это неверно, то $D\tilde{\theta} < d^2(\theta)$, чего быть не может), но тогда $\hat{\theta}_1 \stackrel{\text{п.н.}}{=} \hat{\theta}_2$. ■

Замечание 1.5. Утверждение теоремы переносится на многомерный случай.

Рассмотрим случай, когда θ вектор. Пусть $\theta = (\theta_1, \dots, \theta_k)^T \in \Theta \subset \mathbb{R}^k$, $\hat{\theta} \in K_0$, то есть $E\hat{\theta} = \theta$ для любого $\theta \in \Theta$. Возьмем любой вектор $y \in \mathbb{R}^k$, рассмотрим скалярное произведение:

$$(\hat{\theta}, y) = \sum_{i=1}^k \hat{\theta}_i(X_{[n]})y_i,$$

как оценку для скалярного произведения (θ, y) . Оценка $(\hat{\theta}, y)$ будет несмещенной оценкой:

$$E(\hat{\theta}, y) = (\theta, y)$$

для любого $\theta \in \Theta$ и $y \in \mathbb{R}^k$.

Оценка $\hat{\theta}$ эффективна, если $D(\hat{\theta}, y) \leq D(\tilde{\theta}, y)$ для любой оценки $\tilde{\theta} \in K_0$ и любого $\theta \in \Theta$, и $y \in \mathbb{R}^k$. Вычислим левую и правую части неравенства, получим:

$$D(\hat{\theta}, y) = y^T D\hat{\theta}y,$$

где $D\hat{\theta}$ — ковариационная матрица вектора $\hat{\theta}$,

$$D(\tilde{\theta}, y) = y^T D\tilde{\theta}y,$$

где $D\tilde{\theta}$ — ковариационная матрица вектора $\tilde{\theta}$. Следовательно, справедливо неравенство:

$$y^T (D\tilde{\theta} - D\hat{\theta})y \geq 0.$$

Тогда, так как y — любое, получаем, что матрица коэффициентов квадратичной формы является неотрицательно определенной матрицей, т. е. $D\tilde{\theta} - D\hat{\theta} \succeq 0$. В результате приходим к определению 1.8, которое эквивалентно определению 1.7 для класса несмещенных оценок.

Определение 1.8. Оценка $\hat{\theta}$ эффективна в классе K_0 , или просто эффективна, если $D\tilde{\theta} - D\hat{\theta} \succeq 0$ (неотрицательно определенная матрица), где $\tilde{\theta} \in K_0$ для любого $\theta \in \Theta \subset \mathbb{R}^k$.

4. Асимптотическая нормальность.

Определение 1.9. Пусть оценивается параметр $\theta \in \Theta \subset \mathbb{R}$. Оценка $\hat{\theta}$ называется асимптотически нормальной оценкой параметра θ с коэффициентом рассеивания $\sigma^2(\theta)$, если

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n \rightarrow \infty]{d} \zeta \sim N(0, \sigma^2(\theta)). \quad (1.7)$$

Из этого определения следует, что для любого $x \in \mathbb{R}$ имеет место сходимость:

$$P \left\{ \sqrt{n}(\hat{\theta} - \theta) \leq x \right\} \xrightarrow[n \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}\sigma(\theta)} \int_{-\infty}^x e^{-\frac{y^2}{2\sigma^2(\theta)}} dy.$$

Определение 1.10. Пусть оценивается параметр $\theta \in \Theta \subset \mathbb{R}^m$. Оценка $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_m)$ называется асимптотически нормальной с матрицей рассеивания $\Sigma(\theta)$, если имеет место сходимость по распределению:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n \rightarrow \infty]{d} \eta \sim N(0, \Sigma(\theta)).$$

5. Асимптотическая эффективность.

Определение 1.11. Оценка $\hat{\theta}$ называется асимптотически эффективной в классе K оценок параметра $\theta \in \Theta \subset \mathbb{R}$, если

$$\lim_{n \rightarrow \infty} \frac{E(\hat{\theta} - \theta)^2}{E(\tilde{\theta} - \theta)^2} \leq 1$$

для любого параметра $\theta \in \Theta \subset \mathbb{R}$ и любой оценки $\tilde{\theta} \in K$.

Статистическая оценка считается «хорошей», если она обладает хотя бы некоторыми из свойств 1–5.

§ 2. МЕТОДЫ ПОСТРОЕНИЯ ТОЧЕЧНЫХ ОЦЕНОК

Рассмотрим сначала *метод моментов*. Пусть требуется оценить параметр $\theta \in \Theta \subset \mathbb{R}$ по имеющейся выборке $X_{[n]} = (X_1, \dots, X_n)$. Рассмотрим борелевскую функцию $g(x) : \mathbb{R} \rightarrow \mathbb{R}$ и определим функцию $m(\theta) = \int_{-\infty}^{\infty} g(x) dF_{\xi}(x; \theta)$.

Далее положим, что

$$\int_{-\infty}^{\infty} g(x) dF_n^*(x) = \frac{1}{n} \sum_{i=1}^n g(X_i) = \bar{g}. \quad (2.1)$$

Составим уравнение

$$m(\theta) = \bar{g} = \frac{1}{n} \sum_{i=1}^n g(X_i). \quad (2.2)$$

Предположим, что уравнение (2.2) имеет единственное решение $\hat{\theta}(X_{[n]})$, тогда будем это решение называть оценкой $\hat{\theta}$ неизвестного параметра θ , полученной по методу моментов:

$$\hat{\theta}(X_{[n]}) = m^{-1} \left(\frac{1}{n} \sum_{i=1}^n g(X_i) \right).$$

Введем следующее обозначение: $h(\cdot) = m^{-1}(\cdot)$, оценка по методу моментов при некоторых очевидных условиях является статистикой первого типа. По теореме о предельном поведении статистики первого типа получаем сходимость:

$$\hat{\theta}(X_{[n]}) \xrightarrow[n \rightarrow \infty]{\text{п.н.}} \theta.$$

Свойства оценок, построенных по методу моментов:

1. Если функция $m^{-1}(y)$ непрерывна на всей области определения, то оценка по методу моментов сильно состоятельна.
2. Если $m'(\theta) \neq 0$ для всех $\theta \in \Theta$, тогда оценка по методу моментов асимптотически нормальна с коэффициентом рассеяния $\frac{Dg(\xi)}{(m'(\theta))^2}$, где θ — истинное значение параметра.

Метод моментов легко обобщить на многомерный случай, при этом $g(x) = (g_1(x), \dots, g_k(x))$, где k — число неизвестных параметров, то есть $\theta = (\theta_1, \dots, \theta_k)^T \in \Theta \subset \mathbb{R}^k$. Название «метод моментов» связано с тем, что обычно в качестве функций $g_i(x)$ берутся степенные функции: $g_i(x) = x^i$.

Пример 2.1. Пусть $\xi \sim N(a, \sigma^2)$, тогда $\theta = (a, \sigma^2)^T \in \Theta = \mathbb{R} \times \mathbb{R}_+$. Выберем $g(x) = (x, x^2)$, тогда

$$Eg(\xi) = \begin{pmatrix} E\xi \\ E\xi^2 \end{pmatrix} = \begin{pmatrix} a \\ \sigma^2 + a^2 \end{pmatrix},$$

так как $\sigma^2 = D\xi = E\xi^2 - (E\xi)^2 = E\xi^2 - a^2$. Нетрудно показать, что

$$\bar{g} = \begin{pmatrix} \frac{1}{n} \sum_{k=1}^n X_k \\ \frac{1}{n} \sum_{k=1}^n X_k^2 \end{pmatrix} = \begin{pmatrix} \bar{X} \\ s^2 + \bar{X}^2 \end{pmatrix},$$

так как $s^2 = \frac{1}{n} \sum_{k=1}^n X_k^2 - (\bar{X})^2$. Таким образом, получили систему:

$$\begin{cases} a = \bar{X}; \\ \sigma^2 + a^2 = s^2 + \bar{X}^2. \end{cases}$$

Следовательно, оценки по методу моментов имеют следующий вид:

$$\begin{cases} \hat{a} = \bar{X}; \\ \hat{\sigma}^2 = s^2. \end{cases}$$

Пример 2.2. Рассмотрим равномерно распределенную случайную величину ξ с плотностью распределения:

$$f_{\xi}(x) = \begin{cases} \frac{1}{\theta}, & x \in [0, \theta]; \\ 0, & x \notin [0, \theta]. \end{cases}$$

Так как неизвестный параметр один, то $g(x) = x$. Вычислим математическое ожидание:

$$Eg(\xi) = E\xi = \int_0^{\theta} x \frac{1}{\theta} dx = \frac{1}{2\theta} x^2 = \frac{\theta}{2}.$$

Уравнение имеет вид:

$$\frac{\theta}{2} = \bar{X},$$

откуда получаем оценку:

$$\hat{\theta} = 2\bar{X} = \frac{2}{n} \sum_{k=1}^n X_i.$$

Может оказаться, что $\frac{1}{n} \sum_{k=1}^n X_i > \theta/2$, тогда $\hat{\theta} > \theta$. Данный метод может дать сильно завышенную оценку.

Рассмотрим теперь *метод максимального правдоподобия* построения точечных оценок. Пусть задана генеральная совокупность ξ с функцией распределения F_{ξ} и плотностью распределения f_{ξ} (будем предполагать, что плотность распределения существует). Задана выборка $X_{[n]} = (X_1, \dots, X_n)$. Совместная плотность распределения выборки имеет вид:

$$f_{X_{[n]}}(x_1, \dots, x_n) = \prod_{i=1}^n f_{\xi}(x_i). \quad (2.3)$$

В плотности распределения выборки существует неизвестный параметр θ , поэтому ниже будем рассматривать совместную плотность распределения в виде:

$$f_{X_{[n]}}(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f_{\xi}(x_i, \theta),$$

подставив вместо x_1, \dots, x_n выборку $X_{[n]}$, получим:

$$f_{X_{[n]}}(X_{[n]} | \theta) = \prod_{i=1}^n f_{\xi}(X_i, \theta),$$

где $\theta = (\theta_1, \dots, \theta_m) \in \Theta$.

Определение 2.1. Если генеральная совокупность имеет плотность распределения f_{ξ} , то функцией правдоподобия выборки $X_{[n]}$ будем называть функцию

$$L(X_{[n]}, \theta) = \prod_{i=1}^n f_{\xi}(X_i, \theta).$$

Определение 2.2. Если генеральная совокупность ξ — дискретная случайная величина с возможными значениями $\{z_i\}$ и соответствующими вероятностями $p_{\xi}(z_i, \theta)$, то функцией правдоподобия выборки $X_{[n]}$ будем называть функцию

$$L(X_{[n]}, \theta) = \prod_{i=1}^n p_{\xi}(X_i, \theta).$$

Будем считать функцию правдоподобия функцией неизвестного параметра θ . Для нахождения оценки параметра θ решаем задачу:

$$\max_{\theta \in \Theta} L(X_{[n]}, \theta).$$

Определение 2.3. Оценкой максимального правдоподобия параметра θ называется оценка

$$\hat{\theta}(X_{[n]}) = \arg \max_{\theta \in \Theta} L(X_{[n]}, \theta), \quad (2.4)$$

если решение задачи максимизации существует и единственно.

Свойства оценок максимального правдоподобия:

1. Предположим, что существует взаимно однозначное соответствие $\beta : \Theta \leftrightarrow B$, пусть

$$\hat{b}(X_{[n]}) = \arg \max_{b \in B} L(X_{[n]}, \beta^{-1}(b)). \quad (2.5)$$

Если решение (2.4) существует и единственно, то существует и единственно решение (2.5), причем имеет место равенство:

$$\hat{\theta} = \beta^{-1}(\hat{b}).$$

2. Если функция правдоподобия непрерывно дифференцируема, и выполнены некоторые условия гладкости, то можно доказать, что оценки метода максимального правдоподобия — сильно состоятельны, асимптотически эффективны и асимптотически нормальны [5, 20, 30].

Замечание 2.1. Часто вместо функции $L(X_{[n]}, \theta)$ рассматривают функцию $\ln L(X_{[n]}, \theta)$, поскольку функция $\ln(t)$ является строго возрастающей функцией своего аргумента t и данный переход правомерен.

Пример 2.3. Рассмотрим случайную величину $\xi \sim N(a, \sigma^2)$ с плотностью распределения

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Функция правдоподобия имеет вид:

$$L(X_{[n]}, a, \sigma^2) = \prod_{i=1}^n f_{\xi}(X_i, a, \sigma^2) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{\sum_{i=1}^n (X_i - a)^2}{2\sigma^2}}.$$

Тогда

$$\ln L = \ln \frac{1}{((2\pi)^{\frac{1}{2}} \sigma)^n} - \frac{\sum_{i=1}^n (X_i - a)^2}{2\sigma^2},$$

продифференцируем по a : $\partial \ln L / \partial a = 0$, или $\sum_{i=1}^n X_i - an = 0$, откуда $\hat{a} = \bar{X}$.

Продифференцируем по σ :

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (X_i - a)^2}{\sigma^3} = 0,$$

$$n\sigma^2 = \sum_{i=1}^n (x_i - a)^2,$$

откуда находим решение:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - a)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = s^2.$$

Нетрудно проверить, что \bar{X} и s^2 доставляют максимум функции правдоподобия.

Пример 2.4. Пусть случайная величина ξ подчиняется равномерному распределению с плотностью:

$$f(x, \theta) = \begin{cases} \frac{1}{\theta}, & x \in [0, \theta]; \\ 0, & x \notin [0, \theta]. \end{cases}$$

Запишем функцию правдоподобия:

$$L(X_{[n]}, \theta) = \prod_{i=1}^n f(X_i, \theta) = \begin{cases} \frac{1}{\theta^n}, & \text{если для } \forall i : X_i \in [0, \theta]; \\ 0, & \text{если } \exists i : X_i \notin [0, \theta]. \end{cases}$$

Построим вариационный ряд $X_{(1)} \leq \dots \leq X_{(n)}$. Таким образом, получаем:

$$L(X_{[n]}, \theta) = \begin{cases} \frac{1}{\theta^n}, & X_{(n)} \in [0, \theta]; \\ 0, & \exists k : X_{(k)} \notin [0, \theta]. \end{cases}$$

Очевидно, что оценка максимального правдоподобия $\hat{\theta}(X_{[n]}) = X_{(n)}$.

§ 3. ДОСТАТОЧНЫЕ СТАТИСТИКИ

Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ . Неизвестные параметры составляют вектор $\theta = (\theta_1, \dots, \theta_m) \in \Theta \subset \mathbb{R}^m$. Функция распределения генеральной совокупности ξ имеет вид $F_\xi(\cdot/\theta)$.

Будем рассматривать некоторую статистику

$$S(X_{[n]}) = (S_1(X_{[n]}), \dots, S_k(X_{[n]})) .$$

Обозначим через $P(B/s)$ условное распределение выборки $X_{[n]}$ при условии, что $S(X_{[n]}) = s$.

Распределение $P(B/s)$, где $B \in \mathcal{B}(\mathbb{R}^n)$, называется *условным распределением* выборки $X_{[n]}$ при условии, что $S(X_{[n]}) = s$, если выполнены условия:

1. При любом фиксированном значении статистики s распределение $P(\cdot/s)$ является вероятностной мерой в пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.
2. При любом фиксированном B функция $P(B/\cdot)$ как функция второго аргумента является борелевской и представляет собой условную вероятность $P\{X_{[n]} \in B/S = s\}$.

Определение 3.1. Статистика $S(X_{[n]})$ называется достаточной, если условное распределение выборки $X_{[n]}$ при условии, что статистика $S(X_{[n]})$ принимает значение s , не зависит от неизвестного параметра θ .

Если статистика является достаточной статистикой, то можно говорить о том, что вся информация о параметре θ содержится в значении статистики $S(X_{[n]})$, и оставшийся «разброс» элементов выборки уже не зависит от θ .

Пример 3.1. Пусть генеральная совокупность ξ подчиняется распределению Пуассона: $\xi \sim \frac{\lambda^k}{k!} e^{-\lambda}$, $\lambda > 0$, $k \in \mathbb{Z}_+$, $X_i \in \mathbb{Z}_+$. Тогда достаточная статистика имеет вид:

$$S(X_{[n]}) = \sum_{i=1}^n X_i.$$

Действительно, пусть $x = (x_1, \dots, x_n)$, $x_k \in \mathbb{Z}_+$, тогда

$$\begin{aligned} P\{X_{[n]} = x / S(X_{[n]}) = S(x) = s\} &= \\ = \frac{P\{X_{[n]} = x, S(X_{[n]}) = S(x) = s\}}{P\{S(X_{[n]}) = S(x) = s\}} &= \frac{P\{X_{[n]} = x\}}{P\{S(X_{[n]}) = S(x) = s\}}, \end{aligned}$$

предполагается, что события совместны. Полученное представление всегда верно для любых дискретных совокупностей. Рассмотрим числитель:

$$\begin{aligned} P\{X_{[n]} = x\} &= P\{X_1 = x_1, \dots, X_n = x_n\} = \\ &= \frac{\lambda^{\sum x_i}}{x_1! \dots x_n!} e^{-\lambda n} = \frac{\lambda^s}{x_1! \dots x_n!} e^{-\lambda n}. \end{aligned}$$

Рассмотрим знаменатель:

$$\begin{aligned} P\{S(X_{[n]}) = S(x) = s\} &= \sum_{y: S(y)=s} P\{X_{[n]} = y\} = \\ &= \frac{\lambda^s n^s}{s!} e^{-\lambda n} \sum_{y: y_1 + \dots + y_n = s} \frac{s!}{y_1! \dots y_n!} \cdot \frac{1}{n^s} = \frac{\lambda^s n^s}{s!} e^{-\lambda n}. \end{aligned}$$

Тогда условная вероятность имеет вид:

$$\frac{P\{X_{[n]} = x\}}{P\{S(X_{[n]}) = S(x) = s\}} = \frac{s!}{x_1! \dots x_n!} \frac{1}{n^s}.$$

И, как легко видеть, представляет собой полиномиальное распределение, в котором n исходов, вероятность каждого исхода равна $1/n$ (исходы равновероятны), s — количество испытаний, x_1 — количество исходов с номером 1, ..., x_n — количество исходов с номером n . Вероятность того, что в последовательности таких s испытаний произойдет x_1 исходов с номером 1, ..., x_n исходов с номером n , такова:

$$P\{x_1, \dots, x_n\} = \frac{s!}{x_1! \dots x_n!} \frac{1}{n^s}.$$

Замечание 3.1. Если статистика $S(X_{[n]})$ является достаточной, то взаимно однозначная функция от $S(X_{[n]})$ также достаточная статистика. Следовательно, возвращаясь к примеру, получаем:

$$\tilde{S}(X_{[n]}) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X},$$

также является достаточной статистикой в примере 3.1.

§ 4. ТЕОРЕМА НЕЙМАНА–ФИШЕРА. ТЕОРЕМА КОЛМОГОРОВА

Теорема 4.1 (теорема Неймана–Фишера). Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ с функцией распределения F_ξ . Пусть $\theta \in \Theta$ — вектор неизвестных параметров распределения. Для того, чтобы статистика $S(X_{[n]})$ была достаточной необходимо и достаточно, чтобы функция правдоподобия $L(X_{[n]}, \theta)$ была представима в виде:

$$L(X_{[n]}, \theta) = G(S(X_{[n]}), \theta)H(X_{[n]}),$$

где G, H — измеримые функции, причем функция G зависит от выборки только через значения статистики $S(X_{[n]})$, а функция H не зависит от неизвестного параметра θ .

Доказательство. Теорему Неймана–Фишера докажем только для двух случаев. Рассмотрим дискретный случай. Докажем достаточность. Найдем условную вероятность:

$$\begin{aligned} P\{X_{[n]} = x / S(X_{[n]}) = s\} &= \\ &= \frac{P\{X_{[n]} = x\}}{\sum_{y: S(y)=s} P\{X_{[n]} = y\}} = \frac{P_\xi(x_1) \dots P_\xi(x_n)}{\sum_{y: S(y)=s} P_\xi(y_1) \dots P_\xi(y_n)} = \\ &= \frac{L(x; \theta)}{\sum_{y: S(y)=s} L(y, \theta)} = \frac{G(s, \theta)H(x)}{\sum_{y: S(y)=s} G(s, \theta)H(y)} = \frac{H(x)}{\sum_{y: S(y)=s} H(y)}. \end{aligned}$$

Видим, что условная вероятность не зависит от θ . Теперь докажем необходимость. Пусть условная вероятность не зависит от

неизвестных параметров. Обозначим через $H(x)$ вероятность $P\{X_{[n]} = x/S(X_{[n]}) = S(x) = s\}$, и, как легко заметить,

$$L(x, \theta) = P\{X_{[n]} = x\} = P\{S(X_{[n]}) = s\}H(x).$$

Вероятность

$$P\{S(X_{[n]}) = S(x) = s\} = G(s, \theta)$$

зависит от неизвестных параметров θ и зависит от x только через значение статистики, справедливо равенство:

$$L(x, \theta) = G(s, \theta)H(x).$$

Рассмотрим «гладкий» случай. Будем считать, что существует плотность распределения $f_{\xi}(\cdot)$, борелевские функции $S(x)$ дифференцируемы для всех $x \in \mathbb{R}^n$, $S(x) \in \mathbb{R}^k$. Будем предполагать, что мы можем дополнить статистику S статистикой $T(x) \in \mathbb{R}^{n-k}$, при этом

$$(s, t) = (S(x), T(x)),$$

где $x \in \mathbb{R}^n$. Будем считать, что x взаимно однозначно соответствует (s, t) . Пусть (s, t) дифференцируемо по x , и, наоборот, обратное отображение дифференцируемо по s и t .

Плотность выборки запишем в виде:

$$f_{X_{[n]}}(x) = L(x, \theta) = \prod_{i=1}^n f_{\xi}(x_i, \theta).$$

Рассмотрим, что происходит с плотностью при взаимно однозначном отображении:

$$f_{S,T}(s, t) = f_{X_{[n]}}(x(s, t)) \left| \det \left(\frac{\partial x}{\partial (s, t)} \right) \right|.$$

Докажем достаточность (факторизация выполнена). Пусть

$$f_{X_{[n]}}(x) = G(S(x), \theta)H(x),$$

тогда

$$f_{S,T}(s, t) = G(s, \theta) H(x(s, t)) \left| \det \left(\frac{\partial x}{\partial(s, t)} \right) \right|.$$

Следовательно,

$$f_{T/S}(t/s) = \frac{f_{S,T}(s, t)}{f_S(s)} = \frac{H(x(s, t)) \left| \det \left(\frac{\partial x}{\partial(s, t)} \right) \right|}{\int_{\mathbb{R}^{n-k}} H(x(s, t)) \left| \det \left(\frac{\partial x}{\partial(s, t)} \right) \right| dt}.$$

Как видим, $f_{T/S}(t/s)$ не зависит от θ , но тогда вероятность $P\{X_{[n]} \in B/S(X_{[n]}) = s\} = \int_{\mathbb{R}^{n-k}} I\{x(s, t) \in B\} f_{T/S}(t/s) dt$ не зависит от θ .

Необходимость следует из того факта, что $f_{T/S}(t/s)$ не зависит от θ , но тогда справедливо равенство:

$$f_{X_{[n]}}(x) = G(S(x), \theta) H(x),$$

где $H(x) = f_{T/S}(t(x)/s(x)) \left| \det \left(\frac{\partial(s, t)}{\partial x} \right) \right|$, $G(S(x), \theta) = f_S(s(x))$. ■

Пример 4.1. Предположим, что генеральная совокупность ξ подчиняется нормальному распределению: $\xi \sim N(a, \sigma^2)$, $\theta = (a, \sigma^2)^T \in \mathbb{R} \times \mathbb{R}_+$, тогда

$$\begin{aligned} L(X, \theta) &= \prod_{i=1}^n f_{\xi}(X_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(X_i - a)^2}{2\sigma^2}} = \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a)^2} = G(S(X_{[n]}), \theta), \end{aligned}$$

и функция $H(x) = 1$. В качестве достаточной статистики $S(X_{[n]})$ можно выбрать:

$$\begin{pmatrix} \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i^2 \end{pmatrix} = \begin{pmatrix} s_1(x) \\ s_2(x) \end{pmatrix} = S(x).$$

Сделаем взаимно однозначное преобразование:

$$\begin{cases} \tilde{S}_2 = a_2^* - (a_1^*)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\frac{1}{n} \sum_{i=1}^n X_i)^2, \\ \tilde{S}_1 = \frac{1}{n} \sum_{i=1}^n X_i. \end{cases}$$

Следовательно, (\bar{X}, s^2) — достаточная статистика, и (\bar{X}, \tilde{s}^2) также достаточная статистика.

Пример 4.2. Пусть случайная величина ξ подчиняется экспоненциальному распределению с параметром λ с плотностью:

$$f_{\xi}(y) = \begin{cases} \lambda e^{-\lambda y}, & y > 0; \\ 0, & y \leq 0. \end{cases}$$

Тогда

$$f_{X_{[n]}}(x) = L(x, \theta) = \begin{cases} \lambda^n e^{-\lambda \sum_{i=1}^n x_i}, & \text{если для } \forall i: x_i > 0; \\ 0, & \text{если } \exists i: x_i \leq 0. \end{cases}$$

Из теоремы 4.1 следует, что \bar{X} — достаточная статистика, здесь $H(x) = 1$.

Теорема 4.2 (теорема Колмогорова–Рао–Блекуэлла). Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ . Пусть $\theta \in \Theta \subset \mathbb{R}$ — неизвестный параметр распределения генеральной совокупности, $\hat{\theta}(X_{[n]})$ — несмещенная оценка параметра θ , $S(X_{[n]})$ — достаточная статистика. Введем обозначение

$$\hat{\theta}_S = E(\hat{\theta}/S(X_{[n]})).$$

Справедливы утверждения:

1. $\hat{\theta}_S$ — несмещенная оценка параметра θ .
2. $\hat{\theta}_S = g(S(X_{[n]}))$, т.е. $\hat{\theta}_S$ зависит от выборки $X_{[n]}$ только через достаточную статистику S .
3. $D\hat{\theta}_S \leq D\hat{\theta}$, при этом равенство возможно только, если оценки $\hat{\theta}_S$ и $\hat{\theta}$ почти наверное совпадают.

Доказательство. Заметим, что $\hat{\theta}_s = \int_{\mathbb{R}^n} \hat{\theta}(x)P(dx/S = s)$ не зависит от θ и измеримо по s . Таким образом, $\hat{\theta}_s$ действительно является оценкой. Очевидно, что $\hat{\theta}_S = E(\hat{\theta}/S(X_{[n]}))$ зависит от выборки только через статистику S .

По свойству условных математических ожиданий получаем, что $E\hat{\theta}_s = EE(\hat{\theta}/S) = E\hat{\theta} = \theta$, что говорит о несмещенности оценки $\hat{\theta}_s$.

Воспользуемся свойствами условных математических ожиданий:

$$\begin{aligned} D\hat{\theta} &= E(\hat{\theta} - \theta \pm \hat{\theta}_s)^2 = E((\hat{\theta} - \hat{\theta}_s) + (\hat{\theta}_s - \theta))^2 = \\ &= E(\hat{\theta} - \hat{\theta}_s)^2 + E(\hat{\theta}_s - \theta)^2 + 2E\{(\hat{\theta} - \hat{\theta}_s)(\hat{\theta}_s - \theta)\}. \end{aligned}$$

Рассмотрим отдельно третье слагаемое:

$$E\{(\hat{\theta} - \hat{\theta}_s)(\hat{\theta}_s - \theta)\} = E\{(\hat{\theta}_s - \theta)E[(\hat{\theta} - \hat{\theta}_s)/S]\} = 0.$$

Получаем

$$D\hat{\theta} = E(\hat{\theta} - \hat{\theta}_s)^2 + D\hat{\theta}_s,$$

откуда следует, что

$$D\hat{\theta} \geq D\hat{\theta}_s.$$

Причем $D\hat{\theta} = D\hat{\theta}_s$ равносильно тому, что $E(\hat{\theta} - \hat{\theta}_s)^2 = 0$ или $\hat{\theta}$ почти наверное совпадает с $\hat{\theta}_s$. ■

Замечание 4.1. Теорема 4.2 переносится на многомерный случай.

Рассмотрим еще одно свойство достаточной статистики.

Лемма 4.1. Если достаточная статистика существует, то оценка максимального правдоподобия неизвестного параметра является функцией от достаточной статистики.

Доказательство. Пусть $\theta \in \Theta \subset \mathbb{R}^k$. Запишем функцию правдоподобия:

$$L(X_{[n]}, \theta) = G(S(X_{[n]}), \theta)H(X_{[n]}).$$

Для оценки максимального правдоподобия справедливы равенства:

$$\begin{aligned}\hat{\theta}(X_{[n]}) &= \arg \max_{\theta \in \Theta} L(X_{[n]}, \theta) = \\ &= \arg \max_{\theta \in \Theta} G(S(X_{[n]}), \theta) = \hat{\theta}(S(X_{[n]})),\end{aligned}$$

что доказывает справедливость леммы. ■

§ 5. НЕРАВЕНСТВО РАО–КРАМЕРА

Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ с функцией распределения $F_{\xi}(x, \theta)$ и плотностью распределения $f_{\xi}(x, \theta)$, где $\theta \in \Theta \subset \mathbb{R}$ — неизвестный параметр.

Замечание 5.1. Все результаты этого параграфа можно перенести на дискретный случай.

Функция правдоподобия имеет вид:

$$L(X_{[n]}, \theta) = \prod_{i=1}^n f_{\xi}(X_i, \theta),$$

совместная плотность выборки

$$f_{X_{[n]}}(x, \theta) = L(x, \theta) = \prod_{i=1}^n f_{\xi}(x_i, \theta).$$

Выполняется равенство:

$$\int_{\mathbb{R}^n} L(x, \theta) dx = 1. \quad (5.1)$$

Пусть имеется оценка $\hat{\theta}(X_{[n]})$ неизвестного параметра θ :

$$E\hat{\theta} = \int_{\mathbb{R}^n} \hat{\theta}(x_1, \dots, x_n) L(x, \theta) dx = h(\theta). \quad (5.2)$$

Обозначим через $I_n(\theta)$ математическое ожидание:

$$I_n(\theta) = E \left(\frac{\partial \ln L(X_{(n)}, \theta)}{\partial \theta} \right)^2 = \int_{\mathbb{R}^n} \left(\frac{\partial \ln L(x, \theta)}{\partial \theta} \right)^2 L(x, \theta) dx.$$

Определение 5.1. Величина $I_n(\theta)$, если математическое ожидание существует и конечно, называется информационным количеством Фишера (соответствующим выборке объема n).

Будем предполагать, что выполнены *условия регулярности*:

- Для информационного количества Фишера выполнено неравенство $0 < I_n(\theta) < \infty$ для любого $\theta \in \Theta$.
- Равенства (5.1) и (5.2) можно продифференцировать и получить уравнения:

$$\int_{\mathbb{R}^n} \frac{\partial L(x, \theta)}{\partial \theta} dx = 0, \quad (5.3)$$

$$\int_{\mathbb{R}^n} \hat{\theta}(x_1, \dots, x_n) \frac{\partial L(x, \theta)}{\partial \theta} dx = h'(\theta). \quad (5.4)$$

- Множество $N = \{x \in \mathbb{R}^n : L(x, \theta) = 0\}$ не зависит от θ .

Теорема 5.1 (неравенство Рао–Крамера). Пусть имеет-ся генеральная совокупность ξ с функцией распределения $F_\xi(y, \theta)$, где $\theta \in \Theta \subset \mathbb{R}$. Задана выборка $X_{[n]}$ из генеральной совокупности ξ , и выполнены условия регулярности, тогда справедливо неравенство:

$$D\hat{\theta} \geq \frac{(h'(\theta))^2}{I_n(\theta)}. \quad (5.5)$$

Доказательство. Перепишем (5.3) и (5.4) следующим образом:

$$\int_{\mathbb{R}^n} \frac{\partial \ln L(x, \theta)}{\partial \theta} L(x, \theta) dx = E \left\{ \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \right\} = 0,$$

$$\begin{aligned} \int_{\mathbb{R}^n} \hat{\theta}(x) \frac{\partial \ln L(x, \theta)}{\partial \theta} L(x, \theta) dx &= \\ &= E \left\{ \hat{\theta}(X_{[n]}) \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \right\} = h'(\theta). \end{aligned}$$

Заметим, что множество $N = \{x : L(x, \theta) = 0\}$ — множество меры нуль, так как

$$P\{X_{[n]} \in N\} = \int_N L(x, \theta) dx = 0.$$

Умножим первое равенство на $E\hat{\theta}$ и вычтем из второго:

$$E \left\{ (\hat{\theta}(X_{[n]}) - E\hat{\theta}) \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \right\} = h'(\theta).$$

Сделаем обозначения:

$$\begin{aligned} \eta_1 &= (\hat{\theta}(X_{[n]}) - E\hat{\theta}), \\ \eta_2 &= \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta}. \end{aligned}$$

Воспользуемся неравенством Коши–Шварца–Буняковского:

$$(h'(\theta))^2 = (E\{\eta_1 \eta_2\})^2 \leq E\eta_1^2 E\eta_2^2 = D\hat{\theta} I_n(\theta).$$

Теорема доказана. ■

Замечание 5.2. Неравенство (5.5) выполнено как равенство тогда и только тогда, когда $\eta_1 \stackrel{\text{п.н.}}{=} A(\theta) \eta_2$ (следует из неравенства Коши–Шварца–Буняковского), т. е.

$$(\hat{\theta}(X_{[n]}) - E\hat{\theta}) \stackrel{\text{п.н.}}{=} A(\theta) \frac{\partial \ln L}{\partial \theta}.$$

Замечание 5.3. Если $E\hat{\theta} = \theta$ для любого $\theta \in \Theta$ (то есть оценка — несмещенная), то справедливо неравенство:

$$D\hat{\theta} \geq \frac{1}{I_n(\theta)}.$$

Замечание 5.4. По определению $I_n(\theta) = E \left(\frac{\partial \ln L}{\partial \theta} \right)^2$. Заметим, что

$$E \left\{ \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \right\} = 0.$$

Следовательно, имеют место равенства:

$$\begin{aligned} I_n(\theta) &= E \left(\frac{\partial \ln L}{\partial \theta} \right)^2 = D \left(\frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \right) = \\ &= D \left(\sum_{i=1}^n \frac{\partial \ln f_{\xi}(X_i, \theta)}{\partial \theta} \right) = \sum_{i=1}^n D \frac{\partial \ln f_{\xi}(X_i, \theta)}{\partial \theta} = n I_1(\theta), \end{aligned}$$

где $I_1(\theta)$ — информационное количество Фишера, соответствующее одному наблюдению. Как видим, наблюдается линейный рост информации.

Из неравенства Рао–Крамера можно сделать вывод, что в регулярном случае дисперсия не может убывать быстрее, чем $1/n$. Для «хороших» оценок дисперсия должна убывать, разброс должен становиться меньше с ростом n . Для несмещенных оценок при выполнении условий регулярности оценка эффективна, если неравенство Рао–Крамера выполнено как равенство.

Справедлива еще одна формула для вычисления $I_n(\theta)$ при выполнении дополнительных условий, которые необходимы для корректного проведения всех последующих преобразований. Продифференцируем равенство

$$\int_{\mathbb{R}^n} \frac{\partial \ln L(x, \theta)}{\partial \theta} L(x, \theta) dx = 0$$

еще один раз, получаем:

$$\int_{\mathbb{R}^n} \left(\frac{\partial \ln L(x, \theta)}{\partial \theta} \right)^2 L(x, \theta) dx + \int_{\mathbb{R}^n} \frac{\partial^2 \ln L(x, \theta)}{\partial \theta^2} L(x, \theta) dx = 0,$$

или, что тоже самое:

$$E \left\{ \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \right\}^2 + E \left\{ \frac{\partial^2 \ln L(X_{[n]}, \theta)}{\partial \theta^2} \right\} = 0.$$

Откуда сразу получаем равенство:

$$I_n(\theta) = -E \left\{ \frac{\partial^2 \ln L(X_{[n]}, \theta)}{\partial \theta^2} \right\}.$$

Замечание 5.5. Справедлив аналог неравенства Рао–Крамера для многомерного случая. Пусть $\theta \in \Theta \subset \mathbb{R}^k$, пусть $E\hat{\theta}(X_{[n]}) = \theta$, тогда

$$D\hat{\theta} \geq (I_n(\theta))^{-1},$$

где I_n — информационная матрица Фишера, а $D\hat{\theta}$ — ковариационная матрица:

$$I_n(\theta) = E \left\{ \frac{\partial \ln L(X_{[n]}, \theta)}{\partial \theta} \left(\frac{\partial \ln L}{\partial \theta} \right)^T \right\},$$

$$D\hat{\theta} = E \left\{ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right\}.$$

Таким образом, матрица $D\hat{\theta} - (I_n(\theta))^{-1} \geq 0$ — неотрицательно определенная.

Пример 5.1. Пусть $\xi \sim N(a, \sigma^2)$, методы максимального правдоподобия и моментов дали оценку $\hat{a} = \bar{X}$, которая является сильно состоятельной, несмещенной, асимптотически нормальной оценкой. Выясним, обращается ли неравенство Рао–Крамера в равенство. Действительно, можно заметить, что

$$\frac{\partial \ln L}{\partial a} = \frac{\sum_{i=1}^n (X_i - a)}{\sigma^2} = \frac{n}{\sigma^2} (\bar{X} - a),$$

где

$$L(X_{[n]}, a) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{\sum_{i=1}^n (X_i - a)^2}{2\sigma^2}}.$$

Как следует из замечания 5.2, \bar{X} — эффективная оценка.

ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

§ 1. АСИМПТОТИЧЕСКИЕ ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

Пусть задана генеральная совокупность ξ с функцией распределения $F_\xi(x)$. Имеется выборка $X_{[n]} = (X_1, \dots, X_n)$ из этой генеральной совокупности и неизвестный параметр распределения $\theta \in \Theta \subset \mathbb{R}$.

Определение 1.1. Пусть для некоторого $\varepsilon \in (0, 1)$ существуют статистики $S^-(X_{[n]}, \varepsilon)$ и $S^+(X_{[n]}, \varepsilon)$ такие, что

$$P \{ S^-(X_{[n]}, \varepsilon) < \theta < S^+(X_{[n]}, \varepsilon) \} = 1 - \varepsilon,$$

тогда интервал $(S^-(X_{[n]}, \varepsilon), S^+(X_{[n]}, \varepsilon))$ называется доверительным интервалом для параметра θ с уровнем доверия $(1 - \varepsilon)$.

Определение 1.2. Пусть для некоторого $\varepsilon \in (0, 1)$ существуют статистики $S^-(X_{[n]}, \varepsilon)$ и $S^+(X_{[n]}, \varepsilon)$ такие, что

$$\lim_{n \rightarrow \infty} P \{ S^-(X_{[n]}, \varepsilon) < \theta < S^+(X_{[n]}, \varepsilon) \} = 1 - \varepsilon,$$

тогда интервал $(S^-(X_{[n]}, \varepsilon), S^+(X_{[n]}, \varepsilon))$ называется асимптотическим (приближенным) доверительным интервалом с уровнем доверия $1 - \varepsilon$.

Построение асимптотических доверительных интервалов основано на асимптотически нормальных оценках. Предположим, что оценка $\hat{\theta} = \hat{\theta}(X_{[n]})$ является асимптотически нормальной, т. е. $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow[n \rightarrow \infty]{d} \zeta \sim N(0, \sigma^2)$, где дисперсия

$\sigma^2(\theta)$ — коэффициент асимптотического рассеивания. Предположим, что функция $\sigma^2(\theta)$ непрерывна на Θ и отлична от нуля для любого $\theta \in \Theta$.

Лемма 1.1. Случайный вектор $(\sqrt{n}(\hat{\theta} - \theta), \hat{\theta}) \xrightarrow[n \rightarrow \infty]{d} (\zeta, \theta)$, где ζ подчиняется нормальному распределению $N(0, \sigma^2(\theta))$.

Доказательство. Покажем, что характеристическая функция случайного вектора $(\sqrt{n}(\hat{\theta} - \theta), \hat{\theta})$ удовлетворяет условию:

$$\varphi_{(\sqrt{n}(\hat{\theta} - \theta), \hat{\theta})}(t_1, t_2) \xrightarrow[n \rightarrow \infty]{} \varphi_{(\zeta, \theta)}(t_1, t_2) = Ee^{it_1\zeta + it_2\theta}.$$

Действительно,

$$\begin{aligned} \varphi_{(\sqrt{n}(\hat{\theta} - \theta), \hat{\theta})}(t_1, t_2) &= Ee^{it_1\sqrt{n}(\hat{\theta} - \theta) + i(t_2\hat{\theta} \pm \theta t_2)} = \\ &= Ee^{i\sqrt{n}(\hat{\theta} - \theta)(t_1 + \frac{t_2}{\sqrt{n}})} e^{it_2\theta} = e^{it_2\theta} \varphi_{\sqrt{n}(\hat{\theta} - \theta)}(t_1 + \frac{t_2}{\sqrt{n}}) = \\ &= e^{it_2\theta} \left(\left\{ \varphi_{\sqrt{n}(\hat{\theta} - \theta)}\left(t_1 + \frac{t_2}{\sqrt{n}}\right) - \varphi_{\zeta}\left(t_1 + \frac{t_2}{\sqrt{n}}\right) \right\} + \right. \\ &\quad \left. + \varphi_{\zeta}\left(t_1 + \frac{t_2}{\sqrt{n}}\right) \right). \end{aligned}$$

При этом $\varphi_{\zeta}(t_1 + t_2/\sqrt{n}) \xrightarrow[n \rightarrow \infty]{} \varphi_{\zeta}(t_1)$, так как любая характеристическая функция равномерно непрерывна.

Имеет место сходимость:

$$\varphi_{\sqrt{n}(\hat{\theta} - \theta)}\left(t_1 + \frac{t_2}{\sqrt{n}}\right) - \varphi_{\zeta}\left(t_1 + \frac{t_2}{\sqrt{n}}\right) \xrightarrow[n \rightarrow \infty]{} 0,$$

так как при любом t : $\varphi_{\sqrt{n}(\hat{\theta} - \theta)}(t) \rightarrow \varphi_{\zeta}(t)$ и сходимость равномерна на любом конечном промежутке [22]. Следовательно, выполняется сходимость:

$$\begin{aligned} \varphi_{(\sqrt{n}(\hat{\theta} - \theta), \hat{\theta})}(t_1, t_2) &\xrightarrow[n \rightarrow \infty]{} e^{it_2\theta} \varphi_{\zeta}(t_1) = \\ &= Ee^{it_2\theta + i\zeta t_1} = \varphi_{(\zeta, \theta)}(t_1, t_2). \end{aligned}$$

Лемма доказана. ■

Рассмотрим функцию от двух переменных $H(x_1, x_2) = x_1/\sigma(x_2)$, она непрерывна на $\mathbb{R} \times \Theta$. Случайный вектор $(\zeta, \theta) \in \mathbb{R} \times \Theta$, следовательно, можем воспользоваться теоремой непрерывности:

$$H(\sqrt{n}(\hat{\theta} - \theta), \hat{\theta}) \xrightarrow[n \rightarrow \infty]{d} H(\zeta, \theta) = \frac{\zeta}{\sigma(\theta)} \sim N(0, 1).$$

Таким образом, имеет место сходимость:

$$\frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma(\hat{\theta})} \xrightarrow[n \rightarrow \infty]{d} \eta \sim N(0, 1).$$

Тогда справедливо соотношение:

$$\begin{aligned} P \left\{ -z_{1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma(\hat{\theta})} < z_{1-\frac{\alpha}{2}} \right\} &\xrightarrow[n \rightarrow \infty]{} 1 - \alpha = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-z_{1-\frac{\alpha}{2}}}^{z_{1-\frac{\alpha}{2}}} e^{-\frac{y^2}{2}} dy, \end{aligned}$$

где $z_{1-\frac{\alpha}{2}}$ — квантиль стандартного нормального распределения уровня $1 - \alpha/2$, т. е. $F(z_{1-\frac{\alpha}{2}}) = 1 - \alpha/2$, где $F(x)$ — функция стандартного нормального распределения. Получаем асимптотический доверительный интервал с уровнем доверия $1 - \alpha$:

$$P \left\{ \hat{\theta} - z_{1-\frac{\alpha}{2}} \frac{\sigma(\hat{\theta})}{\sqrt{n}} < \theta < \hat{\theta} + z_{1-\frac{\alpha}{2}} \frac{\sigma(\hat{\theta})}{\sqrt{n}} \right\} \approx 1 - \alpha.$$

Ширина доверительного интервала характеризует точность интервальной оценки.

Рассмотрим схему Бернулли, в которой n испытаний. Пусть m — число успехов. Выборка $X_{[n]} = (a_1, \dots, a_n)$ состоит из последовательности нулей и единиц, тогда функция правдоподобия имеет вид:

$$L(X_{[n]}, p) = p^m q^{n-m}, \quad p \in \Theta = (0, 1),$$

где m — число единиц в выборке. Логарифмическая функция правдоподобия имеет вид:

$$\ln L = m \ln p + (n - m) \ln(1 - p).$$

Найдем оценку максимального правдоподобия:

$$\frac{\partial \ln L}{\partial p} = \frac{m}{p} - \frac{n - m}{1 - p} = \frac{m - mp - np + mp}{p(1 - p)} = 0.$$

Следовательно, получаем оценку:

$$\hat{p} = \frac{m}{n}.$$

Убеждаемся, что \hat{p} максимизирует функцию правдоподобия:

$$\frac{\partial^2 \ln L}{\partial p^2} = -\frac{m}{p^2} - \frac{n - m}{(1 - p)^2} < 0.$$

Следовательно, $\hat{p} = m/n$ — точка максимума или оценка по методу максимального правдоподобия. Нетрудно показать, что оценка \hat{p} асимптотически нормальна:

$$\begin{aligned} \sqrt{n} \left(\frac{m}{n} - p \right) &= \frac{m - np}{\sqrt{n}} = \frac{\sum_{i=1}^n \xi_i - np}{\sqrt{n}} = \\ &= \frac{\sum_{i=1}^n (\xi_i - p)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \zeta \sim N(0, pq), \end{aligned}$$

где $P\{\xi_i = 1\} = p$, $P\{\xi_i = 0\} = q = 1 - p$, $\sigma^2 = pq = p(1 - p)$. Воспользуемся доказанным утверждением:

$$\frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma(\hat{\theta})} \xrightarrow[n \rightarrow \infty]{d} \eta \sim N(0, 1).$$

Тогда имеет место сходимость:

$$\frac{\sqrt{n}(\frac{m}{n} - p)}{\sqrt{\frac{m}{n}(1 - \frac{m}{n})}} \xrightarrow[n \rightarrow \infty]{d} \eta \sim N(0, 1).$$

Следуя приведенным выше рассуждениям, получаем асимптотический доверительный интервал с уровнем доверия $1 - \alpha$ для вероятности p :

$$\left(\frac{m}{n} - z_{1-\frac{\alpha}{2}} \frac{\sqrt{\frac{m}{n}(1-\frac{m}{n})}}{\sqrt{n}}, \frac{m}{n} + z_{1-\frac{\alpha}{2}} \frac{\sqrt{\frac{m}{n}(1-\frac{m}{n})}}{\sqrt{n}} \right).$$

§ 2. РАСПРЕДЕЛЕНИЯ СТАТИСТИК ДЛЯ ВЫБОРОК ИЗ НОРМАЛЬНОЙ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

Пусть имеется генеральная совокупность $\xi \sim N(a, \sigma^2)$ и выборка $X_{[n]}$ из этой генеральной совокупности. Если ξ — гауссова случайная величина, то функция плотности ее распределения имеет вид:

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

а если ξ — гауссов случайный вектор, то

$$f_{\xi}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-a)^T \Sigma^{-1}(x-a)}, \quad x \in \mathbb{R}^m.$$

В первом случае предполагается, что $\sigma^2 > 0$, а во-втором — что $\det |\Sigma| \neq 0$.

Любому распределению взаимно однозначно соответствует характеристическая функция. С помощью характеристических функций легко получить, что компоненты гауссова случайного вектора независимы тогда и только тогда, когда ковариационная матрица Σ диагональна или, другими словами, когда равны нулю попарные ковариации всех компонент.

Лемма 2.1. Пусть ζ — случайный вектор размерности m , подчиняющийся многомерному нормальному распределению $N(a, \Sigma)$, пусть A — любая матрица размерности $n \times m$, b — вектор размерности $n \times 1$. Тогда вектор $\eta = A\zeta + b$ подчиняется нормальному распределению с параметрами $(Aa + b; A\Sigma A^T)$.

Доказательство. Рассмотрим

$$\begin{aligned}\varphi_{\eta}(t) &= E e^{it^T \eta} = e^{it^T b} E e^{i(t^T A)\zeta} = \\ &= e^{it^T b} \varphi_{\zeta}(t^T A) = e^{it^T (Aa+b) - \frac{1}{2} t^T A \Sigma A^T t},\end{aligned}$$

где $Aa + b$ — математическое ожидание, $A \Sigma A^T$ — ковариационная матрица случайного вектора η . ■

Лемма 2.2. Пусть $\xi = (\xi_1, \dots, \xi_n)^T \sim N(0, \sigma^2 E_n)$ и $\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$ — среднее арифметическое. Тогда $\bar{\xi}$ и вектор $(\xi_1 - \bar{\xi}, \dots, \xi_n - \bar{\xi})$ взаимно независимы.

Доказательство. Возьмем любой элемент вектора, например $\xi_k - \bar{\xi}$, и проверим его независимость с $\bar{\xi}$. Рассмотрим разность:

$$\begin{aligned}\xi_k - \bar{\xi} &= \xi_k - \frac{1}{n} \sum_{i=1}^n \xi_i = \\ &= -\frac{1}{n} \xi_1 - \dots - \frac{1}{n} \xi_{k-1} + \frac{n-1}{n} \xi_k - \frac{1}{n} \xi_{k+1} - \dots - \frac{1}{n} \xi_n = \\ &= \left(-\frac{1}{n}, \dots, -\frac{1}{n}, \frac{n-1}{n}, -\frac{1}{n}, \dots, -\frac{1}{n} \right) \xi,\end{aligned}$$

$$\bar{\xi} = \frac{1}{n} \xi_1 + \dots + \frac{1}{n} \xi_n = \left(\frac{1}{n}, \dots, \frac{1}{n} \right) \xi.$$

Следовательно,

$$\left(-\frac{1}{n}, \dots, -\frac{1}{n}, \frac{n-1}{n}, -\frac{1}{n}, \dots, -\frac{1}{n} \right) \xi = \begin{pmatrix} \xi_k - \bar{\xi} \\ \bar{\xi} \end{pmatrix}.$$

Тогда по лемме 2.1 получаем:

$$\begin{pmatrix} \xi_k - \bar{\xi} \\ \bar{\xi} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right),$$

где $\sigma_1^2 = \frac{n-1}{n}\sigma^2$, $\sigma_2^2 = \frac{1}{n}\sigma^2$, внедиагональные элементы равны нулю, поскольку

$$\begin{aligned}\sigma^2 \left(-\frac{1}{n}, \dots, -\frac{1}{n}, \frac{n-1}{n}, -\frac{1}{n}, \dots, -\frac{1}{n} \right) \left(\frac{1}{n}, \dots, \frac{1}{n} \right)^T = \\ = \sigma^2 \frac{-(n-1) + (n-1)}{n^2} = 0.\end{aligned}$$

■

Лемма 2.3. Пусть $\xi = (\xi_1, \dots, \xi_m)^T \sim N(0, E_m)$, $CC^T = C^TC = E_m$ и $\eta = C\xi$. Тогда $\tau = \sum_{k=1}^m \xi_k^2 - \eta_1^2 - \dots - \eta_r^2$ подчиняется распределению хи-квадрат с $m-r$ степенями свободы, и случайные величины η_1, \dots, η_r взаимно независимы с τ .

Доказательство. Из леммы 2.1 следует, что $\eta \sim N(0, E_{m \times m})$. Как легко видеть, имеет место равенство:

$$\sum_{i=1}^m \xi_i^2 = \xi^T \xi = \xi^T C^T C \xi = \eta^T \eta = \sum_{i=1}^m \eta_i^2.$$

Следовательно, справедливо равенство:

$$\sum_{i=1}^m \xi_i^2 - \eta_1^2 - \dots - \eta_r^2 = \sum_{j=r+1}^m \eta_j^2,$$

полученное равенство доказывает лемму. ■

§ 3. РАСПРЕДЕЛЕНИЕ СТЬЮДЕНТА

Определение 3.1. Пусть заданы случайные величины $\zeta \sim N(0, 1)$ и $\tau_k \sim \chi_k^2$. Пусть случайные величины ζ и τ_k взаимно независимы. Распределение случайной величины

$$\xi = \frac{\zeta}{\sqrt{\frac{\tau_k}{k}}}$$

называется распределением Стьюдента с k степенями свободы и обозначается через T_k .

Замечание 3.1. Очевидно, что если $\zeta \sim N(0, 1)$ и ζ_1, \dots, ζ_k — взаимно независимые случайные величины, подчиняющиеся стандартному нормальному распределению, независимые с ζ , тогда

$$\frac{\zeta}{\sqrt{\frac{\zeta_1^2 + \dots + \zeta_k^2}{k}}} \sim T_k.$$

Лемма 3.1. Пусть $\zeta = \eta/\xi$, где η, ξ — взаимно независимые случайные величины. Пусть $\xi >^{\text{п.н.}} 0$, $f_\xi(x)$, $f_\eta(y)$ — плотности распределения ξ и η соответственно, тогда плотность распределения $f_\zeta(z)$ дроби ζ имеет вид:

$$f_\zeta(z) = \int_0^\infty x f_\xi(x) f_\eta(zx) dx.$$

Доказательство. Найдем функцию распределения случайной величины ζ :

$$\begin{aligned} F_\zeta(z) &= P\{\zeta \leq z\} = \int_{\{(x,y): \frac{y}{x} \leq z, x > 0\}} f_\xi(x) f_\eta(y) dx dy = \\ &= \int_0^\infty \int_{-\infty}^{zx} f_\xi(x) f_\eta(y) dy dx = \int_0^\infty f_\xi(x) \left(\int_{-\infty}^z f_\eta(xs) x ds \right) dx. \end{aligned}$$

Положим $s = y/x$, $dy = x ds$. Меняем порядок интегрирования:

$$F_\zeta(z) = \int_{-\infty}^z \left(\int_0^\infty x f_\xi(x) f_\eta(xs) dx \right) ds.$$

Полученное равенство доказывает лемму. ■

Лемма 3.2. Пусть $\eta \sim N(0, 1)$, $\xi = \sqrt{\tau}$, $\tau \sim \chi_k^2$. Пусть случайные величины η и τ взаимно независимы. Тогда случайная величина $\zeta = \eta/\xi$ имеет плотность распределения следующего вида:

$$f_\zeta(z) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{\pi} \Gamma(\frac{k}{2})} \frac{1}{(1+z^2)^{\frac{k+1}{2}}}.$$

Доказательство. Распределение хи-квадрат с k степенями свободы представляет собой гамма-распределение с параметрами формы $k/2$ и масштаба $1/2$:

$$f_{\tau}(x) = \begin{cases} (\frac{1}{2})^{\frac{k}{2}} \frac{x^{\frac{k}{2}-1}}{\Gamma(\frac{k}{2})} e^{-\frac{x}{2}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Как легко заметить, имеет место равенство:

$$P\{\sqrt{\tau} \leq x\} = P\{\tau \leq x^2\},$$

тогда

$$f_{\sqrt{\tau}}(x) = 2xf_{\tau}(x^2) = \begin{cases} \frac{1}{2^{\frac{k}{2}-1}} \frac{x^{k-1}}{\Gamma(\frac{k}{2})} e^{-\frac{x^2}{2}}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

Следовательно, имеют место равенства:

$$\begin{aligned} f_{\zeta}(z) &= \int_0^{\infty} xf_{\xi}(x)f_{\eta}(zx)dx = \\ &= \frac{1}{2^{\frac{k}{2}-1}\Gamma(\frac{k}{2})\sqrt{2\pi}} \int_0^{\infty} x^k e^{-\frac{x^2}{2}} e^{-\frac{z^2x^2}{2}} dx = \\ &= \frac{1}{2^{\frac{k-1}{2}}\Gamma(\frac{k}{2})\sqrt{\pi}} \int_0^{\infty} x^k e^{-\frac{x^2}{2}(z^2+1)} dx = \\ &= \frac{1}{2^{\frac{k-1}{2}}\sqrt{\pi}\Gamma(\frac{k}{2})} \int_0^{\infty} \frac{2^{\frac{k-1}{2}} u^{\frac{k-1}{2}}}{(z^2+1)^{\frac{k-1}{2}}} \frac{1}{z^2+1} e^{-u} du = \\ &= \frac{1}{\sqrt{\pi}\Gamma(\frac{k}{2})} \frac{1}{(z^2+1)^{\frac{k+1}{2}}} \int_0^{\infty} u^{\frac{k+1}{2}-1} e^{-u} du, \end{aligned}$$

где $u = x^2(z^2+1)/2$, $xdx = du/(z^2+1)$, причем $\int_0^{\infty} u^{\frac{k+1}{2}-1} e^{-u} \times$
 $\times du = \Gamma\left(\frac{k+1}{2}\right)$. Лемма доказана. ■

Следствие 3.1. Плотность распределения Стьюдента с k степенями свободы имеет вид:

$$f(z) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{\pi k} \Gamma(\frac{k}{2})} \frac{1}{(1 + \frac{z^2}{k})^{\frac{k+1}{2}}}. \quad (3.1)$$

Доказательство. Для доказательства достаточно заметить, что случайная величина $\sqrt{k}\zeta$ подчиняется распределению Стьюдента с k степенями свободы и $f_{\sqrt{k}\zeta}(z) = \frac{1}{\sqrt{k}} f_{\zeta}(\frac{z}{\sqrt{k}})$. ■

Замечание 3.2. Можно показать, что для плотности $f(z)$ из выражения (3.1) имеет место сходимость:

$$f(z) \xrightarrow[k \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

§ 4. ТОЧНЫЕ ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ ДЛЯ НОРМАЛЬНОЙ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

Теорема 4.1. Пусть задана выборка $X_{[n]}$ из генеральной совокупности $\xi \sim N(a, \sigma^2)$. Справедливы утверждения:

1. Статистика $\frac{\bar{X}-a}{\sigma} \sqrt{n}$ подчиняется стандартному нормальному распределению.
2. Если $\tilde{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, тогда статистика $\frac{\bar{X}-a}{\tilde{s}} \sqrt{n}$ подчиняется распределению Стьюдента с $n-1$ степенью свободы.
3. Статистика $\frac{(n-1)\tilde{s}^2}{\sigma^2}$ подчиняется распределению хи-квадрат с $n-1$ степенью свободы.
4. Если $s_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - a)^2$, тогда статистика $\frac{n s_0^2}{\sigma^2}$ подчиняется распределению хи-квадрат с n степенями свободы.

Доказательство. В лемме 2.1 было показано, что линейное преобразование не меняет тип нормального распределения и изменяет только параметры. Тогда утверждение 1 теоремы очевидно.

Очевидно, что

$$\begin{aligned}\tilde{s}^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n (X_i - a)^2 - n \left(\frac{1}{n} \sum_{i=1}^n (X_i - a) \right)^2 \right) = \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n (X_i - a)^2 - \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j - a) \right)^2 \right).\end{aligned}$$

Тогда справедливо равенство:

$$\frac{(n-1)\tilde{s}^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - a}{\sigma} \right)^2 - \left(\sum_{j=1}^n \frac{1}{\sqrt{n}} \frac{X_j - a}{\sigma} \right)^2.$$

Введем обозначение:

$$\delta = \begin{pmatrix} \frac{X_1 - a}{\sigma} \\ \vdots \\ \frac{X_n - a}{\sigma} \end{pmatrix},$$

при этом δ подчиняется многомерному нормальному распределению с параметрами $(0, E_n)$.

Рассмотрим строку $(1/\sqrt{n}, \dots, 1/\sqrt{n}) = C_1$. При этом нетрудно заметить, что $C_1 C_1^T = 1$, тогда методом ортогонализации Грама–Шмидта последовательно получим $n-1$ строку C_2, \dots, C_n . Строки будут ортогональными и нормированными. Составим матрицу:

$$C = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}.$$

Рассмотрим преобразование $C\delta = \eta$, из лемм 2.1 и 2.3 следует:

$$\frac{(n-1)\tilde{s}^2}{\sigma^2} = \sum_{i=1}^n \delta_i^2 - \eta_1^2 \sim \chi_{n-1}^2,$$

где $\delta_i = \frac{X_i - a}{\sigma}$, $\eta_1 = \sum_{j=1}^n \frac{1}{\sqrt{n}} \frac{X_j - a}{\sigma} = \sum_{j=1}^n \frac{1}{\sqrt{n}} \delta_j = C_1 \delta$. Утверждение 3 теоремы доказано.

В лемме 2.2 было доказано, что $(X_1 - \bar{X}, \dots, X_n - \bar{X})$ и $\bar{X} - a$ взаимно независимы. Рассмотрим дробь Стьюдента:

$$\frac{\frac{\bar{X} - a}{\sigma} \sqrt{n}}{\sqrt{\frac{1}{n-1} \frac{(n-1)\bar{s}^2}{\sigma^2}}} = \frac{\bar{X} - a}{\bar{s}} \sqrt{n} \sim T_{n-1}.$$

Таким образом, доказано утверждение 2 теоремы.

Утверждение 4 следует из определения:

$$\frac{ns_0^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - a)^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - a}{\sigma} \right)^2 \sim \chi_n^2.$$

Теорема доказана. ■

Точные доверительные интервалы для нормальной генеральной совокупности можно построить по следующим правилам:

1. Если a неизвестно, σ^2 известно, тогда

$$P \left\{ \bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\varepsilon}{2}} < a < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\varepsilon}{2}} \right\} = 1 - \varepsilon,$$

где $z_{1-\frac{\varepsilon}{2}}$ — квантиль стандартного нормального распределения.

2. Если a неизвестно, σ^2 неизвестно, то доверительный интервал для a будет иметь вид:

$$P \left\{ \bar{X} - \frac{\tilde{s}}{\sqrt{n}} t_{1-\frac{\varepsilon}{2}} < a < \bar{X} + \frac{\tilde{s}}{\sqrt{n}} t_{1-\frac{\varepsilon}{2}} \right\} = 1 - \varepsilon,$$

где $t_{1-\frac{\varepsilon}{2}}$ — квантиль распределения Стьюдента с $n - 1$ степенью свободы.

3. Предположим, что a неизвестно, σ^2 неизвестно и необходимо построить доверительный интервал для σ^2 . Пусть $\varepsilon = \varepsilon_1 + \varepsilon_2$. Пусть $u_{1-\varepsilon_2}$ — квантиль распределения хи-квадрат с $n - 1$ степенью свободы уровня $1 - \varepsilon_2$, u_{ε_1} —

квантиль распределения хи-квадрат с $n-1$ степенью свободы уровня ε_1 , тогда доверительный интервал для σ^2 будет следующим:

$$P \left\{ \frac{(n-1)\bar{s}^2}{u_{1-\varepsilon_2}} < \sigma^2 < \frac{(n-1)\bar{s}^2}{u_{\varepsilon_1}} \right\} = 1 - \varepsilon.$$

4. Если a известно, σ^2 неизвестно, тогда доверительный интервал строится так же, как и в случае 3, только в качестве статистики рассматривается статистика ns_0^2/σ^2 из пункта 4 теоремы 4.1. Пусть v_{ε_1} — квантиль распределения хи-квадрат с n степенями свободы уровня ε_1 , $v_{1-\varepsilon_2}$ — квантиль распределения хи-квадрат с n степенями свободы уровня $1-\varepsilon_2$, тогда доверительный интервал для σ^2 будет

$$P \left\{ \frac{ns_0^2}{v_{1-\varepsilon_2}} < \sigma^2 < \frac{ns_0^2}{v_{\varepsilon_1}} \right\} = 1 - \varepsilon,$$

здесь $\varepsilon = \varepsilon_1 + \varepsilon_2$.

Обычно при построении доверительных интервалов для дисперсии выбирают $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$.

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

§ 1. ПРОВЕРКА ДВУХ ПРОСТЫХ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Под *статистической гипотезой* принято понимать любое предположение о законе распределения генеральной совокупности. Статистическая гипотеза называется *простой*, если при условии истинности гипотезы закон распределения генеральной совокупности однозначно определен, в противном случае гипотеза называется *сложной*.

Пусть задана выборка $X_{[n]}$ из генеральной совокупности ξ с функцией распределения $F_\xi(x)$. Пусть имеются две простые гипотезы:

- $H_0 : F_\xi(x) = F_0(x)$.
- $H_1 : F_\xi(x) = F_1(x)$.

Причем, функции $F_0(x)$ и $F_1(x)$ полностью известны. По выборке $X_{[n]}$ требуется принять решение об истинности нулевой гипотезы H_0 при альтернативной гипотезе H_1 .

Выборка $X_{[n]}$ — точка из пространства \mathbb{R}^n . Выделим множество $S \subset \mathbb{R}^n$ — критическую область для гипотезы H_0 , тогда можно сформулировать правило проверки гипотезы H_0 при альтернативе H_1 :

- Если $X_{[n]} \in S$, то отвергаем гипотезу H_0 , принимаем H_1 .
- Если $X_{[n]} \notin S$, то принимаем гипотезу H_0 , отвергаем H_1 .

Правило проверки статистической гипотезы при некоторой фиксированной альтернативе принято называть *статистическим критерием*.

За исключением тривиальных ситуаций сформулированное правило не может всегда приводить к правильным решениям. Возможны два типа ошибок:

1. *Ошибка первого рода* — отклонить гипотезу H_0 , когда она верна, вероятность ошибки первого рода $\alpha(S)$ определяется равенством:

$$\alpha(S) = P \{X_{[n]} \in S/H_0\} = P_0 \{X_{[n]} \in S\}.$$

2. *Ошибка второго рода* — принять гипотезу H_0 , когда верна H_1 , вероятность ошибки второго рода $\beta(S)$ определяется равенством:

$$\beta(S) = P \{X_{[n]} \notin S/H_1\} = P_1 \{X_{[n]} \notin S\}.$$

Также будем рассматривать вероятность

$$\gamma(S) = 1 - \beta(S) = P_1 \{X_{[n]} \in S\},$$

вероятность $\gamma(S)$ называют *мощностью критерия*. Если $\gamma(S) < \alpha(S)$, то попасть в S при условии истинности гипотезы H_1 труднее, чем при условии истинности гипотезы H_0 , т. е. S — критическая область скорее для H_1 . Следовательно, неравенство должно иметь вид:

$$\gamma(S) > \alpha(S),$$

т. е. S следует выбирать так, чтобы выполнялось это неравенство.

Определение 1.1. Критерий называется несмещенным, если выполняется условие

$$\alpha(S) \leq \gamma(S) = 1 - \beta(S).$$

В большинстве задач гипотезы H_0 и H_1 не равноправны. Поэтому в дальнейшем изложении будем считать, что H_0 — основная гипотеза, H_1 — альтернативная гипотеза.

Зададим α_0 и будем иметь дело только с такими критериями, где $\alpha_0 \geq \alpha(S)$ (т. е. вероятность ошибки первого рода не

превосходит величины α_0) и дополнительно будем решать задачу: $\beta(S) \rightarrow \min_S$. Получаем две эквивалентные задачи определения критической области S :

$$\begin{cases} \alpha_0 \geq \alpha(S), \\ \beta(S) \rightarrow \min_S; \end{cases}$$

$$\begin{cases} \alpha_0 \geq \alpha(S), \\ \gamma(S) \rightarrow \max_S. \end{cases}$$

Задачи в такой постановке не всегда решаемы, так как требуется ответить точно «да» или «нет». Такие статистические критерии называются *нерандомизированными критериями*.

Можно рассмотреть функцию $\varphi(x) = I\{x \in S\}$. Тогда нерандомизированный критерий примет вид:

- Если $\varphi(X_{[n]}) = 1$, тогда отвергаем гипотезу H_0 , принимаем H_1 .
- Если $\varphi(X_{[n]}) = 0$, тогда принимаем гипотезу H_0 , отвергаем H_1 .

Функцию φ принято называть *критической функцией*.

Пусть теперь φ не является индикатором множества S , и

$$\varphi(x) = P\{\bar{H}_0 / X_{[n]} = x\},$$

тогда $\varphi(X_{[n]}) \in [0, 1]$ — условная вероятность отклонения гипотезы H_0 . При таком определении $\varphi(x)$ приходим к рандомизированному критерию, то есть критерию, который при некоторых значениях x может не давать ответа «да» или «нет» в отношении истинности нулевой гипотезы H_0 . Тогда с вероятностью $1 - \varphi(X_{[n]})$ следует принимать гипотезу H_0 и с вероятностью $\varphi(X_{[n]})$ принимать гипотезу H_1 . При использовании введенного обозначения вероятность ошибки первого рода, вероятность ошибки второго рода и мощность критерия будем обозначать: $\alpha(\varphi)$, $\beta(\varphi)$ и $\gamma(\varphi) = 1 - \beta(\varphi)$ соответственно. Определение 1.1 можно переформулировать в новых терминах, критерий называется несмещенным, если $\alpha(\varphi) \leq \gamma(\varphi)$.

Без ограничения общности будем предполагать, что существует плотность $f_0(x)$ для функции распределения $F_0(x)$

и существует плотность $f_1(x)$ для функции распределения $F_1(x)$. В дискретном случае все результаты аналогичны. В дальнейшем будем считать, что плотности $f_i(x)$, $i = 1, 2$, определены относительно сигма-конечной меры μ . Если в качестве меры μ рассматривается мера Лебега, то плотности распределения представляют собой обычные плотности распределения. Если в качестве меры μ рассматривается «считающая» мера, тогда плотности распределения являются вероятностями, что соответствует дискретному случаю. При этом в любом случае $F_i(x) = \int_{-\infty}^x f_i(t) \mu(dt)$, $i = 0, 1$.

Функция правдоподобия выборки (совместная плотность выборки) при справедливости гипотезы H_0 имеет вид:

$$L_0(X_{[n]}) = \prod_{i=1}^n f_0(X_i).$$

Если верна гипотеза H_1 , то функция правдоподобия имеет вид:

$$L_1(X_{[n]}) = \prod_{i=1}^n f_1(X_i).$$

Для рандомизированного критерия получаем

$$P_0(\bar{H}_0) = \int_{\mathbb{R}^n} \varphi(x) L_0(x) \mu_n(dx) = \alpha(\varphi),$$

$$P_1(H_0) = \int_{\mathbb{R}^n} (1 - \varphi(x)) L_1(x) \mu_n(dx) = \beta(\varphi),$$

$$\gamma(\varphi) = \int_{\mathbb{R}^n} \varphi(x) L_1(x) \mu_n(dx),$$

где μ_n — произведение мер μ . Очевидно, что $\gamma(\varphi) = 1 - \beta(\varphi)$.

Тогда задача построения статистического критерия сводится к нахождению критической функции $\varphi(x)$ и будет формулироваться следующим образом:

$$\begin{cases} \alpha_0 \geq \alpha(\varphi), \\ \beta(\varphi) \rightarrow \min_{\varphi}. \end{cases}$$

Эквивалентная формулировка имеет вид:

$$\begin{cases} \alpha_0 \geq \alpha(\varphi), \\ \gamma(\varphi) \longrightarrow \max_{\varphi}. \end{cases}$$

Таким образом, задача заключается в том, чтобы найти наиболее мощный критерий, когда вероятность ошибки первого рода не превосходит некоторого заданного порогового значения. Решение сформулированных задач дается леммой Неймана–Пирсона.

Лемма 1.1 (лемма Неймана–Пирсона). Пусть $\alpha_0 \in (0, 1)$, тогда при фиксированной вероятности ошибки первого рода α_0 наиболее мощный критерий имеет критическую функцию φ^* вида

$$\varphi^*(x) = \begin{cases} 1, & \text{если } L_1(x) > cL_0(x); \\ \varepsilon, & \text{если } L_1(x) = cL_0(x); \\ 0, & \text{если } L_1(x) < cL_0(x), \end{cases}$$

где $L_0(x) = \prod_{i=1}^n f_0(x_i)$ соответствует гипотезе H_0 , $L_1(x) = \prod_{i=1}^n f_1(x_i)$ — гипотезе H_1 . Константы c и ε — решения уравнения $\alpha(\varphi^*) = \alpha_0$.

Доказательство. Покажем, что константы c и ε могут быть найдены из уравнения $\alpha(\varphi^*) = \alpha_0$. Заметим, что

$$\begin{aligned} \alpha(\varphi^*) &= P_0\{L_1(X_{[n]}) > cL_0(X_{[n]})\} + \\ &\quad + \varepsilon P_0\{L_1(X_{[n]}) = cL_0(X_{[n]})\} = \\ &= P_0\left\{\frac{L_1(X_{[n]})}{L_0(X_{[n]})} > c\right\} + \varepsilon P_0\left\{\frac{L_1(X_{[n]})}{L_0(X_{[n]})} = c\right\}. \end{aligned} \quad (1.1)$$

Может ли быть так, что $L_0(X_{[n]}) = 0$? Да, может, но тогда

$$P_0\{L_0(X_{[n]}) = 0\} = \int_{\{x: L_0(x)=0\}} L_0(x) \mu_n(dx) = 0$$

и, следовательно, равенство (1.1) корректно. Поэтому рассмотрим случайную величину $\eta(X_{[n]}) = L_1(X_{[n]})/L_0(X_{[n]})$. Положим $F_{H_0, \eta}(t) = P_0\{\eta \leq t\}$, тогда

$$\alpha(\varphi^*) = 1 - F_{H_0, \eta}(c) + \varepsilon(F_{H_0, \eta}(c) - F_{H_0, \eta}(c - 0)).$$

Пусть $g(c) = 1 - F_{H_0, \eta}(c)$, константу c_{α_0} можно выбрать так, чтобы было выполнено неравенство:

$$g(c_{\alpha_0}) \leq \alpha_0 \leq g(c_{\alpha_0} - 0),$$

если $g(c_{\alpha_0}) = g(c_{\alpha_0} - 0)$, то положим $\varepsilon_{\alpha_0} = 0$, если $g(c_{\alpha_0}) < g(c_{\alpha_0} - 0)$, то в качестве ε_{α_0} выберем следующую величину:

$$\varepsilon_{\alpha_0} = \frac{\alpha_0 - g(c_{\alpha_0})}{g(c_{\alpha_0} - 0) - g(c_{\alpha_0})} \in [0, 1].$$

Нетрудно заметить, что в обоих случаях выполнено равенство:

$$\alpha_0 = g(c_{\alpha_0}) + \varepsilon_{\alpha_0}(g(c_{\alpha_0} - 0) - g(c_{\alpha_0})) = \alpha(\varphi^*).$$

Докажем, что $\varphi^*(x)$ — критическая функция наиболее мощного критерия. Выберем любую другую критическую функцию $\tilde{\varphi}(x)$ такую, что $\alpha(\tilde{\varphi}) \leq \alpha_0$, и сравним ее с критической функцией φ^* . Заметим, что для любого x справедливо неравенство:

$$(\varphi^*(x) - \tilde{\varphi}(x))(L_1(x) - c_{\alpha_0}L_0(x)) \geq 0.$$

Тогда

$$\int_{\mathbb{R}^n} (\varphi^*(x) - \tilde{\varphi}(x))(L_1(x) - c_{\alpha_0}L_0(x))\mu_n(dx) \geq 0.$$

Раскроем скобки и преобразуем:

$$\begin{aligned} & \int_{\mathbb{R}^n} \varphi^*(x)L_1(x)\mu_n(dx) - \int_{\mathbb{R}^n} \tilde{\varphi}(x)L_1(x)\mu_n(dx) \geq \\ & \geq c_{\alpha_0} \left(\int_{\mathbb{R}^n} \varphi^*(x)L_0(x)\mu_n(dx) - \int_{\mathbb{R}^n} \tilde{\varphi}(x)L_0(x)\mu_n(dx) \right) \geq 0. \end{aligned}$$

Следовательно, $\gamma(\varphi^*) - \gamma(\tilde{\varphi}) \geq c_{\alpha_0}(\alpha(\varphi^*) - \alpha(\tilde{\varphi}))$, откуда получаем неравенство:

$$\gamma(\varphi^*) \geq \gamma(\tilde{\varphi}).$$



§ 2. ПРОСТЫЕ ГИПОТЕЗЫ О ПАРАМЕТРАХ НОРМАЛЬНОГО И БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЙ

Нормальное распределение

Пусть задана генеральная совокупность ξ , выборка $X_{[n]}$ из этой генеральной совокупности, имеются две гипотезы о распределении генеральной совокупности $N(a_0, \sigma^2)$, $N(a_1, \sigma^2)$, где a_0 , a_1 известны. Также считаем, что σ^2 известна. Пусть для определенности $a_1 > a_0$. То есть имеем две гипотезы:

- $H_0 : a = a_0$.
- $H_1 : a = a_1 > a_0$.

Без ограничения общности считаем, что $a_1 > a_0$. Применим критерий Неймана–Пирсона. Выпишем функции правдоподобия для каждой гипотезы:

$$L_0(X_{[n]}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a_0)^2},$$

$$L_1(X_{[n]}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a_1)^2}.$$

Рассмотрим отношение:

$$\frac{L_1(X_{[n]})}{L_0(X_{[n]})} = \exp \frac{1}{2\sigma^2} \{2(a_1 - a_0)n\bar{X} - n(a_1^2 - a_0^2)\}.$$

Нетрудно заметить, что $L_1(X_{[n]})/L_0(X_{[n]}) > c$ тогда и только тогда, когда $\bar{X} > c_1$, поэтому оптимальный критерий Неймана–Пирсона выглядит следующим образом:

$$\varphi^*(x) = \begin{cases} 1, & \bar{X} > c_1; \\ \varepsilon, & \bar{X} = c_1; \\ 0, & \bar{X} < c_1, \end{cases}$$

при этом константы c_1 и ε выбираются при заданном $\alpha_0 \in (0, 1)$ как решение уравнения $\alpha_0 = \alpha(\varphi^*)$. Так как при справедливости гипотезы H_0 распределение статистики \bar{X} является нормальным, то $P\{\bar{X} = c_1\} = 0$, поэтому можно положить

$\varepsilon = 0$, тогда

$$\varphi^*(x) = \begin{cases} 1, & \bar{X} > c_1; \\ 0, & \bar{X} \leq c_1. \end{cases}$$

В этом случае всегда $g(c_{\alpha_0}) = g(c_{\alpha_0} - 0)$, поэтому можно выбрать $\varepsilon = 0$. Оптимальный критерий является нерандомизированным.

Зададим вероятность ошибки первого рода α_0 :

$$\begin{aligned} \alpha_0 &= \alpha(\varphi^*) = P_0\{\bar{X} > c_1\} = \\ &= P_0\left\{\frac{\bar{X} - a_0}{\sigma}\sqrt{n} > \frac{c_1 - a_0}{\sigma}\sqrt{n}\right\} = 1 - \Phi\left(\frac{c_1 - a_0}{\sigma}\sqrt{n}\right), \end{aligned}$$

где $\Phi(x)$ — функция распределения, соответствующая стандартному нормальному распределению,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Таким образом, получили уравнение:

$$\Phi\left(\frac{c_1 - a_0}{\sigma}\sqrt{n}\right) = 1 - \alpha_0.$$

Его решением является квантиль уровня $1 - \alpha_0$ стандартного нормального распределения $z_{1-\alpha_0} = \sqrt{n}(c_1 - a_0)/\sigma$. Равенство

$$\alpha(\varphi^*) = \alpha_0 \tag{2.1}$$

будет выполнено, если выбрать $c_1 = a_0 + z_{1-\alpha_0}\sigma/\sqrt{n}$ и, следовательно, критическая область для нулевой гипотезы H_0 при использовании статистики \bar{X} имеет следующий вид:

$$S = \left(a_0 + \frac{\sigma}{\sqrt{n}}z_{1-\alpha_0}; +\infty\right).$$

Если $\bar{X} \in S$, то гипотезу H_0 следует отклонить, если $\bar{X} \notin S$, то гипотезу H_0 следует принять. Оптимальный критерий в данном случае является нерандомизированным.

Если выберем

$$c_1 \geq a_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha_0}, \quad (2.2)$$

то вероятность ошибки первого рода будет удовлетворять неравенству $\alpha(\varphi^*) \leq \alpha_0$. Введем в рассмотрение ошибку второго рода β . Зададим уровень ошибки второго рода β_0 и выясним условия, когда $\beta(\varphi^*) \leq \beta_0$:

$$\begin{aligned} \beta(\varphi^*) &= P_1\{\bar{X} \leq c_1\} = \\ &= P_1\left\{\frac{\bar{X} - a_1}{\sigma} \sqrt{n} \leq \frac{c_1 - a_1}{\sigma} \sqrt{n}\right\} = \Phi\left\{\frac{c_1 - a_1}{\sigma} \sqrt{n}\right\}. \end{aligned}$$

Неравенство $\beta(\varphi^*) \leq \beta_0$ окажется выполненным, если

$$\frac{c_1 - a_1}{\sigma} \sqrt{n} \leq z_{\beta_0},$$

где z_{β_0} — квантиль стандартного нормального распределения уровня β_0 . Из последнего неравенства получаем:

$$c_1 \leq a_1 + \frac{\sigma}{\sqrt{n}} z_{\beta_0}. \quad (2.3)$$

Константа c_1 , для которой выполнены неравенства (2.2) и (2.3), может быть выбрана, если имеет место неравенство:

$$a_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha_0} \leq a_1 + \frac{\sigma}{\sqrt{n}} z_{\beta_0}.$$

Преобразуя последнее неравенство, получим

$$n \geq \frac{\sigma^2 (z_{1-\alpha_0} - z_{\beta_0})^2}{(a_1 - a_0)^2}. \quad (2.4)$$

Таким образом, при объеме выборки n , удовлетворяющем неравенству (2.4), будут выполнены условия: $\alpha(\varphi^*) \leq \alpha_0$ и $\beta(\varphi^*) \leq \beta_0$.

Предположим, что известно математическое ожидание a . Сформулируем гипотезы:

- $H_0 : \sigma = \sigma_0$.
- $H_1 : \sigma = \sigma_1 > \sigma_0$.

Без ограничения общности считаем, что $\sigma_1 > \sigma_0$, σ_0 и σ_1 известны.

Статистический критерий проверки гипотезы H_0 при альтернативе H_1 основан на статистике отношения правдоподобия, как следует из леммы Неймана–Пирсона. Выпишем функции правдоподобия:

$$L_0(X_{[n]}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_0^n} e^{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (X_i - a)^2},$$

$$L_1(X_{[n]}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_1^n} e^{-\frac{1}{2\sigma_1^2} \sum_{i=1}^n (X_i - a)^2}.$$

Тогда неравенство

$$\frac{L_1(X_{[n]})}{L_0(X_{[n]})} = \left(\frac{\sigma_0}{\sigma_1}\right)^n e^{\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right) \sum_{i=1}^n (X_i - a)^2} > c$$

равносильно неравенству:

$$\sum_{i=1}^n (X_i - a)^2 > c_1.$$

Ранее было показано, что статистика $\frac{1}{\sigma_0^2} \sum_{i=1}^n (X_i - a)^2$ подчиняется распределению хи-квадрат с n степенями свободы при условии справедливости гипотезы H_0 . Следовательно, вероятность $P_0(\sum_{i=1}^n (X_i - a)^2 = c_1)$ равна нулю при любом значении константы c_1 , поэтому в оптимальном критерии можно положить $\varepsilon = 0$. Поэтому критерий имеет вид:

$$\varphi^*(x) = \begin{cases} 1, & \sum_{i=1}^n (X_i - a)^2 > c_1; \\ 0, & \sum_{i=1}^n (X_i - a)^2 \leq c_1, \end{cases}$$

и, оказывается, нерандомизированным. Выберем константу так, чтобы выполнялось равенство: $\alpha(\varphi^*) = \alpha_0$. Нетрудно

заметить, что

$$\begin{aligned}\alpha(\varphi^*) &= P_0 \left\{ \sum_{i=1}^n (X_i - a)^2 > c_1 \right\} = \\ &= P_0 \left\{ \sum_{i=1}^n \left(\frac{X_i - a}{\sigma_0} \right)^2 > \frac{c_1}{\sigma_0^2} \right\} = 1 - F_{\chi_n^2} \left(\frac{c_1}{\sigma_0^2} \right),\end{aligned}$$

где $F_{\chi_n^2}(\cdot)$ — функция распределения закона хи-квадрат с n степенями свободы. Решением уравнения $F_{\chi_n^2}(c_1/\sigma_0^2) = 1 - \alpha_0$ является квантиль распределения χ^2 с n степенями свободы, то есть $u_{1-\alpha_0, n}$.

Критическая область для гипотезы H_0 выглядит следующим образом:

- Если $\sum_{i=1}^n (X_i - a)^2 \in (\sigma_0^2 u_{1-\alpha_0, n}; +\infty) = S$, то гипотеза H_0 отклоняется.
- Если $\sum_{i=1}^n (X_i - a)^2 \in [0; \sigma_0^2 u_{1-\alpha_0, n}]$, то гипотеза H_0 принимается.

Оптимальный критерий является нерандомизированным критерием.

Биномиальное распределение

Теперь рассмотрим случай, когда проверяется гипотеза о параметре биномиального распределения. Рассмотрим схему Бернулли. Выборка $X_{[n]}$ состоит из нулей и единиц, единицы соответствуют успехам. Тогда вероятность того, что в серии из n испытаний произойдет ровно m успехов равна

$$P_n\{\xi = m\} = C_n^m p^m (1-p)^{n-m},$$

где n — число испытаний; p — вероятность успеха; m — число успехов.

Сформулируем гипотезы:

- $H_0 : p = p_0$.
- $H_1 : p = p_1 > p_0$.

Как и в предыдущих примерах, без ограничения общности считаем, что $p_1 > p_0$, p_1 и p_0 известны. Применим оптималь-

ный критерий Неймана–Пирсона. Выпишем функции правдоподобия для каждой гипотезы:

$$L_1(m) = C_n^m p_1^m (1 - p_1)^{n-m},$$

$$L_0(m) = C_n^m p_0^m (1 - p_0)^{n-m}.$$

Неравенство

$$\frac{L_1(m)}{L_0(m)} = \left(\frac{p_1}{p_0}\right)^m \left(\frac{1-p_1}{1-p_0}\right)^{n-m} > c$$

эквивалентно неравенству:

$$\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)^m \left(\frac{1-p_1}{1-p_0}\right)^n > c,$$

или эквивалентно неравенству: $m > c_1$. Следовательно, оптимальную критическую функцию можно записать в виде:

$$\varphi^*(x) = \begin{cases} 1, & m > c_1; \\ \varepsilon, & m = c_1; \\ 0, & m < c_1. \end{cases}$$

Оптимальный критерий является рандомизированным. Константы c_1 и ε нужно выбирать из условия: $\alpha(\varphi^*) = \alpha_0$. Воспользуемся асимптотикой:

$$\alpha(\varphi^*) = P_0\{m > c_1\} + \varepsilon P_0\{m = c_1\} \xrightarrow{n \rightarrow \infty} \xrightarrow{n \rightarrow \infty} 1 - \Phi\left(\frac{c_1 - np_0}{\sqrt{np_0(1-p_0)}}\right).$$

Решением уравнения

$$\Phi\left(\frac{c_1 - np_0}{\sqrt{np_0(1-p_0)}}\right) = 1 - \alpha_0$$

является квантиль стандартного нормального распределения уровня $1 - \alpha_0$:

$$z_{1-\alpha_0} = \frac{c_1 - np_0}{\sqrt{np_0(1-p_0)}}.$$

Из полученного условия можно найти константу c_1 :

$$c_1 = np_0 + \sqrt{np_0(1-p_0)}z_{1-\alpha_0}.$$

Критическая область для гипотезы H_0 имеет вид $(c_1; \infty)$. Таким образом, построен статистический критерий:

- Если $m \in (np_0 + \sqrt{np_0(1-p_0)}z_{1-\alpha_0}; +\infty)$, то гипотеза H_0 отклоняется.
- Если $m \in [0; np_0 + \sqrt{np_0(1-p_0)}z_{1-\alpha_0}]$, то гипотеза H_0 принимается.

Построенный критерий является нерандомизированным, однако, в отличие от предыдущих примеров, построенный критерий не является точным. Нельзя утверждать, что вероятность ошибки первого рода равна α_0 . Для выборок большого объема вероятность ошибки первого рода окажется приближенно равной α_0 .

§ 3. ГИПОТЕЗЫ О ПАРАМЕТРАХ РАСПРЕДЕЛЕНИЙ ДЛЯ СЛОЖНЫХ АЛЬТЕРНАТИВ

Пусть заданы генеральная совокупность ξ и выборка $X_{[n]}$ из генеральной совокупности. Сформулируем две гипотезы о распределении генеральной совокупности: генеральная совокупность ξ подчиняется нормальному распределению $N(a_0, \sigma^2)$, и генеральная совокупность ξ подчиняется нормальному распределению $N(a_1, \sigma^2)$, где a_1 неизвестно, a_0 известно, также считаем, что σ^2 известна. Будем предполагать, что $a_1 > a_0$. Таким образом, имеем две гипотезы:

- $H_0 : a = a_0$.
- $H_1 : a > a_0$.

Гипотезу H_1 можно записать в виде: $a = a_1 > a_0$, a_1 неизвестно. Гипотеза H_1 представляет собой правостороннюю альтернативу. Оптимальный критерий имеет вид:

$$\varphi^*(x) = \begin{cases} 1, & \bar{X} > c_1; \\ 0, & \bar{X} \leq c_1, \end{cases}$$

где c_1 находится из уравнения:

$$\alpha_0 = P_0\{\bar{X} > c_1\},$$

где α_0 — вероятность ошибки первого рода (уровень значимости критерия). Как показано в предыдущем параграфе:

$$c_1 = a_0 + z_{1-\alpha_0}\sigma/\sqrt{n}.$$

Критерий Неймана–Пирсона является равномерно наиболее мощным критерием для проверки гипотезы $H_0 : a = a_0$ при альтернативе $H_1 : a > a_0$, то есть он не зависит от конкретного значения a_1 и является наиболее мощным при любом $a_1 > a_0$.

Результат полностью сохраняется для левосторонней альтернативы $a = a_1 < a_0$ при соответствующих изменениях. Для двусторонней альтернативы не удастся построить равномерно наиболее мощный критерий.

Рассмотрим двустороннюю альтернативу:

- $H_0 : a = a_0$.
- $H_1 : a \neq a_0$.

Как и раньше фиксируем вероятность ошибки первого рода α_0 . В качестве статистики критерия возьмем \bar{X} . В предположении истинности нулевой гипотезы выберем константу c_1 из условия:

$$P_0\{|\bar{X} - a_0| > c_1\} = \alpha_0.$$

Критическая область для гипотезы H_0 при двусторонней альтернативе H_1 при использовании статистики \bar{X} имеет вид:

$$S = \{x : |x - a_0| > c_1\}.$$

Выберем вероятность ошибки первого рода:

$$P\{\bar{X} \in S/H_0\} = \alpha_0.$$

Преобразуем выражение:

$$P_0\left\{\frac{\sqrt{n}|\bar{X} - a_0|}{\sigma} \geq \frac{\sqrt{nc_1}}{\sigma}\right\} = \alpha_0.$$

Пусть $\eta = \sqrt{n}(\bar{X} - a_0)/\sigma$, тогда при условии справедливости гипотезы H_0 случайная величина η подчиняется стандартному

нормальному распределению $N(0, 1)$. Следовательно, в качестве c_1 можно выбрать

$$c_1 = z_{1-\frac{\alpha_0}{2}} \frac{\sigma}{\sqrt{n}},$$

где $z_{1-\frac{\alpha_0}{2}}$ — квантиль стандартного нормального распределения уровня $1 - \alpha_0/2$. Критическая область для нулевой гипотезы при использовании статистики \bar{X} принимает следующий вид:

$$S = (-\infty; a_0 - c_1) \cup (a_0 + c_1; +\infty).$$

Критерий проверки таков:

- Если $\bar{X} \in S$, то гипотеза H_0 отклоняется.
- Если $\bar{X} \notin S$, то гипотеза H_0 принимается.

Теперь рассмотрим случай, когда σ^2 неизвестна. Рассмотрим правостороннюю альтернативу:

- $H_0 : a = a_0$.
- $H_1 : a > a_0$, то есть $a = a_1 > a_0$, a_1 — неизвестно.

Вычислим статистику

$$t = \frac{\bar{X} - a_0}{\tilde{s}} \sqrt{n},$$

где $\tilde{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$.

При справедливости нулевой гипотезы статистика t должна подчиняться распределению Стьюдента с $n - 1$ степенью свободы. Если верна альтернативная гипотеза H_1 , то можно заметить, что статистика будет смещена вправо по отношению к распределению Стьюдента с $n - 1$ степенью свободы.

Критическая область в данном случае для нулевой гипотезы будет иметь вид $S = \{t > c_1\} = (c_1; +\infty)$, где константу c_1 следует выбирать из условия $P_0\{t > c_1\} = \alpha_0$. Таким образом, $c_1 = t_{1-\alpha_0, n-1}$ — квантиль распределения Стьюдента с $n - 1$ степенью свободы уровня $1 - \alpha_0$. Критерий с вероятностью ошибки первого рода α_0 имеет вид:

- Если статистика $t \in (t_{1-\alpha_0, n-1}; +\infty)$, то отклоняем гипотезу H_0 в пользу H_1 .
- Если статистика $t \in (-\infty; t_{1-\alpha_0, n-1}]$, то отклоняем гипотезу H_1 в пользу H_0 .

Для левосторонней альтернативы критическая область для гипотезы H_0 при использовании статистики t будет следующей:

$$S = (-\infty; t_{\alpha_0, n-1}),$$

где $t_{\alpha_0, n-1}$ — квантиль уровня α_0 распределения Стюдента с $n - 1$ степенью свободы.

Для двусторонней альтернативы критическая область с вероятностью ошибки первого рода α_0 для гипотезы H_0 при использовании статистики t будет следующей:

$$S = (-\infty; t_{\frac{\alpha_0}{2}, n-1}] \cup [t_{1-\frac{\alpha_0}{2}, n-1}; +\infty),$$

где $t_{\frac{\alpha_0}{2}, n-1}$ — квантиль уровня $\alpha_0/2$ распределения Стюдента с $n - 1$ степенью свободы; $t_{1-\frac{\alpha_0}{2}, n-1}$ — квантиль уровня $1 - \alpha_0/2$ того же распределения.

§ 4. РАСПРЕДЕЛЕНИЕ ФИШЕРА

Определение 4.1. Пусть случайные величины $\eta \sim \chi_m^2$, $\xi \sim \chi_n^2$ независимы. Будем говорить, что случайная величина

$$\zeta = \frac{\eta/m}{\xi/n} \sim \mathcal{F}_{m,n}$$

подчиняется распределению Фишера со степенями свободы числителя m и знаменателя n .

Лемма 4.1. Плотность распределения случайной величины $\zeta \sim \mathcal{F}_{m,n}$ имеет вид:

$$f_\zeta(z) = \begin{cases} \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \frac{m^{\frac{m}{2}} n^{\frac{n}{2}} z^{\frac{m}{2}-1}}{(n+mz)^{\frac{m+n}{2}}}, & \text{если } z > 0; \\ 0, & \text{если } z \leq 0. \end{cases}$$

Доказательство. Рассмотрим случайную величину $\tilde{\zeta} = \eta/\xi$ и напомним для нее плотность распределения, считая,

что $z > 0$:

$$\begin{aligned} f_{\tilde{\zeta}}(z) &= \int_0^{\infty} x f_{\xi}(x) f_{\eta}(zx) dx = \\ &= \frac{z^{\frac{m}{2}-1}}{2^{\frac{n+m}{2}} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\frac{m+n}{2}-1} e^{-\frac{x}{2}(z+1)} dx. \end{aligned}$$

Далее сделаем замену переменных: $y = x(z+1)/2$, $dx = 2dy/(1+z)$, $x = 2y/(1+z)$. После преобразований получаем формулу плотности распределения для случайной величины $\tilde{\zeta}$:

$$f_{\tilde{\zeta}}(z) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \frac{z^{\frac{m}{2}-1}}{(1+z)^{\frac{m+n}{2}}}, \quad z > 0.$$

Как нетрудно заметить, $\zeta = (n/m)\tilde{\zeta}$, следовательно,

$$f_{\zeta}(z) = \frac{m}{n} f_{\tilde{\zeta}}\left(\frac{m}{n}z\right).$$

■

§ 5. КРИТЕРИИ ОДНОРОДНОСТИ ДЛЯ ДВУХ НЕЗАВИСИМЫХ ВЫБОРОК ИЗ НОРМАЛЬНО РАСПРЕДЕЛЕННОЙ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

Критерий Фишера

Важная задача прикладной статистики — сравнение двух и более выборок, что позволяют осуществить критерии однородности.

Пусть имеются две генеральные совокупности η с выборкой $X_{[m]}$ и ξ с выборкой $Y_{[n]}$. Будем считать, что η подчиняется нормальному распределению $N(a_1, \sigma_1^2)$ и ξ подчиняется нормальному распределению $N(a_2, \sigma_2^2)$, причем математические ожидания a_1 , a_2 и дисперсии σ_1^2 и σ_2^2 неизвестны. Сформулируем нулевую и альтернативную гипотезы:

- $H_0 : \sigma_1^2 = \sigma_2^2$.
- $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Альтернативная гипотеза является двусторонней. Тогда

$$(m-1)\tilde{s}_X^2/\sigma_1^2 \sim \chi_{m-1}^2,$$

$$(n-1)\tilde{s}_Y^2/\sigma_2^2 \sim \chi_{n-1}^2,$$

где

$$\tilde{s}_X^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2,$$

$$\tilde{s}_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

В соответствии с определением распределения Фишера:

$$F_{m-1, n-1} = \frac{\tilde{s}_X^2}{\tilde{s}_Y^2} \frac{\sigma_2^2}{\sigma_1^2} \sim \mathcal{F}_{m-1, n-1}.$$

Тогда при справедливости нулевой гипотезы:

$$\frac{\tilde{s}_X^2}{\tilde{s}_Y^2} \sim \mathcal{F}_{m-1, n-1}.$$

Пусть $u_{\frac{\alpha}{2}}, u_{1-\frac{\alpha}{2}}$ — квантили распределения Фишера $\mathcal{F}_{m-1, n-1}$. Можно сформулировать критерий проверки гипотезы H_0 при альтернативе H_1 с вероятностью ошибки первого рода α (уровнем значимости критерия):

- Если $\frac{\tilde{s}_X^2}{\tilde{s}_Y^2} \in [u_{\frac{\alpha}{2}}, u_{1-\frac{\alpha}{2}}]$, то принимается гипотеза H_0 .
- Если $\frac{\tilde{s}_X^2}{\tilde{s}_Y^2} \notin [u_{\frac{\alpha}{2}}, u_{1-\frac{\alpha}{2}}]$, то принимается гипотеза H_1 .

Если альтернативная гипотеза H_1 односторонняя, т. е. $\sigma_1^2 > \sigma_2^2$, то в качестве критической области для нулевой гипотезы рассматривается $S = (u_{1-\alpha}, +\infty)$, где α — вероятность ошибки первого рода. Случай $\sigma_2^2 > \sigma_1^2$ сводится к предыдущей переменной мест выборок.

Критерий Стьюдента

Пусть заданы случайные величины $\eta \sim N(a_1, \sigma_1^2)$ с выборкой $X_{[m]}$ и $\xi \sim N(a_2, \sigma_2^2)$ с выборкой $Y_{[n]}$. Сформулируем нулевую и альтернативную двустороннюю гипотезы:

- $H_0 : a_1 = a_2$.

- $H_1 : a_1 \neq a_2$.

Рассмотрим два случая:

1. Дисперсии σ_1^2, σ_2^2 известны, тогда используем статистику

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0, 1),$$

которая при условии, что верна гипотеза H_0 , подчиняется стандартному нормальному распределению. Получаем критерий с вероятностью ошибки первого рода α :

- Если справедливо неравенство:

$$\frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \leq u_{1-\frac{\alpha}{2}},$$

то принимается гипотеза H_0 , где $u_{1-\frac{\alpha}{2}}$ — квантиль уровня $1 - \alpha/2$ стандартного нормального распределения.

- Если справедливо неравенство:

$$\frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} > u_{1-\frac{\alpha}{2}},$$

то принимается гипотеза H_1 .

Для правосторонней альтернативной гипотезы $H_1 : a_1 > a_2$ критическая область для H_0 будет выглядеть следующим образом:

$$S = (u_{1-\alpha}, +\infty).$$

Для левосторонней альтернативной гипотезы $H_1 : a_1 < a_2$ критическая область для H_0 будет следующей:

$$S = (-\infty, u_\alpha).$$

2. Дисперсии неизвестны, но есть основания считать, что $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Статистики $(m-1)\tilde{s}_X^2/\sigma^2$, $(n-1)\tilde{s}_Y^2/\sigma^2$, которые были введены ранее, взаимно независимы и имеют распределения χ_{m-1}^2 и χ_{n-1}^2 соответственно, тогда статистика

$$\frac{\tilde{s}_X^2(m-1) + \tilde{s}_Y^2(n-1)}{\sigma^2}$$

подчиняется распределению хи-квадрат с $m + n - 2$ степенями свободы. Если верна гипотеза H_0 , то статистика

$$t_{n+m-2}^{(0)} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{m+n}{mn}} \sqrt{\frac{\bar{s}_X^2(m-1) + \bar{s}_Y^2(n-1)}{m+n-2}}}$$

подчиняется распределению Стьюдента T_{n+m-2} с $n + m - 2$ степенями свободы ($t_{n+m-2}^{(0)}$ — дробь Стьюдента).

Если верна гипотеза H_1 , то статистика

$$t_{n+m-2} = \frac{\bar{X} - \bar{Y} + (a_2 - a_1)}{\sqrt{\frac{m+n}{mn}} \sqrt{\frac{\bar{s}_X^2(m-1) + \bar{s}_Y^2(n-1)}{m+n-2}}}$$

подчиняется распределению Стьюдента T_{n+m-2} с $n + m - 2$ степенями свободы.

Сформулируем критерий с вероятностью ошибки первого рода α :

- Если $|t_{n+m-2}^{(0)}| \leq t_{1-\frac{\alpha}{2}}$, где $t_{1-\frac{\alpha}{2}}$ — квантиль распределения Стьюдента с $n + m - 2$ степенями свободы уровня $1 - \alpha/2$, то принимается гипотеза H_0 .
- Если $|t_{n+m-2}^{(0)}| > t_{1-\frac{\alpha}{2}}$, то принимается гипотеза H_1 .

Для правосторонней альтернативной гипотезы $H_1 : a_1 > a_2$ критическая область для H_0 при использовании статистики $t_{n+m-2}^{(0)}$ будет следующей: $S = (t_{1-\alpha}, +\infty)$. Для левосторонней альтернативной гипотезы $H_1 : a_1 < a_2$ критическая область для H_0 при использовании статистики $t_{n+m-2}^{(0)}$ будет следующей: $S = (-\infty, t_\alpha)$. Границы $t_{1-\alpha}$ и t_α — квантили распределения Стьюдента с $n + m - 2$ степенями свободы уровней $1 - \alpha$ и α соответственно.

§ 6. КРИТЕРИЙ ОДНОРОДНОСТИ ВИЛКОКСОНА И МАННА–УИТНИ

Критерии Вилкоксона и Манна–Уитни неприменимы к дискретным распределениям, и для применения этих критериев не нужно предполагать нормальность распределений генеральных совокупностей.

Пусть имеются две независимые выборки $X_{[n]} = (X_1, \dots, \dots, X_n)$ и $Y_{[m]} = (Y_1, \dots, Y_m)$ из двух генеральных совокупностей с непрерывными функциями распределения равными соответственно F и G . Сформулируем гипотезы:

- $H_0 : F(x) = G(x)$ для всех $x \in \mathbb{R}$.
- $H_1 : F(x) \geq G(x)$ для всех $x \in \mathbb{R}$ — правосторонняя альтернативная гипотеза.
- $H'_1 : F(x) \leq G(x)$ для всех $x \in \mathbb{R}$ — левосторонняя альтернативная гипотеза.
- $H''_1 : F(x) \neq G(x)$ для всех $x \in \mathbb{R}$ — двусторонняя альтернативная гипотеза (т.е. выполнена гипотеза H_1 или H'_1).

Критерий однородности Вилкоксона

Без ограничения общности будем считать, что $m \leq n$. Составим объединенную выборку $Z_{[n+m]} = (X_{[n]}, Y_{[m]})$. Построим вариационный ряд объединенной выборки:

$$z_{(1)} < z_{(2)} < \dots < z_{(m+n)},$$

т.е. упорядочим все элементы объединенной выборки по возрастанию. Элементы получившегося вариационного ряда будем обозначать $z_{(1)} < z_{(2)} < \dots < z_{(m+n)}$. Если распределения генеральных совокупностей непрерывны, то совпадения возможны только с нулевой вероятностью. В дальнейшем будем предполагать, что совпадений нет. В литературе [3, 32] имеются поправочные формулы, когда совпадения возникают, например, вследствие округления. Однако следует заметить, что при большом числе совпадений рассматриваемые критерии неприменимы.

Найдем, какие места занимают в вариационном ряду, построенном по объединенной выборке, элементы выборки $Y_{[m]}$. Назовем эти номера *рангами* элементов выборки $Y_{[m]}$ в объединенной выборке $Z_{[n+m]}$:

$$\text{rank}(Y_1) = s_1, \text{rank}(Y_2) = s_2, \dots, \text{rank}(Y_m) = s_m.$$

Рассмотрим статистику критерия:

$$W = \sum_{i=1}^m s_i.$$

Очевидно, что минимальным значением статистики Вилкоксона может быть величина $W_{\min} = m(m+1)/2$, максимальным — $W_{\max} = mn + m(m+1)/2$. Поэтому статистика W находится в промежутке $[m(m+1)/2; mn + m(m+1)/2]$. Распределение статистики Вилкоксона W является симметричным относительно середины данного промежутка при условии справедливости нулевой гипотезы H_0 .

Если справедлива альтернативная гипотеза H_1 , то чаще будут встречаться события $x_i < y_j$, то есть распределение статистики W перестанет быть симметричным относительно середины и будет сдвинуто вправо. Если справедлива альтернативная гипотеза H'_1 , то распределение статистики W будет сдвинуто влево, так как чаще будут выполняться события $x_i > y_j$.

Если выбрана альтернативой гипотеза H_1 , то критическая область для нулевой гипотезы H_0 будет иметь вид:

$$S = \left[c_1, mn + \frac{m(m+1)}{2} \right].$$

Если выбрана альтернативой гипотеза H'_1 , то критическая область для нулевой гипотезы H_0 будет иметь вид:

$$S = \left[\frac{m(m+1)}{2}, c_2 \right].$$

Если выбрана альтернативой гипотеза H''_1 , то критическая область для нулевой гипотезы H_0 будет иметь вид:

$$S = \left[\frac{m(m+1)}{2}, c_3 \right] \cup \left[c_4, mn + \frac{m(m+1)}{2} \right].$$

При этом константы c_1, c_2, c_3, c_4 следует находить по таблицам распределения статистики Вилкоксона W , рассчитанной при условии справедливости нулевой гипотезы H_0 для

разных m и n . В качестве искоемых констант выбираются квантили распределения. При этом константы c_1 и c_2 симметричны относительно середины промежутка $[m(m+1)/2, mn + m(m+1)/2]$. Также симметрично относительно середины этого промежутка расположены константы c_3 и c_4 . Общее требование заключается в том, что вероятность попадания статистики W в критическую область при условии справедливости нулевой гипотезы H_0 должна быть равна заданному значению α :

$$P_0\{W \in S\} = \alpha.$$

Числовые значения констант могут быть найдены из статистических таблиц [3] и в приложении Д книги.

Критерий Манна–Уитни

Будем проверять те же нулевую и альтернативные гипотезы, что и в критерии Уилкоксона. Запишем статистику критерия Манна–Уитни:

$$U = \sum_{i=1}^n \sum_{j=1}^m I\{X_i < Y_j\},$$

где

$$I\{X_i < Y_j\} = \begin{cases} 1, & X_i < Y_j; \\ 0, & X_i > Y_j. \end{cases}$$

Находим в таблице [34] критические значения распределения статистики Манна–Уитни: $U_{left} = U_{\alpha,n,m}$, $U_{right} = V_{\alpha,n,m}$ — и сравниваем с ними статистику критерия. Здесь $U_{\alpha,n,m}$ и $V_{\alpha,n,m}$ — критические значения уровня α статистики U при условии справедливости гипотезы однородности H_0 , т. е. $P_0\{U \leq U_{\alpha,n,m}\} = \alpha$, $P_0\{U \geq V_{\alpha,n,m}\} = \alpha$, α — вероятность ошибки первого рода.

Если $U \geq U_{right}$, то нулевая гипотеза отвергается в пользу правосторонней альтернативной гипотезы H_1 . Если $U \leq U_{left}$, то нулевая гипотеза отвергается в пользу левосторонней альтернативной гипотезы H'_1 . Для случая двусторонней альтернативы H''_1 находятся точки $U_{right} = V_{\alpha/2,n,m}$ и $U_{left} = U_{\alpha/2,n,m}$,

и нулевая гипотеза H_0 отвергается при выполнении любого из неравенств: $U \geq U_{right}$, $U \leq U_{left}$.

Статистики Вилкоксона и Манна–Уитни связаны между собой следующим соотношением:

$$W = U + \frac{m(m+1)}{2}.$$

Для доказательства справедливости приведенной формулы заметим, что значение статистики Вилкоксона W по существу равно количеству неравенств вида $\{X_i < Y_j\}$ плюс сумма рангов элементов выборки $Y_{[m]}$ в самой выборке $Y_{[m]}$. Последняя сумма представляет собой арифметическую прогрессию чисел от 1 до m шагом 1.

Полученная формула позволяет пересчитывать значение одной статистики в другую и пользоваться тем распределением, которое удобнее для вычислений.

Для больших объемов выборок существуют аппроксимации статистики W , которую можно найти, например, в [3].

§ 7. НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ АНАЛИЗА ПАРНЫХ ПОВТОРНЫХ НАБЛЮДЕНИЙ

В отличие от предыдущего параграфа здесь рассматриваются две зависимые выборки. При этом проверяются те же гипотезы. Выборки имеют одинаковый объем, и зависимость носит следующий характер. Внутри каждой из выборок элементы независимы, но X_i и Y_i — зависимые наблюдения. Чаше всего i означает номер объекта, а X_i и Y_i — два наблюдения над одним и тем же объектом до и после некоторого воздействия.

Пусть имеются две выборки $X_{[n]} = (X_1, \dots, X_n)$ и $Y_{[n]} = (Y_1, \dots, Y_n)$ из генеральных совокупностей с функциями распределения равными соответственно $F(x)$ и $G(x)$, которые считаем непрерывными. Сформулируем гипотезы:

- H_0 : $F(x) = G(x)$ для всех $x \in \mathbb{R}$.
- H_1 : $F(x) \geq G(x)$ для всех $x \in \mathbb{R}$ — правосторонняя альтернативная гипотеза.

- $H_1' : F(x) \leq G(x)$ для всех $x \in \mathbb{R}$ — левосторонняя альтернативная гипотеза.
- $H_1'' : F(x) \neq G(x)$ для всех $x \in \mathbb{R}$ — двусторонняя альтернативная гипотеза.

Критерий знаков

Составим разности $z_i = X_i - Y_i$. Случайные величины z_i , $i = 1, \dots, n$, взаимно независимы и одинаково распределены. Переформулируем гипотезы:

- $H_0 : P\{z_i < 0\} = P\{z_i > 0\} = 1/2$.
- $H_1 : P\{z_i < 0\} > P\{z_i > 0\}$.
- $H_1' : P\{z_i < 0\} < P\{z_i > 0\}$.
- $H_1'' : P\{z_i < 0\} \neq P\{z_i > 0\}$.

Так как в гипотезах фигурируют вероятности, то можем рассматривать схему Бернулли, где событие $z_i < 0$ означает успех. Рассмотрим статистику:

$$L = \sum_{i=1}^n I\{z_i < 0\}.$$

Очевидно, что при выполнении гипотезы H_0 статистика L подчиняется биномиальному распределению с вероятностью успеха $p = 1/2$. Критическая область для нулевой гипотезы при выборе альтернативной гипотезы H_1 имеет вид: $(c_1, n]$. Критическая область для нулевой гипотезы при выборе альтернативной гипотезы H_1' имеет вид: $[0, c_2)$. Критическая область для нулевой гипотезы при выборе альтернативной гипотезы H_1'' имеет вид: $[0, c_3) \cup (c_4, n]$.

Для нахождения констант c_1, c_2, c_3, c_4 можно использовать таблицы вероятностей биномиального распределения. Константы c_1 и c_2 расположены симметрично относительно середины промежутка $[0, n]$, то же относится и к константам c_3 и c_4 . Общее правило, как всегда, заключается в том, что вероятность попадания статистики критерия L в критическую область при условии выполнения гипотезы H_0 равна α .

*Критерий знаковых ранговых сумм
Вилкоксона*

Составим разности $z_i = X_i - Y_i$. Предполагаем, что величины z_i не зависят друг от друга. Рассмотрим случаи, когда $z_i < 0$ и $z_i > 0$. Рассмотрим те же гипотезы:

- $H_0 : P\{z_i < 0\} = P\{z_i > 0\} = 1/2$.
- $H_1 : P\{z_i < 0\} > P\{z_i > 0\}$.
- $H'_1 : P\{z_i < 0\} < P\{z_i > 0\}$.
- $H''_1 : P\{z_i < 0\} \neq P\{z_i > 0\}$.

Построим вариационный ряд из модулей разностей: $|z_1|, \dots, |z_n|$. Сопоставим каждому элементу вариационного ряда ранг: $s_1 = \text{rank}(|z_1|), \dots, s_n = \text{rank}(|z_n|)$, т. е. номер в вариационном ряду. Составим ранговую статистику

$$U = \sum_{i=1}^n \Psi_i s_i,$$

где

$$\Psi_i = \begin{cases} 1, & z_i > 0; \\ 0, & z_i < 0. \end{cases}$$

Составлены статистические таблицы [34] распределения статистики U при условии справедливости гипотезы H_0 .

Аналогично предыдущему критерию используем таблицы [34] для нахождения критических значений $t_{\alpha, n}$ и сравнения их с полученным значением статистики U . При выборе левосторонней альтернативы H_1 критическая область имеет вид: $[0, c_1]$. При выборе правосторонней альтернативы H'_1 критическая область имеет вид: $[c_2, n(n+1)/2]$. При выборе двусторонней альтернативы H''_1 критическая область имеет вид: $[0, c_3] \cup [c_4, n(n+1)/2]$. Константы c_1 и c_2 расположены симметрично относительно середины промежутка $[0, n(n+1)/2]$, тоже относится и к константам c_3 и c_4 . Вероятность попадания статистики критерия U в критическую область при условии выполнения гипотезы H_0 равна α .

§ 8. ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Однофакторный дисперсионный анализ является обобщением T -критерия для двух выборок из генеральной совокупности.

Пусть число выборок $k \geq 2$:

$$\begin{array}{cccc} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & & \vdots \\ X_{n_1 1} & X_{n_2 2} & \dots & X_{n_k k} \\ N(a_1, \sigma^2) & N(a_2, \sigma^2) & \dots & N(a_k, \sigma^2) \end{array}.$$

Пусть все выборки взаимно независимы между собой, при этом выборка $(X_{1i}, \dots, X_{n_i i})$ взята из генеральной совокупности с распределением $N(a_i, \sigma^2)$. Элементы каждой выборки тоже, конечно, взаимно независимы.

Выдвигается гипотеза $H_0 : a_1 = a_2 = \dots = a_k = a$ при альтернативной гипотезе H_1 , которая заключается в отрицании гипотезы H_0 . Рассмотрим следующие величины:

$$\bar{X}_{\cdot j} = \frac{1}{n_j} \sum_{i=1}^{n_j} X_{ij}, \quad \bar{X}_{\cdot \cdot} = \frac{1}{N} \sum_{i=1}^{n_j} \sum_{j=1}^k X_{ij}, \quad N = \sum_{j=1}^k n_j,$$

где j — номер выборки.

Независимо от справедливости гипотезы H_0 :

$$\sum_{i=1}^{n_j} \frac{(X_{ij} - \bar{X}_{\cdot j})^2}{\sigma^2} \sim \chi_{n_j-1}^2 \sim G\left(\frac{1}{2}, \frac{n_j-1}{2}\right).$$

Следовательно,

$$S_1 = \sum_{i=1}^{n_j} \sum_{j=1}^k \frac{(X_{ij} - \bar{X}_{\cdot j})^2}{\sigma^2} \sim \chi_{N-k}^2.$$

Рассмотрим величину

$$\frac{(\bar{X}_{\cdot j} - a_j)}{\sigma} \sqrt{n_j} \sim N(0, 1).$$

Пусть гипотеза H_0 верна. Рассмотрим статистику S_2 следующего вида:

$$S_2 = \frac{1}{\sigma^2} \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - \bar{X}_{..})^2.$$

Прибавим и вычтем величину a , получим:

$$\begin{aligned} \frac{1}{\sigma^2} \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - \bar{X}_{..})^2 &= \\ &= \frac{1}{\sigma^2} \sum_{j=1}^k n_j \left[(\bar{X}_{\cdot j} - a) - \frac{1}{N} \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - a) \right]^2, \end{aligned}$$

так как $\bar{X}_{..} - a = \frac{1}{N} \sum_{i=1}^{n_j} \sum_{j=1}^k (X_{ij} - a) = \frac{1}{N} \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - a)$.

$$\begin{aligned} S_2 &= \frac{1}{\sigma^2} \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - \bar{X}_{..})^2 = \\ &= \frac{1}{\sigma^2} \sum_{j=1}^k n_j \left[(\bar{X}_{\cdot j} - a) - \frac{1}{N} \sum_{j=1}^k n_j (\bar{X}_{\cdot j} - a) \right]^2 = \\ &= \sum_{j=1}^k \frac{(\sqrt{n_j}(\bar{X}_{\cdot j} - a))^2}{\sigma^2} - \frac{\left[\sum_{j=1}^k \sqrt{\frac{n_j}{N}} \sqrt{n_j}(\bar{X}_{\cdot j} - a) \right]^2}{\sigma^2} = \\ &= \sum_{j=1}^k \left[\frac{\sqrt{n_j}(\bar{X}_{\cdot j} - a)}{\sigma} \right]^2 - \left[\sum_{j=1}^k \sqrt{\frac{n_j}{N}} \frac{\sqrt{n_j}(\bar{X}_{\cdot j} - a)}{\sigma} \right]^2 \sim \chi_{k-1}^2, \end{aligned}$$

как следует из леммы 2.3 главы 16. Из леммы 2.2 главы 16 следует, что статистики S_1 и S_2 взаимно независимы. По определению распределения Фишера $\mathcal{F}_{k-1, N-k}$ получаем, что при выполнении гипотезы H_0 статистика

$$F = \frac{S_2/(k-1)}{S_1/(N-k)} \sim \mathcal{F}_{k-1, N-k}$$

подчиняется распределению Фишера с $k-1$ и $N-k$ степенями свободы числителя и знаменателя соответственно. Заметим, что статистика F не зависит от σ^2 . Большие значения статистики F свидетельствуют против нулевой гипотезы, поэтому критическая область S для H_0 с вероятностью ошибки первого рода α имеет вид: $S = (u_{1-\alpha, k-1, N-k}; \infty)$, где $u_{1-\alpha, k-1, N-k}$ — квантиль уровня $1 - \alpha$ распределения Фишера $\mathcal{F}_{k-1, N-k}$.

§ 9. КРИТЕРИЙ ОДНОРОДНОСТИ КОЛМОГОРОВА–СМИРНОВА

Пусть имеется две выборки $X_{[n]} = \{X_1, \dots, X_n\}$ и $Y_{[m]} = \{Y_1, \dots, Y_m\}$ из генеральных совокупностей ξ и η соответственно. По выборке X_1, \dots, X_n построим вариационный ряд $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, а по Y_1, \dots, Y_m построим вариационный ряд $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(m)}$. Объемы выборок могут быть различными, но, не нарушая общности, предположим, что $m \leq n$. Функции распределения этих генеральных совокупностей равны $F(x)$ и $G(x)$ соответственно. Наложим дополнительное ограничение: функции распределения $F(x)$ и $G(x)$ непрерывны.

Критерий Колмогорова–Смирнова проверяет гипотезу о равенстве функций распределения двух генеральных совокупностей ξ и η , из которых извлечены выборки $X_{[n]}$ и $Y_{[m]}$ соответственно, т. е. $H_0 : F(x) = G(x)$ для всех $x \in \mathbb{R}$ при альтернативной гипотезе $H_1 : F(x) \neq G(x)$. Этот критерий основан на использовании эмпирических функций распределения $F_n^*(x)$ и $G_m^*(x)$, построенных по первой и второй выборкам соответственно.

Статистика Смирнова определяется следующей формулой:

$$D_{m,n} = \sup_{|x| < \infty} |G_m^*(x) - F_n^*(x)|. \quad (9.1)$$

Теорема 9.1. Введем обозначение

$$D_{m,n} = \sup_{x \in \mathbb{R}} |G_m^*(x, Y_{[m]}) - F_n^*(x, X_{[n]})|.$$

Тогда

$$P_0 \left\{ \sqrt{\frac{mn}{m+n}} D_{m,n} \leq z \right\} \rightarrow K(z) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2 z^2}, \quad (9.2)$$

если истинная функция распределения $F_0(x) = F(x) = G(x)$ непрерывна.

На практике значение статистики $D_{m,n}$ рекомендуется вычислять по формулам:

$$D_{m,n}^+ = \max_{1 \leq r \leq m} \left[\frac{r}{m} - F_n^*(y(r)) \right] = \max_{1 \leq s \leq n} \left[G_m^*(x(s)) - \frac{s-1}{n} \right], \quad (9.3)$$

$$D_{m,n}^- = \max_{1 \leq r \leq m} \left[F_n^*(y(r)) - \frac{r-1}{m} \right] = \max_{1 \leq s \leq n} \left[\frac{s}{n} - G_m^*(x(s)) \right], \quad (9.4)$$

$$D_{m,n} = \max(D_{m,n}^+, D_{m,n}^-). \quad (9.5)$$

При справедливости нулевой гипотезы и неограниченном увеличении объемов выборок исправленная статистика

$$\sqrt{\frac{mn}{m+n}} D_{m,n} \quad (9.6)$$

асимптотически подчиняется распределению Колмогорова с функцией распределения $K(z)$ из правой части (9.2). Таблицу значений функции распределения Колмогорова можно найти в [3]. Критическая область для гипотезы H_0 при использовании статистики (9.6) имеет вид: $S = (k_{1-\alpha}, \infty)$, где $k_{1-\alpha}$ — квантиль уровня $1 - \alpha$ распределения Колмогорова (9.2).

§ 10. КОЭФФИЦИЕНТЫ РАНГОВОЙ КОРРЕЛЯЦИИ СПИРМЕНА И КЕНДЕЛЛА

Рассматриваемые в этом параграфе коэффициенты вычисляются только для порядковых шкал.

Определение 10.1. Шкалы, в которых существенен лишь взаимный порядок, в котором следуют результаты измерений, а не их количественные значения, называют порядковыми или ординальными шкалами.

Пусть имеется два признака A и B , между которыми мы хотим установить наличие зависимости или независимости. Пусть (X_1, \dots, X_n) — измерение степени выраженности признака A , (Y_1, \dots, Y_n) — измерение степени выраженности признака B . Каждый объект характеризует пара (X_i, Y_j) , $1 \leq i \leq n, 1 \leq j \leq n$.

Коэффициент ранговой корреляции Спирмена

Проранжируем наблюдения X_1, \dots, X_n и Y_1, \dots, Y_n , ранги которых будут соответственно обозначаться r_1, \dots, r_n и s_1, \dots, s_n , то есть r_i — номер наблюдения X_i в вариационном ряду, построенном по наблюдениям X_1, \dots, X_n . Аналогично s_i — номер наблюдения Y_i в вариационном ряду, построенном по наблюдениям Y_1, \dots, Y_n . Будем предполагать, что в выборках нет повторяющихся элементов.

Рассмотрим статистику Спирмена $S = \sum_{i=1}^n (s_i - r_i)^2$. Вычислим коэффициент ранговой корреляции Спирмена:

$$\varrho = 1 - \frac{6S}{n^3 - n}.$$

Нетрудно показать, что $|\varrho| \leq 1$. Если $|\varrho| = 1$, то это означает полную зависимость одного признака от другого, либо, иначе говоря, полную предсказуемость одной выборки по другой. Если ранги признаков совпадают, то $\varrho = 1$. Если последовательности рангов полностью противоположны, то $\varrho = -1$. Оба случая означают полную предсказуемость одной ранговой последовательности по другой, или, другими словами, полную зависимость признаков A и B .

Сформулируем гипотезы:

- H_0 : признаки A и B взаимно независимы.
- H_1 : имеется монотонная положительная связь признаков.
- H'_1 : имеется монотонная отрицательная связь признаков.
- H''_1 : имеется монотонная связь признаков.

Гипотеза H_0 соответствует отсутствию взаимосвязи между признаками или, иначе говоря, независимости признаков. Если гипотеза H_0 справедлива, то распределение статистики ϱ симметрично и концентрируется около нуля. При наличии зависимости распределение окажется другим. Для монотонной положительной зависимости распределение ϱ сдвинуто вправо, для монотонной отрицательной — влево.

Для проверки гипотезы H_0 необходимо обратиться к таблицам распределения коэффициента Спирмена [3], вычисленным в предположении истинности гипотезы H_0 . По заданной вероятности ошибки первого рода α необходимо найти соответствующие пороговые значения, после чего при попадании вычисленного по наблюдениям коэффициента ϱ в критическую область следует отклонить гипотезу H_0 в пользу альтернативной гипотезы. При выборе в качестве альтернативы гипотезы H_1 (положительная монотонная связь) критическую область следует выбрать в виде: $(c_1, 1]$. При выборе альтернативы гипотезы H'_1 (отрицательная монотонная связь) критическую область следует выбрать в виде: $[-1, c_2)$. При выборе альтернативной гипотезы H''_1 критическая область для гипотезы H_0 имеет вид: $[-1, c_3) \cup (c_4, 1]$. Пороговые значения c_1, c_2, c_3, c_4 определяются из статистических таблиц так, чтобы вероятность попадания в критическую область при выполнении гипотезы H_0 была равна α .

Коэффициент ранговой корреляции Кенделла

Коэффициент ранговой корреляции Кенделла применяется для решения тех же задач, что и коэффициент Спирмена. Проверяются те же гипотезы, и алгоритм проверки такой же, как и при использовании коэффициента Спирмена. Разница лишь в том, что используются статистические таблицы

для распределения коэффициента Кендалла [3]. Для вычисления статистики Кендалла достаточно посчитать количество инверсий (минимальное число перестановок соседних объектов), которое надо сделать для того, чтобы одно упорядочение объектов превратилось в другое.

Пусть есть пары наблюдений каждого из признаков $(X_1, Y_1), \dots, (X_n, Y_n)$. Упорядочим наблюдения первого признака и пронумируем их рангами от 1 до n . Затем ранжируем последовательность наблюдений второго признака, при этом объекты перенумерованы в соответствии с рангами первой совокупности. Пусть во втором наборе приписаны каждому наблюдению ранги z_1, \dots, z_n , то есть теперь все объекты характеризуются парами рангов: $(1, z_1), \dots, (n, z_n)$. После перенумерования ранги измерений признаков A представляют собой новые номера самих объектов. Пусть R — число инверсий в выборке $\{z_1, \dots, z_n\}$. Инверсия суть нарушение порядка. Например, в ряду $(4, 3, 1, 2)$ имеется всего 5 инверсий.

Рассмотрим коэффициент ранговой корреляции Кендалла:

$$\tau = 1 - \frac{4R}{n(n-1)}.$$

Нетрудно доказать, что $|\tau| \leq 1$. При этом $|\tau| = 1$ означает полную предсказуемость (зависимость) признаков. При больших n при справедливости гипотезы H_0 случайные величины $\sqrt{n-1}\rho$ и $\tau\sqrt{9n(n+1)/(2(2n+5))}$ приближенно распределены по стандартному нормальному закону $N(0, 1)$, что позволяет проверять гипотезу H_0 , пользуясь указанной асимптотикой.

Для обоих коэффициентов корреляции характерно то обстоятельство, что они обнаруживают лишь монотонную зависимость признаков.

§ 11. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ПИРСОНА

Коэффициент корреляции Пирсона применяется к данным, измеренным в шкале отношений.

Определение 11.1. Шкалой отношений называют такую шкалу с непрерывным множеством числовых значений, в которой

о двух сопоставляемых объектах можно сказать не только, одинаковы они или различны, не только, в каком из них признак выражен сильнее, но и во сколько раз сильнее этот признак выражен.

Предположим, что есть генеральная совокупность, каждый элемент которой обладает двумя количественными признаками. Если случайным образом извлекать объекты, то пусть ξ — значение, которое принимает первый признак, η — значение, которое принимает второй признак. Величины ξ и η — случайные. Корреляция случайных величин ξ и η выражается следующей формулой:

$$\rho(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}}.$$

Если случайные величины ξ и η независимы, то корреляция равна нулю. Обратное утверждение, вообще говоря, неверно. Если $\rho = 1$, то существует положительная линейная связь между величинами ξ и η такая, что $\eta = a + b\xi$, $b > 0$. Если $\rho = -1$, то существует линейная связь между величинами ξ и η такая, что $\eta = a + b\xi$, $b < 0$. Если вектор $(\xi, \eta)^T$ подчиняется совместному нормальному распределению с вектором математических ожиданий $a = (a_1, a_2)^T$ и ковариационной матрицей

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix},$$

то корреляция случайных величин ξ и η равна нулю тогда и только тогда, когда эти случайные величины взаимно независимы, $\sigma_1^2 = D\xi$, $\sigma_2^2 = D\eta$.

Получим оценку коэффициента корреляции — выборочный коэффициент корреляции Пирсона, который определяется выражением:

$$r_{X,Y} = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{s_X s_Y},$$

где $s_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$, $s_Y^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$. Здесь предполагается, что задана двумерная выборка: $(X_1, Y_1), \dots, (X_n, Y_n)$. Совместное распределение случайных величин ξ и η является нормальным.

Сформулируем гипотезы:

- $H_0: \rho = 0$ — гипотеза о независимости.
- $H_1: \rho > 0$.
- $H'_1: \rho < 0$.
- $H''_1: \rho \neq 0$ — двусторонняя альтернатива.

Справедливо утверждение. При сделанных предположениях о распределении случайного вектора $(\xi, \eta)^T$, статистика

$$t = r_{X,Y} \sqrt{n-2} / \sqrt{1-r^2} \quad (11.1)$$

при выполнении гипотезы H_0 подчиняется распределению Стьюдента с $n-2$ степенями свободы.

При использовании статистики (11.1) для альтернативы H_1 критическая область для гипотезы H_0 имеет вид: $S = (t_{1-\alpha, n-2}, \infty)$, для альтернативы H'_1 критическая область для гипотезы H_0 имеет вид: $S = (-\infty, t_{\alpha, n-2})$. Критическая область для нулевой гипотезы H_0 при альтернативе H''_1 будет иметь вид:

$$S = (-\infty, t_{\frac{\alpha}{2}, n-2}) \cup (t_{1-\frac{\alpha}{2}, n-2}, +\infty),$$

где $t_{\beta, n-2}$ — квантиль уровня β распределения Стьюдента с $n-2$ степенями свободы. Если значение статистики $t \in S$, то гипотеза H_0 отклоняется, если $t \notin S$, то гипотеза H_0 принимается. Величина вероятности ошибки первого рода равна α .

§ 12. КРИТЕРИЙ СОГЛАСИЯ ХИ-КВАДРАТ ДЛЯ СЛОЖНЫХ ГИПОТЕЗ

В критерии согласия хи-квадрат реализуется следующая схема. Выдвигаются гипотезы:

- $H_0: F_\xi(x) \equiv F(x)$ — нулевая гипотеза.
- $H_1: F_\xi(x) \neq F(x)$ — альтернативная гипотеза.

В прикладных задачах, как правило, известна не сама функция распределения, а параметрическое семейство, которому

она принадлежит:

$$\{F(\cdot/\theta) : \theta \in \Theta \subset \mathbb{R}^l\}.$$

Таким образом, проверяемая гипотеза принимает вид:

$$H_0 : F_\xi \in \{F(\cdot/\theta) : \theta \in \Theta \subset \mathbb{R}^l\}.$$

Альтернативная гипотеза H_1 примет вид: гипотеза H_0 не верна.

Разобьем числовую ось на k промежутков: $\Delta_1, \dots, \Delta_k$ таким образом, что $\bigcup_i \Delta_i = \mathbb{R}$, $\Delta_i \cap \Delta_j = \emptyset$, $i \neq j$. Получаем набор частот: n_1, \dots, n_k , $\sum_{i=1}^k n_i = n$. Каждому промежутку $\Delta_1, \dots, \Delta_k$ сопоставим вероятности: $p_1(\theta), \dots, p_k(\theta)$. В монографии [47] приведена теорема Фишера.

Теорема 12.1 (теорема Фишера). Пусть Θ — открытое множество в \mathbb{R}^l . Пусть выполнены условия:

1. Для любого $\theta \in \Theta$: $\sum_{i=1}^k p_i(\theta) = 1$.
2. Для любого $\theta \in \Theta$: $p_i(\theta) > c > 0$ для любого $i = \overline{1, k}$.
3. Для любого $\theta \in \Theta$ существуют и непрерывны производные: $\partial p_i(\theta)/\partial \theta_j$, $\partial^2 p_i(\theta)/(\partial \theta_u \partial \theta_v)$ для любого $i = \overline{1, \dots, k}$, $u, v, j = \overline{1, \dots, l}$.
4. Для любого $\theta \in \Theta$ матрица $\left(\frac{\partial p_i(\theta)}{\partial \theta_j}\right)_{i=\overline{1, l}}$ имеет ранг l .

Пусть $\hat{\theta}$ — оценка, найденная методом максимального правдоподобия по выборке n_1, \dots, n_k , т.е. $\hat{\theta} = \arg \max_{\theta \in \Theta} L(\{n_i\}, \theta)$, где

$$L(\{n_i\}, \theta) = \frac{n!}{n_1! \cdot \dots \cdot n_k!} \prod_{i=1}^k p_i^{n_i}(\theta),$$

или $\hat{\theta}$ — оценка по методу минимума хи-квадрат:

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{i=1}^k \frac{(n_i - np_i(\theta))^2}{np_i(\theta)}.$$

Тогда, если гипотеза H_0 верна, то

$$\chi^2(\hat{\theta}) = \sum_{i=1}^k \frac{(n_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})} \xrightarrow[n \rightarrow \infty]{d} \zeta \sim \chi_{k-l-1}^2.$$

Критическая область для гипотезы H_0 при использовании статистики $\chi^2(\hat{\theta})$ имеет вид: $S = (u_{1-\alpha, k-l-1}, \infty)$, где $u_{1-\alpha, k-l-1}$ — квантиль уровня $1-\alpha$ распределения хи-квадрат с $k-l-1$ степенями свободы. Вероятность ошибки первого рода приближенно равна α .

§ 13. ТАБЛИЦЫ СОПРЯЖЕННОСТИ

Определение 13.1. В номинальных шкалах измерения представляют собой метки, обозначающие принадлежность измерения определенной градации измеряемого признака. Никаких содержательных соотношений кроме $x = y$ или $x \neq y$ между значениями в этих шкалах нет.

Для проверки независимости качественных признаков A и B , то есть признаков, измеряемых в номинальных шкалах, применяются таблицы сопряженности.

Пусть имеются два качественных признака A и B . Признак A имеет r градаций: A_1, \dots, A_r , признак B имеет s градаций B_1, \dots, B_s . По выборке из n случайно выбранных объектов можно составить таблицу сопряженности:

	B_1	B_2	\dots	B_s	
A_1	n_{11}	n_{12}	\dots	n_{1s}	m_1
A_2	n_{21}	n_{22}	\dots	n_{2s}	m_2
\dots	\dots	\dots	\dots	\dots	
A_r	n_{r1}	n_{r2}	\dots	n_{rs}	m_r
	n_1	n_2	\dots	n_s	n

где n_{ij} — количество элементов в выборке, обладающих одновременно свойствами A_i и B_j . Приведенная таблица называется *таблицей сопряженности*. Справедливы равенства:

$$\sum_{i=1}^r n_{ij} = n_j,$$

$$\sum_{j=1}^s n_{ij} = m_i,$$

$$\sum_{j=1}^s n_j = \sum_{i=1}^r m_i = n.$$

Пусть $p_i = P(A_i)$, $i = 1, \dots, r$ и $q_j = P(B_j)$, $j = 1, \dots, s$. При этом $\sum_{i=1}^r p_i = 1$, $\sum_{j=1}^s q_j = 1$. Признаки A и B называются независимыми, если при любых i и j выполняется равенство:

$$p_{ij} = P(A_i \cap B_j) = p_i q_j.$$

Очевидно, что $\sum_{i=1}^r p_{ij} = q_j$, $\sum_{j=1}^s p_{ij} = p_i$. Сформулируем гипотезу независимости и альтернативную ей гипотезу:

- H_0 : $P(A_i \cap B_j) = p_i q_j$ для любых i, j .
- H_1 : существует пара (i, j) такая, что $P(A_i \cap B_j) \neq p_i q_j$.

Гипотеза H_0 представляет собой гипотезу независимости двух признаков.

При анализе таблицы сопряженности речь идет о полиномиальной схеме. Построим функцию правдоподобия выборки:

$$L = \frac{n!}{\prod_{\substack{i=1, \dots, r \\ j=1, \dots, s}} n_{ij}!} \prod_{\substack{i=1, \dots, r \\ j=1, \dots, s}} (p_i q_j)^{n_{ij}}.$$

Найдем оценки максимального правдоподобия по выборке частот и воспользуемся теоремой 12.1.

Нетрудно заметить, что, максимизируя $\ln L$ по $p_1, \dots, p_r, q_1, \dots, q_s$ при ограничениях $\sum_{i=1}^r p_i = 1$ и $\sum_{j=1}^s q_j = 1$ методом неопределенных множителей Лагранжа, получим следующие оценки:

$$\hat{p}_i = \frac{m_i}{n}, \quad i = 1, \dots, r,$$

$$\hat{q}_j = \frac{n_j}{n}, \quad j = 1, \dots, s.$$

Статистика χ^2 для данной задачи после подстановки оценок метода максимального правдоподобия имеет вид:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - \frac{m_i n_j}{n})^2}{\frac{m_i n_j}{n}} \xrightarrow{d} \zeta \sim \chi_{(s-1)(r-1)}^2.$$

Число степеней свободы в предельном распределении хи-квадрат в соответствии с теоремой Фишера 12.1 вычисляется как $rs - (r - 1) - (s - 1) - 1 = (r - 1)(s - 1)$. Большие значения статистики хи-квадрат свидетельствуют против нулевой гипотезы H_0 .

Получаем критерий для проверки гипотезы H_0 :

- Если $\chi^2 > \chi_{\text{кр}}^2$, то отвергаем гипотезу H_0 в пользу альтернативной гипотезы H_1 , где $\chi_{\text{кр}}^2$ представляет собой квантиль уровня $1 - \alpha$ распределения хи-квадрат с $(s - 1)(r - 1)$ степенями свободы.
- Если $\chi^2 \leq \chi_{\text{кр}}^2$, то принимаем гипотезу H_0 .

РЕГРЕССИОННЫЙ АНАЛИЗ

§ 1. СПЕЦИФИКАЦИЯ МОДЕЛИ

В главе будет рассматриваться следующая модель наблюдений, связывающая значения некоторого наблюдаемого показателя y и объясняющих переменных $x = (x_1, \dots, x_k)^T$:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

где $\beta^T = (\beta_0, \dots, \beta_k)$ — неизвестные параметры; x_{ij} — значения объясняющих факторов; ε_i — ненаблюдаемая случайная компонента; j — номер переменной; i — номер наблюдения.

Будем предполагать, что имеется n наблюдений (измерений) показателя y_i , точно известны значения объясняющих переменных $x_i = (x_{i1}, \dots, x_{ik})$ в каждом из наблюдений, причем в модель наблюдений (1.1) входит значение ненаблюдаемой случайной компоненты ε_i . Ненаблюдаемая компонента может представлять собой ошибку измерений или заменять совокупное действие других переменных, значения которых не учитываются в модели наблюдений (1.1).

Сформулируем основные предположения регрессионного анализа, которые относятся к случайным компонентам ε_i , $i = 1, \dots, n$.

Первая группа предположений:

- Случайные величины ε_i , $i = 1, \dots, n$ образуют так называемый слабый белый шум, т.е. последовательность центрированных ($E\varepsilon_i = 0$, $i = 1, \dots, n$) и некоррелированных ($E(\varepsilon_l \varepsilon_u) = 0$ при $l \neq u$) случайных величин с одинаковыми дисперсиями σ^2 ($E\varepsilon_i^2 = \sigma^2$, $i = 1, \dots, n$).

Кроме первой группы предположений будем также рассматривать вторую группу, в которой сформулируем предположение о совместном распределении случайных величин ε_i , $i = 1, \dots, n$.

Вторая группа предположений:

- Совместное распределение случайных величин ε_i , $i = 1, \dots, n$ является нормальным распределением с нулевым вектором математических ожиданий и ковариационной матрицей $\sigma^2 E_n$, т.е. случайный вектор $\varepsilon^T = (\varepsilon_1, \dots, \varepsilon_n) \sim N(0, \sigma^2 E_n)$, где E_n — единичная матрица порядка $n \times n$.

Иногда рассматривается частный случай модели (1.1), в которой априори полагают $\beta_0 = 0$. Конечно, рассмотрение такого частного случая требует отдельного обоснования.

Модель наблюдений (1.1) можно записать в матричном виде:

$$Y = X\beta + \varepsilon, \quad (1.2)$$

где $Y = (y_1, \dots, y_n)^T$, $\beta = (\beta_0, \beta_1, \dots, \beta_k)^T$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$,

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}$$

— матрица порядка $n \times (k+1)$.

§ 2. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

В модели наблюдений (1.1), (1.2) неизвестными являются параметры $\sigma, \beta_0, \beta_1, \dots, \beta_k$. Рассмотрим в качестве процедуры оценивания неизвестных параметров метод наименьших квадратов. Метод наименьших квадратов удобнее всего излагать геометрически — на языке векторов. С этой целью введем обозначение: $X_r = (x_{1r}, x_{2r}, \dots, x_{nr})^T$, $r = 1, \dots, k$ — столбец матрицы X , тогда

$$X = (X_0, X_1, \dots, X_k),$$

где $X_0 = (1, 1, \dots, 1)^T$.

Из смысла задачи оценки неизвестных параметров β следует, что оценки разумно выбирать так, чтобы вектор $X\beta$ как можно «лучше» объяснял бы значения наблюдаемого показателя, т.е. вектор Y . Но тогда естественно выбирать β , минимизируя расстояние между векторами или квадрат длины отклонения $Y - X\beta$.

В этом случае в качестве минимизируемого критерия рассматривают

$$(Y - X\beta)^T(Y - X\beta) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_k x_{ik})^2. \quad (2.1)$$

Оценки, получаемые из условия минимума (2.1), называют оценками метода наименьших квадратов. Разумеется, для нахождения оценок β могут быть применены совсем другие процедуры, например, вместо суммы квадратов можно было бы брать сумму абсолютных величин разностей.

Для минимизации выражения (2.1) по неизвестным параметрам β воспользуемся геометрическими представлениями. Рассмотрим в n -мерном пространстве \mathbb{R}^n линейное подпространство $L(X_0, X_1, \dots, X_k)$, в котором вектора X_0, X_1, \dots, X_k являются базисом. При этом, конечно, в дальнейшем будем предполагать, что они линейно независимы. Пусть \hat{Y} представляют собой ортогональную проекцию вектора Y на подпространство $L(X_0, X_1, \dots, X_k)$. Очевидно, что для любого другого вектора $\tilde{Y} \in L(X_0, \dots, X_k)$ будет выполняться

$$(Y - \hat{Y})^T(Y - \hat{Y}) \leq (Y - \tilde{Y})^T(Y - \tilde{Y}),$$

неравенство следует из условия ортогональности вектора $Y - \hat{Y}$ подпространству $L(X_0, \dots, X_k)$, действительно, для любого вектора $\tilde{Y} \in L(X_0, \dots, X_k)$ выполняется $(\tilde{Y} - \hat{Y}) \perp (Y - \hat{Y})$, так как $(\tilde{Y} - \hat{Y}) \in L(X_0, \dots, X_k)$ и $(Y - \hat{Y}) \perp L(X_0, \dots, X_k)$. Но тогда

$$\begin{aligned} (Y - \tilde{Y})^T(Y - \tilde{Y}) &= \\ &= (Y - \hat{Y} - (\tilde{Y} - \hat{Y}))^T(Y - \hat{Y} - (\tilde{Y} - \hat{Y})) = \\ &= (Y - \hat{Y})^T(Y - \hat{Y}) + (\tilde{Y} - \hat{Y})^T(\tilde{Y} - \hat{Y}), \end{aligned}$$

откуда следует

$$(Y - \tilde{Y})^T(Y - \tilde{Y}) \geq (Y - \hat{Y})^T(Y - \hat{Y}),$$

причем равенство возможно, только если $\tilde{Y} = \hat{Y}$. Для того, чтобы вектор $Y - \hat{Y}$ был ортогонален подпространству $L(X_0, X_1, \dots, X_k)$, необходимо и достаточно, чтобы

$$(Y - \hat{Y}) \perp X_i, \quad i = 0, 1, \dots, k$$

или, чтобы

$$(Y - \hat{Y})^T X = 0. \quad (2.2)$$

Так как вектор $\hat{Y} \in L(X_0, X_1, \dots, X_k)$, то существует набор коэффициентов, обозначим их через $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$, что

$$\hat{Y} = \sum_{r=0}^k \hat{\beta}_r X_r = X \hat{\beta},$$

где $\hat{\beta} = (\hat{\beta}_0, \dots, \hat{\beta}_k)^T$. Тогда равенство (2.2) равносильно равенству

$$X^T Y - X^T X \hat{\beta} = 0$$

или равенству

$$(X^T X) \hat{\beta} = X^T Y. \quad (2.3)$$

Так как, по нашему предположению, столбцы матрицы X линейно независимы, то отсюда следует, что $n \geq k + 1$ и определитель $|X^T X| \neq 0$. Следовательно, оценки метода наименьших квадратов $\hat{\beta}$ неизвестных параметров β имеют вид:

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \quad (2.4)$$

Для удобства дальнейших преобразований введем ряд новых матриц и изучим их свойства. Из формулы (2.4) следует, что ортогональная проекция произвольного вектора Y на $L(X_0, \dots, X_K)$ имеет вид $\hat{Y} = X \hat{\beta} = X(X^T X)^{-1} X^T Y$, но тогда матрица

$$P = X(X^T X)^{-1} X^T$$

представляет собой матрицу ортогонального проектирования на пространство $L(X_0, \dots, X_k)$. Рассмотрим подпространство L^\perp ортогональное подпространству $L(X_0, \dots, X_k)$. Любой вектор Y из \mathbb{R}^n однозначно представим в виде:

$$Y = \hat{Y} + (Y - \hat{Y}),$$

где $\hat{Y} \in L(X_0, \dots, X_k)$, $Y - \hat{Y} \in L^\perp$. Из предыдущего ясно, что

$$Y - \hat{Y} = Y - X(X^T X)^{-1} X^T Y = (E_n - X(X^T X)^{-1} X^T)Y.$$

Следовательно, матрица

$$P^\perp = E_n - P$$

является матрицей ортогонального проектирования на подпространство L^\perp . Действительно, для любого вектора $Y \in \mathbb{R}^n$ получаем

$$P^\perp Y = Y - PY = Y - \hat{Y},$$

но вектор $Y - \hat{Y} \perp L(X_0, \dots, X_k)$, следовательно, вектор $Y - \hat{Y} \in L^\perp$.

Нетрудно проверить справедливость следующих свойств:

1. $P^T = P$, $(P^\perp)^T = P^\perp$.
2. $P^2 = P$, $(P^\perp)^2 = P^\perp$.
3. $PP^\perp = P^\perp P = 0$.
4. $P + P^\perp = E_n$.
5. $PX_j = X_j$, $j = 0, 1, \dots, k$, $P^\perp X_j = 0$, $j = 0, 1, \dots, k$.

Для удобства дальнейшего изложения введем вектор

$$\hat{\varepsilon} = Y - \hat{Y} = P^\perp Y, \quad (2.5)$$

который, следуя сложившейся традиции, будем называть вектором остатков. Следует иметь в виду, что после нахождения коэффициентов $\hat{\beta}$ можно рассмотреть следующую функцию, которую собственно и называют линейной регрессией:

$$\hat{y}(x) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k. \quad (2.6)$$

Построенная линейная регрессия представляет собой эмпирическую зависимость между изучаемым показателем y и объясняющими переменными $x = (x_1, \dots, x_k)^T$, поэтому естественно называть остатками разности

$$y_i - \hat{y}(x_{i1}, \dots, x_{ik}) = \hat{\varepsilon}_i, \quad i = 1, \dots, n, \quad (2.7)$$

которые не объясняются построенной эмпирической моделью (2.6). Вектор, составленный из разностей $\hat{\varepsilon}_i$, и представляет собой вектор $\hat{\varepsilon} = (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)$.

Нетрудно заметить, что

$$\hat{\varepsilon} = P^\perp Y = P^\perp (X\beta + \varepsilon) = P^\perp \varepsilon. \quad (2.8)$$

Величины $\hat{\varepsilon}_i$, $i = 1, \dots, n$, можно, следовательно, рассматривать, как оценки ненаблюдаемых величин ε_i , $i = 1, \dots, n$.

§ 3. СВОЙСТВО ОЦЕНОК МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

Как нетрудно заметить, оценки метода наименьших квадратов обладают свойством несмещенности:

$$E\hat{\beta} = (X^T X)^{-1} X^T EY = (X^T X)^{-1} X^T X\beta = \beta.$$

Вычислим дисперсионную (ковариационную) матрицу оценок метода наименьших квадратов, считая, что выполнена первая группа предположений регрессионного анализа:

$$\begin{aligned} D\hat{\beta} &= E\{(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T\} = \\ &= E\{(X^T X)^{-1} X^T \varepsilon \varepsilon^T X (X^T X)^{-1}\} = \\ &= (X^T X)^{-1} X^T \sigma^2 E_n X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}. \end{aligned}$$

Оценки метода наименьших квадратов при условии выполнения предположений из первой группы являются наилучшими линейными несмещенными оценками. Линейность понимается в том смысле, что оценки имеют вид $\hat{\beta} = AY$, где $A = (X^T X)^{-1} X^T$ — матрица коэффициентов, т.е. речь идет о линейности по наблюдениям. Следуя общему определению

эффективности оценок, заметим, что оценки $\hat{\beta}$ являются наилучшими линейными несмещенными оценками в том смысле, что для любых других оценок $\tilde{\beta}$ вида $\tilde{\beta} = CY$, обладающих свойством несмещенности, оказывается, что разность дисперсионных матриц оценок $D\tilde{\beta} - D\hat{\beta}$ представляет собой неотрицательно определенную матрицу, т.е. для любого вектора $\gamma \in \mathbb{R}^{k+1}$ выполняется $\gamma^T(D\tilde{\beta} - D\hat{\beta})\gamma \geq 0$. Отсюда, в частности, сразу следует, что дисперсии оценок $\tilde{\beta}_i$ не меньше, чем дисперсии оценок $\hat{\beta}_i$, т.е. $D\tilde{\beta}_i \geq D\hat{\beta}_i$, $i = 0, 1, 2, \dots, k$.

Теорема 3.1 (теорема Гаусса–Маркова) Пусть выполнены все условия из первой группы предположений регрессионного анализа, тогда оценки метода наименьших квадратов $\hat{\beta}$ являются наилучшими линейными несмещенными оценками.

Доказательство. Пусть $\tilde{\beta} = CY$ — произвольные линейные несмещенные оценки вектора β . Из условия несмещенности следует

$$E\tilde{\beta} = CX\beta = \beta.$$

Так как это равенство должно выполняться для любого вектора β , то условие несмещенности равносильно матричному равенству $CX = E_{k+1}$. Введем в рассмотрение матрицу T , определяемую равенством

$$T = C - (X^T X)^{-1} X^T.$$

Как легко видеть, $TX = 0$. Очевидно, что

$$D\tilde{\beta} = E\{(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)^T\} = E\{C\varepsilon\varepsilon^T C^T\} = \sigma^2 CC^T.$$

Заметим, что

$$\begin{aligned} CC^T &= (T + (X^T X)^{-1} X^T)(T + (X^T X)^{-1} X^T)^T = \\ &= TT^T + (X^T X)^{-1}. \end{aligned}$$

Но тогда для произвольного вектора $\gamma \in \mathbb{R}^{k+1}$:

$$\gamma^T(D\tilde{\beta} - D\beta)\gamma = \sigma^2 \gamma^T TT^T \gamma = \sigma^2 (\gamma^T T)(\gamma^T T)^T \geq 0.$$



Оценим дисперсию одиночного наблюдения σ^2 . Для этого вначале найдем дисперсионную (ковариационную) матрицу $D\hat{\varepsilon}$. Как следует из формулы (2.4) предыдущего параграфа:

$$D\hat{\varepsilon} = E\{P^\perp \varepsilon \varepsilon^T P^\perp\} = \sigma^2 P^\perp.$$

Рассмотрим сумму квадратов:

$$\sum_{i=1}^n \hat{\varepsilon}_i^2 = \hat{\varepsilon}^T \hat{\varepsilon} = \text{tr}\{\hat{\varepsilon} \hat{\varepsilon}^T\}.$$

Вычислим математическое ожидание:

$$E\hat{\varepsilon}^T \hat{\varepsilon} = E\text{tr}\{\hat{\varepsilon} \hat{\varepsilon}^T\} = \text{tr}E\{\hat{\varepsilon} \hat{\varepsilon}^T\} = \text{tr}\{E(P^\perp \varepsilon \varepsilon^T P^\perp)\} = \sigma^2 \text{tr}P^\perp.$$

Заметим, что

$$\begin{aligned} \text{tr}P^\perp &= \text{tr}\{E_n - X(X^T X)^{-1}X^T\} = \\ &= n - \text{tr}\{X(X^T X)^{-1}X^T\} = \\ &= n - \text{tr}\{(X^T X)(X^T X)^{-1}\} = n - k - 1. \end{aligned}$$

При выводе формулы мы воспользовались свойством $\text{tr}(AB) = \text{tr}(BA)$, если оба произведения имеют смысл (т.е. размерности матриц позволяют рассматривать оба произведения).

В результате проведенных выкладок получаем, что статистика

$$S^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - k - 1} = \frac{1}{n - k - 1} \hat{\varepsilon}^T \hat{\varepsilon}$$

является несмещенной оценкой дисперсии σ^2 .

Рассмотрим матрицу $E\{(\hat{\beta} - \beta)\hat{\varepsilon}^T\}$, элементы которой представляют собой ковариации случайных величин $\hat{\beta}_j$ и $\hat{\varepsilon}_i$, $j = 0, 1, \dots, k$, $i = 1, 2, \dots, n$. Как нетрудно видеть,

$$\begin{aligned} E\{(\hat{\beta} - \beta)\hat{\varepsilon}^T\} &= (X^T X)^{-1}X^T E\{\varepsilon \varepsilon^T\}P^\perp = \\ &= \sigma^2 (X^T X)^{-1}X^T P^\perp = 0. \end{aligned}$$

Последнее равенство следует из свойств матрицы P^\perp . Найденная несмещенная оценка S^2 одиночной дисперсии σ^2 позволяет построить несмещенные оценки всех ковариаций вектора оценок $\hat{\beta}$. Действительно, как было получено ранее, $D\hat{\beta} = \sigma^2(X^T X)^{-1}$, но теперь, заменяя σ^2 на S^2 , получаем несмещенные оценки всех ковариаций и дисперсий вектора $\hat{\beta}$, другими словами, элементы матрицы $S^2(X^T X)^{-1}$ являются несмещенными оценками дисперсий и ковариаций оценок $\hat{\beta}_j$, $j = 0, \dots, k$.

Лемма 3.1. Пусть выполнены все условия из первой и второй групп предположений регрессионного анализа, тогда статистика $(n - k - 1)S^2/\sigma^2$ подчиняется распределению χ^2 с $n - k - 1$ степенями свободы.

Доказательство. Ранее было установлено, что $\hat{\varepsilon} = P^\perp \varepsilon$, где P^\perp — матрица ортогонального проектирования на подпространстве L^\perp . Очевидно, что размерность подпространства $L(X_0, \dots, X_k)$ равна $k + 1$. Но тогда размерность подпространства L^\perp равна $n - k - 1$. Выберем в подпространстве L^\perp ортонормированный базис l_1, \dots, l_{n-k-1} . Пусть $l = (l_1, \dots, l_{n-k-1})$. Тогда вектор $\hat{\varepsilon}$ единственным образом представим в базисе l :

$$\hat{\varepsilon} = \sum_{r=1}^{n-k-1} \eta_r l_r = l\eta,$$

где $\eta^T = (\eta_1, \dots, \eta_{n-k-1})$.

Очевидно, что

$$\eta = l^T \hat{\varepsilon}.$$

В силу второй группы предположений вектор $\hat{\varepsilon}$ подчиняется многомерному нормальному распределению, но тогда η также подчиняется многомерному нормальному распределению в виду линейности преобразования.

Нетрудно увидеть, что $E\eta = l^T E\hat{\varepsilon} = 0$,

$$D\eta = l^T D\hat{\varepsilon} l = \sigma^2 l^T P^\perp l = \sigma^2 l^T l = \sigma^2 E_{n-k-1},$$

так как P^\perp матрица ортогонального проектирования на L^\perp , но все $l_i \in L^\perp$, $i = 1, \dots, n - k - 1$. Кроме того,

$$\hat{\varepsilon}^T \hat{\varepsilon} = \eta^T l^T l \eta = \eta^T \eta.$$

Следовательно, статистики $\hat{\varepsilon}^T \hat{\varepsilon} / \sigma^2$ и $\eta^T \eta / \sigma^2$ распределены одинаково, но из определения распределения χ^2 получаем, что статистика $\eta^T \eta / \sigma^2$ подчиняется распределению χ^2 с $n - k - 1$ степенями свободы. ■

Лемма 3.2. Пусть выполнены предположения первой и второй групп, тогда справедливы утверждения:

1. Вектор оценок $\hat{\beta}$ подчиняется многомерному нормальному распределению, $\hat{\beta} \sim N(\beta, \sigma^2(X^T X)^{-1})$.
2. Статистика $(n - k - 1)S^2 / \sigma^2$ подчиняется распределению χ^2 с $(n - k - 1)$ степенями свободы и взаимно независима с вектором оценок $\hat{\beta}$.

Доказательство. Первое утверждение сразу следует из линейности явного выражения $\hat{\beta}$ и второй группы предположений:

$$\hat{\beta} = (X^T X)^{-1} X^T Y,$$

линейное преобразование не меняет тип нормального распределения.

Вектор математических ожиданий и дисперсионная матрица были вычислены ранее. Второе утверждение следует из леммы 3.1 и того факта, что $E\{(\hat{\beta} - \beta)\hat{\varepsilon}^T\} = 0$.

Действительно, как нетрудно заметить,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\varepsilon} \end{pmatrix} = \begin{pmatrix} (X^T X)^{-1} X^T \\ P^\perp \end{pmatrix} Y.$$

Следовательно, вектор $(\hat{\beta}^T, \hat{\varepsilon}^T)$ подчиняется многомерному нормальному распределению, но тогда вектора $\hat{\beta}$ и $\hat{\varepsilon}$ взаимно независимы. ■

§ 4. ПОСТРОЕНИЕ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ И ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

При выполнении всех условий из первой и второй групп предположений регрессионного анализа справедлива лемма 3.2 предыдущего параграфа, из которой сразу следует, что статистика

$$\frac{\hat{\beta}_j - \beta_j}{S\sqrt{[(X^T X)^{-1}]_{(j+1)(j+1)}}} \sim T_{n-k-1}, \quad j = 0, \dots, k, \quad (4.1)$$

где $[(X^T X)^{-1}]_{(j+1)(j+1)}$ — элемент, стоящий на главной диагонали в строке $j + 1$ и столбце $j + 1$ матрицы $(X^T X)^{-1}$, распределение T_{n-k-1} — распределение Стьюдента с $n - k - 1$ степенями свободы.

Доказательство следует из определения распределения Стьюдента и леммы 3.2.

Из формулы (4.1) следует формула для доверительного интервала с уровнем доверия $1 - \alpha$ для любого параметра β_j , $j = 0, 1, \dots, k$. Доверительный интервал имеет вид:

$$P \left\{ \hat{\beta}_j - t_{1-\frac{\alpha}{2}, n-k-1} S \sqrt{[(X^T X)^{-1}]_{(j+1)(j+1)}} < \beta_j < \right. \\ \left. < \hat{\beta}_j + t_{1-\frac{\alpha}{2}, n-k-1} S \sqrt{[(X^T X)^{-1}]_{(j+1)(j+1)}} \right\} = 1 - \alpha,$$

где $t_{1-\frac{\alpha}{2}, n-k-1}$ — квантиль уровня $1 - \frac{\alpha}{2}$ распределения Стьюдента T_{n-k-1} .

Важное значение имеет проверка гипотез статистической значимости найденных оценок $\hat{\beta}_j$, $j = 0, 1, \dots, n$. Если доверительный интервал для параметра β_j с уровнем доверия $1 - \alpha$ содержит нуль, т.е. если концы доверительного интервала имеют разный знак, то нельзя отклонить гипотезу H_0 : $\beta_j = 0$, следовательно, в этом случае найденная оценка $\hat{\beta}_j$ не может быть признана статистически значимой на уровне значимости α . Отсюда следует правило проверки статистической значимости оценки $\hat{\beta}_j$ или, что то же самое, правило проверки гипотезы H_0 : $\beta_j = 0$ (справедливость гипотезы означает статистическую незначимость оценки $\hat{\beta}_j$).

Выберем уровень значимости α и вычислим статистику

$$t_{\beta_j} = \frac{\hat{\beta}_j}{S\sqrt{[(X^T X)^{-1}]_{(j+1)(j+1)}}}.$$

Если $|t_{\beta_j}| > t_{1-\frac{\alpha}{2}, n-k-1}$, то гипотеза H_0 отклоняется и оценка $\hat{\beta}_j$ признается статистически значимой на уровне значимости α . Если $|t_{\beta_j}| \leq t_{1-\frac{\alpha}{2}, n-k-1}$, то гипотеза H_0 не отклоняется и оценка $\hat{\beta}_j$ признается статистически незначимой на уровне значимости α .

Проверка статистической значимости оценки $\hat{\beta}_j$ сводится к проверке гипотезы $H_0: \beta_j = 0$.

Нетрудно аналогичным образом проверить более общую гипотезу вида $H_0: \beta_j = \beta_j^{(0)}$. Рассуждая таким же образом, получаем правило проверки гипотезы H_0 на уровне значимости α :

- Если $\frac{|\hat{\beta}_j - \beta_j^{(0)}|}{S\sqrt{[(X^T X)^{-1}]_{(j+1)(j+1)}}} > t_{1-\frac{\alpha}{2}, n-k-1}$, то гипотеза $H_0: \beta_j = \beta_j^{(0)}$ отклоняется на уровне значимости α .
- Если $\frac{|\hat{\beta}_j - \beta_j^{(0)}|}{S\sqrt{[(X^T X)^{-1}]_{(j+1)(j+1)}}} \leq t_{1-\frac{\alpha}{2}, n-k-1}$, то гипотеза $H_0: \beta_j = \beta_j^{(0)}$ принимается на уровне значимости α .

В линейном регрессионном анализе коэффициентом детерминации R^2 называется квадрат коэффициента корреляции между наблюдаемыми значениями показателя $Y^T = (y_1, \dots, y_n)$ и значениями эмпирической функции $\hat{Y}^T = (\hat{y}_1, \dots, \hat{y}_n)$.

Так как $X_0^T = (1, \dots, 1)$ и $\hat{Y} = PY$, причем P — матрица ортогонального проектирования на подпространство $L(X_0, \dots, X_n)$, то $PX_0 = X_0$ или $X_0^T = X_0^T P$.

Но тогда

$$X_0^T \hat{Y} = X_0^T PY = X_0^T Y$$

или

$$\sum_{j=1}^n \hat{y}_i = \sum_{i=1}^n y_i,$$

следовательно,

$$\bar{\hat{y}} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}.$$

По определению

$$\begin{aligned} R^2 &= \frac{(\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{y}))^2}{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{y})^2} = \\ &= \frac{((Y - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0))^2}{(Y - \bar{y}X_0)^T(Y - \bar{y}X_0)(\hat{Y} - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0)}. \end{aligned}$$

Заметим, что

$$\begin{aligned} (Y - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0) &= ((Y - \hat{Y}) + \\ &+ (\hat{Y} - \bar{y}X_0))^T(\hat{Y} - \bar{y}X_0) = (\hat{Y} - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0), \end{aligned}$$

так как вектор $Y - \hat{Y} \in L^\perp$, а $\hat{Y} - \bar{y}X_0 \in L(X_0, \dots, X_k)$ и, следовательно, вектора $Y - \hat{Y}$ и $\hat{Y} - \bar{y}X_0$ ортогональны, т. е.

$$(Y - \hat{Y})^T(\hat{Y} - \bar{y}X_0) = 0.$$

Но тогда

$$\begin{aligned} R^2 &= \frac{(\hat{Y} - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0)}{(Y - \bar{y}X_0)^T(Y - \bar{y}X_0)} = \\ &= 1 - \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{(Y - \bar{y}X_0)^T(Y - \bar{y}X_0)}. \quad (4.2) \end{aligned}$$

Для доказательства последнего равенства заметим, что

$$\begin{aligned} (\hat{Y} - \bar{y}X_0)^T(\hat{Y} - yX_0) &= \\ &= ((\hat{Y} - Y) + (Y - \bar{y}X_0))^T((\hat{Y} - Y) + (Y - \bar{y}X_0)) = \\ &= (-\hat{\varepsilon} + (Y - \bar{y}X_0))^T(-\hat{\varepsilon} + (Y - \bar{y}X_0)) = \\ &= -\hat{\varepsilon}^T \hat{\varepsilon} + (Y - \bar{y}X_0)^T(Y - \bar{y}X_0), \end{aligned}$$

в виду того, что

$$\begin{aligned}\hat{\varepsilon}^T(Y - \bar{y}X_0) &= \hat{\varepsilon}^TY - \bar{y}\hat{\varepsilon}^TX_0 = \\ &= \hat{\varepsilon}^T(Y - \hat{Y}) + \hat{\varepsilon}^T\hat{Y} - \bar{y}\hat{\varepsilon}^TX_0 = \hat{\varepsilon}^T\hat{\varepsilon},\end{aligned}$$

так как $\hat{\varepsilon} \in L^\perp$, $\hat{Y} \in L(X_0, \dots, X_k)$, $X_0 \in L(X_0, \dots, X_k)$.

Равенство (4.2) удобно в вычислительном отношении, кроме того, из него видно, что если построенная линейная регрессия идеально точно соответствует наблюдениям, то $Y = \hat{Y}$ и, следовательно, $\hat{\varepsilon} = 0$, но тогда $R^2 = 1$. Наоборот, если $\hat{\beta}_1 = \dots = \hat{\beta}_k = 0$, т.е. линейная регрессия не зависит от x_1, \dots, x_k , то $R^2 = 0$. Из определения коэффициента детерминации R^2 сразу следует, что $0 \leq R^2 \leq 1$.

Наиболее важным применением коэффициента R^2 является его использование при проверке гипотезы статистической значимости линейной регрессионной модели в целом, т.е. при проверке гипотезы $H_0: \beta_1 = \dots = \beta_k = 0$. Если указанную гипотезу H_0 нельзя отвергнуть на уровне значимости α , то это означает, что на уровне значимости α построенная линейная регрессия статистически незначима. Малые значения R^2 свидетельствуют против статистической значимости построенной линейной регрессии.

Лемма 4.1. Пусть выполнены обе группы основных предположений линейного регрессионного анализа, тогда в предположении справедливости гипотезы $H_0: \beta_1 = \dots = \beta_k = 0$ статистика

$$F = \frac{R^2}{1 - R^2} \frac{n - k - 1}{k} \sim \mathcal{F}_{k, n-k-1}$$

подчиняется распределению Фишера со степенями свободы k и $n - k - 1$.

Доказательство. Очевидно, что

$$\frac{R^2}{1 - R^2} = \frac{(\hat{Y} - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0)}{\hat{\varepsilon}^T\hat{\varepsilon}}.$$

При доказательстве леммы 3.2 было показано, что вектора $\hat{\beta}$ и $\hat{\varepsilon}$ взаимно независимы, так как $\hat{Y} = X\hat{\beta}$ и $\bar{y} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i$, то

числитель и знаменатель дроби взаимно независимы. Кроме того, уже было ранее показано, что статистика

$$\frac{(n-k-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \hat{\varepsilon}^T \hat{\varepsilon} \sim \chi_{n-k-1}^2$$

подчиняется распределению χ^2 с $n-k-1$ степенями свободы. Найдем распределение числителя в предположении справедливости гипотезы H_0 : $\beta_1 = \dots = \beta_k = 0$. Заметим, что

$$\hat{Y} = PY = X\beta + P\varepsilon = \beta_0 X_0 + P\varepsilon,$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i = \beta_0 + \frac{1}{n} X_0^T P\varepsilon = \beta_0 + \frac{1}{n} X_0^T \varepsilon.$$

Тогда

$$\hat{Y} - \bar{y}X_0 = P\varepsilon - \frac{1}{n} X_0^T \varepsilon X_0 = P\varepsilon - \bar{\varepsilon} P X_0 = P(\varepsilon - \bar{\varepsilon} X_0),$$

где $\bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^n \varepsilon_i$, $X_0 = P X_0$.

Как видим,

$$\eta = \hat{Y} - \bar{y}X_0 = P(E_n - \frac{1}{n} T_n) \varepsilon, \quad (4.3)$$

где $T_n = X_0 X_0^T$ — квадратная матрица порядка $n \times n$, все элементы которой являются единицами. Нетрудно заметить, что $\eta \in L(X_0, \dots, X_k)$, кроме того, вектор η ортогонален вектору X_0 , следовательно, $\eta \in L(\tilde{X}_1, \dots, \tilde{X}_k)$, где базисные вектора определяются следующими равенствами $\tilde{X}_r = X_r - \bar{x}_r X_0$, $\bar{x}_r = \sum_{i=1}^n x_{ir}/n$. Нетрудно убедиться, что все вектора \tilde{X}_r ортогональны вектору X_0 , следовательно, вектор X_0 ортогонален подпространству $L(\tilde{X}_1, \dots, \tilde{X}_k)$. Кроме того, из (4.3) следует, что вектор η подчиняется многомерному нормальному распределению, $E\eta = 0$.

Нетрудно проверить, что матрица $P_1 = P(E_n - T_n/n)$ — матрица ортогонального проектирования на подпространство $L(\tilde{X}_1, \dots, \tilde{X}_k)$. Вектор $\eta = P_1 \varepsilon$, следовательно, $D\eta = \sigma^2 P_1$. Все свойства, которые были сформулированы для матрицы P , выполняются и для матрицы P_1 . Выберем в подпространстве

$L(\tilde{X}_1, \dots, \tilde{X}_k)$ ортонормированный базис $d = (d_1, \dots, d_k)$, тогда

$$\eta = \sum_{r=1}^k \xi_r d_r = d\xi, \quad \xi = d^T \eta,$$

$$\xi = (\xi_1, \dots, \xi_k)^T \sim N(0, \sigma^2 E_k),$$

так как

$$D\xi = E\{(d^T \eta)(\eta^T d)\} = \sigma^2 d^T P_1 d = \sigma^2 d^T d = \sigma^2 E_k.$$

Кроме того,

$$\eta^T \eta = \xi^T d^T d \xi = \xi^T \xi.$$

Следовательно, распределения левой и правой частей совпадают. Но тогда из определения распределения χ^2 получаем, что статистика

$$\frac{1}{\sigma^2}(\hat{Y} - \bar{y}X_0)^T(\hat{Y} - \bar{y}X_0) \sim \chi_k^2$$

подчиняется распределению χ^2 с k степенями свободы. ■

Построенная линейная регрессия статистически значима на уровне α тогда и только тогда, когда гипотеза $H_0: \beta_1 = \dots = \beta_k = 0$ отклоняется на уровне значимости α . Поэтому правило проверки статистической значимости линейной регрессии в целом сформулируем следующим образом:

- Если $F = \frac{R^2}{1-R^2} \frac{n-k-1}{k} > F_{1-\alpha; k, n-k-1}$, то гипотеза $H_0: \beta_1 = \dots = \beta_k = 0$ отклоняется на уровне значимости α и, следовательно, построенная линейная регрессия является статистически значимой.
- Если $F \leq F_{1-\alpha; k, n-k-1}$, то гипотеза H_0 принимается, и, следовательно, построенная линейная регрессия является статистически незначимой, здесь $F_{1-\alpha; k, n-k-1}$ — квантиль уровня $1 - \alpha$ распределения Фишера с k и $n - k - 1$ степенями свободы.

БИНАРНАЯ РЕГРЕССИЯ

§ 1. ДИСКРЕТНЫЕ ДАННЫЕ

В данной главе зависимые переменные принимают дискретные значения, выражающие какие-либо качественные признаки. Объясняющие переменные могут быть как дискретными, так и непрерывными. Например, при приеме на работу кандидата компания рассматривает многие показатели (возраст, образование, уровень владения иностранными языками, семейное положение, опыт работы и т. д.), на основе которых она должна сделать вывод: принять на работу или нет.

Выделим несколько классов задач, в которых зависимые переменные принимают дискретные значения:

1. Переменные — это решения «да» (1) или «нет» (0), т. е. выбор одной из двух альтернатив.

Если имеется только две альтернативы, то результат наблюдения обычно описывается переменной, называемой бинарной. В общем случае при наличии k альтернатив результат выбора можно представить переменной, принимающей значения $1, \dots, k$. Если альтернативы нельзя упорядочить (например, при выборе профессии из списка: учитель, инженер, врач, повар), то их нумерация может быть произвольной. В этих случаях соответствующую переменную называют *номинальной*.

2. Переменные — ранги. Например, «0» означает «категорически против», «1» — «против», «2» — «ни да, ни нет», «3» — «за», «4» — «полностью за». Такая система

ранжирования может быть использована, например, при голосовании.

Соответствующая переменная называется *порядковой*, *ординальной* или *ранговой*.

3. Переменная — количественная целочисленная характеристика. Например, количество торговых точек, приносящих прибыль, количество частных предприятий, зарегистрированных в регионе, и т. д.

Для моделей с дискретными зависимыми переменными, конечно же, возможно формальное применение метода наименьших квадратов, однако, получить удовлетворительные с содержательной точки зрения результаты удастся, как правило, лишь для моделей с количественными целочисленными переменными. В случае порядковых переменных интерпретация оценок коэффициентов при объясняющих переменных значительно затруднена: увеличение на одну единицу порядковой переменной означает переход к следующей по рангу альтернативе, однако далеко не всегда переход от первой альтернативы ко второй численно эквивалентен переходу от второй к третьей. Если же зависимая переменная является номинальной, то результаты оценивания вообще теряют смысл в силу произвольности нумерации альтернатив. Таким образом, стандартная регрессионная схема нуждается в коррекции.

§ 2. МОДЕЛЬ ЛИНЕЙНОЙ ВЕРОЯТНОСТИ

Рассмотрим задачу покупки семьей некоторой квартиры. Зависимая переменная y принимает значение 1, если семья купила квартиру в данный период, и 0 в противном случае. Решение о покупке квартиры зависит от множества факторов (независимых переменных): бюджет семьи, наличие детей, наличие объектов социальной инфраструктуры в шаговой доступности от дома, удаленность от центра города и т. д. Из этих факторов можно сформировать вектор $x^T = (1, x_1, \dots, x_k)$. Если следовать основным идеям регрессионного анализа, то будем считать, что на решение о покупке квартиры влияют также неучтенные факторы, совокупное влияние которых моделируется случайной компонентой. При

различных предположениях о характере зависимости y от x , можно получить различные модели построения зависимости. В настоящей главе рассматриваются модель линейной вероятности, логит и пробит модели бинарного выбора.

Пусть имеется выборка объема n наблюдений (x_i, y_i) , $i = 1, \dots, n$, где $x_i^T = (1, x_{i1}, \dots, x_{ik})^T$, y_i — зависимая переменная, которая может принимать только два значения: ноль и единица. Рассмотрим стандартную модель линейной регрессии:

$$y_i = \beta^T x_i + \varepsilon_i, \quad (2.1)$$

где $\beta^T = (\beta_0, \dots, \beta_k)$ — вектор неизвестных параметров, $\beta \in \mathbb{R}^k$, ε_i — случайная компонента. В предположениях регрессионного анализа считается, что случайная компонента подчиняется нормальному закону распределения с нулевым математическим ожиданием. Учитывая это, получаем, что

$$Ey_i = \beta^T x_i.$$

Так как y_i принимает значения 0 или 1, то для математического ожидания y_i имеем равенство:

$$Ey_i = 1 \cdot P\{y_i = 1\} + 0 \cdot P\{y_i = 0\} = P\{y_i = 1\}. \quad (2.2)$$

Таким образом, получаем равенство:

$$P\{y_i = 1\} = \beta^T x_i, \quad (2.3)$$

которое дало название модели линейной вероятности (linear probability model).

Следует отметить некоторые недостатки этой модели, которые не позволяют успешно применять метод наименьших квадратов для оценивания параметров β и построения прогнозов. Из (2.1) следует, что компонента ε_i в каждом наблюдении может принимать только два значения: $(1 - \beta^T x_i)$ с вероятностью $P\{y_i = 1\}$ и $(-\beta^T x_i)$ с вероятностью $1 - P\{y_i = 1\}$. Это, в частности, не позволяет считать случайную компоненту нормально распределенной случайной величиной или подчиняющейся распределению, близкому к нормальному.

Проверим выполнение условия из первой группы предположений регрессионного анализа о равенстве дисперсий различных наблюдений. Вычислим дисперсию компоненты:

$$D\varepsilon_i = \beta^T x_i (1 - \beta^T x_i).$$

Получается, что дисперсия компоненты ε_i зависит от x_i . Известно, что оценка параметров β , полученная обычным методом наименьших квадратов, в этом случае не является эффективной.

Еще одним серьезным недостатком модели линейной вероятности является тот факт, что прогнозные значения $\hat{y}_i = \hat{\beta}^T x_i$, т. е. прогнозные значения вероятности $P\{y_i = 1\}$, могут лежать вне отрезка $[0, 1]$ (здесь $\hat{\beta}$ — оценка параметра β , полученная методом наименьших квадратов).

§ 3. ЛОГИТ И ПРОБИТ МОДЕЛИ БИНАРНОГО ВЫБОРА

Откажемся от предположения о линейной зависимости вероятности $P\{y_i = 1\}$ от β . Предположим, что

$$P\{y_i = 1\} = F(\beta^T x_i), \quad (3.1)$$

где $F(x)$ — некоторая функция, область значений которой лежит в отрезке $[0, 1]$. В частности, в качестве функции $F(x)$ можно рассмотреть функцию распределения некоторой случайной величины. Возможна следующая интерпретация предположения (3.1). Предположим, что существует некоторая количественная переменная y_i^* , связанная с независимыми переменными x_i линейным регрессионным уравнением:

$$y_i^* = \beta^T x_i + \varepsilon_i, \quad (3.2)$$

где случайные компоненты ε_i независимы и одинаково распределены с нулевым математическим ожиданием и дисперсией σ^2 . Пусть $F(\cdot)$ — функция распределения нормированной случайной величины ε_i/σ . Переменная y_i^* является ненаблюдаемой (латентной), а решение, соответствующее значению $y_i = 1$, принимается тогда, когда y_i^* превосходит некоторое

пороговое значение. Без ограничения общности, если константа включена в число независимых переменных модели, можно считать это пороговое значение равным нулю. Величину y_i^* можно также интерпретировать как разность полезностей альтернативы 1 и альтернативы 0.

Таким образом, получаем

$$y_i = \begin{cases} 1, & \text{если } y_i^* \geq 0, \\ 0, & \text{если } y_i^* < 0. \end{cases} \quad (3.3)$$

Тогда, предполагая, что случайные компоненты ε_i имеют одно и то же симметричное распределение с непрерывной функцией распределения $F(x)$, $F(-x) = 1 - F(x)$, получаем следующие равенства:

$$\begin{aligned} P\{y_i = 1\} &= P\{y_i^* \geq 0\} = \\ &= P\{\beta^T x_i + \varepsilon_i \geq 0\} = P\{\varepsilon_i \geq -\beta^T x_i\} = \\ &= 1 - F\left(\frac{-\beta^T x_i}{\sigma}\right) = F\left(\frac{\beta^T x_i}{\sigma}\right), \end{aligned} \quad (3.4)$$

что с точностью до нормировки совпадает с (3.1).

В модели (3.4) параметры β и σ участвуют только в виде отношения β/σ и не могут быть по отдельности идентифицированы (т. е. оценить можно только отношения β/σ). Поэтому, в данном случае, без ограничения общности можно считать, что $\sigma = 1$.

Наиболее часто в качестве функции $F(x)$ используют:

1. Функцию стандартного нормального распределения

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{z^2}{2}} dz,$$

в этом случае модель принято называть *пробит моделью*.

2. Функцию логистического распределения

$$\Lambda(u) = \frac{e^u}{1 + e^u}, \quad (3.5)$$

тогда модель принято называть *логит моделью*.

Использование функции стандартного нормального распределения представляется естественным в рамках приведенной выше интерпретации. Применение функции логистического распределения объясняется простотой численной реализации процедуры оценивания параметров. У исследователя может возникнуть вопрос о том, какую из моделей (пробит или логит) использовать в конкретном случае. Точного ответа на этот вопрос нет, но можно дать некоторые рекомендации. Например, можно выбрать ту модель, для которой функция правдоподобия имеет большее значение. Также можно отметить, что для значений u , близких по модулю к нулю (в частности, при $u \in [-1.2; 1.2]$), функции $\Phi(u)$ и $\Lambda(u)$ ведут себя примерно одинаково (см. § 4), в то же время «хвосты» логистического распределения значительно «тяжелее» «хвостов» нормального распределения. Практический опыт показывает, что для выборок с небольшим разбросом независимых переменных и при отсутствии существенного преобладания одной альтернативы над другой выводы, получаемые с помощью пробит и логит моделей, будут, как правило, совпадать.

Исходя из предположения (3.1), функция $F(x)$ нелинейна по параметрам β , и интерпретация этих параметров отличается от интерпретации подобных коэффициентов в линейной регрессионной модели.

§ 4. СРАВНЕНИЕ ЗНАЧЕНИЙ ФУНКЦИЙ НОРМАЛЬНОГО И ЛОГИСТИЧЕСКОГО РАСПРЕДЕЛЕНИЙ

Вероятностные распределения, используемые в пробит и логит моделях, — стандартное нормальное и логистическое распределения соответственно. Обе указанные функции распределения симметричны относительно 0 и имеют дисперсии, равные 1 и $\pi^2/3$ соответственно. Рассмотрим модифицированное логистическое распределение с функцией распределения следующего вида:

$$\Lambda_{\delta}(u) = \frac{e^{\delta u}}{1 + e^{\delta u}}. \quad (4.1)$$

Значение параметра δ может быть выбрано таким образом, чтобы значения функции (4.1) были бы достаточно близкими к значениям функции стандартного нормального распределения на большей части области определения. Рассмотрим, например, $\delta = 1.6$. В следующей таблице можно найти значения функций $\Phi(u)$ и $\Lambda_{1.6}(u)$ для различных u . Таблица 19.1 показывает, что функции распределения «очень близки» около 0, но логистическое распределение имеет более тяжелые «хвосты».

u	$\Phi(u)$	$\Lambda_{1.6}(u)$
0.0	0.5	0.5
0.1	0.5398	0.5399
0.2	0.5793	0.5793
0.3	0.6179	0.6177
0.4	0.6554	0.6548
0.5	0.6915	0.6900
0.6	0.7257	0.7231
0.7	0.7580	0.7540
0.8	0.7881	0.7824
0.9	0.8159	0.8085
1.0	0.8413	0.8320
2.0	0.9772	0.9608
3.0	0.9987	0.9918

Таблица 19.1

Таблица значений функций $\Phi(u)$ и $\Lambda_{1.6}(u)$

Из-за близости двух распределений трудно идентифицировать тип распределения при наличии выборки небольшого объема. Таким образом, при построении модели бинарной регрессии не имеет большого значения, будет использоваться пробит или логит модель, исключая случаи, когда большое количество данных расположены в хвостах, что может быть обусловлено спецификой рассматриваемой проблемной области. В моделях множественного дискретного выбора пробит и логит модели отличаются гораздо более существенно.

Предположим, что найдены оценки $\hat{\beta}_\Phi$ и $\hat{\beta}_\Lambda$ для параметров пробит и логит моделей соответственно. Тогда, используя

приближенное равенство $\Lambda_{1.6}(u) \simeq \Phi(u)$, тогда можно записать следующее приближенное равенство для оценок параметров:

$$1.6\hat{\beta}_\Phi \simeq \hat{\beta}_\Lambda. \quad (4.2)$$

Формула (4.2) может быть полезна как быстрый способ сравнения оценок параметров пробит и логит моделей.

§ 5. ОЦЕНИВАНИЕ ПАРАМЕТРОВ В ЛОГИТ И ПРОБИТ МОДЕЛЯХ

Для нахождения оценок параметров β обычно используют метод максимального правдоподобия, предполагая, что наблюдения y_1, \dots, y_n независимы. Так как y_i может принимать значения 0 или 1, то функция правдоподобия примет следующий вид:

$$L(y_1, \dots, y_n) = \prod_{i: y_i=0} (1 - F(\beta^T x_i)) \prod_{i: y_i=1} F(\beta^T x_i). \quad (5.1)$$

Нетрудно заметить, что

$$L(y_1, \dots, y_n) = \prod_{i=1}^n F^{y_i}(\beta^T x_i) (1 - F(\beta^T x_i))^{1-y_i}.$$

Рассмотрим логарифмическую функцию правдоподобия:

$$\begin{aligned} \ln L(y_1, \dots, y_n) = \\ = \sum_{i=1}^n (y_i \ln F(\beta^T x_i) + (1 - y_i) \ln(1 - F(\beta^T x_i))). \end{aligned} \quad (5.2)$$

Дифференцируя равенство (5.2) по вектору β , получаем уравнение правдоподобия, записанное в векторной форме:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^n \left(\frac{y_i f(\beta^T x_i)}{F(\beta^T x_i)} - \frac{(1 - y_i) f(\beta^T x_i)}{1 - F(\beta^T x_i)} \right) x_i = 0, \quad (5.3)$$

где $f(x)$ — плотность распределения, соответствующая функции $F(x)$. Для логит модели уравнение (5.3) можно существенно упростить, если воспользоваться тождеством

$$\Lambda'(u) = \Lambda(u)(1 - \Lambda(u)):$$

$$\sum_{i=1}^n (y_i - \Lambda(\beta^T x_i)) x_i = 0. \quad (5.4)$$

Уравнение правдоподобия есть лишь необходимое условие локального экстремума. Можно показать, что для пробит и логит моделей логарифмическая функция правдоподобия (5.2) является вогнутой по β функцией и, значит, решение уравнения (5.3) дает оценку максимального правдоподобия параметра β [40]. Гессиан для логит модели имеет следующий вид:

$$H = \frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} = - \sum_{i=1}^n \Lambda(\beta^T x_i) (1 - \Lambda(\beta^T x_i)) x_i x_i^T. \quad (5.5)$$

Заметим также, что гессиан в этом случае отрицательно определен [40], т. е. логарифмическая функция правдоподобия вогнута.

Для пробит модели логарифмическую функцию правдоподобия (5.2) можно записать в следующем виде:

$$\ln L = \sum_{i:y_i=0} \ln(1 - \Phi(\beta^T x_i)) + \sum_{i:y_i=1} \ln(\Phi(\beta^T x_i)). \quad (5.6)$$

Тогда условие (5.3) будет следующим:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i:y_i=0} \frac{-\varphi(\beta^T x_i)}{1 - \Phi(\beta^T x_i)} x_i + \sum_{i:y_i=1} \frac{\varphi(\beta^T x_i)}{\Phi(\beta^T x_i)} x_i,$$

где $\varphi(x) = \Phi'(x)$. Учитывая, что нормальное распределение, как и логистическое, симметрично, $1 - \Phi(\beta^T x) = \Phi(-\beta^T x)$, получаем:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^n \frac{q_i \varphi(\beta^T x_i)}{\Phi(q_i \beta^T x_i)} x_i = \sum_{i=0}^n \lambda_i x_i = 0, \quad (5.7)$$

где $q_i = 2y_i - 1$, $\lambda_i = q_i \varphi(\beta^T x_i) / \Phi(q_i \beta^T x_i)$.

Для вычисления гессиана в модели пробит анализа будем использовать свойство стандартного нормального распределения: $d\varphi(u)/du = -u\varphi(u)$. Тогда для пробит модели получим следующее выражение для гессиана:

$$H = \frac{\partial^2 \ln L}{\partial \beta^T \partial \beta} = - \sum_{i=1}^n \lambda_i (\lambda_i + \beta^T x_i) x_i x_i^T. \quad (5.8)$$

Эта матрица также отрицательно определена [46].

Уравнения правдоподобия (5.4) и (5.7) являются системой нелинейных (относительно β) уравнений, аналитическое решение которой невозможно найти в явном виде в общем случае, поэтому при ее решении приходится прибегать к численным методам.

§ 6. ЧИСЛЕННЫЕ МЕТОДЫ НАХОЖДЕНИЯ ОЦЕНОК В ЛОГИТ И ПРОБИТ МОДЕЛЯХ

В предыдущем параграфе мы говорили о том, что в явном виде найти решение уравнения правдоподобия не удастся. Будем использовать численные методы для того, чтобы найти оценку максимального правдоподобия параметра β , который является $(k + 1)$ -мерным вектором. Общая схема численных методов для нахождения оценки максимального правдоподобия имеет следующий вид. Сначала выбираем начальную точку $\beta^{(0)}$, на следующей итерации переходим в точку $\beta^{(1)}$ по следующему правилу:

$$\beta^{(1)} = \beta^{(0)} + \mu_0,$$

на t -ой итерации переходим в точку β^{t+1} по следующему правилу:

$$\beta^{(t+1)} = \beta^{(t)} + \mu_t.$$

Обычно μ_t выбирают в виде $\mu_t = D_t \partial \ln L / \partial \beta^{(t)}$, где $\partial \ln L / \partial \beta^{(t)}$ задает направление «спуска», т. е. направление изменения значений β , D_t — матрица «длин» шага. Итерационный процесс заканчивается на том шаге, на котором выполняется заранее определенное условие остановки. Такая итеративная процедура в результате определяет точку $\hat{\beta}$

(оценку максимального правдоподобия). Задача — определить способ выбора вектора μ_t , поскольку от этого зависит скорость сходимости к искомой точке.

Метод градиентного спуска. Для этого метода предполагается, что $D_t = E$ для любого шага t . Выбор точки на $(t + 1)$ -ой итерации осуществляется по правилу:

$$\beta^{(t+1)} = \beta^{(t)} + \frac{\partial \ln L}{\partial \beta^{(t)}}.$$

В скалярном случае оценка будет возрастать, если градиент положителен, и уменьшаться, если отрицателен. Метод останавливается, когда значение производной будет «близко» к нулю.

К минусам этого метода можно отнести его «нечувствительность» к скорости изменения значений функции. Следующие три метода учитывают, как быстро логарифмическая функция правдоподобия меняется [49]. Нельзя отдать предпочтение ни одному из ниже описанных методов, поскольку скорость работы алгоритмов зависит от обрабатываемых данных.

Метод Ньютона. В качестве матрицы D_t в методе Ньютона выбирается $\left(\frac{\partial^2 \ln L}{\partial \beta^{(t)T} \partial \beta^{(t)}}\right)^{-1}$. Например, гессиан логарифмической функции правдоподобия с двумя параметрами $\beta^T = (\beta_0, \beta_1)$ выглядит следующим образом:

$$\frac{\partial^2 \ln L}{\partial \beta^T \partial \beta} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_0} & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} \\ \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_1} \end{pmatrix}.$$

Если $\partial^2 \ln L / \partial \beta_0 \partial \beta_0$ больше $\partial^2 \ln L / \partial \beta_1 \partial \beta_1$, то градиент меняется быстрее при возрастании β_0 , чем при возрастании на то же самое число аргумента β_1 .

Следующий элемент приближения в методе Ньютона выбирается по правилу:

$$\beta^{t+1} = \beta^{(t)} - \left(\frac{\partial^2 \ln L}{\partial \beta^{(t)T} \partial \beta^{(t)}}\right)^{-1} \frac{\partial \ln L}{\partial \beta^{(t)}},$$

где $\frac{\partial^2 \ln L}{\partial \beta^{(t)T} \partial \beta^{(t)}} = \frac{\partial^2 \ln L}{\partial \beta^T \partial \beta} \Big|_{\beta=\beta^{(t)}}$.

Method scoring. В некоторых случаях математическое ожидание гессiana, известное как информационная матрица, проще вычислить, чем сам гессиан. Метод scoring использует информационную матрицу как матрицу D_t , и следующее приближение выбирается по правилу:

$$\beta^{t+1} = \beta^{(t)} + \left(E \left[\frac{\partial^2 \ln L}{\partial \beta^{(t)} \partial \beta^{(t)T}} \right] \right)^{-1} \frac{\partial \ln L}{\partial \beta^{(t)}}.$$

Метод ВННН (названный по первым буквам авторов E. Berndt, B. Hall, R. Hall, J. Hausman [42]) — численный метод, имеющий широкое распространение в эконометрических задачах. Так как гессиан и информационную матрицу в некоторых случаях сложно вычислить, то в качестве матрицы D_t можно использовать произведение градиентов:

$$\sum_{i=1}^n \frac{\partial \ln L_i}{\partial \beta^{(t)}} \left(\frac{\partial \ln L_i}{\partial \beta^{(t)}} \right)^T,$$

где $\ln L_i$ — значение логарифмической функции правдоподобия, вычисленное в i -ом наблюдении.

Тогда уравнение перехода будет иметь вид:

$$\beta^{(t+1)} = \beta^{(t)} + \left(\sum_{i=1}^n \frac{\partial \ln L_i}{\partial \beta^{(t)}} \left(\frac{\partial \ln L_i}{\partial \beta^{(t)}} \right)^T \right)^{-1} \frac{\partial \ln L}{\partial \beta^{(t)}}.$$

Этот метод также называют модифицированным методом scoring.

Оценки параметров модели, найденные методом максимального правдоподобия, асимптотически подчиняются нормальному распределению [46]:

$$\hat{\beta} \stackrel{a}{\sim} N(\beta, V(\hat{\beta})),$$

где $V(\hat{\beta})$ — ковариационная матрица вектора $\hat{\beta}$. Например, для модели с двумя независимыми переменными ковариационная матрица будет выглядеть следующим образом:

$$V \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} V(\hat{\beta}_0) & Cov(\hat{\beta}_0, \hat{\beta}_1) & Cov(\hat{\beta}_0, \hat{\beta}_2) \\ Cov(\hat{\beta}_1, \hat{\beta}_0) & V(\hat{\beta}_1) & Cov(\hat{\beta}_1, \hat{\beta}_2) \\ Cov(\hat{\beta}_2, \hat{\beta}_0) & Cov(\hat{\beta}_2, \hat{\beta}_1) & V(\hat{\beta}_2) \end{pmatrix},$$

где $Cov(\hat{\beta}_i, \hat{\beta}_j)$ — ковариация оценок $\hat{\beta}_i$ и $\hat{\beta}_j$.

Вышеописанные вычислительные методы позволяют попутно получать оценку для ковариационной матрицы $V(\hat{\beta})$, которая будет необходима для проверки статистических гипотез о значении параметров регрессии, рассмотренных ниже в § 7.

По свойству асимптотической нормальности и асимптотической эффективности оценки максимального правдоподобия $\hat{\beta}$ ковариационная матрица $V(\hat{\beta})$ равна:

$$V(\hat{\beta}) \approx I^{-1}(\beta) = \left(-E \left[\frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} \right] \right)^{-1}, \quad (6.1)$$

т. е. ковариационная матрица равна обратному значению информационной матрицы (ожидаемое значение гесссиана, взятое с противоположным знаком).

Выражение в правой части (6.1) вычисляется в точке β , но, поскольку мы знаем только оценку $\hat{\beta}$ параметра β , то можем найти только оценку для асимптотической ковариационной матрицы, т. е. производные и интегралы вычисляются при $\beta = \hat{\beta}$.

Можно использовать следующие три способа нахождения оценки $\hat{V}(\hat{\beta})$ асимптотической ковариационной матрицы $V(\hat{\beta})$ [46, 49]:

1.

$$\hat{V}_1(\hat{\beta}) = \left(-E \left[\frac{\partial^2 \ln L}{\partial \hat{\beta} \partial \hat{\beta}^T} \right] \right)^{-1}. \quad (6.2)$$

Оценка (6.2) часто используется вместе с методом scoring, т. к. этот численный метод на каждой итерации вычисляет информационную матрицу.

2.

$$\hat{V}_2(\hat{\beta}) = - \left(\sum_{i=1}^n \frac{\partial^2 \ln L_i}{\partial \hat{\beta} \partial \hat{\beta}^T} \right)^{-1}. \quad (6.3)$$

Оценка (6.3) чаще всего используется вместе с численным методом Ньютона.

3.

$$\hat{V}_3(\hat{\beta}) = \left(\sum_{i=1}^n g_i^2 x_i x_i^T \right)^{-1}, \quad (6.4)$$

где $g_i = y_i - \Lambda(\beta x_i)$ для логит модели (см. (5.5)), и $g_i = q_i \varphi(\beta x_i) / \Phi(q_i \beta x_i)$, $q_i = 2y_i - 1$ (см. (5.7)) для пробит модели.

В некоторых случаях возможен альтернативный подход к оцениванию неизвестных параметров. Рассмотрим его на примере логит модели. Предположим, что для каждого набора факторов $x_l^T = (1, x_{1l}, \dots, x_{nl})$ проведено несколько наблюдений, или все наблюдения сгруппированы таким образом, что внутри каждой группы значения факторов меняются мало. Тогда можно заменить различные наборы факторов внутри каждой группы некоторыми средними значениями и, соответственно, все наблюдения внутри группы рассматривать как наблюдения, соответствующие выбранным средним значениям факторов. В обоих случаях появляется возможность оценить эмпирическую вероятность появления единичного значения для соответствующего набора факторов — пусть это будет оценка \hat{p}_l , получаемая как относительная частота (доля) наблюдений, равных единице. Тогда можно применить так называемое логит-преобразование $z_l = \ln(\hat{p}_l / (1 - \hat{p}_l))$, после которого, учитывая модель (3.1) с логистической функцией распределения (3.5), получаем новую модель наблюдений:

$$z_i = \beta^T x_i + \zeta_i, \quad (6.5)$$

где относительно случайной компоненты ζ_i сделаем традиционные предположения о взаимной независимости наблюдений.

Модель (6.5) можно анализировать как модель наблюдений множественной регрессии, и для оценивания коэффициентов модели применить метод наименьших квадратов. Трудности применения метода наименьших квадратов к модели (6.5) связаны с тем, что дисперсии наблюдений не постоянны, и выражения для дисперсий случайных компонент носят приближенный характер. Аналогичный подход возможен и для случая пробит модели, где вместо логит-преобразования применяют обратную функцию к функции стандартного нормального распределения. Возникающие проблемы аналогичны выше рассмотренным.

Сравнивая два подхода к оценке неизвестных параметров модели (3.1), можно согласиться с тем, что метод максималь-

ного правдоподобия выглядит предпочтительнее. Каким бы способом ни оценивались параметры модели (3.1), о качестве построенной модели можно судить по ее способности правильно прогнозировать имеющиеся наблюдения. Подставляя в модель оценку $\hat{\beta}$ и значения факторов x_i , находим оценку вероятности появления единицы: $\hat{P}\{y_i = 1\} = F(\hat{\beta}^T x)$. Если наблюдение оказалось равным единице, то для правильного прогноза найденная вероятность должна принимать значение, большее 0,5. Перебирая имеющиеся наблюдения и определяя соответствие наблюдения вычисленной вероятности, можно оценить качество построенной модели.

§ 7. ПРОВЕРКА ГИПОТЕЗ О ЗНАЧИМОСТИ ПАРАМЕТРОВ ЛОГИТ И ПРОБИТ МОДЕЛЕЙ БИНАРНОГО ВЫБОРА

Для логит и пробит моделей проверка гипотез о наличии ограничений на коэффициенты, в частности, гипотез о значимости одного или группы коэффициентов, может проводиться с помощью любого из трех критериев — Вальда, отношения правдоподобия, множителей Лагранжа [45, 46].

Рассмотрим нулевую гипотезу в виде системы уравнений:

$$H_0 : Q\beta = r, \quad (7.1)$$

где $\beta^T = (\beta_0, \beta_1, \dots, \beta_k)$, Q — матрица констант, r — вектор констант, которые формируются определенным образом в зависимости от того, какую гипотезу необходимо проверить. Например, рассмотрим пробит модель $P\{y = 1\} = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$. Для проверки нулевой гипотезы $H_0 : \beta_1 = 0$ система уравнений (7.1) примет следующий вид:

$$(0 \quad 1 \quad 0) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = (0).$$

Для проверки гипотезы $\beta_1 = \beta_2 = 0$ система уравнений (7.1) примет следующий вид:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Сначала рассмотрим случай, когда $q = 1$, где q — число строк матрицы Q . Пусть в строке все элементы, кроме одного равного единице, равны нулю. В этом случае исследователю требуется проверить гипотезу о значении одного параметра модели, т.е. гипотезу $H_0: \beta_i = \beta^*$, где β^* — некоторое число (часто равное 0). Так как среднеквадратичное отклонение оценки $\hat{\beta}_i$ параметра β_i неизвестно, то для него может быть найдена оценка методами, описанными в § 6.

Будем предполагать, что оценки неизвестных параметров асимптотически нормальны. В качестве статистики критерия рассмотрим функцию:

$$z = \frac{\hat{\beta}_i - \beta^*}{\hat{\sigma}(\hat{\beta}_i)}, \quad (7.2)$$

которая в случае справедливости гипотезы H_0 асимптотически подчиняется стандартному нормальному распределению, если оценки параметров в модели производятся методом максимального правдоподобия [49].

Алгоритм критерия:

1. Выдвигаем нулевую гипотезу $H_0: \beta_i = \beta^*$. Сформулируем альтернативную гипотезу: $H_1: \beta_i \neq \beta^*$.
2. Задаем уровень значимости α .
3. Вычисляем значение статистики z по формуле (7.2).
4. Находим критическую область — это объединение интервалов $(-\infty; -u_{1-\frac{\alpha}{2}}) \cup (u_{1-\frac{\alpha}{2}}; \infty)$, где $u_{1-\alpha/2}$ — квантиль стандартного нормального распределения уровня $1 - \frac{\alpha}{2}$.
5. Если значение статистики (7.2) попадет в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований ее отвергнуть при уровне значимости α .

Замечание 7.1. В некоторых прикладных статистических пакетах в качестве асимптотического распределения статистики (7.2) используется распределение Стьюдента с числом степеней свободы, равным $n - k$. Кроме того, следует отметить, что критерий носит асимптотический характер, истинный уровень значимости близок к α .

Далее рассмотрим основные критерии, используемые для проверки значимости коэффициентов бинарной регрессии для случая, когда $q \geq 1$.

§ 8. КРИТЕРИЙ ВАЛЬДА

Выдвинем нулевую гипотезу $H_0 : Q\beta = r$ при альтернативной гипотезе $H_1 : Q\beta \neq r$.

Пусть мы нашли оценку максимального правдоподобия $\hat{\beta}$ для неизвестного параметра β , и $\hat{V}(\hat{\beta})$ — состоятельная оценка для асимптотической ковариационной матрицы $V(\hat{\beta})$. Статистика критерия Вальда выглядит следующим образом:

$$W = (Q\hat{\beta} - r)^T (Q\hat{V}(\hat{\beta})Q^T)^{-1} (Q\hat{\beta} - r). \quad (8.1)$$

При справедливости нулевой гипотезы статистика (8.1) асимптотически подчиняется распределению χ^2 с числом степеней свободы, равным количеству тестируемых параметров, т.е. равным q [45].

Пример 8.1. Для пробит модели $P\{y = 1\} = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$ проверим гипотезу $H_0 : \beta_1 = \beta^*$ при альтернативной гипотезе $H_1 : \beta_1 \neq \beta^*$ при заданном уровне значимости α . Найдем выражение для W из (8.1) для этого случая. Выражение $Q\hat{\beta} - r$ примет следующий вид:

$$\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} - \beta^* = \hat{\beta}_1 - \beta^*.$$

Выражение $(Q\hat{V}(\hat{\beta})Q^T)^{-1}$ в статистике (8.1) будет следующим:

$$\left[\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \hat{V}(\hat{\beta}) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right]^{-1} = \frac{1}{\hat{V}(\hat{\beta}_1)}.$$

Получаем выражение для статистики (8.1):

$$W = \frac{(\hat{\beta}_1 - \beta^*)^2}{\hat{V}(\hat{\beta}_1)} = \left(\frac{\hat{\beta}_1 - \beta^*}{\hat{\sigma}(\hat{\beta}_1)} \right)^2, \quad (8.2)$$

где $\hat{\beta}_1$ — оценка максимального правдоподобия, асимптотически нормальная и сильно состоятельная; $\hat{\sigma}(\hat{\beta}_1)$ — состоятельная оценка среднеквадратического отклонения $\hat{\beta}_1$. О способах нахождения $\hat{\beta}_1$ и $\hat{\sigma}(\hat{\beta}_1)$ говорилось в § 6. Статистика (8.2) асимптотически подчиняется χ^2 -распределению с одной степенью свободы в случае справедливости гипотезы H_0 . Нетрудно заметить, что статистика W — квадрат значения статистики z в (7.2). В рассматриваемом случае гипотезу H_0 следует отвергнуть при уровне значимости приближенно равном α в случае, когда $W \in [\chi_{1-\alpha,1}^2; \infty)$, где $\chi_{1-\alpha,1}^2$ — квантиль уровня $1 - \alpha$ распределения χ^2 с одной степенью свободы.

Пример 8.2. Для пробит модели $P\{y = 1\} = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$ проверим гипотезу $H_0 : \beta_1 = \beta_2 = 0$ при альтернативной гипотезе H_1 , которая говорит о том, что хотя бы один из параметров (β_1, β_2) отличен от нуля, при уровне значимости приближенно равном α . Гипотеза H_0 может быть записана в следующем виде:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

где матрица $Q\hat{\beta} - r$ и $(Q\hat{V}(\hat{\beta})Q^T)^{-1}$ приобретают вид:

$$Q\hat{\beta} - r = (\hat{\beta}_1, \hat{\beta}_2)^T,$$

$$(Q\hat{V}(\hat{\beta})Q^T)^{-1} = \left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \hat{V}(\hat{\beta}) \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \right]^{-1}.$$

Предположив, что оценки $\hat{\beta}_1, \hat{\beta}_2$ не коррелируют (в общем случае это неверно), получаем равенства:

$$\begin{aligned} (Q\hat{V}(\hat{\beta})Q^T)^{-1} &= \begin{pmatrix} \hat{\sigma}^2(\hat{\beta}_1) & 0 \\ 0 & \hat{\sigma}^2(\hat{\beta}_2) \end{pmatrix}^{-1} = \\ &= \begin{pmatrix} 1/\hat{\sigma}^2(\hat{\beta}_1) & 0 \\ 0 & 1/\hat{\sigma}^2(\hat{\beta}_2) \end{pmatrix}. \end{aligned}$$

Тогда статистика Вальда (8.1) примет вид:

$$W = \sum_{i=1}^2 \left(\frac{\hat{\beta}_i}{\hat{\sigma}(\hat{\beta}_i)} \right)^2 = z_{\hat{\beta}_1}^2 + z_{\hat{\beta}_2}^2.$$

Здесь также статистика Вальда равна сумме квадратов статистик z , вычисленных для тестируемых параметров. В случае, когда оценки параметров коррелируют (что чаще всего встречается в реальных задачах), окончательный вид статистики W имеет более сложный вид.

Алгоритм критерия Вальда:

1. Выдвигаем нулевую гипотезу $H_0 : Q\beta = r$. Сформулируем альтернативную гипотезу $H_1 : Q\beta \neq r$.
2. Задаем уровень значимости критерия α .
3. Находим оценку $\hat{\beta}$ для неизвестного параметра β и оценку $\hat{V}(\hat{\beta})$ для асимптотической ковариационной матрицы.
4. Вычисляем значение статистики W по формуле (8.1).
5. Находим критическую область — интервал $(\chi_{1-\alpha, q}^2; \infty)$, где $\chi_{1-\alpha, q}^2$ — квантиль уровня $1 - \alpha$ распределения χ^2 с q степенями свободы.
6. Если численное значение статистики W попадет в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований ее отвергнуть при уровне значимости приближенно равном α .

Критерий Вальда носит асимптотический характер, и поэтому уровень значимости критерия должен быть близок к α при больших объемах наблюдений.

§ 9. КРИТЕРИЙ ОТНОШЕНИЯ ПРАВДОПОДОБИЯ

Часто для проверки адекватности пробит и логит моделей бинарных регрессий используют критерии, основанные на сравнении значений функции правдоподобия в случае, когда максимизация проводится по всем неизвестным параметрам, и при условии, что $Q\beta = r$.

Пусть $\ln L_1$ — максимальное значение логарифмической функции правдоподобия (5.2) при условии, что максимизация производится по параметру β без ограничений на этот параметр; $\ln L_0$ — максимальное значение логарифмической функции правдоподобия (5.2) при условии, что $Q\beta = r$. Очевидно, что $\ln L_1 \geq \ln L_0$. Статистика отношения правдоподобия (*likelihood ratio*) выглядит следующим образом:

$$LR = 2(\ln L_1 - \ln L_0), \quad (9.1)$$

которая при справедливости нулевой гипотезы асимптотически подчиняется распределению χ^2 с числом степеней свободы, равным q [45, 46].

Алгоритм критерия правдоподобия:

1. Выдвигаем нулевую гипотезу $H_0 : Q\beta = r$. Сформулируем альтернативную гипотезу $H_1 : Q\beta \neq r$.
2. Задаем уровень значимости α .
3. Находим значение функции правдоподобия $\ln L_1$ в точке $\hat{\beta}$, которая является оценкой максимального правдоподобия для неизвестного параметра β в задаче без ограничений и $\ln L_0$.
4. Вычисляем значение статистики по формуле (9.1).
5. Находим критическую область — интервал $(\chi_{q,1-\alpha}^2; \infty)$, где $\chi_{1-\alpha,q}^2$ — квантиль уровня $1 - \alpha$ распределения χ^2 с q степенями свободы, где q — число строк матрицы Q .
6. Если численное значение статистики (9.1) попадет в критическую область, то нулевая гипотеза H_0 отвергается, в противном случае нет оснований ее отвергнуть при уровне значимости приближенно равном α .

Критерий отношения правдоподобия носит асимптотический характер, уровень значимости критерия должен быть близок к α при больших n .

Подробное описание критерия множителей Лагранжа для проверки гипотезы $H_0: Q\beta = r$ можно найти в [45, 46, 49].

§ 10. КРИТЕРИИ АДЕКВАТНОСТИ МОДЕЛЕЙ БИНАРНОЙ РЕГРЕССИИ

В настоящее время предложено большое количество мер адекватности для моделей бинарной регрессии [46, 49, 51], приведем некоторые из них.

Сумма квадратов остатков SSR вычисляется по формуле:

$$SSR = \sum_{i=1}^n (y_i - \hat{F}_i)^2, \quad (10.1)$$

где $\hat{F}_i = F(\hat{\beta}^T x_i)$. Значение *SSR* является часто используемой мерой, поскольку она используется для вычисления коэффициента детерминации R^2 в моделях линейной регрессии. Тем не менее, использование этой меры не может быть математически строго обосновано, поскольку модели бинарной регрессии не удовлетворяют условию равенства дисперсий [39]. В. Efron предложил аналог R^2 следующего вида [44]:

$$R_{Ef}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{F}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (10.2)$$

где $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

Взвешенная сумма квадратов WSSR для моделей бинарной регрессии может быть вычислена по формуле:

$$WSSR = \sum_{i=1}^n \frac{(y_i - \hat{F}_i)^2}{\hat{F}_i(1 - \hat{F}_i)}. \quad (10.3)$$

Как утверждается в работе [39], критерий (10.3) более предпочтителен, чем критерий (10.1).

Квадратичный коэффициент корреляции SCC вычисляется по формуле:

$$SCC = \frac{\left[\sum_{i=1}^n (y_i - \bar{y}) \hat{F}_i \right]^2}{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{F}_i - \bar{F})^2}, \quad (10.4)$$

где $\bar{F} = \sum_{i=1}^n \hat{F}_i / n$. В случае линейной регрессии коэффициенты (10.2) и (10.4) совпадают, но это не является верным для случая бинарной регрессии.

Существует еще одна мера адекватности моделей бинарной регрессии [41, 48]:

$$R_{BL}^2 = \frac{1}{n} \sum_{i=1}^n \left(y_i \hat{F}_i + (1 - y_i)(1 - \hat{F}_i) \right), \quad (10.5)$$

которая представляет собой среднюю вероятность правильного предсказания в соответствии с полученным правилом.

Существуют меры адекватности моделей бинарной регрессии, основанные на сравнении значений функции правдоподобия при различных ограничениях. Например, D. MacFadden предложил индекс отношения правдоподобия следующего вида [50]:

$$LRI = 1 - \frac{\ln L(\hat{\beta})}{\ln L_0}, \quad (10.6)$$

где $\ln L_0$ — максимальное значение логарифмической функции правдоподобия при $\beta_1 = \dots = \beta_k = 0$.

С практической точки зрения полезно составить таблицу 2×2 истинных и ложных прогнозов по следующему правилу:

$$\hat{y} = \begin{cases} 1, & \text{если } \hat{F} > F^*, \\ 0, & \text{в противном случае.} \end{cases} \quad (10.7)$$

Обычно F^* выбирают равным 0,5, это означает, что мы прогнозируем 1 в случае, когда модель «говорит», что 1 более вероятна, чем 0. Например, в таблице 19.2 количество верных прогнозов равно $271 + 58 = 329$, т. е. модель, которой соответствует таблица 19.2, корректно прогнозирует 91% данных.

	$\hat{y}_i = 1$	$\hat{y}_i = 0$	Кол-во
$y_i = 1$	271	14	285
$y_i = 0$	17	58	75
	288	72	360

Таблица 19.2
Результаты прогнозирования

Замечание 10.1. Различные скалярные меры адекватности моделей бинарной регрессии дают различные результаты [39]. Оценка максимального правдоподобия, на которой основаны все выше перечисленные скалярные меры адекватности для моделей бинарной регрессии, не выбирается из условия максимизации критерия адекватности, в отличие от классической модели линейной регрессии (коэффициенты регрессии, найденные методом наименьших квадратов, максимизируют коэффициент детерминации R^2). В случае бинарной регрессии оценка максимального правдоподобия $\hat{\beta}$ максимизирует совместную плотность распределения наблюдаемых случайных величин. Возникает вопрос для исследователя: выбрать лучшую оценку параметров при возможно низком уровне достоверного прогноза или получить наилучшую оценку параметров, максимизирующую выбранную скалярную меру адекватности модели, которая чаще всего не будет являться оценкой максимального правдоподобия?

ТАБЛИЦА ЗНАЧЕНИЙ ФУНКЦИИ СТАНДАРТНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Приложение Б.

ТАБЛИЦА КВАНТИЛЕЙ $\chi^2_{\alpha,n}$ УРОВНЯ α
РАСПРЕДЕЛЕНИЯ χ^2 С n СТЕПЕНЯМИ СВОБОДЫ

$$\frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \int_0^{\chi^2_{\alpha,n}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = \alpha$$

α	0.001	0.005	0.010	0.025	0.050	0.10	0.125	0.200	0.250	0.333	0.500
n											
1	0.000	0.000	0.000	0.001	0.004	0.016	0.025	0.064	0.102	0.186	0.455
2	0.002	0.010	0.020	0.051	0.103	0.211	0.267	0.446	0.575	0.811	1.386
3	0.024	0.072	0.115	0.216	0.352	0.584	0.692	1.005	1.213	1.568	2.366
4	0.091	0.207	0.297	0.484	0.711	1.064	1.219	1.649	1.923	2.378	3.357
5	0.210	0.412	0.554	0.831	1.145	1.610	1.808	2.343	2.675	3.216	4.351
6	0.381	0.676	0.872	1.237	1.635	2.204	2.441	3.070	3.455	4.074	5.348
7	0.598	0.989	1.239	1.690	2.167	2.833	3.106	3.822	4.255	4.945	6.346
8	0.857	1.344	1.646	2.180	2.733	3.490	3.797	4.594	5.071	5.826	7.344
9	1.152	1.735	2.088	2.700	3.325	4.168	4.507	5.380	5.899	6.716	8.343
10	1.479	2.156	2.558	3.247	3.940	4.865	5.234	6.179	6.737	7.612	9.342
11	1.834	2.603	3.053	3.816	4.575	5.578	5.975	6.989	7.584	8.514	10.341
12	2.214	3.074	3.571	4.404	5.226	6.304	6.729	7.807	8.438	9.420	11.340
13	2.617	3.565	4.107	5.009	5.892	7.042	7.493	8.634	9.299	10.331	12.340
14	3.041	4.075	4.660	5.629	6.571	7.790	8.266	9.467	10.165	11.245	13.339
15	3.483	4.601	5.229	6.262	7.261	8.547	9.048	10.307	11.037	12.163	14.339
16	3.942	5.142	5.812	6.908	7.962	9.312	9.837	11.152	11.912	13.083	15.338
17	4.416	5.697	6.408	7.564	8.672	10.085	10.633	12.002	12.792	14.006	16.338
18	4.905	6.265	7.015	8.231	9.390	10.865	11.435	12.857	13.675	14.931	17.338
19	5.407	6.844	7.633	8.907	10.117	11.651	12.242	13.716	14.562	15.859	18.338
20	5.921	7.434	8.260	9.591	10.851	12.443	13.055	14.578	15.452	16.788	19.337
21	6.447	8.034	8.897	10.283	11.591	13.240	13.873	15.445	16.344	17.720	20.337
22	6.983	8.643	9.542	10.982	12.338	14.041	14.695	16.314	17.240	18.653	21.337
23	7.529	9.260	10.196	11.689	13.091	14.848	15.521	17.187	18.137	19.587	22.337
24	8.085	9.886	10.856	12.401	13.848	15.659	16.351	18.062	19.037	20.523	23.337
25	8.649	10.520	11.524	13.120	14.611	16.473	17.184	18.940	19.939	21.461	24.337
26	9.222	11.160	12.198	13.844	15.379	17.292	18.021	19.820	20.843	22.399	25.336
27	9.803	11.808	12.879	14.573	16.151	18.114	18.861	20.703	21.749	23.339	26.336
28	10.391	12.461	13.565	15.308	16.928	18.939	19.704	21.588	22.657	24.280	27.336
29	10.986	13.121	14.256	16.047	17.708	19.768	20.550	22.475	23.567	25.222	28.336
30	11.588	13.787	14.953	16.791	18.493	20.599	21.399	23.364	24.478	26.165	29.336
35	14.688	17.192	18.509	20.569	22.465	24.797	25.678	27.836	29.054	30.894	34.336
40	17.916	20.707	22.164	24.433	26.509	29.051	30.008	32.345	33.660	35.643	39.335
45	21.251	24.311	25.901	28.366	30.612	33.350	34.379	36.884	38.291	40.407	44.335
50	24.674	27.991	29.707	32.357	34.764	37.689	38.785	41.449	42.942	45.184	49.335
55	28.173	31.735	33.570	36.398	38.958	42.060	43.220	46.036	47.610	49.972	54.335
60	31.738	35.534	37.485	40.482	43.188	46.459	47.680	50.641	52.294	54.770	59.335

α	0.600	0.667	0.750	0.800	0.875	0.900	0.950	0.975	0.990	0.995	0.999
n											
1	0.708	0.936	1.323	1.642	2.354	2.706	3.841	5.024	6.635	7.879	10.828
2	1.833	2.197	2.773	3.219	4.159	4.605	5.991	7.378	9.210	10.597	13.816
3	2.946	3.405	4.108	4.642	5.739	6.251	7.815	9.348	11.345	12.838	16.266
4	4.045	4.579	5.385	5.989	7.214	7.779	9.488	11.143	13.277	14.860	18.467
5	5.132	5.730	6.626	7.289	8.625	9.236	11.070	12.833	15.086	16.750	20.515
6	6.211	6.867	7.841	8.558	9.992	10.645	12.592	14.449	16.812	18.548	22.458
7	7.283	7.992	9.037	9.803	11.326	12.017	14.067	16.013	18.475	20.278	24.322
8	8.351	9.107	10.219	11.030	12.636	13.362	15.507	17.535	20.090	21.955	26.125
9	9.414	10.215	11.389	12.242	13.926	14.684	16.919	19.023	21.666	23.589	27.877
10	10.473	11.317	12.549	13.442	15.198	15.987	18.307	20.483	23.209	25.188	29.588
11	11.530	12.414	13.701	14.631	16.457	17.275	19.675	21.920	24.725	26.757	31.264
12	12.584	13.506	14.845	15.812	17.703	18.549	21.026	23.337	26.217	28.300	32.910
13	13.636	14.595	15.984	16.985	18.939	19.812	22.362	24.736	27.688	29.819	34.528
14	14.685	15.680	17.117	18.151	20.166	21.064	23.685	26.119	29.141	31.319	36.123
15	15.733	16.761	18.245	19.311	21.384	22.307	24.996	27.488	30.578	32.801	37.697
16	16.780	17.840	19.369	20.465	22.595	23.542	26.296	28.845	32.000	34.267	39.252
17	17.824	18.917	20.489	21.615	23.799	24.769	27.587	30.191	33.409	35.718	40.790
18	18.868	19.991	21.605	22.760	24.997	25.989	28.869	31.526	34.805	37.156	42.312
19	19.910	21.063	22.718	23.900	26.189	27.204	30.144	32.852	36.191	38.582	43.820
20	20.951	22.133	23.828	25.038	27.376	28.412	31.410	34.170	37.566	39.997	45.315
21	21.991	23.201	24.935	26.171	28.559	29.615	32.671	35.479	38.932	41.401	46.797
22	23.031	24.268	26.039	27.301	29.737	30.813	33.924	36.781	40.289	42.796	48.268
23	24.069	25.333	27.141	28.429	30.911	32.007	35.172	38.076	41.638	44.181	49.728
24	25.106	26.397	28.241	29.553	32.081	33.196	36.415	39.364	42.980	45.559	51.179
25	26.143	27.459	29.339	30.675	33.247	34.382	37.652	40.646	44.314	46.928	52.620
26	27.179	28.520	30.435	31.795	34.410	35.563	38.885	41.923	45.642	48.290	54.052
27	28.214	29.580	31.528	32.912	35.570	36.741	40.113	43.195	46.963	49.645	55.476
28	29.249	30.639	32.620	34.027	36.727	37.916	41.337	44.461	48.278	50.993	56.892
29	30.283	31.697	33.711	35.139	37.881	39.087	42.557	45.722	49.588	52.336	58.301
30	31.316	32.754	34.800	36.250	39.033	40.256	43.773	46.979	50.892	53.672	59.703
35	36.475	38.024	40.223	41.778	44.753	46.059	49.802	53.203	57.342	60.275	66.619
40	41.622	43.275	45.616	47.269	50.424	51.805	55.758	59.342	63.691	66.766	73.402
45	46.761	48.510	50.985	52.729	56.052	57.505	61.656	65.410	69.957	73.166	80.077
50	51.892	53.733	56.334	58.164	61.647	63.167	67.505	71.420	76.154	79.490	86.661
55	57.016	58.945	61.665	63.577	67.211	68.796	73.311	77.380	82.292	85.749	93.168
60	62.135	64.147	66.981	68.972	72.751	74.397	79.082	83.298	88.379	91.952	99.607

Приложение В.

ТАБЛИЦА КВАНТИЛЕЙ $T_{\alpha,n}$ УРОВНЯ α РАСПРЕДЕЛЕНИЯ СТЬЮДЕНТА С n СТЕПЕНЯМИ СВОБОДЫ

$$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \int_{-\infty}^{t_{\alpha,n}} \frac{dx}{(1 + \frac{x^2}{n})^{\frac{n+1}{2}}} = \alpha$$

α	0.600	0.667	0.750	0.800	0.875	0.900	0.950	0.975	0.990	0.995	0.999
n											
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.256	0.436	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.256	0.435	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.256	0.435	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.256	0.435	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.256	0.435	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385
35	0.255	0.434	0.682	0.852	1.170	1.306	1.690	2.030	2.438	2.724	3.340
40	0.255	0.434	0.681	0.851	1.167	1.303	1.684	2.021	2.423	2.704	3.307
45	0.255	0.434	0.680	0.850	1.165	1.301	1.679	2.014	2.412	2.690	3.281
50	0.255	0.433	0.679	0.849	1.164	1.299	1.676	2.009	2.403	2.678	3.261
55	0.255	0.433	0.679	0.848	1.163	1.297	1.673	2.004	2.396	2.668	3.245
60	0.254	0.433	0.679	0.848	1.162	1.296	1.671	2.000	2.390	2.660	3.232
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Приложение Г.

ТАБЛИЦА КВАНТИЛЕЙ $F_{\alpha,n,m}$ УРОВНЯ α
РАСПРЕДЕЛЕНИЯ ФИШЕРА С n И m СТЕПЕНЯМИ
СВОБОДЫ

$$\int_0^{F_{\alpha,n,m}} \frac{\sqrt{\frac{(nx)^n m^m}{(nx+m)^{n+m}}}}{xB\left(\frac{n}{2}, \frac{m}{2}\right)} dx = \alpha, \quad F_{\alpha,n,m} \equiv \frac{1}{F_{1-\alpha,m,n}}$$

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
1	0.500	1.50	1.71	1.82	1.89	1.94	1.98	2.00	2.04	2.07	2.09	2.12	2.15	2.17	2.20
	0.600	2.63	2.93	3.09	3.20	3.27	3.32	3.36	3.41	3.45	3.48	3.52	3.56	3.59	3.64
	0.667	4.00	4.42	4.64	4.78	4.88	4.95	5.00	5.08	5.13	5.18	5.24	5.29	5.33	5.39
	0.750	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.32	9.41	9.50	9.58	9.67	9.74	9.85
	0.800	12.0	13.1	13.6	14.0	14.3	14.4	14.6	14.8	14.9	15.0	15.2	15.3	15.4	15.6
2	0.500	1.00	1.13	1.21	1.25	1.28	1.30	1.32	1.35	1.36	1.38	1.39	1.41	1.42	1.44
	0.600	1.50	1.64	1.72	1.76	1.80	1.82	1.84	1.86	1.88	1.89	1.91	1.92	1.94	1.96
	0.667	2.00	2.15	2.22	2.27	2.30	2.33	2.34	2.37	2.38	2.40	2.42	2.43	2.45	2.47
	0.750	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.38	3.39	3.41	3.43	3.44	3.46	3.48
	0.800	4.00	4.16	4.24	4.28	4.32	4.34	4.36	4.38	4.40	4.42	4.43	4.45	4.47	4.48
3	0.500	0.88	1.00	1.06	1.10	1.13	1.15	1.16	1.18	1.20	1.21	1.23	1.24	1.25	1.27
	0.600	1.26	1.37	1.43	1.47	1.49	1.51	1.52	1.54	1.55	1.56	1.57	1.58	1.59	1.60
	0.667	1.62	1.72	1.77	1.80	1.82	1.83	1.84	1.86	1.87	1.88	1.89	1.90	1.90	1.91
	0.750	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.45	2.46	2.46	2.47	2.47	2.47
	0.800	2.89	2.94	2.96	2.97	2.97	2.97	2.98	2.98	2.98	2.98	2.98	2.98	2.98	2.98
4	0.500	0.83	0.94	1.00	1.04	1.06	1.08	1.09	1.11	1.13	1.14	1.15	1.16	1.18	1.19
	0.600	1.16	1.26	1.31	1.34	1.36	1.37	1.38	1.40	1.41	1.42	1.43	1.43	1.44	1.45
	0.667	1.46	1.55	1.58	1.61	1.62	1.63	1.64	1.65	1.65	1.66	1.67	1.67	1.68	1.68
	0.750	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
	0.800	2.47	2.48	2.48	2.48	2.47	2.47	2.47	2.46	2.46	2.45	2.44	2.44	2.43	2.43
5	0.500	0.80	0.91	0.96	1.00	1.02	1.04	1.05	1.07	1.09	1.10	1.11	1.12	1.13	1.15
	0.600	1.11	1.20	1.24	1.27	1.29	1.30	1.31	1.32	1.33	1.34	1.34	1.35	1.36	1.37
	0.667	1.38	1.45	1.48	1.50	1.51	1.52	1.53	1.53	1.54	1.54	1.54	1.55	1.55	1.55
	0.750	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.88	1.88	1.88	1.87
	0.800	2.26	2.25	2.24	2.23	2.22	2.21	2.20	2.19	2.18	2.18	2.17	2.16	2.15	2.13
6	0.500	0.78	0.89	0.94	0.98	1.00	1.02	1.03	1.05	1.06	1.07	1.08	1.10	1.11	1.12
	0.600	1.07	1.16	1.20	1.22	1.24	1.25	1.26	1.27	1.28	1.29	1.29	1.30	1.31	1.31
	0.667	1.33	1.39	1.42	1.44	1.44	1.45	1.45	1.46	1.46	1.47	1.47	1.47	1.47	1.47
	0.750	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.76	1.76	1.75	1.75	1.74
	0.800	2.13	2.11	2.09	2.08	2.06	2.05	2.04	2.03	2.02	2.01	2.00	1.98	1.97	1.95
7	0.500	0.77	0.87	0.93	0.96	0.98	1.00	1.01	1.03	1.04	1.05	1.07	1.08	1.09	1.10
	0.600	1.05	1.13	1.17	1.19	1.21	1.22	1.23	1.24	1.24	1.25	1.26	1.26	1.27	1.27
	0.667	1.29	1.35	1.38	1.39	1.40	1.40	1.41	1.41	1.41	1.41	1.41	1.42	1.42	1.42
	0.750	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.69	1.68	1.68	1.67	1.66	1.66	1.65
	0.800	2.04	2.02	1.99	1.97	1.96	1.94	1.93	1.92	1.91	1.89	1.88	1.86	1.85	1.83
8	0.500	0.76	0.86	0.91	0.95	0.97	0.99	1.00	1.02	1.03	1.04	1.05	1.07	1.07	1.09
	0.600	1.03	1.11	1.15	1.17	1.19	1.20	1.20	1.21	1.22	1.22	1.23	1.24	1.24	1.25
	0.667	1.26	1.32	1.35	1.36	1.36	1.37	1.37	1.37	1.37	1.38	1.38	1.38	1.37	1.37
	0.750	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.62	1.62	1.61	1.60	1.59	1.58
	0.800	1.98	1.95	1.92	1.90	1.88	1.87	1.86	1.84	1.83	1.81	1.80	1.78	1.76	1.74

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
9	0.500	0.75	0.85	0.91	0.94	0.96	0.98	0.99	1.01	1.02	1.03	1.04	1.05	1.06	1.08
	0.600	1.02	1.10	1.13	1.15	1.17	1.18	1.18	1.19	1.20	1.21	1.21	1.22	1.22	1.22
	0.667	1.24	1.30	1.32	1.33	1.34	1.34	1.34	1.34	1.35	1.35	1.35	1.34	1.34	1.34
	0.750	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.58	1.57	1.56	1.55	1.54	1.53
	0.800	1.93	1.90	1.87	1.85	1.83	1.81	1.80	1.78	1.76	1.75	1.73	1.71	1.70	1.67
10	0.500	0.74	0.85	0.90	0.93	0.95	0.97	0.98	1.00	1.01	1.02	1.03	1.05	1.06	1.07
	0.600	1.01	1.08	1.12	1.14	1.15	1.16	1.17	1.18	1.18	1.19	1.19	1.20	1.20	1.21
	0.667	1.23	1.28	1.30	1.31	1.32	1.32	1.32	1.32	1.32	1.32	1.32	1.32	1.32	1.31
	0.750	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.55	1.54	1.53	1.52	1.51	1.50	1.48
	0.800	1.90	1.86	1.83	1.80	1.78	1.77	1.75	1.73	1.72	1.70	1.68	1.66	1.65	1.62
11	0.500	0.74	0.84	0.89	0.93	0.95	0.96	0.98	0.99	1.01	1.02	1.03	1.04	1.05	1.06
	0.600	1.00	1.07	1.11	1.13	1.14	1.15	1.16	1.17	1.17	1.18	1.18	1.18	1.19	1.19
	0.667	1.22	1.27	1.29	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.29
	0.750	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.45
	0.800	1.87	1.83	1.80	1.77	1.75	1.73	1.72	1.69	1.68	1.66	1.64	1.62	1.60	1.57
12	0.500	0.73	0.84	0.89	0.92	0.94	0.96	0.97	0.99	1.00	1.01	1.02	1.03	1.04	1.06
	0.600	0.99	1.07	1.10	1.12	1.13	1.14	1.15	1.16	1.16	1.17	1.17	1.17	1.18	1.18
	0.667	1.21	1.26	1.27	1.28	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.28	1.28	1.27
	0.750	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.45	1.44	1.42
	0.800	1.85	1.80	1.77	1.74	1.72	1.70	1.69	1.66	1.65	1.63	1.61	1.59	1.57	1.54
13	0.500	0.73	0.83	0.88	0.92	0.94	0.96	0.97	0.98	1.00	1.01	1.02	1.03	1.04	1.05
	0.600	0.98	1.06	1.09	1.11	1.13	1.13	1.14	1.15	1.15	1.16	1.16	1.16	1.17	1.17
	0.667	1.20	1.25	1.26	1.27	1.28	1.28	1.28	1.28	1.28	1.28	1.27	1.27	1.27	1.26
	0.750	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.43	1.42	1.40
	0.800	1.83	1.78	1.75	1.72	1.69	1.68	1.66	1.64	1.62	1.60	1.58	1.56	1.54	1.51
14	0.500	0.73	0.83	0.88	0.91	0.94	0.95	0.96	0.98	0.99	1.00	1.01	1.03	1.04	1.05
	0.600	0.98	1.05	1.09	1.11	1.12	1.13	1.13	1.14	1.14	1.15	1.15	1.16	1.16	1.16
	0.667	1.19	1.24	1.26	1.26	1.27	1.27	1.27	1.27	1.27	1.26	1.26	1.26	1.25	1.24
	0.750	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.41	1.40	1.38
	0.800	1.81	1.76	1.73	1.70	1.67	1.65	1.64	1.62	1.60	1.58	1.56	1.53	1.51	1.48
15	0.500	0.73	0.83	0.88	0.91	0.93	0.95	0.96	0.98	0.99	1.00	1.01	1.02	1.03	1.05
	0.600	0.97	1.05	1.08	1.10	1.11	1.12	1.13	1.13	1.14	1.14	1.15	1.15	1.15	1.15
	0.667	1.18	1.23	1.25	1.25	1.26	1.26	1.26	1.26	1.26	1.25	1.25	1.25	1.24	1.23
	0.750	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.41	1.40	1.38	1.36
	0.800	1.80	1.75	1.71	1.68	1.66	1.64	1.62	1.60	1.58	1.56	1.54	1.51	1.49	1.46
16	0.500	0.72	0.82	0.88	0.91	0.93	0.95	0.96	0.97	0.99	1.00	1.01	1.02	1.03	1.04
	0.600	0.97	1.04	1.08	1.10	1.11	1.12	1.12	1.13	1.13	1.14	1.14	1.14	1.14	1.14
	0.667	1.18	1.22	1.24	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.24	1.24	1.23	1.22
	0.750	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.43	1.41	1.40	1.38	1.37	1.34
	0.800	1.78	1.74	1.70	1.67	1.64	1.62	1.61	1.58	1.56	1.54	1.52	1.49	1.47	1.43
17	0.500	0.72	0.82	0.87	0.91	0.93	0.94	0.96	0.97	0.98	0.99	1.01	1.02	1.03	1.04
	0.600	0.97	1.04	1.07	1.09	1.10	1.11	1.12	1.12	1.13	1.13	1.13	1.14	1.14	1.14
	0.667	1.17	1.22	1.23	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.23	1.23	1.22	1.21
	0.750	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.41	1.40	1.39	1.37	1.36	1.33
	0.800	1.77	1.72	1.68	1.65	1.63	1.61	1.59	1.57	1.55	1.53	1.50	1.48	1.46	1.42

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
18	0.500	0.720	0.820	0.870	0.900	0.930	0.940	0.950	0.970	0.980	0.990	1.000	1.020	1.020	1.040
	0.600	0.960	1.040	1.070	1.090	1.100	1.110	1.110	1.120	1.120	1.130	1.130	1.130	1.130	1.130
	0.667	1.170	1.210	1.230	1.240	1.240	1.240	1.240	1.240	1.230	1.230	1.230	1.220	1.220	1.210
	0.750	1.500	1.490	1.480	1.460	1.450	1.440	1.430	1.420	1.400	1.390	1.380	1.360	1.340	1.320
	0.800	1.760	1.710	1.670	1.640	1.620	1.600	1.580	1.550	1.530	1.510	1.490	1.460	1.440	1.400
19	0.500	0.720	0.820	0.870	0.900	0.920	0.940	0.950	0.970	0.980	0.990	1.000	1.010	1.020	1.040
	0.600	0.960	1.030	1.070	1.090	1.100	1.100	1.110	1.120	1.120	1.120	1.130	1.130	1.130	1.130
	0.667	1.160	1.210	1.220	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.220	1.220	1.210	1.200
	0.750	1.490	1.490	1.470	1.460	1.440	1.430	1.420	1.410	1.400	1.380	1.370	1.350	1.330	1.300
	0.800	1.750	1.700	1.660	1.630	1.610	1.580	1.570	1.540	1.520	1.500	1.480	1.450	1.430	1.390
20	0.500	0.720	0.820	0.870	0.900	0.920	0.940	0.950	0.970	0.980	0.990	1.000	1.010	1.020	1.030
	0.600	0.960	1.030	1.060	1.080	1.090	1.100	1.110	1.110	1.120	1.120	1.120	1.120	1.120	1.120
	0.667	1.160	1.210	1.220	1.230	1.230	1.230	1.230	1.230	1.220	1.220	1.220	1.210	1.200	1.190
	0.750	1.490	1.480	1.470	1.450	1.440	1.430	1.420	1.400	1.390	1.370	1.360	1.340	1.320	1.290
	0.800	1.750	1.700	1.650	1.620	1.600	1.580	1.560	1.530	1.510	1.490	1.470	1.440	1.410	1.370
21	0.500	0.720	0.810	0.870	0.900	0.920	0.940	0.950	0.960	0.980	0.990	1.000	1.010	1.020	1.030
	0.600	0.960	1.030	1.060	1.080	1.090	1.100	1.100	1.110	1.110	1.120	1.120	1.120	1.120	1.120
	0.667	1.160	1.200	1.220	1.220	1.220	1.220	1.220	1.220	1.220	1.220	1.210	1.200	1.200	1.190
	0.750	1.480	1.480	1.460	1.440	1.430	1.420	1.410	1.390	1.380	1.370	1.350	1.330	1.320	1.280
	0.800	1.740	1.690	1.650	1.610	1.590	1.570	1.550	1.520	1.500	1.480	1.460	1.430	1.400	1.360
22	0.500	0.720	0.810	0.870	0.900	0.920	0.930	0.950	0.960	0.970	0.990	1.000	1.010	1.020	1.030
	0.600	0.960	1.030	1.060	1.080	1.090	1.100	1.100	1.110	1.110	1.110	1.120	1.120	1.120	1.120
	0.667	1.160	1.200	1.210	1.220	1.220	1.220	1.220	1.220	1.210	1.210	1.210	1.200	1.190	1.180
	0.750	1.480	1.470	1.450	1.440	1.420	1.410	1.400	1.390	1.370	1.360	1.340	1.320	1.310	1.280
	0.800	1.730	1.680	1.640	1.610	1.580	1.560	1.540	1.510	1.490	1.470	1.450	1.420	1.390	1.350
23	0.500	0.710	0.810	0.860	0.900	0.920	0.930	0.950	0.960	0.970	0.980	1.000	1.010	1.020	1.030
	0.600	0.950	1.020	1.060	1.070	1.090	1.090	1.100	1.100	1.110	1.110	1.110	1.110	1.110	1.110
	0.667	1.150	1.200	1.210	1.220	1.220	1.220	1.220	1.210	1.210	1.210	1.200	1.190	1.190	1.170
	0.750	1.470	1.470	1.450	1.430	1.420	1.410	1.400	1.380	1.370	1.350	1.340	1.320	1.300	1.270
	0.800	1.730	1.680	1.630	1.600	1.570	1.550	1.530	1.510	1.490	1.460	1.440	1.410	1.380	1.340
24	0.500	0.710	0.810	0.860	0.900	0.920	0.930	0.940	0.960	0.970	0.980	0.990	1.010	1.010	1.030
	0.600	0.950	1.020	1.060	1.070	1.080	1.090	1.100	1.100	1.100	1.110	1.110	1.110	1.110	1.110
	0.667	1.150	1.190	1.210	1.210	1.210	1.210	1.210	1.210	1.210	1.200	1.200	1.190	1.180	1.170
	0.750	1.470	1.460	1.440	1.430	1.410	1.400	1.390	1.380	1.360	1.350	1.330	1.310	1.290	1.260
	0.800	1.720	1.670	1.630	1.590	1.570	1.550	1.530	1.500	1.480	1.460	1.430	1.400	1.380	1.330
25	0.500	0.710	0.810	0.860	0.890	0.920	0.930	0.940	0.960	0.970	0.980	0.990	1.000	1.010	1.030
	0.600	0.950	1.020	1.050	1.070	1.080	1.090	1.090	1.100	1.100	1.110	1.110	1.110	1.110	1.110
	0.667	1.150	1.190	1.210	1.210	1.210	1.210	1.210	1.210	1.200	1.200	1.190	1.190	1.180	1.160
	0.750	1.470	1.460	1.440	1.420	1.410	1.400	1.390	1.370	1.360	1.340	1.330	1.310	1.290	1.250
	0.800	1.720	1.660	1.620	1.590	1.560	1.540	1.520	1.490	1.470	1.450	1.420	1.390	1.370	1.320
26	0.500	0.710	0.810	0.860	0.890	0.910	0.930	0.940	0.960	0.970	0.980	0.990	1.000	1.010	1.030
	0.600	0.950	1.020	1.050	1.070	1.080	1.090	1.090	1.100	1.100	1.100	1.100	1.100	1.110	1.100
	0.667	1.150	1.190	1.200	1.210	1.210	1.210	1.210	1.200	1.200	1.200	1.190	1.180	1.180	1.160
	0.750	1.460	1.450	1.440	1.420	1.410	1.390	1.380	1.370	1.350	1.340	1.320	1.300	1.280	1.250
	0.800	1.710	1.660	1.620	1.580	1.560	1.530	1.520	1.490	1.470	1.440	1.420	1.390	1.360	1.310

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
27	0.500	0.710	0.810	0.860	0.890	0.910	0.930	0.94	0.96	0.97	0.9a	0.99	1.00	1.01	1.03
	0.600	0.95	1.02	1.05	1.07	1.08	1.08	1.09	1.10	1.10	1.10	1.10	1.10	1.10	1.10
	0.667	1.14	1.19	1.20	1.21	1.21	1.21	1.20	1.20	1.20	1.19	1.19	1.18	1.17	1.16
	0.750	1.46	1.45	1.43	1.42	1.40	1.39	1.38	1.36	1.35	1.33	1.32	1.30	1.28	1.24
	0.800	1.71	1.66	1.61	1.58	1.55	1.53	1.51	1.48	1.46	1.44	1.41	1.3a	1.35	1.30
28	0.500	0.710	0.810	0.860	0.890	0.910	0.930	0.94	0.96	0.97	0.98	0.99	1.00	1.01	1.02
	0.600	0.95	1.02	1.05	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.10	1.10	1.10
	0.667	1.14	1.18	1.20	1.20	1.20	1.20	1.20	1.20	1.20	1.19	1.19	1.18	1.17	1.15
	0.750	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.36	1.34	1.33	1.31	1.29	1.27	1.24
	0.800	1.71	1.65	1.61	1.57	1.55	1.52	1.51	1.48	1.46	1.43	1.41	1.37	1.35	1.30
29	0.500	0.710	0.810	0.860	0.890	0.910	0.930	0.94	0.96	0.97	0.98	0.99	1.00	1.01	1.02
	0.600	0.95	1.02	1.05	1.06	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.10	1.10	1.10
	0.667	1.14	1.18	1.20	1.20	1.20	1.20	1.20	1.20	1.19	1.19	1.18	1.17	1.17	1.15
	0.750	1.45	1.45	1.43	1.41	1.40	1.38	1.37	1.35	1.34	1.32	1.31	1.29	1.27	1.23
	0.800	1.70	1.65	1.60	1.57	1.54	1.52	1.50	1.47	1.45	1.43	1.40	1.37	1.34	1.29
30	0.500	0.71	0.8a	0.860	0.890	0.910	0.930	0.94	0.96	0.97	0.98	0.99	1.00	1.01	1.02
	0.600	0.94	1.01	1.05	1.06	1.07	1.08	1.08	1.09	1.09	1.10	1.10	1.10	1.10	1.09
	0.667	1.14	1.18	1.19	1.20	1.20	1.20	1.20	1.19	1.19	1.19	1.18	1.17	1.16	1.15
	0.750	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.35	1.34	1.32	1.30	1.28	1.26	1.23
	0.800	1.70	1.64	1.60	1.57	1.54	1.52	1.50	1.47	1.45	1.42	1.39	1.36	1.34	1.28
60	0.500	0.700	0.800	0.850	0.880	0.900	0.920	0.93	0.94	0.96	0.97	0.98	0.99	1.00	1.01
	0.600	0.93	1.00	1.03	1.04	1.05	1.06	1.06	1.07	1.07	1.07	1.07	1.07	1.07	1.06
	0.667	1.12	1.16	1.17	1.17	1.17	1.17	1.17	1.16	1.16	1.15	1.14	1.13	1.12	1.10
	0.750	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.30	1.29	1.27	1.25	1.22	1.20	1.15
	0.800	1.65	1.59	1.55	1.51	1.48	1.46	1.44	1.41	1.38	1.35	1.32	1.29	1.25	1.18
80	0.500	0.700	0.800	0.850	0.880	0.900	0.910	0.93	0.94	0.95	0.96	0.97	0.99	1.00	1.01
	0.600	0.930	0.99	1.02	1.04	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.06	1.06	1.05
	0.667	1.11	1.15	1.16	1.17	1.17	1.16	1.16	1.16	1.15	1.14	1.13	1.12	1.11	1.08
	0.750	1.41	1.40	1.38	1.36	1.34	1.32	1.31	1.29	1.27	1.26	1.23	1.21	1.18	1.12
	0.800	1.64	1.58	1.53	1.50	1.47	1.44	1.42	1.39	1.37	1.34	1.31	1.27	1.23	1.16
100	0.500	0.700	0.790	0.840	0.880	0.900	0.910	0.92	0.94	0.95	0.96	0.97	0.98	0.99	1.01
	0.600	0.920	0.99	1.02	1.04	1.05	1.05	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.04
	0.667	1.11	1.15	1.16	1.16	1.16	1.16	1.16	1.15	1.15	1.14	1.13	1.12	1.10	1.07
	0.750	1.41	1.39	1.37	1.35	1.33	1.32	1.30	1.28	1.27	1.25	1.23	1.20	1.17	1.11
	0.800	1.64	1.58	1.53	1.49	1.46	1.43	1.41	1.38	1.36	1.33	1.30	1.26	1.22	1.14
120	0.500	0.700	0.790	0.840	0.880	0.900	0.910	0.92	0.94	0.95	0.96	0.97	0.98	0.99	1.01
	0.600	0.920	0.99	1.02	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.06	1.06	1.05	1.04
	0.667	1.11	1.15	1.16	1.16	1.16	1.16	1.15	1.15	1.14	1.13	1.13	1.11	1.10	1.06
	0.750	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.28	1.26	1.24	1.22	1.19	1.16	1.10
	0.800	1.63	1.57	1.52	1.48	1.45	1.43	1.41	1.37	1.35	1.32	1.29	1.25	1.21	1.12
∞	0.500	0.690	0.790	0.840	0.870	0.890	0.910	0.92	0.93	0.95	0.96	0.97	0.98	0.99	1.00
	0.600	0.920	0.98	1.01	1.03	1.04	1.04	1.04	1.05	1.05	1.05	1.05	1.04	1.04	1.00
	0.667	1.10	1.13	1.14	1.15	1.14	1.14	1.14	1.13	1.13	1.12	1.11	1.09	1.07	1.00
	0.750	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.25	1.24	1.22	1.19	1.16	1.13	1.00
	0.800	1.61	1.55	1.50	1.46	1.43	1.40	1.38	1.34	1.32	1.29	1.25	1.21	1.16	1.00

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
1	0.900	49.5	53.6	55.8	57.2	58.2	59.1	59.7	60.5	61.0	61.5	62.0	62.6	63.0	63.3
	0.950	199.	216.	225.	230.	234.	237.	239.	242.	244.	246.	248.	250.	252.	254.
	0.975	800.	864.	900.	922.	937.	948.	957.	969.	977.	985.	993.			
	0.990														
	0.999														
2	0.900	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.39	9.41	9.43	9.44	9.46	9.47	9.49
	0.950	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5
	0.975	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5
	0.990	99.0	99.2	99.2	99.3	99.3	99.4	100.	100.	100.	100.	100.	100.	100.	99.5
	0.999	999.	999.												
3	0.900	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.23	5.22	5.20	5.18	5.17	5.15	5.13
	0.950	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.79	8.74	8.70	8.66	8.62	8.58	8.53
	0.975	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.4	14.3	14.3	14.2	14.1	14.0	13.9
	0.990	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.2	27.1	26.9	26.7	26.5	26.4	26.1
	0.999	149.	141.	137.	135.	133.	132.	131.	129.	128.	127.	126.	125.	125.	123.
4	0.900	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.92	3.90	3.87	3.84	3.82	3.79	3.76
	0.950	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.91	5.86	5.80	5.75	5.70	5.63
	0.975	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.84	8.75	8.66	8.56	8.46	8.38	8.26
	0.990	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.5	14.4	14.2	14.0	13.8	13.7	13.5
	0.999	61.2	56.2	53.4	51.7	50.5	49.7	49.0	48.0	47.4	46.8	46.1	45.4	44.9	44.1
5	0.900	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.30	3.27	3.24	3.21	3.17	3.15	3.10
	0.950	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.68	4.62	4.56	4.50	4.44	4.36
	0.975	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.62	6.52	6.43	6.33	6.23	6.14	6.02
	0.990	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.1	9.89	9.72	9.55	9.38	9.24	9.02
	0.999	37.1	33.2	31.1	29.8	28.8	28.2	27.6	26.9	26.4	25.9	25.4	24.9	24.4	23.8
6	0.900	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.94	2.90	2.87	2.84	2.80	2.77	2.72
	0.950	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.94	3.87	3.81	3.75	3.67
	0.975	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.46	5.37	5.27	5.17	5.07	4.98	4.85
	0.990	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.87	7.72	7.56	7.40	7.23	7.09	6.88
	0.999	27.0	23.7	21.9	20.8	20.0	19.5	19.0	18.4	18.0	17.6	17.1	16.7	16.3	15.7
7	0.900	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.70	2.67	2.63	2.59	2.56	2.52	2.47
	0.950	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.57	3.51	3.44	3.38	3.32	3.23
	0.975	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.76	4.67	4.57	4.47	4.36	4.28	4.14
	0.990	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.62	6.47	6.31	6.16	5.99	5.86	5.65
	0.999	21.7	18.8	17.2	16.2	15.5	15.0	14.6	14.1	13.7	13.3	12.9	12.5	12.2	11.7
8	0.900	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.54	2.50	2.46	2.42	2.38	2.35	2.29
	0.950	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.28	3.22	3.15	3.08	3.02	2.93
	0.975	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.29	4.20	4.10	4.00	3.89	3.81	3.67
	0.990	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.81	5.67	5.52	5.36	5.20	5.07	4.86
	0.999	18.5	15.8	14.4	13.5	12.9	12.4	12.0	11.5	11.2	10.8	10.5	10.1	9.80	9.33

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
9	0.900	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.42	2.38	2.34	2.30	2.25	2.22	2.16
	0.950	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.14	3.07	3.01	2.94	2.86	2.80	2.71
	0.975	5.71	5.08	4.72	4.48	4.32	4.20	4.10	3.96	3.87	3.77	3.67	3.56	3.47	3.33
	0.990	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.26	5.11	4.96	4.81	4.65	4.52	4.31
	0.999	16.4	13.9	12.6	11.7	11.1	10.7	10.4	9.89	9.57	9.24	8.90	8.55	8.26	7.81
10	0.900	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.32	2.28	2.24	2.20	2.16	2.12	2.06
	0.950	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.98	2.91	2.84	2.77	2.70	2.64	2.54
	0.975	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.72	3.62	3.52	3.42	3.31	3.22	3.08
	0.990	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.85	4.71	4.56	4.41	4.25	4.11	3.91
	0.999	14.9	12.6	11.3	10.5	9.93	9.52	9.20	8.75	8.45	8.13	7.80	7.47	7.19	6.76
11	0.900	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.25	2.21	2.17	2.12	2.08	2.04	1.97
	0.950	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.85	2.79	2.72	2.65	2.57	2.51	2.40
	0.975	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.53	3.43	3.33	3.23	3.12	3.03	2.88
	0.990	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.54	4.40	4.25	4.10	3.94	3.81	3.60
	0.999	13.8	11.6	10.3	9.58	9.05	8.66	8.35	7.92	7.63	7.32	7.01	6.68	6.42	6.00
12	0.900	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.19	2.15	2.10	2.06	2.01	1.97	1.90
	0.950	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.75	2.69	2.62	2.54	2.47	2.40	2.30
	0.975	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.37	3.28	3.18	3.07	2.96	2.87	2.72
	0.990	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.30	4.16	4.01	3.86	3.70	3.57	3.36
	0.999	13.0	10.8	9.63	8.89	8.38	8.00	7.71	7.29	7.00	6.71	6.40	6.09	5.83	5.42
13	0.900	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.14	2.10	2.05	2.01	1.96	1.92	1.85
	0.950	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.67	2.60	2.53	2.46	2.38	2.31	2.21
	0.975	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.25	3.15	3.05	2.95	2.84	2.74	2.60
	0.990	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.10	3.96	3.82	3.66	3.51	3.37	3.17
	0.999	12.3	10.2	9.07	8.35	7.86	7.49	7.21	6.80	6.52	6.23	5.93	5.63	5.37	4.97
14	0.900	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.10	2.05	2.01	1.96	1.91	1.87	1.80
	0.950	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.60	2.53	2.46	2.39	2.31	2.24	2.13
	0.975	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.15	3.05	2.95	2.84	2.73	2.64	2.49
	0.990	6.51	5.56	5.04	4.69	4.46	4.28	4.14	3.94	3.80	3.66	3.51	3.35	3.22	3.00
	0.999	11.8	9.73	8.62	7.92	7.44	7.08	6.80	6.40	6.13	5.85	5.56	5.25	5.00	4.60
15	0.900	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.06	2.02	1.97	1.92	1.87	1.83	1.76
	0.950	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.54	2.48	2.40	2.33	2.25	2.18	2.07
	0.975	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.06	2.96	2.86	2.76	2.64	2.55	2.40
	0.990	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.80	3.67	3.52	3.37	3.21	3.08	2.87
	0.999	11.3	9.34	8.25	7.57	7.09	6.74	6.47	6.08	5.81	5.53	5.25	4.95	4.70	4.31
16	0.900	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.03	1.99	1.94	1.89	1.84	1.79	1.72
	0.950	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.42	2.35	2.28	2.19	2.12	2.01
	0.975	4.69	4.08	3.73	3.50	3.34	3.22	3.12	2.99	2.89	2.79	2.68	2.57	2.47	2.32
	0.990	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.69	3.55	3.41	3.26	3.10	2.97	2.75
	0.999	11.0	9.01	7.94	7.27	6.80	6.46	6.19	5.81	5.55	5.27	4.99	4.70	4.45	4.06
17	0.900	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.00	1.96	1.91	1.86	1.81	1.76	1.69
	0.950	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.45	2.38	2.31	2.23	2.15	2.08	1.96
	0.975	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.92	2.82	2.72	2.62	2.50	2.41	2.25
	0.990	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.59	3.46	3.31	3.16	3.00	2.87	2.65
	0.999	10.7	8.73	7.68	7.02	6.56	6.22	5.96	5.58	5.32	5.05	4.77	4.48	4.24	3.85

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
18	0.900	2.62	2.42	2.29	2.20	2.13	2.08	2.04	1.98	1.93	1.89	1.84	1.78	1.74	1.66
	0.950	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.41	2.34	2.27	2.19	2.11	2.04	1.92
	0.975	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.87	2.77	2.67	2.56	2.44	2.35	2.19
	0.990	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.51	3.37	3.23	3.08	2.92	2.78	2.57
	0.999	10.4	8.49	7.46	6.81	6.35	6.02	5.76	5.39	5.13	4.87	4.59	4.30	4.06	3.67
19	0.900	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.96	1.91	1.86	1.81	1.76	1.71	1.63
	0.950	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.38	2.31	2.23	2.16	2.07	2.00	1.88
	0.975	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.82	2.72	2.62	2.51	2.39	2.30	2.13
	0.990	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.43	3.30	3.15	3.00	2.84	2.71	2.49
	0.999	10.2	8.28	7.27	6.62	6.18	5.85	5.59	5.22	4.97	4.70	4.43	4.14	3.90	3.51
20	0.900	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.94	1.89	1.84	1.79	1.74	1.69	1.61
	0.950	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.35	2.28	2.20	2.12	2.04	1.97	1.84
	0.975	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.77	2.68	2.57	2.46	2.35	2.25	2.09
	0.990	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.37	3.23	3.09	2.94	2.78	2.64	2.42
	0.999	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.08	4.82	4.56	4.29	4.00	3.76	3.38
21	0.900	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.92	1.87	1.83	1.78	1.72	1.67	1.59
	0.950	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.32	2.25	2.18	2.10	2.01	1.94	1.81
	0.975	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.73	2.64	2.53	2.42	2.31	2.21	2.04
	0.990	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.31	3.17	3.03	2.88	2.72	2.58	2.36
	0.999	9.77	7.94	6.95	6.32	5.88	5.56	5.31	4.95	4.70	4.44	4.17	3.88	3.64	3.26
22	0.900	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.90	1.86	1.81	1.76	1.70	1.65	1.57
	0.950	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.30	2.23	2.15	2.07	1.98	1.91	1.78
	0.975	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.70	2.60	2.50	2.39	2.27	2.17	2.00
	0.990	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.26	3.12	2.98	2.83	2.67	2.53	2.31
	0.999	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.83	4.58	4.33	4.06	3.78	3.54	3.15
23	0.900	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.89	1.84	1.80	1.74	1.69	1.64	1.55
	0.950	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.27	2.20	2.13	2.05	1.96	1.88	1.76
	0.975	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.67	2.57	2.47	2.36	2.24	2.14	1.97
	0.990	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.21	3.07	2.93	2.78	2.62	2.48	2.26
	0.999	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.73	4.48	4.23	3.96	3.68	3.44	3.05
24	0.900	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.88	1.83	1.78	1.73	1.67	1.62	1.53
	0.950	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.25	2.18	2.11	2.03	1.94	1.86	1.73
	0.975	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.64	2.54	2.44	2.33	2.21	2.11	1.94
	0.990	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.17	3.03	2.89	2.74	2.58	2.44	2.21
	0.999	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.64	4.39	4.14	3.87	3.59	3.36	2.97
25	0.900	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.87	1.82	1.77	1.72	1.66	1.61	1.52
	0.950	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.24	2.16	2.09	2.01	1.92	1.84	1.71
	0.975	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.61	2.51	2.41	2.30	2.18	2.08	1.91
	0.990	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.13	2.99	2.85	2.70	2.54	2.40	2.17
	0.999	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.56	4.31	4.06	3.79	3.52	3.28	2.89
26	0.900	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.86	1.81	1.76	1.71	1.65	1.59	1.50
	0.950	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.22	2.15	2.07	1.99	1.90	1.82	1.69
	0.975	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.59	2.49	2.39	2.28	2.16	2.05	1.88
	0.990	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.09	2.96	2.81	2.66	2.50	2.36	2.13
	0.999	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.48	4.24	3.99	3.72	3.44	3.21	2.82

$m \backslash n$	α	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
27	0.900	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.85	1.80	1.75	1.70	1.64	1.58	1.49
	0.950	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.20	2.13	2.06	1.97	1.88	1.81	1.67
	0.975	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.57	2.47	2.36	2.25	2.13	2.03	1.85
	0.990	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.06	2.93	2.78	2.63	2.47	2.33	2.10
	0.999	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.41	4.17	3.92	3.66	3.38	3.14	2.75
28	0.900	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.84	1.79	1.74	1.69	1.63	1.57	1.48
	0.950	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.19	2.12	2.04	1.96	1.87	1.79	1.65
	0.975	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.55	2.45	2.34	2.23	2.11	2.01	1.83
	0.990	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.03	2.90	2.75	2.60	2.44	2.30	2.06
	0.999	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.35	4.11	3.86	3.60	3.32	3.09	2.69
29	0.900	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.83	1.78	1.73	1.68	1.62	1.56	1.47
	0.950	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.18	2.10	2.03	1.94	1.85	1.77	1.64
	0.975	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.53	2.43	2.32	2.21	2.09	1.99	1.81
	0.990	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.00	2.87	2.73	2.57	2.41	2.27	2.03
	0.999	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.29	4.05	3.80	3.54	3.27	3.03	2.64
30	0.900	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.82	1.77	1.72	1.67	1.61	1.55	1.46
	0.950	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.16	2.09	2.01	1.93	1.84	1.76	1.62
	0.975	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.51	2.41	2.31	2.20	2.07	1.97	1.79
	0.990	5.39	4.51	4.02	3.70	3.47	3.30	3.17	2.98	2.84	2.70	2.55	2.39	2.25	2.01
	0.999	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.24	4.00	3.75	3.49	3.22	2.98	2.59
60	0.900	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.71	1.66	1.60	1.54	1.48	1.41	1.29
	0.950	3.15	2.76	2.53	2.37	2.25	2.17	2.10	1.99	1.92	1.84	1.75	1.65	1.56	1.39
	0.975	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.27	2.17	2.06	1.94	1.82	1.70	1.48
	0.990	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.63	2.50	2.35	2.20	2.03	1.88	1.60
	0.999	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.54	3.32	3.08	2.83	2.55	2.32	1.89
80	0.900	2.37	2.15	2.02	1.92	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.44	1.38	1.24
	0.950	3.11	2.72	2.49	2.33	2.21	2.13	2.06	1.95	1.88	1.79	1.70	1.60	1.51	1.32
	0.975	3.86	3.28	2.95	2.73	2.57	2.45	2.35	2.21	2.11	2.00	1.88	1.75	1.63	1.40
	0.990	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.55	2.42	2.27	2.12	1.94	1.79	1.49
	0.999	7.54	5.97	5.12	4.58	4.20	3.92	3.70	3.39	3.16	2.93	2.68	2.41	2.16	1.72
100	0.900	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.66	1.61	1.56	1.49	1.42	1.35	1.21
	0.950	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.93	1.85	1.77	1.68	1.57	1.48	1.28
	0.975	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.18	2.08	1.97	1.85	1.71	1.59	1.35
	0.990	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.50	2.37	2.22	2.07	1.89	1.74	1.43
	0.999	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.30	3.07	2.84	2.59	2.32	2.08	1.62
120	0.900	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.65	1.60	1.54	1.48	1.41	1.34	1.19
	0.950	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.91	1.83	1.75	1.66	1.55	1.46	1.25
	0.975	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.16	2.05	1.94	1.82	1.69	1.56	1.31
	0.990	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.47	2.34	2.19	2.03	1.86	1.70	1.38
	0.999	7.32	5.78	4.95	4.42	4.04	3.77	3.55	3.24	3.02	2.78	2.53	2.26	2.02	1.54
∞	0.900	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.60	1.55	1.49	1.42	1.34	1.26	1.00
	0.950	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.83	1.75	1.67	1.57	1.46	1.35	1.00
	0.975	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.05	1.94	1.83	1.71	1.57	1.43	1.00
	0.990	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.32	2.18	2.04	1.88	1.70	1.52	1.00
	0.999	6.91	5.42	4.62	4.10	3.74	3.47	3.27	2.96	2.74	2.51	2.27	1.99	1.73	1.00

Приложение Д. ТАБЛИЦЫ КВАНТИЛЕЙ СТАТИСТИКИ ВИЛКОКСОНА

m	n	0.005	0.010	0.025	0.050	0.100	0.200	$m(n + m + 1)$
4	4	.	.	10	11	13	14	36
4	5	.	10	11	12	14	15	40
4	6	10	11	12	13	15	17	44
4	7	10	11	13	14	16	18	48
4	8	11	12	14	15	17	20	52
4	9	11	13	14	16	19	21	56
4	10	12	13	15	17	20	23	60
4	11	12	14	16	18	21	24	64
4	12	13	15	17	19	22	26	68
5	5	15	16	17	19	20	22	55
5	6	16	17	18	20	22	24	60
5	7	16	18	20	21	23	26	65
5	8	17	19	21	23	25	28	70
5	9	18	20	22	24	27	30	75
5	10	19	21	23	26	28	32	80
5	11	20	22	24	27	30	34	85
5	12	21	23	26	28	32	36	90
6	6	23	24	26	28	30	33	78
6	7	24	25	27	29	32	35	84
6	8	25	27	29	31	34	37	90
6	9	26	28	31	33	36	40	96
6	10	27	29	32	35	38	42	102
6	11	28	30	34	37	40	44	108
6	12	30	32	35	38	42	47	114
7	7	32	34	36	39	41	45	105
7	8	34	35	38	41	44	48	112
7	9	35	37	40	43	46	50	119
7	10	37	39	42	45	49	53	126
7	11	38	40	44	47	51	56	133
7	12	40	42	46	49	54	59	140
8	8	43	45	49	51	55	59	136
8	9	45	47	51	54	58	62	144
8	10	47	49	53	56	60	65	152
8	11	49	51	55	59	63	69	160
8	12	51	53	58	62	66	72	168
9	9	56	59	62	66	70	75	171
9	10	58	61	65	69	73	78	180
9	11	61	63	68	72	76	82	189
9	12	63	66	71	75	80	86	198
10	10	71	74	78	82	87	93	210
10	11	73	77	81	86	91	97	220
10	12	76	79	84	89	94	101	230
11	11	87	91	96	100	106	112	253
11	12	90	94	99	104	110	117	264
12	12	105	109	115	120	127	134	300

Приложение Е.
ТАБЛИЦА ЗНАЧЕНИЙ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ
КОЛМОГОРОВА

$$K(x) = 1 + 2 \sum_{m=1}^{\infty} (-1)^m e^{-2m^2 x^2}$$

x	0	1	2	3	4	5	6	7	8	9
0,2	0,000000	000000	000000	000000	000000	000000	000000	000000	000001	000004
0,3	0,000009	000021	000046	000091	000171	000303	000511	000826	001285	001929
0,4	0,002808	003972	005476	007377	009730	012589	016005	020022	024682	030017
0,5	0,036055	042814	050306	058534	067497	077183	087577	098656	110394	122760
0,6	0,135718	149229	163255	177752	192677	207987	223637	239582	255780	272188
0,7	0,288765	305471	322265	339114	355981	372833	389640	406372	423002	439505
0,8	0,455858	472039	488028	503809	519365	534682	549745	564545	579071	593315
0,9	0,607269	620928	634285	647337	660081	672515	684836	696445	707941	719126
1,0	0,730000	740566	750825	760781	770436	779794	788860	797637	806130	814343
1,1	0,822282	829951	837356	844502	851395	858040	864443	870610	876546	882258
1,2	0,887750	893030	898102	903973	907648	912134	916435	920557	924506	928288
1,3	0,931908	935371	938682	941847	944871	947758	950514	953144	955651	958041
1,4	0,960318	962487	964551	966515	968383	970159	971846	973448	974969	976413
1,5	0,977782	979080	980310	981475	982579	983623	984610	985544	986427	987261
1,6	0,988048	988791	989492	990154	990777	991364	991917	992438	992928	993389
1,7	0,993823	994230	994612	994972	995309	995625	995922	996200	996460	996704
1,8	0,996932	997146	997346	997533	997707	997870	998023	998165	998297	998421
1,9	0,998536	998644	998744	998837	998924	999004	999079	999149	999213	999273
2,0	0,999329	999381	999429	999473	999514	999553	999588	999620	999651	999679
2,1	0,999705	999728	999750	999771	999790	999807	999823	999837	999851	999863
2,2	0,999874	999886	999895	999904	999912	999920	999927	999933	999939	999944
2,3	0,999949	999954	999958	999961	999965	999968	999971	999974	999976	999978
2,4	0,999980	999982	999984	999985	999987	999988	999989	999990	999991	999992

ЛИТЕРАТУРА

1. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. — М.: ЮНИТИ, 1998.
2. Биллингсли П. Сходимость вероятностных мер. — М.: Наука, 1977.
3. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. — М.: Наука, 1983.
4. Боровков А. А. Теория вероятностей. — М.: Наука, 1986.
5. Боровков А. А. Математическая статистика. — М.: Наука, 1984.
6. Бочаров П. П., Печинкин А. В. Теория вероятностей. Математическая статистика. — М.: Гардарика, 1998.
7. Буре В. М., Евсеев Е. А. Основы эконометрики. — СПб.: СПбГУ, 2004.
8. Буре В. М., Кирпичников Б. К. Вероятностные модели продолжительности функционирования сложных систем. — СПб.: СПбГУ, 1993. — 92 с.
9. Буре В. М., Парилина Е. М., Седаков А. А., Шевкопляс Е. В. Прикладная статистика в R, STATISTICA и Excel. Описательная статистика. Оценка параметров. Статистические критерии. — СПб.: СПбГУ, 2011.
10. Ван дер Варден Б. Л. Математическая статистика. — М.: ИЛ, 1960.
11. Валландер С. С. Лекции по статистике и эконометрике. — СПб.: Изд. Европейского университета в СПб, 2005.
12. Емельянов Г. В., Скитович В. П. Задачник по теории вероятностей и математической статистике. — 2-е изд., стер. — СПб.: Лань, 2007. — 336 с.
13. Ермаков С. М. Метод Монте-Карло и смежные вопросы. — М.: Наука, 1975.
14. Ермаков С. М., Жиглявский А. А. Математическая теория оптимального эксперимента. — М.: Наука, 1987.
15. Ермаков С. М., Михайлов Г. А. Курс статистического моделирования. — М.: Наука, 1976.

16. Зубков А. М., Севастьянов Б. А., Чистяков В. П. Сборник задач по теории вероятностей: Учеб. пособие для вузов. — 2-е изд., испр. и доп. — М.: Наука, 1989. — 320 с.
17. Ибрагимов И. А., Хасьминский Р. З. Асимптотическая теория оценивания. — М.: Наука, 1979.
18. Ивченко Г. И., Медведев Ю. И. Математическая статистика. — М.: Высшая школа, 1984. — 248 с.
19. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. — 4-е изд. — М.: Наука, 1976. — 544 с.
20. Крамер Г. Математические методы статистики. — М.: Мир, 1976.
21. Лозв М. Теория вероятностей. — М.: Изд. ИЛ, 1962.
22. Лукач Е. Характеристические функции. — М.: Наука, 1979. — 423 с.
23. Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс. — М.: Дело, 2004. — 576 с.
24. Неве Ж. Математические основы теории вероятностей. — М.: Мир, 1969.
25. Невзоров В. Б. Рекорды. Математическая теория. — М.: ФАЗИС Стохастика, 2000.
26. Никитин Я. Ю. Асимптотическая эффективность непараметрических критериев. — М.: ФИЗМАТЛИТ, 1995.
27. Парасарати К. Введение в теорию вероятностей и теорию меры. — М.: Мир, 1983.
28. Петров В. В. Предельные теоремы для сумм независимых случайных величин. — М.: Наука, 1987.
29. Рао С. Р. Линейные статистические методы и их применение. — М.: Наука, 1968.
30. Севастьянов Б. А. Курс теории вероятностей и математической статистики. — М.: Наука, 1982.
31. Секей Г. Парадоксы в теории вероятностей и математической статистике. — М.: Мир, 1990.
32. Тюрин Ю. Н., Макаров А. А. Статистический анализ опытных данных на компьютере. / Под ред. В. Э. Фигурнова. — М.: ИНФРА-М, 1998.
33. Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1, 2. — М.: Мир, 1984.
34. Холлендер М., Вулф Д. Непараметрические методы статистики. — М.: Финансы и статистика, 1983. — 518 с.

35. Ширяев А. Н. Вероятность. — М.: Наука, 1989.
36. Якушев В. П., Буре В. М. Подходы к обнаружению статистических зависимостей. — СПб.: СПбГУ, 2003.
37. Якушев В. П., Буре В. М., Якушев В. В. Построение и анализ эмпирических зависимостей. — СПб.: СПбГУ, 2005.
38. Якушев В. П., Буре В. М. Статистический анализ опытных данных. Непараметрические критерии. — СПб.: АФИ РАСХН, 2001.
39. Amemiya T. Qualitive Response Models: A Survey // Journal of Economic Literature. — 1981. — Vol. XIX. — P. 1483–1536.
40. Amemiya T. Advanced Econometrics. — Cambridge: Harvard University Press, 1985.
41. Ben-Akiva M., Lerman S. Discrete choice analysis. — The MIT Press, Cambridge Massachusetts, 1985.
42. Berndt E., Hall B., Hall R., Hausman J. Estimation and Inference in Nonlinear Structural Models // Annals of Economic and Social Measurement. — 1974. — Vol. 3. — P. 653–665.
43. Davidson R., MacKinnon J. G. Estimation and Inference in Econometrics. — New York, Oxford University Press, 1993.
44. Efron B. Regression and ANOVA with Zero-One Data: Measures of Residual Variation // Journal of American Statistical Association. — 1978. — Vol. 73. — P. 113–121.
45. Engle R. F. Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics // Handbook of Econometrics. II. — Elsevier, 1983. — P. 796–801.
46. Greene W. H. Econometric Analysis, 5th edition. — New Jearsey: Pearson Education, 2003.
47. Greenwood P. E., Nikulin M. S. A Guide to Chi-Squared Testing. — New York, John Wiley & Sons, Inc., 1996.
48. Kay R., Little S. Assessing the Fit of the Logistic Model: A Case Study of Children with Haemolytic Uraemic Syndrome // Applied Statistics. — 1986. — 35. — P. 16–30.
49. Long J. S. Regression models for categorial and limited dependent variables. — Thousand Oaks: Sage Publ., 1997.
50. MacFadden D. The Measurement of Urban Travel Demand // Journal of Public Economics. — 1978/ — 3. — P. 303–328.
51. Maddala G. S. Introduction to Econometrics. — 2nd ed. — Macmillan, 1992.
52. Rao C. Information and Accuracy Attainable in Estimation of Statistical Parameters // Bulletin of the Calcutta Mathematical Society. — 1945. — 37. — P. 81–91.

ОГЛАВЛЕНИЕ

<i>Предисловие</i>	3
<i>Глава 1.</i> Вероятностное пространство	5
§ 1. Случайные события	5
§ 2. Аксиоматика	7
§ 3. Свойства вероятностей	10
§ 4. Классическое определение вероятности	15
§ 5. Геометрические вероятности. Задача о встрече. Парадокс Бертрана. Задача Бюффона	19
§ 6. Условные вероятности	24
§ 7. Независимость событий	26
§ 8. Формула полной вероятности. Формулы Байеса	28
<i>Глава 2.</i> Схема Бернулли	29
§ 1. Схема Бернулли. Биномиальное и полиномиальное распределения	29
§ 2. Теорема Пуассона	31
§ 3. Локальная и интегральная теоремы Муавра–Лапласа	32
§ 4. Закон больших чисел	37
§ 5. Метод Монте-Карло. Вычисление интегралов	38
§ 6. Моделирование случайных величин	41
<i>Глава 3.</i> Цепи Маркова	46
§ 1. Последовательности зависимых испытаний. Цепи Маркова	46
§ 2. Теорема о предельных вероятностях для цепей Маркова	52
<i>Глава 4.</i> Основные вероятностные пространства	56
§ 1. Полуалгебры. Теоремы о продолжении меры	56
§ 2. Примеры измеримых пространств	59
§ 3. Вероятностное пространство $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$	61
§ 4. Классификация вероятностных мер	65
§ 5. Конечномерное вероятностное пространство $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$	70
§ 6. Вероятностное пространство $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty), P)$	73

<i>Глава 5.</i>	Случайные величины	75
§ 1.	Распределения случайных величин	75
§ 2.	Свойство измеримости	81
§ 3.	Случайные элементы со значениями в конечномерном пространстве \mathbb{R}^k	83
§ 4.	Функции распределения в конечномерных пространствах	85
§ 5.	Независимые случайные величины	88
<i>Глава 6.</i>	Математическое ожидание как интеграл Лебега	92
§ 1.	Определение и свойства математического ожидания	92
§ 2.	Предельный переход под знаком интеграла Лебега	107
§ 3.	Теорема Лебега о мажорируемой сходимости	111
§ 4.	Теорема о замене переменной под знаком интеграла Лебега	112
§ 5.	Формулы для вычисления математического ожидания	117
§ 6.	Моменты случайных величин. Дисперсия	121
§ 7.	Неравенство Чебышева. Другие неравенства	124
§ 8.	Ковариация, корреляция	128
§ 9.	Задача о наилучшем линейном прогнозе	130
<i>Глава 7.</i>	Условные распределения	133
§ 1.	Условные математические ожидания простых случайных величин	133
§ 2.	Условные математические ожидания произвольных случайных величин	139
§ 3.	Многомерное нормальное распределение	144
<i>Глава 8.</i>	Некоторые распределения	149
§ 1.	Дискретные распределения	149
§ 2.	Непрерывные распределения	152
<i>Глава 9.</i>	Виды сходимостей случайных величин	167
§ 1.	Сходимости по вероятности, почти наверное, по распределению, в среднем, в основном	167
§ 2.	Критерий сходимости почти наверное. Теорема Бореля–Кантелли	169
§ 3.	Эквивалентность сходимости в основном и сходимости по распределению. Свойства сходимости в основном	176
<i>Глава 10.</i>	Характеристические функции	182
§ 1.	Свойства характеристических функций	182
§ 2.	Примеры характеристических функций	194
§ 3.	Формулы обращения	195
§ 4.	Характеристические функции для нормального распределения	199
§ 5.	Метод характеристических функций в доказательстве предельных теорем	202

<i>Глава 11.</i>	Предельные теоремы	209
§ 1.	Закон больших чисел для независимых одинаково распределенных случайных величин	209
§ 2.	Центральная предельная теорема для независимых одинаково распределенных случайных величин и случайных векторов	211
§ 3.	Закон больших чисел для независимых произвольно распределенных случайных величин. Схема серий	215
§ 4.	Центральная предельная теорема для независимых произвольно распределенных случайных величин. Схема серий	221
§ 5.	Теорема Линдберга. Теорема Ляпунова	227
§ 6.	Усиленный закон больших чисел для произвольно распределенных независимых случайных величин . . .	228
§ 7.	Усиленный закон больших чисел для одинаково распределенных независимых случайных величин . . .	234
<i>Глава 12.</i>	Выборочное пространство	240
§ 1.	Выборка. Выборочное пространство	240
§ 2.	Эмпирическая вероятностная мера. Гистограмма	242
§ 3.	Теорема Гливенко–Кантелли. Теорема о предельном распределении эмпирических вероятностей	244
§ 4.	Описательная статистика	248
<i>Глава 13.</i>	Статистики. Предельные распределения	251
§ 1.	Статистики первого типа	251
§ 2.	Теоремы непрерывности	254
§ 3.	Предельное распределение статистик первого типа . .	259
§ 4.	Предельное распределение статистики Пирсона	261
§ 5.	Гамма-распределение	264
<i>Глава 14.</i>	Критерии согласия для простых гипотез	270
§ 1.	Критерий согласия Пирсона	270
§ 2.	Критерий согласия Колмогорова	273
<i>Глава 15.</i>	Точечные оценки	278
§ 1.	Свойства точечных оценок	278
§ 2.	Методы построения точечных оценок	285
§ 3.	Достаточные статистики	291
§ 4.	Теорема Неймана–Фишера. Теорема Колмогорова . . .	293
§ 5.	Неравенство Рао–Крамера	298
<i>Глава 16.</i>	Доверительные интервалы	303
§ 1.	Асимптотические доверительные интервалы	303
§ 2.	Распределения статистик для выборок из нормальной генеральной совокупности	307
§ 3.	Распределение Стьюдента	309

§ 4.	Точные доверительные интервалы для нормальной генеральной совокупности	312
<i>Глава 17.</i>	<i>Проверка статистических гипотез</i>	<i>316</i>
§ 1.	Проверка двух простых статистических гипотез	316
§ 2.	Простые гипотезы о параметрах нормального и биномиального распределений	322
§ 3.	Гипотезы о параметрах распределений для сложных альтернатив	328
§ 4.	Распределение Фишера	331
§ 5.	Критерии однородности для двух независимых выборок из нормально распределенной генеральной совокупности	332
§ 6.	Критерий однородности Вилкоксона и Манна–Уитни	335
§ 7.	Непараметрические критерии анализа парных повторных наблюдений	339
§ 8.	Однофакторный дисперсионный анализ	342
§ 9.	Критерий однородности Колмогорова–Смирнова	344
§ 10.	Коэффициенты ранговой корреляции Спирмена и Кенделла	346
§ 11.	Коэффициент корреляции Пирсона	348
§ 12.	Критерий согласия хи-квадрат для сложных гипотез	350
§ 13.	Таблицы сопряженности	352
<i>Глава 18.</i>	<i>Регрессионный анализ</i>	<i>355</i>
§ 1.	Спецификация модели	355
§ 2.	Метод наименьших квадратов	356
§ 3.	Свойство оценок метода наименьших квадратов	360
§ 4.	Построение доверительных интервалов и проверка статистических гипотез	365
<i>Глава 19.</i>	<i>Бинарная регрессия</i>	<i>371</i>
§ 1.	Дискретные данные	371
§ 2.	Модель линейной вероятности	372
§ 3.	Логит и пробит модели бинарного выбора	374
§ 4.	Сравнение значений функций нормального и логистического распределений	376
§ 5.	Оценивание параметров в логит и пробит моделях	378
§ 6.	Численные методы нахождения оценок в логит и пробит моделях	380
§ 7.	Проверка гипотез о значимости параметров логит и пробит моделей бинарного выбора	385
§ 8.	Критерий Вальда	387
§ 9.	Критерий отношения правдоподобия	390
§ 10.	Критерии адекватности моделей бинарной регрессии	391

Приложение А. Таблица значений функции стандартного нормального распределения	394
Приложение Б. Таблица квантилей $\chi^2_{\alpha,n}$ уровня α распределения χ^2 с n степенями свободы	395
Приложение В. Таблица квантилей $t_{\alpha,n}$ уровня α распределения Стьюдента с n степенями свободы	397
Приложение Г. Таблица квантилей $F_{\alpha,n,m}$ уровня α распределения Фишера с n и m степенями свободы	398
Приложение Д. Таблицы квантилей статистики Вилкоксона	406
Приложение Е. Таблица значений функции распределения Колмогорова	407
Литература	408

*Владимир Мансурович БУРЕ,
Елена Михайловна ПАРИЛИНА*

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

У ч е б н и к

Зав. редакцией физико-математической
литературы *О. Ю. Краснокутская*
Верстка *А. Г. Сандомирская*
Выпускающие *Н. В. Черезова, Е. П. Королькова*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lanbook.ru; www.lanbook.com
192029, Санкт-Петербург, Общественный пер., 5.
Тел./факс: (812) 412-29-35, 412-05-97, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

ГДЕ КУПИТЬ

ДЛЯ ОРГАНИЗАЦИЙ:

*Для того, чтобы заказать необходимые Вам книги, достаточно обратиться
в любую из торговых компаний Издательского Дома «ЛАНЬ»:*

по России и зарубежью
«ЛАНЬ-ТРЕЙД». 192029, Санкт-Петербург, ул. Крупской, 13
тел.: (812) 412-85-78, 412-14-45, 412-85-82; тел./факс: (812) 412-54-93
e-mail: trade@lanbook.ru; ICQ: 446-869-967
www.lanpbl.spb.ru/price.htm

в Москве и в Московской области
«ЛАНЬ-ПРЕСС». 109263, Москва, 7-я ул. Текстильщиков, д. 6/19
тел.: (499) 178-65-85; e-mail: lanpress@lanbook.ru

в Краснодаре и в Краснодарском крае
«ЛАНЬ-ЮГ». 350072, Краснодар, ул. Жлобы, д. 1/1
тел.: (861) 274-10-35; e-mail: lankrd98@mail.ru

ДЛЯ РОЗНИЧНЫХ ПОКУПАТЕЛЕЙ:

интернет-магазины:

Издательство «Лань»: <http://www.lanbook.com>
«Сова»: <http://www.symplex.ru>; «Ozon.ru»: <http://www.ozon.ru>
«Библион»: <http://www.biblion.ru>

Подписано в печать 25.03.13.
Бумага офсетная. Гарнитура Школьная. Формат 84×108^{1/32}.
Печать офсетная. Усл. п. л. 21,84. Тираж 1000 экз.

Заказ № .

Отпечатано в полном соответствии
с качеством предоставленных диапозитивов
в ОАО «Издательско-полиграфическое предприятие «Правда Севера».
163002, г. Архангельск, пр. Новгородский, д. 32.
Тел./факс (8182) 64-14-54; www.ippps.ru