# Лекции по теории вероятностей и математической статистике

А.А. Соловьев

draft 1.12.03

# Оглавление

# Глава 1

# Вероятностное пространство

## 1.1. Предмет теории вероятностей

Закономерные события – это события, которые всегда происходят как только создаются определенные условия. Закономерные же явления – это система закономерных событий.

Математика, как и любая другая наука, изучает математические модели закономерных явлений окружающего нас мира.

Случайные же события – это события, которые при одних и тех же условиях происходят или нет. Массовые случайные события – это события, для которых можно создать одни и те же условия, при которых случайное событие может произойти или нет.

Однако и случайные события подчиняются закономерностям, которые называют вероятностными закономерностями.

Пусть при проведении эксперимента $N$ раз событие $A$ осуществляется $N(A)$ раз. Число $N(A)$ называется частотой события, а отношение $N(A)/N$ – относительной частотой события $A$. Оказывается, что при больших $N$ для массовых случайных событий относительная частота $N(A)/N$ обладает свойством устойчивости, которое состоит в следующем: в одних и тех же условиях в нескольких сериях достаточно большего числа $N_1, \dots, N_s$ наблюдений события $A$ как правило имеет место приближенное равенство

$$\frac{N_1(A)}{N_1} \approx \frac{N_2(A)}{N_2} \approx \dots \approx \frac{N_s(A)}{N_s}.$$

Таким образом, относительная частота события $A$ колеблется около числа, которое характеризует данное случайное событие $A$. Это число $\mathbf{P}(A)$ в соответствующей математической моделе называют вероятностью события $A$.

Например, пусть событие $A$ – это выпадение герба при одном бросании монеты. Если монета симметричная и однородная, то $\mathbf{P}(A) = 1/2$.

Устойчивость частот – это объективное свойство массовых случайных явлений. Это свойство лежит в основе теории вероятностей как математической науки.

Теория вероятностей изучает математические модели случайных явлений и устанавливает связи между вероятностями случайных событий в математических моделях, которые позволяют вычислять вероятности сложных событий по вероятностям более простых событий.

Основным в теории вероятностей является понятие случайного события.

Примеры: Выпадение герба не менее $k$ раз при $n$ бросаниях монеты. При подбрасывании игральной кости выпадет четное число очков. Бросание монеты до первого выпадения герба.

Достоверным событием называется событие, которое всегда происходит (будем обозначать достоверное событие, как правило, символом $\Omega$).

Невозможное событие $\emptyset$ – событие, которое никогда не происходит.

Противоположное событие $\bar{A}$ – событие, которое происходит тогда и только тогда, когда не происходит событие $A$.

Суммой или объединением событий $A$ и $B$ называется событие, обозначаемое $A \cup B$, которое происходит тогда и только тогда, когда происходит или $A$, или $B$ (или оба вместе).

Произведением или пересечением событий $A$ и $B$ назовем событие, обозначаемое $A \cap B$ или $AB$, которое происходит тогда и только тогда, когда происходят события $A$ и $B$ вместе.

Разностью $A \backslash B$ событий $A$ и $B$ называется событие, которое происходит тогда и только тогда, когда происходит $A$ и не происходит $B$.

События $A$ и $B$ называются несовместными, если $AB = \emptyset$. Сумму попарно несовместных событий $A$ и $B$ будем обозначать через $A + B$.

Будем писать $A \subseteq B$ и говорить, что событие $A$ влечет за собой событие $B$, если из наступления события $A$ следует наступление события $B$.

Если $A \subseteq B$ и $B \subseteq A$, то события $A$ и $B$ будем называть равносильными и писать $A = B$.

Возможные исходы эксперимента называются элементарными событиями. Общепризнанным считается подход, когда события отождествляются с подмножествами множества всех возможных исходов (элементарных событий). Тогда объединение, пересечение и противоположное событие совпадают с соответствующими теоретико-множественными операциями, а класс всех событий, связанных с экспериментом и включающих интересующее нас событие, должен быть замкнут относительно этих операций. Вероятность же есть числовая характеристика класса событий, свойства которой должны быть аналогичны свойствам относительной частоты осуществления события.

## 1.2. Вероятностное пространство

Начнем с определений.

**Определение 1.** Назовем класс $\mathcal{A}$ подмножеств пространства $\Omega$ *алгеброй множеств*, если
1) $\emptyset \in \mathcal{A}$, $\Omega \in \mathcal{A}$;
2) из $A \in \mathcal{A}$ следует $\bar{A} \in \mathcal{A}$;
3) из $A, B \in \mathcal{A}$ следует $A \cup B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$.

**Определение 2.** Алгебру множеств $\mathcal{A}$ называется $\sigma$-алгеброй, если из $A_n \in \mathcal{A}$, $n = 1, 2, \ldots$, следует

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}, \quad \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}.$$

**Определение 3.** Тройку $(\Omega, \mathcal{U}, \mathbf{P})$, где $\Omega$ – пространство элементарных событий, $\mathcal{U}$ – $\sigma$-алгебра подмножеств $\Omega$, называемых случайными событиями, $\mathbf{P}$ – числовая функция, определенная на $\mathcal{U}$ и называемая вероятностью, будем называть вероятностным пространством, если $\mathbf{P}$ удовлетворяет следующим аксиомам:

1°. $\mathbf{P}(A) \geq 0$ для всех $A \in \mathcal{A}$ (неотрицательность $\mathbf{P}$);
2°. $\mathbf{P}(\Omega) = 1$ (нормированность $\mathbf{P}$);
3°. $\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B)$, если $A \cap B = \emptyset$ (аддитивность $\mathbf{P}$);
4°. Если события $A_n$ в последовательности $\{A_n\}$ попарно несовместны, то

$$\mathbf{P}\Big(\sum_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)\,.$$

Свойства вероятностной функции на $\sigma$-алгебре $\mathcal{U}$:
1) $\mathbf{P}(\emptyset) = 0$;
2) Если $A \subseteq B$, то $\mathbf{P}(A) \leqslant \mathbf{P}(B)$;
3) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$;
4) $\mathbf{P}(\cup_k A_k) \leqslant \sum_k \mathbf{P}(A_k)$;
5) Если $A_k \uparrow A$, т.е. $A_1 \subseteq A_2 \subseteq \ldots$ и $\bigcup_{n=1}^{\infty} A_n = A$, то $\mathbf{P}(A) = \lim \mathbf{P}(A_k)$;
6) Если $A_k \downarrow A$, т.е. $A_1 \supseteq A_2 \supseteq \ldots$ и $\bigcap_{n=1}^{\infty} A_n = A$, то $\mathbf{P}(A) = \lim \mathbf{P}(A_k)$.

**Определение 4.** Если некоторое свойство $Q$ выполняется на вероятностном пространстве $(\Omega, \mathcal{U}, \mathbf{P})$ с вероятностью 1, то говорят, что свойство $Q$ имеет место почти наверное (п.н.).

## 1.3. Дискретное вероятностное пространство

Пусть пространство элементарных событий $\Omega$ – не более чем счетное множество, $\sigma$-алгебра случайных событий $\mathcal{U}$ – система всех подмножеств множества $\Omega$ и вероятностная функция определяется заданием вероятности $\mathbf{P}(\omega)$ каждого исхода $\omega \in \Omega$, удовлетворяя соотношению $\sum_{\omega\in\Omega}\mathbf{P}(\omega) = 1$. Тройка $\big(\Omega, \mathcal{U}, \mathbf{P}\big)$ называется дискретным вероятностным пространством. Вероятность события $A \in \Omega$ задается равенством

$$\mathbf{P}(A) = \sum_{\omega\in A}\mathbf{P}(\omega).$$

Частным случаем дискретного вероятностного пространства является так называемое классическое вероятностное пространство, когда вероятности всех исходов равны друг другу. Классическая модель вероятностного пространства используется тогда, когда элементарные события равновозможны. Обозначим символом $|A|$ число элементов множества $A$. Так как $1 = \sum_{\omega\in\Omega}\mathbf{P}(\omega) = \mathbf{P}(\omega)|\Omega|$, то

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|}\,.$$

**Элементы комбинаторики**

Установленный в конечном множестве порядок называется перестановкой. Множество с заданным порядком называется упорядоченным множеством.

Конечное упорядоченное множество называется размещением. Число размещений $A_n^k$ объема $k$ в множестве из $n$ элементов равно

$$A_n^k = \frac{n!}{(n-m)!}.$$

Конечные подмножества конечного множества называются сочетаниями. Число сочетаний объема $m$ в множестве из $n$ элементов равно

$$C_n^m = \frac{n!}{m!(n-m)!}$$

**Урновая схема**

*Выборка без возвращения.* Пусть имеется урна с $N$ одинаковыми неразличимыми на ощупь шарами, занумерованных числами $1, 2, \ldots, N$. Предположим, что шары с номерами $1, 2, \ldots, M$ белого цвета, остальные – черного. Выборка без возвращения состоит в последовательном извлечении $n$ шаров, не возвращая их обратно. В этом случае за пространство элементарных событий $\Omega$ можно принять множество всех упорядоченных наборов $(\omega_1, \omega_2, \ldots, \omega_n)$ чисел $\omega_k$, $\ 1 \leqslant \omega_k \leqslant N$, не равных друг другу. Общее число исходов в этом случае равно $A_N^n$ – числу размещений из $N$ элементов по $n$. Вычислим вероятность события $B_m$, состоящего в том, что среди извлеченных $n$ шаров имеется ровно $m$ белых. Число благоприятных исходов определяется как произведение числа способов выбора $m$ координат из общего их количества $n$, на которые мы разместим номера белых шаров, $A_M^m$ – числа различных способов разместить белые шары на $m$ мест и $A_{N-M}^{n-m}$ – числа различных способов разместить черные шары на оставшиеся $n-m$ мест. В итоге, получаем

$$\mathbf{P}(B_m) = \frac{C_n^m A_M^m A_{N-M}^{n-m}}{A_N^n}\,.$$

Пользуясь комбинаторными формулами, вероятность $\mathbf{P}(B_m)$ можно представить в виде

$$\mathbf{P}(B_m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\,.$$

Если нас не интересует порядок извлечения шаров, а только состав выборки, то выбрать $m$ белых шаров из $n$ шаров можно $C_M^m C_{N-M}^{n-m}$ способами при общем количестве исходов, равном $C_N^n$. Поэтому и в этом случае вероятность события $B_m$ будет равна

$$\mathbf{P}(B_m) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\,.$$

Таким образом, в приведенном примере выборки без возвращения элементарными событиями можно считать как упорядоченные, так и не упорядоченные выборки.

*Выборки с возвращением.* Пусть имеется та же урна, из которой последовательно извлекается по одному шару, фиксируется номер шара, а сам шар возвращается обратно в урну. В этом случае пространство элементарных событий состоит из векторов $(\omega_1, \omega_2, \ldots, \omega_n)$, координаты которых $1 \leqslant \omega_j \leqslant N$. В этом случае

$$|\Omega| = N^n,$$

а вероятность события $A_m$ равна

$$\mathbf{P}(A_m) = C_n^m \frac{M^m (N-M)^{n-m}}{N^n} = C_n^m \left(\frac{M}{N}\right)^m \left(1 - \frac{M}{N}\right)^{n-m}.$$

Если же нас интересует только состав шаров, а не последовательность, в которой эти шары извлекаются, элементарными событиями будут векторы $(n_1, \ldots, n_N)$, где $n_1 + \cdots + n_N = n$. Здесь $n_k$ число извлеченных шаров с номером $k$. В этом случае вероятность события $A_m$ равна

$$\mathbf{P}(A_m) = \frac{C_{M+m-1}^m C_{N-M+n-m-1}^{n-m}}{C_{N+n-1}^n}\,.$$

Этот результат легко выводится из леммы

**Лемма 1.1.** *Число возможных способов решить уравнение*

$$r_1 + \cdots + r_n = r, \;\; r_k \geqslant 0,$$

*равно* $C_{n+r-1}^r$.

### Схема Бернулли

Пусть $(\Omega, \mathcal{U}, \mathbf{P})$ вероятностное пространство, в котором элементарными событиями являются всевозможные последовательности из нулей и единиц объема $n$. Если $\omega \in \Omega$, положим $\mathbf{P}(\omega) = p^k(1-p)^{n-k}$, если $\omega$ содержит $k$ единиц. Ясно, что $\mathbf{P}(\Omega) = \sum_{k=0}^r C_n^k p^k (1-p)^{n-k} = 1$. Пусть событие $A_k$ состоит из всех исходов, содержащих $k$ единиц. Вероятности каждого такого исхода равна $p^k(1-p)^{n-k}$, поэтому вероятность события $A_k$ равна

$$\mathbf{P}(A_k) = C_n^k p^k (1-p)^{n-k}\,.$$

Набор вероятностей $\left\{C_n^k p^k (1-p)^{n-k}\right\}_{k=0}^n$ называется биномиальным распределением, а вероятностная схема – схемой Бернулли.

### Геометрические вероятности

Пусть $\Omega$ – область евклидова пространства $\mathbf{R}^n$ с конечным $n$-мерным объемом. Событиями назовем измеримые подмножества $\Omega$. За алгебру событий примем $\sigma$-алгебру $\mathcal{B}$ борелевских подмножеств $\Omega$. За вероятность события $A \in \mathcal{B}$ примем

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|},$$

где $|V|$ обозначает $n$-мерный объем множества $V$. Тогда тройка $(\Omega, \mathcal{B}, \mathbf{P})$ является вероятностным пространством.

*Пример 1. Задача Бюффона.* На плоскости, расчерченной параллельными прямыми, находящимися на расстоянии $a$ друг от друга, случайно бросается игла длины $\ell < a$. Найти вероятность $\mathbf{P}(A)$ события $A = \{$игла пересечет какую-нибудь из параллельных прямых$\}$.

Обозначим через $y$ расстояние от середины иглы до ближайшей прямой, $x$ – острый угол между иглой и перпендикуляром к параллельным прямым. Координаты $(x, y)$ определяют положение иглы относительно параллельных прямых и удовлетворяют условиям $0 \leqslant x \leqslant \pi/2, \quad 0 \leqslant y \leqslant \ell/2$. На плоскости $(x, y)$ они образуют прямоугольник $\Omega$. Игла пересечет одну из параллельных прямых, если $x$ и $y$ удовлетворяют неравенству

$$0 \leqslant y \leqslant \frac{\ell}{2} \cos x\,.$$

Искомая вероятность равна

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{\int\limits_0^{\pi/2} \frac{\ell}{2} \cos x dx}{a/2 \cdot \pi/2} = \frac{2\ell}{a\pi}\,.$$

## 1.4. Условные вероятности, независимость

### Условные вероятности

Пусть при $N$ испытаниях события $A$, $B$ и $AB$ произошли с частотами $N_A$, $N_B$ и $N_{AB}$. Если имеет место устойчивость частот

$$\frac{N_A}{N} \approx \mathbf{P}(A), \quad \frac{N_B}{N} \approx \mathbf{P}(B) \text{ и } \frac{N_{AB}}{N} \approx \mathbf{P}(AB)$$

и $\mathbf{P}(B) > 0$, то относительная частота $N_{AB}/N_B$ события $A$ при условии, что произошло событие $B$ также устойчива

$$\frac{N_{AB}}{N_B} = \frac{N_{AB}/N}{N_B/N} \approx \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

**Определение 5.** Пусть $\mathbf{P}(B) > 0$. Условной вероятностью $\mathbf{P}(A|B)$ события $A$ при условии, что произошло событие $B$, назовем отношение

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Условная вероятность $\mathbf{P}(A|B)$ обозначается также $\mathbf{P}_B(A)$.

Если случайное событие $B$ из некоторого вероятностного пространства $(\Omega, \mathcal{U}, \mathbf{P})$ фиксировано, то условная вероятность $\mathbf{P}_B(A)$ как функция случайного события $A$ определяет новое вероятностное пространство $(\Omega, \mathcal{U}, \mathbf{P}_B)$.

**Теорема 1.1** (теорема умножения). *Пусть события $A_1, \dots, A_n$ таковы, что $\mathbf{P}(A_1 \dots A_{n-1}) > 0$. Тогда*

$$\mathbf{P}(A_1 \dots A_n) = \mathbf{P}(A_1)\mathbf{P}_{A_1}(A_2)\mathbf{P}_{A_1 A_2}(A_3) \dots \mathbf{P}_{A_1 \dots A_{n-1}}(A_n)\,. \tag{1}$$

Доказывается методом математической индукции.

### формула полной вероятности

**Определение 6.** Систему событий $A_1, A_2, \dots, A_n$ будем называть конечным разбиением, если они попарно несовместны и

$$A_1 + A_2 + \cdots + A_n = \Omega\,.$$

**Теорема 1.2** ( Формула полной вероятности). *Если* $A_1, \dots A_n$ *– разбиение и все* $\mathbf{P}(A_k) > 0$, *то для любого события* $B$ *имеет место формула*

$$\mathbf{P}(B) = \sum_{k=1}^{n} \mathbf{P}(A_k)\mathbf{P}(B|A_k)\,, \tag{2}$$

*называемая формулой полной вероятности.*

*Доказательство.* Ясно, что

$$B = B\Omega = BA_1 + B_2 + \cdots + BA_n\,,$$

поэтому

$$\mathbf{P}(B) = \sum_{k=1}^{n} \mathbf{P}(BA_k)\,.$$

Применяя к $\mathbf{P}(BA_k)$ теорему умножения получаем (2). □

**Формула Байеса**

**Теорема 1.3.** *Если выполнены условия теоремы о полной вероятности и* $\mathbf{P}(B) > 0$, *то имеет место формула*

$$\mathbf{P}(A_k|B) = \frac{\mathbf{P}(A_k)\mathbf{P}(B|A_k)}{\sum_{i=1}^{n} \mathbf{P}(A_i)\mathbf{P}(B|A_i)}\,.$$

*Доказательство.* По теореме умножения

$$\mathbf{P}(A_kB) = \mathbf{P}(A_k)\mathbf{P}(B|A_k) = \mathbf{P}(B)\mathbf{P}(A_k|B)\,,$$

поэтому

$$\mathbf{P}(A_k|B) = \frac{\mathbf{P}(A_k)\mathbf{P}(B|A_k)}{\mathbf{P}(B)}$$

Осталось применить к знаменателю $\mathbf{P}(B)$ формулу полной вероятности. □

Назовем события $A_k$ гипотезами. Пусть событие $B$ результат некоторого эксперимента. Вероятности $\mathbf{P}(A_k)$ – это *априорные* вероятности гипотез, вычисляемые до проведения опыта. Условные вероятности $\mathbf{P}(A_k|B)$ – это *апостериорные* вероятности, вычисляемые после проведения опыта.

**Независимые события**

Пусть вероятность события $\mathbf{P}(B)$ не равна нулю. Естественно считать, что событие $A$ не зависит от события $B$, если $\mathbf{P}(A|B) = \mathbf{P}(A)$. Если при этом и $\mathbf{P}(A) > 0$, то

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(AB)}{\mathbf{P}(A)} = \frac{\mathbf{P}(B)\mathbf{P}(A|B)}{\mathbf{P}(A)} = \frac{\mathbf{P}(B)\mathbf{P}(A)}{\mathbf{P}(A)} = \mathbf{P}(B)\,,$$

т.е. $B$ не зависит от $A$. Из теоремы умножения тогда следует, что для независимых друг от друга событий $A$ и $B$ имеет место равенство $\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B)$. Это приводит к следующему определению независимости.

**Определение 7.** События $A$ и $B$ называются независимыми, если

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

Если равенство не выполняется, то события будем называть зависимыми.

Обычно при построении вероятностной модели независимость событий не устанавливается, а постулируется на основе причинной независимости реальных прообразов.

**Определение 8.** События $A_1$, $A_2$, $\dots$ $A_n$ называются независимыми, если для любых $1 \leqslant i_1 < i_2 < \dots < i_m \leqslant n$, $2 \leqslant m \leqslant n$, выполняются равенства

$$\mathbf{P}(A_{i_1}A_{i_2}\dots A_{i_m}) = \mathbf{P}(A_{i_1})\mathbf{P}(A_{i_2})\dots\mathbf{P}(A_{i_m}).$$

В противном случае события называются зависимыми.

Из попарной независимости

$$\mathbf{P}(A_kA_m) = \mathbf{P}(A_k)\mathbf{P}(A_m) \ \ k \neq m\,.$$

не следует независимости.

*Пример 2.* Пусть из чисел 2, 3, 5, 30 выбирается одно, причем все исходы равновозможны и имеют вероятность 1/4. Введем через $A_k = \{$выбранное число делится на $k\}$. Легко видеть, что события $A_2$, $A_3$, $A_5$ попарно независимы, но зависимы в совокупности, так как

$$\mathbf{P}(A_2) = \mathbf{P}(A_3) = \mathbf{P}(A_5) = 1/2,$$

$$\mathbf{P}(A_2A_3) = \mathbf{P}(A_2A_5) = \mathbf{P}(A_3A_5) = 1/4 \ \text{ и } \ \mathbf{P}(A_2A_3A_5) = 1/4\,.$$

Из совместной независимости

$$\mathbf{P}(A_1A_2\dots A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2)\dots\mathbf{P}(A_n)\,.$$

не следует независимости.

*Пример 3.* Рассмотрим пространство $\Omega$, состоящее из 36 упорядоченных пар $(i,j)$, где $i,j = 1,\dots,6$. Будем считать, что все пары равновозможны. Тогда для событий $A = \{(i,j) : j = 1,2,5\}$, $B = \{(i,j) : j = 4,5,6\}$ и $C = \{(i,j) : i+j = 9\}$ имеют место соотношения

$$\mathbf{P}(AB) = \frac{1}{6} \neq \frac{1}{4} = \mathbf{P}(A)\mathbf{P}(B),$$

$$\mathbf{P}(AC) = \frac{1}{36} \neq \frac{1}{18} = \mathbf{P}(A)\mathbf{P}(C),$$

$$\mathbf{P}(BC) = \frac{1}{12} \neq \frac{1}{18} = \mathbf{P}(B)\mathbf{P}(C).$$

Но в то же время

$$\mathbf{P}(ABC) = \frac{1}{36} = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C).$$

**Лемма 1.2.** $1^\circ$ *Если события $A$ и $B$ независимы, то события $\bar{A}$ и $B$ также независимы;* $2^\circ$ *Если $A_1$ и $B$, $A_2$ и $B$ независимы и $A_1A_2 = \emptyset$, то $A_1 + A_2$ и $B$ независимы.*

**Определение 9.** Алгебры ($\sigma$-алгебры) событий $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n \subset \mathcal{U}$ в вероятностном пространстве $(\Omega, \mathcal{U}, \mathbf{P})$ называются независимыми, если для любых $A_i \in \mathcal{U}_i$, $i = \overline{1,n}$, выполняется равенство

$$\mathbf{P}(A_1A_2\dots A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2)\dots\mathbf{P}(A_n)\,.$$

## 1.5. Независимые испытания

Под испытанием понимаем некоторый эксперимент, исходами которого служат случайные события. Поэтому моделью испытания является вероятностное пространство. Пусть проведены $n$ испытаний, то есть даны вероятностные пространства

$$(\Omega_1, \mathcal{U}_1, \mathbf{P}_1), \ldots, (\Omega_n, \mathcal{U}_n, \mathbf{P}_n).$$

Если эти испытания причинно независимы, то $\sigma$-алгебры должны быть независимыми. Мы можем об этом говорить только рассматривая $\sigma$-алгебры $\mathcal{U}_k$ как $\sigma$-подалгебры $\sigma$-алгебры некоторого общего вероятностного пространства. Такое вероятностное пространство всегда можно построить. Проделаем это в случае, когда вероятностные пространства $(\Omega_k, \mathcal{U}_k, \mathbf{P}_k)$ конечны.

Пусть $(\Omega_k, \mathcal{U}_k, \mathbf{P}_k)$ – конечные вероятностные пространства, $\Omega_k = \{\omega_k\}$, $\mathcal{U}_k$ – алгебра всех подмножеств из $\Omega_k$, $\mathbf{P}_k(A) = \sum\limits_{\omega_k \in A} \mathbf{P}(\omega_k)$. Положим $\Omega = \Omega_1 \times \cdots \times \Omega_n = \{\omega = (\omega_1, \ldots, \omega_n)\}$, где $\omega_k \in \Omega_k$, $\mathcal{U}$ – алгебра всех подмножеств множества $\Omega$,

$$\mathbf{P}(\omega) = \mathbf{P}_1(\omega_1) \cdot \cdots \cdot \mathbf{P}_n(\omega_n), \;\; \mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega)\,.$$

Введенная вероятность $\mathbf{P}$ называется прямым произведением вероятностей $\mathbf{P}_k$ и обозначается $\mathbf{P} = \mathbf{P}_1 \times \cdots \times \mathbf{P}_n$. Аналогично, $\mathcal{U} = \mathcal{U}_1 \times \cdots \times \mathcal{U}_n$ есть прямое произведение алгебр. В построенном вероятностном пространстве событие вида $A = A_1 \times \cdots \times A_n$ будем называть прямоугольником. Из определения вероятности следует, что вероятность прямоугольника $A$ равна

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega) = \sum_{\omega_1 \in A_1} \mathbf{P}_1(\omega_1) \cdot \cdots \cdot \sum_{\omega_n \in A_n} \mathbf{P}_n(\omega_n) = \prod_{k=1}^{n} \mathbf{P}_k(A_k)\,.$$

Пусть $\mathcal{U}'_k$ обозначает подалгебру всех прямоугольников вида

$$A'_k = \Omega_1 \times \cdots \times \Omega_{k-1} \times A_k \times \Omega_{k+1} \times \cdots \times \Omega_n \in \mathcal{U}'_k\,.$$

Между событиями $A'_i \in U'_i$ и $A_i \in U_i$ устанавливается естественный изоморфизм $A'_i \sim A_i$, сохраняющий вероятности $\mathbf{P}(A'_i) = \mathbf{P}_i(A_i)$. Поэтому события $A_i$ из вероятностного пространства $(\Omega_i, U_i, \mathbf{P}_i)$ будем отождествлять с изоморфными событиями $A'_i$ из $(\Omega, U, \mathbf{P})$. Так как $A = A_1 \times \cdots \times A_n = \bigcap\limits_{k=1}^{n} A'_k$, то из определения вероятности $P$ получаем

$$\mathbf{P}(\bigcap_{k=1}^{n} A'_k) = \prod_{k=1}^{n} \mathbf{P}(A'_k),$$

т.е. алгебры $U'_1, \ldots, U'_n$ независимы.

### Еще раз о схеме Бернулли

Схема Бернулли — частный случай независимых испытаний с 2 исходами в каждом испытании.

Пусть вероятность пространства $(\Omega_i, U_i, \mathbf{P}_i)$ такова, что $\Omega_i = \{0, 1\}$, $A_i = \{\emptyset, \{0\}, \{1\}, \Omega_i\}$, $\mathbf{P}\{0\} = p$, $\mathbf{P}\{1\} = q = 1 - p$, $i = 1, \ldots, n$.

В прямом произведении $(\Omega, U, \mathbf{P})$ элементарный исход $\omega \in \Omega$ имеет вид $\omega = (\omega_1, \ldots, \omega_n)$, $\omega_i = 0, 1$. Вероятность $\mathbf{P}\{\omega\}$ равна $\prod\limits_{k=1}^{n} p^{\omega_i}(1-p)^{1-\omega_i}$. Исход $\{1\}$ будем называть *успехом*, противоположный исход $\{0\}$ - *неуспехом.* Обозначим через $B_k = \{\omega : \omega_1 + \cdots + \omega_n = k\}$ событие, состоящее в том, что при $n$ независимых испытаниях в схеме Бернулли произошло ровно $k$ успехов. Поскольку для любого исхода $\omega \in B_k$ имеем $\mathbf{P}\{\omega\} = p^k (1-q)^{n-k}$, то

$$\mathbf{P}(B_k) = C_n^k p^k q^{n-k}, k = 0, 1, \ldots, n.$$

Напомним, что набор вероятностей $\{C_n^k p^k q^{n-k}\}_{k=0}^n$ называется *биномиальным распределением.*

**Полиномиальная схема**

В каждом независимом испытании возможность появления одного из $r$ попарно несовместных исходов. Пусть вероятностное пространство $(\Omega_i, U_i, \mathbf{P}_i)$, $i = 1, \ldots, n$, таково, что $\Omega_i = \{1, \ldots, r\}$, $p_1, \ldots, p_r$ - вероятности исходов, $p_1 + \cdots + p_r = 1$, и $U_i$ состоит из всех подмножеств из $\Omega_i$.

В прямом произведении $(\Omega, U, \mathbf{P})$ элементарное событие $\omega \in \Omega$ имеет вид $\omega = (\omega_1, \ldots, \omega_n)$, где $\omega_i$ – номер исхода при $i$-ом испытании. Полагаем $\mathbf{P}\{\omega\} = p_{\omega_1} p_{\omega_2} \ldots p_{\omega_n}$.

Введем событие $B_{n_1,\ldots,n_r} = \{$ в $n$ независимых испытаниях произошло ровно $n_k$ $k$-х исходов, $n_1 + n_2 + \cdots + n_r = n\}$. Так как для любого $\omega \in B_{n_1,\ldots,n_r}$

$$p\{\omega\} = \prod_{k=1}^{r} p_k^{n_k}$$

и количество элементов в $B_{n_1,\ldots,n_r}$ равно $\frac{n!}{n_1! \ldots n_r!}$, то $\mathbf{P}(B_{n_1 \ldots n_r}) = \frac{n!}{n_1! \ldots n_r!}\, p_1^{n_1} \ldots p_r^{n_r}$. Набор вероятностей $\{\mathbf{P}(B_{n_1 \ldots n_r})\}$, $n_1 + \cdots + n_r = n$, называется *полиномиальным распределением.* Описанная схема с $r$ исходами называется *полиномиальной.*

# Глава 2

# Случайные величины

## 2.1. Случайные величины (конечная схема)

Начнем с определения случайной величины.

**Определение 10.** Рассмотрим конечное вероятностное пространство $(\Omega, U, \mathbf{P})$. Числовую функцию от элементарного события $\xi = \xi(\omega)$, $\omega \in \Omega$ будем называть случайной величиной.

*Пример 1.* Число успехов при $n$ испытаниях в схеме Бернулли — случайная величина.

*Пример 2.* Пусть в урне имеется $N$ шаров, из них $M$ белых, остальные черные. По схеме выбора без возвращения из урны извлекаются $n$ шаров. Число извлеченных белых шаров в выборке объема $n$ является случайной величиной.

*Пример 3.* С каждым событием $A \in U$ можно связать случайную величину

$$I_A = I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A, \end{cases}$$

называемую индикатором события $A$.

Легко проверить следующие свойства индикаторов:
$I_\emptyset = 0$, $I_\Omega = 1$, $I_{AB} = I_A I_B$, $I_{\bar{A}} = 1 - I_A$.

Если события $A_1, \ldots, A_n$ попарно несовместны, то

$$I_{\sum\limits_{k=1}^{n} A_k} = \sum_{k=1}^{n} I_{A_k}.$$

**Определение 11.** Пусть $x_1 < x_2 < \cdots < x_n$ – значения, принимаемые случайной величиной $\xi$. Через $A_i$ обозначим случайное событие $A_i = \{\omega \in \Omega_i, \xi(\omega) = x_i\}$. Тогда $\alpha_\xi = \{A_i\}_{i=1}^{n}$ есть *разбиение* $\Omega$. Разбиение $\alpha_\xi$ порождает алгебру событий $U_\xi$, элементами которой являются события

$$\{\xi \in B\} = \{\omega : \xi(\omega) \in B\},$$

где $B$ - любое числовое множество. Разбиение $\alpha_\xi$ и алгебру $U_\xi$ будем называть *порожденными случайной величиной* $\xi$. Любое случайное событие $\{\xi \in B\}$ представимо в виде сумм $\sum\limits_i A_i$, где суммирование ведется по тем $i$, для которых $x_i \in B$.

Нетрудно видеть, что случайная величина $\xi$ представляется в виде

$$\xi = \xi(\omega) = \sum_{i=1}^{k} x_i I_{A_i}(\omega).$$

**Определение 12.** *Законом распределения* случайной величины будем называть функцию числового множества $B : B \to \mathbf{P}\{\xi \in B\}$. Закон распределения $\xi$ определяется значениями $x_1, \ldots, x_n$ и вероятностями $\mathbf{P}\{\xi = x_i\} = p_i$, $\quad i = \overline{1, n}$, $\quad \sum\limits_i p_i = 1$. Закон распределения удобно записать в форме таблицы:

| $\xi$ | $x_1$ | $x_2$ | $\dots$ | $x_n$ |
|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | $\dots$ | $p_n$ |

В теории вероятности часто говорят о случайной величине с законом распределения, не указывая вероятностного пространства $(\Omega, U, \mathbf{P})$. Предполагается, что такое вероятностное пространство существует.

Простейшим вероятностным пространством, связанным с законом распределения $\mathbf{P}\{\xi \in B\}$ будет множество элементарных событий $\Omega = \{x_1, \dots, x_n\}$ с элементарными вероятностями $\mathbf{P}(x_i) = p_i$. Случайная величина определяется соотношением $\xi(x_i) = x_i$.

Примеры законов распределения.

1. *Биномиальный закон* для числа успехов $\mu$ в $n$ независимых испытаниях в схеме Бернулли:

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

2. *Равномерное распределение* на $\{1, 2, \dots, N\}$:

$$\mathbf{P}\{\xi = m\} = \frac{1}{N}, \quad m = 1, 2, \dots, N.$$

### Математическое ожидание

**Определение 13.** Математическое ожидание случайной величины $\xi = \xi(\omega)$ на вероятностном пространстве $(\Omega, U, \mathbf{P})$ определяется как сумма

$$M\xi = \sum_{\omega \in \Omega} \xi(\omega)\mathbf{P}(\omega)$$

Перечислим *свойства математического ожидания.* Пусть случайная величина $\xi$ принимает значения $x_1, \dots, x_k$ с вероятностями $p_1 = \mathbf{P}\{\xi = x_1\}, \dots, p_k = \mathbf{P}\{\xi = x_k\}$.

1) $MI_A = \mathbf{P}(A)$, так как

$$MI_A = \sum_{\omega \in \Omega} I_A(\omega)\mathbf{P}(\omega) = \sum_{\omega \in A} \mathbf{P}(\omega) = \mathbf{P}(A).$$

2) аддитивность: $M(\xi + \eta) = M\xi + M\eta$. По индукции получаем, что $M(\xi_1 + \dots + \xi_k) = M\xi_1 + \dots \dots + M\xi_k$.

3) для любой константы $c$

$$M(c\xi) = cM\xi, \quad Mc = c.$$

4) если $\xi \geq \eta$, то $M\xi \geq M\eta$ , в частности, если $\xi \geq 0$ и $M\xi = 0$, то $\mathbf{P}\{\xi = 0\} = 1$.

Действительно, если $\xi \geq 0$ и $M\xi = 0$, то $\forall \omega \in \Omega$ имеем $\xi(\omega)\mathbf{P}(\omega) = 0$, откуда из $\mathbf{P}(\omega) > 0$ следует $\xi(\omega) = 0$.

5) математическое ожидание $\xi$ выражается через закон распределения случайной величины $\xi$ формулой

$$M\xi = \sum_{i=1}^{k} x_i \mathbf{P}\{\xi = x_i\} = \sum_{i=1}^{k} x_i p_i.$$

В самом деле, $\xi = \sum\limits_{i=1}^{n} x_i I_{\{\xi = x_i\}}$. Из свойств 1 – 3 следует

$$M\xi = \sum_{i=1}^{k} x_i M I_{\{\xi = x_i\}} = \sum_{i=1}^{k} x_i \mathbf{P}\{\xi = x_i\}.$$

6) если $g(x)$ - некоторая функция и $\eta = g(\xi)$ случайная величина, то

$$M\eta = Mg(\xi) = \sum_{i=1}^{k} g(x_i)\mathbf{P}\{\xi = x_i\}.$$

Доказывается как и свойство 5, надо только воспользоваться равенством

$$g(\xi) = \sum_{i=1}^{k} g(x_i)I_{\{\xi = x_i\}}.$$

Полагая $g(x) = x^n$, получаем

$$M\xi^n = \sum_{i=1}^{k} x_i^n\mathbf{P}\{\xi = x_i\}.$$

Пусть в ходе эксперимента, осуществляемого $N$ раз, случайная величина $\xi$ значение $x_1$ принимает $N_1$ раз, ..., $x_k - N_k$ раз. Среднее значение, принимаемое случайной величиной имеет вид

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{k} x_iN_i = \sum_{i=1}^{k} x_i\frac{N_i}{N}.$$

Если $N$ велико и эксперимент проводится в одних и тех же условиях, то относительная частота $\frac{N_i}{N}$ близка к вероятностям $p_i = \mathbf{P}\{\xi = x_i\}$. Поэтому среднее значение $\bar{x}$ колеблется около $M\xi$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{k} x_iN_i \approx \sum_{i=1}^{k} x_ip_i = M\xi$$

Таким образом, математическое ожидание статистически истолковывается как среднее значение случайной величины.

**Определение 14.** Математическое ожидание $M\xi^n$ называется $n$-м *моментом случайной величины* $\xi$. *Абсолютным $n$-м моментом* называется $M|\xi|^n$. Обозначим $M\xi$ через $m$. *Центральным моментом $n$-го порядка* называется $M(\xi - m)^n$ и *абсолютным центральным моментом $n$-го порядка* — $M \mid \xi - m \mid^n$.

Центральный момент второго порядка $M(\xi - m)^2$ называется *дисперсией* и обозначается через $D\xi$. Корень квадратный из дисперсии $\sqrt{D\xi}$ называется *среднеквадратическим отклонением.*

Дисперсия обладает сследующими свойствами:
1) $D\xi = M\xi^2 - (M\xi)^2$;
2) $D\xi \geq 0$ и $D\xi = 0$ тогда и только тогда, когда $\mathbf{P}\{\xi = c\} = 1$;
3) для любой константы $c$ имеют место равенства

$$D(c\xi) = c^2D\xi, \quad D(\xi + c) = D\xi\,;$$

4) Неравенство Коши-Буняковского. Для любых двух величин $\xi$ и $\eta$

$$|M\xi\eta| \leqslant \sqrt{M\xi^2\, M\eta^2}\,.$$

**Теорема 2.1** (вероятность суммы событий).

$$\mathbf{P}(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n}\mathbf{P}(A_k) - \sum_{1\leqslant k_1<k_2\leqslant n}\mathbf{P}(A_{k_1}A_{k_2}) + \sum_{1\leqslant k_1<k_2<k_3\leqslant n}\mathbf{P}(A_{k_1})A_{k_1}A_{k_3}) - \cdots +$$

$$(-1)^{n-1}\mathbf{P}(A_1A_2\ldots A_n)$$

*Доказательство.* Для индикатора объединения событий $A_1, \dots, A_n$ имеет место равенство

$$I_{\bigcup\limits_{k=1}^{n} A_k} = 1 - \prod_{k=1}^{n}(1 - I_{A_k}) = \sum_{k=1}^{n} I(A_k) -$$

$$\sum_{1 \leqslant k_1 < k_2 \leqslant n} I_{A_{k_1} A_{k_2}} + \sum_{1 \leqslant k_1 < k_2 < k_3 \leqslant n} I_{A_{k_1} A_{k_1} A_{k_3}} - \dots + (-1)^{n-1} I_{A_1 A_2 \dots A_n}.$$

Осталось вычислить математическое ожидание обеих частей равенства и воспользоваться его аддитивностью. □

### Независимость случайных величин

Если $\xi$ - случайная величина и $\{x_1, \dots, x_n\}$ принимаемые ею значения, то система случайных событий $A_i = \{\xi = x_i\}$, $i = 1, \dots, n$, является разбиением. Порождаемую случайной величиной $\xi$ алгебру будем обозначать через $\alpha_\xi$.

**Определение 15.** Случайные величины $\xi_1, \dots, \xi_m$ независимы, если независимы порожденные ими алгебры $\alpha_{\xi_1}, \dots, \alpha_{\xi_m}$, т.е. для любых $B_1, \dots, B_n \subset \mathbb{R}$ выполняется равенство

$$\mathbf{P}\{\xi_1 \in B_1, \dots, \xi_n \in B_n\} = \prod_{k=1}^{n} \mathbf{P}\{\xi_k \in B_k\}$$

Дадим эквивалентное определение независимости.

**Определение 16.** Пусть $\xi_i$ принимаемые значения $\{x_{i1}, \dots, x_{im_i}\}$. Случайные величины $\xi_1, \dots, \xi_m$ независимы, если для любых $x_{1j_1}, \dots, x_{nj_n}$, $1 \leqslant j_k \leqslant m_k$,

$$\mathbf{P}\{\xi_1 = x_{1j_1}, \dots, \xi_n = x_{nj_n}\} = \prod_{k=1}^{n} \mathbf{P}\{\xi_k = x_{kj_k}\}.$$

**Теорема 2.2.** *Если случайные величины $\xi_1, \dots, \xi_n$ независимы и $g_i(x)$ - числовые функции, то случайные величины $\eta_i = g_i(\xi_i)$ также независимы.*

*Доказательство.* Действительно,

$$\{\zeta_i \in B_i\} = \{g_i(\xi_i) \in B_i\} = \{\xi_i \in g_i^{-1}(B_i)\}.$$

Поэтому $A_{g_i(\xi_i)} \subset A_{\xi_i}$. □

### Мультипликативное свойство математического ожидания

Доказательство следующей теоремы опирается на лемму

**Лемма 2.1.** *Пусть случайные величины $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_m$ независимы. Для любых числовых функций $g_1(x_1, \dots, x_n)$ и $g_2(y_1, \dots, y_2)$ случайные величины $\zeta_1 = g_1(\xi_1, \dots, \xi_n)$ и $\zeta_2 = g_2(\eta_1, \dots, \eta_m)$ независимы.*

**Теорема 2.3.** *Если случайные величины $\xi_1, \dots, \xi_n$ независимы, то*

$$M\big(\xi_1 \dots \xi_n\big) = \prod_{i=1}^{n} M\xi_i$$

*Доказательство.* Пусть $\xi$ и $\eta$ независимы и

$$\xi = \sum_{i=1}^{k} x_i I_{A_i}, \quad \text{где} \quad A_i = \{\xi = x_i\}$$

$$\eta = \sum_{j=1}^{k} y_j I_{B_j}, \quad \text{где} \quad B_j = \{\eta = y_j\}.$$

В силу аддитивности математического ожидания имеем

$$\xi\eta = \sum_{i=1}^{k}\sum_{j=1}^{m} x_i y_j I_{A_i B_j}, \quad M\xi\eta = \sum_{i=1}^{k}\sum_{j=1}^{m} x_i y_j \mathbf{P}(A_i B_j) =$$

$$\sum_{i=1}^{k}\sum_{j=1}^{m} x_i y_j \mathbf{P}(A_i)\mathbf{P}(B_j) = \sum_{i=1}^{k} x_i \mathbf{P}(A_i) \sum_{j=1}^{m} y_j \mathbf{P}(B_j) = M\xi * M\eta$$

Общий случай можно доказать по индукции, если положить $\xi = \xi_1 \ldots \xi_{n-1}$ и $\eta = \xi_n$ и воспользоваться независимостью $\xi$ и $\eta$ (см. лемму 2.1). $\square$

**Теорема 2.4** (аддитивное свойство дисперсии). *Если случайные величины* $\xi_1, \ldots, \xi_n$ *независимы, то*

$$D(\xi_1 + \cdots + \xi_n) = D\xi_1 + \cdots + D\xi_n.$$

*Доказательство.* Пусть $\xi$ и $\eta$ независимы. Имеем

$$D(\xi + \eta) = M((\xi + \eta) + M(\xi + \eta))^2 = M((\xi - M\xi) + (\eta - M\eta))^2 =$$

$$M(\xi - M\xi)^2 + M(\eta - M\eta)^2 + 2M(\xi - M\xi)(\eta - M\eta).$$

Учитывая, что $\xi - M\xi$ и $\eta - M\eta$ независимы, имеем

$$M(\xi - M\xi)(\eta - M\eta) = 0$$

Отсюда следует $D(\xi + \eta) = D\xi + D\eta$. Общий случай получается по индукции. $\square$

*Пример 4. Равномерное распределение* на $\{1, 2, \ldots, N\}$

$$\mathbf{P}\{\xi = m\} = \frac{1}{N}, \quad n = 1, 2, \ldots, N$$

$$M\xi = \frac{N+1}{2} \quad D\xi = \frac{N^2 - 1}{12}, \quad \text{так как}$$

$$\sum_{k=1}^{n} k^2 = \frac{1}{3}(n+1)^3 - \frac{1}{2}(n+1)^2 + \frac{1}{6}(n+1)\,.$$

*Пример 5. Биномиальное распределение* для числа успехов $\mu_n$ в $n$ независимых испытаниях

$$\mathbf{P}\{\mu_n = m\} = C_n^m p^m (1-p)^{n-m}.$$

Представим $\eta_n$ в виде суммы $\xi_1 + \cdots + \xi_n$, где $\xi_k$ индикатор успеха в $k$-ом испытании. Тогда

$$M\eta_n = \sum_{k=1}^{n} M\xi_k = np \quad D\eta_n = \sum_{k=1} nD\xi_k = npg, \quad \text{так как}$$

$$D\xi_k = M\xi_k^2 - (M\xi)^2 = p - p^2 = pq.$$

*Пример 6. Гипергеометрическое распределение.*
Пусть $\xi$ - число белых шаров в выборке объема $n$ из урны с $N$ шарами, из которых $M$ шаров белого цвета. Тогда

$$\mathbf{P}\{\xi = m\} = \frac{C_n^m C_{N-n}^{M-m}}{C_N^M}, \quad m \leqslant min\{n, M\}$$

.

Представим $\xi$ в виде сумм $\xi_1 + \dots + \xi_n$, где $\xi_k$ – индикатор события: *$k$-м извлечен белый шар.* Имеем $M\xi_k = \frac{M}{N}$. Поэтому $M\xi = \sum\limits_{k=1}^{n} M\xi_k = n\frac{M}{N}$. Далее, $\xi^2 = \sum\limits_{k=1}^{n} \xi_k + \sum\limits_{k \neq l} \xi_k \xi_l$. Ясно, что $M\xi_k^2 = M\xi_k$. В предположении $k < l$ имеем:

$$M\xi_k\xi_l = MI_{\{\xi_k=1, \xi_l=1\}} = \mathbf{P}\{\xi_k = 1, \xi_l = 1\} = \mathbf{P}\{\xi_k = 1\}\mathbf{P}\{\xi_l = 1 | \xi_k = 1\} = \frac{M}{N}\frac{(M-1)}{(N-1)}$$

В результате,

$$M\xi^2 = M\xi + \sum_{k \neq l} M\xi_k\xi_l = n\frac{M}{N} + n(n-1)\frac{M(M-1)}{N(N-1)}$$

и

$$D\xi = M\xi^2 - (M\xi)^2 = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

### Совместное распределение случайных величин, случайные векторы

Пусть $\xi, \eta$ случайные величины на $(\Omega, U, \mathbf{P})$, $\xi$ принимает значения $\{x_1, \dots, x_n\}$, $\eta$ — значения $\{y_1, \dots, y_m\}$. Набор чисел $p_{ij} = \mathbf{P}\{\xi = x_i, \eta = y_j\}$, $i = 1, \dots, n$ и $j = 1, \dots, m$, будем называть *совместным распределением* случайных величин $\xi$ и $\eta$. Совместным законом распределения случайных величин $\xi$ и $\eta$ будем называть вероятность

$$\mathbf{P}\{(\xi, \eta) \in B\} = \sum_{x_i, y_j \in B} \mathbf{P}\{\xi = x_i, \eta = y_j\} = \sum_{x_i, y_j \in B} p_{ij},$$

где $B$ любое подмножество в $\mathbb{R}^2$.

Распределение вероятностей определяется таблицей

| | $y_1$ | $y_2$ | $\dots$ |
|---|---|---|---|
| $x_1$ | $p_{11}$ | $p_{12}$ | $\dots$ |
| $x_2$ | $p_{21}$ | $p_{22}$ | $\dots$ |
| $\dots$ | $\dots$ | $\dots$ | $\dots$ |

Из двумерного закона распределения можно получить одномерный закон распределения для случайной величины $\xi$: $\mathbf{P}\{\xi = x_i\} = \sum\limits_j p_{ij}$ и для случайной величины $\eta$: $\mathbf{P}\{y = y_j\} = \sum\limits_i p_{ij}$.

Аналогичным образом для $n$ случайных величин определяется $n$-мерных закон распределения

$$\mathbf{P}_{\xi_1, \dots, \xi_n}(B) = \mathbf{P}\{(\xi_1, \dots, \xi_n) \in B\},$$

где $B$- множество точек $n$-мерного пространства $\mathbb{R}^n$.

Случайные величины $\xi_1, \dots, \xi_n$ будем трактовать как компоненты случайного вектора $\xi = (\xi_1, \dots, \xi_n)$.

Определение независимости для случайных величин

$$\mathbf{P}\{\xi_1 \in B_1, \dots, \xi_n \in B_n\} = \prod_{i=1}^{n} \mathbf{P}\{\xi_i \in B_i\}$$

обобщается на случайные вектора $\xi_i = (\xi_{i1}, \dots, \xi_{im_i})$, где $B_1 \in \mathbb{R}^{m_i}$. Для таких независимых случайных векторов будет справедлива теорема.

**Теорема 2.5.** *Если случайные вектора* $\xi_1, \dots, \xi_n$ *независимы и* $g_k(x)$ *— числовые функции, то случайные величины* $\eta_k = g_k(\xi_k)$ *независимы.*

### Зависимые случайные величины. Коэффициент корреляции

Случайная величина $\xi$ называется нормированной, если $M\xi = 0$ и $D\xi = 1$. Пусть $\xi$ имеет конечное математическое ожидание и дисперсию. Случайная величина

$$\xi_1 = \frac{\xi - M\xi}{\sqrt{D\xi}}$$

с нулевым математическим ожиданием и единичной дисперсией называется нормированной.

Пусть $\xi$ и $\eta$ — случайные величины и $\xi_1$ и $\eta_1$ – их нормировки. Коэффициент корреляции $\rho(\xi, \eta)$ между $\xi$ и $\eta$ определяется соотношением

$$\rho(\xi, \eta) = M(\xi_1 \eta_1) = \frac{M(\xi - M\xi)(\eta - M\eta)}{\sqrt{D\xi \cdot D\eta}}.$$

Числитель носит название ковариации между $\xi$ и $\eta$ и обозначается

$$Cov(\xi, \eta) = M(\xi - M\xi)(\eta - M\eta).$$

Из неравенства Коши-Буняковского $(M\xi_1\eta_1)^2 \leqslant M\xi_1^2 M\eta_1^2$ следует, что всегда $|\rho(\xi, \eta)| \leqslant 1$. Это же свойство можно доказать непосредственно:

$$0 \leqslant D(\xi_1 \pm \eta_1) = M(\xi_1 \pm \eta_1)^2 = 2 \pm 2\rho(\xi, \eta).$$

Если $\xi$ и $\eta$ независимы, то $\rho(\xi, \eta) = 0$. Обратное неверно. Пусть $\xi$ и $\eta$ — независимые случайные величины и $M\xi = M\eta = 0$. Определим $\zeta = \xi\eta$. Тогда $\xi$ и $\zeta$ зависимы, вообще говоря. Но

$$M\xi\zeta = M\xi^2\eta = M\xi^2 M\eta = 0 = M\xi M\eta\,.$$

*Пример 7.* Пусть случайные величины $\xi$ и $\eta$ равномерно распределены на $\{1, 2, 3, \}$ и независимы. Положим $\zeta = \xi\eta$. Имеем

$$\mathbf{P}\{\zeta < 3, \xi < 2\} = \mathbf{P}\{\xi = 1, \eta = 1\} + \mathbf{P}\{\xi = 1, \eta = 2\} = \frac{2}{9};$$

$$\mathbf{P}\{\zeta < 3\} = \frac{1}{3}, \quad \mathbf{P}\{\xi < 2\} = \frac{1}{3}.$$

Поэтому $\mathbf{P}\{\zeta < 3, \xi < 2\} \neq \mathbf{P}\{\zeta < 3\}\mathbf{P}\{\xi < 2\}$, т.е. $\zeta$ и $\xi$ зависимы.

Если $\mathbf{P}(\xi, \eta) = 0$, то случайные величины $\xi$ и $\eta$ называются некоррелированными.

Коэффициент корреляции равен единице тогда и только тогда, когда существуют числа $a \neq 0$ и $b$ такие, что $\mathbf{P}\{\eta = a\xi + b\} = 1$.

*Необходимость.* Пусть $\mathbf{P}\{\eta = a\xi + b\} = 1$, $M\xi = \alpha$ и $D\xi = \beta^2$. Тогда

$$\rho\{\xi, \eta\} = M\left(\frac{\xi - \alpha}{\beta}\,\frac{a\xi + b - a\alpha - b}{|a|\beta}\right) = sign(a).$$

*Достаточность.* a) Пусть $\rho(\xi, \eta) = 1$. Тогда $D(\xi_1 - \eta_1) = 2(1 - \rho(\xi, \eta)) = 0$. Поэтому $\mathbf{P}\{\xi_1 - \eta_1 = const\} = 1$.

b) Пусть $\rho(\xi, \eta) = -1$. Тогда $D(\xi_1 + \eta_1) = 0$ и $\mathbf{P}\{\xi_1 + \eta_1 = const\} = 1$.

### Линейная среднеквадратическая регрессия

Пусть $\xi$ и $\eta$ случайные величины и $g$ вещественная функция на $\mathbb{R}$. Следующее понятие и нижеследующая теорема носят общий характер и не зависят от конечности рассматриваемой схемы.

**Определение 17.** Случайная величина $g(\xi)$ называется наилучшим приближением $\eta$ в смысле наименьших квадратов, если $M(\eta - g(\xi))^2$ принимает наименьшее значение. Случайная величина $g(\xi)$ называется среднеквадратической регрессией случайной величины $\eta$ на случайную величину $\xi$.

**Теорема 2.6.** *Линейная среднеквадратическая регрессия $g(\xi)$ величины $\eta$ на величину $\xi$ имеет вид*

$$g(\xi) = m_2 + \rho\frac{\sigma_2}{\sigma_1}(\xi - m_1).$$

*Здесь $m_1 = M\xi$, $m_2 = M\eta$, $\sigma_1^2 = D\xi$, $\sigma_2^2 = D\eta$, $\rho = \rho(\xi,\eta)$ — коэффициент корреляции.*

*Доказательство.* Найдем $\alpha$ и $\beta$ такие что $\Phi(\alpha,\beta) = M(\eta-\alpha-\beta\xi)^2$ принимает наименьшее возможное значение. Имеем

$$\Phi(\alpha,\beta) = M\eta^2 + \alpha^2 + \beta^2 M\xi^2 - 2\alpha M\eta - 2\beta M\xi\eta + 2\alpha\beta M\xi =$$

$$\sigma_2^2 + m_2^2 + \alpha^2 + \beta^2(\sigma_1^2 + m_1^2) - 2\alpha m_2 - 2\beta(cov(\xi,\eta) + m_1 m_2) + 2\alpha\beta m_1.$$

Исследуем функцию $\Phi(\alpha,\beta)$ на экстремум:

$$\frac{\partial\Phi}{\partial\alpha} = -2(m_2 - \alpha - \beta m_1) = 0.$$

$$\frac{\partial\Phi}{\partial\beta} = 2[\beta(\sigma_1^2 + m_1^2) - (cov(\xi,\eta) + m_1 m_2) + \alpha m_1] = 0.$$

Отсюда находим, что

$$\beta = \frac{cov(\xi,\eta)}{\sigma_1^2} = \rho\frac{\sigma_2}{\sigma_1}, \quad \alpha = m_2 - \beta m_1 = m_2 - \rho\frac{\sigma_2}{\sigma_1}m_1.$$

В это же время

$$\begin{vmatrix} \frac{\partial^2\Phi}{\partial\alpha^2} & \frac{\partial^2\Phi}{\partial\alpha\partial\beta} \\ \frac{\partial^2\Phi}{\partial\alpha\partial\beta} & \frac{\partial^2\Phi}{\partial\beta^2} \end{vmatrix} = \begin{vmatrix} 2 & 2m_1 \\ 2m_1 & 2(\sigma_1^2 + m_1^2) \end{vmatrix} = 4\sigma_1^2 > 0.$$

Таким образом, при найденных $\alpha$ и $\beta$ функция $\Phi(\alpha,\beta)$ принимает наименьшее значение. Значит,

$$g(\xi) = m_2 + \rho\frac{\sigma_2}{\sigma_1}(\xi - m_1)$$

является линейной регрессией $\eta$ на $\xi$. $\square$

Коэффициент $\beta = \rho\frac{\sigma_2}{\sigma_1}$ называется *коэффициентом регрессии* $\eta$ на $\xi$. Прямая

$$\frac{y - m_2}{\sigma_2} = \rho\frac{x - m_1}{\sigma_1}$$

называется *прямой регрессии* $\eta$ на $\xi$. Нетрудно найти наименьшее значение $M(\eta - \alpha - \beta\xi)^2$;

$$\min_{\alpha,\beta} M(\eta - \alpha - \beta\xi)^2 = \sigma_2^2(1 - \rho^2).$$

Эта величина называется *остаточной дисперсией* величины $\eta$ относительно $\xi$. Разность $\eta_1 = \eta - g(\xi)$ называется остатком величины $\eta$ относительно величины $\xi$. Случайные величины $\eta_1$ и $\xi$ являются некоррелированными. В самом деле,

$$M(\xi - m_1)(\eta - m_2 - \rho\frac{\sigma_2}{\sigma_1}(\xi - m_1)) =$$

$$M(\xi - m_1)(\eta - m_2) - \rho\frac{\sigma_2}{\sigma_1}M(\xi - m_1)^2 = cov(\xi,\eta) - \rho\sigma_1\sigma_2 = 0.$$

Аналогично можно рассмотреть линейную регрессию величины $\xi$ на величину $\eta$:

$$L(\eta) = m_1 + \rho\frac{\sigma_1}{\sigma_2}(\eta - m_2).$$

Коэффициент регрессии $\xi$ на $\eta$ равен $\rho\frac{\sigma_1}{\sigma_2}$, и прямая линейной регрессии $\xi$ на $\eta$ определяется соотношением

$$\frac{y - m_2}{\sigma_2} = \frac{1}{\rho}\frac{x - m_1}{\sigma_1}$$

Остаточная дисперсия величины $\xi$ относительно $\eta$ равна $\sigma_1^2(1-\rho^2)$.

Обе прямые регрессии проходят через точку $(m_1, m_2)$ и совпадают только при $\rho = \pm 1$.

### Неравенство Чебышева

**Теорема 2.7.** *Для любого $x > 0$ имеет место неравенства*

$$\mathbf{P}\{|\xi| \geqslant x\} \leqslant \frac{M|\xi|}{x}$$

$$\mathbf{P}\{|\xi - M\xi| \geqslant x\} \leqslant \frac{D\xi}{x^2}.$$

*Доказательство.* Воспользуемся неравенством

$$|\xi| = |\xi| I_{\{|\xi| \geqslant x\}} + |\xi| I_{\{|\xi| < x\}} \geqslant |\xi| I_{\{|\xi| \geqslant x\}} \geqslant x I_{\{|\xi| \geqslant x\}}.$$

Вычисляя математическое ожидание от обеих частей этого неравенства, получаем

$$M|\xi| \geqslant x M I_{\{|\xi| \geqslant x\}} = x\mathbf{P}\{|\xi| \geqslant x\}.$$

Другое неравенство получается из первого, если его применить к случайной величине $\eta = (\xi - M\xi)^2$, помня, что $M\eta = D\xi$. □

Рассмотрим последовательность случайных величин $\xi_1, \xi_2, \ldots$. Будем считать, что $\xi_1, \xi_2, \ldots$ определены на вероятностном пространстве $(\Omega_n, U_n, \mathbf{P}_n)$, причем при каждом $k < n$ случайные величины $\xi_k$ имеют распределение вероятностей, не зависящее от $n$.

*Пример 8.* $\xi_k$ - индикатор успеха $k$-го испытания при испытаниях в схеме Бернулли, $k \leqslant n$.

На самом деле, любую последовательность независимых случайных величин можно определить на одном бесконечном вероятностном пространстве.

Следующая теорема имеет общих характер.

**Теорема 2.8** (Чебышев). *Если $\xi_1, \xi_2, \ldots$ независимы и существует такая константа $c > 0$, что $D\xi_n \leqslant c$, $n = 1, 2, \ldots$, при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left\{\left|\frac{\xi_1 + \ldots \xi_n}{n} - \frac{M\xi_1 + \cdots + M\xi_n}{n}\right| < \varepsilon\right\} = 1.$$

*Доказательство.* Обозначим через $\zeta_n = \xi_1 + \cdots + \xi_n$ и применим к $\frac{\zeta_n}{n}$ второе неравенство Чебышева

$$1 \geqslant \mathbf{P}\left\{\left|\frac{\xi_n}{n} - \frac{M\xi_n}{n}\right| < x\right\} \geqslant 1 - \frac{D\xi_n}{x^2 n^2} \geqslant 1 - \frac{c}{x^2 n},$$

так как $D\xi_n = \sum\limits_{k=1}^{n} D\xi_k \leqslant nc$. В пределе получим нужное равенство. □

Следующая теорема является следствием предыдущей.

**Теорема 2.9** (закон больших чисел в форме Чебышева). *Если* $\xi_1, \xi_2, \dots$ *независимы и одинаково распределены,* $M\xi_n = a$, $D\xi_n = \sigma^2 < \infty$, *то при любом* $x > 0$

$$\lim_{n\to\infty} \mathbf{P}\Big\{\Big|\frac{\xi_1 + \dots + \xi_n}{n} - a\Big| < x\Big\} = 1\,.$$

В частности, для схемы Бернулли верна

**Теорема 2.10** (Бернулли). *Пусть* $\mu_n$ *- число успехов при* $n$ *испытаниях в схеме Бернулли с вероятностью успеха* $p \in (0,1)$ *в каждом испытании. Тогда при любом* $x > 0$

$$\lim_{n\to\infty} \mathbf{P}\Big\{\Big|\frac{\mu_n}{n} - p\Big| \leqslant x\Big\} = 1.$$

*Доказательство.* Представим $\mu_n$ в виде суммы независимых слагаемых $\xi_1 + \dots + \xi_n$, где $\xi_i$ – индикатор успеха при $i$-ом испытании. Поскольку $M\xi_i = p$ и $D\xi_i = p(1-p)$, то к $\mu_n$ применимо следствие. □

### Условное математического ожидания

Рассмотрим случайные величины $\xi$ и $\eta$, принимающие значения $\{x_1, \dots, x_m\}$ и $\{y_1, \dots, y_n\}$. Условным законом распределения $\eta$ при заданном значении $\xi = x_k$ назовем набор условных вероятностей

$$\mathbf{P}\{\eta = y_l | \xi = x_k\} = \frac{\mathbf{P}\{\eta = y_l, \xi = x_k\}}{\mathbf{P}\{\xi = x_k\}},\ l = 1, \dots, n.$$

Условным математическим ожиданием $\eta$ при заданном $\xi = x_k$ будет тогда сумма

$$M(\eta|\xi = x_k) = \sum_{l=1}^{n} y_l \mathbf{P}\{\eta = y_l | \xi = x_k\} = \frac{\sum\limits_{l=1}^{n} y_l \mathbf{P}\{\eta = y_l, \xi = x_k\}}{\mathbf{P}\{\xi = x_k\}}.$$

Можем считать $M\{\eta|\xi = x_k\}$ значениями случайной величины $M(\eta|\xi)$, являющейся функцией от $\xi$. Случайную величину $M(\eta|\xi)$ будем называть условным математическим ожиданием $\eta$ при заданной случайной величине $\xi$.

Математическое ожидание случайной величины $M(\eta|\xi)$ тогда равно

$$M[M(\eta|\xi)] = \sum_{k=1}^{m} \mathbf{P}\{\xi = x_k\} M\{\eta|\xi = x_k\}.$$

Имеет место следующее утверждение.

**Теорема 2.11.** *Верно равенство:*

$$M[M(\eta|\xi)] = M\eta.$$

*Доказательство.* Имеем

$$M[M(\eta|\xi)] = \sum_{k=1}^{m} \mathbf{P}\{\xi = x_k\} M\{\eta|\xi = x_k\} =$$

$$\sum_{k=1}^{m} \sum_{l=1}^{n} y_l \mathbf{P}\{\eta = y_l, \xi = x_k\} = \sum_{l=1}^{n} y_l \mathbf{P}\{\eta = y_l\} = M\eta. \quad \square$$

В результате приходим к формуле

$$M\eta = \sum_{k=1}^{m} \mathbf{P}\{\xi = x_k\} M\{\eta|\xi = x_k\},$$

частным случаем которой является формула полной вероятности.

### Локальная предельная теорема Муавра-Лапласа

**Теорема 2.12.** *Пусть $\mu$ - биномиально распределенная случайная величина с вероятностью успеха $p$, $0 < p < 1$. Если $\sigma = \sqrt{npq} \to 0$, то для любого $c > 0$ равномерно по всем $x$, удовлетворяющим $|x| \leqslant c$, вида $x = \frac{m-np}{\sqrt{npq}}$, где $m$ - целое неотрицательное, имеет место равенство*

$$\mathbf{P}\{\mu = m\} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2}}(1 + O(\frac{1}{\sigma}))$$

*Доказательство.* Пусть $m = np + x\sigma$. Имеем

$$\mathbf{P}\{\mu = m\} = \mathbf{P}\{\frac{\mu - np}{\sigma} = x\} = \frac{n!}{m!(n-m)!} p^m q^{n-m}, \quad q = 1 - p.$$

Воспользуемся формулой Стирлинга

$$n! = \sqrt{2\pi n}\; n^n e^{-n} e^{\theta_n}, \quad \text{где} \quad |\theta_n| < \frac{1}{12n}.$$

Ясно, что $n$, $m = np + x\sqrt{npq} = np(1 + \frac{xq}{\sigma})$ и $n - m = np - x\sqrt{npq} = np(1 - \frac{xq}{\sigma})$ стремится к $+\infty$ при $n \to +\infty$. Поэтому

$$\log \mathbf{P}\{\mu = m\} = \log n! - \log m! - \log(n-m)! + m \log p + (n-m)\log q =$$

$$n \log n - m \log m - (n-m)\log(n-m) + m\log p + (n-m)\log q + \frac{1}{2}\log\frac{n}{2\pi m(n-m)} + \theta_n + \theta_m - \theta_{n-m}.$$

Так как $\log \frac{n}{m(n-m)} = \log\frac{1}{npq} - \log(1 + \frac{xq}{\sigma}) - \log(1 - \frac{xq}{\sigma}) = 2\log\frac{1}{\sigma} + O(\frac{1}{\sigma})$, находим, что

$$\log \mathbf{P}\{\mu = m\} = -m \log\frac{m}{np} - (n-m)\log\frac{n-m}{nq} + \log\frac{1}{\sigma\sqrt{2\pi}} + O(\frac{1}{\sigma}) =$$

$$-(np + x\sigma)\log(1 + \frac{xq}{\sigma}) - (nq - x\sigma)\log(1 - \frac{xp}{\sigma}) + \log\frac{1}{\sigma\sqrt{2\pi}} + O(\frac{1}{\sigma}) =$$

$$\log\frac{1}{\sigma\sqrt{2\pi}} - (np + x\sigma)\Big(\frac{xq}{\sigma} - \frac{x^2q^2}{2\sigma^2} + O\Big(\frac{1}{\sigma^3}\Big)\Big) - (nq - x\sigma)\Big(-\frac{xq}{\sigma} - \frac{x^2q^2}{2\sigma^2} + O\Big(\frac{1}{\sigma^3}\Big)\Big) + O\Big(\frac{1}{\sigma}\Big) =$$

$$\log\frac{1}{\sigma\sqrt{2\pi}} - \frac{x^2}{2} + O(\frac{1}{\sigma}). \;\square$$

Аппроксимация в локальной теореме Муавра-Лапласа удовлетворительная, если

$$\frac{m - np}{\sqrt{npq}} = o((npq)^{\frac{1}{6}})$$

и тем хуже, чем больше $|\frac{1}{2} - p|$.

### Интегральная предельная теорема Муавра-Лапласа

**Теорема 2.13.** *При $\sigma = \sqrt{npq} \to \infty$ равномерно по $-\infty \leqslant a < b \leqslant +\infty$.*

$$\mathbf{P}\{a \leqslant \frac{\mu - np}{\sqrt{npq}} \leqslant b\} - \frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}} dx \to 0.$$

*Здесь $\mu$ - число успехов в схеме Бернулли при $n$ независимых испытаниях с вероятностью успеха $p$ в каждом испытании, $q = 1 - p$.*

*Доказательство.* Выберем положительное $c > 0$ и пусть $|a| \leqslant c$ и $|b| \leqslant c$.

Через $m_1$ и $m_2$ обозначим соответственно $]np+a\sqrt{npq}[$ и $[np+b\sqrt{npq}]$, где $]x[$ - такое наименьшее целое, что $x \leqslant ]x[$, а $[x]$ такое наибольшее целое, что $[x] \leqslant x$.

Положим $m = np + x_m\sqrt{npq}$, тогда $\Delta x_m = x_{m+1} - x_m = \frac{1}{\sqrt{npq}}$. Согласно локальной предельной теореме

$$\mathbf{P}\{a \leqslant \frac{\mu - np}{\sqrt{npq}} \leqslant b\} = \sum_{m=m_1}^{m_2} \mathbf{P}\{\mu = m\} = \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x_m^2}\ \Delta x_m(1 + o(\frac{1}{\sigma})).$$

Справа в равенстве стоит интегральная сумма, сходящаяся равномерно по $a$ и $b$ при $\sigma \to \infty$ к интегралу $\frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}dx$, если $|a| \leqslant c$ и $|b| \leqslant c$.

Так как

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 1,$$

верно равенство

$$\frac{1}{\sqrt{2\pi}}\int\limits_{-c}^{c} e^{-\frac{x^2}{2}}dx = 1 - \frac{1}{\sqrt{2\pi}}\int\limits_{|x|>c} e^{-\frac{x^2}{2}}dx.$$

В общем случае пусть $\xi_n = \frac{\mu - np}{\sqrt{npq}}$. Учитывая также, что

$$\mathbf{P}\{|\xi_n| > c\} = 1 - \mathbf{P}\{|\xi_n| \leqslant c\}.$$

Поэтому

$$\Big|\mathbf{P}\{|\xi_n| > c\} - \frac{1}{\sqrt{2\pi}}\int\limits_{|x|>c} e^{-\frac{x^2}{2}}dx\Big| = \Big|\mathbf{P}\{|\xi_n| \leqslant c\} - \frac{1}{\sqrt{2\pi}}\int\limits_{-c}^{c} e^{-\frac{x^2}{2}}dx\Big|.$$

Выберем $\epsilon > 0$. Найдем $c > 0$ такое, что

$$\frac{1}{\sqrt{2\pi}}\int\limits_{|x|<c} e^{-\frac{x^2}{2}}dx < \frac{\epsilon}{8}$$

и зафиксируем его. По только что доказанному найдется $n_1$, что для всех $n \geqslant n_1$

$$\Big|\mathbf{P}\{|\xi_n| \leqslant c\} - \frac{1}{\sqrt{2\pi}}\int\limits_{-c}^{c} e^{-\frac{x^2}{2}}dx\Big| < \frac{\epsilon}{8}.$$

Поэтому при тех же $n \geqslant n_1$ имеем

$$\mathbf{P}\{|\xi_n| > c\} \leqslant \frac{\epsilon}{4}.$$

Выберем теперь любой интервал $[a, b]$ и обозначим $[A, B] = [a, b] \cap [-c, c]$. Согласно доказанному найдется $n_2$ такое, что для всех $n \geqslant n_2$ имеет место неравенство

$$\Big|\mathbf{P}\{\xi_n \in [A, B]\} - \frac{1}{\sqrt{2\pi}}\int\limits_{A}^{B} e^{-\frac{x^2}{2}}dx\Big| < \frac{\epsilon}{2}.$$

Из неравенства

$$\Big|\mathbf{P}\{\xi_n \in [a, b]\} - \frac{1}{\sqrt{2\pi}}\int\limits_{a}^{b} e^{-\frac{x^2}{2}}dx\Big| \leqslant \mathbf{P}\{|\xi_n| > c\} + \frac{1}{\sqrt{2\pi}}\int\limits_{|x|>c} e^{-\frac{x^2}{2}}dx$$

$$+\left|\mathbf{P}\{\xi_n \in [A,B]\} - \frac{1}{\sqrt{2\pi}}\int\limits_A^B e^{-\frac{x^2}{2}}dx\right|$$

получаем, что при $n \geqslant n_0 = max(n_1, n_2)$

$$\left|\mathbf{P}\{\xi_n \in [a,b]\} - \frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}dx\right| \leqslant \epsilon$$

равномерно по всем $a \leqslant b$. $\square$

Неравенство Берри-Эссена

$$\sup_x |F_n(x) - \Phi(x)| \leqslant \frac{p^2+q^2}{\sqrt{npq}},$$

где $F_n(x) = \mathbf{P}\{\frac{\mu-np}{\sqrt{npq}} \leqslant x\}$ и $\Phi(x) = \int\limits_{-\infty}^{\infty} e^{-\frac{u^2}{2}}du$, позволяет оценить точность приближения в интегральной теореме Муавра-Лапласа.

Без доказательства сформулируем аналог теоремы Муавра - Лапласа для полиномиальной схемы.

Пусть $\mu_i$ – число появлений $i$ – го события в $n$ испытаниях, $i = 1, \ldots, k$, с вероятностью успеха $p_i$, $\mu_1 + \cdots + \mu_k = n$. Заметим, что для

$$x_i = \frac{\mu_i - np_i}{\sqrt{np_iq_i}}$$

выполняется соотношение

$$\sum_{i-1}^{k} x_i\sqrt{np_iq_i} = 0\,.$$

**Теорема 2.14.** *Если в полиномиальной схеме вероятность исходов $p_i$, $0 < p_i < 1$, то какова бы ни была измеримая по Жордану область $G$ на гиперплоскости $\sum_{i=1}^k x_i\sqrt{np_iq_i} = 0$ равномерно относительно $G$ при $n \to \infty$ имеет место соотношение*

$$\mathbf{P}(G) \to \frac{\sqrt{q_1 \cdot \ldots \cdot q_n}}{\sqrt{(2\pi)^{k-1}\sum_{i=1}^k p_iq_i}}\int\limits_G \exp\Big\{-\frac{1}{2}\sum_{i=1}^k q_ix_i^2\Big\}dx\,,$$

*где $dx$ – элемент объема.*

### Предельная теорема Пуассона

Пусть $\mu$ – биномиально распределенная случайная величина,

$$\mathbf{P}\{\mu = m\} = c_n^m p^m q^{n-m}, \quad q = 1-p.$$

**Теорема 2.15** (Пуассона). *Если $n \to \infty$ и $p \to 0$ так, что $np \to a$, то для любого фиксированного $m = 0, 1, \ldots$ имеет место соотношение*

$$\mathbf{P}\{\mu = m\} = C_n^m p^m q^{n-m} \to \frac{a^m}{m!}\,e^{-a}\,.$$

*Доказательство.* Действительно,

$$\mathbf{P}\{\mu = m\} = \frac{(np)^m}{m!}\,(1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{m-1}{n})(1-p)^{n-m}$$

При $n \to \infty$ и $np \to a$ предел $(1-p)^n$ равен $e^{-a}$. Поэтому

$$\mathbf{P}\{\mu = m\} \to \frac{a^m}{m!}\,e^{-a}. \quad \Box$$

Можно показать, что имеет место следующая оценка

$$|\mathbf{P}\{\mu = m\} - \frac{a^m}{m!}\,e^{-a}| \leqslant np^2 = \frac{a^2}{n},$$

где $a = np$.

**Применение предельных теорем**

Приближение, задаваемое теоремой Пуассона, называется *пуассоновским*. Приближение с помощью теорем Муавра-Лапласа называется *нормальным*. Для распределения Пуассона

$$\mathbf{P}\{\mu = m\} = \frac{a^m}{m!}e^{-a}$$

и интеграла

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{u^2}{2}}du,$$

называемого интегралом Лапласа, имеются таблицы.

Пусть, например, нам нужно вычислить вероятность $\mathbf{P}\{m_1 \leqslant \mu \leqslant m_2\}$ в схеме Бернулли в $n$ независимых испытаниях с вероятностью успеха $p$.

Вычислим $x_{m_1} = \frac{m_1-np}{\sqrt{npq}}$, $x_{m_2} = \frac{m_2-np}{\sqrt{npq}}$ и положим

$$\mathbf{P}\{m_1 \leqslant \mu \leqslant m_2\} \approx \frac{1}{\sqrt{2\pi}}\int\limits_{x_{m_1}}^{x_{m_2}} e^{-\frac{x^2}{2}}dx = \Phi_0(x_{m_2}) - \Phi(x_{m_1}).$$

При этом мы допускаем погрешность, которую можно оценить по формуле Берри-Эссена. Вместе с тем эту погрешность можно уменьшить, если в правой части приближенного равенства немного изменить пределы интегрирования, полагая

$$\mathbf{P}\{m_1 \leqslant \mu \leqslant m_2\} \approx \Phi_0(x_{m_2+\frac{1}{2}}) - \Phi_0(x_{m_1-\frac{1}{2}}),$$

где $x_{m_2+\frac{1}{2}} = \frac{m_2+1/2-np}{\sqrt{npq}}$ и $x_{m_1-\frac{1}{2}} = \frac{m_1-1/2-np}{\sqrt{npq}}$.

Предельная теорема Муавра–Лапласа позволяет дать приближенный ответ на следующие вопросы:

Пусть $\mu_n$ – число успехов при $n$ испытаниях в схеме Бернулли с вероятностью успеха $p$, $0 < p < 1$, в каждом испытании.

I. Чему равна вероятность того, что частота наступления события отклоняется от вероятности не более чем на $\alpha$?

$$\mathbf{P}\Big\{\Big|\frac{\mu_n}{n} - p\Big| \le \alpha\Big\} = \mathbf{P}\Big\{-\alpha\sqrt{\frac{n}{pq}} \le \frac{\mu_n - np}{\sqrt{npq}} \le \alpha\sqrt{\frac{n}{pq}}\Big\} \approx$$

$$\approx \frac{1}{\sqrt{2\pi}}\int\limits_{-\alpha\sqrt{\frac{n}{pq}}}^{\alpha\sqrt{\frac{n}{pq}}} \exp\Big\{-\frac{x^2}{2}\Big\}dx = 2\Phi_0\Big(\alpha\sqrt{\frac{n}{pq}}\Big).$$

II. Какое наименьшее число испытаний нужно произвести для того, чтобы с вероятностью, не меньшей $\beta$, частота наступления события отклонялась от вероятности не более чем на $\alpha$?

$$\mathbf{P}\Big\{\big|\frac{\mu_n}{n} - p\big| \le \alpha\Big\} \approx 2\Phi_0\Big(\alpha\sqrt{\frac{n}{pq}}\Big) \ge \beta\,.$$

III. При данной вероятности $\beta$ и числе испытаний $n$ определить границу возможных значений $|\frac{\mu_n}{n} - p|$?

$$\mathbf{P}\Big\{\big|\frac{\mu_n}{n} - p\big| \le \alpha\Big\} \approx 2\Phi_0\Big(\alpha\sqrt{\frac{n}{pq}}\Big) = \beta\,.$$

## 2.2. Случайные величины (общий случай)

### Определение случайной величины

Пусть $(\Omega, \mathcal{U}, \mathbf{P})$ – произвольное вероятностное пространство.

**Определение 18.** Функция $\xi = F(\omega)$, определенная на $\Omega$ и принимающая вещественные значения, называется случайной величиной, если для любого $x \in \mathbb{R}$ множество $\{\omega \in \Omega : \xi(\omega) \le x\} = \{\xi \le x\}$ является случайным событием, то есть принадлежит $\mathcal{U}$.

**Определение 19.** Функция $F(x) = F_\xi(x) = \mathbf{P}\{\xi \le x\}$ называется функцией распределения случайной величины $\xi$.

По свойству непрерывности вероятности $\mathbf{P}$ имеем $\mathbf{P}(\xi < x) = F(x-0)$. Поэтому

**1)** $\mathbf{P}(x_1 < \xi \le x_2) = F(x_2) - F(x_1)$;

**2)** $\mathbf{P}(\xi = x) = F(x) - F(x-0)$;

**3)** $\mathbf{P}(x_1 \le \xi \le x_2) = F(x_2) - F(x_1 - 0)$;

**4)** $\mathbf{P}(x_1 < \xi < x_2) = F(x_2 - 0) - F(x_1)$;

**5)** $\mathbf{P}(x_1 \le \xi < x_2) = F(x_2 - 0) - F(x_1 - 0)$.

**Теорема 2.16.** *Функция $F(x)$ не убывает, $F(x)$ непрерывна справа и удовлетворяет соотношениям*

$$F(+\infty) = \lim_{x\to+\infty} F(x) = 1, \quad F(-\infty) = \lim_{x\to-\infty} F(x) = 0\,.$$

*Доказательство.* Ясно, что $B_n = \{x < \xi \le x + \frac{1}{n}\} \downarrow \emptyset$. Поэтому $\mathbf{P}(B_n) = F(x+\frac{1}{n}) - F(x) \to 0$. Введем $A_n = \{n-1 < \xi \le n\}$. Имеем $\Omega = \sum_{-\infty}^{\infty} A_n$. Поэтому

$$1 = \lim_{N\to\infty} \sum_{-N}^{N} \mathbf{P}(A_n) = \lim_{N\to\infty} \big[F(N) - F(-N)\big]\,. \quad \square$$

**Определение 20.** $\sigma$-алгебра $\mathcal{B}$ подмножеств $\mathbb{R}$, порожденная интервалами вида $a < x \le b$, называется борелевской. Элементы борелевской $\sigma$-алгебра называются борелевскими множествами.

Если $\xi$ – случайная величина, то система множеств $\{\xi^{-1}(B)\,,\, B \in \mathcal{B}\}$ является $\sigma$-алгеброй, обозначается символом $\mathcal{U}_\xi$ и называется $\sigma$-алгеброй, порожденной случайной величиной $\xi$. $\mathcal{U}_\xi$ является $\sigma$-подалгеброй $\sigma$-алгебры $\mathcal{U}$.

**Теорема 2.17** (теорема о продолжении функции вероятности). *Пусть на алгебре $\mathcal{V}_0$ подмножеств $\Omega$ определена вероятность $\mathbf{P}$, удовлетворяющая аксиомам*

1° $\mathbf{P}(A) \ge 0$ *для всех* $A \in \mathcal{V}_0$*;*

2° $\mathbf{P}(\Omega) = 1$*;*

3° $\mathbf{P}(A + B) = \mathbf{P}(A) + \mathbf{P}(B)$*, если* $A \cap B = \emptyset$*;*

4° *Если события* $A_n$ *в последовательности* $\{A_n\}$ *попарно несовместны и* $\sum_{n=1}^{\infty} A_n \in \mathcal{V}_0$ *, то*

$$\mathbf{P}\big(\sum_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)\,.$$

*Тогда вероятность* $\mathbf{P}$ *можно единственным образом продолжить на минимальную* $\sigma$*-алгебру* $\mathcal{V}$*, порожденную* $\mathcal{V}_0$*.*

Система промежутков $\{(x_1, x_2]\}$ порождает алгебру $\mathcal{B}_0$, состоящую из всевозможных конечных сумм вида $\sum_{k=1}^{n}(x_{1k}, x_{2k}]$. Соотношение $\mathbf{P}_\xi(B) = \mathbf{P}\{\xi \in B\} = F(x_1) - F(x_2)$ определяет на алгебре $\mathcal{B}_0$ $\sigma$-аддитивную функцию множества, удовлетворяющую условиям 1° – 4°. По теореме о продолжении вероятности функция множества $\mathbf{P}_\xi$ продолжается на минимальную $\sigma$-алгебру, совпадающую с $\sigma$-алгеброй борелевских множеств $\mathcal{B}$. Вероятность $\mathbf{P}_\xi(B) = \mathbf{P}\{\xi \in B\}$, $B \in \mathcal{B}$, представляется интегралом Лебега-Стилтьеса

$$\mathbf{P}_\xi(B) = \int\limits_B dF_\xi(x)\,.$$

Тройка $\left(\mathbb{R}, \mathcal{B}, \mathbf{P}_\xi\right)$ является новым вероятностным пространством, порожденным случайной величиной $\xi$.

**Определение 21.** Функция $\mathbf{P}_\xi(B)$, определенная для всех $B \in \mathcal{B}$, называется распределением вероятностей случайной величины $\xi$.

Пусть $F(x)$ имеет не более счетного числа точек разрыва $x_1, x_2, \ldots$, $\mathbf{P}\{\xi = x_k\} = p_k$ и $\sum_k p_k = 1$. В этом случае $\xi$ называется дискретной случайной величиной, говорят также, что $\xi$ имеет дискретное распределение.

*Пример 9. Биномиальное распределение*

$$\mathbf{P}\{\xi = k\} = C_n^k p^k q^{n-k}, \;\; k = 0, \ldots, n\,, \;\; p \in (0,1), \;\; q = 1 - p\,.$$

*Пример 10. Пуассоновское распределение*

$$\mathbf{P}\{\xi = k\} = \frac{a^k}{k!} e^{-a}, \;\; a > 0, \;\; k = 0, 1, \ldots.$$

*Пример 11. Геометрическое распределение*

$$\mathbf{P}\{\xi = k\} = p\,(1-p)^{k-1}, \;\; k = 1, 2, \ldots, \;\; p \in (0,1)\,.$$

Суммируемая функция $p_\xi(x)$ называется плотность случайной величины $\xi$, если

$$\mathbf{P}\{x_1 < \xi \le x_2\} = F_\xi(x_2) - F_\xi(x_1) = \int\limits_{x_1}^{x_2} p_\xi(x)dx\,,$$

Ясно, что

1°. $F'_\xi(x) = p_\xi(x)$ в точках непрерывности $p_\xi(x)$;

2°. $F_\xi(x) = \int_{-\infty}^{x} p_\xi(x)dx$.

Из свойств функции распределения следует, что

$$p_\xi(x) \ge 0, \quad \int\limits_{-\infty}^{\infty} p_\xi(x)dx = 1\,.$$

Случайная величина $\xi$, имеющая плотность, называется абсолютно непрерывной случайной величиной, говорят также, что $\xi$ имеет абсолютно непрерывное распределение. При этом

$$\mathbf{P}_\xi(B) = \int\limits_B p_\xi(x)dx\,.$$

*Пример 12. Нормальное (гауссовское) распределение.*

Случайная величина $\xi$ имеет нормальное распределение с параметрами $(a, \sigma)$, $-\infty < a < \infty$, $\sigma > 0$, если $\xi$ абсолютно непрерывна и ее плотность имеет вид

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

*Пример 13. Равномерное распределение.*

Случайная величина $\xi$ равномерно распределена на отрезке $[a, b]$, если $\xi$ абсолютно непрерывна и ее плотность имеет вид

$$p_\xi(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b]. \end{cases}$$

*Пример 14. $\gamma$-распределение.*

Случайная величина $\xi$ имеет $\gamma$-распределение, если ее плотность представляется в виде

$$p_\xi(x) = \begin{cases} 0, & x < 0 \\ \dfrac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, & x \geq 0 \end{cases}$$

с параметрами $\alpha > 0$ и $\lambda > 0$.

В частности, при $\alpha = 1$ получаем плотность показательного распределения. Тогда случайная величина $\xi$ называется показательно распределенной.

Если существует такое случайное событие $B \in \mathcal{B}$ лебеговой меры нуль, что $\mathbf{P}_\xi(B) = 1$, то случайная величина $\xi$ называется сингулярной или имеющей сингулярное распределение.

В общем случае любая функция распределения $F(x)$ представима в виде

$$F(x) = a_1 F_1(x) + a_2 F_2(x) + a_3 F_3(x),$$

где $a_i \geq 0$, $a_1 + a_2 + a_3 = 1$, $F_1(x)$ – дискретная функция распределения, $F_2(x)$ – функция распределения, имеющая плотность, $F_3(x)$ – сингулярная функция распределения.

**Определение 22.** Функция $g(u)$ называется борелевской функцией, если для любого $B \in \mathcal{B}$ полный прообраз $g^{-1}(B)$ принадлежит $\mathcal{B}$.

**Теорема 2.18.** *Если $\xi$ – случайная величина и $g(u)$ – борелевская функция, то $\tau = g(\xi)$ является случайной величиной.*

*Доказательство.* Пусть $B \in \mathcal{B}$. Тогда $g^{-1}(B) = B_1 \in \mathcal{B}$. Так как $\tau^{-1} = \xi^{-1}(B_1) \in \mathcal{U}$, то $\tau$ – случайная величина. □

*Пример 15.* Пусть $g(x) = x^2$, $F_\xi(x)$ – абсолютно непрерывная функция распределения с плотностью $p_\xi(x)$ и $\tau = \xi^2$. При $x \geq 0$ из соотношения

$$F_\eta(x) = \mathbf{P}\{\xi^2 \leq x\} = \mathbf{P}\{-\sqrt{x} \leq \xi \leq \sqrt{x}\} = F_\xi(\sqrt{x}) - F_\xi(-\sqrt{x})$$

получаем

$$p_\eta(x) = \frac{1}{2\sqrt{x}} \big(p_\xi(\sqrt{x}) + p_\xi(-\sqrt{x})\big).$$

### Многомерные распределения

Пусть $\xi_1, \ldots, \xi_n$ – случайные величины. Вектор $\xi$ со случайными компонентами $\xi_1, \ldots, \xi_n$ будем называть случайным вектором. Функция вида

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \mathbf{P}\{\xi_1 \leq x_1, \ldots, \xi_n \leq x_n\}$$

называется функцией распределения случайного вектора. Сформулируем некоторые свойства функции распределения:

1° $\lim\limits_{x\to+\infty} F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n) = F_{\xi_1,\dots,\xi_{n-1}}(x_1,\dots,x_{n-1})$ (согласованность).

2° $\lim\limits_{x\to-\infty} F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n) = 0$ ..

3° $F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n)$ не убывает и непрерывна справа по каждому аргументу.

4° $F_{\xi_1,\dots,\xi_n}(+\infty,\dots,+\infty) = 1$.

5° $\mathbf{P}\{x_1 < \xi_1 \le x_1 + h_1,\dots,x_n < \xi_n \le x_n + h_n\} = \Delta_{h_1,\dots,h_n} F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n) =$
$\sum\limits_{\theta_1,\dots,\theta_n=0}^{1} (-1)^{n+\theta_1+\dots+\theta_n} F_{\xi_1,\dots,\xi_n}(x_1+\theta_1 h_1,\dots,x_n+\theta_n h_n)$.
Последнее соотношение доказывается методом математической индукции.

Рассмотрим алгебру $\mathcal{B}_0^n$ всевозможных конечных объединений прямоугольников вида $\Pi_{i=1}^n\{a_i < \xi_i \le b_i\}$. Наименьшая $\sigma$-алгебра, содержащая $\mathcal{B}_0^n$, называется $\sigma$-алгеброй $\mathcal{B}^n$ борелевских множеств. $\sigma$-алгебра $\mathcal{B}^n$ является замыканием алгебры $\mathcal{B}_0^n$ относительно операций $\cup$ и $\cap$, взятых в счетном числе. Поэтому $\sigma$-алгебра $\mathcal{U}_\xi = \{\xi \in B,\, B \in \mathcal{B}^n\} = \{U = \xi^{-1}(B),\, B \in \mathcal{B}^n\}$, порожденная случайным вектором $\xi = (\xi_1,\dots,\xi_n)$ содержится в исходной $\sigma$-алгебре $\mathcal{U}$ вероятностного пространства $(\Omega, \mathcal{U}, \mathbf{P})$, $\mathcal{U}_\xi \subset \mathcal{U}$.

Элементы алгебры $\mathcal{B}_0^n$ представляются в виде объединения попарно непересекающихся прямоугольников вида $\Pi_{i=1}^n\{a_i < \xi_i \le b_i\}$. Пусть $B = \sum_k B_k$, где $B_k$ – попарно непересекающиеся прямоугольники. Определим функцию множества, положив $\mathbf{P}_\xi(B) = \sum_k \mathbf{P}\{\xi \in B_k\}$ (см. 5°). Доказывается, что так определенная функция множества $\mathbf{P}_\xi$ не зависит от представления $B \in \mathcal{B}_0^n$ в виде объединения попарно непересекающихся прямоугольников и $\sigma$-аддитивна на алгебре $\mathcal{B}_0^n$. С помощью теоремы о продолжении вероятности функция $\mathbf{P}_\xi$ продолжается единственным образом до функции вероятности на $\sigma$-алгебре $\mathcal{B}^n$ и полностью определяется функцией распределения $F_\xi$. Так как $\mathcal{U}_\xi \subset \mathcal{U}$, имеет место равенство $\mathbf{P}_\xi(B) = \mathbf{P}\{\xi \in B\}$.

Тройка $(\mathbb{R}^n, \mathcal{B}^n, \mathbf{P}_\xi)$ является новым вероятностным пространством. Вероятностная функция $\mathbf{P}_\xi$ называется распределением вероятностей случайного вектора $\xi = (\xi_1,\dots,\xi_n)$.

*Дискретное вероятностное пространство* задается конечным или счетным набором точек $x^k = (x_1,\dots,x_n)$ и неотрицательных чисел $p_k = \mathbf{P}\{\xi = x^k\}$, удовлетворяющих $\sum_k p_k = 1$. Вероятность случайного события $B \in \mathcal{B}^n$ определяется соотношением

$$\mathbf{P}_\xi(B) = \sum_{x^k \in B} p_k\,.$$

*Многомерной плотностью распределения* называется такая функция $p_\xi(x)$, $x \in \mathbb{R}^n$, что $\mathbf{P}_\xi(B) = \int_B p_\xi(x)dx$, $B \in \mathcal{B}^n$. Плотность распределения неотрицательна, $p_\xi(x) \ge 0$ и нормирована условием $\int\limits_{\mathbb{R}^n} p_\xi(x)dx = 1$. Плотность распределения $p_\xi$ и функция распределения $F_\xi$ связаны соотношением

$$p_\xi(x_1,\dots,x_n) = \frac{\partial^n F_\xi(x_1,\dots,x_n)}{\partial x_1 \dots \partial x_n}$$

в точках непрерывности $p_\xi$.

*Пример 16.* Распределение (распределение вероятностей) вектора $(\xi_1,\dots,\xi_n)$ называется *невырожденным нормальным*, если оно имеет плотность вида

$$p(x_1,\dots,x_n) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}Q(x_1,\dots,x_n)\}\,,$$

где $Q(x_1,\dots,x_n) = \sum\limits_{i,j=1}^{n} a_{ij}x_i x_j$ – положительно определенная невырожденная форма, $|A\,|$ – определитель матрицы $\|\, a_{ij}\,\|$.

*Пример 17.* Распределение случайного вектора $\xi = (\xi_1, \ldots, \xi_n)$ называется равномерным в области $\Omega \subset \mathbb{R}^n$, если его плотность распределения равна

$$p(x_1, \ldots, x_n) = \begin{cases} |\Omega|^{-1}, & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases}$$

В этом случае для любого борелевского множества $B \subset \mathbb{R}^n$ имеем

$$\mathbf{P}_\xi(B) = \frac{|B \cap \Omega|}{|\Omega|}.$$

### Математическое ожидание случайной величины

Пусть $F_\xi(x) = \mathbf{P}\{\xi \le x\}$ — функция распределения случайной величины $\xi$ на вероятностном пространстве $(\Omega, U, \mathbf{P})$ и $\mathbf{P}_\xi(B)$, $B \in \mathcal{B}$, распределение вероятностей случайной величины $\xi$.

**Определение 23.** Математическим ожиданием случайной величины $\xi$ называется число $M\xi$, равное

$$M\xi = \int_\Omega \xi(\omega)\mathbf{P}(d\omega),$$

если последний интеграл существует.

Пусть $g$ борелевская функция на $\mathbb{R}$ и $\int_\mathbb{R} |g(x)|\mathbf{P}_\xi(dx) < \infty$. Так как функции $g(x)$ $x \in \mathbb{R}$ и $g(\xi(\omega))$, $\omega \in \Omega$, равноизмеримы в пространствах с мерой $(\mathbb{R}, \mathcal{B}, \mathbf{P}_\xi)$ и $(\Omega, \mathcal{U}, \mathbf{P})$, то есть $\mathbf{P}_\xi\{g \leqslant t\} = \mathbf{P}\{g(\xi) \leqslant t\}$, то имеет место равенство

$$\int_\Omega g(\xi(\omega))\mathbf{P}(d\omega) = \int_\mathbb{R} g(x)\mathbf{P}_\xi(dx).$$

Мера $\mathbf{P}_\xi$ порождается функцией распределения $F_\xi$, поэтому

$$Mg(\xi) = \int_\Omega g(\xi)(\omega)\mathbf{P}(d\omega) = \int_\mathbb{R} g(x)\mathbf{P}_\xi(dx) = \int_\mathbb{R} g(x)F_\xi(dx).$$

Последний интеграл носит название интеграла Лебега-Стилтьеса. В частности,

$$M\xi = \int_\Omega \xi(\omega)\mathbf{P}(d\omega) = \int_\mathbb{R} x\mathbf{P}_\xi(dx) = \int_\mathbb{R} xF_\xi(dx)$$

Для дискретной случайной величины $\xi$ и борелевской функции $g$ имеем

$$\int_\mathbb{R} g(x)dF_\xi(x) = \sum_k g(x_k)[F_\xi(x_k) - F_\xi(x_k - 0)] = \sum_k g(x_k)\mathbf{P}\{\xi = x_k\}.$$

Для абсолютно непрерывной случайной величины $\xi$ имеем

$$\int_\mathbb{R} g(x)dF_\xi(x) = \int_\mathbb{R} g(x)p_\xi(x)dx,$$

где $p_\xi$ – плотность распределения случайной величины $\xi$.

Из определения интеграла Лебега следует, что математическое ожидание $M\xi$ конечно тогда и только тогда, когда $M|\xi| < +\infty$ ($M\xi$ может существовать, но не быть конечным).

Таким образом, если $\xi$ – дискретная случайная величина, то $M\xi = \sum_k x_k\mathbf{P}\{\xi = x_k\}$. Если $F_\xi$ имеет плотность, то $M\xi = \int_\mathbb{R} xp_\xi(x)dx$.

*Свойства математического ожидания*:

**1.** Если $\mathbf{P}\{\xi = c\} = 1$, то $M\xi = c$;

**2.** $M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2$, если существуют два из трех математических ожиданий;

**3.** Если $\xi_1 \le \xi_2$, то $M\xi_1 \le M\xi_2$, в частности, $|M\xi_1| \le M|\xi_2|$;

**4.** Если $I_A$ – индикатор события $A$, то $MI_A = \mathbf{P}(A)$;

**5.** Если $\xi \ge 0$ и $M\xi = 0$, то $\xi = 0$ с вероятностью 1;

*Доказательство.* Ясно, что $\{\xi > 0\} = \sum\limits_n \{\xi \ge \frac{1}{n}\}$ и $\{\xi \ge \frac{1}{n}\} \subset \{\xi \ge \frac{1}{n+1}\}$. Поэтому

$$0 = M\xi = M\xi I_{\{\xi=0\}} + M\xi I_{\{\xi>0\}} = M\xi I_{\{\xi>0\}} \ge M\xi I_{\{\xi\ge\frac{1}{n}\}} \ge \frac{1}{n} M I_{\{\xi\ge\frac{1}{n}\}} = \frac{1}{n}\mathbf{P}\{\xi \ge \frac{1}{n}\}\,.$$

Из свойства непрерывности вероятности получаем, что $\mathbf{P}\{\xi > 0\} = 0$. □

**6.** Для дисперсии $D\xi = M(\xi - M\xi)^2$ имеет место равенство $D\xi = M\xi^2 - (M\xi)^2$.

*Пример 18.* Пусть $\xi$ – нормально распределенная случайная величина с параметрами $(a, \sigma)$. Тогда

$$M\xi = \int\limits_{\mathbb{R}} (t-a)\frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(t-a)^2}{2\sigma^2}\Big\}dt + a\int\limits_{\mathbb{R}} \exp\Big\{-\frac{(t-a)^2}{2\sigma^2}\Big\}dt = a\,,$$

и

$$D\xi = \sigma\,.$$

Если $\xi$ нормально распределена с параметрами $(0,1)$, то $\eta = \sigma\xi + a$ нормально распределена с параметрами $(a, \sigma)$.

*Пример 19.* Пусть $\xi$ – случайная величина, распределенная по закону Пуассона с параметром $\mu$. Тогда

$$M\xi = \sum_n n\frac{\mu^n}{n!}e^{-\mu} = \mu$$

и

$$D\xi = \mu\,.$$

*Пример 20.* Пусть случайная величина $\xi$ равномерно распределена на отрезке $[a, b]$. Тогда

$$M\xi = \frac{1}{a-b}\int\limits_a^b x dx = \frac{a+b}{2} \;\text{ и }\; D\xi = \frac{(b-a)^2}{12}\,.$$

**7.** Если $0 \le \xi_n \nearrow \xi$, то $M\xi_n \to M\xi$ (теорема о монотонной сходимости).

**Следствие.** *Если ряд* $\sum\limits_{n=1}^{\infty} \xi_n$ *состоит из неотрицательных случайных величин, то*

$$M\Big(\sum_n \xi_n\Big) = \sum_n M\xi_n\,.$$

**8.** Если $\lim\limits_{n\to\infty} \xi_n = \xi$ и для любого $\omega \in \Omega$ выполняется неравенство $|\xi_n(\omega)| \le \eta(\omega)$, где $M\eta \le +\infty$, то $\lim\limits_{n\to\infty} M\xi_n = M\xi$ (теорема Лебега о мажорируемой сходимости).

**9.** Если случайная величина $\xi$ имеет плотность $p_\xi(x)$, $g(x)$ – борелевская функция и

$$\int\limits_{\mathbb{R}} |g(x)|p_\xi(x)dx < +\infty\,,$$

то

$$Mg(\xi) = \int\limits_{\mathbb{R}} g(x)\,p_\xi(x)dx\,.$$

### Независимые случайные величины

Одним из важнейших понятий теории вероятностей является понятие независимости случайных величин.

**Определение 24.** Случайные величины $\xi_1, \ldots, \xi_n$ называются *независимыми*, если

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n) .$$

**Определение 25.** Случайные величины последовательности $\{\xi_n\}_{n=1}^{\infty}$, заданные на вероятностном пространстве $\big(\Omega, \mathcal{U}, \mathbf{P}\big)$, называются независимыми, если для любого числа $k$

$$F_{\xi_1,\ldots,\xi_k}(x_1, \ldots, x_k) = F_{\xi_1}(x_1) \cdots F_{\xi_k}(x_k) .$$

Отметим также, что для последовательности независимых случайных величин будут независимыми любые два события, относящиеся к непересекающимся группам случайных величин этой последовательности.

**Теорема 2.19.** *Пусть $\xi_1, \ldots, \xi_n$ – независимые случайные величины. Если $F_{\xi_i}(x)$ есть функция распределения абсолютно непрерывной случайной величины $\xi_i$ с плотностью $p_{\xi_i}(x)$, $i = 1, \ldots, n$, то $F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$ имеет плотность $p_{\xi_1,\ldots,\xi_k}(x_1, \ldots, x_k)$, равную $p_{\xi_1}(x_1) \cdots p_{\xi_n}(x_n)$. Обратно, если $F_{\xi_1,\ldots,\xi_k}(x_1, \ldots, x_k)$ абсолютно непрерывна, то и $F_{\xi_i}(x)$ тоже абсолютно непрерывны для всех $i$. При этом*

$$p_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = p_{\xi_1}(x_1) \cdots p_{\xi_n}(x_n)$$

*почти всюду.*

*Доказательство.* Если функция распределения каждой случайной величины абсолютно непрерывна и

$$F_{\xi_i}(x_i) = \int\limits_{-\infty}^{x_i} p_{\xi_i}(t_i) dt_i ,$$

то совместная функция распределения имеет вид

$$\begin{aligned} F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) &= F_{\xi_1}(x_1) \cdots F_{\xi_n}(x_n) = \\ &= \int\limits_{-\infty}^{x_1} p_{\xi_1}(t_1) dt_1 \cdots \int\limits_{-\infty}^{x_n} p_{\xi_n}(t_n) dt_n = \\ &= \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_n} p_{\xi_1}(t_1) \cdots p_{\xi_n}(t_n) dt_1 \ldots dt_n . \end{aligned}$$

Обратно, предположим, что

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_n} p_{\xi_1,\ldots,\xi_n}(t_1, \ldots, t_n) dt_1 \ldots dt_n ,$$

Тогда

$$\begin{aligned} F_{\xi_1}(x_1) &= F_{\xi_1,\ldots,\xi_n}(x_1, \infty, \ldots, \infty) = \\ &= \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} p_{\xi_1,\ldots,\xi_n}(t_1, \ldots, t_n) dt_1 \ldots dt_n , \end{aligned}$$

так что

$$p_{\xi_1}(t_1) = \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} p_{\xi_1,\ldots,\xi_n}(t_1, t_2, \ldots, t_n) dt_2 \ldots dt_n \,.$$

Равенства имеет место почти всюду, поэтому функции $p_{\xi_i}(x_i)$ определены почти всюду. □

Дадим другое определение независимости случайных величин.

**Определение 26.** Случайные величины $\xi_1, \ldots, \xi_n$ называется независимыми, если для любых борелевских $B_i \in \mathcal{B}$, $i = 1, \ldots, n$, имеет месть равенство

$$\mathbf{P}\{\xi_1 \in B_1, \ldots \xi_n \in B_n\} = \mathbf{P}\{\xi_1 \in B_1\} \cdot \ldots \cdot \mathbf{P}\{\xi_n \in B_n\} \,.$$

Приведенные определения независимости случайных величин эквивалентны. Для того, чтобы доказать эквивалентность воспользуемся следующей теоремой теории меры.

**Определение 27.** Алгебры событий $\mathcal{A}_1$ и $\mathcal{A}_2$ называются независимыми, если для любых событий $A_1 \in \mathcal{A}_1$ и $A_2 \in \mathcal{A}_2$ имеет место равенство

$$\mathbf{P}(A_1 A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2) \,.$$

**Теорема 2.20.** *$\sigma$-алгебры $\mathcal{U}_1$ и $\mathcal{U}_2$, порожденные соответственно независимыми алгебрами событий $\mathcal{A}_1$ и $\mathcal{A}_2$ независимы.*

Теперь мы можем сформулировать утверждение, доказывающую эквивалентность определений 25 и 26.

**Теорема 2.21.** *Случайные величины $\xi_1$ и $\xi_2$ независимы тогда и только тогда, когда независимы $\sigma$-алгебры $\mathcal{U}_{\xi_1}$ и $\mathcal{U}_{\xi_2}$, порожденные этими случайными величинами.*

*Доказательство.* Пусть $\xi_1$ и $\xi_2$ независимы и $\Delta$ и $\Lambda$ – полуинтервалы вида $(a_1, a_2]$ и $(b_1, b_2]$. Имеет место цепочка равенств

$$\begin{aligned}
\mathbf{P}\{\xi_1 \in \Delta,\ \xi_2 \in \Lambda\} &= \mathbf{P}\{\xi_1 \in (a_1, a_2], \xi_2 \in (b_1, b_2]\} = \\
&= F_{\xi_1,\xi_2}(a_2, b_2) - F_{\xi_1,\xi_2}(a_1, b_2) - F_{\xi_1,\xi_2}(a_2, b_1) + F_{\xi_1,\xi_2}(a_1, b_1) = \\
&= \big(F_{\xi_1}(a_2) - F_{\xi_1}(a_1)\big)\big(F_{\xi_2}(b_2) - F_{\xi_1}(b_1)\big) = \mathbf{P}\{\xi \in \Delta\}\mathbf{P}\{\xi_2 \in \Lambda\} \,.
\end{aligned}$$

Далее, если $\Delta_i$, $i = 1, \ldots, n$, и $\Lambda_j$, $j = 1, \ldots, m$ – две системы непересекающихся интервалов, то

$$\begin{aligned}
\mathbf{P}\{\xi_1 \in \cup_{i=1}^n \Delta_i,\ \xi_2 \in \cup_{j=1}^m \Lambda_j\} &= \\
&= \sum_{i,j} \mathbf{P}\{\xi_1 \in \Delta_i,\ \xi_2 \in \Lambda_j\} = \sum_{i,j} \mathbf{P}\{\xi_1 \in \Delta_i\}\mathbf{P}\{\xi_2 \in \Lambda_j\} = \\
&= \mathbf{P}\{\xi_1 \in \cup_{i=1}^n \Delta_i\}\mathbf{P}\{\xi_2 \in \cup_{i=1}^n \Lambda_j\} \,.
\end{aligned}$$

Таким образом, алгебры, порожденные полуинтервалами $\Delta_i$ и $\Lambda_j$ независимы. Из теоремы 2.20 следует независимость $\mathcal{U}_{\xi_1}$ и $\mathcal{U}_{\xi_2}$.

Обратное утверждение следует из равенств

$$\begin{aligned}
F_{\xi_1,\xi_2}(x_1, x_2) &= \mathbf{P}\{\xi_1 \le x_1,\ \xi_2 \le x_2\} = \\
&= \mathbf{P}\{\xi_1 \le x_1\}\mathbf{P}\{\xi_2 \le x_2\} = F_{\xi_1}(x_1)F_{\xi_2}(x_2) \,. \ \square
\end{aligned}$$

Эта теорема справедлива для любого числа случайных величин.

**Следствие.** *Пусть $\varphi_1$, $\varphi_2$ – борелевские функции и $\xi_1$ и $\xi_2$ – независимые случайные величины. Тогда $\eta_1 = \varphi_1(\xi_1)$ и $\eta_2 = \varphi_2(\xi_2)$ – также независимые случайные величины.*

*Доказательство.* Очевидно, что $\mathcal{U}_{\eta_1} \subset \mathcal{U}_{\xi_1}$ и $\mathcal{U}_{\eta_2} \subset \mathcal{U}_{\xi_2}$ Поэтому из независимости $\sigma$-алгебр $\mathcal{U}_{\xi_1}$ и $\mathcal{U}_{\xi_2}$ следует независимость $\mathcal{U}_{\eta_1}$ и $\mathcal{U}_{\eta_2}$. □

*Замечание.* Если $\xi = (\xi_1, \dots, \xi_n)$, $\eta = (\eta_1, \dots, \eta_m)$ — независимые случайные вектора $\varphi_1(x_1, \dots, x_n)$, $\varphi_2(y_1, \dots, y_n)$ —борелевские функции, то случайные величины $\varphi_1(\xi_1, \dots, \xi_n)$, $\varphi_2(\eta_1, \dots, \eta_n)$ независимы.

*Замечание.* Из теоремы 2.21 следует, что если компоненты $\xi_1$ и $\xi_2$ случайного вектора $\xi = (\xi_1,\, \xi_2)$ независимы, то для борелевских множеств $B = B_1 \times B_2 = \{(x_1, x_2) : x_1 \in B_1,\ x_2 \in B_2\}$, $B_1,\, B_2 \in \mathcal{B}$, выполняется равенство

$$\mathbf{P}\{\xi \in B\} = \mathbf{P}\{\xi_1 \in B_1,\, \xi_2 \in B_2\} = \mathbf{P}\{\xi_1 \in B_1\}\mathbf{P}\{\xi_2 \in B_2\}\,.$$

В это случае говорят, что мера $\mathbf{P}_\xi$ равна произведению мер $\mathbf{P}_{\xi_1}$ и $\mathbf{P}_{\xi_2}$. Интеграл

$$\int\limits_B g(x_1, x_2)\mathbf{P}_{\xi_1,\xi_2}(dx_1, dx_2)$$

в этом случае может быть записан в виде повторного интеграла:

$$\iint\limits_{B_1 \times B_2} g(x_1, x_2)\mathbf{P}_{\xi_1}(dx_1)\mathbf{P}_{\xi_2}(dx_2)\,.$$

Если $g(x_1, x_2) = g_1(x_1)g_2(x_2)$, тогда

$$\int\limits_B g(x_1, x_2)\mathbf{P}_{\xi_1,\xi_2}(dx_1, dx_2) = \int\limits_{B_1} g_1(x_1)\mathbf{P}_{\xi_1}(dx_1)\cdot \int\limits_{B_2} g_2(x_2)\mathbf{P}_{\xi_2}(dx_2)$$

при условии, что интегралы сходятся.

**Теорема 2.22** (формула композиции). *Пусть $\xi$ и $\eta$ – независимые абсолютно непрерывные случайные величины, $p_\xi$ и $p_\eta$ – их плотности. Тогда*

$$p_{\xi+\eta}(x) = \int\limits_{\mathbb{R}} p_\xi(x-y)p_\eta(y)dy = \int\limits_{\mathbb{R}} p_\xi(y)p_\eta(x-y)dy\,.$$

*Доказательство.* Плотность $p_{\xi,\eta}(x,y)$ совместного распределения случайного вектора с компонентами $\xi$ и $\eta$ равна произведению плотностей $p_\xi(x)$ и $p_\eta(y)$. Тогда

$$F_{\xi+\eta}(z) = \mathbf{P}\{\xi+\eta \le z\} = \int\limits_{x+y\le z} p_\xi(x)p_\eta(y)dxdy = \quad \text{(по теореме Фубини)}$$

$$= \int\limits_{-\infty}^{\infty} p_\eta(x)\Big(\int\limits_{-\infty}^{z-x} p_\eta(y)dy\Big)dx = \int\limits_{-\infty}^{\infty} F_\eta(z-x)p_\eta(x)dx =$$

$$= \int\limits_{-\infty}^{\infty} p_\eta(x)\Big(\int\limits_{-\infty}^{z} p_\eta(y-x)dy\Big)dx = \int\limits_{-\infty}^{z} dz \int\limits_{-\infty}^{\infty} p_\eta(x)p_\eta(y-x)dx$$

Таким образом,

$$F_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} F_\eta(z-x)p_\eta(x)dx$$

и

$$p_{\xi+\eta}(z) = \int\limits_{-\infty}^{\infty} p_\eta(z-x)p_\eta(x)dx\,. \ \square$$

*Пример 21.* Пусть $\xi_1$ и $\xi_2$ независимы и равномерно распределены на отрезке $[0,1]$. Тогда

$$p_{\xi_1}(x) = p_{\xi_2}(x) = \begin{cases} 1, & x \in [0,1]\,, \\ 0, & x \notin [0,1] \end{cases}$$

и

$$p_{\xi_1+\xi_2}(z) = \int_0^1 p_{\xi_1}(x-t)dt = \begin{cases} 0, & x \notin [0,2]\,, \\ x, & x \in [0,1]\,, \\ 2-z & x \in [1,2]\,. \end{cases}$$

**Мультипликативное свойство**

**Теорема 2.23.** *Если случайные величины $\xi$ и $\eta$ независимы и имеют конечное математическое ожидание $M\xi$ и $M\eta$, то*

$$M\xi\eta = M\xi \cdot M\eta\,.$$

Приведем *набросок доказательства.* Случайную величину $\xi$ будем называть простой, если она представима в виде

$$\xi = \sum_{k=1}^{m} c_k I_{A_k}\,,$$

где $A_k$ – попарно несовместные случайные события.

1. Если $\xi \geq 0$, то существует такая последовательность простых неотрицательных случайных величин $\xi_n$, $\xi_n \geq 0$, что $\xi_n \nearrow \xi$ почти наверное. В качестве $\xi_n$ можно рассмотреть

$$\xi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{\{\frac{k-1}{2^n} < \xi \leq \frac{k}{2^n}\}}\,.$$

Ясно, что $0 \leq \xi_n \leq \xi_{n+1} \leq \xi$ и при $\xi_n(\omega) \leq n$ имеет место неравенство $\xi(\omega) \leq \xi_n(\omega) + \frac{1}{2^n}$.

2. Пусть $\xi$ и $\eta$ независимые простые случайные величины, принимающие значения $x_1 < \cdots < x_m$ и $y_1 < \cdots < y_n$. Тогда

$$\xi = \sum_{k=1}^{m} x_k I_{\{\xi = x_k\}}, \quad \eta = \sum_{k=1}^{n} y_k I_{\{\eta = y_k\}}$$

и

$$\mathbf{P}\{\xi = x_k, \eta = y_l\} = \mathbf{P}\{\xi = k_k\} \cdot \mathbf{P}\{\eta = y_l\}\,.$$

В этом случае, как нам известно, $M\xi\eta = M\xi \cdot M\eta$ (см. теорему 2.3).

3. Напомним, что если $\xi$ и $\eta$ независимы и $g_1$ и $g_2$ – борелевские функции, то случайные величины $g_1(\xi)$ и $g_2(\eta)$ также независимы.

В обозначениях пункта 1 имеем

$$M(\xi\eta) = \lim_{n\to\infty} M(\xi_n\eta_n) = \lim_{n\to\infty} M\xi_n \cdot \lim_{n\to\infty} M\eta_n = M\xi \cdot M\eta\,.$$

4. В общем случае $\xi = \xi^+ - \xi^-$ и $\eta = \eta^+ - \eta^-$, где $\xi^{\pm}$ и $\eta^{\pm}$ независимые как функции $\xi$ и $\eta$ и неотрицательные случайные величины. Поэтому

$$M(\xi\eta) = M(\xi^+ - \xi^-)(\eta^+ - \eta^-) = M(\xi^+ - \xi^-) \cdot M(\eta^+ - \eta^-) = M\xi \cdot M\eta\,. \ \square$$

**Следствие.** *Если случайные величины $\xi_1, \ldots, \xi_n$ независимы и имеют конечное математическое ожидание, то*

$$M(\xi_1 \cdot \cdots \cdot \xi_n) = M\xi_1 \cdot \cdots \cdot M\xi_n\,.$$

### Условное математическое ожидание

Пусть $(\Omega,\mathcal{U},\mathbf{P})$ – вероятностное пространство, $B\subset\mathcal{U}$ и $\mathbf{P}(B)>0$. Рассмотрим $(\Omega,\mathcal{U},\mathbf{P}_B)$, где $\mathbf{P}_B(A)=\mathbf{P}(A|B)$. Пусть $\xi$ случайная величина на $(\Omega,\mathcal{U},\mathbf{P})$, а значит и на $(\Omega,\mathcal{U},\mathbf{P}_B)$.

**Определение 28.** Условным математическим ожиданием случайной величины $\xi$ относительно случайного события $B$, называется число, равное

$$M_B\xi = M(\xi|B) = \int_\Omega \xi(\omega)\mathbf{P}_B(d\omega)\,.$$

Имеем

$$M(\xi|B) = \int_\Omega \xi(\omega)\mathbf{P}(d\omega|B) = \frac{1}{\mathbf{P}(B)}\int_\Omega \xi(\omega)\mathbf{P}(d\omega\cap B) =$$

$$= \frac{1}{\mathbf{P}(B)}\int_B \xi(\omega)\mathbf{P}(d\omega) = \frac{M(\xi;B)}{\mathbf{P}(B)}\,,$$

где

$$M(\xi;B) = \int_B \xi(\omega)\mathbf{P}(d\omega)\,.$$

Пусть $\{B_n\}$ – разбиение $\Omega$. Тогда

$$M\xi = \sum_n \int_{B_n} \xi(\omega)\mathbf{P}(d\omega) = \sum_n M(\xi;B_n) = \sum_n \mathbf{P}(B_n)M(\xi|B_n)\,. \tag{2.1}$$

Эта формула является обобщением формулы полной вероятности для математического ожидания.

*Пример 22.* Рабочий обслуживает $n$ однотипных станков, расположенных прямолинейно на расстоянии 1 друг от друга. Рабочий обслуживает их подходя к ним в порядке очередности. Найти средний переход между станками.

Через $B_k$ обозначим случайное событие – рабочий находится у станка с номером $k$. Тогда $\mathbf{P}(B_k)=1/n$. Вероятность того, что от станка $k$ рабочий направляется к станку $i$, $i-1,\dots,n$, также равна $1/n$. Длина перехода при этом составляет $|i-k|$ и

$$M(\xi|B_k) = \sum_{i=1}^{n}\frac{1}{n}(k-i) + \sum_{i=1}^{n}\frac{1}{n}(i-k) = \frac{k(n-k)}{2n} + \frac{(n-k)(n-k+1)}{2n}$$

По формуле (2.1)

$$M\xi = \sum_{k=1}^{n}\mathbf{P}(B_k)M(\xi|B_k) = \sum_{k=1}^{n}\frac{1}{n^2}k(k-1)\,.$$

Так как

$$\sum_{k=1}^{n}k(k-1) = \frac{1}{3}(n-1)n(n+1)\,,$$

то

$$M\xi = \frac{n-1}{3}\Big(1+\frac{1}{n}\Big)\,.$$

Таким образом, если обслуживается $n$ станков, то средняя длина перехода рабочего составляет приблизительно $\frac{n-1}{3}$.

## 2.3. Характеристические функции

### Определение и простейшие свойства характеристических функций

Пусть $\zeta = \xi + i\eta$, где $\xi$ и $\eta$ действительные случайные величины, у которой существуют и конечны $M\xi$ и $M\eta$. Математическое ожидание комплексной случайной величины $\zeta$ определим соотношением

$$M\zeta = M\xi + iM\eta\,.$$

Основные свойства математического ожидания, включая свойства аддитивности и мультипликативности остаются в силе. Докажем неравенство

$$|M\zeta\,| \le M|\zeta\,|\,.$$

Существует действительное $\alpha$ такое, что

$$|M\zeta\,| = M\big(e^{i\alpha}\zeta\big)\,.$$

Пусть $e^{i\alpha}\zeta = \xi_1 + i\eta_1$, где $\xi_1$ и $\eta_1$ – действительные случайные величины. Имеем

$$|M\zeta\,| = M\xi_1 \le M|\xi_1\,| \le M|e^{i\alpha}\zeta\,| = M|\zeta\,|\,.$$

**Определение 29.** Характеристической функцией случайной величины $\xi$ называется функция $f_\xi(t)$ действительного переменного $t$, равная

$$f_\xi(t) = Me^{it\xi} = M\cos(\xi t) + iM\sin(\xi t) = \int\limits_{-\infty}^{\infty} e^{itx}dF_\xi(x)\,.$$

Если случайная величина $\xi$ дискретна, то

$$f_\xi(t) = \sum_k e^{itx_k}\mathbf{P}\{\xi = x_k\}\,.$$

Если случайная величина абсолютно непрерывна, то

$$f_\xi(t) = \int\limits_{-\infty}^{\infty} e^{itx}p_\xi(x)dx\,,$$

то есть характеристическая функция $f_\xi(t)$ в этом случае совпадает с преобразованием Фурье плотности $p_\xi(x)$ случайной величины $\xi$.

*Некоторые свойства характеристических функций:*

**1.** $|f_\xi(t)\,| \le 1$ для всех $t \in \mathbb{R}$ и $f_\xi(0) = 1$.

**2.** $f_\xi(t)$ равномерно непрерывна по $t$.

Предварительно докажем лемму.

**Лемма 2.2.** *При действительных $x$ и любом целом $n \ge 1$ имеет место равенство*

$$\Big|\, e^{ix} - \sum_{k=0}^{n-1}\frac{(ix)^k}{k!}\,\Big| \le \frac{|x\,|^n}{n!}$$

*Доказательство.* Доказательство по индукции. Имеем

$$|e^{ix}-1\,|=\Big|\int\limits_0^x e^{iu}du\,\Big|\le\int\limits_0^{|x\,|}du\le|x\,|\,.$$

Предположим, что неравенство верно при $m=n$. Докажем его справедливость при $m=n+1$. Воспользуемся соотношением

$$\int\limits_0^x\Big(e^{iu}-\sum_{k=0}^{n-1}\frac{(iu)^k}{k!}\Big)du=\frac{1}{i}\Big(e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\Big)\,.$$

Поэтому

$$\Big|\,e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\,\Big|=\Big|\,\int\limits_0^x\Big(e^{iu}-\sum_{k=0}^{n-1}\frac{(iu)^k}{k!}\Big)du\,\Big|\le\int\limits_0^{|x\,|}\frac{u^n}{n!}du=\frac{|x\,|^{n+1}}{(n+1)!}\,.\ \Box$$

Теперь нетрудно доказать свойство 2. Положим $A=\{|\xi\,|\le x\}$. Тогда

$$\begin{aligned}|f(t+h)-f(t)\,|=|M\,e^{it\xi}\big(e^{ih\xi}-1\big)|\le M\,|e^{ih\xi}-1|I_A+M\,|e^{ih\xi}-1|I_{\bar A}\le\\ \le 2M\,I_{\bar A}+|h\,|M|\xi\,|I_A\le x|h\,|+2\mathbf{P}\{|\xi\,|>x\}\,.\end{aligned}$$

Фиксируем $\varepsilon>0$ и выберем $x$ так, чтобы $\mathbf{P}\{|\xi\,|>x\}<\frac{\varepsilon}{4}$. Положим теперь $\delta=\frac{\varepsilon}{2x}$. Тогда для всех $h$, для которых $|h\,|<\delta$ выполняется неравенство $|f(t+h)-f(t)\,|<\varepsilon$. $\Box$

**3.** Если $\eta=a\,\xi+b$, где $a$ и $b$ действительные константы, то $f_\eta(t)=e^{itb}f_\xi(at)$.

**Определение 30.** Комплекснозначные случайные величины $\xi=\xi_1+i\xi_2$ и $\eta=\eta_1+i\eta_2$ независимы, если независимы $\sigma$-алгебры $\mathcal{A}_{\xi_1,\xi_2}$ и $\mathcal{A}_{\eta_1,\eta_2}$. Иначе, для любых борелевских $B_1,\ B_2\subset\mathbb{R}^2$ выполняется равенство

$$\mathbf{P}\{(\xi_1,\xi_2)\in B_1,\ (\eta_1,\eta_2)\in B_2\}=\mathbf{P}\{(\xi_1,\xi_2)\in B_1\}\,\mathbf{P}\{(\eta_1,\eta_2)\in B_2\}$$

Данное определение независимости случайных величин легко обобщается на случай конечного числа комплекснозначных случайных величин.

**4.** Если $\xi_1,\dots,\xi_n$ независимы, то

$$f_{\xi_1+\dots+\xi_n}(t)=\prod_{k=1}^n f_{\xi_k}(t)\,.$$

Доказывается с помощью мультипликативного свойства математического ожидания.

**5.** Обозначим через $m_n=M\xi^n$. Если $m_n$ конечно, то существуют все производные $f_\xi^{(k)}(t)$ порядка $k\le n$ и

$$f_\xi^{(k)}(0)=i^k\,M\xi^k\,. \tag{2.2}$$

Кроме того,имеет место разложение

$$f_\xi(t)=\sum_{k=0}^{n}\frac{(it)^k}{k!}m_k+R_n(t)\,, \tag{2.3}$$

где $R_n(t)=o\big(|t\,|^n\big)$ при $t\to 0$.

*Доказательство.* Продиффиренцировав формально характеристическую функцию $f_\xi(t)$ $k$ раз получим

$$f_\xi^{(k)}(t) = i^k M\xi^k e^{it\xi} = i^k \int\limits_{-\infty}^{\infty} x^k e^{itx} dF_\xi(x)\,. \tag{2.4}$$

Обосновать законность дифференцирования под знаком интеграла можно, рассуждая по индукции. Пусть соотношение (2.4) имеет место при $k < n$. Поскольку

$$\frac{f_\xi^{(k)}(t+h) - f_\xi^{(k)}(t)}{h} = i^k M\xi^k \frac{e^{it\xi}\left(e^{ih\xi} - 1\right)}{h}\,, \tag{2.5}$$

$$\left|\xi^k e^{it\xi}\frac{\left(e^{ih\xi}-1\right)}{h}\right| \le |\xi\,|^{k+1} \;\text{ и }\; M|\xi^{k+1}\,| < \infty\,,$$

то в правой части (2.5) по теореме Лебега о мажорируемой сходимости можно перейти к пределу при $h \to 0$. Полагая в (2.4) $t = 0$, приходим к (2.2). Разложение (2.3) следует из теоремы Тейлора. □

**6.** Имеет место равенство $f_\xi(-t) = \overline{f_\xi(t)}$.

**7.** Если $\xi \ge 0$, то $f_\xi(\lambda)$ определена при $\mathrm{Im}\lambda \ge 0$. При этом $|f_\xi(\lambda)| \le 1$, $f_\xi(\lambda)$, регулярна в $\mathrm{Im}\lambda > 0$ и непрерывна, включая границу $\mathrm{Im}\lambda = 0$,

$$f_\xi(\lambda) = \int\limits_0^{\infty} e^{i\lambda x} dF_\xi(x)\,.$$

Непрерывность доказывается как в свойстве 2. Регулярность следует из возможности дифференцирования под знаком интеграла.

Дискретную случайную величину $\xi$, принимающую только целые неотрицательнве значения, будем называть *целочисленной* случайной величиной. Закон распределения целочисленной случайной величины определяется вероятностями

$$p_n = \mathbf{P}\{\xi = n\},\;\; n = 0, 1, \ldots\,,$$

для которой $\sum_{n=0}^{\infty} p_n = 1$.

**Определение 31.** *Вероятностной производящей функцией* называется сумма ряда

$$\varphi_\xi(s) = \sum_{n=0}^{\infty} p_n s^n\,, \tag{2.6}$$

равная $Ms^\xi$. Ряд в (2.6) абсолютно сходится при $|s\,| \le 1$.

**8.** Если $\varphi_\xi(s) = Ms^\xi$ – производящая функция целочисленной случайной величины, то

$$f_\xi(t) = \varphi_\xi\left(e^{it}\right).$$

Здесь же укажем еще одно свойство производящих функций Пусть $\xi_1, \xi_2, \ldots$ – последовательность целочисленных независимых одинаково распределенных случайных величин с производящей функции $\varphi_\xi(s)$ и $\nu$ – независимая от них целочисленная случайная величина с производящей функцией $\varphi_\nu(s)$. Через $\zeta_\nu$ обозначим сумму случайного числа случайных величин:

$$\zeta_\nu = \xi_1 + \cdots + \xi_\nu,\;\; \nu \ge 1 \;\text{ и }\; \zeta_0 = 0\,.$$

**Теорема 2.24.** *Производящая функция $\varphi_{\zeta_\nu}$ случайной величины $\zeta_\nu$ равна суперпозиции*

$$\varphi_{\zeta_\nu}(s) = \varphi_\nu\big(\varphi_\xi(s)\big)\,.$$

*Доказательство.* Имеем

$$\varphi_{\zeta_\nu}(s) = Ms^{\zeta_\nu} = M\big[Ms^{\zeta_\nu}|\nu\big] = M\big[\varphi_\xi(s)\big]^\nu = \varphi_\nu\big(\varphi_\xi(s)\big)\,,$$

так как

$$M\{s^{\xi_1+\dots+\xi_\nu}\mid \nu = n\} = M\,s^{\xi_1+\dots+\xi_n} = \big[\varphi_\xi(s)\big]^n. \ \Box$$

Рассмотрим несколько примеров.

*Пример 23.* Если $\mathbf{P}\{\xi = c\} = 1$, то $f_\xi(t) = e^{itc}$.

*Пример 24.* Для нормально распределенной случайной величины $\xi$ с параметрами $(0,1)$ имеем

$$f_\xi(t) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{itx}e^{-\frac{x^2}{2}}dx = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}dx = I$$

Положим $z = x - it$ и продолжим выкладку, сославшись на теорему Коши,

$$I = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\int\limits_{-\infty-it}^{\infty-it} e^{-\frac{z^2}{2}}dz = e^{-\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{u^2}{2}}du = e^{-\frac{t^2}{2}}\,.$$

Пусть теперь случайная величина $\eta$ нормально распределена с параметрами $(a,\sigma)$. Тогда случайная величина $\xi = (\eta - a)/\sigma$ нормально распределена с параметрами $(0,1)$. Поэтому

$$f_\eta(t) = e^{-ita}e^{-\frac{(t\sigma)^2}{2}} = e^{ita-\frac{t^2\sigma^2}{2}}\,.$$

*Пример 25.* Пусть $\xi$ равномерно распределена на отрезке $[0,1]$. Тогда

$$f_\xi(t) = \int\limits_0^1 e^{itx}dx = \frac{e^{it}-1}{it}\,.$$

*Пример 26.* Для случайной величины $\xi$, распределенной по Пуассону с параметром $\mu$ имеем

$$\varphi_\xi(s) = \sum_{k=0}^{\infty} s^k\frac{\mu^k}{k!}e^{-\mu} = e^{-\mu}e^{s\mu} = e^{[\mu(s-1)]}\,.$$

*Пример 27.* Для биномиально распределенной случайной величины $\xi$ с законом распределения $\mathbf{P}\{\xi = m\} = C_n^m p^m(1-p)^{n-m}$ имеем

$$\varphi_\xi(s) = \sum_{m=0}^{n} C_n^m p^m(1-p)^{n-m}s^m = (ps+1-p)^n$$

### Формула обращения

Далеко не всякая функция $f(t)$, $t \in \mathbb{R}$, может быть характеристической функцией некоторой случайной величины.

**Определение 32.** Непрерывная функция $f(t)$ вещественного аргумента $t$ называется положительно определенной на $\mathbb{R}$, если каковы бы ни были вещественные $t_1,\dots,t_n$ и комплексные $\zeta_1,\dots,\zeta_n$ выполняется неравенство

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f(t_j-t_k)\zeta_j\overline{\zeta_k} \ge 0\,.$$

**Теорема 2.25** (Бохнер–Хинчин). *Для того, чтобы непрерывная функция $f(t)$ такая, что $f(0)=1$, была характеристической необходимо и достаточно, чтобы она была положительно определенной.*

Необходимость положительной определенности проверяется несложно:

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f(t_j-t_k)\zeta_j\overline{\zeta_k} = \sum_{k=1}^{n}\sum_{j=1}^{n}\Big\{\int_{\mathbb{R}} e^{ix(t_j-t_k)}dF(x)\Big\}\zeta_j\overline{\zeta_k} =$$

$$= \int_{\mathbb{R}} \sum_{k=1}^{n}\sum_{j=1}^{n} e^{ix(t_j-t_k)}\zeta_j\overline{\zeta_k}dF(x) = \int_{\mathbb{R}} \Big(\sum_{k=1}^{n} e^{ixt_k}\zeta_k\Big)\overline{\Big(\sum_{j=1}^{n} e^{ixt_j}\zeta_j\Big)}dF(x) =$$

$$= \int_{\mathbb{R}} \Big|\sum_{k=1}^{n} e^{ixt_k}\zeta_k\Big|^2 dF(x) \ge 0. \quad (2.7)$$

Сформулируем здесь же теорему Герглотца, дискретный вариант теоремы Бохнера-Хинчина.

**Теорема 2.26.** *Если последовательность $c_n$, $n=0,\pm1,\dots$, обладает тем свойством, что при любом выборе комплексных чисел $\zeta_1,\dots,\zeta_N$ и произвольном $N$*

$$\sum_{k=1}^{N}\sum_{j=1}^{N} c_{j-k}\zeta_j\overline{\zeta_k} \ge 0$$

*то последовательность $c_n$ может быть записана в форме*

$$c_n = \int_{-\pi}^{\pi} e^{inx}d\sigma(x)\,,$$

*где $\sigma(x)$ – неубывающая функция с ограниченной вариацией.*

Если случайная величина $\xi$ имеет плотность $p_\xi(x)$, то характеристическая функция $f_\xi(t)$, является преобразованием Фурье плотности $p_\xi(x)$. Если предположить, что интеграл $\int_{-\infty}^{\infty}|f_\xi(t)|dt$ конечен, то имеет место формула обращения для преобразования Фурье

$$p_\xi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}f_\xi(t)dt.$$

Докажем формулу обращения в общем случае.

**Теорема 2.27** (формула обращения). *Если $F(x)$ – функция распределения случайной величины $\xi$, а $f(t)$ – характеристическая функция для $\xi$, то для любых точек $x,y$ непрерывности $F(x)$ имеет место равенство*

$$F(y)-F(x) = \frac{1}{2\pi}\lim_{A\to\infty}\int_{-A}^{A}\frac{e^{-itx}-e^{-ity}}{it}f(t)dt\,.$$

*Доказательство.* Для определенности будем считать, что $x < y$. Имеем

$$I_A = \frac{1}{2\pi}\int\limits_{-A}^{A} \frac{e^{-itx} - e^{-ity}}{it} f(t)dt = \frac{1}{2\pi}\int\limits_{-A}^{A}\int\limits_{-\infty}^{\infty} \frac{e^{it(u-x)} - e^{it(u-y)}}{it} dF(u)dt\,.$$

Подинтегральная функция в повторном интеграле ограничена. Поэтому по теореме Фубини можно переставить местами интегралы:

$$I_A = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\Big[\int\limits_{-A}^{A} \frac{e^{it(u-x)} - e^{it(u-y)}}{it}dt\Big]dF(u) =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\Big[\int\limits_{0}^{A} \frac{e^{it(u-x)} - e^{-it(u-x)} - e^{it(u-y)} + e^{-it(u-y)}}{it}dt\Big]dF(u) =$$

$$= \frac{1}{\pi}\int\limits_{-\infty}^{\infty}\Big[\int\limits_{0}^{A}\Big(\frac{\sin(u-x)t}{t} - \frac{\sin(u-y)t}{t}\Big)dt\Big]dF(u) =$$

$$= \frac{1}{\pi}\int\limits_{-\infty}^{\infty}\Big[\int\limits_{A(u-y)}^{A(u-x)} \frac{\sin t}{t}dt\Big]dF(u)\,.$$

Из курса анализа известно, что

$$\lim_{\substack{A\to\infty\\ B\to\infty}} \frac{1}{\pi}\int\limits_{-B}^{A} \frac{sint}{t}dt = 1\,.$$

Отсюда следует, что функция

$$S_A(u) = \frac{1}{\pi}\int\limits_{A(u-y)}^{A(u-x)} \frac{\sin t}{t}dt$$

ограничена. Поэтому

$$\lim_{A\to\infty} I_A = \int\limits_{-\infty}^{\infty} \lim_{A\to\infty} S_A(u)dF(u)\,.$$

Так как

$$\lim_{A\to\infty} S_A(u) = \begin{cases} 1, & x < u < y \\ \dfrac{1}{2}, & u = x \text{ или } u = y \\ 0, & u < x \text{ или } u > y\,, \end{cases}$$

получаем

$$\lim_{A\to\infty} I_A = \int\limits_{x}^{y} dF(u) = F(y) - F(x)\,. \ \square$$

**Следствие.** *Характеристическая функция случайной величины однозначно определяет функцию распределения этой случайной величины.*

Приведем несколько полезных примеров применения доказанной теоремы.

*Пример 28.* Пусть $\xi_1$ и $\xi_2$ независимые случайные величины, распределенные по нормальному закону с параметрами $(a_1, \sigma_1)$ и $(a_2, \sigma_2)$. Тогда случайная величина $\xi_1 + \xi_2$ также нормально распределена с параметрами $(a_1 + a_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

Действительно,

$$f_{\xi_1+\xi_2}(t) = f_{\xi_1}(t)+f_{\xi_2}(t) = \\ = e^{ita_1 - \frac{t^2\sigma_1^2}{2}} e^{ita_2 - \frac{t^2\sigma_2^2}{2}} = e^{it(a_1+a_2) - \frac{t^2(\sigma_1^2+\sigma_2^2)}{2}}$$

*Пример 29.* Сумма двух биномиально распределенных независимых случайных величин $\xi_1$ и $\xi_2$ с одинаковым параметром $p$ является биномиально распределенной случайной величиной.

В самом деле,

$$\varphi_{\xi_1+\xi_2}(s) = \varphi_{\xi_1}\varphi_{\xi_2} = (ps+q)^{n_1}(ps+q)^{n_2} = (ps+q)^{n_1+n_2} .$$

*Пример 30.* Сумма двух независимых пуассоновских случайных величин $\xi_1$ и $\xi_2$ является пуасоновской случайной величиной.

Имеем

$$\varphi_{\xi_1+\xi_2}(s) = \varphi_{\xi_1}\varphi_{\xi_2} = e^{\lambda_1(s-1)}e^{\lambda_2(s-1)} = e^{(\lambda_1+\lambda_2)(s-1)} .$$

Как мы видели раньше, если $\xi_1$ и $\xi_2$, распределены равномерно на $[0, 1]$ и независимы, то $F_{\xi_1}$ и $F_{\xi_1+\xi_2}$ существенно разные функции.

### Теорема непрерывности

Таким образом, между множеством функций распределения и множеством их характеристических функций имеется взаимно однозначное соответствие. Ограничимся здесь формулировкой теоремы, их которой следует, что это соответствие не только взаимно однозначное, но и взаимно непрерывное.

**Теорема 2.28** (теорема непрерывности). *Пусть $f_n(t)$ есть характеристическая функция случайной величины $\xi_n$ и $F_{\xi_n}(x) = F_n(x)$. Если при $n \to \infty$ $f_n(t) \to f(t)$ в каждой точке $t \in \mathbb{R}$ и $f(t)$ непрерывна при $t = 0$, то*

*1. $f(t)$ является характеристической функцией некоторой функции распределения $F(x)$;*

*2. $F_n(x) \to F(x)$ в каждой точке непрерывности предельной функции.*

*Обратно, если $F_n(x) \to F(x)$ в каждой точке непрерывности функции $F(x)$ и $F(x)$ является функцией распределения, то $f_n(t) \to f(t)$ в каждой точке $t \in \mathbb{R}$, при этом $f(t)$ есть характеристическая функция для функции распределения $F(x)$.*

Для целочисленных случайных величин теорема непрерывности формулируется следующим образом:

**Теорема 2.29.** *Пусть*

$$\varphi_r(s) = \sum_{n=0}^{\infty} p_n^{(r)} s^n , \quad r = 1, 2, \dots ,$$

*есть последовательность вероятностных производящих функций и*

$$\varphi(s) = \sum_{n=0}^{\infty} p_n s^n ,$$

*является производящей функцией последовательности $\{p_n\}$. Для того, чтобы при каждом $n$*

$$\lim_{r\to\infty} p_n^{(r)} = p_n ,$$

*необходимо и достаточно, чтобы при всех* $s \in [0,1)$

$$\lim_{r\to\infty} \varphi_r(s) = \varphi(s)\,.$$

*Пример 31.* Воспользуемся теоремой непрерывности для доказательства предельной теоремы Пуассона.

Пусть $a_n = np_n \to a$. Тогда

$$\big(sp_n + 1 - p_n\big)^n = \Big(s\frac{a_n}{n} + 1 - \frac{a_n}{n}\Big)^n = \Big(1 + \frac{a_n}{n}(s-1)\Big)^n \to e^{a(s-1)}\,.$$

Поэтому

$$\lim_{n\to\infty} C_n^m \Big(\frac{a_n}{n}\Big)^m \Big(1 - \frac{a_n}{n}\Big)^{n-m} = \frac{a^m}{m!}e^{-a}\,.$$

## 2.4. Центральная предельная теорема

Ранее была доказана интегральная теорема Муавра - Лапласа утверждающая, что распределение числа успехов $\mu$ в схеме Бернулли при $n \to \infty$ и постоянном $0 < p < 1$ обладает следующим предельным свойством

$$\lim_{n\to\infty} \mathbf{P}\Big\{\frac{\mu - M\mu}{\sqrt{D\mu}} \le x\Big\} \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{u^2}{2}}du\,.$$

Предельной функцией является функция нормального распределения вида

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{u^2}{2}}du\,,$$

которая связана с интегралом Лапласа

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-\frac{u^2}{2}}du$$

соотношением $\Phi(x) = \frac{1}{2} + \Phi_0(x)$.

Этот результат является частным случаем общей *центральной предельной теоремы.* Пусть $\xi_n$, $n = 1,\ldots,$ независимы, $M\xi_n = a$, $D\xi_n = \sigma^2$. Положим $S_n = \xi_1 + \cdots + \xi_n$ и введем

$$\zeta_n = \frac{S_n - an}{\sigma\sqrt{n}} = \frac{\frac{S_n}{n} - a}{\sigma\sqrt{n}}\,.$$

**Теорема 2.30.** *Если* $0 < \sigma^2 < \infty$, *то при* $n \to \infty$

$$\mathbf{P}\{\zeta_n \le x\} \to \Phi(x)$$

*равномерно относительно* $x \in (-\infty, \infty)$.

*Доказательство.* Можем считать, что $M\xi_k = 0$, иначе можно было бы рассмотреть последовательность $\{\xi_n' = \xi_n - a\}_{n=1}^{\infty}$, при этом последовательность $\{\zeta_n\}$ не изменилась бы. Согласно теореме непрерывности достаточно доказать, что

$$f_{\zeta_n}(t) \to e^{-\frac{t^2}{2}},\ \ t \in (-\infty, \infty)\,.$$

Положим $f_\xi(t) = f(t)$. Имеем

$$f_{\zeta_n}(t) = f^n\Big(\frac{t}{\sigma\sqrt{n}}\Big)(t).$$

Так как

$$f(t) = f(0) + tf'(0) + \frac{t^2}{2}f''(0) + o(t^2) = 1 - \frac{t^2\sigma^2}{2} + o(t^2)\,,$$

поэтому

$$\begin{aligned}\log f_{\zeta_n}(t) =& n\log\Big[1 - \frac{\sigma^2}{2}\Big(\frac{t}{\sigma\sqrt{n}}\Big)^2 + o\Big(\frac{t^2}{n}\Big)\Big] \sim \\ \sim& n\Big(-\frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)\Big) = -\frac{t^2}{2} + o(t^2) \to -\frac{t^2}{2}\,.\end{aligned}$$

Если последовательность функций распределения $F_n(x)$ сходится к непрерывной функции распределения $F(x)$, то сходимость $F_n(x)$ к $F(x)$ равномерна. Достаточно показать, что у каждой точки расширенной числовой прямой $\overline{\mathbb{R}}$ имеется окрестность, в которой сходимость равномерна. Пусть $-\infty < x < y < \infty$ и $\varepsilon > 0$. Найдется $n_0$ такое, что для всех $n \geqslant n_0$ выполняются неравенства

$$\begin{aligned}F_n(x) - \varepsilon < F(x) < F_n(x) + \varepsilon\,,\\ F_n(y) - \varepsilon < F(y) < F_n(y) + \varepsilon\,.\end{aligned}$$

Тогда для всех $t \in [x, y]$ имеем

$$\begin{aligned}F(x) - \varepsilon < F(x) \leqslant F(t) \leqslant F(y) < F(y) + \varepsilon\,,\\ F(x) - \varepsilon < F_n(x) \leqslant F_n(t) \leqslant F_n(y) < F(y) + \varepsilon\,.\end{aligned}$$

Отсюда следует, что для всех $n \geqslant n_0$ и $t \in [x, y]$

$$|F(t) - F_n(t)| < 2\varepsilon\,.$$

Для бесконечно удаленной точки $x = -\infty$ имеем

$$\begin{aligned}0 \leqslant F(t) \leqslant F(y) < F(y) + \varepsilon\,,\\ 0 \leqslant F_n(t) \leqslant F_n(y) < F(y) + \varepsilon\,.\end{aligned}$$

Поэтому для всех $t \in (-\infty, y]$ при $n \geqslant n_0$ верно неравенство

$$|F(t) - F_n(t)| < \varepsilon\,.$$

Аналогично рассматривается случай другой бесконечно удаленной точки $y = \infty$. □

Центральная предельная теорема имеет место также при некоторых условиях и для произвольной последовательности независимых случайных величин. Сформулируем одну из подобного рода теорем.

Пусть теперь $\xi_k$, $k = 1, \ldots$, независимы и имеют не обязательно одно и то же распределение. Введем обозначения

$$M\xi_k = a_k,\;\; D\xi_k = b_k^2,\;\; M|\xi_k - a_k|^3 = c_k^3$$

и

$$A_k = \sum_{k=1}^{n} a_k,\;\; B_n^2 = \sum_{k=1}^{n} b_k^2,\;\; C_n^3 = \sum_{k=1}^{n} c_k^3.$$

**Теорема 2.31** (Теорема Ляпунова). *Если случайные величины $\xi_1, \xi_2, \ldots$ независимы, $a_k$, $b_k$, $c_k$ конечны и $C_n/B_n \to 0$, то*

$$\lim_{n\to\infty} \mathbf{P}\Big\{\frac{\xi_1 + \cdots + \xi_n - A_n}{B_n} \le x\Big\} = \Phi(x)\,.$$

Хорошо известной иллюстрацией сформулированной теоремы является рассеивание артиллерийских снарядов при стрельбе по цели. На траекторию полета снаряда действует большое количество независимых факторов, причем влияние каждого фактора само по себе невелико. Это отклонения в количестве заряда, отклонения в весе и размерах снаряда, отклонения во влажности и температуре воздуха, направление и сила ветра на разных высотах и т.д. В результате, отклонение снаряда от точки прицеливания удивительно точно описывается нормальным законом.

Нужно не забывать, что центральная предельная теорема не доказывает нормальность рассеяния снаряда, а дает только гипотезе нормальности некоторое теоретическое обоснование. Ведь математические теории строятся не на реальных явлениях, а на математических моделях. Поэтому в применении теории вероятностей всегда следует заботиться о том, чтобы рассматриваемая модель правильно отражала соответствующее явление.

*Пример 32.* (Ошибки измерения). При измерении некоторой величины $a$ мы получаем приближенное значение $\xi$. Сделанная ошибка $\delta = \xi - a$ может быть представлена в виде суммы двух ошибок

$$\delta = (\xi - M\xi) + (M\xi - a)\,,$$

первая из которых $\xi - M\xi$ называется случайной ошибкой, а вторая $M\xi - a$ – систематической ошибкой. Хорошие методы измерения не должны иметь систематической ошибки. Будем полагать $M\xi = a$. Случайная ошибка $\delta$ имеет нулевое математическое ожидание, $M\delta = 0$. Пусть $D\delta = \sigma^2$. Для уменьшения этой ошибки производят $n$ независимых измерений $\xi_1, \ldots, \xi_n$ и принимают за оценку измеримой величины $a$ среднее арифметическое

$$\widehat{a} = \frac{\xi_1 + \cdots + \xi_n}{n}\,.$$

Допускаемую при этом погрешность можно оценить с помощью центральной предельной теоремы. Согласно этой теореме сумма $\xi_1 + \cdots + \xi_n$ одинаково распределенных независимых случайных величин с $M\xi_n = a$ и $D\xi_n = \sigma^2 > 0$ асимптотически нормально распределена с параметрами $(an, \sigma\sqrt{n})$. Поэтому $\widehat{a}$ при больших $n$ асимптотически нормальна с параметрами $(a, \sigma/\sqrt{n})$ и

$$\mathbf{P}\{|\widehat{a} - a| \le \varepsilon\} \approx \frac{1}{\sqrt{2\pi}} \int\limits_{-\varepsilon\sqrt{n}/\sigma}^{\varepsilon\sqrt{n}/\sigma} e^{-\frac{x^2}{2}} du$$

Дисперсия $D(\widehat{a} - a)$ погрешности $\widehat{a} - a$ уменьшится в сравнении с дисперсией $\xi$ и станет равной

$$D(\widehat{a} - a) = \frac{D\xi_1 + \cdots + D\xi_n}{n^2} = \frac{\sigma^2}{n}\,.$$

## 2.5. Многомерные характеристические функции

### Определение и свойства

Пусть $\xi = (\xi_1, \ldots, \xi_n)$ — случайный вектор и $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$.

**Определение 33.** Характеристическая функция $f_\xi(t)$ случайного вектора $\xi$ определяется соотношением

$$f_\xi(t) = Me^{i(t,\xi)} = Me^{i\sum_{k=1}^n t_k\xi_k}\,.$$

Перечислим свойства характеристических функций.

**1.** При всех $t \in \mathbb{R}^n$ имеем $|f_\xi(t)| \le 1$, $f_\xi(0) = 1$.

**2.** $f_\xi(-t) = \overline{f_\xi(t)}$.

**3.** $f_\xi(t)$ равномерно непрерывна по $t$.

*Доказательство.* Обозначим через $A$ событие $A = \{|\xi_k| \le x, \ \ k = 1, \dots, n\}$ и пусть $h \in \mathbb{R}^n$. Оценим приращение

$$|f_\xi(t+h) - f_\xi(t)| = |Me^{i(t,\xi)}\big(e^{i(h,\xi)} - 1\big)| \le M|e^{i(h,\xi)} - 1| =$$
$$= M|e^{i(h,\xi)} - 1|I_A + M|e^{i(h,\xi)} - 1|I_{\bar A} \le M|(h,\xi)|I_A + 2MI_{\bar A} \le$$
$$\le x|h| + 2\mathbf{P}\{\xi \notin [-x,x]^n\},$$

где $|h| = \sum_{k=1}^n |h_k|$ и $[-x,x]^n$ – прямоугольник в $\mathbb{R}^n$:

$$\{t = (t_1, \dots, t_n) \in \mathbb{R}^n \ : \ |t_k| \le x, \ \ k = 1, \dots, n\}\,.$$

Пусть дано $\varepsilon > 0$. Выберем $x$ так, чтобы выполнялось неравенство $\mathbf{P}\{\xi \notin [-x,x]^n\} < \varepsilon/4$. Тогда при $|h| < \varepsilon/(2x)$ имеем $|f_\xi(t+h) - f_\xi(t)| < \varepsilon$. $\square$

**4.** Если $\xi_1, \dots, \xi_n$ – независимые случайные величины и $\zeta = \xi_1 + \dots + \xi_n$, то

$$f_\zeta(t) = \prod_{k=1}^n f_{\xi_k}(t)\,.$$

**5.** Если известна характеристическая функция $f_\xi(t)$ случайного вектора $\xi = (\xi_1, \dots, \xi_n)$, то характеристическая функция случайного вектора $(\xi_{k_1}, \dots, \xi_{k_j})$ можно получить, положив все $t_k$, за исключением $t_{k_1}, \dots, t_{k_j}$, равными нулю.

**6.** Если $\eta = \xi C$ – линейное преобразование с матрицей $C = \| c_{\alpha\beta} \|$, $\alpha = 1, \dots, m$; $\beta = 1, \dots, k$, то

$$f_\eta(t) = f_\xi(tC')\,,$$

где $C'$ – сопряженная с $C$ матрица.

*Доказательство.*

$$f_\eta(t) = Me^{i(t,\eta)} = Me^{i(t,\xi C)} = Me^{it(C'\xi\prime)} = Me^{i(tC')\xi'} = Me^{i(tC',\xi)} = f_\xi(tC')\,,$$

здесь $\xi'$ обозначает вектор-столбец. $\square$

*Пример 33.* Если $m = k$, определитель $\det C$ матрицы $C$ отличен от нуля и случайная величина $\xi$ имеет плотность $p_\xi(x)$, то $\eta = \xi C$ имеет плотность

$$p_\eta(y) = \frac{1}{\det C}\, p_\xi\big(yC^{-1}\big)\,.$$

*Доказательство.* В самом деле, для любого $A \in \mathcal{B}^k$ имеем

$$\mathbf{P}\{\xi \in A\} = \int\limits_A p_\xi(x)dx = \int\limits_{\{y=xC;x\in A\}} p_\xi\big(yC^{-1}\big)\frac{1}{|\det C\,|}dy =$$
$$= \int\limits_{\{y=xC;x\in A\}} p_\eta(y)dy = \mathbf{P}\{\eta \in \{xC; x \in A\}\}\,. \ \square$$

*Пример 34.* Если $\eta = \xi C + b$ – аффинное преобразование, то

$$f_\eta(t) = e^{i(t,b)} f_\xi(tC')\,.$$

**7.** Если конечен момент $M(|\xi_1| + \dots + |\xi_n|)^k$, то $f_\xi(t)$ имеет все производные порядка $k$ и для всех натуральных $k_1, \dots, k_n$, для которых $k = k_1 + \dots + k_n$, выполняются соотношения

$$\frac{\partial^k f_\xi(0, \dots, 0)}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} = i^k\, M\xi_1^{k_1} \dots \xi_n^{k_n}\,.$$

и

$$f_\xi(t) = \sum_{\alpha=0}^{k} i^\alpha \sum_{\alpha_1+\dots+\alpha_n=\alpha} \frac{t_1^{\alpha_1}\dots t_n^{\alpha_n}}{\alpha_1!\dots\alpha_n!} M\xi_1^{k_1}\dots\xi_n^{k_n} + o\big(|t|^k\big)$$

при $|t_1|+\dots+|t_k|\to 0$.
Свойство 7 доказывается так же, как и в одномерном случае.

**Определение 34.** Пусть $\zeta = (\xi_1,\dots,\xi_r)$ – случайный вектор с целочисленными случайными компонентами и $p_\alpha = \mathbf{P}\{\zeta=\alpha\}$, где $\alpha = (\alpha_1,\dots,\alpha_r)$. Многомерная производящая функция $\varphi_\zeta(s_1,\dots,s_r)$ определяется соотношением

$$\varphi_\zeta(s_1,\dots,s_r) = Ms_1^{\xi_1}\dots s_r^{\xi_r} = \sum_\alpha p_\alpha s_1^{\alpha_1}\dots s_r^{\alpha_r}\,.$$

Характеристическая функция $f_\zeta(t_1,\dots,t_r)$ и производящая функция $\varphi_\zeta(s_1,\dots,s_r)$ связаны соотношением

$$f_\zeta(t_1,\dots,t_r) = \varphi_\zeta(e^{it_1},\dots,e^{it_r})\,.$$

**Определение 35.** Случайный вектор $\zeta = (\xi_1,\dots,\xi_r)$ называется полиномиально распределенным, если

$$\mathbf{P}\{\xi=\alpha\} = \frac{n!}{\alpha_1!\dots\alpha_r!}p_1^{\alpha_1}\dots p_r^{\alpha_r},\ \ p_1+\dots+p_r=1\,.$$

Для полиномиального распределения

$$\varphi_\zeta(s_1,\dots,s_r) = \big(p_1s_1+\dots+p_rs_r\big)^n\,.$$

*Пример 35.* Пусть $\xi_1,\dots,\xi_k$ – нормально распределенные независимые случайные величины с $M\xi_\alpha = a_\alpha$ и $D\xi_\alpha = b_\alpha$. Тогда

$$f_{\xi_1\dots\xi_k}(t_1,\dots,t_k) = e^{i\sum\limits_{\alpha=1}^{k} t_\alpha a_\alpha - \frac{1}{2}\sum\limits_{\alpha=1}^{k} b_\alpha t_\alpha^2}\,.$$

Пусть случайный вектор $\xi = (\xi_1,\dots,\xi_n)$ имеет непрерывную плотность $p_\xi(x)$ и суммируемую характеристическую функцию $f_\xi(t)\in L_1$. Тогда плотность $p_\xi(x)$ восстанавливается по характеристической функции $f_\xi(t)$ однозначно при помощи обратного преобразования Фурье

$$p_\xi(x) = \frac{1}{(2\pi)^n}\int\limits_{\mathbb{R}^n} e^{-i(t,x)}f_\xi(t)dt\,.$$

Формула обращения в общем случае является обобщением своего одномерного аналога.

**Теорема 2.32.** *Пусть $\Delta = \{x : a_k < x_k < b_k\}$, $k=1,\dots,n$, и функция распределения $F_\xi(x)$ непрерывна на гранях $\Delta$, тогда*

$$\mathbf{P}\{\xi\in\Delta\} = \lim_{T\to\infty}\frac{1}{(2\pi)^n}\int\limits_{-T}^{T}\cdots\int\limits_{-T}^{T}\prod_{k=1}^{n}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}f_\xi(t)dt_1\dots dt_n\,.$$

Следствием этой теоремы является утверждение о взаимнооднозначном соответствии между характеристическими функциями и функциями распределения.

Приведем здесь еще один критерий независимости случайных величин.

**8.** Для независимости $\xi_1,\dots,\xi_n$ необходимо и достаточно, чтобы

$$f_{\xi_1,\dots,\xi_n}(t_1,\dots,t_n) = \prod_{k=1}^{n} f_{\xi_k}(t_k)\,.$$

Необходимость следует из мультипликативного свойства математического ожидания. Достаточность следует из доказанной формулы обращения (см. теорему 27).

Как и в одномерном случае приведем только формулировку *теоремы непрерывности.*

**Теорема 2.33.** *Для того, чтобы последовательность многомерных функций распределения $F^{(n)}(x)$ сходилась к некоторой функции распределения $F(x)$ в каждой точке непрерывности предельной функции, необходимо и достаточно, чтобы при любом $t$ последовательность соответствующих характеристических функций $f^{(n)}(t)$ сходилась к пределу $f(t)$, непрерывному в точке $t = 0$. При этом $f(t)$ есть характеристическая функция распределения $F(x)$.*

## 2.6. Многомерное нормальное распределение

### Невырожденное нормальное распределение

**Определение 36.** Определим невырожденное центрированное многомерное нормальное распределение случайного вектора $\xi = (\xi_1, \ldots, \xi_n)$ как распределение с плотностью

$$p_\xi(x) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} e^{-\frac{1}{2}Q(x)}, \quad x \in \mathbb{R}^n$$

где $Q(x) = x\,A\,x' = \sum\limits_{i,j=1}^{n} a_{ij}x_ix_j$, $|A\,|$ – определитель положительно определенной матрицы $\|\, a_{i,j} \,\|$. Здесь $x'$ обозначает вектор-столбец. Для центрированного распределения ниже будет доказано, что $M\xi = 0$ . Распределение случайного вектора $\xi + a$, где $a \in \mathbb{R}^n$ – постоянный вектор, также будем называть невырожденным нормальным распределением.

Найдем характеристическую функцию случайного вектора $\xi$. Покажем, что

$$f_\xi(t) = e^{-\frac{1}{2}t\,\overline{M}\,t'}\,,$$

где $\overline{M} = A^{-1}$ есть матрица, обратная к $A$, совпадающая с матрицей $\|\, m_{ij} \,\|$ вторых моментов распределения $\xi$:

$$m_{ij} = M\,\xi_i\xi_j\,.$$

В самом деле,

$$f_\xi(t) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} e^{itx' - \frac{1}{2}x\,A\,x'} dx_1 \ldots dx_n$$

Выберем ортогональную матрицу $C$ так, чтобы $C\,A\,C'$ была бы диагональной матрицей, и обозначим через $\mu_1, \ldots, \mu_n$ значения диагональных элементов. Сделаем замену переменных, положив $x = y\,C$ и $t = v\,C$. Тогда

$$|A\,| = |E\,| = \prod_{k=1}^{n} \mu_k\,,$$

$$itx' - \frac{1}{2}x\,A\,x' = ivy' - \frac{1}{2}y\,E\,y' = \\ = i\sum_{k=1}^{n} v_k y_k - \frac{1}{2}\sum_{k=1}^{n} \mu_k y_k^2$$

Поэтому

$$f_\xi(t) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \prod_{k=1}^{n} \int\limits_{-\infty}^{\infty} e^{iv_k y_k - \frac{1}{2}\mu_k y_k^2} dy_k = \\ = \sqrt{|A|} \prod_{k=1}^{n} e^{-\frac{1}{2}\frac{v_k^2}{\mu_k}} = e^{-\frac{1}{2}v\,E^{-1}\,v'} = \\ = e^{-\frac{1}{2}t\,C'E^{-1}C\,t'} = e^{-\frac{1}{2}t\,A^{-1}\,t'}\,.$$

С другой стороны, так как все моменты случайного вектора $\xi$ существуют, то в окрестности точки $t=0$ имеет место равенство

$$f_\xi(t) = 1 - \frac{1}{2} t A^{-1} t' + o\big(|t|^2\big) =$$

$$= 1 + it\,(M\xi)' + \frac{i^2}{2} t\,\overline{M}\,t' + 0\big(|t|^2\big)\,.$$

Отсюда следует, что $M\xi = 0$, $A^{-1} = \overline{M}$.

Из доказанного вытекает очень важное свойство нормальных распределений: *компоненты нормально распределенного вектора $\xi = (\xi_1, \ldots, \xi_n)$ независимы тогда и только тогда, когда коэффициенты корреляции $\rho(\xi_i, \xi_j) = 0$ при всех $i \neq j$.*

*Доказательство.* Действительно, если $\overline{M}$ – диагональная матрица, то матрица $A = \overline{M}^{-1}$ также диагональная и $f_\xi(x)$ равна произведению характеристических функций компонент вектора $\xi$. Обратно, если $\xi_1, \ldots, \xi_n$ независимы, то $A$ диагональна и, значит, диагональна матрица $M$. □

Пусть $\xi = (\xi_1, \ldots, \xi_n)$ невырожденный центрированный нормально распределенный случайный вектор с плотностью $p_\xi(x)$

$$p_\xi(x) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} e^{-\frac{1}{2} x A x'}, \quad x \in \mathbb{R}^n\,.$$

Выберем ортогональную матрицу $B$ такую, что $B'AB$ является диагональной матрицей. В этом случае, как мы видели, компоненты случайного вектора $\zeta = \xi B$ нормально распределены с нулевым математическим ожиданием и независимы. Поэтому $\xi = \zeta B'$ и, легко видеть, найдется невырожденная $(n \times n)$-матрица $C$ и случайный вектор $\eta$, компоненты которого независимы и нормально распределены с параметрами $(0, 1)$, такие, что $\xi \;=\; \eta C$. В общем случае многомерный нормально распределенный случайный вектор $\xi$ имеет представление

$$\xi = \eta C + a,$$

где компоненты случайного вектора $\eta$ нормально распределены с параметрами $(0, 1)$ и независимы, $C$ – невырожденная $(n \times n)$-матрица и $a$ – постоянный вектор, равный $M\xi$.

**Определение 37.** Распределение $\xi = (\xi_1, \ldots, \xi_n)$ с плотностью

$$p_\xi(x) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2} \sum\limits_{k=1}^{n} x_k^2}$$

называется сферически нормальным распределением. Характеристическая функция сферически нормального распределения имеет вид

$$f_\xi(t) = e^{-\frac{\sigma^2}{2}(t,t)}\,.$$

Сферически нормальное распределение инвариантно относительно ортогональных преобразований $\eta \;=\; \xi C$, так как

$$f_\eta(t) = f_\xi(tC') = e^{-\frac{\sigma^2}{2}(tC', tC')} = e^{-\frac{\sigma^2}{2}(t,t)}\,.$$

### Многомерное нормальное распределение ( общий случай)

Пусть $\xi = (\xi_1, \ldots, \xi_r)$ – $r$-мерный случайный вектор. Положим

$$\eta_i = \sum_{\ell=1}^{r} c_{\ell i} \xi_\ell + a_i, \quad i = 1, \ldots, m,$$

или в векторной форме

$$\eta = \xi\, C + a\,, \tag{2.8}$$

где $\eta = (\eta_1, \ldots, \eta_m)$, $a = (a_1, \ldots, a_m)$ и $C = \| \; c_{ij} \; \|$ – $(r \times m)$-матрица. Здесь вектор $a$ и матрица $C$ постоянны. Через $D[\xi] = D[\xi_1, \ldots, \xi_r]$ будем обозначать ковариационную матрицу $\| \operatorname{cov}(\xi_i, \xi_j) \, \|$.

**Теорема 2.34.** *Для любой постоянной матрицы $C$ в (2.8) имеет место равенство*

$$D[\eta] = C'D[\xi]C\,. \tag{2.9}$$

*Доказательство.* Из (2.8) следует, что $\eta_i - M\eta_i = \sum_{\ell=1}^{r} c_{\ell i}(\xi_\ell - M\xi_\ell)$ и

$$(\eta_i - M\eta_i)(\eta_j - M\eta_j) = \sum_{k,\ell=1}^{r} c_{\ell i}c_{kj}(\xi_\ell - M\xi_\ell)(\xi_k - M\xi_k)\,.$$

Вычисляя от обеих частей последнего равенства математическое ожидание, получаем

$$\mathrm{cov}(\xi_i, \xi_j) = \sum_{\ell,k=1}^{r} c_{ki}\mathrm{cov}(\xi_\ell, \xi_k)c_{\ell j} = \sum_{\ell,k=1}^{r} c'_{ik}\mathrm{cov}(\xi_k, \xi_\ell)c_{\ell j}\,,$$

где $c'_{ik} = c_{ki}$. □

**Определение 38.** Многомерным нормальным распределением $m$ случайных величин назовем распределение вектора $\eta = (\eta_1, \dots, \eta^m)$, где вектор $\eta$ определен формулой (2.8), а компоненты $\xi_1, \dots, \xi_r$ случайного вектора $\xi$ независимы и каждая из них распределена нормально с параметрами $(0, 1)$.

Отметим некоторые общие свойства многомерного нормального распределения. Пусть вектор $\eta = (\eta_1, \dots, \eta_m$ распределен нормально. Имеем:

1. *Одномерные распределения координат $\eta_i$ являются нормальными, если $D\eta_i > 0$.*

2. *Случайный вектор $\zeta$, представимый в виде $\zeta = A\eta$, где $A$ – постоянная матрица, имеет многомерное нормальное распределение.*

3.*Любая линейная комбинация $\zeta = a_1\eta_1 + \dots + a_n\eta_n$ компонент вектора $\eta$ имеет нормальное распределение, если $D\zeta > 0$.*

Эти свойства непосредственно следуют из определения и того факта, что сумма независимых нормально распределенных случайных величин нормально распределена.

**Теорема 2.35.** *Для того, чтобы случайный вектор $\eta = (\eta_1, \dots, \eta_m)$ был нормально распределен, необходимо и достаточно, чтобы характеристическая функция случайного вектора $\eta$ имела вид*

$$f_\eta(t) = e^{ita' - \frac{1}{2}tBt'}\,, \tag{2.10}$$

*где $a = (a_1, \dots, a_m)$ – постоянный вектор, $B = \| b_{\alpha\beta} \|$ – симметричная неотрицательно определенная $(m \times m)$-матрица.*

*Доказательство.* Найдем характеристическую функцию нормально распределенного случайного вектора. Согласно (37) и определению характеристической функции

$$f_\eta(t) = Me^{it\eta'} = M\exp\Big\{i\sum_{\ell=1}^{r}\Big(\sum_{k=1}^{m} t_k c_{k\ell}\Big)\xi_\ell + i\sum_{k=1}^{m} t_k a_k\Big\}$$

Используя независимость случайных величин $\xi_1, \dots, \xi_r$, получаем

$$\begin{aligned} f_\eta(t) &= \exp\Big\{i\sum_{k=1}^{m} t_k a_k\Big\} \cdot \prod_{\ell=1}^{r} M\exp\Big\{i\Big(\sum_{k=1}^{m} t_k c_{k\ell}\Big)\xi_\ell\Big\} = \\ &= \exp(ita')\prod_{\ell=1}^{r}\exp\Big\{-\frac{1}{2}\Big(\sum_{k=1}^{m} t_k c_{k\ell}\Big)^2\Big\} = \\ &= \exp(ita')\exp\Big\{-\frac{1}{2}\sum_{k_1,k_2=1}^{m} t_{k_1}t_{k_2}\Big(\sum_{\ell=1}^{r} c_{k_1\ell}c_{k_2\ell}\Big)\Big\}\,. \end{aligned}$$

Так как

$$b_{k_1k_2} = \operatorname{cov}(\eta_{k_1}, \eta_{k_2}) = \sum_{\ell=1}^{r} c_{k_1\ell}c_{k_2\ell}, \quad \text{(см. теорему 2.34)},$$

то окончательно получаем

$$f_\eta(t) = \exp\left\{ita' - \frac{1}{2}\sum_{k_1,k_2=1}^{m} b_{k_1k_2}t_{k_1}t_{k_2}\right\}.$$

Пусть теперь (2.10) имеет место. Тогда каждая компонента $\eta_k$ имеет характеристическую функцию вида

$$f_{\eta_k}(t) = e^{ita_k - \frac{b_{kk}}{2}t^2},$$

то есть нормально распределена с $M\eta_k = a_k$ и $D\eta_k = b_{kk}$. Далее удобнее будет перейти к центрированному вектору $\eta - a$. Сохраняя обозначения, будем считать, что $M\eta_k = 0$. Поскольку конечны все моменты $M\eta_\alpha^2$, то конечны и смешанные моменты $M\eta_\alpha\eta_\beta$, поэтому их можно вычислить, пользуясь равенством

$$\operatorname{cov}(\eta_\alpha, \eta_\beta) = M\eta_\alpha\eta_\beta = -\frac{\partial^2 f_\eta}{\partial t_\alpha \partial t_\beta}(0) = b_{\alpha\beta}.$$

Таким образом, $B = \| b_{\alpha\beta} \|$ является ковариационной матрицей случайного вектора $\eta$.

Для любой $(m \times m)$-матрицы $S = \| s_{\alpha\beta} \|$ случайный вектор $\zeta = \eta S$ имеет нулевое математическое ожидание $M\zeta = 0$, ковариационную матрицу $\| \operatorname{cov}(\zeta_\alpha, \zeta_\beta) \| = S'BS$ и характеристическую функцию

$$f_\zeta(t) = f_\eta(tS') = e^{-\frac{1}{2}t\,S'BS\,t'}.$$

Пусть $S$ – такая ортогональная матрица, что

$$S'BS = \begin{Vmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{mm} \end{Vmatrix} = D, \quad d_{\alpha\alpha} \geq 0.$$

Тогда

$$f_\zeta(t) = e^{-\frac{1}{2}\sum_{\alpha=1}^{m} d_{\alpha\alpha}t_\alpha^2},$$

поэтому $\zeta_1, \dots, \xi_m$ независимы, причем для тех $\alpha$, для которых $d_{\alpha\alpha} > 0$ случайные величины $\zeta_\alpha$ нормально распределены с параметрами $(0, \sqrt{d_{\alpha\alpha}})$, а при $d_{\alpha\alpha} = 0$ с вероятностью 1 имеем $\zeta_\alpha = 0$.

Если ранг $r$ матрицы $B$ меньше $m$, то можем считать, что $d_{\alpha\alpha} \neq 0$ для $\alpha = 1, \dots, r$ и $d_{r+1,r+1} = \dots = d_{mm} = 0$. В этом случае $\mathbf{P}\{\zeta_{r+1} = \dots = \zeta_m = 0\} = 1$, то есть случайные величины $\eta_1, \dots, \eta_m$ сосредоточены на пространстве, определяемом равенствами

$$\sum_{\beta=1}^{m} s_{\alpha\beta}\eta_\beta = 0, \quad \alpha = r+1, \dots, k.$$

Тогда

$$\eta_i = \sum_{k=1}^{r} s_{ki}\zeta_k, \quad i = 1, \dots, m.$$

Положим $\xi_\alpha = \zeta_\alpha / \sqrt{d_{\alpha\alpha}}$ и введем $(r \times m)$-матрицу $C$ с компонентами $c_{\alpha\beta} = \sqrt{d_{\alpha\alpha}}s_{\beta\alpha}$. Случайные величины $\xi_\alpha$ независимы и нормально распределены с параметрами $(0, 1)$. В результате, имеем

$$\eta = \xi\, C,$$

то есть

$$\eta_i = \sum_{k=1}^{r} c_{\alpha i}\xi_\alpha, \quad i = 1, \dots, m. \ \square$$

**Теорема 2.36.** *Если $r = m$ и определитель $\det C \neq 0$, то $r$-мерное нормальное распределение $\eta$ является невырожденным абсолютно непрерывным, и его плотность распределения имеет вид*

$$p_\eta(x_1, \ldots, x_r) = \frac{1}{(2\pi)^{r/2}\sqrt{\det B}} e^{\frac{1}{2}Q(x-a)} \quad \textit{(см. определение 36)},$$

*где $B = \| b_{k\ell} \|$ – невырожденная $(r \times r)$-матрица, $b_{k\ell} = \operatorname{cov}(\eta_k, \eta_\ell)$, $x = (x_1, \ldots, x_r)$, $a = (a_1, \ldots, a_r)$, $a_k = M\eta_k$, $Q(x) = \sum b^*_{k\ell} x_k x_\ell$ – квадратичная форма, коэффициенты которой составляют матрицу $\| b^*_{k\ell} \|$, обратную к матрице $B$.*

*Доказательство.* Плотность распределения случайного вектора $\xi$ из теоремы 2.35 имеет вид

$$p_\xi(x_1, \ldots, x_r) = \frac{1}{(\sqrt{2\pi})^r} e^{\frac{1}{2}\sum\limits_{k=1}^{r} x_k^2}.$$

Так как $\det C \neq 0$ преобразование вида

$$y_i = g_i(x_1, \ldots, x_r) = \sum_{k=1}^{r} c_{i\ell} x_\ell + a_i \quad i = 1, \ldots, m,$$

является взаимнооднозначным и его якобиан $J$ равен $\det C$. Пусть $x_i = \sum_{\ell=1}^{r} c^*_{i\ell}(y_\ell - a_\ell)$, где $c^*_{i\ell}$ – элементы матрицы, обратной к $C$. Тогда

$$p_\eta(y_1, \ldots, y_r) = |J|^{-1} p_\xi(g^{-1}(y)) = (\sqrt{2\pi})^{-r} |J|^{-1} e^{\frac{1}{2}Q(y-a)},$$

где

$$Q(y-a) = \sum_{k=1}^{r} x_k^2 = \sum_{k=1}^{r} \Big( \sum_{\ell=1}^{r} c^*_{k\ell}(y_\ell - a_\ell) \Big)^2 = \sum_{\ell_1, \ell_2 = 1}^{r} (y_{\ell_1} - a_{\ell_1})(y_{\ell_2} - a_{\ell_2}) \Big( \sum_{k=1}^{r} c^*_{k\ell_1} c^*_{k\ell_2} \Big)$$

Так как $D[\xi] = E$ (E – единичная матрица) имеем $B = C' D[\xi] C = C'C$. Следовательно,

$$\| b^*_{\ell_1 \ell_2} \| = B^{-1} = C^{-1}(C')^{-1} = \| \sum_{k=1}^{r} c^*_{k\ell_1} c^*_{k\ell_2} \|$$

и $\det B = |\det C|^2$. Заменяя $|J|$ на $\sqrt{\det B}$ и $\sum_{k=1}^{r} c^*_{k\ell_1} c^*_{k\ell_2}$ на $b^*_{\ell_1\ell_2}$, получим утверждение теоремы. □

### Предельная теорема для сумм независимых одинаково распределенных случайных векторов

Докажем здесь предельную теорему для сумм независимых одинаково распределенных случайных векторов.

**Теорема 2.37.** *Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых одинаково распределенных случайных векторов $\xi_n = (\xi_{n1}, \ldots, \xi_{nk})$ с $M\xi_{nk} = 0$ и с конечными ковариациями $\operatorname{cov}(\xi_{n\alpha}, \xi_{n\beta}) = b_{\alpha\beta}$. Обозначим $\zeta = \xi_1 + \cdots + \xi_n$. Тогда функция распределения случайного вектора*

$$\zeta' = \frac{\zeta_n - na}{\sqrt{n}}$$

*слабо сходится к нормальной функции распределения с нулевым математическим ожиданием и матрицей ковариации $B = \| b_{\alpha\beta} \|$.*

*Доказательство.* Обозначим через $f(t)$ характеристическую функцию случайного вектора $\widetilde{\xi}_n = \xi_n - a$. Поскольку $M\widetilde{\xi}_n = 0$ и $M\widetilde{\xi}_{n\alpha}\widetilde{\xi}_{n\beta} = b_{\alpha\beta}$, то по свойству характеристической функции

$$f\Big(\frac{t}{\sqrt{n}}\Big) = 1 - \frac{1}{2n}\sum_{\alpha,\beta=1}^{k} b_{\alpha\beta}t_\alpha t_\beta + o\Big(\frac{1}{n}\Big).$$

Поэтому при любом $t$ при $n \to \infty$

$$f_{\zeta_n'(t)} = \Big[f\big(\frac{t}{\sqrt{n}}\big)\Big]^n \to e^{-\frac{1}{2}\sum_{\alpha,\beta=1}^{k} b_{\alpha\beta}t_\alpha t_\beta},$$

отсюда и из теоремы 2.33 следует доказываемое утверждение. □

Важным частным случаем этой теоремы является следующее утверждение

**Теорема 2.38.** *Пусть $\xi = (\xi_1, \dots, \xi_k)$ имеет полиномиальное распределение в $n$ испытаниях с вероятностями исходов $p = (p_1, \dots, p_k)$ в отдельном испытании. Распределение вектора $(\xi - np)/\sqrt{n}$ при $n \to \infty$ слабо сходится к нормальному с нулевым средним и матрицей ковариации $\| \delta_{\alpha\beta}p_\alpha - p_\alpha p_\beta \|$, где $\delta_{\alpha\beta}$ – символ Кронекера.*

*Доказательство.* Случайный вектор $\xi$ представим в виде суммы $\eta_1 + \dots + \eta_n$ независимых векторов $\eta_\alpha = (\eta_{\alpha 1}, \dots, \eta_{\alpha_k})$, где $\eta_{\alpha\beta} = 1$, если при $\alpha$-м испытании произошел исход $\beta$ и $\eta_{\alpha\beta} = 0$ в противоположном случае. Совместный закон распределения случайных величин $\eta_{\alpha\beta}$ и $\eta_{\alpha\gamma}$ имеет вид

$$\begin{aligned} &\mathbf{P}\{\eta_{\alpha\beta} = 1,\ \eta_{\alpha\gamma} = 1\} = 0,\ \ \mathbf{P}\{\eta_{\alpha\beta} = 1,\ \eta_{\alpha\gamma} = 0\} = p_\beta, \\ &\mathbf{P}\{\eta_{\alpha\beta} = 0,\ \eta_{\alpha\gamma} = 1\} = p_\gamma,\ \ \mathbf{P}\{\eta_{\alpha\beta} = 0,\ \eta_{\alpha\gamma} = 0\} = 1 - p_\beta - p_\gamma. \end{aligned}$$

Учитывая, что $M\eta_{\alpha\beta} = p_\beta$ и $\mathrm{cov}(\eta_{\alpha\beta}, \eta_{\alpha\gamma}) = p_\beta\delta_{\beta\gamma} - p_\beta p_\gamma$, в силу предыдущей теоремы получаем требуемое утверждение. □

## 2.7. Распределения, связанные с многомерным нормальным распределением

### $\chi^2$-распределение

Пусть $\xi = (\xi_1, \dots, \xi_k)$ – сферически нормальное распределение. Тогда $\xi_i$, $i = 1, \dots, k$, являются нормально распределенными с параметрами $(0, 1)$ независимыми случайными величинами. Положим $\chi_k^2 = \xi_1^2 + \dots + \xi_n^2$ и $\chi_k = \sqrt{\xi_1^2 + \dots + \xi_k^2}$. Найдем функцию распределения и плотность случайной величины $\chi_k$:

$$\mathbf{P}\{\chi_k \le r\} = \frac{1}{(2\pi)^{k/2}}\int\limits_{|x|\le r} e^{-\frac{1}{2}x\,x'}dx_1\dots dx_k = c_k\int\limits_0^r t^{k-1}e^{-\frac{1}{2}t^2}dt$$

Поэтому

$$p_{\chi_k}(x) = c_k r^{k-1}e^{-\frac{r^2}{2}},\ \ x \ge 0.$$

Постоянная $c_k$ находится из условия

$$\int\limits_0^\infty p_{\chi_k}(x)dx = 1.$$

Имеем

$$c_k \int\limits_0^\infty x^{k-1} e^{-\frac{x^2}{2}} dx = c_k\, 2^{\frac{k}{2}} \int\limits_0^\infty x^{k-1} e^{-x^2} dx = c_k\, 2^{\frac{k}{2}-1} \int\limits_0^\infty x^{\frac{k-2}{2}} e^{-x} dx =$$

$$= c_k\, 2^{\frac{k}{2}-1} \int\limits_0^\infty x^{\frac{k}{2}-1} e^{-x} dx = c_k\, 2^{\frac{k}{2}-1} \Gamma\big(\frac{k}{2}\big) = 1\,.$$

Поэтому

$$p_{\chi_k}(x) = \frac{x^{k-1}}{2^{\frac{k}{2}-1}\Gamma\big(\frac{k}{2}\big)} e^{-\frac{x^2}{2}}, \;\; x \geq 0\,.$$

Плотность распределения случайной величины $\chi_k^2$ имеет вид

$$p_{\chi_k^2}(x) = \frac{1}{2\sqrt{x}} p_{\chi_k}(\sqrt{x}) = \frac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}} \Gamma\big(\frac{1}{2}\big) e^{-\frac{x}{2}}\,, \;\; x \geq 0\,. \tag{2.11}$$

Случайная величина с плотностью (2.11) называется $\chi^2$-*распределенной* с $k$ степенями свободы.

### Распределение Стьюдента

Пусть случайные величины $\xi_0, \xi_1, \ldots, \xi_k$ независимы и нормально распределены с параметрами $(0,1)$. Случайная величина

$$\tau_k = \frac{\xi_0}{\sqrt{\frac{1}{k}\sum\limits_{i=1}^k \xi_i^2}} = \frac{\xi_0\sqrt{k}}{\chi_k}$$

называется отношением Стьюдента с $k$ степенями свободы. Найдем функцию распределения и плотность $\tau_k$:

$$S_k(x) = \mathbf{P}\{\tau_k \leq x\} = \iint\limits_{\substack{\frac{u}{v}\leq x\\ v>0}} \mathbf{P}_{\xi_0\sqrt{k}}(u) \mathbf{P}_{\chi_k}(v) du dv = a_k \iint\limits_{\substack{\frac{u}{v}\leq x\\ v>0}} v^{k-1} e^{-\frac{1}{2}(\frac{u^2}{k}+v^2)} du dv\,, \tag{2.12}$$

где

$$a_k = \frac{1}{\sqrt{2\pi k}} \frac{1}{2^{\frac{k}{2}-1}\Gamma(\frac{k}{2})}\,.$$

Сделаем замену переменных $u = yz$, $v = z$. Учитывая, что

$$\frac{\partial(u,v)}{\partial(y,z)} = z\,,$$

получаем

$$(2.12) = a_k \int\limits_{-\infty}^x dy \int\limits_0^\infty z^k e^{-\frac{z^2}{2}(\frac{y^2}{k}+1)} dz\,. \tag{2.13}$$

После замены

$$w = z\sqrt{\frac{y^2}{k}+1}$$

получаем

$$(2.13) = a_k \int\limits_{-\infty}^x \big(\frac{y^2}{k}+1\big)^{-\frac{k+1}{2}} dy \int\limits_0^\infty w^k e^{-\frac{w^2}{2}} dw = a_k 2^{\frac{k-1}{2}} \Gamma\big(\frac{k+1}{2}\big) \int\limits_{-\infty}^x \big(\frac{y^2}{k}+1\big)^{-\frac{k+1}{2}} dy\,.$$

Поэтому плотность $p_{\tau_k}(x)$ равна

$$p_{\tau_k}(x)=\frac{1}{\sqrt{k\pi}}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}},\ \ -\infty<x<\infty\,. \tag{2.14}$$

Случайная величина с плотностью (2.14) называется распределенной по Стьюденту с $k$ степенями свободы.

Так как $\Gamma(1+x)\sim\sqrt{2\pi x}x^xe^{-x}$ то нетрудно убедиться, что для каждого $x\in\mathbb{R}$

$$p_{\tau_k}(x)\to\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\ \text{ при }\ k\to\infty\,.$$

Распределение Стьюдента ввел англичанин Госсет, работавший на фирме, занимавшейся пивоварением.

## F-распределение

Пусть $\xi_1,\dots,\xi_p,\eta_1,\dots,\eta_q$ – независимые $(0,1)$-нормально распределенные случайные величины. Введем случайную величину

$$F_{pq}=\frac{\frac{1}{p}\sum\limits_{\alpha=1}^{p}\xi_\alpha^2}{\frac{1}{q}\sum\limits_{\beta=1}^{q}\eta_\beta^2}\,.$$

Имеем

$$\mathbf{P}\left\{\frac{p}{q}F_{pq}\le x\right\}=\frac{1}{2^{\frac{p+q}{2}}\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}\iint\limits_{\substack{\frac{u}{v}\le x\\ u\ge0,v\ge0}}u^{\frac{p}{2}-1}v^{\frac{q}{2}-1}e^{-\frac{u+v}{2}}du dv\,. \tag{2.15}$$

Положим

$$a_{pq}=\frac{1}{2^{\frac{p+q}{2}}\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}\,.$$

Сделаем замену переменных $u=yz$, $v=z$. Тогда

$$(2.15)=a_{pq}\int\limits_0^x y^{\frac{p}{2}-1}dy\int\limits_0^\infty z^{\frac{p+q}{2}-1}e^{-\frac{z(1+y)}{2}}dz=a_{pq}2^{\frac{p+q}{2}}\Gamma\left(\frac{p+q}{2}\right)\int\limits_0^x\frac{y^{\frac{p}{2}-1}}{(1+y)^{\frac{p+q}{2}}}dy\,.$$

Поэтому

$$\mathbf{P}\left\{\frac{p}{q}F_{pq}\le x\right\}=\frac{\Gamma\left(\frac{p+q}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}\int\limits_0^x\frac{y^{\frac{p}{2}-1}}{(1+y)^{\frac{p+q}{2}}}dy\,.$$

Плотность распределения $F_{pq}$ имеет вид

$$p_{F_{pq}}(x)=p^{\frac{p}{2}}q^{\frac{q}{2}}\frac{\Gamma\left(\frac{p+q}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}\frac{x^{\frac{p}{2}-1}}{(q+px)^{\frac{p+q}{2}}}\,,x\ge0\,.$$

Распределения с такой плотностью называются $F$-распределениями Фишера.

## 2.8. Закон больших чисел

**О сходимости случайных величин**

Пусть $(\Omega, \mathcal{U}, \mathbf{P})$ есть вероятностное пространство, $\zeta_n$, $n = 1, \ldots,$ и $\zeta$ – случайные величина.

**Определение 39.** Последовательность случайных величин $\{\zeta_n\}$ сходится по вероятности к $\zeta$ ($\zeta_n \xrightarrow{P} \zeta$), если для любого $\varepsilon > 0$

$$\mathbf{P}\{|\zeta_n - \zeta| > \varepsilon\} \to 0 \text{ при } n \to \infty\,.$$

*Пример 36.* Пусть $\mu_n$ – число успехов в серии $n$ независимых испытаний с вероятностью успеха $p$ в одном испытании. Согласно теореме Бернулли $\frac{\mu_n}{n} \xrightarrow{P} p$.

**Определение 40.** Последовательность случайных величин $\{\zeta_n\}$ сходится почти наверное к $\zeta$, пишут $\zeta_n \xrightarrow{\text{п.н.}} \zeta$, если $\zeta(\omega) \to \zeta(\omega)$ на $A \in \mathcal{U}$ и $\mathbf{P}(\Omega \backslash A) = 0$.

**Теорема 2.39.** *Из сходимости $\{\zeta_n\}$ к $\zeta$ п.н. следует сходимость по вероятности.*

*Доказательство.* Пусть $\zeta_n(\omega) \to \zeta(\omega)$ на $A \in \mathcal{U}$. Для любого положительного $\varepsilon$ имеет место включение

$$A \subset \bigcup_n \bigcap_{k \geq n} \{|\zeta_k - \zeta| \leq \varepsilon\}\,.$$

Поэтому верно соотношение

$$\Omega \backslash A \supset \bigcap_n \bigcup_{k \geq n} \{|\zeta_k - \zeta| > \varepsilon\}\,.$$

Так как

$$0 = \mathbf{P}(\Omega \backslash A) \geq \lim_{n \to \infty} \mathbf{P}\{\bigcup_{k \geq n} \{|\zeta_k - \zeta| > \varepsilon\}\}\,,$$

из свойства монотонности меры следует, что

$$\lim_{n \to \infty} \mathbf{P}\{|\zeta_n - \zeta| > \varepsilon\} = 0\,. \;\square$$

Обратное утверждение неверно. Но из сходящейся по вероятности последовательности случайных величин можно выбрать подпоследовательность, сходящуюся почти наверное.

**Определение 41.** Последовательность случайных величин $\{\zeta_n\}$ сходится к случайной величине $\zeta$ по распределению, или слабо сходится, если $F_{\zeta_n}(x) \to F_\zeta(x)$ в каждой точке непрерывности предельной функции распределения. Слабая сходимость $F_{\zeta_n}$ к $F_\zeta$ обозначается через $F_{\zeta_n} \Rightarrow F_\zeta$.

**Теорема 2.40.** *Из сходимости $\{\zeta_n\}$ к $\zeta$ по вероятности следует сходимость по распределению.*

*Доказательство.* Введем событие $A_n = \{|\zeta_n - \zeta| \leq \varepsilon\}$. Так как при $\omega \in A_n$

$$\zeta(\omega) - \varepsilon \leq \zeta_n(\omega) \leq \zeta(\omega) + \varepsilon\,,$$

то при каждом $x$ имеем

$$\{\zeta_n \leq x\} \subseteq \{\zeta \leq x + \varepsilon\} \bigcup \overline{A}_n$$

и

$$\{\zeta \leq x - \varepsilon\} \subseteq \{\zeta_n \leq x\} \bigcup \overline{A}_n$$

Отсюда следует, что

$$\mathbf{P}\{\zeta \le x-\varepsilon\} - \mathbf{P}(\overline{A}_n) \le \mathbf{P}\{\zeta_n \le x\} \le \mathbf{P}\{\zeta \le x+\varepsilon\} + \mathbf{P}(\overline{A}_n)\,.$$

Поэтому

$$\mathbf{P}\{\zeta \le x-\varepsilon\} \le \liminf_{n\to\infty} \mathbf{P}\{\zeta_n \le x\} \le \limsup_{n\to\infty} \mathbf{P}\{\zeta_n \le x\} \le \mathbf{P}\{\zeta \le x+\varepsilon\} \tag{2.16}$$

Если $x$ есть точка непрерывности $F_\zeta(x)$, то из (2.16) получаем

$$\lim_{n\to\infty} F_{\zeta_n}(x) = F_\zeta(x)\,. \ \square$$

Из сходимости по распределению, вообще говоря, не следует сходимость по вероятности. Но если $F_{\zeta_n}(x)$ слабо сходится к вырожденному распределению, то имеет место обратное утверждение.

**Теорема 2.41.** *Если $F_{\zeta_n} \Rightarrow F_\zeta$ и $F_\zeta$ вырождено в точке $c$, то есть*

$$F_\zeta(x) = \begin{cases} 0, & x < c \\ 1, & x \ge c\,, \end{cases}$$

*то $\zeta_n \xrightarrow{P} c$.*

*Доказательство.* Так как $F_{\zeta_n}(c+\varepsilon) \to 1$ и $F_{\zeta_n}(c-\varepsilon) \to 0$, то $\mathbf{P}\{c-\varepsilon < \zeta_n \le c+\varepsilon\} \to 1$, то есть $\mathbf{P}\{|\zeta_n - c| > \varepsilon\} \to 0$. $\square$

Следующая теорема нам потребуется при изучении математической статистики.

**Теорема 2.42.** *Пусть $\zeta_n = \xi_n + \eta_n$, $n = 1, 2, \ldots$ – случайные величины. Если $\eta_n \xrightarrow{P} 0$ и последовательность функций распределения $F_{\xi_n}(x)$ слабо сходится к функции распределения $F(x)$, то последовательность функций распределения $F_{\zeta_n}(x)$ также слабо сходится к $F(x)$.*

*Доказательство.* Пусть $x$ – точка непрерывности $F(x)$ и $\varepsilon > 0$. Положим

$$A_n = \{\xi_n + \eta_n < x\}, \quad B_n = \{|\eta_n| < \varepsilon\}.$$

Учитывая, что

$$A_n = A_nB_n + A_n\overline{B}_n \subset A_nB_n + \overline{B}_n\,.$$

получаем

$$\mathbf{P}(A_nB_n) \le \mathbf{P}(A_n) \le \mathbf{P}(A_nB_n) + \mathbf{P}(\overline{B}_n)\,.$$

Отсюда и из соотношения

$$\{(\xi_n < x-\varepsilon) \cap B_n\} \subset A_nB_n \subset \{\xi_n < x+\varepsilon\}\,.$$

находим

$$\mathbf{P}\{(\xi_n < x-\varepsilon) \cap B_n\} \le \mathbf{P}(A_n) \le \mathbf{P}\{\xi_n < x+\varepsilon\} + \mathbf{P}(\overline{B}_n)\,. \tag{2.17}$$

Из (2.17) и неравенства

$$\mathbf{P}\{(\xi_n < x-\varepsilon) \cap B_n\} \ge \mathbf{P}\{\xi_n < x-\varepsilon\} - \mathbf{P}(\overline{B}_n)$$

следует, что

$$F_{\xi_n}(x-\varepsilon) - \mathbf{P}(\overline{B}_n) \le \mathbf{P}(A_n) \le F_n(x+\varepsilon) + \mathbf{P}(\overline{B}_n)\,.$$

По условию теоремы $\mathbf{P}(\overline{B}_n) \to 0$ при $n \to \infty$, поэтому

$$F(x-\varepsilon) \leqslant \liminf_{n\to\infty} \mathbf{P}\{\zeta_n \leqslant x\} \leqslant \limsup_{n\to\infty} \mathbf{P}\{\zeta_n \leqslant x\} \leqslant F(x+\varepsilon)\,.$$

Отсюда при $\varepsilon \to 0$ получаем утверждение теоремы, так как $x$ – точка непрерывности функции $F(x)$. $\square$

## Закон больших чисел в форме Хинчина

**Теорема 2.43.** *Пусть* $\{\xi_n\}_{n=1}^{\infty}$ *– последовательность независимых одинаково распределенных случайных величин, имеющих математическое ожидание* $M\xi_n = a$. *Тогда*

$$\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{P} a \;\; \textit{при} \;\; n \to \infty\,.$$

*Доказательство.* Через $f(t)$ обозначим характеристическую функцию случайной величины $\xi_1$, а через $S_n$ – сумму $\xi_1 + \dots + \xi_n$. Достаточно доказать, что для каждого $t$

$$f_{\frac{S_n}{n}}(t) \to e^{iat}\,.$$

В некоторой окрестности нуля выполняется неравенство

$$|f(t) - 1| < \frac{1}{2}\,.$$

Для таких $t$ определена функция $\ell(t) = \log f(t)$. Поэтому для любого $t$ и достаточно больших $n$ имеет место представление

$$f_{\frac{S_n}{n}}(t) = f^n\big(\frac{t}{n}\big) = e^{\ell\left(\frac{t}{n}\right)n}\,.$$

Так как

$$\ell'(0) = \frac{f'(0)}{f(0)} = ia, \;\; \ell(0) = 0\,,$$

то

$$e^{\ell\left(\frac{t}{n}\right)n} = \exp\left\{t\frac{\ell\left(\frac{t}{n}\right) - \ell(0)}{\frac{t}{n}}\right\} \to e^{\ell'(0)t} = e^{ita}\,. \;\; \square$$

**Усиленный закон больших чисел**

Сформулируем здесь две теоремы. Первая теорема об усиленном законе больших чисел для независимых равно распределенных случайных величин.

**Теорема 2.44.** *Пусть* $\xi_1, \xi_2, \dots$ *независимы,* $M\xi_k = 0$ *и* $D\xi_k = \sigma_k^2$. *Если*

$$\sum_{k=1}^{\infty} \frac{\sigma_k^2}{n^2} \leq \infty,$$

*тогда*

$$\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{\textit{п.н.}} 0\,.$$

Доказательство этой теоремы опирается на неравенство Колмогорова, усиливающего неравенство Чебышева.

**Теорема 2.45** (неравенство Колмогорова). *Пусть* $\xi_1, \xi_2, \dots$ *независимы и имеют конечные математическое ожидание и дисперсию. Тогда*

$$\mathbf{P}\{\max_{1\leq k\leq n} |\zeta_k - M\zeta_k| \geq x\} \leq \frac{D\zeta_n}{x^2}\,,$$

*где* $\zeta_k = \xi_1 + \dots + \xi_k$.

Для независимых одинаково распределенных случайных величин справедлив критерий сходимости почти всюду.

**Теорема 2.46** (усиленный закон больших чисел Колмогорова). *Пусть* $\xi_1, \xi_2, \dots$ *независимы и одинаково распределены. Для того, чтобы*

$$\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{\textit{п.н.}} a\,.$$

*необходимо и достаточно, чтобы существовало конечное* $M\xi_n = a$.

**Следствие.** *В схеме Бернулли для числа успехов имеет место не только закон больших чисел*

$$\frac{\mu_n}{n} \xrightarrow{P} p\,,$$

*но и усиленный закон больших чисел*

$$\frac{\mu_n}{n} \xrightarrow{\text{п.н.}} p\,.$$

Следствие вытекает из критерия Колмогорова, так как $\mu_n = \xi_1 + \dots + \xi_n$, где $\xi_1, \dots, \xi_n$ независимы и $\mathbf{P}\{\xi_k = 1\} = p$, $\mathbf{P}\{\xi_k = 0\} = 1 - p$.

# Глава 3

# Математическая статистика

## 3.1. Предмет математической статистики

Теория вероятностей изучает математические модели случайных явлений. Математическая статистика решает обратные задачи: разрабатывает различные методы, которые позволяют по статистическим данным, которые носят случайный характер, подобрать подходящую теоретико – вероятностную модель.

Множество объектов, подлежащих контролю называется генеральной совокупностью. Множество отобранных объектов называется выборкой. Число элементов выборки называется объемом выборки. Повторная выборка — это выборка с возвращением. Бесповторная выборка — выборка без возвращения. Выборка должна быть представительной и случайной. Объекты генеральной совокупности описываются одним или несколькими числовыми параметрами. Допустим, что объекты генеральной совокупности описываются случайной величиной $\xi$, принимающей на каждом объекте некоторое числовое значение.

Сделав выборку объема $n$, мы определяем ряд значений $\xi$. По этой последовательности нам следует получить представление о функции распределения, оценить математическое ожидание, дисперсию и, возможно, другие параметры случайной величины. Наблюдаемые в эксперименте значения $x_k$ случайной величины $\xi$ носят случайный характер. Эти $x_k$ принимают те же значения как и $\xi$ и распределены как $\xi$. На них следует смотреть как на $n$ случайных величин, распределенных как $\xi$, в частности, $Mx_k = M\xi = m$ и $Dx_k = D\xi = \sigma^2$.

Случайный вектор $(x_1, \ldots, x_n)$, где $x_k$ независимые и одинаково распределенные случайные величины с функцией распределения $\mathbf{P}\{x_k \leqslant x\} = F_\xi(x) = F(x)$, является математической моделью независимых измерений, проводимых в равных условиях. Случайная величина $\xi$ является подлежащей определению случайной величиной .

Мы будем рассматривать только повторные выборки. Если выборка составляет незначительную часть генеральной совокупности, то различие между повторной и бесповторной выборкой несущественно, а в предельном случае это различие исчезает.

## 3.2. Эмпирическая функция распределения

Эмпирическая функция распределения $F_n^*$ определяется соотношением

$$F_n^*(x) = \frac{\mu_n(x)}{n}\,,$$

где $\mu_n$ – число $x_k$ не превосходящих $x$. Заметим, что $F_n^*$ является случайной величиной.

**Теорема 3.1.** *Для каждого $x$ и любого $\varepsilon > 0$ выполняется равенство*

$$\lim_{n\to\infty} \mathbf{P}\{|F_n^*(x) - F(x)| < \varepsilon\} = 1\,.$$

*Доказательство.* Назовем случайное событие $\{x_k < x\}$ успехом. Тогда $F(x) = \mathbf{P}\{x_k \leqslant x\} = p$ – вероятность успеха, $\mu_n$ – число успехов и $F_n^*(x)$ – относительная частота успехов. Из теоремы Бернулли следует утверждение теоремы. □

Разобьем точками $-\infty = z_0 < z_1 < \cdots < z_m < z_{m+1} = \infty$ числовую ось на промежутки $\cup_{k=1}^{m}[z_k, z_{k+1})$. Тогда относительная частота попадания $x_i$ и промежуток $[z_k, z_{k+1})$ равна $p_k^* = F_n^*(z_{k+1} - F^*(z_k)$. При большом объеме выборки $p_k^*$ близки к вероятностям

$$p_k = F(z_{k+1}) - F(z_k) = \int\limits_{z_k}^{z_{k+1}} p(x)dx\,,$$

где $p(x)$ – плотность распределения случайной величины $\xi$. Над каждым промежутком $[z_k, z_{k+1})$ строится прямоугольник, площадь которого пропорциональна $p_k^*$. Построенная фигура называется гистограммой. Гистограмма дает представление о плотности распределения плотности $p(x)$, но чувствительна к разбиению.

Оценка неизвестной функции распределения может быть получена на основании теоремы Колмогорова.

**Теорема 3.2.** *Если $F(x)$ непрерывна, то при $n \to \infty$*

$$\mathbf{P}\{\sqrt{n}(\sup_{-\infty<x<\infty} |F_n^*(x) - F(x)|) < z\} \to K(z) =$$

$$= \begin{cases} 0, & \text{\textit{при} } z \leqslant 0 \\ \sum_{k=-\infty}^{\infty}(-1)^k e^{-2k^2z^2}, & \text{\textit{при} } z > 0 \end{cases}$$

Если $\widehat{z}_\alpha$ – решение уравнения $1 - K(\widehat{z}_\alpha) = \alpha$, то с вероятностью $1 - \alpha$ для всех $x$ выполняется неравенство

$$F_n^*(x) - \frac{\widehat{z}_\alpha}{\sqrt{n}} < F(x) < F_n^*(x) + \frac{\widehat{z}_\alpha}{\sqrt{n}}.$$

Следующая теорема утверждает, что с вероятностью 1 эмпирическая функция распределения $F_n^*$ равномерно сходится к $F$.

**Теорема 3.3** (Гливенко). *Пусть $F(x)$ функция распределения случайной величины $\xi$ и $F_n^*(x)$ эмпирическая функция распределения результатов $n$ независимых наблюдений над случайной величиной $\xi$. Тогда при $n \to \infty$*

$$\mathbf{P}\{\lim_{n\to\infty} \sup_{-\infty<x<\infty} |F_n^*(x) - F(x)| = 0\} = 1\,.$$

## 3.3. Выборочный метод

Выборочное среднее определяется как среднее арифметическое элементов выборки

$$\overline{x} = m_1^* = \frac{1}{n}\sum_{k=1}^{n} x_k\,,$$

выборочная дисперсия вводится следующим образом:

$$s^2 = \alpha_2^* = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2\,.$$

Выборочные моменты порядка $j$ определяются соотношениями

$$m_j^* = \frac{1}{n}\sum_{k=1}^{n} x_k^j \;\text{ и }\; \alpha_j^* = \sum_{k=1}^{n}(x_k - \overline{x})^j\,,$$

Выборочные моменты и центральные выборочные моменты являются случайными величинами.

Теоретические моменты $M\xi^k$ порядка $k$ будем обозначать через $m_k$, а центральный момент $M(\xi - M\xi)^k$ порядка $k$ будем обозначать через $\alpha_k$. В частности, $\alpha_2$ есть дисперсия, обозначаемая также через $\sigma^2$.

**Теорема 3.4.** *Выборочные моменты сходятся по вероятности к соответствующим теоретическим моментам при* $n \to \infty$.

*Доказательство.* Положим
$$\xi_1 = x_1^k, \ldots, \xi_n = x_n^k.$$
Случайные величины $\xi_i$ независимы и одинаково распределены, $M\xi_i = M\xi^k = m_k$. Согласно закону больших чисел
$$\lim_{n\to\infty} \mathbf{P}\Big\{\Big|\frac{\xi_1 + \cdots + \xi_n}{n} - M\xi_i\Big| > \varepsilon\Big\} = 0\,.$$
Из соотношений
$$m_k^* = \frac{\xi_1 + \cdots + \xi_n}{n}, \quad M\xi_i = m_k\,,$$
следует, что
$$m_k^* \xrightarrow{P} m_k\,. \ \square$$

**Теорема 3.5.** *Выборочные центральные моменты сходятся по вероятности к соответствующим теоретическим моментам.*

Предварительно докажем лемму.

**Лемма 3.1.** *Пусть* $f(t) = f(t_1, \ldots, t_k)$ – *непрерывная функция по совокупности переменных в точке* $(m_1, \ldots, m_k)$ *и последовательность случайных векторов* $(\xi_1^{(n)}, \ldots, \xi_k^{(n)})$ *сходится при* $n \to \infty$ *по вероятности к вектору* $(m_1, \ldots, m_k)$. *Тогда последовательность случайных величин* $f(\xi_1^{(n)}, \ldots, \xi_k^{(n)})$ *сходится по вероятности к* $f(m_1, \ldots, m_k)$.

*Доказательство.* По определению непрерывности
$$(\forall\varepsilon > 0)(\exists\delta > 0)(|t_i - m_i| < \delta, \ \ i = \overline{1,k} \Rightarrow |f(t_1, \ldots, t_k) - f(m_1, \ldots, m_k)| < \varepsilon).$$
Положим
$$A_i^{(n)} = \{|\xi_i^{(n)} - m_i| < \delta\} \ \text{ и } \ B^{(n)} = \{|f(\xi_1^{(n)}, \ldots, \xi_k^{(n)}) - f(m_1, \ldots, m_k)| < \varepsilon\}.$$
Тогда
$$\bigcap_{i=1}^{k} A_i^{(n)} \subset B^{(n)} \Rightarrow \overline{B^{(n)}} \subset \bigcup_{i=1}^{k} \overline{A_i^{(n)}}\,.$$
Отсюда следует, что
$$\mathbf{P}\{|f(\xi_1^{(n)}, \ldots, \xi_k^{(n)}) - f(m_1, \ldots, m_k)| > \varepsilon\} \leqslant \sum_{i=1}^{k} \mathbf{P}\{|\xi_i^{(n)} - m_i| \geqslant \delta\} \to 0\,. \ \square$$

*Доказательство теоремы.* Введем функцию
$$f(y) = f(y_0, \ldots, y_k) = \sum_{j=0}^{k} (-1)^{k-j} C_k^j y_1^{k-j} y_j\,.$$
Заметим, что
$$\alpha_k = M(\xi - M\xi)^k = M\Big(\sum_{j=0}^{k} (-1)^{k-j} C_k^j (M\xi)^{k-j} \xi^j\Big) =$$
$$\sum_{j=0}^{k} (-1)^{k-j} C_k^j (M\xi)^{k-j} M\xi^j = \sum_{j=0}^{k} (-1)^{k-j} C_k^j m_1^{k-j} m_j = f(1, m_1, \ldots, m_k)$$

и

$$\alpha_k^* = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^k = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k}(-1)^{k-j}C_k^j x_i^j \overline{x}^{k-j}$$

$$\sum_{j=1}^{k}(-1)^{k-j}C_k^j\overline{x}^{k-j}\Big(\frac{1}{n}\sum_{i=1}^{n}x_i^j\Big) = \sum_{j=1}^{k}(-1)^{k-j}C_k^j(m_1^*)^{k-j}m_j^* = f(1,m_1^*,\ldots,m_k^*)$$

Так как $m_i^* \xrightarrow{P} m_i$, из леммы следует, что

$$f(1,m_1^*,\ldots,m_k^*) \xrightarrow{P} f(1,m_1,\ldots,m_k), \quad \text{то есть } \alpha_k^* \xrightarrow{P} \alpha_k. \ \Box$$

## 3.4. Понятие оценки

Пусть $\mathcal{F}$ – множество распределений и $\theta$ – числовая характеристика, однозначно определяемая по распределению. В таком случае $\theta$ называется параметром распределения. Например, $\theta = \int_{\mathbf{R}} x dF(x)$. Соответствие не является взаимно однозначным. Задача оценки параметра состоит в нахождении функции $\theta^* = \theta(x_1,\ldots,x_n)$ от элементов выборки, которая в каком-либо смысле близка к параметру $\theta = \theta(F_\xi)$. Так как $x_k$ – независимые случайные величины, то $\theta^*$ – случайная величина, определяемая распределением $F_\xi$. Любая функция от элементов выборки называется статистикой. Таким образом, оценка параметра является статистикой. Близость статистики к оцениваемому параметру обеспечивается условиями несмещенности, состоятельности и эффективности.

1. **Несмещенность** означает отсутствие систематической ошибки, то есть

$$M\theta^* = \theta\,.$$

В качестве примера рассмотрим выборочное среднее и выборочную дисперсию.

**Теорема 3.6.** *Пусть $(x_1,\ldots,x_n)$ – выборка из распределения $F_\xi(x)$ с математическим ожиданием $m$ и дисперсией $\sigma^2$. Тогда*

$$M\overline{x} = m, \quad D\overline{x} = \frac{\sigma^2}{n}, \quad Ms^2 = \frac{n-1}{n}\sigma^2$$

*и*

$$Ds^2 = \frac{\alpha_4-\alpha_2^2}{n} - 2\frac{\alpha_4-2\alpha_2^2}{n^2} + \frac{\alpha_4-3\alpha_2^2}{n^3} = O\big(\frac{1}{n}\big)\,,$$

*если конечен момент $\alpha_4$.*

*Доказательство.* Первые два утверждения легко проверяются. Остановимся на доказательстве двух последних равенств. Имеем

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

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

Поэтому

$$Ms^2 = \frac{1}{n}\sum_{i=1}^{n}M(x_i-m)^2 - M(\overline{x}-m)^2 = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2\,.$$

Введем $y_k = x_k - m$. Тогда

$$\overline{y} = \frac{1}{n}\sum_{k=1}^{n} y_k = \frac{1}{n}\sum_{k=1}^{n} x_k - m\,.$$

Поэтому

$$s^2 = \alpha_2^* = \frac{1}{n}\sum_{k=1}^{n}(y_k - \overline{y})^2 = \frac{1}{n}\sum_{k=1}^{n} y_k^2 - \overline{y}^2$$

и

$$(\alpha_2^*)^2 = \frac{1}{n^2}\Big(\sum_{k=1}^{n} y_k^2\Big)^2 - \frac{2}{n}\overline{y}^2\Big(\sum_{k=1}^{n} y_k^2\Big) + \overline{y}^4\,.$$

Заметим, что $My_k = 0$ и $y_k$ независимы. Поэтому в сумме

$$M\overline{y}^4 = \frac{1}{n^4}\sum_{\ell_1,\ell_2,\ell_3,\ell_4=1}^{n} M(y_{\ell_1}y_{\ell_2}y_{\ell_3}y_{\ell_4})$$

отличны от нуля только $n$ слагаемых вида $My_\ell^4 = \alpha_4$ и $C_n^2 C_4^2 = 3n(n-1)$ слагаемых вида $My_\ell^2 y_k^2 = \alpha^2$, $k \neq \ell$. В результате, имеем $M\overline{y}^4 = (\alpha_4 + 3(n-1)\alpha_2^2)/n^3$. Далее,

$$M\Big[\big(\sum_{k=1}^{n} y_k^2\big)^2/n\Big] = M\Big(\frac{1}{n}\big[\sum_{k=1}^{n} y_k^4 + 2\sum_{k\neq\ell} y_k^2 y_\ell^2\big]\Big) = \alpha_4 + 2\frac{1}{n}\frac{n(n-1)}{2}\alpha_2^2 = \alpha_4 + (n-1)\alpha_2^2\,.$$

Представим $\overline{y}^2$ в виде

$$\overline{y}^2 = \frac{1}{n^2}\Big[\sum_{k=1}^{n} y_k^2 + 2\sum_{k\neq\ell} y_k y_\ell\Big]\,.$$

Тогда

$$\frac{1}{n}\overline{y}^2\big(\sum_{k=1}^{n} y_k^2\big) = \frac{1}{n^3}\big(\sum_{k=1}^{n} y_k^2\big)^2 + 2\frac{1}{n^3}\big(\sum_{k\neq\ell} y_k y_\ell\big)\big(\sum_{k=1}^{n} y_k^2\big)$$

Поэтому

$$n^2 M\Big[\frac{1}{n}\overline{y}^2\big(\sum_{k=1}^{n} y_k^2\big)\Big] = M\Big[\frac{1}{n}\big(\sum_{k=1}^{n} y_k^2\big)^2\Big] = \alpha_4 + (n-1)\alpha_2^2\,,$$

так как $M(y_k y_\ell^3) = M(y_k y_\ell y_r^2) = 0$ при $k \neq \ell \neq r$. Легко теперь убедиться, что,

$$Ds^2 = D\alpha_2^* = M(\alpha_2^*)^2 - (M\alpha_2^*)^2 = \frac{\alpha_4 - \alpha_2^2}{n} - 2\frac{\alpha_4 - 2\alpha_2^2}{n^2} + \frac{\alpha_4 - 3\alpha_2^2}{n^3}\,.\ \square$$

Согласно доказанной теореме выборочное среднее являются несмещенной оценкой математического ожидания, а выборочная дисперсия является смещенной оценкой. Несмещенной оценкой дисперсии является статистика

$$\hat{s}^2 = \frac{n-1}{n}s^2\,,$$

так называемая исправленная выборочная дисперсия.

В общем случае, можно показать, что

$$M\alpha_k^* = \alpha_k + O\Big(\frac{1}{n}\Big) \text{ и } M(\alpha_k^* - \alpha_k)^2 = O\Big(\frac{1}{n}\Big)\,.$$

Вторая из этих формул верна, если конечен момент $\alpha_{2k}$. Отсюда следует, что

$$D\alpha_k^* = O\Big(\frac{1}{n}\Big)\,.$$

**2. Состоятельность** оценки. Естественно потребовать, чтобы оценка $\theta_n^*$ при $n \to \infty$ приближалась к оцениваемому параметру. Оценка $\theta_n^*$ называется состоятельной, если

$$\theta_n^* \xrightarrow{P} \theta \,.$$

**Теорема 3.7.** *Если при $n \to \infty$ $M\theta_n^* \to \theta$ и $D\theta_n^* \to 0$, тогда оценка $\theta_n^*$ состоятельна.*

*Доказательство.* По неравенству Чебышева

$$\mathbf{P}\{|\theta_n^* - M\theta_n^*| > \varepsilon\} \leqslant \frac{D\theta_n^*}{\varepsilon^2} \to 0 \,.$$

Воспользуемся неравенством $|\theta_n^* - \theta| \leqslant |\theta_n^* - M\theta_n^*| + |M\theta_n^* - \theta|$. Так как

$$(\forall \varepsilon > 0)(\exists n_0)(\forall n \geqslant n_0 \Rightarrow |M\theta_n^* - \theta| \leqslant \varepsilon/2),$$

из $|\theta_n^* - \theta| > \varepsilon$ следует, что $|\theta_n^* - M\theta_n^*| > \varepsilon/2$. Поэтому

$$\mathbf{P}\{|\theta_n^* - \theta| > \varepsilon\} \leqslant \mathbf{P}\{|\theta_n^* - M\theta_n^*| > \varepsilon/2\} \to 0 \,,$$

то есть $\theta_n^* \xrightarrow{P} \theta$. □

**Теорема 3.8.** *Начальные и центральные моменты $m_k^*$ и $\alpha_k^*$, $k = 1, 2 \ldots$, являются состоятельными оценками начальных и центральных моментов $m_k$ и $\alpha_k$ при условии существования и конечности $m_{2k}$ и $\alpha_{2k}$.*

**3. Эффективность.** Мерой рассеяния является дисперсия. Естественно вопрос о нахождении оценок с наименьшей дисперсией.

Пусть имеется семейство абсолютно непрерывных одномерных распределений

$$\mathcal{F} = \{p(u, \theta)\} \,,$$

зависящих от одного параметра $\theta$ и пусть $x_1, \ldots, x_n$ – выборка из распределения с плотностью $p(u, \theta)$ с неизвестным параметром $\theta$.

Введем совместную плотность распределения

$$f(u, \theta) = \prod_{i=1}^{n} p(u_i, \theta), \quad u = (u_1, \ldots, u_n) \,.$$

Ясно, что

$$\int\limits_{\mathbf{R}^n} f(u, \theta) du = 1 \,. \qquad (*)$$

Пусть $\theta_n^* = \varphi(x_1, \ldots, x_n, \theta) = \varphi(x_1, \ldots, x_n)$ – оценка параметра $\theta$. Положим

$$g(\theta) \stackrel{\text{def}}{=\!=} M\theta^* = \int\limits_{\mathbf{R}^n} \varphi(u) f(u, \theta) du \,. \qquad (**)$$

**Определение 42.** Информацией Фишера относительно семейства плотностей $\{f(u, \theta)\}$ называется величина

$$J(\theta) = \int\limits_{\mathbf{R}^n} \Big(\frac{\partial \ln f(u, \theta)}{\partial \theta}\Big)^2 f(u, \theta) dx \,.$$

**Теорема 3.9** (неравенство Рао – Крамера). *Если семейство плотностей* $\{f(u,\theta)\}$, $u = (u_1, \dots, u_n)$, *и оценка* $\theta_n^* = \varphi(x_1, \dots, x_n)$ *таковы, что интегралы в равенствах (*) и (**) можно дифференцировать по параметру* $\theta$, *то имеет место неравенство*

$$D\theta_n^* \geqslant \frac{[g'(\theta)]^2}{J(\theta)}.$$

*Знак равенства при фиксированном* $\theta$ *достигается тогда и только тогда, когда существует число* $K$, *не зависящее от* $u_1, \dots, u_n$, *но может быть зависящее от* $\theta$ *такое, что*

$$\frac{\partial \ln f}{\partial \theta} = K(\varphi - g(\theta)).$$

*Доказательство.* Согласно предположениям теоремы

$$\int\limits_{\mathbf{R}^n} \frac{\partial f(u,\theta)}{\partial \theta} du = 0 \quad \text{и} \quad g'(\theta) = \int\limits_{\mathbf{R}^n} \varphi(u) \frac{\partial f(u,\theta)}{\partial \theta} du.$$

Отсюда следует, что

$$\int\limits_{\mathbf{R}^n} f(u,\theta) \frac{\partial \ln f(u,\theta)}{\partial \theta} du = 0 \quad \text{и} \quad g'(\theta) = \int\limits_{\mathbf{R}^n} \varphi(u) f(u,\theta) \frac{\partial \ln f(u,\theta)}{\partial \theta} du.$$

Умножим первое равенство на $g(\theta)$ и вычтем результат из второго равенства

$$g'(\theta) = \int\limits_{\mathbf{R}^n} [\varphi(u) - g(\theta)] f(u,\theta) \frac{\partial \ln f(u,\theta)}{\partial \theta} du.$$

Положим

$$\varphi_1 = (\varphi - g)\sqrt{f} \quad \text{и} \quad \varphi_2 = \frac{\partial \ln f}{\partial \theta} \sqrt{f}.$$

С помощью неравенства Коши–Буняковского получим неравенство Рао–Крамера

$$[g'(\theta)]^2 = \Big( \int\limits_{\mathbf{R}^n} (\varphi - g) \frac{\partial \ln f}{\partial \theta} f du \Big)^2 \leqslant \int\limits_{\mathbf{R}^n} \Big(\frac{\partial \ln f}{\partial \theta}\Big)^2 f du \int\limits_{\mathbf{R}^n} (\varphi - g)^2 f du \leqslant J(\theta)\, D\theta_n^*.$$

Знак равенства в неравенстве Коши–Буняковского достигается тогда и только тогда, когда $\varphi_1$ и $\varphi_2$ пропорциональны

$$\frac{\partial \ln f}{\partial \theta} \sqrt{f} = K(\varphi - g(\theta))\sqrt{f}. \ \square$$

*Замечания.*

**1.** Информация Фишера есть не что иное как

$$J(\theta) = M\Big(\frac{\partial \ln f(x,\theta)}{\partial \theta}\Big)^2, \quad \text{здесь } x = (x_1, \dots, x_n) - \text{случайная выборка}.$$

**2.** Так как

$$M\Big(\frac{\partial \ln f(x,\theta)}{\partial \theta}\Big) = \int\limits_{\mathbf{R}^n} \frac{\partial \ln f}{\partial \theta} f du = 0,$$

то

$$J(\theta) = D\Big(\frac{\partial \ln f(x,\theta)}{\partial \theta}\Big).$$

**3.** Так как

$$f(x,\theta)=\prod_{i=1}^{n}p(x_i,\theta)$$

и случайные величины

$$\frac{\partial\ln p(x_i,\theta)}{\partial\theta}$$

независимы и одинаково распределены, поэтому

$$J(\theta)=D\Big(\sum_{i=1}^{n}\frac{\partial\ln p(x_i,\theta)}{\partial\theta}\Big)=\sum_{i=1}^{n}D\Big(\frac{\partial\ln p(x_i,\theta)}{\partial\theta}\Big)=nJ_1(\theta)\,,$$

где

$$J_1(\theta)=\int\limits_{-\infty}^{\infty}\Big(\frac{\partial\ln p(t,\theta)}{\partial\theta}\Big)^2p(t,\theta)dt$$

есть информация Фишера одного наблюдения.

*Замечания.* Для дискретных случайных величин интеграл в информации Фишера заменяется суммой.

Отношение

$$e(\theta_n^*)=\frac{[g'(\theta)]^2}{D(\theta_n^*)J(\theta)}$$

называется эффективностью оценки. Если $e(\theta_n^*)<1$ и $e(\theta_n^*)\to 1$, то говорят об асимптотической эффективности.

*Пример 1.* Пусть $(x_1,\ldots,x_n)$ – выборка из нормально распределенной генеральной совокупности и известной дисперсией $\sigma^2$ и с оцениваемым математическим ожиданием $m$,

$$p(u,m)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(u-m)^2}{2\sigma^2}}$$

$$J_1(m)=\int\limits_{-\infty}^{\infty}\Big(\frac{\partial\ln p(u,m)}{\partial m}\Big)^2p(u,m)du=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\frac{(u-m)^2}{\sigma^4}e^{-\frac{(u-m)^2}{2\sigma^2}}du=\frac{1}{\sigma^2}\,.$$

Для несмещенной оценки математического ожидания неравенство Рао – Крамера принимает вид

$$Dm^*\geqslant\frac{\sigma^2}{n}\,.$$

Если в качестве оценки $m^*$ взять выборочное среднее $m^*=\overline{x}$, то $Mm^*=m$, и $Dm^*=\frac{\sigma^2}{n}$. Поэтому оценка $m^*=\overline{x}$ является эффективной.

*Пример 2.* Пусть $(x_1,\ldots,x_n)$ – выборка из нормально распределенной генеральной совокупности с известным математическим ожиданием $m$ и с оцениваемой дисперсией $\theta$, имеющая плотность

$$p(u,m)=\frac{1}{\sqrt{2\pi\theta}}\,e^{-\frac{(u-m)^2}{2\theta}}$$

$$J_1(\theta)=\frac{1}{\sqrt{2\pi\theta}}\int\limits_{-\infty}^{\infty}\Big(\frac{(u-m)^2}{2\theta^2}-\frac{1}{2\theta}\Big)^2e^{-\frac{(u-m)^2}{2\theta}}du=\frac{1}{2\theta^2}.$$

Поэтому для несмещенной оценки имеет место неравенство

$$D\theta^*\geqslant\frac{2\theta^2}{n}\,.$$

Рассмотрим три оценки

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2, \;\; \widehat{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, \;\; s_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2.$$

Для нормально распределенной случайной величины имеем

$$\alpha_4 = 4\sigma^4.$$

Поэтому

$$Ms^2 = \frac{n-1}{n}\sigma^2, \;\; Ds^2 = \frac{2(n-1)}{n^2}\sigma^4 = \frac{2(n-1)}{n^2}\theta^2 \;\; \text{(см. теорему 3.6)}.$$

Так что

$$e(s^2) = \frac{n-1}{n} < 1,$$

то оценка $\theta^* = s^2$ не является эффективной, но является асимптотически эффективной.
Оценка $\theta^* = \widehat{s}^2$ является несмещенной оценкой и

$$D\widehat{s}^2 = \frac{2\theta^2}{n-1}\,.$$

Поэтому эффективность этой оценки меньше единицы

$$e(\widehat{s}^2) = \frac{1}{D\widehat{s}^2 nJ_1(\theta)} = \frac{n-1}{n} < 1\,.$$

Для оценки $s_0^2$ выполняются равенства

$$Ms_0^2 = \theta \;\text{ и }\; Ds_0^2 = \frac{2\theta^2}{n},$$

то есть $s_0^2$ является несмещенной и эффективной оценкой дисперсии $\theta$.

*Пример 3.* Пусть выборка $(x_1, \ldots, x_n)$ – биномиально распределена

$$k \to P_N(k) = C_N^k p^k (1-p)^{N-k}$$

с известным $N$ и параметром $\theta = p$.

$$J_1(p) = \sum_{i=0}^{N}\Big(\frac{\partial \ln P_N(i)}{\partial p}\Big)^2 P_N(i) = \sum_{i=0}^{N}\Big(\frac{i}{p} - \frac{N-i}{1-p}\Big)^2 C_N^i p^i (1-p)^{N-i} = \frac{N}{p(1-p)}\,.$$

Для несмещенной оценки неравенство Рао – Крамера принимает вид

$$Dp^* \geqslant \frac{p(1-p)}{nN}\,.$$

Положим

$$p^* = \frac{\overline{x}}{N} = \frac{1}{nN}\sum_{i=1}^{n} x^i\,.$$

Тогда

$$Mp^* = p \;\text{ и }\; Dp^* = \frac{p(1-p)}{nN},$$

то есть $p^*$ – эффективная оценка.

*Пример 4.* Пусть выборка $(x_1, \dots, x_n)$ имеет распределение Пуассона с параметром $\lambda$

$$P(k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Тогда

$$J_1(\lambda) = \sum_{i=0}^{\infty} \Big(\frac{\partial \ln P(i)}{\partial \lambda}\Big)^2 P(i) = \sum_{i=0}^{\infty} \Big(\frac{i}{\lambda} - 1\Big)^2 \frac{\lambda^i}{i!} e^{-\lambda} = \frac{1}{\lambda}.$$

Для несмещенной оценки $\lambda^*$ неравенство Рао – Крамера принимает вид

$$D\lambda^* \geqslant \frac{\lambda}{n}.$$

Пусть $\lambda^* = \overline{x}$. Тогда $M\lambda^* = Mx_i = \lambda$ и $D\lambda^* = \frac{\lambda}{n}$. Поэтому оценка $\lambda^* = \overline{x}$ является несмещенной.

## 3.5. Асимптотическая нормальность выборочных моментов

**Определение 43.** Если последовательность случайных величин $\zeta_n$ такова, что при некоторых $A_n$ и $B_n$

$$\lim_{n\to\infty} \mathbf{P}\Big\{\frac{\zeta_n - A_n}{B_n} \leqslant x\Big\} = \Phi(x),$$

то говорят, что последовательность $\{\zeta_n\}$ *асимптотически нормальна* с параметрами $(A_n, B_n)$ или просто $(A_n, B_n)$-*асимптотически нормальна.*

**Теорема 3.10.** *Если конечен теоретический момент $m_{2k}$, то при $n \to \infty$ выборочный момент $m_k^*$ асимптотически нормален с параметрами $\big(m_k, \frac{m_{2k} - m_k^2}{n}\big)$.*

*Доказательство.* Случайная величина

$$nm_k^* = \sum_{i=1}^{n} x_i^k$$

является суммой независимых одинаково распределенных случайных величин. Если конечен теоретический момент $m_{2k} = Mx_i^{2k}$, то к этой сумме можно применить центральную предельную теорему. Так как $Mx_i^k = m_k$ и $Dx_i^k = m_{2k} - m_k^2$, то случайная величина

$$\frac{nm_k^* - nm_k}{\sqrt{(m_{2k} - m_k^2)n}} = \frac{m_k^* - m_k}{\sqrt{\frac{m_{2k} - m_k^2}{n}}}$$

асимптотически нормальна с параметрами $(0, 1)$. □

Асимптотическая нормальность центральных выборочных моментов $\alpha_k^*$ доказывается сложнее. Ограничимся случаем $k = 2$.

**Теорема 3.11.** *Если конечен теоретический момент $\alpha_4$, то при $n \to \infty$ выборочная дисперсия $\alpha_2^* = s^2$ асимптотически нормальна с параметрами $\big(\alpha_2, \frac{\alpha_4 - \alpha_2^2}{n}\big)$.*

*Доказательство.* Положим $y_k = x_k - Mx_k$, $k = 1, \dots, n$. Перепишем формулу для $s^2$ в виде

$$s^2 = \frac{1}{n}\sum_{k=1}^{n} (y_k - \overline{y})^2 = \frac{1}{n}\sum_{k=1}^{n} y_k^2 - \overline{y}^2,$$

где $\overline{y} = (y_1 + \ldots, y_n)/n$. Случайную величину

$$\zeta_n = \sqrt{n}\Big(\frac{1}{n}\sum_{k=1}^{n} y_k^2 - \overline{y}^2 - My_1^2\Big)$$

представим в виде $\xi_n + \eta_n$, где

$$\xi_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}(y_k^2 - My_k^2), \quad \eta_n = -\overline{y}^2\sqrt{n}.$$

Напомним, что $y_k^2$, $k = 1, \ldots, n$ – независимые одинаково распределенные случайные величины. Так как

$$M|\eta_n| = \sqrt{n}M\overline{y}^2 = \frac{Dx_1}{\sqrt{n}} = \frac{\alpha_2}{\sqrt{n}}\,,$$

из неравенства Чебышева следует, что для любого $\varepsilon > 0$

$$\mathbf{P}\{|\eta_n| > \varepsilon\} \leqslant \frac{M|\eta_n|}{\varepsilon} = \frac{\alpha_2}{\varepsilon\sqrt{n}} \to 0\,.$$

Отсюда и из теоремы 2.42 следует, что предельное распределение $\zeta_n$ совпадает с предельным распределением $\xi_n$. Согласно центральной предельной теореме случайная величина $\xi_n$ асимптотически нормально распределена с параметрами $(0, \alpha_4 - \alpha_2^2)$. Поэтому центральный выборочный момент $\alpha_2 = s^2$ при $n \to \infty$ асимптотически нормально распределен с параметрами $\big(\alpha_2, \frac{\alpha_4 - \alpha_2^2}{n}\big)$. $\square$

**Теорема 3.12.** *Допустим, что* $\{\xi_n\}_{n=1}^{\infty}$ *— асимптотически нормально распределенная последовательность случайных величин с параметрами* $\big(m_n, \sigma_n\big)$*, где* $M\xi_n = m_n \to m$*,* $D\xi_n = \sigma_n^2 \to 0$ *и* $\frac{m - m_n}{\sigma_n} \to 0$*. Пусть функция* $g(x)$ *определена в окрестности точки* $m$ *и имеет в этой окрестности непрерывные частные производные второго порядка и* $g'(m) \neq 0$*. Тогда последовательность случайных величин* $\eta_n = g(\xi_n)$ *асимптотически нормальна с параметрами* $\big(g(m), |g'(m)|\sigma_n\big)$*, то есть при* $n \to \infty$

$$\mathbf{P}\Big\{\frac{(\eta_n - g(m))}{|g'(m)|\sigma} < x\Big\} \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}du\,.$$

*Доказательство.* При $|x - m| < \varepsilon$ имеем

$$g(x) = g(m) + g'(m)(x - m) + r(x)(x - m)^2\,,$$

где $|r(x)| \leqslant K$, $K < \infty$. Существует $n_0$ такое, что для всех $n \geqslant n_0$ выполняется неравенство $\sigma_n < \varepsilon$. Введем случайное событие

$$A_n = \{\omega : |\xi_n(\omega) - m| < \sigma_n^{1/4}\}$$

Представим $\eta_n$ в виде

$$\eta_n = g(\xi_n)I_{A_n} + g(\xi_n)I_{\overline{A}_n} =$$
$$\big(g(m) + g'(m)(\xi_n - m) + r(\xi_n)(\xi_n - m)^2\big)I_{A_n} + g(\xi_n)I_{\overline{A}_n}\,.$$

Положим $\eta_n = \eta_n' + \delta_n$, где

$$\eta_n' = g(m) + g'(m)(\xi - m)$$
$$\delta_n = -g(m)I_{\overline{A}_n} - g'(m)(\xi_n - m)I_{\overline{A}_n} + r(\xi_n)(\xi_n - m)^2 I_{A_n} + g(\xi_n)I_{\overline{A}_n}\,.$$

Рассмотрим

$$\frac{\eta_n' - g(m)}{|g'(m)|\sigma_n} = \frac{g'(m)}{|g'(m)|}\frac{\xi_n - m}{\sigma_n}$$

Согласно теореме 2.42 последовательность случайных величин

$$\frac{\xi_n - m_n}{\sigma_n} = \frac{\xi_n - m}{\sigma_n} + \frac{m - m_n}{\sigma_n}$$

асимптотически нормальна с параметрами $(0,1)$. Поэтому последовательность $\eta_n'$ асимптотически нормальна с параметрами $(g(m), |g'(m)|\sigma_n)$. Покажем, что $\delta_n/\sigma_n \xrightarrow{P} 0$. Из неравенства Чебышева следует, что

$$\mathbf{P}(\overline{A}_n) \leqslant \frac{\sigma_n^2}{\sqrt{\sigma_n}} = \sigma_n^{3/2} \to 0\,.$$

Поэтому

$$\mathbf{P}\{|g(m)|I_{\overline{A}_n} > \varepsilon\sigma_n\} \leqslant \frac{K\mathbf{P}(\overline{A}_n)}{\varepsilon\sigma_n} \leqslant \frac{K}{\varepsilon}\sqrt{\sigma_n} \to 0$$

Аналогично доказывается, что

$$\mathbf{P}\{|g(\xi_n)|I_{\overline{A}_n} > \varepsilon\sigma_n\} \to 0\,.$$

Далее, применяя неравенство Коши–Буняковского, получим

$$\mathbf{P}\{|g'(m)||\xi_n - m|I_{\overline{A}_n} > \varepsilon\sigma_n\} \leqslant \frac{|g'(m)|M(|\xi_n - m|I_{\overline{A}_n})}{\varepsilon\sigma_n} \leqslant$$
$$\frac{|g'(m)|\sqrt{M(\xi_n - m)^2 M I_{\overline{A}_n}}}{\varepsilon\sigma_n} \leqslant \frac{|g'(m)|\sigma_n\sigma_n^{3/4}}{\varepsilon\sigma_n} = \frac{|g'(m)|\sigma_n^{3/4}}{\varepsilon} \to 0\,.$$

И наконец,

$$\mathbf{P}\{I_{\overline{A}_n}|r(\xi_n)(\xi_n - m)^2| > \varepsilon\sigma_n\} \leqslant \frac{KM(\xi_n - m)^2}{\varepsilon\sigma_n} = \frac{K\sigma_n}{\varepsilon} \to 0\,.$$

Поэтому последовательность случайных величин $\delta_n/\sigma_n$ сходится по вероятности к нулю. По теореме 2.42 последовательность

$$\frac{\eta_n - g(m)}{|g'(m)|\sigma_n} = \frac{\eta_n' - g(m)}{|g'(m)|\sigma_n} + \frac{\delta_n}{|g'(m)|\sigma_n}$$

асимптотически нормальна с параметрами $(0,1)$. □

Приведем здесь формулировку схожего утверждения для последовательности асимптотически нормальных векторов.

**Теорема 3.13.** *Пусть $\xi_n = (\xi_{n1}, \ldots, \xi_{nr})$, $n = 1, 2, \ldots$ – последовательность асимптотически нормальных векторов с моментами*

$$M\xi_{n\ell} = m_{m\ell} \to m_\ell \;\; (\ell = 1, \ldots, r),$$

$$\operatorname{cov}(\xi_{n\ell}, \xi_{nk}) = \frac{\sigma_{\ell k}(n)}{n} \;\; (k, \ell = 1, \ldots, r)\,,$$

*где $(m_{n\ell} - m_\ell)\sqrt{n} \to 0$ и $\sigma_{\ell k}(n) \to \sigma_{\ell k}$ при $n \to \infty$ и пусть функция $g(x) = g(x_1, \ldots, x_r)$ определена в окрестности точки $m = (m_1, \ldots, m_r)$ и имеет в этой окрестности непрерывные частные производные второго порядка. Если $\eta_n = g(\xi_{n1}, \ldots, \xi_{nr})$ и величина $\sigma^2 = \sum_{k,\ell=1}^r g_k g_\ell \sigma_{k\ell}$, где $g_i = \frac{\partial g}{\partial x_i}(m)$, положительна, то при $n \to \infty$*

$$\mathbf{P}\Big\{\frac{\sqrt{n}(\eta_n - g(m))}{\sigma} < x\Big\} \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{\frac{u^2}{2}}\, du\,.$$

## 3.6. Методы нахождения оценок

### Метод моментов

Пусть $(x_1, \dots, x_n)$ – выборка из семейства распределений $F(u, \theta)$, $u \in \mathbf{R}$, зависящего от неизвестного $r$-мерного параметра $\theta = (\theta_1, \dots, \theta_r)$ из множества $\Theta \subset \mathbf{R}^r$. Пусть вектор–функцию $g(x) = (g_1(x), \dots, g_r(x))$ такова, что уравнение

$$m_g(\theta) = t,$$

где $t = (t_1, \dots, t_r)$, $m_g(\theta) = \big(m_{g1}(\theta), \dots, m_{gr}(\theta)\big)$ и $m_{gj}(\theta) = Mg_j(\xi) = \int_{-\infty}^{\infty} g_j(u) dF(u, \theta)$, однозначно и непрерывно разрешимо относительно $\theta$ в области $m_g(\Theta)$ значений $\theta \in \Theta$, то есть $\theta = m_g^{-1}(t)$. Допустим, что вектор

$$\bar{g} = (\bar{g}_1, \dots, \bar{g}_r) = \Big(\frac{1}{n}\sum_{i=1}^n g_1(x_i), \dots, \frac{1}{n}\sum_{i=1}^n g_r(x_i)\Big)$$

принадлежит области $m_g(\Theta)$.

**Определение 44.** Оценка $\theta^* = m_g^{-1}(\bar{g})$ называется оценкой, полученной по методу моментов.

Подобного рода оценки $\theta^*$ состоятельны.

**Теорема 3.14.** *Оценка $\theta^* = (\theta_1^*, \dots, \theta_r^*)$ получаемая как решение уравнения $m_g(\theta) = \bar{g}$ является состоятельной.*

*Доказательство.* Имеем

$$\theta_k^* = m_{gk}^{-1}(\bar{g}_1, \dots, \bar{g}_r)\,,$$

где $\bar{g}_k \xrightarrow{P} m_{gk}$, $k = 1, \dots, r$ и $m_{gk}^{-1}$ – непрерывные функции. По лемме 3.1

$$\theta_k^* = m_{gk}^{-1}(\bar{g}_1, \dots, \bar{g}_r) \xrightarrow{P} m_{gk}^{-1}(m_{g1}, \dots, m_{gr}) = \theta_k\,. \ \square$$

Если в качестве функций $g_k(x)$, $k = 1, \dots, r$, выбрать функции $g_k(x) = x^k$, то наше векторное уравнение превращается в систему уравнений для обычных моментов. Свое название "метод моментов" получил потому, что суть его состоит в приравнивании "теоретических" и эмпирических моментов.

В частном наиболее важном в приложениях случае $r = 2$, $g_1(x) = x$ и $g_2(x) = x^2$ асимптотическая нормальность оценки $\theta^*$ является следствием теоремы 3.13.

**Теорема 3.15.** *Пусть $\eta_n = g(\bar{x}, s^2)$, где функция $g(u_1, u_2)$ определена в окрестности точки $(m, \sigma^2)$ и имеет в этой окрестности непрерывные частные производные второго порядка. Если конечен момент $\alpha_4$ и $\sigma^2 > 0$, то случайная величина $\eta_n$ при $n \to \infty$ асимптотически нормальна с параметрами $(g_0, \sigma_g^2/n)$, где*

$$g_0 = g(m, \sigma^2)$$

*и*

$$\sigma_g^2 = g_{11}\alpha_2 + 2g_{12}\alpha_3 + g_{22}(\alpha_4 - \alpha_2^2), \quad \text{где } g_{ij} = \frac{\partial g}{\partial u_i}(m, \sigma^2)\frac{\partial g}{\partial u_j}(m, \sigma^2) \ \ (i, j = 1, 2)\,.$$

*Напомним, что через $\alpha_k$ обозначается центральный момент порядка $k$.*

*Доказательство.* Согласно теореме 3.6 имеем

$$M\bar{x} = m, \ \ Ms^2 = \sigma^2 + O\Big(\frac{1}{n}\Big),$$

$$D\bar{x} = \frac{\sigma^2}{n}, \ \ Ds^2 = \frac{\alpha_4 - \alpha_2^2}{n} + O\Big(\frac{1}{n^2}\Big)\,.$$

Положим $y_k = x_k - Mx_k = x_k - m$ и перепишем $s^2$ в виде

$$s^2 = \frac{1}{n}\sum_{k=1}^{n}(y_k - \overline{y})^2 = \frac{1}{n}\sum_{k=1}^{n} y_k^2 - \overline{y}^2,$$

где $\overline{y} = (y_1 + \cdots + y_k)/n$. В этих обозначениях

$$\operatorname{cov}(\overline{x}, s^2) = M\Big[\overline{y}\Big(\frac{1}{n}\sum_{k=1}^{n} y_k^2 - \overline{y}^2 - \alpha_2\Big)\Big] =$$

$$\frac{1}{n}\sum_{k=1}^{n} M[\overline{y}y_k^2] - M\overline{y}^3 = \frac{1}{n^2}\sum_{k=1}^{n} My_k^3 - \frac{1}{n^3}\sum_{k=1}^{n} My_k^3 = \frac{n-1}{n^2}\alpha_3\,.$$

Таким образом, условия теоремы 3.13 выполнены. □

*Пример 5.* Рассмотрим семейство показательных распределений с плотность вероятности $\{p(u,\lambda)\} = \lambda e^{-\lambda u}$, $u \geqslant 0$, с неизвестным параметром $\lambda$. Построим оценку для $\lambda$ сначала с функцией $g(x) = x$, а затем с функцией $g(x) = x^2$. Справедливо следующее равенство

$$m_1(\lambda) = Mx_1 = \int_0^\infty u\lambda e^{-\lambda u} du = \frac{1}{\lambda}.$$

Из уравнения $m_1(\lambda) = \overline{x}$ найдем оценку по методу моментов

$$\lambda_1^* = \frac{1}{\overline{x}}\,.$$

Для второй функции $g(x) = x^2$ имеем

$$m_2(\lambda) = Mx_1^2 = \int_0^\infty u^2\lambda e^{-\lambda u} du = \frac{2}{\lambda^2},$$

Решением уравнения $m_2(\lambda) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ является оценка

$$\lambda_2^* = \Big(\frac{1}{2n}\sum_{i=1}^{n} x_i^2\Big)^{-1/2}.$$

Из приведенной выше теоремы 3.14 следует, что обе оценки $\lambda_1^*$ и $\lambda_2^*$ являются состоятельными.

*Пример 6.* Пусть $(x_1, \ldots, x_n)$ – нормально распределенная выборка с неизвестными параметрами $m$ и $\sigma^2$

$$p(u, m, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}\,.$$

Рассмотрим $g_1(x) = x$, $g_2(x) = x^2$. получим следующие уравнения для метода моментов

$$m = \overline{x}, \quad \sigma^2 + \alpha^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2,$$

из которых находим оценки

$$m^* = \overline{x}, \quad (\sigma^2)^* = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\overline{x})^2 = s^2\,.$$

**Метод наибольшего правдоподобия**

Пусть $(x_1, \ldots, x_n)$ — выборка из распределения с плотностью $p(u, \theta)$, $u \in \mathbf{R}$, зависящая, вообще говоря, от векторного параметра $\theta$. Функция

$$L(x_1, \ldots, x_n, \theta) = \prod_{k=1}^{n} p(x_k, \theta)$$

как функция $\theta$ называется функцией правдоподобия. Оценкой наибольшего правдоподобия называется статистика $\theta_n^* = \theta^*(x_1, \ldots, x_n)$ для которой

$$L(x_1, \ldots, x_n, \theta_n^*) = \max_{\theta} L(x_1, \ldots, x_n, \theta) \,. \tag{3.1}$$

Оценка наибольшего правдоподобия определяется тем условием, что вероятность получить данную конкретную выборку $(x_1, \ldots, x_n)$ должна быть наибольшей. На практике вместо $\max L$ ищут максимум $\log L$.

**Теорема 3.16.** *В условиях теоремы Рао–Крамера, если для параметра $\theta$ существует несмещенная и эффективная оценка $\theta^*$, то уравнение правдоподобия имеет единственное решение $\theta^*$.*

*Доказательство.* Если $\theta^*$ эффективная несмещенная оценка, то неравенство Рао–Крамера обращается в равенство. Что возможно только при выполнении условия

$$\frac{d \log L(x_1, \ldots, x_n, \theta)}{d\theta} = k(\theta^* - \theta),$$

где константа $k$ зависит от $\theta$. Уравнение правдоподобия эквивалентно уравнению

$$\theta^* - \theta = 0 \,,$$

которое имеет единственное решение, а именно, статистику $\theta^*(x_1, \ldots, x_n)$, удовлетворяющую (3.1). □

*Пример 7.* Нормальное распределение имеет плотность

$$p(u, m, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(u-m)^2}{2\sigma^2}}, \quad -\infty < m < \infty, \;\; \sigma > 0.$$

Положим $\theta = (m, \sigma^2)$. Функция правдоподобия запишется в виде

$$L(x_1, \ldots, x_n, m, \sigma) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} \exp\Big\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m)^2 \Big\}.$$

Так как $L$ и $\log L$ достигают максимума при одних и тех же $\theta$, для нахождения точки максимума решим систему

$$\frac{\partial \log L}{\partial m} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - m) = 0 \,,$$

$$\frac{\partial \log L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x_i - m)^2 = 0 \,.$$

Получим

$$m^* = \overline{x} \,, \;\; (\sigma^2)^* = s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 \,.$$

Так как второй дифференциал функции $\log L$ в точке $(\overline{x}, (\sigma^2)^*)$ отрицательно определен, то в этой точке действительно достигается максимум этой функции.

*Пример 8.* Для биномиально распределенной выборки объема $n$ имеем

$$L(x_1, \ldots, x_n, p) = p^\nu (1-p)^{n-\nu},$$

где $p = \mathbf{P}(x_i = 1)$, $\mathbf{P}(x_i = 0) = 1 - p$, и $\nu$ – число элементов выборки, равных единице. Поэтому

$$\log L(x_1, \ldots, x_n, p) = \nu \log p + (n - \nu)\log(1-p),$$

и

$$\frac{\partial L}{\partial p} = \frac{\nu}{p} - \frac{n-\nu}{1-p}.$$

Решением уравнения является оценка

$$p^* = \frac{\nu}{n}.$$

Нетрудно проверить, что в этой точке действительно достигается максимум функции $L$.

*Пример 9.* Пусть выборка объема $n$ равномерно распределена на отрезке $[0, \theta]$. В этом случае

$$p(u, \theta) = \begin{cases} \theta^{-1}, & z \in [0, \theta], \\ 0, & z \notin [0, \theta]. \end{cases}$$

Поэтому

$$L(x_1, \ldots, x_n, \theta) = \begin{cases} \theta^{-n}, & \text{если } \ x_i \in [0, \theta] \ \text{ при всех } \ i = 1, \ldots, n. \\ 0, & \text{в противном случае}. \end{cases}$$

Обозначим через $x_{(n)} = \max x_i$. Тогда $L(x_1, \ldots, x_n, \theta) = 0$ при $\theta \in [0, x_{(n)})$, и $L(x_1, \ldots, x_n, \theta) = \theta^{-n}$ при $\theta \in (x_{(n)}, \infty))$. Поэтому максимум функции $L$ достигается в точке $\theta^* = x_{(n)}$.

Пусть $x_1, \ldots, x_n$ – выборка из распределения с плотностью $p(u, \theta)$, зависящей от параметра $\theta$. Приведем без доказательства теорему об асимптотической нормальности оценки наибольшего правдоподобия. Будем предполагать, что

1) Истинное значение параметра $\theta_0$ лежит внутри интервала $(\theta_1, \theta_2)$ и в этом интервале существуют производные $\frac{\partial p}{\partial \theta}$, $\frac{\partial^2 p}{\partial \theta^2}$ и $\frac{\partial^3 p}{\partial \theta^3}$.

2) Интеграл $\int p(u, \theta) dx$ можно два раза дифференцировать под знаком интеграла.

3) Информация Фишера одного наблюдения $J_1(\theta) = \int \left(\frac{\partial \log p}{\partial \theta}\right)^2 p(u) du$ в точке $\theta_0$ положительна и существует мажоранта $H(u)$ семейства функций $\frac{\partial^3 \log p(u,\theta)}{\partial \theta^3}$

$$\left|\frac{\partial^3 \log p(u, \theta)}{\partial \theta^3}\right| \leqslant H(u),$$

удовлетворяющая соотношению

$$M_\theta H(\xi) = \int H(x) p(x, \theta) dx \leqslant M,$$

где $M$ не зависит от $\theta$.

Все рассмотренные ранее распределения удовлетворяют условиям 1, 2 и 3.

**Теорема 3.17.** *При выполнении условий 1, 2 и 3 уравнение правдоподобия* $\frac{\partial L(x_1, \ldots, x_n, \theta)}{\partial \theta} = 0$ *имеет решение* $\theta_n^*$, *которое при* $n \to \infty$ *сходится по вероятности к* $\theta_0$. *Эта оценка наибольшего правдоподобия асимптотически нормальна с параметрами* $(\theta_0, \frac{1}{\sqrt{J_1(\theta_0) n}})$ *и асимптотически эффективна. Точнее,* $\theta_n^* \xrightarrow{P} \theta_0$, $e(\theta_n^*) \to 1$ *и*

$$\lim_{n \to \infty} \mathbf{P}\left\{ \left(\theta_n^* - \theta_0\right) \sqrt{J_1(\theta_0) n} \leqslant x \right\} = \Phi(x).$$

*Пример 10.* Рассмотрим Г-распределение с плотностью

$$p(u,\alpha)=\frac{\alpha^\lambda}{\Gamma(\lambda)}u^{\lambda-1}e^{-\alpha u},\quad u\geqslant 0,\quad \alpha>0\,,$$

в случае когда параметр $\lambda$ известен. Имеем

$$L(x_1,\dots,x_n,\alpha)=\lambda n\log\alpha-n\log\Gamma(\lambda)+(\lambda-1)\sum_{i=1}^n\log x_i-\alpha\sum_{i=1}^n x_i\,,$$

$$\frac{\partial L}{\partial\alpha}=\frac{\lambda n}{\alpha}-\overline{x}n,\quad \alpha_n^*=\frac{\lambda}{\overline{x}}\,.$$

Согласно доказанной теореме полученная оценка $\alpha_n^*$ является состоятельной, асимптотически эффективной и асимптотически нормальной.

## 3.7. Доверительные интервалы

Пусть $x_1,\dots,x_n$ – выборка из распределения с плотностью $p(u,\theta)$, зависящего от параметра $\theta$. Интервал $(\theta_1,\theta_2)$ со случайными границами $\theta_i(x_1,\dots,x_n)$ называется *доверительным интервалом* для параметра $\theta$ с *надежностью* $\gamma=1-\alpha$, где $\alpha$ – *уровень значимости*, если $\mathbf{P}\{\theta_1<\theta<\theta_2\}=\gamma$.

Пусть $X(x_1,\dots,x_n)$ – некоторая статистика с функцией распределения $F(x,\theta)=\mathbf{P}\{X\leqslant x\}$. Решение $\widehat{x}_\gamma=\widehat{x}_\gamma(\theta)$ уравнения $F(x,\theta)=1-\gamma$ называется *квантилью.* Пусть $\alpha_1$ и $\alpha_2$ – *разбиение* $\alpha$, то есть $\alpha_1$, $\alpha_2>0$, $\alpha_1+\alpha_2=\alpha$, и пусть $\widehat{x}_{\alpha_1}$, $\widehat{x}_{1-\alpha_2}$ квантили порядка $\alpha_1$ и $1-\alpha_2$. Тогда неравенство

$$\widehat{x}_{1-\alpha_2}\leqslant X\leqslant\widehat{x}_{\alpha_1} \tag{3.2}$$

выполняется с вероятностью $\gamma$, так как

$$\mathbf{P}\{\widehat{x}_{1-\alpha_2}\leqslant X\leqslant\widehat{x}_{\alpha_1}\}=F(\widehat{x}_{\alpha_1},\theta)-F(\widehat{x}_{1-\alpha_2},\theta)=1-\alpha_1-\alpha_2=1-\alpha=\gamma\,.$$

Предположим, что $F(x,\theta)$ есть убывающая (возрастающая) функция от $\theta$, тогда квантиль $\widehat{x}_\alpha$ является убывающей (возрастающей) функцией параметра $\theta$. Монотонность по $\theta$ позволяет разрешить неравенство (3.2)

$$\theta_1(x_1,\dots,x_n,\alpha_1)\leqslant\theta\leqslant\theta_2(x_1,\dots,x_n,1-\alpha_2)\,.$$

Таким образом, $(\theta_1,\theta_2)$ является доверительным интервалом для параметра $\theta$ с надежностью $\gamma$.

Разбиение $\alpha_1$, $\alpha_2$ можно варьировать. Средняя длина $M[\theta_2-\theta_1]$ доверительного интервала при этом, вообще говоря, меняется. Желательно среди всех доверительных интервалов с надежностью $\gamma$ выбрать тот, который имеет наименьшую среднюю длину. Для симметричных распределений, то есть для распределений с четной плотностью $p(u,\theta)$, оптимальным разбиением, как будет видно ниже, является $\alpha_1=\alpha_2=\alpha/2$. Для $\chi^2$-распределенной случайной величины имеются таблицы выбора разбиения $\alpha$ наилучшим образом.

### Доверительные интервалы для параметров нормального распределения

Пусть $x_1,\dots,x_n$ – выборка из нормального распределения с параметрами $(m,\sigma)$.

**Теорема 3.18.** *а) Выборочное среднее $\overline{x}$ и исправленная выборочная дисперсия $\widehat{s}^2$ взаимно независимы; $\overline{x}$ распределена нормально с параметрами $(m,\sigma/\sqrt{n})$ и $(n-1)\widehat{s}^2/\sigma^2$ имеет распределение $\chi^2$ с $(n-1)$-й степенью свободы.*

*б) Случайная величина $\frac{\overline{x}-m}{\widehat{s}}\sqrt{n}$ подчиняется распределению Стьюдента с $(n-1)$ степенью свободы.*

*Доказательство.* Случайные величины $y_i = (x_i - m)/\sigma$ независимы и распределены нормально с параметрами $(0,1)$. Введем $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$ и $s_1^2 = \frac{1}{n-1}\sum_{i=1}^n (y_i - \overline{y})^2$. Достаточно показать, что случайная величина $(n-1)s_1^2 = (n-1)\widehat{s}^2/\sigma^2$ имеет распределение $\chi^2$ с $(n-1)$ степенью свободы и случайные величины $\overline{y}$ и $s_1^2$ независимы. Пусть $C = \big(c_{ij}\big)$ – ортогональная матрица такая, что $c_{1j} = \frac{1}{\sqrt{n}}$, $j = 1, \dots, n$, и пусть $(z_1, \dots, z_n)^t = C(y_1, \dots, y_n)^t$ – ортогональное преобразование заданное соотношениями

$$z_1 = \frac{1}{\sqrt{n}}\sum_{j=1}^n y_j = \sqrt{n}\overline{y}, \quad z_i = \sum_{j=1}^n c_{ij} y_j, \quad i = 2, \dots, n\,.$$

Так как $\sum_{j=1}^n c_{ij} c_{jk} = \delta_{ik}$, имеем

$$\sum_i z_i^2 = \sum_j y_j^2$$

и

$$Mz_i = \sum_{j=1}^n c_{ij} My_j = 0, \quad M[z_i z_j] = \sum_{j=1} c_{ij} c_{jk} M y_j^2 = \sum_{j=1} c_{ij} c_{jk} = 0, \quad \text{если } \; i \neq k\,,$$

$$Dz_i = D\big(\sum_{j=1}^n c_{ij} y_j\big) = \sum_{j=1}^n c_{ij}^2 D y_j = \sum_{j=1}^n c_{ij^2} = 1\,.$$

Таким образом, $z_1, \dots, z_n$ распределены по нормальному закону с параметрами $(0,1)$ и некоррелированы, а значит, $z_1, \dots, z_n$ независимы. Учитывая, что $z_1 = \sqrt{n}\,\overline{y}$, случайная величина

$$(n-1)s_1^2 = \sum_{i=1}^n (y_i - \overline{y})^2 = \sum_{i=1}^n y_i^2 - n\overline{y}^2 = \sum_{j=1}^n z_j^2 - z_1^2 = \sum_{j=2}^n z_j^2$$

имеет распределение $\chi^2$ с $(n-1)$ степенью свободы как и $(n-1)\widehat{s}^2/\sigma^2$. Отсюда и выше сказанного следует, что $z_1$ и $z_2^2 + \dots + z_n^2$ независимы (см. замечание на стр. 34). Поэтому $\overline{y} = z_1/\sqrt{n}$ и $s_1^2 = \sum_{j=2}^n z_j^2/(n-1)$ независимы. Для того, чтобы завершить доказательство теоремы заметим, что случайная величина

$$Y = \frac{\overline{x} - m}{\widehat{s}}\sqrt{n} = \frac{\sigma\overline{y}\sqrt{n}}{\sigma s_1} = \frac{\overline{y}\sqrt{n}}{s_1} = z_1\sqrt{\frac{n-1}{\sum_{j=1}^n z_j^2}}$$

является отношением Стьюдента с $(n-1)$ степенью свободы. □

Пусть независимая выборка $x_1, \dots, x_n$ взята из нормального распределения с параметрами $(m, \sigma)$.

**а) Доверительный интервал для параметра $m$ при известном $\sigma$.** В качестве оценки возьмем среднее арифметическое $\overline{x}$. Статистика $\overline{x}$ отвечает всем требованиям, предъявляемым к оценкам и имеет нормальное распределение с параметрами $(m, \frac{\sigma}{\sqrt{n}})$. Обозначим через $\widehat{u}_\gamma$ квантиль нормального распределения, то есть решение уравнения

$$1 - \Phi(\widehat{u}_\gamma) = \gamma\,.$$

Пусть $\alpha = \alpha_1 + \alpha_2$. Так как

$$\Phi(\widehat{u}_{1-\gamma}) = 1 - \Phi(\widehat{u}_\gamma) = \frac{1}{\sqrt{2\pi}}\int_{\widehat{u}_\gamma}^{\infty} e^{-\frac{u^2}{2}}du = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\widehat{u}_\gamma} e^{-\frac{u^2}{2}}du = \Phi(-\widehat{u}_\gamma),$$

имеем $\widehat{u}_{1-\gamma} = -\widehat{u}_\gamma$. Статистика $Y = \frac{\overline{x}-m}{\sigma}\sqrt{n}$ имеет нормальное распределение с параметрами $(0,1)$. Поэтому неравенство $-\widehat{u}_\gamma \leqslant Y \leqslant \widehat{u}_\gamma$, то есть

$$m - \widehat{u}_{\alpha_2}\frac{\sigma}{\sqrt{n}} \leqslant \overline{x} \leqslant m + \widehat{u}_{\alpha_1}\frac{\sigma}{\sqrt{n}}, \tag{3.3}$$

выполняется с вероятностью $1-\alpha$. Из (3.3) находим доверительный интервал для параметра $m$

$$\overline{x}-\widehat{u}_{\alpha_1}\frac{\sigma}{\sqrt{n}}\leqslant m\leqslant \overline{x}+\widehat{u}_{\alpha_2}\frac{\sigma}{\sqrt{n}}\,.$$

Длина доверительного интервала равна $(\widehat{u}_{\alpha_1}+\widehat{u}_{\alpha_2})\frac{\sigma}{\sqrt{n}}$. Покажем, что эта длина будет наименьшей, если $\alpha_1=\alpha_2=\alpha/2$. Пусть, например, $\alpha_2>\alpha_1$ и выберем $\Delta>0$ так, чтобы $\alpha_2-\Delta>\alpha_1+\Delta$; тогда $\widehat{u}_{\alpha_1}>\widehat{u}_{\alpha_1+\Delta}>\widehat{u}_{\alpha_2-\Delta}>\widehat{u}_{\alpha_2}$. Из неравенств

$$\Delta=1-\Phi(\widehat{u}_{\alpha_2})-\big(1-\Phi(\widehat{u}_{\alpha_2-\Delta})\big)=$$
$$\frac{1}{\sqrt{2\pi}}\int\limits_{\widehat{u}_{\alpha_2}}^{\widehat{u}_{\alpha_2-\Delta}}e^{-\frac{x^2}{2}}dx\geqslant\frac{1}{\sqrt{2\pi}}e^{-\frac{\widehat{u}^2_{\alpha_2-\Delta}}{2}}\big(\widehat{u}_{\alpha_2-\Delta}-\widehat{u}_{\alpha_2}\big)\,,$$

$$\Delta=1-\Phi(\widehat{u}_{\alpha_1+\Delta})-\big(1-\Phi(\widehat{u}_{\alpha_1})\big)=$$
$$\frac{1}{\sqrt{2\pi}}\int\limits_{\widehat{u}_{\alpha_1+\Delta}}^{\widehat{u}_{\alpha_1}}e^{-\frac{x^2}{2}}dx\leqslant\frac{1}{\sqrt{2\pi}}e^{-\frac{\widehat{u}^2_{\alpha_1+\Delta}}{2}}\big(\widehat{u}_{\alpha_1}-\widehat{u}_{\alpha_1+\Delta}\big)\,,$$

следует, что

$$\widehat{u}_{\alpha_2-\Delta}-\widehat{u}_{\alpha_2}\leqslant\Delta\sqrt{2\pi}e^{\frac{\widehat{u}^2_{\alpha_2-\Delta}}{2}}<\Delta\sqrt{2\pi}e^{\frac{\widehat{u}^2_{\alpha_1+\Delta}}{2}}\leqslant\widehat{u}_{\alpha_1}-\widehat{u}_{\alpha_1+\Delta}\,,$$

то есть

$$\widehat{u}_{\alpha_2-\Delta}+\widehat{u}_{\alpha_1+\Delta}<\widehat{u}_{\alpha_1}+\widehat{u}_{\alpha_2}\,.$$

Таким образом, при $\alpha_1\neq\alpha_2$ доверительный интервал не имеет наименьшей длины.

**б) Доверительный интервал для параметра $m$ при неизвестном $\sigma$.** Для построения доверительного интервала воспользуемся отношением Стьюдента. Пусть $S_n(t)$ – функция распределения Стьюдента с $n$ степенями свободы. Обозначим через $\widehat{t}_\alpha(n)$ квантиль распределения $S_n(t)$, то есть решение уравнения

$$S_n(t)=1-\alpha\,.$$

Так как плотность распределения Стьюдента четная функция, то $t_{1-\gamma}(n)=-t_\gamma(n)$ и при построении доверительного интервала как и в пункте а) следует брать $\alpha_1=\alpha_2=\alpha/2$. По теореме 3.18 неравенство

$$-\frac{s}{\sqrt{n}}\widehat{t}_{\alpha/2}(n-1)\leqslant\overline{x}-m\leqslant\frac{s}{\sqrt{n}}\widehat{t}_{\alpha/2}(n-1)$$

выполняется с вероятностью $1-\alpha$. Отсюда находим доверительный интервал для параметра $m$

$$\overline{x}-\frac{s}{\sqrt{n}}\widehat{t}_{\alpha/2}(n-1)\leqslant m\leqslant\overline{x}+\frac{s}{\sqrt{n}}\widehat{t}_{\alpha/2}(n-1)$$

**в) Доверительный интервал для параметра $\sigma$ при известном $m$.** Воспользуемся статистикой

$$\frac{\eta}{\sigma^2}=\sum_{k=1}^{n}\Big(\frac{x_k-m}{\sigma}\Big)^2\,,$$

имеющей $\chi^2$-распределение с $n$ степенями свободы. Обозначим через $K_n(x)$ функцию распределения случайной величины $\eta/\sigma^2$ и через $\widehat{k}_\alpha(n)$ – квантиль $K_n(x)$, то есть корень уравнения

$$K_n(x)=1-\alpha\,.$$

Пусть $\alpha=\alpha_1+\alpha_2$. Тогда неравенство

$$\widehat{k}_{1-\alpha_1}(n)\leqslant\frac{\eta}{\sigma^2}\leqslant\widehat{k}_{\alpha_2}(n)$$

выполняется с вероятностью $1-\alpha$. Разрешая это неравенство относительно $\sigma$ находим доверительный интервал для параметра $\sigma$

$$\sqrt{\frac{\eta}{\widehat{k}_{\alpha_2}(n)}} \leqslant \sigma \leqslant \sqrt{\frac{\eta}{\widehat{k}_{1-\alpha_1}(n)}}$$

Можно доказать, что этот интервал имеет наименьшую среднюю длину среди всех доверительных интервалов этого вида с доверительной вероятностью $1-\alpha$, если $\alpha_1$ и $\alpha_2$ выбраны так, что плотность $k_n(x) = K_n'(x)$ удовлетворяет соотношению

$$k_n(\widehat{k}_{1-\alpha_1}(n)) = k_n(\widehat{k}_{\alpha_2}(n)).$$

**г) Доверительный интервал для параметра $\sigma$ при неизвестном $m$.** Воспользуемся теоремой 3.18, согласно которой статистика $\frac{s^2}{\sigma^2}(n-1)$ имеет $\chi^2$-распределение с $(n-1)$-й степенью свободы. Как и в пункте в) это приводит к доверительному интервалу

$$s\sqrt{\frac{n-1}{\widehat{k}_{\alpha_2}(n-1)}} \leqslant \sigma \leqslant s\sqrt{\frac{n-1}{\widehat{k}_{1-\alpha_1}(n-1)}}$$

с доверительной вероятностью $1-\alpha$ $(\alpha = \alpha_1 + \alpha_2)$.

### Общий подход к доверительному оцениванию

При большом объеме выборки оценки метода моментов и наибольшего правдоподобия, как правило, асимптотически нормальны. Остановимся подробно на рассмотрении оценки, полученной методом наибольшего правдоподобия. В условиях теоремы 3.17 оценка $\theta_n^*$ асимптотически нормальна с параметрами $\left(\theta, \frac{1}{n\,J_1(\theta)}\right)$, то есть

$$\mathbf{P}\left\{\left(\theta_n^* - \theta_0\right)\sqrt{J_1(\theta_0)n} \leqslant x\right\} \to \Phi(x)\,.$$

В случае, когда $\sigma^2(\theta) = \frac{1}{J_1(\theta)}$ не зависит от $\theta$ и равна $\sigma^2$, будем считать, что $\theta_n^*$ нормально распределенной с параметрами $(\theta, \sigma^2/n)$. При этом допущении доверительный интервал имеет вид

$$\theta_n^* - \widehat{u}_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leqslant \theta \leqslant \theta_n^* + \widehat{u}_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

с приблизительной надежностью $\gamma = 1-\alpha$, где $\widehat{u}_\alpha$ – квантиль порядка $\alpha$ нормального распределения с параметрами $(0,1)$.

Если $\sigma^2(\theta)$ зависит от $\theta$, будем искать дифференцируемую функцию $\eta = g(\theta)$ такую, что

1) Оценка $\eta_n^* = g(\theta_n^*)$ параметра $g(\theta)$ асимптотически нормальна с параметрами $\left(\theta, \frac{[g'(\theta)\sigma(\theta)]^2}{n}\right)$, причем $g'(\theta)\sigma(\theta)$ не зависит от неизвестного параметра $\theta$.

2) Функция $g$ строго монотонна на множестве допустимых значений параметра $\theta$.

Пусть $g(\theta)$ такова, что $g'(\theta)\sigma(\theta) = 1$, то есть

$$g(\theta) = \int \frac{d\theta}{\sigma(\theta)}\,.$$

В этом случае $g(\theta_n^*)$ асимптотически нормальна с параметрами $(\theta, 1/\sqrt{n})$. Поэтому неравенство

$$\eta_n^* - \widehat{u}_{\alpha/2}\frac{1}{\sqrt{n}} \leqslant g(\theta) \leqslant \eta_n^* + \widehat{u}_{\alpha/2}\frac{1}{\sqrt{n}}$$

выполняется с приблизительной надежностью $\gamma = 1 - \alpha$. В результате, интервал

$$g^{-1}\Big(\eta_n^* - \widehat{u}_{\alpha/2}\frac{1}{\sqrt{n}}\Big) \leqslant \theta \leqslant g^{-1}\Big(\eta_n^* + \widehat{u}_{\alpha/2}\frac{1}{\sqrt{n}}\Big)$$

накроет параметр $\theta$ с приблизительной вероятностью $\gamma$.

*Пример 11.* Доверительный интервал для вероятности успеха в схеме Бернулли. Пусть $\mu$ – число успехов с серии $n$ независимых испытаний с неизвестной вероятностью $p$ успеха в одном испытании. Оценка $p^* = \frac{\mu}{n}$ параметра $p$ получена методом наибольшего правдоподобия. Оценка $p^*$ асимптотически нормальна по теореме Муавра–Лапласа с параметрами $\Big(p, \frac{p(1-p)}{n}\Big)$. Выберем функцию $g$, равную

$$g(p) = \int \frac{dp}{\sqrt{p(1-p)}} = 2\arcsin\sqrt{p}\,.$$

Тогда оценка $2\arcsin\sqrt{\frac{\mu}{n}}$ асимптотически нормальна с параметрами $\big(2\arcsin\sqrt{p}, \frac{1}{\sqrt{n}}\big)$. Неравенство

$$2\arcsin\sqrt{\frac{\mu}{n}} - \frac{\widehat{u}_{\alpha/2}}{\sqrt{n}} \leqslant 2\arcsin\sqrt{p} \leqslant 2\arcsin\sqrt{\frac{\mu}{n}} + \frac{\widehat{u}_{\alpha/2}}{\sqrt{n}}$$

выполняется с приблизительной вероятностью $\gamma = 1 - \alpha$. Отсюда получаем приближенный доверительный интервал для параметра $p$

$$\sin^2\Big(\arcsin\sqrt{p^*} - \frac{\widehat{u}_{\alpha/2}}{2\sqrt{n}}\Big) \leqslant p \leqslant \sin^2\Big(\arcsin\sqrt{p^*} + \frac{\widehat{u}_{\alpha/2}}{2\sqrt{n}}\Big)\,.$$

## 3.8. Непараметрические критерии проверки гипотез. Критерий значимости

Предположим, что гипотетическое распределение полностью задано заранее. И необходимо проверить статистическую гипотезу, что наша выборка получена из генеральной совокупности и этим распределением. Если гипотеза верна, то эмпирическая функция распределения $F_n^*(x)$ является приближением в каком-то смысле к функции $F(x)$ при $n \to \infty$. Предположим, что построена мера отклонения $F_n^*(x)$ от $F(x)$ – статистика $S$, которая является функцией от выборки и при справедливости основной гипотезы распределение этой статистики известно точно или асимптотически точно. Выберем число $S_0$ так, чтобы $\mathbf{P}\{S > S_0\} = \alpha$. Если $\alpha$ мало, то можем считать *практически достоверным фактом*, что событие с вероятностью $\alpha$ не произойдет в единичном опыте. По выборочным значениям вычислим величину $S$. Если окажется, что $S < S_0$, то гипотезу признаем разумной интерпретацией имеющихся данных. В противном случае гипотезу отвергаем. Описанный критерий называется *критерием значимости или согласия*. Параметр $\alpha$ называется уровнем значимости. Подчеркнем особо, что проверка гипотезы не есть доказательство гипотезы.

Одним из наиболее известных критериев согласия является $\chi^2$-критерий Пирсона.

### Критерий $\chi^2$ в случае полностью определенного гипотетического распределения ($\chi^2$-критерий Пирсона)

Рассмотрим гипотезу $H$, состоящую в том, что данные $x_1, \ldots, x_n$ образуют выборку из n значений случайной величины с данной функцией распределения $F = F_\xi$. Определим меру расхождения теоретического и гипотетического распределений. Множество значений случайной величины $\xi$ разобьем на $r$ промежутков $I_1, I_2, \ldots, I_r$. По известной функции $F$ вычислим вероятности событий $\mathbf{P}\{\xi \in I_i\} = p_i$. Через $\nu_i$ обозначим количество выборочных значений, попадающих в промежуток $I_i$.

Ясно, что $\sum_{i=1}^r \nu_i = n$. Относительная частота попадания значения выборки в промежуток $I_i$ равна $\frac{\nu_i}{n}$. За меру расхождения между выборочным распределением и гипотетическим примем

$$\sum_{i=1}^r c_i \Big(\frac{\nu_i}{n} - p_i\Big)^2 .$$

Пирсон предложил $c_i$ принять равным $n/p_i$. Тогда статистика, на основе которой строится критерий

$$\widehat{\chi}_n^2 = \sum_{i=1}^r \frac{(\nu_i - np_i)^2}{np_i} = \sum_{i=1}^r \frac{\nu^2}{np_i} - n\,,$$

просто выражается через эмпирические частоты $\nu_i$ и гипотетические частоты $np_i$.

**Теорема 3.19.** *При справедливости гипотезы $H$ распределение $\widehat{\chi}_n^2$ при $n \to \infty$ слабо сходится к $\chi_n^2$ распределению с $(r-1)$-й степенью свободы.*

*Доказательство.* Если гипотеза $H$ верна, то случайные величины $\nu_1, \dots, \nu_r$ имеют полиномиальное распределение

$$P\{\nu_i = n_i,\ i = 1, \dots, r\} = \frac{n!}{n_1! \dots n_r!} p_1^{n_1} \dots p_r^{n_r}\,, \quad n_1 + \dots + n_r = n\,.$$

Из теоремы 2.38 следует, что случайный вектор $\eta^{(n)} = (\eta_1^{(n)}, \dots, \eta_r^{(n)})$ с компонентами

$$\eta_k^{(n)} = \frac{\nu_k - np_k}{\sqrt{np_k}}$$

сходится при $n \to \infty$ к нормальному распределению с нулевыми средними и матрицей ковариации

$$\| \operatorname{cov}(\eta_k, \eta_\ell) \| = \| \delta_{k\ell} - \sqrt{p_k p_\ell} \|\,.$$

Пусть $\eta_k$ обозначают компоненты случайные величины с предельным нормальным распределением. Характеристическая функция случайного вектора $\eta$ равна

$$e^{-\frac{1}{2} R_\eta(t)}\,,$$

где

$$R_\eta(t) = \sum_{k=1}^r t_k^2 - \sum_{k,\ell} t_k t_\ell \sqrt{p_k p_\ell} = \sum_{k=1}^r t_k^2 - \Big(\sum_{k=1}^r \sqrt{p_k} t_k\Big)^2 .$$

Пусть $C$ – ортогональная матрица вида

$$\begin{Vmatrix} \sqrt{p_1} & c_{21} & \cdots & c_{r1} \\ \sqrt{p_2} & c_{22} & \cdots & c_{r2} \\ \dots & \dots & \dots & \dots \\ \sqrt{p_r} & c_{2r} & \cdots & c_{rr} \end{Vmatrix}$$

Положим $(\xi_1 \dots, \xi_r) = \xi = \eta\, C = (\eta_1, \dots, \eta_r)\, C$ и $(u_1, \dots, u_r) = u = t\, C = (t_1, \dots, t_r)\, C$. Характеристические функции $f_\xi$ и $f_\eta$ связаны соотношением $f_\xi(u) = f_\eta(u\, C')$. Поэтому $f_\xi(u) = e^{-R_\xi(u)/2}$, где

$$R_\xi(u) = R_\eta(u\, C') = \sum_{k=1}^r t_k^2 - u_1^2 = \sum_{k=1}^r u_k^2 - u_1^2 = \sum_{k=2}^r u_k^2\,,$$

т.е. $D\xi_1 = 0$ и случайные величины $\xi_2, \dots, \xi_r$ распределены нормально с параметрами $(0,1)$ и независимы. Из предельной теоремы теории меры и интеграла и теоремы Хелли (формулировки которых приведена ниже) следует, что последовательность распределений случайных величин $\widehat{\chi}_n^2 = \eta_{n1}^2 + \dots + \eta_{nr}^2$

сходится к распределению случайной величины $\eta_1^2+\dots+\eta_r^2$. В силу ортогональности матрицы $C$ имеет место равенство $\eta_1^2+\dots+\eta_r^2=\xi_1^2+\dots+\xi_r^2=\xi_2^2+\dots+\xi_r^2$, учитывая, что $\xi_1=0$ п.н. Утверждение теоремы следует теперь из определения $\chi^2$- распределения. $\square$

На практике задаются уровнем значимости $\alpha$. Согласно теореме 3.19 при больших $n$ с вероятностью, приближенно равной $\alpha$, выполняется неравенство

$$\widehat{\chi}_n^2=\sum_{k=1}^{r}\frac{(\nu_k-np_k)^2}{np_k}\geqslant k_\alpha(r-1),$$

где $k_\alpha(r-1)$ – $\alpha$-квантиль $\chi^2$-распределения с $(r-1)$-й степенью свободы. Если гипотеза $H$ верна, то практически невозможно встретить в единичной выборке значение $\widehat{\chi}_n^2$, превышающее $k_\alpha(r-1)$. Поэтому основная гипотеза считается принятой или выборка совместима с гипотезой, если $\widehat{\chi}_n^2<k_\alpha(r-1)$ и отвергнутой, если выполняется противоположное неравенство.

Выбор точек $z_1,\dots,z_{r-1}$ на практике осуществляется таким образом, чтобы удовлетворить неравенствам

$$np_k\geqslant 10,\quad \nu_k\geqslant 10\,.$$

*Пример 12.* В последовательности $n$ независимых испытаний событие $A$ произошло $\nu$ раз. Совместимо ли это с гипотезой, что событие $A$ имеет заданную вероятность?

Каждую компоненту выборки будем рассматривать как индикатор события $A$. Гипотеза $H$ формулируется следующим образом: вероятности событий $A$ и $\overline{A}$ равны $p$ и $q$ соответственно. Построим статистику Пирсона

$$\widehat{\chi}_n^2=\frac{(\nu-np)^2}{np}+\frac{(n-\nu-nq)^2}{nq}=\frac{\nu^2}{np}+\frac{(n-\nu)^2}{nq}-n=\frac{(\nu-np)^2}{npq}\,.$$

Зададимся уровнем значимости $\alpha$ и по таблице значений $\chi^2$-распределения с одной степенью свободы найдем квантиль $k_\alpha(1)$. Сравним значение статистики $\widehat{\chi}_n^2$ с квантилью $k_\alpha(1)$. Если $\widehat{\chi}_n^2<k_\alpha(1)$, то будем считать выборку совместимой с гипотезой, в противном случае отвергнем ее.

Пусть $F_n(x)$, $x=(x_1,\dots,x_m)$ – функция распределения случайного вектора $\xi_n=(\xi_{n1},\dots$ $\dots,\xi_{nm})$ и $F(x)$ – функция распределения случайного вектора $\xi=(\xi_1,\dots,\xi_m)$. Последовательность $F_n(x)$ слабо сходится к $F(x)$ ($F_n(x)\Rightarrow F(x)$), если $F_n(x)\to F(x)$ в каждой точке непрерывности предельной функции.

**Теорема 3.20** (Хелли). *Если $g(x)$ – непрерывная ограниченная функция на пространстве $\mathbf{R}^m$ и $F_n(x)\Rightarrow F(x)$, то*

$$\lim_{n\to\infty}\int\limits_{\mathbf{R}^m}g(x)dF_n(x)=\int\limits_{\mathbf{R}^m}g(x)dF(x)\,.$$

Через $\mathbf{P}_{\xi_n}$ и $\mathbf{P}_\xi$ обозначим вероятностные меры, порожденные функциями распределения $F_n(x)$ и $F(x)$.

**Теорема 3.21.** *Если $F_n(x)\Rightarrow F(x)$, то $\lim_{n\to\infty}\mathbf{P}_{\xi_n}(A)=\mathbf{P}_\xi(A)$ для любого борелевского множества $A$, граница которого имеет $P$-меру нуль.*

## Критерий $\chi^2$ в случае, когда по выборке оцениваются некоторые параметры

Предположим теперь, что функция распределения $F(x,\theta_1,\dots,\theta_s)$ зависит от $s$ параметров. Гипотеза $H$, подлежащая проверке, состоит в том, что выборка $x_1,\dots,x_n$ извлечена из генеральной совокупности с функцией распределения $F$ при некоторых значениях параметра.

Как и в предыдущем пункте будем считать, что выборка разбита на $r$ групп, соответствующих $r$ непересекающимся промежуткам $I_1,\dots,I_r$, Пусть $\nu_i$ количество выборочных значений, попавших в $I_i$.

При истинных значениях параметров $\theta_1, \ldots, \theta_s$ и $p_i(\theta_1, \ldots, \theta_s) = \mathbf{P}\{x_k \in I_i\}$ – вероятности попадания случайной величины $x_k$ в промежуток $I_i$. Параметры $\theta_i$ неизвестны и должны быть оценены по выборке. Но если $\theta_i$ заменить их оценками, то тогда $p_i$ станут случайными величинами и мы не сможем применить теорему Пирсона. Однако оказывается, если в качестве оценок параметров $\theta_i$ взять оценки максимального правдоподобия $\theta^*_{in}$, то как показал Р.Фишер статистика $\widehat{\chi}^2_n$ имеет в пределе при $n \to \infty$ распределение $\chi^2$ с $r - s - 1$ степенями свободы.

**Теорема 3.22** (Р.Фишер). *Статистика*

$$\widehat{\chi}^2_n = \sum_{i+1}^{n} \frac{\left[\nu_i - np_i(\theta^*_{1n}, \ldots, \theta^*_{sn})\right]^2}{np_i(\theta^*_{1n}, \ldots, \theta^*_{sn})}$$

*имеет распределение* $\chi^2$ *с* $r - s - 1$ *степенями свободы.*

### Критерий Колмогорова

Другим непараметрическим критерием является *критерий Колмогорова.*

Этот критерий основан на статистике

$$D_n = \sup_{-\infty < x < \infty} \left| F^*_n(x) - F(x) \right|,$$

где $F^*_n(x)$ – эмпирическая функция распределения

$$F^*_n(x) = \frac{1}{n} \sum_{k=1}^{n} I_{\{x_n \leqslant x\}}.$$

**Теорема 3.23.** *Если распределение* $F(x)$ *непрерывно, то распределение статистики* $D_n$ *не зависит от* $F(x)$.

*Доказательство.* Покажем, что статистика $D_n$ для любой непрерывной $F(x)$ имеет такое же распределение, как и в случае, когда $F(x)$ определяет равномерное распределение на $[0, 1]$. Пусть $x_1, \ldots, x_n$ – выборка с распределением $F(x)$. Предположим, что $F(a) = 0$, $F(b) = 1$ и $F(x)$ строго возрастает на $(a, b)$, причем $a$ и $b$ могут быть бесконечными. Введем случайные величины $y_k = F(x_k)$. Так как $x_1, \ldots, x_n$ независимы, то и $y_1, \ldots, y_n$ независимы.

Покажем, что $y_1, \ldots, y_n$ равномерно распределены на $[0, 1]$. При любом $y \in (0, 1)$ имеем

$$\mathbf{P}\{y_k \leqslant y\} = \mathbf{P}\{x_k \leqslant F^{-1}(y)\} = F(F^{-1}) = y$$

Для $y = F(x)$ случайные события $\{y_k \leqslant y\}$ и $\{x_k \leqslant x\}$ равносильны. Поэтому

$$F^*_n(x, x_1, \ldots x_n) = \frac{1}{n} \sum_{k=1}^{n} I_{\{x_k \leqslant x\}} = \frac{1}{n} \sum_{k=1}^{n} I_{\{y_k \leqslant y\}} = F^*_n(y, y_1, \ldots y_n).$$

Поэтому

$$D_n = \sup_{a < x < b} \left| F^*_n(x, x_1, \ldots x_n) - F(x) \right| = \sup_{0 < y < 1} \left| F^*_n(y_1, \ldots, y_n) - y \right|.$$

В общем случае в качестве $F^{-1}(y)$ следует взять $\sup\{x : F(x) = y\}$. $\square$

А.Н.Колмогоров доказал, что

$$\lim_{n \to \infty} \mathbf{P}\{\sqrt{n} D_n \leqslant x\} = K(x) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 x^2}, \;\; x > 0.$$

На основе этого предельного равенства строится критерий Колмогорова. Пусть $k_\alpha$ – $\alpha$-квантиль предельного распределения
$$1 - K(k_\alpha) = \alpha\,.$$
Тогда гипотеза о том, что выборка взята из распределения $F(x)$, принимается, если $\sqrt{n}D_n \leqslant k_\alpha$, и отвергается, если $\sqrt{n}D_n > k_\alpha$. Уровень значимости этого критерия приблизительно равен $\alpha$.

### Критерий Смирнова

Пусть $x_1, \dots, x_n$ и $y_1, \dots, y_m$ – две независимые выборки. Первая имеет распределениеt $F(x)$, вторая выборка – $G(x)$. Введем статистику
$$D_{n,m} = \sup_{-\infty<x<\infty} |F_n^*(x, x_1, \dots, x_n) - F_m^*(x, y_1, \dots, y_m)|\,.$$
Н.В. Смирнов доказал, что если $F(x) = G(x)$ и непрерывны, то при $n, m \to \infty$ так, что $\frac{n}{m} \to \tau$, $0 < \tau < \infty$, случайная величина
$$\sqrt{\frac{nm}{n+m}}D_{n,m}$$
стремится к распределению $K(x)$. Эта предельная теорема позволяет строить критерий проверки гипотезы, что выборки $x_1, \dots, x_n$ и $y_1, \dots, y_m$ взяты из одного и того же распределения.

### Критерий значимости для параметров

1) Пусть $x_1, \dots, x_n$ и $y_1, \dots, y_m$ выборки из нормальной совокупности с неизвестным средними. Для проверки гипотезы $H : \sigma_x^2 = \sigma_y^2$ используется статистика
$$F_{n,m} = \frac{\widehat{S}_x^2}{\widehat{S}_y^2} = \frac{m-1}{n-1}\frac{\sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma_x}\right)^2}{\sum_{i=1}^m \left(\frac{y_i - \bar{y}}{\sigma_y}\right)^2}\,,$$
которая имеет распределение Фишера с $(n-1, m-1)$ степенями свободы. Пусть $\widehat{F}_\alpha(n,m)$ обозначим квантиль распределения Фишера, то есть решение уравнения
$$\mathbf{P}\{F_{n,m} > \widehat{F}_\alpha(n,m)\} = 1 - \alpha\,.$$
Заметим, что для квантилей $\widehat{F}_\alpha(n,m)$ выполняется соотношение
$$\widehat{F}_\alpha(n,m) = \frac{1}{\widehat{F}_{1-\alpha}(m,n)}\,.$$
Имеем
$$\begin{aligned}\mathbf{P}\{\widehat{F}_{1-\alpha/2}(n,m) < F_{n,m} \leqslant \widehat{F}_{\alpha/2}(n,m)\} &= \mathbf{P}\{F_{n,m} \leqslant \widehat{F}_{\alpha/2}(n,m)\} - \mathbf{P}\{\widehat{F}_{1-\alpha/2}(n,m) < F_{n,m}\}\\ &=1 - \alpha/2 - (1 - (1-\alpha/2)) = 1 - \alpha/2 - \alpha/2 = 1 - \alpha\,.\end{aligned}$$
Положим
$$C_1 = \widehat{F}_{1-\alpha/2}(n,m) \ \text{ и } \ C_2 = \widehat{F}_{\alpha/2}(n,m).$$
Если для полученных выборок выполняется неравенство
$$C_1 \leqslant \frac{\widehat{S}_x^2}{\widehat{S}_y^2} \leqslant C_2$$
то гипотеза принимается, в противном случае гипотеза отклоняется.

2) Для проверки гипотезы о равенстве дисперсий ($H : \sigma_x^2 = \sigma_y^2$) для выборок $x_1, \ldots, x_n$ и $y_1, \ldots, y_m$ из нормальной совокупности с известными средними используется статистика

$$F_{n,m} = \frac{s_{0x}^2}{s_{0y}^2},$$

Которая имеет распределение Фишера с $(n, m)$ степенями свободы.

3) Гипотеза о равенстве средних ($H : m_x = m_y$) при известных дисперсиях $\sigma_x^2$ и $\sigma_y^2$ проверяется с использованием статистики

$$Z = \frac{\overline{x} - \overline{y}}{\sqrt{D(\overline{x} - \overline{y})}} = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}\,.$$

Статистика $Z$ имеет нормальное распределение с параметрами $(0, 1)$. Гипотеза принимается, если $|Z| < \widehat{u}_{\alpha/2}$ и отклоняется в противном случае. Здесь $\widehat{u}_{\alpha/2}$ – квантиль нормального распределения с параметрами $(0, 1)$.

4) Обсудим здесь гипотезу о равенстве средних при неизвестной дисперсии в предположении равенства дисперсий $\sigma_x^2 = \sigma_y^2 = \sigma$. В общем случае получено только приближенное решение задачи. Статистики

$$(n-1)\frac{\widehat{s}_x^2}{\sigma_x^2} = \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma_x}\right)^2 \ \text{ и } \ (m-1)\frac{\widehat{s}_y^2}{\sigma_y^2} = \left(\frac{y_i - \overline{y}}{\sigma_y}\right)^2$$

распределены по закону $\chi^2$ с $(n-1)$ и $(m-1)$ степенями свободы. Тогда статистика

$$(n-1)\frac{\widehat{s}_x^2}{\sigma_x^2} + (m-1)\frac{\widehat{s}_y^2}{\sigma_y^2}$$

распределена по закону $\chi^2$ с $(n+m-2)$ степенями свободы. Статистика

$$\frac{(n-1)\widehat{s}_x^2 + (m-1)\widehat{s}_y^2}{n+m-2}$$

является несмещенной оценкой дисперсии, так как

$$M\Big[\frac{(n-1)\widehat{s}_x^2 + (m-1)\widehat{s}_y^2}{n+m-2}\Big] = \frac{n-1}{n+m-2}\sigma^2 + \frac{m-1}{n+m-2}\sigma^2 = \sigma^2\,.$$

Далее, математическое ожидание случайной величины $\overline{x} - \overline{y}$ равно $m_x - m_y$ и дисперсия при наших допущениях равна $\sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)$. Поэтому случайная величина

$$\frac{(\overline{x} - \overline{y}) - (m_x - m_y)}{\sigma\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)}}$$

имеет нормальное распределение с параметрами $(0, 1)$. Случайная величина

$$\frac{(\overline{x} - \overline{y}) - (m_x - m_y)}{\sqrt{\left((n-1)\widehat{s}_x^2 + (m-1)\widehat{s}_y^2\right)}}\sqrt{n+m-2}\sqrt{\frac{mn}{n+m}}$$

не зависит от $m_x$, $m_y$ и $\sigma^2$ и распределена по Стьюденту с $n+m-2$ степенями свободы. Для проверки гипотезы $H : m_x = m_y$ используется статистика

$$\tau_{n+m-2} = \frac{(\overline{x} - \overline{y})}{\sqrt{\left((n-1)\widehat{s}_x^2 + (m-1)\widehat{s}_y^2\right)}}\sqrt{n+m-2}\sqrt{\frac{mn}{n+m}}\,,$$

распределенная по Стьюденту с $(n+m-2)$ степенями свободы.

**Оценка коэффициента корреляции**

Пусть $(x_1, y_1), \ldots, (x_n, y_n)$ – независимые наблюдения над двумерной случайной величиной $(X, Y)$. Введем статистику

$$r^* = \frac{m_{11}}{s_{01}s_{02}},$$

где

$$m_{11} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}), \quad s_{01}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \text{ и } s_{02}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2.$$

Как показал Фишер статистика

$$\frac{1}{2}\log\frac{1+r^*}{1-r^*}$$

асимптотически нормально распределена с параметрами

$$\left(\frac{1}{2}\log\frac{1+r}{1-r}, \frac{1}{\sqrt{n-3}}\right),$$

где $r$ – коэффициент корреляции случайных величин $X$ и $Y$. Напомним, что функция

$$\frac{1}{2}\log\frac{1+r}{1-r}$$

является обратной к гиперболическому тангенсу. Если $\widehat{x}_\alpha$ – квантиль нормального распределения с параметрами $(0, 1)$, то доверительный интервал для коэффициента корреляции $r$ с приближенной надежностью $1 - \alpha$ имеет вид

$$\tanh\left(\frac{1}{2}\log\frac{1+r^*}{1-r^*} - \widehat{x}_{\alpha/2}\frac{1}{\sqrt{n-3}}\right) < r < \tanh\left(\frac{1}{2}\log\frac{1+r^*}{1-r^*} + \widehat{x}_{\alpha/2}\frac{1}{\sqrt{n-3}}\right).$$

## 3.9. Статистические гипотезы. Критерий Неймана – Пирсона

Пусть $\xi$ – случайная величина из семейства распределений с плотностей $p(t, \theta)$, описывающая генеральную совокупность и $x_1, \ldots, x_n$ – независимая выборка из генеральной совокупности с совместной плотностью $p(u, \theta) = \prod_{k=1}^{n} p(u_k, \theta)$, $u = (u_1, \ldots, u_n)$.

Предположения о параметре $\theta$ будем называть статистическими гипотезами. Ставится задача: проверить, что при данных значениях $\theta$ выборка согласуется с гипотезой или нет. Как правило строится критическое множество $S$ в пространстве всех значений принимаемых выборкой. Гипотеза отвергается, если конкретная выборка попадает в $S$ и принимается в противном случае. Подобного рода статистический критерий называется $S$-критерием. Относительно параметра $\theta$ имеются основная (проверяемая) гипотеза $H_0$ и часто конкурирующая (альтернативная) гипотеза $H_1$. Гипотезы бывают простыми, $\theta = \theta_0$, и сложными, $\theta \in \Theta_0$.

*Пример 13.*

$H_0$ : выборка из нормального распределения с параметрами $(a_0, \sigma_0)$ – простая гипотеза;
$H_1$ : выборка из нормального распределения с параметрами $(a, \sigma_0)$, $a \neq a_0$ – сложная гипотеза.

Рассмотрим случай двух простых гипотез. Пусть $H_0 : \theta = \theta_0$ и $H_1 : \theta = \theta_1$ – две простые гипотезы. С каждым критерием связаны ошибки двух родов.

Ошибка первого рода – отвергается гипотеза $H_0$, когда она верна.

Ошибка второго рода – принимается гипотеза $H_0$, когда верна гипотеза $H_1$.

Если $\mathbf{P}_{H_i}(B) = \mathbf{P}_i(B) = \int_B p(u, \theta_i)du$, то $\alpha = \mathbf{P}_0(S)$ – вероятность ошибки первого рода и $\beta = \mathbf{P}_1(\overline{S}) = 1 - \mathbf{P}_1(S)$ — вероятность ошибки второго рода, а $1 - \beta = \mathbf{P}_1(S)$ – вероятность отвергнуть гипотезу $H_0$, когда верна гипотеза $H_1$.

Вероятность $\alpha$ ошибки первого рода называется уровнем значимости, а вероятность $\mathbf{P}_1(S)$, равная $1-\beta$, обозначается через $W(S,\theta_1)$ и называется мощностью критерия. Среди всех $S$-критериев с уровнем значимости $\alpha$ ищется наиболее мощный критерий $S^*$. Если $\mathcal{L}_\alpha$ – множество всех $S$-критериев с уровнем значимости $\alpha$, то

$$W(S^*,\theta) = \int\limits_{S^*} p(u,\theta_1)du = \max_{S\in\mathcal{L}_\alpha} \int\limits_{S} p(u,\theta_1)du\,.$$

Отметим, что оптимальный критерий не всегда существует.

### Теорема Неймана-Пирсона

Предположим, что по выборке $x_1,\dots,x_n$ нужно различить две простые гипотезы $H_0$ и $H_1$, согласно которым выборка имеет плотность распределения $p_0(u)$ или $p_1(u)$, $u=(u_1,\dots,u_n)$. Введем критическое множество $S_c = \{u=(u_1,\dots,u_n)\,:\,p_1(u)\geqslant cp_0(u)\}$. Ограничимся рассмотрением случая, когда для любого $\alpha$ существует постоянная $c$ такая, что $\mathbf{P}_{H_0}(S_c)=\alpha$.

**Теорема 3.24.** *Среди всех критериев, различающих гипотезы $H_0$ и $H_1$ с заданной ошибкой 1-го рода $\alpha$, наиболее мощный критерий определяется критическим множеством $S_c$.*

*Доказательство.* Пусть $S_c$ и $S$ – критические множества с $\mathbf{P}_{H_0}(S_c)=\alpha$ и $\mathbf{P}_{H_0}(S)=\alpha$. Имеем

$$\mathbf{P}_{H_0}(S_c\backslash S_cS) = \alpha - \mathbf{P}_{H_0}(S_cS) = \mathbf{P}_{H_0}(S\backslash SS_c)\,.$$

Отсюда находим, что

$$\begin{aligned}\mathbf{P}_{H_1}(S_c\backslash S_cS) &= \int\limits_{S_c\backslash S_cS} p_1(u)du \geqslant c\int\limits_{S_c\backslash S_cS} p_0(u)du = c\mathbf{P}_{H_0}(S_c\backslash S_cS) = \\ &= c\mathbf{P}_{H_0}(S\backslash SS_c) = c\int\limits_{S\backslash SS_c} p_0(u)du > \int\limits_{S\backslash SS_c} p_1(u)du = \mathbf{P}_{H_1}(S\backslash SS_c)\,,\end{aligned} \tag{3.4}$$

так как на дополнении к $S_c$ выполняется неравенство $p_1(u) < cp_0(u)$. Добавим к обеим частям равенства (3.4) слагаемое $\mathbf{P}_{H_1}(SS_c)$. В результате, получим

$$\mathbf{P}_{H_1}(S_c) \geqslant \mathbf{P}_{H_1}(S)\,. \ \square$$

### Оптимальный критерий для проверки гипотезы о параметрах нормального распределения

Пусть $x_1,\dots,x_n$ – независимая выборка из нормального распределения с параметрами $(a,\sigma)$. Пусть $\sigma$ известна, а относительно параметра $a$ имеются две гипотезы:

- гипотеза $H_0\,:\ a=a_0$;
- гипотеза $H_1\,:\ a_1>a_0$.

Отношение правдоподобия $p_1(u)/p_0(u)$ имеет вид

$$\frac{p_1(u)}{p_0(u)} = \exp\big\{\frac{n}{\sigma^2}\overline{u}(a_1-a_0) - \frac{n}{2\sigma^2}(a_1^2-a_0^2)\big\}\,,\ \ u=(u_1,\dots.u_n)\,,$$

где $\overline{u} = \frac{1}{n}\sum_{i=1}^{n} u_i$. Ясно, что $\{p_1(u)/p_0(u) > c\} = \{\overline{u} > c_1\}$ при некотором $c_1$. Выборочное среднее $\overline{x}$ распределено нормально с параметрами $(a, \sigma/\sqrt{n})$. Найдем ошибки первого и второго рода:

$$\alpha = \mathbf{P}\{\overline{x} > c_1 \,|\, H_0\} = \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{c_1 - a_0}{\sigma}\sqrt{n}}^{\infty} e^{-\frac{u^2}{2}} du = 1 - \Phi\big(\frac{c_1 - a_0}{\sigma}\sqrt{n}\big) \tag{3.5}$$

$$\beta = \mathbf{P}\{\overline{x} < c_1 \,|\, H_1\} = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{c_1 - a_1}{\sigma}\sqrt{n}} e^{-\frac{u^2}{2}} du = \Phi\big(\frac{c_1 - a_1}{\sigma}\sqrt{n}\big) \tag{3.6}$$

Обозначим через $u_\gamma$ квантиль нормального распределения

$$1 - \Phi(u_\gamma) = \gamma\,.$$

Напомним, что $u_\gamma = -u_{1-\gamma}$. Тогда из (3.5) вытекает

$$\frac{c_1 - a_0}{\sigma}\sqrt{n} = u_\alpha\,.$$

С помощью (3.6) находим квантиль $u_\beta$, а по таблице значений функции $\Phi(x)$ определяем ошибку второго рода $\beta$. Далее сравнивая

$$a_0 + u_\alpha \frac{\sigma}{\sqrt{n}} = a_1 + u_\beta \frac{\sigma}{\sqrt{n}} = c_1\,,$$

находим объем выборки $n$, который при оптимальном критерии обеспечивает ошибки первого и второго рода $\alpha$ и $\beta$, а именно,

$$n = \frac{\sigma^2(u_\alpha + u_\beta)^2}{(a_1 - a_0)^2}\,.$$

Пусть теперь параметр $a$ известен. Рассмотрим следующие две гипотезы:

- гипотеза $H_0 :\ a = 0,\ \sigma = \sigma_0$;
- гипотеза $H_1 :\ a = 0,\ \sigma = \sigma_1 > \sigma_0$.

Отношение правдоподобия

$$\frac{p_1(u)}{p_0(u)} = \frac{\sigma_0^n}{\sigma_1^n} \exp\big\{\frac{1}{2}\big(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\big)\sum_{i-1}^{n} x_i^2\big\}\,,\quad u = (u_1, \ldots, u_n)$$

приводит к критическому множеству

$$\sum_{i=1}^{n} x_i^2 \geqslant c_1\,.$$

Случайная величина $\sum_{i=1}^{n} \frac{x_i^2}{\sigma^2}$ имеет $\chi^2$-распределение с $n$ степенями свободы с функцией распределения

$$K_n(x) = \int\limits_0^x \frac{1}{2^{n/2}\Gamma\big(\frac{n}{2}\big)} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} dt\,,\quad x > 0\,.$$

Ошибки первого и второго рода определяются из равенств

$$\alpha = \mathbf{P}\big\{\frac{1}{\sigma_0^2}\sum_{i=1}^{n} x_i^2 \geqslant \frac{c_1}{\sigma_0^2}\big\} = 1 - K_n\big(\frac{c_1}{\sigma_0^2}\big)\,,$$

$$\beta = \mathbf{P}\Big\{\frac{1}{\sigma_1^2}\sum_{i=1}^{n} x_i^2 \geqslant \frac{c_1}{\sigma_1^2}\Big\} = 1 - K_n\Big(\frac{c_1}{\sigma_1^2}\Big)\,.$$

Величины $\alpha$ и $\beta$ выбираются в зависимости от степени нежелательности ошибочных решений.

**Оптимальный критерий в схеме Бернулли**

Пусть $0 < p_0 < p_1 < 1$. Рассмотрим две гипотезы:

- $H_0:\ p_0(k) = C_n^k p_0^k (1-p_0)^{n-k}$;
- $H_1:\ p_1(k) = C_n^k p_1^k (1-p_1)^{n-k}$.

Оптимальный критерий для проверки гипотезы $H_0$ против $H_1$ строится исходя из неравенства

$$\frac{p_1(k)}{p_0(k)} = \Big[\frac{p_1(1-p_0)}{p_0(1-p_1)}\Big]^k \frac{(1-p_1)^n}{(1-p_0)^n} \geqslant c\,,$$

которое равносильно неравенству $k \geqslant c_1$ при некотором $c_1$. Для вычисления ошибок первого и второго рода воспользуемся асимптотической нормальностью случайной величины $\mu_n$ с параметрами $(np, \sqrt{np(1-p)})$. Имеем

$$\alpha = \mathbf{P}\{\mu_n \geqslant c_1 | H_0\} = \mathbf{P}\Big\{\frac{\mu_n - np_0}{\sqrt{np_0(1-p_0)}} \geqslant \frac{c_1 - np_0}{\sqrt{np_0(1-p_0)}} \Big| H_0\Big\},$$

$$\beta = \mathbf{P}\{\mu_n < c_1 | H_1\} = \mathbf{P}\Big\{\frac{\mu_n - np_1}{\sqrt{np_1(1-p_1)}} \geqslant \frac{c_1 - np_1}{\sqrt{np_1(1-p_1)}} \Big| H_1\Big\}.$$

Задавшись ошибкой первого рода $\alpha$, найдем из первого равенства квантиль $u_\alpha$ нормального распределения с параметрами $(0,1)$ и границу $c_1$. Из второго равенства находим ошибку второго рода $\beta$ Для вероятностей ошибок $\alpha$ и $\beta$ граница $c_1$ представляется в виде

$$c_1 \approx np_0 + u_\alpha\sqrt{np_0(1-p_0)} \approx np_1 - u_\beta\sqrt{np_1(1-p_1)}\,.$$

Отсюда находим необходимый приблизительный объем выборки

$$n \approx \frac{\big(u_\alpha\sqrt{p_0(1-p_0)} + u_\beta\sqrt{p_1(1-p_1)}\big)^2}{(p_0 - p_1)^2},$$

гарантирующий ошибку первого рода $\alpha$ и ошибку второго рода $\beta$.

## Рекомендуемая литература

### ОСНОВНАЯ:

1. Севастьянов Б.А. Курс теории вероятностей и математической статистики. М.: Наука, 1982.

2. Боровков А.А. Теория вероятностей. М.: Эдиториал УРСС, 1999.

3. Ширяев А.Н. Вероятность. М.: Наука, 1989.

4. Ивченко Г.И., Медведев Ю.И. Математическая статистика. М.: Высшая школа, 1992.

5. Боровков А.А. Математическая статистика. М.: Эдиториал УРСС, 2000.

6. Севастьянов Б.А., Чистяков В.П., Зубков А.М. Сборник по теории вероятностей. М.: Наука, 1980.

7. Ивченко Г.И., Медведев Ю.И., Чистяков А.В. Сборник задач по математической статистике. М.: Высшая школа, 1989.

### ДОПОЛНИТЕЛЬНАЯ:

8. Гнеденко Б.В. Курс теории вероятностей. М.: Наука, 1988.

9. Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1998.

10. Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. М.: Наука, 1989.

11. Феллер В. Теория вероятностей и ее приложения. т. 1. М.: Мир, 1984.

12. Крамер Г. Математические методы статистики. М.: Мир, 1975.

13. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М.: Высшая школа, 1998.

14. Прохоров А.В., Ушаков В.Г., Ушаков Н.Г. Задачи по теории вероятностей: Основные понятия. Предельные теоремы. Случайные процессы. М.: Наука, 1986.