

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Агафонова Н.Ю.

СТОХАСТИЧЕСКИЙ АНАЛИЗ

Учебное пособие для студентов механико-математического факультета

Саратов, 2019

Содержание

1	ТЕОРИЯ ВЕРОЯТНОСТЕЙ	3
1.1	Случайные события	3
1.2	Случайные величины	8
1.3	Числовые характеристики основных законов распределения	12
1.4	Условные распределения. Функция регрессии	16
1.5	Характеристические функции	20
1.6	Законы больших чисел. Сходимость последовательностей случайных величин	24
2	МАТЕМАТИЧЕСКАЯ СТАТИСТИКА	28
2.1	Вариационные ряды и их характеристики	28
2.2	Оценки параметров распределения	32
2.3	Эффективные оценки. Неравенство Рао-Крамера.	34
2.4	Методы построения оценок параметров распределения	38
2.5	Доверительные интервалы	41
2.6	Проверка статистических гипотез.	47

1 ТЕОРИЯ ВЕРОЯТНОСТЕЙ

1.1 Случайные события

Понятия множества элементарных исходов эксперимента и случайного события

Назовем случайным экспериментом такой эксперимент, в результате которого может наступить один из множества результатов. При этом множество результатов эксперимента известно *до* проведения эксперимента и не меняется в *процессе* его проведения. Результат эксперимента определяется только после проведения эксперимента.

Определение. Элементарным исходом эксперимента называется непосредственный, "самый мелкий" результат, которым может закончиться эксперимент. Элементарный исход обозначается ω .

Множество элементарных исходов эксперимента обозначается Ω .

Определение. Случайным событием называется любое подмножество A множества Ω , $A \in \Omega$. Элементарные исходы, из которых состоит событие A , называются исходами, благоприятствующими событию A .

Как аксиому примем, что пустое множество \emptyset является подмножеством любого множества, а само множество является собственным подмножеством.

Определение. Невозможным называется событие \emptyset , которому не благоприятствует ни один исход эксперимента, т.е. невозможное событие в результате данного эксперимента наступить не может.

Определение. Достоверным называется событие Ω , которому благоприятствует любой исход эксперимента, т.е. достоверное событие в результате данного эксперимента наступает обязательно.

Определение. Если множество элементарных исходов эксперимента состоит из конечного числа равновозможных исходов, тогда говорят, что опре-

делена классическая вероятность события

$$P(A) = k/n,$$

где k — количество исходов, благоприятствующих событию A , n — количество всех возможных исходов эксперимента.

Определение. Пусть множество элементарных исходов $\Omega \in \mathbb{R}^n$ является множеством из пространства \mathbb{R}^n $A \in \Omega$ — случайное событие. *Геометрической вероятностью* события A называется число

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

где $\mu(A)$ — мера множества A в пространстве \mathbb{R}^n , т.е. длина в \mathbb{R}^1 , площадь в \mathbb{R}^2 , объем в \mathbb{R}^3 .

Вероятность может рассматриваться как функция, область определения которой — \mathcal{F} , множество значений — отрезок $[0, 1]$. Вероятность невозможного события $P(\emptyset) = 0$, вероятность достоверного события $P(\Omega) = 1$, вероятность некоторого события $0 \leq P(A) \leq 1$. (\mathcal{F} — множество всех возможных подмножеств Ω , или множество всех возможных в данном эксперименте случайных событий)

Операции сложения и умножения случайных событий.

Определение. Сумма событий A и B есть событие C , которому благоприятствуют и исходы события A и исходы события B . Другими словами, событие C наступает, если наступает хотя бы одно из событий A или B .

Определение. Произведение событий A и B есть событие C , которому благоприятствуют исходы событий A и B одновременно. Другими словами событие C наступает, если наступают оба события A и B .

Сумма событий обозначается $C = A \cup B$, произведение $C = A \cap B$.

Полной группой событий называется последовательность событий таких, что в результате опыта непременно должно произойти хотя бы одно из них.

Другими словами, сумма всех этих событий является достоверным событием

$$A_1 \cup A_2 \cdots \cup A_n = \Omega$$

Определение. Несовместными событиями называются такие события, которые в данном опыте не могут появиться вместе.

Это означает также, что появление одного из несовместных событий исключает появление другого. Произведение несовместных событий есть невозможное событие $A \cap B = \emptyset$.

Определение. Событием, противоположным событию A , называется событие \bar{A} , состоящее из исходов, не благоприятствующих событию A . Сумма противоположных событий $A \cup \bar{A} = \Omega$, произведение $A \cap \bar{A} = \emptyset$, т.е. противоположные события всегда несовместны.

Вероятность суммы противоположных событий $P(A \cup \bar{A}) = P(\Omega) = 1$, поэтому $P(A) = 1 - P(\bar{A})$, $P(\bar{A}) = 1 - P(A)$.

Теоремы сложения вероятностей случайных событий.

Теорема сложения вероятностей для несовместных событий. Пусть A и B — несовместные случайные события. Тогда вероятность суммы этих событий равна сумме вероятностей этих событий, т.е.

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

Аксиома теории вероятностей. Если A_1, A_2, \dots, A_n — последовательность попарно несовместных событий, то вероятность суммы событий равна сумме вероятностей

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

Теорема сложения вероятностей для несовместных событий. Пусть

A и B — несовместные случайные события. Тогда вероятность суммы событий равна сумме вероятностей этих событий, т.е.

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

Теорема сложения вероятностей для совместных событий. Пусть A и B — совместные случайные события. Тогда вероятность суммы событий равна сумме вероятностей этих событий минус вероятность их произведения, т.е.

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

Пусть A , B и C — совместные случайные события. Тогда вероятность суммы событий находится по формуле:

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

Теоремы умножения вероятностей случайных событий.
Независимость случайных событий.

Определение. Случайные события A и B называются независимыми, если $P(A \cap B) = P(A) \cdot P(B)$.

Теорема. (умножения вероятностей)

Пусть A , B — случайные события, $P\{A\} > 0$.

Тогда $P(A \cap B) = P(A) \cdot P(B|A)$.

Теорема. (критерий независимости событий) Пусть A , B — случайные события, $P\{A\} > 0$. События A и B независимы тогда и только тогда, когда $P(B|A) = P(B)$.

Определение. Последовательность событий $\{A_i\}_{i=1}^{\infty}$, удовлетворяющая условиям

1. $P(A_i) > 0, \forall i \in \mathbb{Z}_+$;

2. $\bigcap_{i=0}^{\infty} A_i = \Omega$.

называется полной группой событий.

Теорема.(Формула полной вероятности)

Пусть A — некоторое событие, а $\{A_i\}_{i=1}^{\infty}$ — полная группа попарно несовместных событий, $A_i \cap A_j = \emptyset$, $i \neq j$. Пусть известны вероятности $P(A_i) > 0$ и $P(A|A_i) \geq 0$.

Тогда

$$P(A) = \sum_{i=1}^{\infty} P(A_i) P(A|A_i).$$

Аксиомы Колмогорова А.Н.

Пусть Ω — множество элементарных исходов эксперимента. Некоторое его подмножество $A \subset \Omega$ называется случайным событием.

Определение. σ -алгеброй событий называется множество \mathcal{F} подмножеств $A \subset \Omega$, удовлетворяющее условиям

1. $\Omega \in \mathcal{F}$;
2. Если $A \in \mathcal{F}$, то $\bar{A} \in \mathcal{F}$;
3. Если $\{A_i\}_{i=1}^{\infty} \in \mathcal{F}$, то $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Определение.

Вероятностной мерой называется функция $P : \mathcal{F} \rightarrow \mathbb{R}$, удовлетворяющая условиям

1. $P(\Omega) = 1$;
2. $P(A) \geq 0$ для любого $A \in \mathcal{F}$;
3. Для любой последовательности $\{A_i\}_{i=1}^{\infty} \in \mathcal{F}$ такой, что при всех $i \neq j$ $A_i \cap A_j = \emptyset$, $P\left(\bigcap_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

1.2 Случайные величины

Вероятностным пространством называется тройка (Ω, \mathcal{F}, P) , где: Ω – это множество элементарных исходов, элементы $\omega \in \Omega$ которого называются элементарными исходами; \mathcal{F} – сигма-алгебра подмножеств множества Ω . Подмножества $A \in \mathcal{F}$ называются случайными событиями; P – вероятностная мера или вероятность, то есть сигма-аддитивная конечная мера, такая что $P(\Omega) = 1$.

Случайной величиной называется действительнзначная функция $\xi(\omega) : \mathcal{F} \rightarrow \mathbb{R}$, т.е. такая что $\forall x \in \mathbb{R}$ множество

$$\{\omega : \xi(\omega) < x\} \in \mathcal{F}.$$

Функцией распределения случайной величины ξ называется функция $F_\xi(x)$, определяющая для каждого значения x вероятность того, что случайная величина ξ примет значение меньшее, чем x , то есть

$$F_\xi(x) = P(\xi < x), \quad x \in \mathbb{R}.$$

Основные свойства функции распределения $F_\xi(x)$:

1. $0 \leq F_\xi(x) \leq 1, \forall x \in \mathbb{R}$;
2. $F_\xi(x)$ – неубывающая, непрерывная слева функция;
3. $\lim_{x \rightarrow -\infty} F_\xi(x) = 0, \lim_{x \rightarrow +\infty} F_\xi(x) = 1$.
4. $P\{a \leq \xi < b\} = F(b) - F(a)$.

Случайные величины делятся на два класса: дискретные и непрерывные. Дискретной называется случайная величина ξ , принимающая конечное или счетное множество значений. Дискретную случайную величину обычно задают таблицей вида:

ξ	x_1	x_2	\dots	x_k	\dots
p	p_1	p_2	\dots	p_k	\dots

Здесь $p_k = P(\xi = x_k)$

Непрерывной называется случайная величина ξ , функция распределения которой дифференцируема, т.е. существует производная $f(x) = F'_\xi(x)$, называемая плотностью распределения случайной величины ξ .

Случайная величина называется абсолютно непрерывной, если $\forall x \in \mathbb{R}$ ее функцию распределения можно представить в виде

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

Случайные величины ξ и η называются независимыми, если для любых $x, y \in \mathbb{R}$ справедливо равенство:

$$P(\xi < x, \eta < y) = P(\xi < x)P(\eta < y),$$

или, в терминах функций распределения:

$$F_{\xi\eta}(x, y) = F_\xi(x)F_\eta(y).$$

Математическим ожиданием дискретной случайной величины ξ называется ряд

$$E\xi = \sum_k x_k p_k$$

при условии его абсолютной сходимости.

Математическим ожиданием непрерывной случайной величины ξ с плотностью распределения $p(x)$ называется интеграл

$$E\xi = \int_{-\infty}^{+\infty} x f(x) dx$$

также при условии, что он абсолютно сходится.

Пусть ξ, η – произвольные случайные величины, a, b – константы.

Свойства математического ожидания:

1. $EC = C$, где $C = const$.
2. $E(a\xi + b\eta) = aE\xi + bE\eta$.
3. Если случайные величины ξ, η имеют конечные математические ожидания, то $E(\xi + \eta) = E\xi + E\eta$.
4. Математическое ожидание произведения независимых случайных величин, имеющих конечные математические ожидания, равно произведению их математических ожиданий, т.е.

$$E[\xi\eta] = E\xi E\eta.$$

5. Если ξ – дискретная случайная величина, функция $g(x)$ непрерывна на \mathbb{R} , то для случайной величины $\eta = g(x)$ справедливо равенство

$$E\eta = \sum_k g(x_k) p_k.$$

Если ξ – непрерывная случайная величина, функция $g(x)$ является непрерывной на \mathbb{R} , то для случайной величины $\eta = g(x)$ справедливо

$$E\eta = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Дисперсией случайной величины называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D\xi = E(\xi - E(\xi))^2.$$

Дисперсия характеризует рассеяние (разбросанность) значений случайной величины около ее математического ожидания.

Свойства дисперсии:

1. $D\xi = E\xi^2 - (E(\xi))^2$.

2. $D\xi \geq 0$.
3. Дисперсия константы $D(C) = 0$.
4. Если $D(C) = 0$, то $P\{\xi = C\} = 1$, то есть ξ константа почти наверное.
5. $D(C\xi) = C^2 D\xi$.

Ковариация двух случайных величин – это числовая характеристика зависимости случайных величин, определяемая следующим образом:

$$\text{cov}(\xi, \eta) = E[(\xi - E(\xi))(\eta - E(\eta))].$$

Свойства ковариации:

1. Если ξ, η являются независимыми случайными величинами, то

$$\text{cov}(\xi, \eta) = 0.$$

2. Ковариация случайной величины с собой равна дисперсии:

$$\text{cov}(\xi, \xi) = D\xi.$$

3. Ковариация симметрична:

$$\text{cov}(\xi, \eta) = \text{cov}(\eta, \xi).$$

4. $\text{cov}(\xi + a, \eta + b) = \text{cov}(\xi, \eta)$.

Коэффициент корреляции $r(\xi, \eta)$ – это статистический показатель зависимости двух случайных величин ξ и η , который вычисляется по формуле:

$$r(\xi, \eta) = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$$

Основные свойства коэффициента корреляции:

1. Если случайные величины ξ, η независимы, то $r(\xi, \eta) = 0$.

$$2. |r(\xi, \eta)| \leq 1.$$

3. Если случайные величины ξ, η линейно зависимы, т.е. $\eta = a\xi + b$, то $r(\xi, \eta) = \pm 1$.

1.3 Числовые характеристики основных законов распределения

Распределение Бернулли $\varepsilon \sim \text{Bern}(p)$.

Закон распределения $P\{\varepsilon = k\} = p^k q^{n-k}$, где $k = 0, 1$:

Сл.в. ε принимает значение 1 в случае наступления события A — 'успех', значение 0 — в случае наступления события \bar{A} — "неуспех". Вероятность успеха p — параметр распределения.

Математическое ожидание сл.в. ε

$$E\varepsilon = 0 \cdot q + 1 \cdot p = p.$$

Математическое ожидание сл.в. ε^2 равно

$$E\varepsilon^2 = 0 \cdot q + 1 \cdot p = p.$$

Дисперсия равна

$$D\xi = E\xi^2 - (E\xi)^2 = p - p^2 = p(1 - p) = pq.$$

□

Биномиальное распределение $\xi \sim \text{Bin}(n, p)$

Сл.в. ξ — количество успехов в n независимых одинаковых испытаниях Бернулли.

Параметры распределения: n — количество испытаний, p — вероятность успеха в одном испытании.

Закон распределения $P\{\xi = k\} = C_n^k p^k q^{n-k}$, $k = 0, 1, \dots, n$.

Для вычисления числовых характеристик применим свойства математического ожидания и дисперсии. Сл.в. ξ может быть представлена в виде суммы сл. величин $\varepsilon_i \sim \text{Bern}(p)$, каждая из которых выражает результат i -го испытания

$$\xi = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n.$$

Математическое ожидание сл.в. ξ

$$E\xi = E(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n) = E(\varepsilon_1) + E(\varepsilon_2) + \dots + E(\varepsilon_n) = p + p + \dots + p = np$$

Дисперсия сл.в. ξ

$$D\xi = D(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n) = D(\varepsilon_1) + D(\varepsilon_2) + \dots + D(\varepsilon_n) = pq + pq + \dots + pq = npq$$

□

Пуассоновское распределение $\xi \sim \text{Pois}(\lambda)$

Сл.в. ξ – количество успехов в n независимых одинаковых испытаниях Бернулли, где n – велико.

Параметр распределения: λ – среднее количество успехов в большом количестве испытаний n .

Закон распределения

$$P\{\xi = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$$

, где $k = 0, 1, \dots$

Математическое ожидание сл.в. ξ

$$E\xi = \sum_{k=0}^{\infty} k \frac{\lambda e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Математическое ожидание сл.в. ξ^2

$$E\xi^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{m=0}^{\infty} (m+1) \frac{\lambda^m}{m!} = \dots = \lambda^2 + \lambda.$$

Дисперсия равна

$$D\xi = E\xi^2 - (E\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Равномерное распределение $\xi \sim R[a, b]$.

Функция плотности имеет вид

$$f(x) = \begin{cases} 1/(b-a), & x \notin [a, b]; \\ 0, & x \in [a, b]. \end{cases}$$

Математическое ожидание сл.в. ξ равно

$$E\xi = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}.$$

Математическое ожидание сл.в. ξ^2 равно

$$E\xi^2 = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}.$$

Дисперсия равна

$$D\xi = E\xi^2 - (E\xi)^2 = \frac{(b+a)^2}{4} - \frac{b^2 + ab + a^2}{3} = \frac{(b-a)^2}{12}.$$

Показательное (экспоненциальное) распределение $\xi \sim \Pi(\lambda)$.

Функция плотности имеет вид

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

Математическое ожидание сл.в. ξ равно (используем интегрирование по частям)

$$\begin{aligned} E\xi &= \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx = -xe^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} \lambda e^{-\lambda x} dx = \\ &= -\left(\lim_{x \rightarrow +\infty} xe^{-\lambda x} - e^0 \right) + \frac{-1}{\lambda} \left(\lim_{x \rightarrow +\infty} xe^{-\lambda x} - e^0 \right) = \frac{1}{\lambda}. \end{aligned}$$

Математическое ожидание сл.в. ξ^2 равно

$$\begin{aligned} E\xi^2 &= \int_0^{\infty} x^2 \cdot \lambda e^{-\lambda x} dx = -x^2 e^{-\lambda x} \Big|_0^{\infty} + 2 \int_0^{\infty} x e^{-\lambda x} dx = \\ &= 0 + 2 \cdot \frac{1}{\lambda^2} = \frac{2}{\lambda^2}. \end{aligned}$$

Дисперсия равна

$$D\xi = E\xi^2 - (E\xi)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

□

Таким образом, вероятностный смысл параметра λ показательного распределения состоит в следующем: параметр λ есть величина, обратная математическому ожиданию сл.в. Также следует отметить характеристическое свойство показательного распределения: среднее значение сл.в совпадает со средним квадратическим отклонением.

Нормальное распределение (распределение Гаусса) $\xi \sim N(a, \sigma^2)$.

Функция плотности сл.в. $\xi \sim N(a, \sigma^2)$ имеет вид

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

Функция плотности сл.в. $\xi_0 \sim N(0, 1)$ имеет вид

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

Найдем характеристики стандартной нормальной случайной величины $\xi_0 \sim N(0, 1)$, а затем воспользуемся свойствами математического ожидания и дисперсии.

Математическое ожидание сл.в. ξ_0 равно

$$E\xi_0 = \int_{-\infty}^{+\infty} x \cdot e^{-\frac{x^2}{2}} dx = \int_{-\infty}^0 x \cdot e^{-\frac{x^2}{2}} dx + \int_0^{+\infty} x \cdot e^{-\frac{x^2}{2}} dx (=)$$

Введем замену переменной $t = \frac{x^2}{2}$, тогда $dt = x dx$.

$$(=) - \int_0^{+\infty} e^{-t} dt + \int_0^{+\infty} e^{-t} dt = 0.$$

Математическое ожидание сл.в. ξ_0^2 равно

$$E\xi_0^2 = \int_{-\infty}^{+\infty} x^2 \cdot e^{-\frac{x^2}{2}} dx (=)$$

Интегрирование по частям: $u = x$, $du = dx$, $dv = xe^{-\frac{x^2}{2}} dx$,

$$v = \int xe^{-\frac{x^2}{2}} dx = \int e^{-\frac{x^2}{2}} d\left(\frac{x^2}{2}\right) = -e^{-\frac{x^2}{2}}$$

$$(\text{=}) \frac{1}{\sqrt{2\pi}} \left(x \cdot e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} \right) + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 0 + \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1.$$

Здесь использовался интеграл Пуассона $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$.

Дисперсия равна

$$D\xi_0 = E\xi_0^2 - (E\xi_0)^2 = 1 - 0 = 1.$$

Таким образом, для стандартной нормальной величины $\xi_0 \sim N(0, 1)$ $E\xi = 0$, $D\xi = 1$, для нормальной величины $\xi \sim N(a, \sigma^2)$ $E\xi = E(\sigma\xi_0 + a) = a$, $D\xi = D(\sigma\xi_0 + a) = \sigma^2$.

□

1.4 Условные распределения. Функция регрессии

Рассмотрим двумерное распределение (ξ, η) .

Определение. Функцией распределения (ф.р.) двумерного вектора или функцией совместного распределения случайных величин ξ и η называют функцию

$$F_{\xi, \eta}(x, y) = P\{\omega : \xi(\omega) < x, \eta(\omega) < y\}.$$

Свойства ф.р. двумерного вектора:

1. Для любых x, y из \mathbb{R} $0 \leq F_{\xi, \eta}(x, y) \leq 1$.

2. При фиксированном x_0 функция $F_{\xi,\eta}(x_0, y)$ и при фиксированном y_0 функция $F_{\xi,\eta}(x, y_0)$ являются неубывающими, непрерывными слева функциями соответствующей переменной.
3. $\lim_{x \rightarrow \infty} F_{\xi,\eta}(x, y) = F_{\eta}(y)$, $\lim_{y \rightarrow \infty} F_{\xi,\eta}(x, y) = F_{\xi}(x)$,
 $\lim_{x \rightarrow \infty, y \rightarrow \infty} F_{\xi,\eta}(x, y) = 1$,
 $\lim_{x \rightarrow -\infty, y \rightarrow -\infty} F_{\xi,\eta}(x, y) = \lim_{x \rightarrow -\infty} F_{\xi,\eta}(x, y) = \lim_{y \rightarrow -\infty} F_{\xi,\eta}(x, y) = 0$.

Сл. вектор называют **абсолютно непрерывным**, если его ф.р. представима в виде

$$F_{\xi,\eta}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy,$$

где $f(x, y)$ – функция плотности распределения вероятностей сл.вектора (ξ, η) .

Функция $f(x, y)$ обладает следующими свойствами:

1. Для любых x, y из \mathbb{R} $f(x, y) \geq 0$.
2. $\frac{\partial^2 F_{\xi,\eta}(x, y)}{\partial x \partial y} = f(x, y)$ п.в. по $x, y \in \mathbb{R}$.
3. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.
4. $\iint_S f(x, y) dS = P\{\xi \in S\}$.

Функции плотности распределения вероятностей сл.величин ξ и η могут быть восстановлены по функции плотности совместного распределения:

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f(x, y) dy, \quad f_{\eta}(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Пусть $F_{\xi,\eta}(x, y)$ – функция совместного распределения вероятностей сл.величин ξ и η .

Определение. Условное распределение сл. в. η относительно сл. в. ξ задается функцией распределения

$$F_{\eta|\xi}(x, y) = \frac{F_{\xi, \eta}(x, y)}{F_{\xi}(x)},$$

при $F_{\xi}(x) > 0$, и $F_{\eta|\xi}(x, y) = 0$, при $F_{\xi}(x) = 0$.

Определение. Условным математическим ожиданием сл. в. η относительно сл. в. ξ называется сл.в.

$$E_{\eta|\xi}(x, y) = \int_{-\infty}^{\infty} y dF_{\eta|\xi}(x, y),$$

с функцией распределения $F_{\xi}(x)$.

Определение. Функцией регрессии сл.в. η на сл.в. ξ называется функция $g(\xi) = E_{\eta|\xi}$. Если $g(\xi) = E_{\eta|\xi} = a\xi + b$, то говорят о линейной регрессии.

Построение уравнения линейной регрессии методом наименьших квадратов (МНК).

Предположим, что в двумерном распределении (ξ, η) возможно получение информации о сл.в. ξ , т.е. мы можем наблюдать её значения, либо знаем распределение. Требуется по имеющейся информации о сл.в. ξ сделать выводы о сл.в. η .

Пусть связь между величинами выражается некоторой линейной функцией $g(x) = ax + b$. Оценим коэффициенты методом наименьших квадратов (МНК).

Обозначим η – наблюдаемое значение η , $\hat{\eta}$ – вычисленные (прогнозные) значения.

Согласно МНК требуется найти такие значения оценок параметров \hat{a} и \hat{b} , чтобы была минимальной сумма квадратов отклонений прогнозных значений от наблюдаемых:

$$L(\hat{a}, \hat{b}) = E(\eta - \hat{\eta})^2 \rightarrow \min.$$

Значит, для нахождения оценки параметров парной регрессионной модели МНК необходимо найти экстремум (минимум) функции двух аргументов.

Запишем необходимые условия экстремума:

$$\begin{cases} \frac{\partial L}{\partial \hat{a}} = -2E(\eta - \hat{a} - \hat{b}\xi) = 0, \\ \frac{\partial L}{\partial \hat{b}} = -2E\xi(\eta - \hat{a} - \hat{b}\xi) = 0, \end{cases}$$

Раскрывая скобки получим:

$$\begin{cases} E\eta - \hat{a}n - \hat{b}E\xi = 0, \\ E\xi\eta - \hat{a}E\xi - \hat{b}E\xi^2 = 0. \end{cases} \quad (1)$$

Из первого уравнения системы имеем оценку параметра a :

$$\hat{a} = E\eta - \hat{b}E\xi. \quad (2)$$

Преобразуем второе уравнение системы и подставим полученную оценку \hat{a}

$$E\xi\eta - \hat{a}E\xi - \hat{b}E\xi^2 = E\xi\eta - E\xi E\eta - \hat{b}(E\xi^2 - (E\xi)^2) = 0,$$

$$\text{cov}(\xi, \eta) - \hat{b}D\xi = 0.$$

Отсюда получаем оценку параметра b :

$$\hat{b} = \frac{\text{cov}(\xi, \eta)}{D\xi} = \frac{\text{cov}(\xi, \eta) \sigma_\eta}{\sigma_\xi \sigma_\eta \sigma_\xi} = r \frac{\sigma_\eta}{\sigma_\xi}. \quad (3)$$

Таким образом, решение системы уравнений имеет вид

$$\begin{cases} \hat{b} = r \frac{\sigma_\eta}{\sigma_\xi}; \\ \hat{a} = E\eta - \hat{b}E\xi. \end{cases}$$

Уравнение регрессии η на ξ имеет вид:

$$\eta = E\eta + r \frac{\sigma_\eta}{\sigma_\xi} (\xi - E\xi).$$

Уравнение регрессии ξ на η имеет вид

$$\xi = E\xi + 1/r \frac{\sigma_\eta}{\sigma_\xi} (\eta - E\eta)$$

Заметим, что каждое из уравнений имеет $E\xi$ своим решением, т.е. графики проходят через точку $(E\xi, E\eta)$ – центр совместного распределения величин ξ и η .

1.5 Характеристические функции

Введем понятие комплекснозначной случайной величины.

Определение. Пусть ξ_1 и ξ_2 – две сл.в., определенные на одном вероятностном пространстве (Ω, \mathcal{F}, P) . Комплексной случайной величиной называется величина $\eta = \xi_1 + i\xi_2$.

Определение Математическим ожиданием комплексной сл.в. называется число вида $E\eta = E\xi_1 + iE\xi_2$, если определены $E\xi_1$, $E\xi_2$.

Определение. Характеристической функцией (х.ф.) сл.в. ξ называется функция $\varphi(t)$

$$\varphi(t) = Ee^{it\xi} = \int_{-\infty}^{+\infty} e^{itx} dF_\xi(x),$$

$F_\xi(x)$ – функция распределения сл.в. ξ .

Пусть сл.в. ξ непрерывна. Тогда $dF_\xi(x) = f(x)dx$, и

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) d(x).$$

Пусть сл.в. ξ дискретна. Тогда $dF_\xi(x)$ – ряд распределения, и

$$\varphi(t) = \sum_{-\infty}^{+\infty} e^{itx_k} p_k,$$

где $p_k = P\{\xi(\omega) = x_k\}$.

Теорема(о существовании х.ф.)

Любая случайная величина имеет х.ф.

Доказательство

Т.к.

$$E|e^{it\xi}| = \int_{-\infty}^{\infty} |e^{it\xi}| dF_{\xi}(x) = \int_{-\infty}^{\infty} |e^{it\xi}| dF_{\xi}(x) = 1,$$

то $Ee^{it\xi} < +\infty$, и, следовательно, х.ф. существует.

Свойства х.ф.

$$1) \varphi(0) = 1, \quad |\varphi(t)| \leq 1, \quad \forall t \in \mathbb{R};$$

Доказательство

$$|Ee^{it\xi}| \leq E|e^{it\xi}| = 1.$$

$$2) \varphi(-t) = \overline{\varphi(t)};$$

Доказательство

$$\begin{aligned} \varphi(-t) &= Ee^{-it\xi} = E(\cos t\xi - i \sin t\xi) = E \cos t\xi - iE \sin t\xi = \\ &= \overline{E \cos t\xi + iE \sin t\xi} = \overline{\varphi(t)} \end{aligned}$$

□

$$3) \text{Х.ф. } \varphi \text{ равномерно непрерывна на } \forall t \in \mathbb{R};$$

Доказательство

По определению требуется показать, что $\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) > 0$ т.ч.

$$\forall |t_1 - t_2| < \delta \quad |\varphi(t_1) - \varphi(t_2)| < \varepsilon.$$

Рассмотрим

$$\begin{aligned} |\varphi(t_1) - \varphi(t_2)| &= \left| \int_{-\infty}^{\infty} (e^{it_1x} - e^{it_2x}) dF_{\xi}(x) \right| \leq \int_{-\infty}^{\infty} |e^{it_1x} - e^{it_2x}| dF_{\xi}(x) = \\ &= \int_{-\infty}^{\infty} |e^{it_2x}| \cdot |e^{i(t_1-t_2)x} - 1| dF_{\xi}(x) \leq \int_{|x| \leq A} |e^{i(t_1-t_2)x} - 1| dF_{\xi}(x) + 2 \int_{|x| > A} dF_{\xi}(x) = \\ &= I_1 + I_2 \end{aligned}$$

Для заданного $\varepsilon > 0$ выберем достаточно большое A , такое чтобы

$$\int_{|x|>A} dF_{\xi}(x) = P\{|x| \leq A\} < \varepsilon/4.$$

Фиксируем найденное A , воспользуемся равномерной непрерывностью функции $e^{i(t_1-t_2)x}$ на ограниченном множестве $|x| \leq A$, и выберем такое $\delta = \delta(\varepsilon)$, что $|e^{i(t_1-t_2)x} - 1| < \varepsilon/2$ при $|t_1 - t_2| < \delta$.

Тогда имеем

$$|\varphi(t_1) - \varphi(t_2)| \leq \frac{\varepsilon}{2} P\{|x| \leq A\} + \frac{\varepsilon}{2} \leq \varepsilon.$$

Таким образом, $\varphi(t)$ равномерно непрерывна.

□

4) Если существуют $E|\xi^n| < \infty$, то $\varphi(t)$ имеет n непрерывных первых производных и

$$\varphi^{(n)}(t) = i^n E\xi^n.$$

Доказательство

Доказательство проводится по индукции. Пусть $n = 1$. По условию теоремы $E|\xi| = \int_{-\infty}^{\infty} |x| dF(x) < \infty$.

Рассмотрим формальную производную х.ф. $\varphi'(t) = \int ix e^{itx} dF(x)$. Теорема о дифференцировании под знаком интеграла требует, чтобы интеграл сходился равномерно. Это требование выполнено, в силу того, что $|e^{itx}|$ ограничена, а $\int |x| dF(x)$ конечен по условию. Таким образом, дифференцирование возможно, и следовательно:

$$\varphi'(t) = i \int x e^{itx} dF(x),$$

$$\varphi'(0) = i \int x dF(x) = iE\xi.$$

□

5) Если ξ_1, \dots, ξ_n — независимые сл.в., то х.ф. их суммы

$$\xi = \xi_1 + \dots + \xi_n$$

имеет вид:

$$\varphi_\xi(t) = \prod_{i=1}^n \varphi_{\xi_i}(t).$$

6) Если $\eta = a\xi + b$ и $\varphi_\xi(t)$ х.ф. сл.в. ξ , то

$$\varphi_\eta(t) = e^{itb} \varphi_\xi(ta).$$

Теорема единственности.

Пусть $F(x)$ и $G(x)$ — функции распределения с одной и той же х.ф. $\varphi(t)$, т.е.

$$\varphi(t) = \int_{-\infty}^{\infty} |x| dF(x), \quad \varphi(t) = \int_{-\infty}^{\infty} |x| dG(x).$$

Тогда $F(x) \equiv G(x)$ в смысле $\mathbb{P}\{\omega : F(x) \neq G(x)\} = 0$.

Теорема (формула обращения).

Пусть $\varphi(t)$ — х.ф. функции распределения $F(x)$.

Тогда для любых $a < b$

$$F(b) - F(a) = \frac{1}{2\pi} \lim_{C \rightarrow \infty} \int_{-C}^C \frac{e^{ita} - e^{itb}}{it} \varphi(t) dt.$$

Теорема Бохнера–Хинчина.

Пусть $\varphi_\xi(t)$, $t \in \mathbb{R}$, такая, что $\varphi_\xi(0) = 1$. Для того, чтобы функция $\varphi_\xi(t)$ являлась х.ф. необходимо и достаточно, чтобы $\forall t_1, \dots, t_n \in \mathbb{R}$, $\lambda_1, \dots, \lambda_n \in \mathbb{C}$

$$\sum_{i,j=1}^n \varphi_\xi(t_i - t_j) \lambda_i \bar{\lambda}_j \geq 0.$$

Теорема Поля.

Пусть непрерывная, четная и выпуклая вниз функция $\varphi_\xi(t)$ такова, что $\varphi_\xi(t) \geq 0$, $\varphi_\xi(0) = 1$, $\varphi_\xi(t) \rightarrow 0$ при $t \rightarrow \infty$. Тогда $\varphi_\xi(t)$ — х.ф.

Теорема Марцинкевича.

Если х.ф. имеет вид $\exp(P_n(t))$, $P_n(t) = a_0 t^n + a_1 t^{n-1} + \dots + a_{n+1} t^0$, то степень полинома не может быть больше 2.

Метод х.ф. для доказательства теорем основан на использовании теоремы непрерывности.

Теорема непрерывности (Леви).

Пусть $\{F_n(x)\}_{n=1}^\infty$ — последовательность функций распределений, а $\{\varphi_n(x)\}_{n=1}^\infty$ — последовательность соответствующих характеристических функций, $t, x \in \mathbb{R}$.

1) Если $F_n(x) \xrightarrow{\omega} F(x)$, то $\varphi_n(x) \rightarrow \varphi(x)$, $t \in \mathbb{R}$;

2) Если для любого $t \in \mathbb{R}$ существует $\lim_{n \rightarrow \infty} \varphi_n(x)$ и функция $\varphi(x) = \lim_{n \rightarrow \infty} \varphi_n(x)$ непрерывна в $t = 0$, то $\varphi(x)$ является х.ф. распределения $F(x)$, такого что $F_n(x) \xrightarrow{\omega} F(x)$.

1.6 Законы больших чисел. Сходимость последовательностей случайных величин

Пусть $\{\xi_n\}_{n=1}^\infty$ — последовательность случайных величин, заданных на одном и том же вероятностном пространстве (Ω, \mathcal{F}, P) .

Определение. Последовательность случайных величин $\{\xi_n\}_{n=1}^\infty$ называется сходящейся **по вероятности** к случайной величине ξ , если для любого $\varepsilon > 0$

$$P\{|\xi_n - \xi| < \varepsilon\} \rightarrow 1, \quad n \rightarrow \infty.$$

Сходимость **по вероятности** обозначается $\{\xi_n\} \xrightarrow{p} \xi$.

Определение. Последовательность случайных величин $\{\xi_n\}_{n=1}^\infty$ называется сходящейся **почти наверное** к случайной величине ξ , если для лю-

бого $\varepsilon > 0$

$$P\{\omega : |\xi_n(\omega) - \xi(\omega)|\} \rightarrow 0, \quad n \rightarrow \infty,$$

или

$$P\{\omega : \lim_{n \rightarrow \infty} \xi_n = \xi\} = 1, \quad n \rightarrow \infty.$$

Сходимость **почти наверное** обозначается $\{\xi_n\} \xrightarrow{\text{п.н.}} \xi$

Определение. Последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ называется сходящейся **в среднем** к случайной величине ξ , если для любого n существуют $E\xi_n$, $E\xi$ и

$$\lim_{n \rightarrow \infty} E|\xi_n - \xi| = 0, \quad n \rightarrow \infty.$$

Сходимость **в среднем** обозначается $\{\xi_n\} \xrightarrow{L} \xi$.

Определение. Последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ называется сходящейся **по распределению** к случайной величине ξ , если для любого $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \quad n \rightarrow \infty,$$

где $F_n(x) = P\{\xi_n < x\}$, $F(x) = P\{\xi < x\}$ — функции распределения случайных величин.

Сходимость **по распределению** обозначается $\{\xi_n\} \xrightarrow{d} \xi$.

Теорема.

Из сходимости почти наверное вытекает сходимость по вероятности,

$$\{\xi_n\} \xrightarrow{\text{п.н.}} \xi \Rightarrow \{\xi_n\} \xrightarrow{p} \xi.$$

Из сходимости по вероятности вытекает сходимость по распределению,

$$\{\xi_n\} \xrightarrow{p} \xi \Rightarrow \{\xi_n\} \xrightarrow{d} \xi.$$

Из сходимости в среднем вытекает сходимость по вероятности,

$$\{\xi_n\} \xrightarrow{L} \xi \Rightarrow \{\xi_n\} \xrightarrow{p} \xi.$$

Определение. Говорят, что последовательность случайных величин $\{\xi_i\}_{i=1}^{\infty}$ удовлетворяет **закону больших чисел** (ЗБЧ), если последовательность $\{S_n\}_{n=1}^{\infty}$, составленная из частных сумм $S_n = \xi_1 + \xi_2 + \dots + \xi_n$, при любом ε удовлетворяет

$$P\left\{\left|\frac{S_n}{n} - \frac{ES_n}{n}\right| < \varepsilon\right\} \rightarrow 1, \quad \text{при } n \rightarrow \infty.$$

Таким образом, если последовательность удовлетворяет ЗБЧ, то при достаточно большом количестве слагаемых среднее из наблюдаемых значений случайных величин и среднее из их мат.ожиданий практически равны с вероятностью 1.

В терминах сходимости сл.величин ЗБЧ означает, что

$$\frac{S_n}{n} - \frac{ES_n}{n} \xrightarrow{p} 0.$$

Теорема. (Закон больших чисел в форме Чебышёва.)

Пусть $\{\xi_i\}_{i=1}^{\infty}$ — последовательность независимых случайных величин со средними $E\xi_i = a_i < \infty$, при всех $i \in \mathbb{Z}_+$, и дисперсиями $D\xi_i \leq C$, при всех $i \in \mathbb{Z}_+$, где $C = \text{const}$.

Тогда последовательность $\{\xi_i\}_{i=1}^{\infty}$ удовлетворяет ЗБЧ, и при любом ε удовлетворяет неравенству

$$P\left\{\left|\frac{S_n}{n} - \frac{ES_n}{n}\right| > \varepsilon\right\} \leq \frac{C}{n\varepsilon^2}.$$

Напомним, что $\xi_0 \sim N(0, 1)$ — стандартная нормальная случайная величина с $E\xi_0 = 0$ и $D\xi_0 = 1$. Обозначим также через $F_{\xi_0}(x)$ — функцию распределения вероятностей случайной величины $\xi_0 \sim N(0, 1)$.

Центральная предельная теорема (ЦПТ).

Пусть $\{\xi_i\}_{i=1}^{\infty}$ — последовательность независимых одинаково распределенных случайных величин с $E\xi_i = a < \infty$ и $D\xi_i = \sigma^2$, при каждом $i \in \mathbb{Z}_+$.

Тогда для любого $x \in \mathbb{R}$ при $n \rightarrow \infty$

$$P\left\{\frac{S_n/n - a}{\sigma/\sqrt{n}} < x\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

Замечание.

Введем в рассмотрение сл. величины

$$Z_n = \frac{S_n - ES_n}{\sqrt{DS_n}} = \frac{S_n/n - a}{\sigma/\sqrt{n}}.$$

Тогда выражение в левой части предела есть функция распределения сл. в. Z_n , а справа записана функция распределения величины $\xi_0 \sim N(0, 1)$. Следовательно, ЦПТ означает, что "суммы" большого количества независимых и одинаково распределенных сл. в. "ведут" себя по нормальному закону, а именно

$$F_{Z_n}(x) \rightarrow F_{\xi_0}(x), n \rightarrow \infty, \quad \text{или} \quad Z_n \xrightarrow{d} \xi_0$$

2 МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

2.1 Вариационные ряды и их характеристики

Рассмотрим постановку задачи математической статистики: по результатам наблюдения за некоторой случайной величиной ξ требуется сделать выводы о неизвестном законе распределения этой величины $\mathcal{L}(x, \theta)$ либо о неизвестных параметрах $\theta_1, \dots, \theta_n$ известного распределения.

Пусть ξ — случайная величина (сл.в.) с некоторой (теоретической) функцией распределения $F_\xi(x) = P\{\xi < x\}$, $x \in \mathbb{R}$.

Определение. Совокупность n независимых одинаково распределенных случайных величин X_1, X_2, \dots, X_n называется выборкой, извлеченной из распределения случайной величины ξ .

Определение. Набор n значений x_1, x_2, \dots, x_n сл. величин X_1, X_2, \dots, X_n называется реализацией выборки или выборкой из генеральной совокупности значений случайной величины ξ .

Под генеральной совокупностью понимается множество всех возможных значений случайной величины ξ . Объем совокупности есть количество всех ее элементов, объем выборки или выборочной совокупности обозначается n , генеральной совокупности — N .

Определение. Выборочным пространством называется множество $\mathbf{X}_n = \{\bar{\mathbf{X}}_n \equiv (X_1, X_2, \dots, X_n)\}$ всех возможных выборок объема n , извлеченных из распределения $\mathcal{L}(x, \theta)$ случайной величины ξ .

В силу того, что наблюдать все выборки из выборочного пространства мы не можем, статистические выводы о наблюдаемой случайной величине приходится делать по имеющейся выборке. *Выборочным методом* в статистике называется метод, на основании которого, за свойства генеральной совокупности принимаются свойства выборочной совокупности. Для того, чтобы полученные при таком подходе оценки были более точными, выборка должна

быть в первую очередь *репрезентативной*, т.е. отражать истинные свойства генеральной совокупности. Обеспечивается репрезентативность, как правило, с помощью методов отбора единиц наблюдения в выборку.

Первым этапом статистического анализа данных является группировка полученных в результате наблюдения за процессом или явлением данных.

Определение. Вариационным рядом называется последовательность расположенных в порядке неубывания элементов выборки $x_1^* \leq x_2^* \leq \dots \leq x_n^*$.

Элементы вариационного ряда называются вариантами.

Определение. Точечным вариационным рядом называется таблица вида

x_i	x_1	x_2	\dots	x_m
n_i	n_1	n_2	\dots	n_m

где x_i — варианты, n_i — частоты, m — количество групп (различных значений), $n = \sum_{i=1}^m n_i$ — объем выборки.

Для графического представления точечных вариационных рядов используется полигон частот — ломаная с вершинами в точках (x_i, n_i) .

Определение. Интервальным вариационным рядом называется таблица вида

x_i	$[x_1, x_2]$	$(x_2, x_3]$	\dots	$(x_m, x_{m+1}]$
n_i	n_1	n_2	\dots	n_m

где x_i — варианты, n_i — частоты, m — количество групп (интервалов), $n = \sum_{i=1}^m n_i$ — объем выборки.

Для графического представления интервальных вариационных рядов используется гистограмма частот — фигура, составленная из прямоугольников, одной стороной которых служат интервалы $(x_i, x_{i+1}]$, а длина второй равна n_i .

Пусть X_1, X_2, \dots, X_n выборка из распределения сл.в. ξ с теоретической функцией распределения $F_\xi(x)$.

Определение. Эмпирической функцией распределения (ЭФР) называется функция

$$\widetilde{F}_n(x) = \frac{1}{n} \sum_{i=1}^n e(x - X_i), \quad (4)$$

где $e(x) = 1$, при $x > 0$, $e(x) = 0$, при $x \leq 0$.

Таким образом, если $X_i < x$, то $e(x) = 1$, если $X_i \geq x$, то $e(x) = 0$, а сумма $e(x - X_i)$ будет равна количеству элементов выборки, которые приняли значение, строго меньше некоторого $x \in \mathbb{R}$.

Пусть x_1, x_2, \dots, x_n — реализация выборки X_1, X_2, \dots, X_n , т.е. наблюдавшиеся значения сл.в. ξ . Обозначим $\mu(x)$ — число элементов выборки, строго меньших $x \in \mathbb{R}$.

Тогда эмпирическая функция распределения $\widetilde{F}_n(x)$ может быть определена как

$$\widetilde{F}_n(x) = \frac{\mu(x)}{n}. \quad (5)$$

С помощью полигона и гистограммы частот можно оценить вид функции плотности наблюдаемого распределения, ЭФР является оценкой (теоретической) функции распределения. С помощью анализа графиков можно сделать предположение (выдвинуть статистическую гипотезу) о типе наблюдаемого распределения.

Числовые характеристики вариационных рядов

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{L}(x, \theta)$ — выборка из распределения сл.в. ξ .

Определение. Выборочным средним называется величина

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (6)$$

Если данные представлены в виде точечного или интервального вариационного ряда, то для вычисления используют формулу:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^m x_i * n_i, \quad (7)$$

где m – количество групп в точечном или интервалов в интервальном вариационном ряду, n_i – частота, т.е. количество элементов выборки, принадлежащих i -той группе или i -тому интервалу, x_i – варианта для точечного ряда и середина i -того интервала для интервального ряда.

Определение. Выборочной дисперсией (смещенной) называется величина

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2. \quad (8)$$

Она характеризует среднее из квадратов отклонений наблюдаемой величины от выборочного среднего. Величина $S = \sqrt{S^2}$ называется выборочным средним квадратическим отклонением (смещенным) величин выборки от выборочного среднего.

Определение. Выборочной дисперсией (несмещенной) называется величина

$$\tilde{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2. \quad (9)$$

Аналогично, величина $\tilde{\sigma} = \sqrt{\tilde{\sigma}^2}$ называется выборочным несмещенным средним квадратическим отклонением. Очевидно, что смещенная и несмещенная выборочные дисперсии связаны формулой

$$\tilde{\sigma}^2 = \frac{n}{n-1} S^2.$$

Кроме того, при любом n $\tilde{\sigma}^2 \geq S^2$, а $\lim_{n \rightarrow \infty} S^2 = \tilde{\sigma}^2$

Если данные представлены в виде точечного или интервального вариационного ряда, то для вычисления используют формулу:

$$S^2 = \frac{1}{n} \sum_{i=1}^m (x_i - \bar{x})^2 n_i. \quad (10)$$

или

$$S^2 = \left(\frac{1}{n} \sum_{i=1}^m x_i^2 n_i \right) - \bar{x}^2. \quad (11)$$

Здесь m – количество групп в точечном или интервалов в интервальном вариационных рядах, n_i – частота, т.е. количество элементов выборки,

принадлежащих i -той группе или i -тому интервалу, x_i – варианта для точечного ряда и середина i -того интервала для интервального ряда.

2.2 Оценки параметров распределения

Пусть X_1, X_2, \dots, X_n выборка, извлеченная из распределения $\mathcal{L}(x, \theta)$ случайной величины ξ . Пусть требуется получить оценку неизвестного значения параметра θ .

Например, известно, что рост человека распределен по нормальному закону, т.е. $\xi \sim N(a, \sigma^2)$, где a – среднее значение величины, и σ^2 – дисперсия, являются параметрами распределения. По результатам измерения роста 100 человек можно получить выборочное среднее значение, которое и будет оценкой параметра a .

При построении оценок необходимо, чтобы они удовлетворяли определенным свойствам.

Определение. Оценкой параметра θ распределения $\mathcal{L}(x, \theta)$ называется величина $\tilde{\theta}_n = f(X_1, X_2, \dots, X_n)$, где $f(t_1, t_2, \dots, t_n)$ некоторая непрерывная функция.

Заметим, что $\tilde{\theta}_n$ как функция от случайных величин также есть случайная величина.

Определение. Оценка $\tilde{\theta}_n$ параметра θ называется **несмещенной**, если

$$E\tilde{\theta}_n = \theta.$$

(математическое ожидание оценки равно оцениваемому параметру).

Определение. Оценка $\tilde{\theta}_n$ параметра θ называется **состоятельной**, если при всех $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|\tilde{\theta}_n - \theta| < \varepsilon\} = 1. \quad (12)$$

(т.е. если оценка состоятельная, то при достаточно большом объеме выборки оценка параметра с высокой вероятностью практически равна параметру).

Также можно заметить, что состоятельность оценки означает сходимость по вероятности последовательности оценок $\tilde{\theta}_n$ к параметру $\theta: \{\theta_n\} \xrightarrow{P} \theta$.

Оценка \bar{x} является несмещенной и состоятельной оценкой математического ожидания $E\xi$, ЭФР $\widetilde{F}_n(x)$ — несмещенной и состоятельной оценкой функции распределения $F_\xi(x) = P\{\xi < x\}$.

Теорема. (о несмещенной и состоятельной оценке математического ожидания)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{L}(x, \theta)$ — выборка из распределения сл.в. ξ с конечным математическим ожиданием $E\xi = a < \infty$. Выборочное среднее \bar{x} является несмещенной и состоятельной оценкой математического ожидания.

Доказательство.

Докажем несмещенность $E\bar{x} = E\xi = a$. Действительно

$$\begin{aligned} E\bar{x} &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \\ &= \frac{1}{n} \sum_{i=1}^n E\xi = \frac{1}{n} n * E\xi = E\xi = a. \end{aligned}$$

Докажем состоятельность, т.е. что для любого $\varepsilon > 0$ имеет место

$$\lim_{n \rightarrow \infty} P\{|\bar{x} - a| < \varepsilon\} = 1.$$

По неравенству Чебышёва для любого $\varepsilon > 0$

$$\begin{aligned} P\{|\bar{x} - a| < \varepsilon\} &\geq \frac{D(\bar{x})}{\varepsilon^2} = \frac{1}{\varepsilon^2} D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \\ &= \frac{1}{\varepsilon^2 n^2} \sum_{i=1}^n D(X_i) = \frac{1}{\varepsilon^2 n^2} n \cdot D\xi = \frac{1}{\varepsilon^2 n} D\xi. \end{aligned}$$

С другой стороны, по свойству вероятности имеем

$$P\{|\bar{x} - a| < \varepsilon\} \leq 1.$$

Таким образом, доказано, что при любом $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|\bar{x} - a| < \varepsilon\} = 1,$$

и оценка \bar{x} является состоятельной.

Теорема. (о несмещенной и состоятельной оценке функции распределения)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{L}(x, \theta)$ — выборка из распределения сл.в. ξ с функцией распределения $F_\xi(x) = P\{\xi < x\}$. Эмпирическая функция распределения (ЭФР) $\widetilde{F}_n(x)$ является несмещенной и состоятельной оценкой функции распределения $F_\xi(x)$.

Теорема. (о несмещенной оценке дисперсии)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{L}(x, \theta)$ — выборка из распределения сл.в. ξ с конечным математическим ожиданием $E\xi = a < \infty$ и дисперсией $D\xi = \sigma^2$. Выборочная дисперсия

$$\tilde{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2$$

является несмещенной оценкой дисперсии.

2.3 Эффективные оценки. Неравенство Рао-Крамера.

Рассмотрим параметрическую статистическую модель

$$\{\mathbf{X}_n; F_\xi(x, \theta) : \theta \in \Theta\},$$

где \mathbf{X}_n — выборочное пространство, $F_\xi(x, \theta)$ — функция распределения наблюдаемой сл.в. ξ , известная с точностью до параметра θ , Θ — множество значений параметра θ .

Определение. Оценка $\tilde{\theta}^*$ параметра θ называется эффективной, если $D\tilde{\theta}^* = \inf_{\theta \in \Theta} D\tilde{\theta}$.

Теорема.(о единственности эффективной оценки)

Пусть $\hat{\theta}(\bar{X}_n)$ и $\tilde{\theta}(\bar{X}_n)$ две эффективные оценки для параметра θ параметрической модели $\{\mathbf{X}_n; F_\xi(x, \theta) : \theta \in \Theta\}$.

Тогда $\hat{\theta}(\bar{X}_n) = \tilde{\theta}(\bar{X}_n)$, где $P\{\bar{X}_n \in \{\bar{x}_n : \hat{\theta}(\bar{x}_n) \neq \tilde{\theta}(\bar{x}_n)\}\} = 0$.

Доказательство.

Рассмотрим оценку

$$\theta^*(\bar{X}_n) = \frac{\hat{\theta}(\bar{X}_n) + \tilde{\theta}(\bar{X}_n)}{2}.$$

Т.к. $\hat{\theta}(\bar{X}_n)$ и $\tilde{\theta}(\bar{X}_n)$ эффективные, то $D\hat{\theta}(\bar{X}_n) = D\tilde{\theta}(\bar{X}_n)$, и $E\hat{\theta}(\bar{X}_n) = E\tilde{\theta}(\bar{X}_n) = \theta$.

Найдем характеристики оценки $\theta^*(\bar{X}_n)$.

$$E\theta^*(\bar{X}_n) = E\frac{\hat{\theta}(\bar{X}_n) + \tilde{\theta}(\bar{X}_n)}{2} = \frac{1}{2}(E\hat{\theta}(\bar{X}_n) + E\tilde{\theta}(\bar{X}_n)) = \theta.$$

$$\begin{aligned} D\theta^*(\bar{X}_n) &= D\frac{\hat{\theta}(\bar{X}_n) + \tilde{\theta}(\bar{X}_n)}{2} = \frac{1}{4}(D\hat{\theta}(\bar{X}_n) + D\tilde{\theta}(\bar{X}_n) + 2cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n))) = \\ &= \frac{1}{2}D\hat{\theta}(\bar{X}_n) + \frac{1}{2}cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n)). \end{aligned}$$

Имеем

$$\begin{aligned} |cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n))| &= |E(\theta - \hat{\theta}(\bar{X}_n))(\theta - \tilde{\theta}(\bar{X}_n))| \leq \\ &\leq \sqrt{E(\theta - \hat{\theta}(\bar{X}_n))^2 E(\theta - \tilde{\theta}(\bar{X}_n))^2} = \sqrt{D\hat{\theta}(\bar{X}_n) D\tilde{\theta}(\bar{X}_n)} = D\hat{\theta}(\bar{X}_n) \end{aligned}$$

Тогда

$$\begin{aligned} D\theta^* &= |D\theta^*| = \frac{1}{2}|D\hat{\theta}(\bar{X}_n) + cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n))| \leq \\ &\leq \frac{1}{2}|D\hat{\theta}(\bar{X}_n)| + \frac{1}{2}|cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n))| \leq D\hat{\theta}(\bar{X}_n). \end{aligned}$$

Таким образом, θ^* несмещенная оценка, и ее дисперсия $D\theta^*(\bar{X}_n) = D\hat{\theta}(\bar{X}_n)$, и следовательно, оценка $\theta^*(\bar{X}_n)$ также является эффективной.

Получаем уравнение

$$D\theta^*(\bar{X}_n) = \frac{1}{2}D\hat{\theta}(\bar{X}_n) + \frac{1}{2}cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n)),$$

откуда следует $cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n)) = D\hat{\theta}(\bar{X}_n)$.

Вычислим коэффициент корреляции

$$r(\hat{\theta}, \tilde{\theta}) = \frac{cov(\hat{\theta}(\bar{X}_n), \tilde{\theta}(\bar{X}_n))}{\sqrt{D\hat{\theta}(\bar{X}_n)D\tilde{\theta}(\bar{X}_n)}} = \frac{D\hat{\theta}(\bar{X}_n)}{D\hat{\theta}(\bar{X}_n)} = 1.$$

Следовательно, величины $\hat{\theta}(\bar{X}_n)$ и $\tilde{\theta}(\bar{X}_n)$ линейно связаны

$$\hat{\theta}(\bar{X}_n) = a\tilde{\theta}(\bar{X}_n) + b.$$

Пользуясь несмещенностью, получаем

$$E\hat{\theta}(\bar{X}_n) = aE\tilde{\theta}(\bar{X}_n) + b$$

$$\theta = a\theta + b.$$

Отсюда следует, что $a = 1$, $b = 0$, $\hat{\theta}(\bar{X}_n) = \tilde{\theta}(\bar{X}_n)$.

САРАТОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО

Неравенство Рао-Крамера. (Информационное неравенство)

Рассмотрим параметрическую статистическую модель

$$\{\mathbf{X}_n; F_\xi(x, \theta) : \theta \in \Theta\},$$

где \mathbf{X}_n – выборочное пространство, $F_\xi(x, \theta)$ – функция распределения наблюдаемой сл.в. ξ , известная с точностью до параметра θ , Θ – множество значений параметра θ .

Обозначим $f(x, \theta)$ функцию плотности для абсолютно непрерывных сл.в. и закон распределения для дискретных сл.в.

Определение. Параметрическая модель называется **регулярной**, если она удовлетворяет условиям регулярности.

Условия регулярности.

1. параметрическое множество Θ – открытый интервал;
2. носитель распределения – множество $A = \{x : f(x, \theta) > 0\}$ не зависит от параметра;
3. для любого $\theta \in \Theta$ и $x \in A$ существует и конечна производная

$$\frac{\partial f(x, \theta)}{\partial \theta};$$

4. для любого $\theta \in \Theta$

$$E \frac{\partial \ln f(x, \theta)}{\partial \theta} = 0; \quad E \left(\frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2 < \infty;$$

5. в интеграле $\int f(x, \theta) dx$ допустимо дифференцирование по параметру под знаком интеграла дважды.

Теорема.(Неравенство Рао-Крамера)

Пусть параметрическая модель $\{\mathbf{X}_n; F_\xi(x, \theta) : \theta \in \Theta\}$ регулярна, и $\tilde{\theta}(\bar{X}_n)$ несмещенная оценка параметра θ .

Тогда

$$D(\tilde{\theta}(\bar{X}_n)) \geq \frac{1}{nI(\theta)},$$

где

$$I(\theta) = \left(E \frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2$$

– количество информации по Фишеру в одном наблюдении.

2.4 Методы построения оценок параметров распределения

К основным методам относятся **метод моментов** и **метод максимального правдоподобия**.

Начальным моментом порядка k , $k \in \mathbb{Z}_+$ распределения случайной величины ξ называется величина

$$m_k = E\xi^k, \quad \text{где } E\xi^k < \infty.$$

Центральным моментом порядка k , $k \in \mathbb{Z}_+$ распределения случайной величины ξ называется величина

$$\mu_k = E(\xi - E\xi)^k, \quad \text{где } E\xi < \infty.$$

Таким образом, $m_1 = E\xi$ – математическое ожидание сл. величины ξ , $m_2 = E\xi^2$, $\mu_1 = E(\xi - E\xi) = 0$, $\mu_2 = E(\xi - E\xi)^2 = D\xi$ – дисперсия сл.в. ξ .

Выборочным начальным моментом порядка k называется величина

$$\tilde{m}_k = \frac{1}{n} \sum_{i=1}^n X_i^k. \quad (13)$$

Выборочным центральным моментом порядка k называется величина

$$\tilde{\mu}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^k. \quad (14)$$

Таким образом, $\tilde{m}_1 = \frac{1}{n} \sum_{i=1}^n X_i = \bar{x}$ – выборочное среднее (несмещенная оценка математического ожидания сл. величины ξ), $\tilde{\mu}_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2 = S^2$ – смещенная оценка дисперсии сл.в. ξ .

Метод моментов состоит в том, что за оценку параметров распределения принимается решение системы уравнений

$$\begin{cases} m_k = \tilde{m}_k, \\ \mu_k = \tilde{\mu}_k. \end{cases}$$

Количество и тип уравнений в системе выбирается в зависимости от распределения.

Пусть наблюдаемая случайная величина имеет распределение с неизвестным параметром θ . Обозначим $P(x, \theta)$ закон распределения этой величины, подразумевая, что $P(x, \theta) = P\{\xi = x\}$, если величина дискретна, $P(x, \theta) = f(x)$, если $P(x, \theta) = f(x)$ величина абсолютно непрерывна.

Функцией правдоподобия называется функция

$$L(\bar{X}_n, \theta) = \prod_{i=1}^n P(X_i, \theta).$$

Таким образом, для выборки из распределения дискретной случайной величины

$$L(\bar{X}_n, \theta) = P(X_1, \theta) \cdot P(X_2, \theta) \cdot \dots \cdot P(X_n, \theta),$$

для выборки из распределения абсолютно непрерывной случайной величины

$$L(\bar{X}_n, \theta) = f(X_1, \theta) \cdot f(X_2, \theta) \cdot \dots \cdot f(X_n, \theta).$$

Метод максимального правдоподобия (ММП) состоит в том, что за оценку параметра θ принимается точка максимума функции правдоподобия $L(x, \theta)$.

Таким образом, алгоритм ММП таков:

1. Составить функцию правдоподобия $L(x, \theta)$;

2. Взять натуральный логарифм $\ln L(x, \theta)$;
3. Найти производную $\frac{\partial \ln L(x, \theta)}{\partial \theta}$ и решение уравнения $\frac{\partial \ln L(x, \theta)}{\partial \theta} = 0$;
4. Доказать, что решение уравнения является точкой максимума

$$\frac{\partial^2 \ln L(x, \theta)}{\partial \theta^2} < 0.$$

Рассмотрим применение ММП.

Задача. Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma)$ выборка из нормального распределения с математическим ожиданием a и дисперсией σ^2 . Оценим методом максимального правдоподобия параметры нормального распределения.

Составим функцию правдоподобия

$$L(\bar{X}_n, a, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-(X_i - a)^2 / 2\sigma^2} = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a)^2 \right).$$

Возьмем \ln

$$\ln L(\bar{X}_n, a, \sigma) = -\frac{n}{2} \ln 2\pi - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a)^2.$$

Вычислим производные и решим соответствующие уравнения

$$\begin{cases} \frac{\partial \ln L(\bar{X}_n, a, \sigma)}{\partial a} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - a) = 0, \\ \frac{\partial \ln L(\bar{X}_n, a, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (X_i - a)^2 = 0. \end{cases}$$

Получаем:

$$\begin{cases} n\bar{x} - na = 0, \\ -n + \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - a)^2 = 0. \end{cases} \quad \begin{cases} a = \bar{x}, \\ -n + \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{x})^2 = 0. \end{cases}$$

Решение этой системы, таким образом, имеет вид $a = \bar{x}$, $\sigma^2 = S^2$.

Докажем, что эти значения доставляют максимум функции правдоподобия.

$$\left\{ \begin{array}{l} \frac{\partial^2 \ln L(\bar{X}_n, a, \sigma)}{\partial a^2} = -\frac{n}{\sigma^2} \Big|_{\sigma^2=S^2} = -\frac{n}{S^2} < 0, \\ \frac{\partial^2 \ln L(\bar{X}_n, a, \sigma)}{\partial \sigma^2} = \frac{n}{\sigma^2} - 3\frac{1}{\sigma^4} \sum_{i=1}^n (X_i - a)^2 \Big|_{a=\bar{x}, \sigma^2=S^2} = -\frac{2n}{S^2} < 0, \\ \frac{\partial^2 \ln L(\bar{X}_n, a, \sigma)}{\partial a \partial \sigma} = -\frac{2}{\sigma^3} \sum_{i=1}^n (X_i - a) \Big|_{a=\bar{x}, \sigma^2=S^2} = -\frac{2}{\sigma^3} (n\bar{x} - n\bar{x}) = 0. \end{array} \right.$$

Таким образом, оценками, полученными ММП, параметров нормального распределения являются $\hat{a}_{\text{ММП}} = \bar{x}$, $\hat{\sigma}^2_{\text{ММП}} = S^2$.

2.5 Доверительные интервалы

Если по некоторым причинам невозможно построить выборку достаточного объема, то даже для несмещенных и состоятельных оценок равенство $\tilde{\theta} \approx \theta$ не может быть выполнено. В таких случаях оценка параметра θ строится в виде интервала.

Определение. Доверительным интервалом надежности γ называется интервал со случайными концами $(\tilde{\theta}_1, \tilde{\theta}_2)$, который покрывает неизвестное значение параметра θ с вероятностью, не меньшей γ , т.е.

$$P\{\theta \in (\tilde{\theta}_1, \tilde{\theta}_2)\} \geq \gamma. \quad (15)$$

Вероятность γ называется также доверительной вероятностью, ее значения обычно выбирают близкими к единице: 0,9; 0,95; 0,99 и т.д.

Точность интервальной оценки существенно зависит от того, что известно о наблюдаемом распределении. Так, если нет никакой информации о типе наблюдаемого распределения, то построенная оценка будет "грубой". Если же известен закон распределения, то доверительный интервал можно построить более точно.

Прежде чем рассматривать построение доверительных интервалов, изучим некоторые важные распределения.

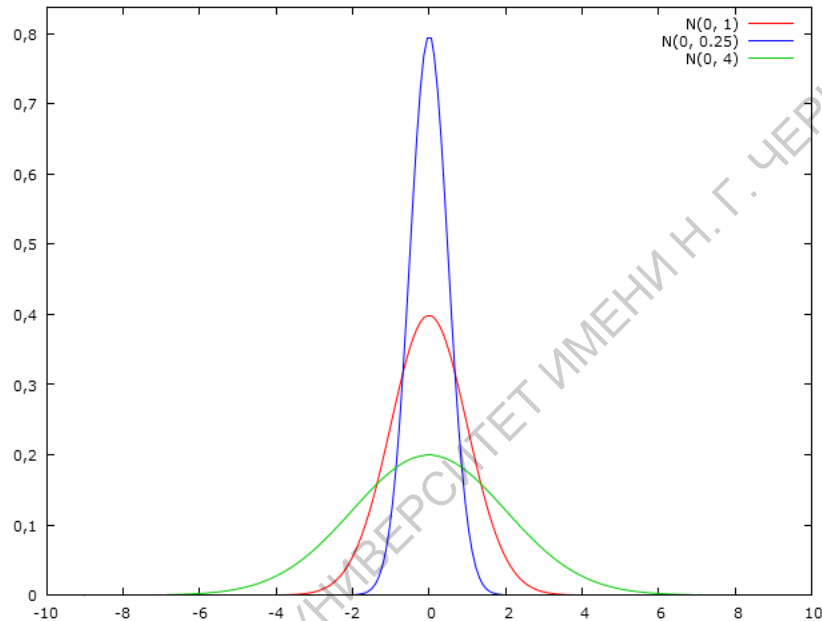
Стандартное нормальное распределение $\mathcal{N}(0, 1)$

Сл. в. $\xi_0 \sim \mathcal{N}(0, 1)$ имеет функцию плотности

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

Числовые характеристики стандартной нормальной величины $E\xi = 0$, $D\xi = \sigma^2 = 1$.

Графики функции плотности $\xi \sim \mathcal{N}(a, \sigma^2)$ представлены на рисунке



Распределение хи-квадрат $\chi^2(n)$

Рассмотрим n независимых одинаково распределенных случайных величин со стандартным нормальным распределением $\xi_1, \xi_2, \dots, \xi_n$.

Случайной величиной с распределением $\chi^2(n)$ с n степенями свободы называется величина $\chi^2(n) = \xi_1^2 + \xi_2^2 + \dots + \xi_n^2$.

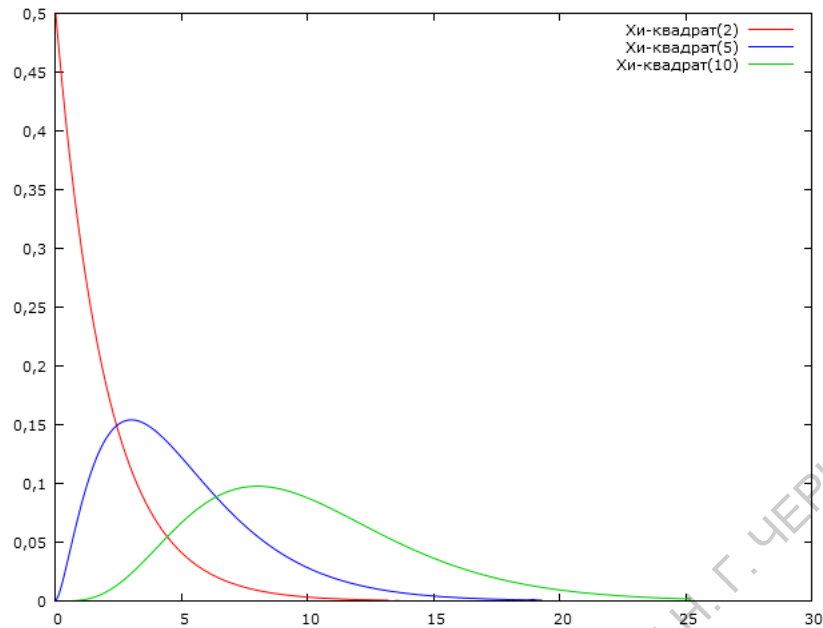
Функция плотности этой сл. в. имеет вид

$$f(x) = \frac{x^{n/2-1} e^{-x/2}}{\Gamma(n/2) 2^{n/2}}, \quad \text{при } x \geq 0, \quad f(x) = 0, \quad \text{при } x < 0.$$

Здесь $\Gamma(\lambda) = \int_0^{\infty} t^{\lambda-1} e^{-t} dt$, $\lambda > 0$, — гамма-функция.

Числовые характеристики $E\xi = n$, $D\xi = \sigma^2 = 2n$.

На рисунках приведены графики распределения хи-квадрат при различных значениях степеней свободы.



Распределение Стьюдента

Рассмотрим независимые одинаково распределенные случайных величин со стандартным нормальным распределением $\eta, \xi_1, \xi_2, \dots, \xi_n$.

Случайной величиной с распределением Стьюдента или T – распределением с n степенями свободы называется величина

$$t(n) = \frac{\eta}{\sqrt{\chi^2(n)/n}} = \frac{\eta}{\sqrt{(\xi_1^2 + \xi_2^2 + \dots + \xi_n^2)/n}}$$

Функция плотности этой сл. в. имеет вид

$$f(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}. \quad (16)$$

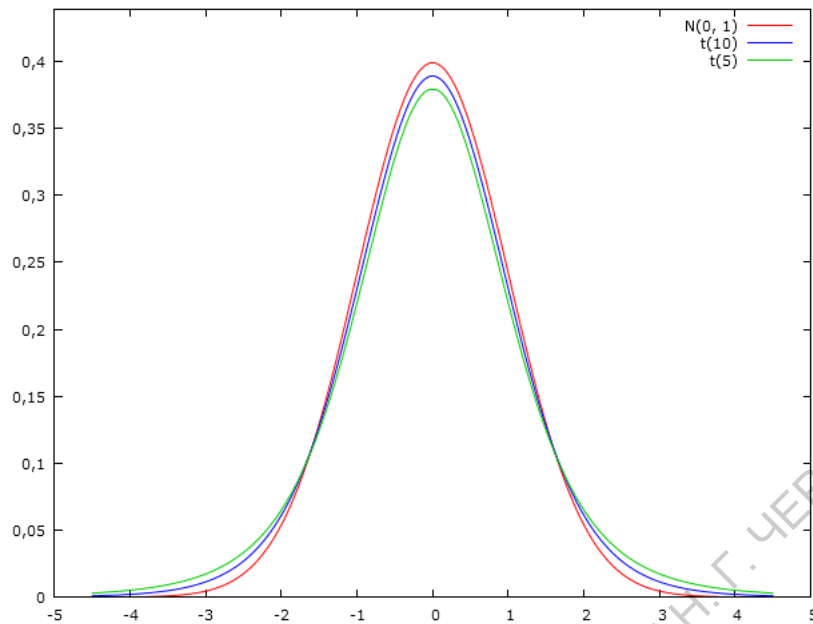
При $n = 1$ функция плотности (16) принимает вид

$$f(x) = \frac{1}{\sqrt{\pi}(1+x^2)},$$

т.е. получаем распределение Коши.

Числовые характеристики при $n > 2$ $E\xi = 0$, $D\xi = \sigma^2 = n/(n-2)$.

На рисунке приведен графики функции плотности распределения Стьюдента для числа степеней свободы $n = 5, 10$, и график функции плотности стандартного нормального распределения.



Лемма Фишера(следствие)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma)$ выборка из нормального распределения математическим ожиданием a и дисперсией σ^2 .

Тогда сл. величины \bar{x} и S^2 независимы, а сл.в

$\frac{\bar{x}-a}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$ распределена по стандартному нормальному закону,

$\frac{\bar{x}-a}{\bar{\sigma}/\sqrt{n}} \sim T(n-1)$ распределена по закону Стьюдента,

$\frac{\sqrt{n}\sigma^2}{\bar{\sigma}^2} \sim \chi^2(n-1)$ распределена по закону хи-квадрат.

Пусть известно, что наблюдаемая выборка извлечена из нормального распределения. Рассмотрим построение доверительных интервалов для параметров нормального распределения.

1. Построение доверительного интервала для параметра a нормального распределения (при известном σ)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma)$ – выборка из нормального распределения, где a – неизвестное математическое ожидание, а σ^2 – известная дисперсия. Пусть задана доверительная вероятность или надежность доверительного интервала γ . Требуется построить интервал $(\tilde{a}_1, \tilde{a}_2)$ так, чтобы $P\{a \in (\tilde{a}_1, \tilde{a}_2)\} \geq \gamma$.

Доверительный интервал имеет вид

$$\bar{x} - t \frac{\sigma}{\sqrt{n}} < a < \bar{x} + t \frac{\sigma}{\sqrt{n}}, \quad (17)$$

где t - значение аргумента, при котором $2\Phi(t) = \gamma$.

Решение:

По лемме Фишера имеем

$$\frac{\bar{x} - a}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1).$$

Из курса теории вероятностей известно, что для любой с.в. ξ ,

$$\mathbb{P}\{a \leq \xi < b\} = F(b) - F(a),$$

а для величины $\xi_0 \sim \mathcal{N}(0, 1)$, справедливо представление $F_{\xi_0}(x) = 0.5 + \Phi(x)$, при любом $x \in \mathbb{R}$, где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$$

— функция Лапласа.

В силу симметричности нормального распределения относительно математического ожидания будем строить интервал симметрично относительно нуля. Для заданной доверительной вероятности γ найдем такое t , чтобы $\mathbb{P}\{-t < \xi_0 < t\} = \gamma$. Для этого воспользуемся представлением с помощью функции Лапласа:

$$\begin{aligned} \mathbb{P}\{-t < \xi_0 < t\} &= F_{\xi_0}(t) - F_{\xi_0}(-t) = 0.5 + \Phi(t) - 0.5 - \Phi(-t) = \\ &= \Phi(t) + \Phi(t) = 2\Phi(t) = \gamma. \end{aligned}$$

Тогда с использованием найденного t получаем

$$\mathbb{P}\{-t < \xi_0 < t\} = \mathbb{P}\left\{-t < \frac{\bar{x} - a}{\sigma/\sqrt{n}} < t\right\} = \mathbb{P}\left\{\bar{x} - t \frac{\sigma}{\sqrt{n}} < a < \bar{x} + t \frac{\sigma}{\sqrt{n}}\right\}.$$

Интервал построен.

2. Построение доверительного интервала для параметра a нормального распределения (при неизвестном σ)

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma)$ – выборка из нормального распределения, где a – неизвестное математическое ожидание, а σ^2 – неизвестная дисперсия. Пусть задана доверительная вероятность или надежность доверительного интервала γ . Требуется построить интервал $(\tilde{a}_1, \tilde{a}_2)$ так, чтобы $P\{a \in (\tilde{a}_1, \tilde{a}_2)\} \geq \gamma$.

Решение.

По лемме Фишера имеем

$$\frac{\bar{x} - a}{\tilde{\sigma}/\sqrt{n}} \sim T(n - 1).$$

В силу симметричности распределения Стьюдента относительно математического ожидания также как и в предыдущем случае будем строить интервал симметричным относительно нуля. Воспользуемся функцией плотности распределения Стьюдента (16) и для заданной доверительной вероятности γ найдем такое t_γ , чтобы

$$P\{-t_\gamma < t_n < t_\gamma\} = \int_{-t_\gamma}^{t_\gamma} f_{t_n}(x) dx = \gamma.$$

Тогда для найденного t_γ доверительный интервал надежности γ имеет вид

$$\bar{x} - t_\gamma \frac{\tilde{\sigma}}{\sqrt{n}} < a < \bar{x} + t_\gamma \frac{\tilde{\sigma}}{\sqrt{n}}.$$

2.6 Проверка статистических гипотез.

Статистической гипотезой называется предположение о виде наблюдаемого неизвестного распределения либо о неизвестных параметрах известного распределения.

Выдвинутую гипотезу называют основной или нулевой гипотезой и обозначают H_0 . Предположения, противоречащие нулевой гипотезе, включают в конкурирующую или противоположную гипотезу и обозначают H_1 .

Различают простые гипотезы, состоящие из одного предположения, и сложные гипотезы. Примеры простых гипотез: $H_0 : "E\xi = 0"$, $H_0 : "\xi \sim \mathcal{N}(0, 1)"$. Примеры сложных гипотез: $H_1 : "E\xi > 0"$, $H_1 : "\xi \sim \mathcal{N}(a, \sigma^2)"$.

Критерием проверки гипотез называют правило, согласно которому выдвинутую гипотезу принимают или отвергают.

Критерием назовем случайную величину $K = f(X_1, X_2, \dots, X_n)$ — функцию от элементов выборки, закон распределения которой известен, и которую используют для проверки гипотез.

Прежде чем переходить к построению критериев, введем некоторые понятия.

Очевидно, что выдвинутая гипотеза H_0 может быть как ложной, так и истинной. Если гипотеза H_0 истинна, и ее приняли, либо гипотеза H_0 ложная и ее отвергли, то ошибки не произошло. В оставшихся случаях возникает одна из двух ошибок.

Ошибка первого рода при принятии гипотезы заключается в том, что нулевая гипотеза H_0 была истинна, но ее отвергли. Вероятность ошибки первого рода обозначают α , и называют **уровнем значимости** критерия.

Итак, уровень значимости критерия означает вероятность отвергнуть верную нулевую гипотезу.

Уровень значимости выбирают близким к нулю, как правило 0.05 или 0.01.

Ошибка второго рода при принятии гипотезы состоит в том, что нулевая гипотеза H_0 была ложной, но ее приняли.

Вероятность ошибки второго рода обозначается β .

Мощностью критерия называют величину $1 - \beta$, то есть вероятность отвергнуть ложную нулевую гипотезу.

Рассмотрим выборочное пространство $\mathbf{X}_n = \{\bar{\mathbf{X}}_n = (X_1, X_2, \dots, X_n)\}$ — множество выборок объема n из распределения наблюдаемой случайной величины. Разобьем множество \mathbf{X}_n на два подмножества W и \bar{W} , $W \cup \bar{W} = \mathbf{X}_n$.

Если наблюдалась выборка $\bar{X}_n = (X_1, X_2, \dots, X_n) \in W$, то гипотезу H_0 отвергаем, если наблюдалась выборка $\bar{X}_n = (X_1, X_2, \dots, X_n) \in \bar{W}$, гипотезу H_0 принимаем. Множество W назовем критическим множеством, множество \bar{W} — областью принятия гипотезы.

Основная задача при решении задачи принятия гипотезы сводится к тому, чтобы построить критическую область. Другими словами, требуется указать правило, согласно которому наблюдавшиеся выборки относят к критическому множеству.

Решается эта задача с использованием понятия ошибки первого рода.

Обозначим α вероятность ошибки первого рода, т.е. события

$$\{\bar{X}_n = (X_1, X_2, \dots, X_n) \in W | H_0\}$$

— наблюдалась выборка из критической области, в то время как была верна нулевая гипотеза H_0 .

Определение. Уровнем значимости критерия называется

$$\alpha = P\{\bar{X}_n = (X_1, X_2, \dots, X_n) \in W | H_0\}.$$

Определение. Мощностью критерия называется

$$1 - \beta = P\{\bar{X}_n = (X_1, X_2, \dots, X_n) \in W | H_1\}.$$

Критерий $K = f(X_1, X_2, \dots, X_n)$ как функция отображает выборочное пространство \mathbf{X}_n на множество действительных чисел \mathbb{R} . Тогда при условии монотонности функции $f(X_1, X_2, \dots, X_n)$ критическая область W и область принятия гипотезы \overline{W} отображаются в непересекающиеся множества на оси \mathbb{R} и имеют границу, которую обозначим k_α . Тогда критерий как правило принятия решения может иметь вид (левосторонняя критическая область): если $K_{\text{набл}} = f(x_1, x_2, \dots, x_n) < k_\alpha$ гипотезу H_0 следует отвергнуть, иначе — принять.

Критерий Неймана-Пирсона

Пусть известно наблюдаемое распределение, но неизвестны его параметры. Пусть относительно параметра распределения выдвинуты две простые гипотезы $H_0 : \theta = \theta_0$ и $H_1 : \theta = \theta_1$.

Построим критерий для проверки истинности простой гипотезы H_0 против простой гипотезы H_1 .

Как и ранее, $P(x, \theta)$ — закон распределения наблюдаемой случайной величины, $\overline{X}_n = (X_1, X_2, \dots, X_n)$ — выборка.

Функцией правдоподобия называется функция

$$L(\overline{X}_n, \theta) = \prod_{i=1}^n P(X_i, \theta).$$

Отношением правдоподобия называется функция

$$\varphi(\overline{X}_n) = \frac{L(\overline{X}_n, \theta_1)}{L(\overline{X}_n, \theta_0)} = \frac{\prod_{i=1}^n P(X_i, \theta_1)}{\prod_{i=1}^n P(X_i, \theta_0)}.$$

Критерий Неймана-Пирсона состоит в том, что за границу критической области принимается такое значение C_φ , при котором

$$P\{\varphi(\overline{X}_n) \geq C_\varphi = \alpha\}.$$

Правило принятия решения имеет вид:

если наблюдаемое значение $\varphi(\bar{x}_n) < C_\varphi$, принимают гипотезу H_0 , если $\varphi(\bar{x}_n) \geq C_\varphi$ принимают H_1 (отвергают H_0). Здесь $\bar{x}_n = (x_1, x_2, \dots, x_n)$ – наблюдавшаяся выборка.

Пример Построение оптимального критерия Неймана-Пирсона для проверки двух простых гипотез $H_0 : "a = a_0"$ против $H_1 : "a = a_1"$, $a_0 < a_1$, относительно параметра a нормального распределения. Дисперсия σ^2 известна.

Решение. Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma^2)$.

Функция правдоподобия имеет вид:

$$L(\bar{X}_n, a) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-(X_i - a)^2 / 2\sigma^2} = \left(\frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - a)^2 \right).$$

Составим отношение правдоподобия

$$\varphi(\bar{X}_n) = \frac{L(\bar{X}_n, a_1)}{L(\bar{X}_n, a_0)} = \exp\left(\frac{a_1 - a_0}{\sigma^2} n\bar{X} - \frac{n(a_1^2 - a_0^2)}{2\sigma^2} \right).$$

Согласно критерию Неймана-Пирсона граница критической области должна удовлетворять неравенству $\varphi(\bar{x}_n) \geq C_\varphi$.

Подставим найденное представление отношения правдоподобия и разрешим неравенство относительно сл. величины $n\bar{X} = \sum_{i=1}^n X_i$:

$$\sum_{i=1}^n X_i \geq \frac{\sigma^2}{a_1 - a_0} \left(\ln C_\varphi - \frac{n(a_1^2 - a_0^2)}{2\sigma^2} \right) = C.$$

Рассматривая в качестве критерия случайную величину $n\bar{X} = \sum_{i=1}^n X_i$ приходим к выводу, что она имеет распределение $\mathcal{N}(na, n\sigma^2)$. В предположении, что верна нулевая гипотеза $H_0 : "a = a_0"$ имеем

$$\frac{\sum_{i=1}^n X_i - na_0}{n\sigma^2} \sim \mathcal{N}(0, 1),$$

а, следовательно, для заданного уровня значимости α можно определить

$$P\left\{ \sum_{i=1}^n X_i \geq C \right\} = P\left\{ \frac{\sum_{i=1}^n X_i - na_0}{n\sigma^2} \geq \frac{C - na_0}{n\sigma^2} \right\} =$$

$$= 1 - F_{\xi_0} \left(\frac{C - na_0}{n\sigma^2} \right) = 1 - F_{\xi_0}(u_{1-\alpha}) = \alpha.$$

Таким образом, из соотношения

$$\frac{C - na_0}{n\sigma^2} = u_{1-\alpha}$$

определяем границу критической области в виде

$$C = na_0 + u_{1-\alpha}\sigma\sqrt{n}.$$

Проверка сложных гипотез

Рассмотрим проверку сложных гипотез относительно параметров распределения.

Сложными гипотезами называются гипотезы вида $H_0 : \theta \in \Theta_0$, $\Theta_0 \in \mathbb{R}$, $H_0 : a > a_0$, $H_0 : a \neq a_0$, и т.п.

Пусть нулевая гипотеза имеет вид $H_0 : \theta \in \Theta_0$, а конкурирующая $H_1 : \theta \in \Theta_1$, где два $\Theta_0 \cap \Theta_1 = \emptyset$ – непересекающиеся множества из \mathbb{R} .

Тогда вероятности ошибок первого и второго рода можно рассматривать как заданные на соответствующих множествах функции:

$$\alpha(\theta) = P\{\bar{X}_n = (X_1, X_2, \dots, X_n) \in W | \theta \in \Theta_0\};$$

$$\beta(\theta) = P\{\bar{X}_n = (X_1, X_2, \dots, X_n) \in \bar{W} | \theta \in \Theta_1\}.$$

Определение. Размером критерия называют величину $\alpha = \max_{\theta \in \Theta_0} \alpha(\theta)$.

Определение. Функцией мощности критерия называют вероятность отвергнуть нулевую гипотезу при любых возможных значения параметра:

$$m(\theta) = P\{\bar{X}_n = (X_1, X_2, \dots, X_n) \in W | \theta \in \Theta\}$$

Таким образом, на множестве Θ_0 $m(\theta) = \alpha(\theta)$, на множестве Θ_1 $m(\theta) = 1 - \beta(\theta)$.

Пусть α – некоторый фиксированный размер критерия. Предположим, что критерий максимизирует функцию мощности по всем возможным критериям на множестве Θ_1 . Тогда этот критерий называют **равномерно наиболее мощным**.

То есть равномерно наиболее мощный критерий имеет наименьшую вероятность ошибки второго рода при фиксированной вероятности ошибки первого рода.

Равномерно наиболее мощные критерии можно построить только в некоторых частных случаях для проверки простых гипотез вида " $\theta = \theta_0$ ".

Пример построения равномерно наиболее мощного критерия.

Пусть $X_1, X_2, \dots, X_n \sim \mathcal{N}(a, \sigma^2)$, $H_0 : "a = a_0"$, $H_1 : "a > a_0"$, σ – известно. Воспользуемся результатом построения критической области для проверки гипотезы $H_0 : "a = a_0"$ против гипотезы $H_1 : "a = a_1"$, $a_1 > a_0$ с помощью критерия Неймана-Пирсона.

Было получено, что критическая область задается неравенством

$$W = \left\{ \bar{X}_n : \sum_{i=1}^n X_i \geq C_\varphi = na_0 + u_{1-\alpha} \sigma \sqrt{n} \right\}. \quad (18)$$

Здесь $u_{1-\alpha}$ значение аргумента функции распределения $F_{\xi_0}(x)$ при котором достигается вероятность $1 - \alpha$, $F(u_{1-\alpha}) = 1 - \alpha$. Величина $u_{1-\alpha}$ квантиль распределения уровня $1 - \alpha$.

Замечая, что область $a > a_0$ состоит из величин, удовлетворяющих условию $a_1 > a_0$ гипотезы $H_1 : "a = a_1"$, а также, что C_φ в (*) не зависит от a_1 , приходим к выводу, что при заданном уровне значимости α для всех a из области $a > a_0$ критерий задается условием (18), а значит, является равномерно наиболее мощным.

Критерии согласия

Предыдущие критерии предназначались для проверки гипотез относительно параметров некоторого известного распределения, т.е. была известна

функция распределения наблюдаемой сл.в., но не были известны значения параметров распределения. Этот факт влиял на построение критерия (такие критерии называют также параметрическими).

Пусть теперь нет информации о типе наблюдаемого распределения, т.е. неизвестна функция распределения. Тогда по-прежнему можно выдвигать гипотезы относительно числовых характеристик распределения, но для проверки использовать другие методы (непараметрической статистики). Более распространенной в этом случае является задача проверки гипотезы о типе наблюдаемого распределения, для решения которой используют критерии согласия.

Пусть X_1, X_2, \dots, X_n – наблюдаемая выборка, и пусть на основании некоторых статистических и вероятностных фактов выдвинута гипотеза $H_0 : \xi \sim \mathcal{L}(x, \theta)$, или H_0 : ”наблюдаемая сл.в. имеет распределение с ф.р. $F_0(x)$ ” , или ”выдвинутая гипотеза согласуется с наблюдаемыми данными”.

Рассмотрим два основных критерия проверки таких гипотез. Критерий Колмогорова основан на анализе различий между предполагаемой, гипотетической функцией распределения $F_0(x)$ и наблюдаемой, эмпирической функцией распределения $\widetilde{F}_n(x)$. Критерий Пирсона (критерий χ^2) основан на анализе различий между предполагаемой, гипотетической функцией плотности распределения $f_0(x)$ и наблюдаемой, эмпирической функцией плотности распределения $\widetilde{f}_n(x)$, которые задаются соответственно теоретическими и наблюдаемыми частотами.

Критерий Колмогорова.

Проверяется простая гипотеза $H_0 : F_\xi(x) = F_0(x)$ против сложной гипотезы $H_0 : F_\xi(x) \neq F_0(x)$, α – заданный уровень значимости.

Для проверки гипотезы используется сл.в.(статистика) $D(\overline{X}_n)$, наблюдаемые значения которой вычисляются

$$D_{\text{набл}} = D(\overline{x}_n) = \sup_{x \in \mathbb{R}} |\widetilde{F}_n(x) - F_0(x)|.$$

Если $D_{\text{набл}} \geq D_{\text{крит}}$, гипотеза H_0 отклоняется, иначе, если $D_{\text{набл}} < D_{\text{крит}}$ гипотеза H_0 принимается, т.е. наблюдаемые значения согласуются с выдвинутой гипотезой. Здесь $D_{\text{крит}} = D_{1-\alpha}$ – квантиль уровня $1 - \alpha$ распределения величины $D(\bar{X}_n)$.

Определение. Квантилью уровня α , $0 < \alpha < 1$ распределения $F_\xi(x)$ называется число u_α , такое что $F_\xi(u_\alpha) = P\{\xi < u_\alpha\} = \alpha$.

Критерий Колмогорова А.Н. основан на предельном соотношении

$$\lim_{n \rightarrow \infty} P\{\sqrt{n}D(\bar{X}_n) < x\} = F_K(x), \quad x > 0. \quad (19)$$

Здесь $F_K(x)$ – функция распределения сл.в K с распределением Колмогорова

$$F_K(x) = \begin{cases} \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 x^2}, & x \geq 0; \\ 0, & x < 0. \end{cases} \quad (20)$$

Критерий Пирсона.

Проверяется простая гипотеза $H_0 : "F_\xi(x) = F_0(x)"$ против сложной гипотезы $H_0 : "F_\xi(x) \neq F_0(x)"$, α – заданный уровень значимости.

Для проверки гипотезы используется сл.в.(статистика) χ^2 , наблюдаемые значения которой вычисляются следующим образом:

$$\chi_{\text{набл}}^2 = \sum_{i=1}^m \frac{(n_i - n'_i)^2}{n'_i}, \quad (21)$$

где m – количество групп в вариационном ряду, n_i – наблюдаемые частоты, n'_i – теоретические частоты, вычисленные в соответствии с выдвинутой гипотезой. Критическое значение $\chi_{\text{крит}}^2 = \chi^2(1 - \alpha, m - r - 1)$ – квантиль уровня $1 - \alpha$ распределения сл. в. χ^2 с $m - r - 1$ степенями свободы. Здесь r – число параметров гипотетического распределения, оцениваемых по выборке. (Замечание. При использовании таблиц распределения часто пишут для простоты $\chi_{\text{крит}}^2 = \chi^2(\alpha, m - r - 1)$)

Если $\chi_{\text{набл}}^2 < \chi_{\text{крит}}^2$, то гипотезу о согласии наблюдаемого распределения с гипотетическим принимают, в противном случае гипотезу H_0 отвергают.

Список литературы

- [1] Ширяев А.Н. Вероятность. — М: МЦНМО; 2017. — 552 с.
- [2] Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. — Киев: Издательское объединение "Вища школа"; 1979. — 408 с.
- [3] Математическая статистика: Учеб. для вузов / В.Б. Горяинов, И.В. Павлов, Г.М. Цветкова и др.; Под ред. В.С. Зарубина, А.П. Крищенко. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. - 424 с.
- [4] Боровков А.А. Математическая статистика. — Новосибирск: Наука; Издательство института математики, 1997. — 772 с.
- [5] Лагутин М.Б. Наглядная математическая статистика: Учебное пособие / М.Б. Лагутин.—М.Ж БИНОМ, Лаборатория знаний, 2007.— 472 с.: ил.