
51 M34

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методические указания по курсу «Теория вероятностей и математическая статистика» для студентов 3 курса ФПМИ (направление 010400.62 – "Прикладная математика и информатика" дневного отделения)

Методические указания предназначены для проведения практических занятий по курсу «Теория вероятностей и математическая статистика» (направление 010400.62 — «Прикладная математика и информатика»). Указания содержат теоретические сведения, необходимые для решения задач по оцениванию параметров и проверке статистических гипотез. В работе приведены задачи, решаемые на практических занятиях, и разобраны примеры их решения.

Составители: доктор техн. наук, доцент С.Н. Постовалов,

доктор техн. наук, доцент Е.В. Чимитова,

канд. техн. наук В.С. Карманов

Рецензент:

Работа подготовлена на кафедре теоретической и прикладной информатики

© Новосибирский государственный технический университет, 2020

Тема 1. Выборки и статистики

Пусть $X_n = \{X_1,...,X_n\}$ — выборка объема n, получаемая в результате n наблюдений случайной величины ξ . Будем считать, что наблюдения (случайные величины) $X_1,...,X_n$ независимы и имеют одну и ту же функцию распределения $F_{\varepsilon}(x)$.

Группирование — это такое преобразование выборки, когда область значений случайной величины ξ разбивается на k непересекающихся интервалов граничными точками $t_0 < t_1 < ... < t_{k-1} < t_k$, где t_0 — нижняя граница области определения, t_k — верхняя грань. В соответствии с заданным разбиением подсчитывается число n_i выборочных значений, попавших в i-й интервал. Для дискретной случайной величины вместо интервалов могут быть взяты ее значения.

Реализацией выборки 2 называется множество реализаций случайной величины ξ .

Определение. *Статистикой* называется измеримая функция от выборки.

Важными примерами статистик являются выборочные моменты. Для *выборочного среднего* используется обозначение

 $^{^1}$ В пособии будет предполагаться, что функция распределения непрерывна справа, т.е. $F_{\varepsilon}(x) = P\big\{\xi \leq x\big\}$

² В дальнейшем не будет делаться различий между выборкой и ее реализацией: это будет определяться контекстом. Например, если дана выборка в виде набора чисел, то это реализация выборки.

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_{i} ,$$

для выборочной дисперсии S^2 и несмещенной выборочной дисперсии S_0^2 используются обозначения

$$S^{2} = \frac{1}{n} \sum_{i=1}^{n} (X_{i} - \overline{X})^{2} \text{ M } S_{0}^{2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i} - \overline{X})^{2}.$$

Задачи

- 1.1. Пусть $X_1, X_2, ..., X_n$ выборка из нормального распределения с параметрами a и σ^2 .
- а) Вычислить математическое ожидание и дисперсию статистики \overline{X}
- б) Вычислить математическое ожидание статистик S^2 и S_0^2 .
 - в) Вычислить дисперсию статистики S^2 .
- 1.2. Пусть $X_1, X_2, ..., X_n$ выборка из распределения Пуассона с параметром λ . Вычислить математическое ожидание и дисперсию статистики \overline{X} . Имеет ли статистика \overline{X} распределение Пуассона? Нормальное распределение?
- 1.3. Пусть $X_1, X_2, ..., X_n$ выборка из равномерного распределения на отрезке [a,b], a < b. Вычислить математическое ожидание и дисперсию статистики \overline{X} . Имеет ли статистика \overline{X} равномерное распределение? Нормальное распределение?
- 1.4. Пусть $X_1, X_2, ..., X_n$ выборка из показательного распределения с параметром 3. Найти распределение выборки $Y_1, ..., Y_n$, где $Y_i = 1 e^{-3X_i}$.
- 1.5. Пусть $X_1, X_2, ..., X_n$ выборка из некоторого распределения, у которого функция распределения F(y) непрерывна и строго возрастает. Какое распределение имеет выборка $Y_1, ..., Y_n$, где $Y_i = F(X_i)$?

- 1.6. В результате пяти измерений длины стержня одним прибором получены следующие значения (в мм): 92; 94; 103; 105; 106. Найти выборочное среднее и выборочную дисперсию длины стержня.
- $1.7.\$ Найти выборочное среднее и выборочную дисперсию по группированной выборке из дискретного распределения объема n=10 :

$$t_i$$
 1250 1270 1280

$$n_i$$
 2 5 3

1.8. Найти выборочное среднее и выборочную дисперсию по группированной выборке из дискретного распределения объема n=20:

$$t_i \quad 2560 \quad 2600 \quad 2620 \quad 2650 \quad 2700$$

$$n_i$$
 2 3 10 4 1

Тема 2. Порядковые статистики

Определение. Если все элементы выборки $X_1, X_2, ..., X_n$ расположены в порядке неубывания их величины: $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq ... \leq X_{(n)}$, то каждый из $X_{(i)}$ называется i-ой порядковой статистикой, а соответствующая неубывающая последовательность — вариационным рядом.

Порядковая статистика $X_{(i)}$ имеет распределение, отличное от распределения X_i , и ее функция распределения равна

$$\sum_{k=i}^{n} C_{n}^{k} F_{\xi}(x)^{k} \left(1 - F_{\xi}(x)\right)^{n-k} . \tag{2.1}$$

Определение. Выборочной *медианой* называется статистика

$$X_{med} = egin{cases} X_{(m)}, \ ext{если} \ n = 2m-1; \ X_{(m)} + X_{(m+1)} \ ext{, если} \ n = 2m. \end{cases}$$

Определение. Выборочной *квантилью уровня* δ называется статистика $X_{([n\delta]+1)}$.

Задачи:

- 2.1. Для выборки с функцией распределения $F_{\xi}(x)$ найти функцию распределения
 - а) максимального члена вариационного ряда $X_{\scriptscriptstyle(n)}$
- б) минимального члена вариационного ряда $X_{(1)}$ не используя формулу (2.1). Совпадают ли полученные выражения с (2.1) при i=n и i=1?

- 2.2. Для выборки с функцией распределения $F_{\xi}(x)$ и плотностью $f_{\varepsilon}(x)$ найти функцию плотности распределения
 - а) максимального члена вариационного ряда $X_{(n)}$;
 - б) минимального члена вариационного ряда $X_{(1)}$;
 - в) k -й порядковой статистики $X_{(k)}$.

<u>Указание</u>. Для упрощения полученного выражения после дифференцирования воспользуйтесь формулами:

$$kC_n^k = nC_{n-1}^{k-1}$$
 и $(n-k)C_n^k = nC_{n-1}^k$

- 2.3. Для выборки из равномерного распределения на $[0,\theta]$ найти математическое ожидание и дисперсию
 - а) максимального члена вариационного ряда $X_{(n)}$;
 - б) минимального члена вариационного ряда $X_{(1)}$;
 - в) k -й порядковой статистики $X_{(k)}$.

Тема 3. Непараметрическое оценивание функции распределения

3.1. Эмпирическая функция распределения

Введем вспомогательную случайную функцию: $\mu_n(x)$ – количество наблюдений в выборке случайной величины ξ , не превосходящих x. Событие $\mu_n(x) = k$ означает, что в интервал $(-\infty, x]$ попало ровно k наблюдений, а в интервал $(x, +\infty)$ попало (n-k) наблюдений.

Определение. Функция $F_n(x) = \frac{\mu_n(x)}{n}$ называется эмпирической функцией распределения.

По определению эмпирическая функция распределения является случайной функцией; $\forall x \in R$, $F_n(x)$ — дискретная случайная величина, принимающая значения

$$0 = \frac{0}{n}, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, \frac{n}{n} = 1$$

при этом

$$P\left\{F_{n}(x) = \frac{k}{n}\right\} = P\left\{\mu_{n}(x) = k\right\} = C_{n}^{k} F_{\xi}^{k}(x) \left(1 - F_{\xi}(x)\right)^{n-k}$$
 (3.1)

Эмпирическую функцию распределения можно представить в виде

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} h(x - X_i), \qquad (3.2)$$

где $h(x) = \begin{cases} 0, x < 0 \\ 1, x \ge 0 \end{cases}$ — функция Хевисайда (единичного скачка).

Функция Хевисайда обладает следующими свойствами:

- $h^2(x) = h(x)$;
- h(ax) = h(x), если a > 0.

Задачи:

3.1. Пусть (3; 0; 4; 3; 6; 0; 3; 1; 2; 1) — наблюдавшиеся значения выборки. Построить график эмпирической функцию распределения и проверить утверждения, что $F_n(1) = \frac{2}{5}$,

$$F_n(3) = \frac{3}{8} \text{ M } F_n(5) = \frac{9}{10}.$$

3.2. Для выборки с функцией распределения $F_{\xi}(x)$ найти математическое ожидание и дисперсию статистики $F_{n}(y)$.

 $\underline{\text{У казание}}$: задачу можно решить двумя способами. Первый способом основан на (3.1), а второй — на (3.2).

3.3. Найти две разные выборки, соответствующие эмпирической функции распределения, график которой представлен на рис. 3.1.

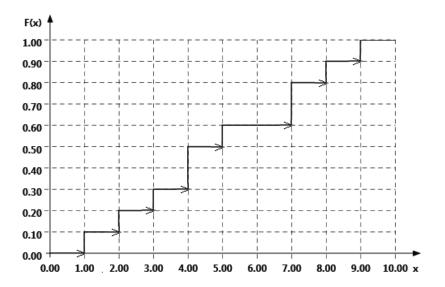


Рис. 3.1. Эмпирическая функция распределения

- 3.4. $F_n(x)$ – эмпирическая Пусть функция распределения, построенная по выборке $X_1, X_2, ..., X_n$, а $G_n(x)$ эмпирическая функция распределения, построенная $Y_1, Y_2, ..., Y_n$ выборке того объема. Является же ЛИ эмпирической функцией распределения функция $(F_{n}(x)+G_{n}(x))/2$? Если «да», то какой выборке она соответствует?
- 3.5. Пусть a>0 и b два фиксированных действительных числа. Пусть $F_n(x)$ эмпирическая функция распределения, построенная по выборке $X_1, X_2, ..., X_n$, а $G_n(x)$ эмпирическая функция распределения, построенная по выборке $Y_1, Y_2, ..., Y_n$, где $Y_i = aX_i + b$. Доказать, что при всех x имеет место равенство $G_n(x) = F_n\left(\frac{x-b}{a}\right)$.

3.2. Цензурированная выборка. Оценка Каплана-Мейера

 $\begin{align*}{ll} \begin{align*}{ll} \begin{ali$

цензурирования, который равен единице, если i-е наблюдение полное, нулю — если цензурированное.

Определение. Выборка называется *случайно цензурированной справа*, если T_i и C_i представляют собой независимые случайные величины, причем T_i принадлежит закону распределения вероятностей с функцией распределения F(x), а C_i — закону с функцией распределения $F_C(x)$.

Обозначим через $a_1 < a_2 < ... < a_k = \tau$, $k \le n$ моменты времени, в которые были зафиксированы системные события $(X_i, \delta_i = 1)$. Тогда *оценку Каплана-Мейера* можно вычислить по формуле:

$$\hat{F}_n(x) = \begin{cases} 0, & x < a_1 \\ 1 - \prod_{a_i \le x} \left(1 - \frac{d_i}{r_i} \right), & x \ge a_1 \end{cases},$$

где $d_i = \sum_{X_j = a_i} \mathcal{S}_j$ - количество полных наблюдений в точке a_i , а

 r_{i} — количество наблюдений в выборке, таких что $X_{j} \geq a_{i}$, j = 1, ..., n .

Оценка Каплана-Мейера является аналогом эмпирической функции распределения для случайно цензурированной справа выборки.

<u>Пример</u>. Построить оценку Каплана-Мейера для выборки (3; 0; 4+; 3; 6; 0; 3; 1; 2; 1) и ее график. Наблюдение (4+) является цензурированным.

<u>Решение</u>. Для построения оценки Каплана-Мейера можно проделать вычисления по следующей таблице.

a_i	d_{i}	r_i	$1 - \frac{d_i}{r_i}$	$\prod_{j=1}^{i} \left(1 - \frac{d_j}{r_j} \right)$	$\hat{F}_n(a_i)$
0	2	10	$\frac{8}{10}$	$\frac{8}{10}$	$\frac{2}{10}$
1	2	8	$\frac{6}{8}$	$\frac{6}{10}$	$\frac{4}{10}$
2	1	6	$\frac{5}{6}$	$\frac{5}{10}$	$\frac{5}{10}$
3	3	5	$\frac{2}{5}$	$\frac{2}{10}$	$\frac{8}{10}$
6	1	1	0	0	1

График строится следующим образом. На интервале $x < a_1$ функция равна 0, на интервале $x \in [a_i, a_{i+1})$ функция постоянна и равна $\hat{F}_n(a_i)$, на интервале $x \ge a_k$ функция постоянна и равна $\hat{F}_n(a_k)$. Цензурированные наблюдения на графике отмечаются крестиком.

Задачи:

- 3.6. Построить оценку Каплана-Мейера для выборки (3; 0; 4; 3; 6+; 0; 3; 1; 2; 1) и ее график. Наблюдение (6+) является цензурированным.
- 3.7. Построить оценку Каплана-Мейера для выборки (3; 0; 4; 3; 6; 0; 3; 1; 2; 1+) и ее график. Наблюдение (3; 0; 4; 3; 6; 0; 3; 1; 2; 1+) и ее график. Наблюдение (3; 0; 4; 3; 6; 0; 3; 1; 2; 1+) и ее график.
- 3.8. Построить оценку Каплана-Мейера для выборки (3; 0; 4+; 3; 6; 0; 3; 1; 2; 1+) и ее график. Наблюдения «1+» и «4+» являются цензурированными.

Тема 4. Непараметрическое оценивание функции плотности

4.1. Гистограмма

Разобьем область значений случайной величины ξ на k интервалов $t_0 < t_1 < t_2 < ... < t_k$, и подсчитаем количество наблюдений, попавших в каждый интервал: n_i — количество наблюдений в интервале $\left[t_i, t_{i-1}\right)$, при этом $\sum_{i=1}^k n_i = n$.

На графике для каждого интервала строим столбцы гистограммы шириной $w_i = t_i - t_{i-1}$ и высотой $h_i = \frac{n_i}{nw_i}$.

Если область значений случайной величины бесконечная, то для построения гистограммы можно взять такие границы t_0, t_k , чтобы в интервал $\left[t_0, t_k\right]$ попадали все наблюдения.

Гистограмма довольно грубый способ оценивания плотности распределения, связанный с неопределенностью выбора числа интервалов k, границ интервалов, а также потерей информации при группировании.

Задачи:

- 4.1. Построить гистограмму по выборке (3; 0; 4; 3; 6; 0; 3; 1; 2; 1).
- 4.2. Построить гистограмму по группированной выборке:

Номер i интервала i	Границы интервала	n_i	
1	2-7	5	
2	7-12	10	

3	12-17	25
4	17-22	6
5	22-27	4

4.3. Построить гистограмму по группированной выборке:

Номер интервала	Границы интервала	n_{i}
1	3-5	4
2	5-7	6
3	7-9	20
4	9-11	40
5	11-13	20
6	13-15	4
7	15-17	6

4.2. Ядерная оценка плотности

Пусть дана ядерная функция g(t), удовлетворяющая условиям:

$$g(t) \ge 0; \qquad g(t) = g(-t); \qquad \int_{-\infty}^{+\infty} g(t)dt = 1; \qquad \int_{-\infty}^{+\infty} t^2 g(t)dt = 1;$$

$$\int_{-\infty}^{+\infty} t^m g(t)dt < \infty, 3 \le m < \infty.$$

Тогда ядерная оценка функции плотности имеет вид

$$\hat{f}_n(x) = \frac{1}{n\lambda_n} \sum_{i=1}^n g\left(\frac{x - X_i}{\lambda_n}\right),\,$$

где λ_n — параметр размытости ядерной функции, который для сходимости оценки к функции плотности должен удовлетворять условиям $\lim_{n\to\infty}\lambda_n=0$, $\lim_{n\to\infty}n\lambda_n=\infty$.

Задачи:

- 4.4. Чему равен интеграл $\int_{-\infty}^{+\infty} t \cdot g(t) dt$?
- 4.5. Проверить выполнение условий регулярности следующих ядер
- а) треугольное:

$$g(x) = \left(\frac{1}{\sqrt{6}} - \frac{|x|}{6}\right) h\left(\sqrt{6} - |x|\right), \quad h(x) = \begin{cases} 1, x \ge 0 \\ 0, x < 0 \end{cases};$$

б) плотность стандартного нормального распределения:

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

в) ядро Епанечникова:

$$g(x) = \left(\frac{3}{4\sqrt{5}} - \frac{3}{20\sqrt{5}}x^2\right)h(\sqrt{5} - |x|).$$

4.6*. Построить на компьютере график ядерной оценки функции плотности по выборке (3; 0; 4; 3; 6; 0; 3; 1; 2; 1), используя различные ядра и параметры размытости.

Тема 5. Свойства оценок. Состоятельность и несмещенность

Пусть имеется выборка $X_n = \{X_1, ..., X_n\}$ распределения случайной величины ξ с функцией распределения $F(x;\theta)$, где $\theta \in \Theta$. В общем случае задача заключается чтобы. используя оценивания В TOM. информацию, доставляемую выборкой, статистическую сделать статистические выводы об истинном значении неизвестного параметра θ .

Определение. Точечной оценкой неизвестного параметра θ по выборке \mathbf{X}_n называется значение некоторой статистики $T_n = T(\mathbf{X}_n)$, которое приближенно равно значению параметра $\theta: \hat{\theta} = T(\mathbf{X}_n)$.

Определение. Статистика $T(\mathbf{X}_n)$ называется *несмещенной* оценкой параметра θ , если выполняется условие: $M[T(\mathbf{X}_n)] = \theta, \forall \theta \in \Theta$.

Определение. Несмещенной оценкой с равномерно минимальной дисперсией (НОРМД) называется такая оценка $T^*(\mathbf{X}_n)$, что $D[T^*(\mathbf{X}_n)] \leq D[T(\mathbf{X}_n)]$: $\forall \theta \in \Theta$, $\forall T(\mathbf{X}_n) : M[T(\mathbf{X}_n)] = \theta$.

Определение. Оценка $T(\mathbf{X}_n)$ некоторой функции $\tau(\theta)$ называется cocmosmeльной, если $T(\mathbf{X}_n) \stackrel{P}{\longrightarrow} \tau(\theta)$, $\forall \theta \in \Theta$, при $n \to \infty$. То есть $\forall \varepsilon > 0$: $P\{|T(\mathbf{X}_n) - \tau(\theta)| > \varepsilon\} \to 0$, при $n \to \infty$.

Свойство состоятельности обязательно для любого правила оценивания, однако оно является асимптотическим и не связано со свойствами оценки при фиксированном объеме выборки (в отличие от свойств несмещенности и минимальной дисперсии).

Критерий состоятельности. Пусть

$$M_{\theta}T_{n} = \tau(\theta) + \varepsilon_{n}, D_{\theta}T_{n} = \delta_{n}$$
 и
$$\varepsilon_{n} = \varepsilon_{n}(\theta) \rightarrow 0, \delta_{n} = \delta_{n}(\theta) \rightarrow 0$$
 при $n \rightarrow \infty$.

Тогда $T(\mathbf{X}_{\scriptscriptstyle n})$ — состоятельная оценка функции $\tau = \tau(\theta)$.

Пример 5.1

Пусть $X_1,...,X_n$ – выборка из распределения Максвелла

с функцией плотности
$$f(x;\theta) = \frac{2x^2}{\theta^3 \sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\theta^2}\right\},$$
 $x>0, \theta>0.$

Требуется проверить оценку $\hat{\theta} = \frac{\sqrt{2\pi}}{4} \, \overline{X}$ на несмещенность и состоятельность.

Решение:

1. Несмещенность.

$$M\left(\frac{\sqrt{2\pi}}{4}\bar{X}\right) = \frac{\sqrt{2\pi}}{4}M\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}\right) = \frac{\sqrt{2\pi}}{4n}\sum_{i=1}^{n}MX_{i} = \frac{\sqrt{2\pi}}{4n}\sum_{i=1}^{n}\frac{4\theta}{\sqrt{2\pi}} = \theta,$$

 \Rightarrow оценка $\hat{\theta} = \frac{\sqrt{2\pi}}{4} \, \bar{X}$ является несмещенной оценкой параметра θ .

2. Состоятельность.

Поскольку $\hat{\theta}$ является несмещенной, то нам достаточно исследовать дисперсию оценки $D(\hat{\theta})$.

$$D(\hat{\theta}) = \frac{2\pi}{16n^2} \sum_{i=1}^{n} DX_i = \frac{\pi}{8n} \cdot \frac{3\pi - 8}{\pi} \theta^2 \to 0, \ n \to \infty,$$

 \Rightarrow по критерию состоятельности, оценка $\hat{\theta} = \frac{\sqrt{2\pi}}{4} \, \bar{X}$ является состоятельной.

Задачи:

- 5.1 Пусть $X_1, X_2, ..., X_n$ выборка из равномерного распределения на отрезке $[0, \theta]$. Проверить состоятельность и несмещенность следующих оценок параметра θ .
 - a) $X_{(n)}$
 - б) $2\bar{X}$;
 - B) $\bar{X} + X_{(n)} / 2$;
 - Γ) $(n+1)X_{(1)}$;
 - $\chi(1) = X_{(1)} + X_{(n)};$
 - e) $\frac{n+1}{n}X_{(n)}$.

<u>Указание</u>. При доказательстве состоятельности оценки можно доказать, что дисперсия оценки ограничена сверху некоторой сходящейся к нулю функцией. Для этого можно воспользоваться неравенством Коши-Буняковского $\text{cov}^2(X,Y) \leq DX \cdot DY$.

- 5.2 В партии из n изделий оказалось m бракованных. Неизвестная вероятность p появления бракованного изделия оценивается величиной m/n. Проверить состоятельность и несмещенность этой оценки.
- 5.3 Пусть $X_1, X_2, ..., X_n$ выборка из некоторого распределения с конечным математическим ожиданием. Доказать, \overline{X} что является несмещенной и состоятельной оценкой параметра $m_1 = MX_1$.
- 5.4~ Пусть $X_1, X_2, ..., X_n$ выборка из некоторого распределения с конечной дисперсией. Доказать, что статистика

$$S^{2} = \frac{1}{n} \sum_{i=1}^{n} (X_{i} - \overline{X})^{2}$$

является состоятельной оценкой параметра $\sigma^2 = DX_1$. Является ли S^2 несмещенной оценкой дисперсии σ^2 ? Построить оценку, являющуюся одновременно состоятельной и несмещенной оценкой параметра.

- 5.5 Пусть $X_1, X_2, ..., X_n$ выборка из равномерного распределения на отрезке [a,b]. Является ли оценка $\theta_n^* = X_{(n)} X_{(1)}$ несмещенной оценкой длины отрезка b-a? Состоятельной?
- 5.6 Пусть $X_1, X_2, ..., X_n$ выборка из показательного распределения с параметром α .

$$f(x) = \begin{cases} \alpha \exp^{-\alpha x}, x \ge 0, \\ 0, x < 0. \end{cases}$$

Будет ли оценка $\alpha_n^* = 1/\bar{X}$ несмещенной? Если «нет», найти смещение. Является ли оценка состоятельной?

Тема 6. Метод моментов

Существует множество различных методов построения оценок неизвестных параметров закона распределения случайной величины по выборке \mathbf{X}_n . Рассмотрим наиболее простые методы.

Приравнивая теоретические $M \, \xi^k$ и выборочные моменты, можно найти точечные оценки неизвестных параметров.

$$M\xi^{k}(\theta)\Big|_{\theta=\hat{\theta}} = \frac{1}{n}\sum_{i=1}^{n}X_{i}^{k}, k=1,2,...$$

Такой метод называется методом моментов. Если в векторе θ содержится r неизвестных параметров, то необходимо взять столько таких уравнений, чтобы можно было выразить неизвестные параметры.

Пример 6.1

Пусть $X_1,...,X_n$ — выборка из гамма-распределения с функцией плотности:

$$f(x;\alpha,\beta) = \frac{x^{\beta-1}e^{-x/\alpha}}{\Gamma(\beta)\alpha^{\beta}}, x > 0, \alpha,\beta > 0.$$

Требуется найти оценку по методу моментов векторного параметра (α, β) .

Решение:

Найдем первый и второй теоретические моменты:

$$MX_{1} = \int_{0}^{\infty} x \frac{x^{\beta-1} e^{-\frac{\gamma}{\alpha}}}{\Gamma(\beta)\alpha^{\beta}} dx = \frac{\alpha}{\Gamma(\beta)} \int_{0}^{\infty} y^{\beta} e^{-y} dy = \frac{\alpha}{\Gamma(\beta)} \Gamma(\beta+1) = \alpha \cdot \beta$$

$$MX_{1}^{2} = \int_{0}^{\infty} x^{2} \frac{x^{\beta-1} e^{-\frac{\gamma}{\alpha}}}{\Gamma(\beta)\alpha^{\beta}} dx = \frac{\alpha^{2}}{\Gamma(\beta)} \int_{0}^{\infty} y^{\beta+1} e^{-y} dy =$$

$$= \frac{\alpha^{2}}{\Gamma(\beta)} \Gamma(\beta+2) = \alpha^{2} \cdot \beta \cdot (\beta+1).$$

Приравнивая теоретические и выборочные моменты, получим:

$$\begin{cases} \hat{\alpha}\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n}X_{i} = A_{1} \\ \hat{\alpha}^{2}\hat{\beta}(\hat{\beta}+1) = \frac{1}{n}\sum_{i=1}^{n}X_{i}^{2} = A_{2} \end{cases} = > \qquad \hat{\alpha} = \frac{A_{2} - A_{1}^{2}}{A_{1}}$$

$$\hat{\beta} = \frac{A_{1}^{2}}{A_{2} - A_{1}^{2}}$$

Задачи:

- 6.1. Пусть дана выборка из гамма-распределения с параметрами $\alpha > 0$ и $\beta > 0$. Построить оценки по методу моментов
 - а) параметра α , если значение β известно;
 - б) параметра β , если значение α известно.
- 6.2. Пусть имеется выборка из нормального распределения с параметрами μ и $\sigma > 0$. Построить оценки параметров по методу моментов
 - а) параметра μ , если значение σ известно;
 - б) параметра σ , если значение μ известно.
 - в) векторного параметра (μ, σ) ,
- 6.3. Пусть имеется выборка из распределения Парето с параметрами β и θ . Построить оценки по методу моментов
 - а) параметра $\beta > 1$, если значение $\theta > 0$ известно;
 - б) параметра $\theta > 0$, если значение $\beta > 1$ известно;
 - в) векторного параметра (β, θ) , где $\beta > 2$ и $\theta > 0$.
- 6.4. Используя метод моментов, оценить параметр $p \in (0,1)$ геометрического распределения.
- 6.5. Случайная величина X (число нестандартных изделий в партии изделий) распределена по закону Пуассона. Ниже приведено число нестандартных изделий в 200 партиях в группированном виде

$$t_i$$
 0 1 2 3 4 n_i 132 43 20 3 2

Найти методом моментов точечную оценку неизвестного параметра λ распределения Пуассона.

Найти методом моментов по выборке X_1, X_2, \dots, X_n точечные оценки неизвестных параметров нормального распределения.

6.6. Случайная величина X (ошибка измерения радиодальномером) подчинена равномерному закону распределения с неизвестными параметрами a и b. Ниже приведено n=200 измерений дальности в группированном виде

$$t_i$$
 2-4 4-6 6-8 8-10 10-12 12-14 14-16 16-18 18-20 20-22 n_i 21 16 15 26 22 14 21 22 18 25

Найти методом моментов точечные оценки неизвестных параметров равномерного распределения.

Тема 7. Метод максимального правдоподобия

Определение. Оценкой максимального правдоподобия (ОМП) параметра θ называется точка параметрического множества Θ , в которой функция максимального правдоподобия $L(\mathbf{X}_n,\theta) = \prod_{i=1}^n f(X_i,\theta)$ достигает наибольшего значения: $L(\mathbf{X}_n,\hat{\theta}) = \sup_{\theta \in \Theta} L(\mathbf{X}_n,\theta)$.

Если для любой выборки \mathbf{X}_n из выборочного пространства максимум $L(\mathbf{X}_n,\theta)$ достигается во внутренней точке θ , и $L(\mathbf{X}_n,\theta)$ дифференцируема по θ , то ОМП θ удовлетворяет уравнению $\frac{\partial \ln L(\mathbf{X}_n,\theta)}{\partial \theta_i} = 0, i = 1,...,r$, которое называется уравнением правдоподобия.

Пример 7.1

Построить оценку максимального правдоподобия параметра p распределения Бернулли: $P\{\xi=k\}=p^k(1-p)^{1-k}$, k=0,1.

Решение:

Логарифмическая функция правдоподобия равна

$$\ln L(\mathbf{X}_{n}; p) = \sum_{i=1}^{n} \left(X_{i} \ln p + (1 - X_{i}) \ln(1 - p) \right) =$$

$$= \sum_{i=1}^{n} X_{i} \ln p + \left(n - \sum_{i=1}^{n} X_{i} \right) \ln(1 - p) ;$$

$$\frac{\partial \ln L}{\partial p} = \frac{1}{p} \sum_{i=1}^{n} X_{i} - \frac{1}{1 - p} \left(n - \sum_{i=1}^{n} X_{i} \right) = 0 \implies \hat{p} = \bar{X} ,$$

где \overline{X} – среднее выборочное значение.

$$\begin{split} \frac{\partial^2 \ln L}{\partial p^2} &= -\frac{1}{p^2} \sum_{i=1}^n X_i - \frac{1}{(1-p)^2} \left(n - \sum_{i=1}^n X_i \right) = \\ &= \left(\frac{1}{(1-p)^2} - \frac{1}{p^2} \right) \sum_{i=1}^n X_i - \frac{n}{(1-p)^2} = \frac{2p-1}{p^2 (1-p)^2} \sum_{i=1}^n X_i - \frac{n}{(1-p)^2} \end{split}$$

Проверим знак второй производной при $p = \overline{X}$:

$$\frac{\partial^{2} \ln L}{\partial p^{2}} \bigg|_{p=\bar{X}} = -\frac{n\bar{X}}{p^{2}} - \frac{1}{(1-p)^{2}} \Big(n - n\bar{X} \Big) = -\frac{n\bar{X}}{\bar{X}^{2}} - \frac{n}{(1-\bar{X})^{2}} \Big(1 - \bar{X} \Big) = \\
= -\frac{n}{\bar{X}} - \frac{n}{(1-\bar{X})} = -\frac{n(1-\bar{X}+\bar{X})}{\bar{X}(1-\bar{X})} = -\frac{n}{\bar{X}(1-\bar{X})} < 0.$$

Таким образом, при $p = \overline{X}$ функция правдоподобия достигает максимума.

Пример 7.2

Построить оценку максимального правдоподобия параметра $\theta > 0$ равномерного распределения на отрезке $[0,\theta].$

Решение:

Функция правдоподобия выборки равна

$$L(\mathbf{X}_{_{n}};\,\theta) = \begin{cases} \theta^{^{-n}}, \text{ если все } X_{_{j}} \in [0,\,\theta] \\ 0, \text{ если хотя бы одно } X_{_{j}} \not\in [0,\,\theta] \end{cases} = \\ = \begin{cases} \theta^{^{-n}}, \text{ если } X_{_{(n)}} \leq \theta \\ 0, \text{ если } X_{_{(n)}} > \theta \end{cases},$$

где $X_{(n)}$ – максимальная порядковая статистика.

При фиксированных значениях выборки (и, следовательно, при фиксированном значении $X_{(n)}$) зависимость $L(\mathbf{X}_n;\theta)$ от θ показана на рисунке 7.1. Максимум функции правдоподобия достигается в точке

 $\theta = X_{\scriptscriptstyle (n)}.$ Поэтому искомая оценка максимального правдоподобия есть $\hat{\theta} = X_{\scriptscriptstyle (n)}.$

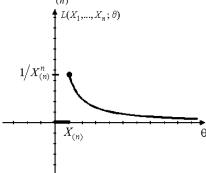


Рис. 7.1. Функция правдоподобия

Задачи:

7.1. Найти оценку максимального правдоподобия параметра масштаба $\theta > 0$ экспоненциального закона распределения.

$$f(x,\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x \ge 0; \\ 0, & x < 0. \end{cases}$$

7.2. Найти оценку максимального правдоподобия параметра сдвига θ экспоненциального закона распределения.

$$f(x,\theta) = \begin{cases} e^{-(x-\theta)}, x \ge \theta; \\ 0, x < \theta. \end{cases}$$

7.3. Найти оценку максимального правдоподобия параметра $\theta > 0$ равномерного распределения на отрезке $[-\theta,0]$.

- 7.4. Найти оценку максимального правдоподобия двумерного параметра (a,b) равномерного распределения на отрезке [a,b].
- 7.5. Найти оценку θ (или (θ_1, θ_2)) методом максимального правдоподобия:
 - а) нормального распределения $N(\theta_1, \theta_2)$;
 - б) гамма-распределения $G(\theta, \alpha)$;
 - в) гамма-распределения $G(\lambda, \theta)$;
 - Γ) гамма-распределения $G(\theta_1, \theta_2)$
 - д) распределения Вейбулла $W(\theta, \alpha)$
- 7.6. Найти оценку максимального правдоподобия параметра сдвига $\mu \in \mathbf{R}$ распределения Лапласа с плотностью

$$f_{\mu}(y) = e^{-|y-\mu|}/2$$
.

Тема 8. Информационное количество Фишера, неравенство Рао-Крамера и критерий эффективности

Определение. Семейство $\{F(x;\theta), \theta \in \Omega\}$ является *регулярным*, если выполняются следующие условия:

- 1) для любого θ , $\theta \in \Omega$, плотность $f(x;\theta)$ дифференцируема по θ , то есть существует $\frac{\partial}{\partial \theta} f(x,\theta)$;
 - 2) множество $\{x: f(x,\theta)=0\}$ не зависит от θ .

Определение. Функция

$$U(\mathbf{X}_{n}, \theta) = \frac{\partial \ln L(\mathbf{X}_{n}, \theta)}{\partial \theta}$$

называется вкладом выборки.

Определение. *Информацией Фишера* о параметре θ содержащейся в выборке \mathbf{X}_n называется дисперсия вклада выборки:

$$I_n(\theta) = D[U(\mathbf{X}_n, \theta)] = n \cdot i(\theta),$$

где $i(\theta)$ — это информация Фишера по выборке из одного наблюдения, которая может быть найдена по формулам:

$$i(\theta) = M \left[\left[\frac{\partial \ln f(x,\theta)}{\partial \theta} \right]^2 \right] = \int_{R} \left(\frac{\partial \ln f(x,\theta)}{\partial \theta} \right)^2 f(x,\theta) dx =$$

$$= -M \frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} = -\int_{R} \frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} f(x,\theta) dx$$

Неравенство Рао-Крамера. Если выполняются условия регулярности, то для любой несмещенной оценки $T(\mathbf{X}_n)$ параметрической функции $\tau(\theta)$ справедливо неравенство:

$$D[T(\mathbf{X}_n)] \ge \frac{[\tau'(\theta)]^2}{ni(\theta)},$$

где $i(\theta)$ — информационное количество Фишера. Оценка, при которой достигается нижняя граница неравенства Рао-Крамера, называется эффективной.

Критерий эффективности. $T(\mathbf{X}_n)$ — эффективная оценка $\tau(\theta)$, если

$$T(\mathbf{X}_n) - \tau(\theta) = a(\theta)U(\mathbf{X}_n, \theta),$$
 (8.1)

где $a(\theta)$ – некоторая функция от θ .

Пример 8.1

Пусть $X_1,...,X_n$ – выборка из распределения Максвелла

с функцией плотности
$$f(x;\theta) = \frac{2x^2}{\theta^3 \sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\theta^2}\right\},$$
 $x>0, \theta>0.$

Требуется проверить, является ли оценка $\hat{\theta} = \frac{\sqrt{2\pi}}{4} \, \bar{X}$ эффективной.

Решение:

Проверим, достигается ли нижняя граница в неравенстве Рао-Крамера.

Найдем информационное количество Фишера:

$$i(\theta) = -M \left(\frac{\partial^2 \ln f(x; \theta)}{\partial \theta^2} \right) = -M \left(\frac{3}{\theta^2} - \frac{3x^2}{\theta^4} \right) =$$
$$= \frac{3}{\theta^2} \left(\frac{MX^2}{\theta^2} - 1 \right) = \frac{3}{\theta^2} \left(\frac{3\theta^2}{\theta^2} - 1 \right) = \frac{6}{\theta^2}$$

$$D(\hat{\theta}) = \frac{2\pi}{16n^2} \sum_{i=1}^n DX_i = \frac{\pi}{8n} \cdot \frac{3\pi - 8}{\pi} \theta^2;$$

$$\frac{1}{ni(\theta)} = \frac{\theta^2}{6n} \neq \frac{3\pi - 8}{8n} \theta^2 = D(\hat{\theta}), \quad \Rightarrow \quad \hat{\theta} \quad \text{ не } \quad \text{является}$$
 эффективной оценкой θ .

Пример 8.2

Найти функцию $\tau(\theta)$, допускающую эффективную оценку для параметра масштаба распределения Вейбулла:

$$f(x,\theta) = \frac{\alpha x^{\alpha-1}}{\theta^{\alpha}} \exp\left\{-\left(\frac{x}{\theta}\right)^{\alpha}\right\}, \ x \ge 0, \ \theta > 0, \ \alpha > 0.$$

Решение:

Вероятностная модель является регулярной, так как область значений случайной величины не зависит от параметров и функция плотности дифференцируема по θ . Поэтому можно воспользоваться критерием эффективности (8.1). Логарифмическая функция правдоподобия и её производная имеют вид:

$$\ln L(\mathbf{X}_{n}, \theta) = \sum_{i=1}^{n} \left[\ln \alpha \cdot X_{i}^{\alpha-1} - \alpha \ln \theta - \frac{X_{i}^{\alpha}}{\theta^{\alpha}} \right],$$

$$U(\mathbf{X}_{n}, \theta) = \frac{\partial \ln L(\mathbf{X}_{n}, \theta)}{\partial \theta} = -\frac{n\alpha}{\theta} + \frac{\alpha}{\theta^{\alpha+1}} \sum_{i=1}^{n} X_{i}^{\alpha}.$$

Отсюда

$$\frac{\theta^{\alpha+1}}{n\alpha}U(\mathbf{X}_n,\theta) = \frac{1}{n}\sum_{i=1}^n X_i^{\alpha} - \theta^{\alpha} = T(\mathbf{X}_n) - \tau(\theta).$$

Таким образом, оценка $T(\mathbf{X}_n) = \frac{1}{n} \sum_{i=1}^n X_i^{\alpha}$ является эффективной оценкой функции $\tau(\theta) = \theta^{\alpha}$.

Задачи:

8.1. Заполнить таблицу и сделать выводы о

выполнении неравенства Рао-Крамера

выполнении неравенетва тао-крамера						
Модель	$i(\theta)$	эффе	итерию ктив- $\hat{\tau}(\theta)$	$Mig[\hat{ au}(heta)ig]$	$D[\hat{ au}(heta)]$	$\frac{\left[\tau'(\theta)\right]^2}{ni(\theta)}$
Нормальное						
$N(\theta,\sigma)$						
Нормальное						
$N(\mu, \theta)$						
Гамма						
$\Gamma(\theta,\lambda)$						
Биномиально						
e						
$Bi(k,\theta)$						
Пуассона						
$\Pi(\theta)$						
Отрицатель-						
ное						
биномиальное						
$Bi(r,\theta)$						

- 8.2. Найти информационную матрицу Фишера для $N(\theta_1,\theta_2)$.
 - 8.3*. Найти $i(\theta)$ для распределения Коши.

Тема 9. Доверительные интервалы

Так как любая статистика является случайной величиной (имеющей некоторое распределение $G_{T_n}(x)$), то для каждой новой реализации выборки \mathbf{X}_n будет получаться другое значение оценки, в общем случае отличное от истинного значения параметра θ .

Определение. Интервальной оценкой (или доверительным интервалом) параметра θ называют интервал $[T_1(\mathbf{X}_n), T_2(\mathbf{X}_n)]$, содержащий истинное значение параметра θ с вероятностью γ .

9.1. Построение доверительного интервала с использованием распределения точечной оценки параметров

Если имеется некоторая точечная оценка $T_n = T(\mathbf{X}_n)$ для параметра θ и известна ее функция распределения $F_T(t,\theta)$, непрерывная и монотонная по θ , то доверительный интервал можно построить, основываясь на этой функции:

- 1. Вычисляем точечную оценку $T_n = T(\mathbf{X}_n)$.
- 2. Решаем относительно $\tilde{\theta}_1$, $\tilde{\theta}_2$ уравнения

$$F_T\left(T_n, \tilde{\theta}_1\right) = \frac{1-\gamma}{2}, F_T\left(T_n, \tilde{\theta}_2\right) = \frac{1+\gamma}{2}.$$

3. Определяем границы доверительного интервала:

$$T_1 = \min(\tilde{\theta}_1, \tilde{\theta}_2), T_2 = \max(\tilde{\theta}_1, \tilde{\theta}_2).$$

9.2. Построение доверительного интервала с использованием центральной статистики

Определение. Статистика $G(\mathbf{X}_n, \theta)$ называется *центральной статистикой*, если распределение $G(\mathbf{X}_n, \theta)$ не зависит от θ , и при любом фиксированном θ статистика $G(\mathbf{X}_n, \theta)$ непрерывна и строго монотонна по θ .

С помощью центральной статистики можно построить доверительный интервал. Пусть $f_G(g)$ плотность распределения статистики $G(\mathbf{X}_n, \theta)$.

- 1. Найдем такие значения $g_1,g_2,$ что $P \Big\{ g_1 < G \big(X_n, \theta \big) < g_2 \Big\} = \int\limits_0^{g_2} f_G \Big(g \Big) dg = \gamma \ .$
 - 2. Решим относительно \tilde{T}_1, \tilde{T}_2 уравнения

$$G(X_n, \tilde{T}_1) = g_1, G(X_n, \tilde{T}_2) = g_2.$$

3. Определяем границы доверительного интервала:

$$T_1 = \min\left\{\tilde{T}_1, \tilde{T}_2\right\}, T_2 = \max\left\{\tilde{T}_1, \tilde{T}_2\right\},$$

Для нахождения центральной статистики для параметра сдвига нормального закона будут полезны следующие теоремы.

Теорема 1. Пусть $\mathbf{X}_n = (X_1,...,X_n), \, X_i \in N(\mu,\sigma)$ и параметр σ известен. Тогда $G\big(\mathbf{X}_n,\theta\big) = \sqrt{n} \, \frac{\overline{X} - \theta}{\sigma}$ подчиняется стандартному нормальному закона N(0,1) .

Теорема 2. Пусть $\mathbf{X}_n = (X_1,...,X_n), X_i \in N(\mu,\sigma)$ и параметр σ неизвестен. Тогда $G\big(\mathbf{X}_n,\theta\big) = \sqrt{n-1}\frac{\overline{X}-\mu}{S}$ подчиняется распределению Стьюдента с n-1 степенью свободы, где $S^2 = \sum_{i=1}^n \left(X_i - \overline{X}\right)^2$ — выборочная дисперсия.

Для построения доверительного интервала с помощью центральной статистики основная проблема заключается в нахождении этой центральной статистики. Можно выделить класс моделей, для которых центральная статистика существует и имеет простой вид.

Пусть $F(x,\theta)$ — функция распределения наблюдаемой случайной величины, *монотонная* по параметру θ . Можно положить в качестве центральной статистики функцию $G(\mathbf{X}_n,\theta) = -\sum_{i=1}^n \ln F(X_i,\theta)$, которая подчинена гаммараспределению с параметром формы n.

Пример 9.1

Построить точный γ -доверительный интервал по выборке $X_1,...,X_n$ для параметра θ экспоненциального распределения $f(x;\theta)=\frac{1}{\theta}\exp\left\{-\frac{x}{\theta}\right\},\ x\geq 0$.

Решение:

Функция распределения $F(x;\theta) = 1 - \exp\left\{-\frac{x}{\theta}\right\}$ является монотонной (возрастающей) по параметру θ ($F_{\theta}^{'}(x;\theta) = \frac{x}{\theta^2}e^{-\frac{x}{\theta}} > 0$), следовательно, в качестве центральной статистики можно взять $G(\mathbf{X}_n,\theta) = -\sum_{i=1}^n \ln F(X_i,\theta)$, которая подчинена гамма-распределению с функцией плотности $f(x) = \frac{x^{n-1}e^{-x}}{\Gamma(n)}$, x > 0, n — объем выборки.

Тогда границы γ -доверительного интервала (T_1,T_2) определяются при численном решении уравнений: $G\left(\mathbf{X}_n,T_1\right)=g_1,\ G\left(\mathbf{X}_n,T_2\right)=g_2,\$ где g_1 и g_2 выбираются такими, что $P\left\{g_1 < G\left(\mathbf{X}_n,\theta\right) < g_2\right\} = \int\limits_{g_1}^{g_2} \frac{x^{n-1}e^{-x}}{\Gamma(n)} dx = \gamma$.

Задачи:

- 7.1. Построить у-доверительный интервал для оценки неизвестного математического ожидания нормальной случайной величины, если дисперсия известна.
- 7.2. С целью изучения размеров дневной выручки в сфере мелкого частного бизнеса была произведена 10%-я случайная бесповторная выборка из 1000 торговых киосков города. В результате были получены данные о средней дневной выручке, которая составила 500 у.е. В каких пределах с доверительной вероятностью 0,95 может находиться средняя точек всех торговых дневная выручка изучаемой совокупности, по предыдущим если исследованиям стандартное отклонение (σ) составило 150 (Предположить, что выручка подчинена нормальному закону).

- 7.3. Построить у-доверительный интервал для оценки неизвестного математического ожидания нормальной случайной величины, если дисперсия неизвестна.
- 7.4. Для оценки остаточных знаний по общеэкономическим предметам были протестированы 25 студентов 2-го курса факультета. Получены следующие результаты в баллах: (107, 90, 114, 88, 117, 110, 103, 120, 96, 122, 93, 100, 121, 110, 135, 85, 120, 89, 100, 126, 90, 94, 99, 116, 111). По этим данным найдите 95%-й доверительный интервал для оценки среднего балла тестирования всех студентов 2-го курса факультета, предполагая распределение полученного балла нормальным.
- 7.5. Построить γ -доверительный интервал для оценки неизвестной дисперсии нормальной случайной величины, если математическое ожидание известно.
- 7.6. Построить у-доверительный интервал для оценки неизвестного параметра масштаба экспоненциального распределения.
- 7.7. Выборка содержит следующие наблюдения: (1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23). Найти 99%-доверительный интервал для оценки неизвестного параметра масштаба экспоненциального распределения.
- 7.8. С помощью статистики X_1 по выборке объема 1 из равномерного распределения на отрезке $[0,\theta]$ построить точный доверительный интервал уровня γ для параметра θ .
- 7.9. Пусть $X_1,...,X_n$ выборка из равномерного распределения на отрезке $[0,\theta]$. С помощью статистики $X_{(n)}$ построить точный доверительный интервал уровня γ для параметра θ .

Тема 10. Асимптотические доверительные интервалы

Оценки максимального правдоподобия при достаточно общих условиях являются асимптотически эффективными и асимптотически нормальными, следовательно

$$P\left\{\left|\hat{\theta}_n-\theta\right|\sqrt{ni\left(\hat{\theta}_n\right)}\leq c_\gamma\right\}\to\Phi\left(c_\gamma\right)-\Phi\left(-c_\gamma\right)=2\Phi\left(c_\gamma\right)-1=\gamma$$
 где
$$\Phi\left(x\right) \ -\ \text{функция распределения стандартного нормального}$$
 закона,
$$i(\theta)=\int\limits_{-\infty}^{\infty}\left(\frac{\partial \ln f\left(x,\theta\right)}{\partial \theta}\right)^2f\left(x,\theta\right)dx \ -\ \text{информационное}$$
 количество Фишера,
$$\hat{\theta}_n \ -\ \text{ОМП. Отсюда} \quad c_\gamma=\Phi^{-1}\left(\frac{\gamma+1}{2}\right),$$
 тогда
$$\left(\theta_n-\frac{c_\gamma}{\sqrt{ni\left(\hat{\theta}_n\right)}};\theta_n+\frac{c_\gamma}{\sqrt{ni\left(\hat{\theta}_n\right)}}\right) \ -\ \text{асимптотически}$$

кратчайший γ -доверительный интервал для θ .

Пример 10.1

Пусть $X_1,...,X_n$ — выборка из гамма-распределения $\Gamma(\lambda,\alpha)$ с функцией плотности $f(x)=\frac{1}{\lambda^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\lambda}, \ x\geq 0.$ Построить асимптотический γ -доверительный интервал для параметра масштаба λ , считая, что $\alpha>0$ — известно.

Решение:

Пусть $\gamma = 0.99$.

Оценкой максимального правдоподобия параметра λ при известном параметре формы α имеет вид: $\hat{\lambda}_n = \frac{\overline{X}}{\alpha}$. ОМП параметров гамма-распределения являются асимптотически

нормальными, поэтому $\left|\hat{\lambda}_n - \lambda\right| \sqrt{ni\left(\hat{\lambda}_n\right)}$ сходится к стандартному нормальному распределению. По таблице из приложения 2:

$$c_{\gamma} = \Phi^{-1} \bigg(\frac{\gamma + 1}{2} \bigg) = \Phi^{-1}(0.995) = 2.57 \;,\; i \bigg(\hat{\lambda}_n \bigg) = \alpha \cdot \hat{\lambda}_n^2 \;.$$
 Следовательно, случайный интервал
$$\bigg(\frac{\overline{X}}{\alpha} - \frac{2.57 \sqrt{\alpha}}{\sqrt{n} \, \overline{X}} \;; \frac{\overline{X}}{\alpha} + \frac{2.57 \sqrt{\alpha}}{\sqrt{n} \, \overline{X}} \bigg) \;$$
 является асимптотическим 99% -

Пример 10.2

доверительным интервалом.

С помощью оценки \overline{X} построить асимптотический γ - доверительный интервал для неизвестного параметра p распределения Бернулли.

Решение:

По центральной предельной теореме распределение случайной величины

$$\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}}$$

слабо сходится к стандартному нормальному закону, а \overline{X} сходится по вероятности к p . Поэтому

$$G(\mathbf{X}_n, p) = \frac{\sqrt{n}(\overline{X} - p)}{\sqrt{\overline{X}(1 - \overline{X})}}$$

слабо сходится также к стандартному нормальному закону. Рассматривая $G(\mathbf{X}_n,p)$ как центральную статистику (см. тему 9) получаем случайный интервал

$$\left(\overline{X} - \frac{c_{\gamma}\sqrt{\overline{X}(1-\overline{X})}}{\sqrt{n}}, \overline{X} + \frac{c_{\gamma}\sqrt{\overline{X}(1-\overline{X})}}{\sqrt{n}}\right)$$

который является асимптотическим $^{\gamma}$ -доверительным интервалом, где $c_{\gamma} = \Phi^{-1} \bigg(\frac{\gamma+1}{2} \bigg)$ — квантиль порядка $\frac{\gamma+1}{2}$ стандартного нормального распределения.

Задачи:

- 10.1. В результате проверки 400 электрических лампочек 40 штук оказались бракованными. Найти 99%-доверительный интервал для вероятности брака.
- 10.2. С помощью оценки \bar{X} построить асимптотический γ -доверительный интервал для неизвестного параметра p биномиального распределения (значение параметра известно)
- $10.3.~\mathrm{C}$ помощью оценки \overline{X} построить асимптотический γ -доверительный интервал для неизвестного параметра λ распределения Пуассона
- 10.4. Пусть $X_1, X_2, ..., X_n$ выборка из распределения F_{θ} с конечной дисперсией, $E_{\theta}X_{_1} = \theta$ и $D_{\theta}X_{_1} = \sigma^2(\theta)$, где $\sigma^2(\theta)$ непрерывная по $\sigma(\theta)$ функция. С помощью оценки \overline{X} построить асимптотический γ -доверительный интервал для параметра θ .
- 10.5.~ Пусть $X_1, X_2, ..., X_n$ выборка из распределения Максвелла. Построить асимптотический γ -доверительный интервал для неизвестного параметра масштаба распределения Максвелла.

Тема 11. Проверка гипотезы о виде распределения по критерию хи-квадрат

Пусть имеется выборка $X_n = \{X_1, ..., X_n\}$ наблюдаемой случайной величины с функцией распределения $F_{\xi}(x)$.

Определение. Простой гипотезой является утверждение $H_0: F_{\varepsilon}(x) = F(x)$, где F(x) полностью задана.

Определение. Сложной гипотезой является утверждение $H_0: F_{\xi}(x) \in F = \{F(x,\theta), \theta \in \Theta\}.$

Процедура проверки гипотез о согласии с применением критерия типа χ^2 предусматривает группирование наблюдений. Область значений случайной величины разбивается на k непересекающихся интервалов граничными точками $t_0 < t_1 < ... < t_{k-1} < t_k$ где t_0 — нижняя граница области определения, t_k — верхняя грань. В соответствии с заданным разбиением подсчитывается число n_i выборочных значений, попавших в i-й интервал и вычисляют вероятность попадания

в интервал $P_i(\theta) = F(t_i) - F(t_{i-1}) = \int\limits_{t_{i-1}}^{t_i} f(x,\theta) dx$. При этом

$$\sum_{i=1}^{k} n_i = n; \sum_{i=1}^{k} P_i(\theta) = 1.$$

Статистика критерия χ^2 Пирсона имеет вид

$$S_{\chi^2} = n \sum_{i=1}^k \frac{\left(\frac{n_i}{n} - P_i(\theta)\right)^2}{P_i(\theta)}.$$

В случае проверки простой гипотезы при $n\to\infty$ статистика S_{χ^2} имеет распределение χ^2_{k-1} с k-1 степенями свободы. В случае проверки сложной гипотезы и при условии, что оценки находятся по методу минимума χ^2 статистика S_{χ^2} при $n\to\infty$ имеет распределение χ^2_{k-m-1} степенями свободы, где m — число оцениваемых параметров.

Пример 11.1 Пана выборка объема n = 30:

	F			
X	1	2	3	4
n_{i}	15	8	4	3

Требуется проверить гипотезу о согласии данной выборки с законом Пуассона.

Решение:

Зададимся уровнем значимости $\alpha = 0.05$.

Поскольку распределение случайных величин является дискретным, для проверки гипотезы о согласии воспользуемся критерием χ^2 Пирсона.

$$P\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \ k = 0, 1, 2, ...$$

ОМП для параметра λ является $\hat{\lambda}_n=\overline{X}$. Для данной выборки $\hat{\lambda}_n=1.83$. Тогда $P_0=P\{\xi=0\}=0.16$, $P_1=P\{\xi=1\}=0.29$, $P_2=P\{\xi=2\}=0.27$, $P_3=P\{\xi=3\}=0.16$, $P_4=P\{\xi=4\}=0.08$, $P_5=\sum_{k=5}^{\infty}P\{\xi=k\}=0.04$, $\sum_{i=0}^{5}P_i=1$. Статистика Пирсона:

$$X_n^2 = \sum_{i=0}^5 \frac{(n_i - nP_i)^2}{nP_i} = \frac{(0 - 30 \cdot 0.16)^2}{30 \cdot 0.16} + \frac{(15 - 30 \cdot 0.29)^2}{30 \cdot 0.29} + \frac{(8 - 30 \cdot 0.27)^2}{30 \cdot 0.27} + \frac{(15 - 30 \cdot 0.29)^2}{30 \cdot 0.29} + \frac{(15 -$$

$$+\frac{(4-30\cdot0.16)^2}{30\cdot0.16} + \frac{(3-30\cdot0.08)^2}{30\cdot0.08} + \frac{(0-30\cdot0.04)^2}{30\cdot0.04} = 10.85$$

В случае оценивания по данной выборке m параметров распределения, статистика X_n^2 Пирсона подчиняется χ^2 распределению с K-m-1 степенью свободы, где K — число групп. В данном случае число степеней свободы равно 6-1-1=4. Находим по таблице из приложения 3 критическое значение статистики Пирсона при $\alpha=0.05$: $S_\alpha=9.49$. Поскольку $X_n^2>S_\alpha$, то гипотеза о согласии данной выборки с распределением Пуассона отвергается. Отметим, что если $\alpha=0.01$, то гипотеза о согласии не отвергается.

Задачи:

- 11.1. Производитель некоторого вида продукции утверждает, что 95% выпускаемой продукции не имеют дефектов. Случайная выборка 100 изделий показала, что только 92 из них свободны от дефектов. Проверьте справедливость утверждения производителя продукции на уровне значимости $\alpha = 0.05$.
- 11.2. Владелец фирмы считает, что добиться более высоких финансовых результатов ему помешала неравномерность поставок комплектующих по месяцам года, несмотря на то, что поставщик в полном объеме выполнил свои обязательства за год. Поставщик утверждает, что поставки были не так уж неравномерны. Распределение поставок по месяцам года имеет следующий вид:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12
Объем поставок,	19	23	26	18	20	20	20	20	32	27	35	40
единиц												

На уровне значимости $\alpha = 0.05$ определите, кто прав: владелец фирмы или поставщик?

<u>Указание</u>: рассмотрите случайную величину "Месяц поставки единицы продукции", которая принимает дискретные значения от 1 до 12.

11.3. В таблице приведены данные о моментах поступления пациентов в отделение интенсивной терапии с понедельника 4 февраля 1963 г. по среду 18 марта 1964 г. сгруппированные по дням недели

День	Пн	Вт	Ср	\mathbf{q}_{T}	Пт	Сб	Вс
недели							
Число	37	53	35	27	30	44	28
пациентов							

- а) Проверить гипотезу о том, что пациенты попадают в отделение с равной вероятностью в любой из семи дней недели.
- б) Проверить гипотезу о том, что пациенты попадают в отделение с равной вероятностью в любой из семи дней недели, кроме вторника.
- 11.4. Через равные промежутки времени в тонком слое раствора золота регистрировалось число частиц золота, попавших в поле зрения микроскопа. По данным наблюдений, приведенных в следующей таблице проверить гипотезу о том, что число частиц золота является пуассоновской случайной величиной.

Число	0	1	2	3	4	5	6	7	Итого
частиц									
n_i	112	168	130	68	32	5	1	1	517

$$P\{\xi = k\} = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, ...$$

- 11.5. Цифры 0, 1, 2, ...,9 среди 800 первых десятичных знаков числа π появились 74, 92, 83, 79, 80, 73, 77, 75, 76, 91 раз соответственно. Проверить гипотезу о согласии этих данных с равномерным на множестве $\{0,1,\ldots,9\}$ законом распределения.
- 11.6. По официальным данным в Швеции в 1935г. родилось 88 273 ребенка, причем в январе родилось 7280 детей, в феврале 6957, марте 7883, апреле 7884, мае 7892, июне 7609, июле 7585, августе 7393, сентябре 7203, октябре 6903, ноябре 6552 и в декабре 7132 ребенка. Совместимы ли эти данные с гипотезой, что день рождения наудачу выбранного человека с равной вероятностью приходится на любой из 365 дней года?
- 11.7. Ниже приводятся результаты 4096 опытов, состоящих в одновременном подбрасывании 12 костей (данные Уэлдона). В каждом из опытов подсчитывалось число костей, выпавших кверху шестеркой (гранью с шестью очками). Проверить гипотезу правильности костей.

Число	0	1	2	3	4	5	6	≥ 7	Всего
шестерок									
n_i	447	1145	1181	796	380	115	24	8	4096

Тема 12. Проверка гипотезы о виде распределения по критерию Колмогорова

Распределение статистики имеет вид

$$D_n = \sup_{|x| < \infty} |F_n(x) - F(x, \theta)|,$$

где $F_n(x)$ — эмпирическая функция распределения, $F(x,\theta)$ — теоретическая функция распределения. При $n\to\infty$ статистика $\sqrt{n}D_n$ сходится равномерно к распределению Колмогорова с функцией распределения

$$K(S) = \sum_{-\infty}^{\infty} (-1)^k e^{-2k^2 S^2}$$
.

Наиболее часто в критерии используется статистика Колмогорова с поправкой Большева

$$S_k = \frac{6nD_n + 1}{6\sqrt{n}},$$

где
$$D_n = \max(D_n^+, D_n^-)$$
, $D_n^+ = \max_{1 \le i \le \infty} \left(\frac{i}{n} - F\left(X_{(i)}\right) \right)$,

$$D_{n}^{-} = \max_{1 \le i \le \infty} \left(F\left(X_{(i)}\right) - \frac{i-1}{n} \right).$$

При проверке простой гипотезы, гипотеза отклоняется если $S > K^{-1}(1-\alpha)$

При проверке сложной гипотезы распределение статистики критерия меняется и начинает зависеть от закона распределения случайной величины и гипотеза отклоняется, если $S>G_S^{-1}(1-\alpha)$, где $G_S(t)$ - функция распределения статистики при верной основной гипотезе.

Примечание. Критерий Колмогорова нельзя использовать для дискретных случайных величин и по группированным данным!

Пример 12.1

В следующей таблице представлены результаты измерений длин чайных ложечек в сантиметрах.

9,65	9,05	9,20	9,79	6,69	9,14	9,93	11,95	10,20	10,21
8,58	9,82	11,75	9,05	12,31	10,47	10,10	8,40	10,77	10,19
8,78	10,36	7,30	11,03	12,47	11,06	10,31	7,43	9,87	10,29
9,41	10,37	9,52	10,15	5,36	11,02	8,52	8,34	10,94	9,33
10,01	9,87	9,43	8,27	10,34	9,48	9,61	10,95	10,01	9,86

Требуется проверить гипотезу о согласии данной выборки с распределением Лапласа.

Решение:

Зададимся уровнем значимости $\alpha = 0.05$.

Поскольку мы имеем непрерывную случайную величину, то для проверки гипотезы о согласии воспользуемся критерием типа Колмогорова, статистика которого имеет вид: $S_K = \frac{6nD_n+1}{6\sqrt{n}}$, где $D_n = \max(D_n^+, D_n^-)$,

$$D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}, \qquad D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}. \qquad \text{Объем}$$

выборки n=50, $x_1,x_2,...,x_n$ – упорядоченные по возрастанию выборочные значения, $F(x,\theta)$ – функция распределения Лапласа.

Для нахождения ОМП параметров распределения воспользуемся программной системой ISW 4.0 https://ami.nstu.ru/~headrd/ISW.htm [10]: $\hat{a}_n = 9,87$, $\hat{\theta}_n = 0,94$.

Вычисляем значение статистики Колмогорова $S_K=0,55$. Находим по таблице из приложения 5 критическое значение статистики Колмогорова при $\alpha=0,05$: $S_\alpha=0,95$. Поскольку $S_K < S_\alpha$, то гипотеза о согласии данной выборки с распределением Лапласа не отвергается.

Задачи:

12.1. Наблюдалось положение секундной стрелки на часах в случайный момент времени. В результате была получена следующая выборка:

23	15	55	34	2	47	39	54	12	9
4	54	21	27	11	49	32	31	7	12

Проверить гипотезу о том, что:

- а) Выборка подчиняется равномерному распределению с параметрами (0,60);
- б) Выборка подчиняется экспоненциальному распределению с неизвестным параметром масштаба.
- 12.2. Мука расфасовывается в пакеты так, чтобы вес одного пакета был 10 кг. В результате измерения веса 20 случайно выбранных пакетов получены измерения:

9,473	10,016	9,703	10,306	10,102
10,002	9,655	10,077	10,084	9,775
10,225	9,861	9,887	9,964	9,974
10,132	10,182	9,788	10,225	9,757

Проверить гипотезу, что:

- а) выборка подчиняется равномерному распределению на интервале от 9,5 до $10,5~{\rm KF};$
- б) выборка подчиняется распределению Лапласа с параметром сдвига 10 и параметром масштаба 0,25;
- в) выборка подчиняется нормальному распределению с математическим ожиданием 10 и стандартным отклонением 0,25;
- г) выборка подчиняется нормальному распределению с неизвестными параметрами;
- д) выборка подчиняется логистическому распределению с неизвестными параметрами (найти ОМП с момощью ISW).

Тема 13. Проверка гипотезы независимости, проверка гипотезы однородности

13.1. Гипотеза независимости

В эксперименте наблюдается двумерная случайная величина $\xi=(\xi_1,\xi_2)$ с неизвестной функцией распределения $F_\xi(x,y)$, и есть основания предполагать, что ξ_1 и ξ_2 независимы. В этом случае нужно проверить гипотезу независимости:

$$H_0: F_{\xi}(x, y) = F_{\xi_1}(x)F_{\xi_2}(y),$$

где $F_{\xi_1}(x)$ и $F_{\xi_2}(y)$ некоторые одномерные функции распределения.

Для проверки гипотезы независимости используется критерий χ^2 Пирсона. Если исходные данные негруппированы, то предварительно производится группировка наблюдений.

Пусть случайная величина ξ_1 принимает значения $c_1,...,c_s$, а ξ_2 — $b_1,...,b_k$. Обозначим v_{ij} количество наблюдений $\left(c_i,b_j\right),$ $\sum_{i=1}^s\sum_{i=1}^k v_{ij}=n$.

ξ_1 ξ_2	v_{1i}	 v_{ki}	v_{i^*}
v_{Ij}	$ u_{11}$	 $ u_{1k}$	
			v_{I^*}
•••	•••	•••	
v_{sj}	$ u_{s1}$	 ${m u}_{sk}$	
			${\mathcal V}_{{\mathcal S}}*$
v_{*j}	v_{*I}	 v_{*k}	

Статистика критерия независимости χ^2 Пирсона

$$\chi_n^2 = n \sum_{i=1}^s \sum_{j=1}^k \frac{\left(v_{ij} - v_{i,} v_{,j} / n\right)^2}{v_{i,} v_{,j}} = n \left(\sum_{i=1}^s \sum_{j=1}^k \frac{v_{ij}^2}{v_{i,} v_{,j}} - 1\right)$$

имеет распределение $\chi^2_{((s-1)(k-1))}$ при $n \to \infty$.

Пример 13.1

В следующей таблице представлены значения показателя Y и значения показателя X в течение 12 лет.

Год	Y	X	Год	Y	\boldsymbol{X}
1986	152	170	1992	177	200
1987	159	179	1993	179	207
1988	162	187	1994	184	215
1989	165	189	1995	186	216
1990	170	193	1996	190	220
1991	172	199	1997	191	225

Проверить гипотезу о независимости величин X и Y . Решение:

Для проверки гипотезы независимости воспользуемся критерием независимости χ^2 . Зададимся уровнем значимости $\alpha=0.05$. Составим таблицу сопряженности двух признаков: $i=\overline{1,s}$, $j=\overline{1,k}$:

X Y (151,161]	(161,171]	(171,181]	(181,191]	V_{i} .	
---------------	-----------	-----------	-----------	-----------	--

(165,180]	2	0	0	0	2
(180,195]	0	3	0	0	3
(195,210]	0	0	3	0	3
(210,225]	0	0	0	4	4
$\overline{v_{\bullet j}}$	2	3	3	4	12

Статистика критерия независимости χ^2 : $X_n^2 = n \left(\sum_{i,j} \frac{V_{ij}^2}{V_i V_{i,j}} - 1 \right)$ имеет χ^2 -распределение с числом степеней свободы (s-1)(k-1). Вычислим значение статистики: $X_n^2 = 36$, число степеней свободы (s-1)(k-1) = 9. Находим по таблице из приложения 3 критическое значение статистики Пирсона при $\alpha = 0.05$: $S_\alpha = 16.9$. Поскольку $X_n^2 > S_\alpha$, то гипотеза о независимости признаков X и Y отвергается.

13.2. Гипотеза однородности

Пусть произведено k серий независимых наблюдений $X_{n_1}=\{x_1^1,...,x_{n_1}^1\},\,X_{n_2}=\{x_1^2,...,x_{n_2}^2\},...,\,X_{n_k}=\{x_1^k,...,x_{n_k}^k\}$ и пусть $F_i(x)$ — функция распределения i-й серии. Чтобы проверить менялось ли распределение от серии к серии, можно сформулировать гипотезу однородности:

$$H_0: F_1(x) = F_2(x) = \dots = F_k(x)$$
,

при этом само распределение F(x) может быть неизвестным.

Для проверки гипотезы однородности используется критерий Смирнова (если выборки негруппированы) и χ^2 Пирсона (если выборки группированы).

13.3. Критерий однородности χ^2 Пирсона

Пусть осуществляется k последовательных серий независимых наблюдений, состоящих из $n_1, n_2, ..., n_k$ наблюдений. Пусть v_{ij} — число наблюдений i -го исхода в j -й серии. Пусть p_{ij} — неизвестная вероятность появления i -го исхода в j -й серии. (i=1,...,s;j=1,...,k).

 $H_0: \left(p_{1j}, \dots, p_{sj}\right) = \left(p_1, \dots, p_s\right), \ j=1,\dots,k. \qquad \text{Так} \qquad \text{как}$ $M\left(v_{ij} \mid H_0\right) = n_j p_i, \ \text{то, следуя принципу } \chi^2 \quad \text{в качестве меры}$ отклонения оптимальных данных от гипотетических следует выбрать статистику $\sum_{i=1}^s \sum_{j=1}^k \frac{\left(v_{ij} - n_j p_i\right)^2}{n_i p_i} \ .$

 p_i неизвестны и их следует предварительно оценить: $p_i = \frac{v_{ij}}{n} \,,$ где $n = n_1 + \ldots + n_k \,-$ общее число наблюдений.

Тогда стаьтистика

$$\chi_n^2 = n \sum_{i=1}^s \sum_{j=1}^k \frac{\left(v_{ij} - n_j v_i / n\right)^2}{n_j v_i} = n \left(\sum_{i=1}^s \sum_{j=1}^k \frac{{v_{ij}}^2}{n_j v_i} - 1\right).$$

при $n \to \infty$ имеет распределение $\chi^2_{((s-1)(k-1))}$.

13.4. Критерий Смирнова

 $D_{nm} = \sup_{x} \left| F_{1n} \left(x \right) - F_{2n} \left(x \right) \right|, \quad \text{где} \quad F_{in} \left(x \right) \quad - \quad \text{эмпирическая}$ функция распределения. Статистика $S_{nm} = \sqrt{\frac{nm}{n+m}} D_{nm}$ при

больших n и m и верной гипотезе H_0 имеет распределение Колмогорова. Поэтому критическая область строится следующим образом:

$$X_{1\alpha} = \left\{ X_n, Y_m \middle| S_{nm} \left(X_n, Y_m \right) \ge c_\alpha \middle| H_0 \right\}$$
$$P\left\{ X_{1\alpha} \right\} = \alpha \Longrightarrow c_\alpha = K^{-1} (1 - \alpha).$$

Пример 13.2

Проверить гипотезу об однородности двух выборок:

X:	3,49	3,5	3,52	3,62	3,79	3,8	3,81	3,99	4,01	4,05
Y:	3,8	3,81	3,83	3,85	3,86	3,9	4,1	4,38	4,66	4,96

Решение:

Так как выборка является негруппированной, то для проверки гипотезы однородности выборок X и Y можно воспользоваться критерием однородности Смирнова. Зададимся уровнем значимости $\alpha = 0,05$.

Статистика критерия однородности Смирнова: $S_{nm} = D_{nm} \sqrt{\frac{nm}{n+m}} \;, \; \; \text{где} \quad D_{nm} = \sup_x \left| F_n^1(x) - F_m^2(x) \right| \; \; \text{подчиняется}$ распределению Колмогорова K(S) . $F_n^1(x)$ — эмпирическая функция распределения по первой выборке, $F_m^2(x)$ — по второй. Проводя вычисления, получаем: $D_{nm} = 0.5$, n = m = 10 , $S_{nm} = 1,118$. Находим по таблице из приложения 4 критическое значение статистики Смирнова при $\alpha = 0,05$: $S_\alpha = 1,36$. Поскольку $S_{nm} < S_\alpha$, то нет оснований для отклонения гипотезы об однородности выборок X и Y.

Задачи:

13.1. Из 300 абитуриентов, поступивших в институт, 97 человек имели балл 5 в школе и 48 получили 5 на вступительных экзаменах по тому же предмету, причем только 18 человек имели 5 и в школе и на экзамене. С уровнем

значимости 0,1 проверить гипотезу о независимости оценок 5 в школе и на экзамене.

13.2. В следующей таблице приведены 818 случаев, классифицированных по двум признакам: наличию прививки против холеры (признак A) и отсутствию заболевания (признак B):

	В	\overline{B}	Σ
A	276	3	279
\overline{A}	473	66	539
Σ	749	69	818

Проверить гипотезу о независимости признаков A и B.

 $13.3.~\mathrm{B}$ следующей таблице приведены данные о распределении доходов (в тыс. крон) заводских мастеров Швеции в $1930~\mathrm{\Gamma}$. для возрастных групп $40-50~\mathrm{net}$ и $50-60~\mathrm{net}$.

доходы \	40 – 50 лет	50 – 60 лет
возраст		
0 - 1	71	54
1 - 2	430	324
2 - 3	1072	894
3 – 4	1609	1202
4 – 6	1178	903
> 6	158	112

Требуется проверить гипотезу о том, что доходы заводских мастеров возрастной группы 40-50 лет и доходы заводских мастеров возрастной группы 50-60 лет распределены одинаково.

13.4. В таблице приведены результаты основных и контрольных анализов на содержание никеля в процентах по одному из уральских месторождений.

Основной	0,62	0,68	0,20	0,57	0,31	0,10	0,09	0,58	0,43
анализ									
Контрольный	0,72	0,68	0,37	0,71	0,52	0,23	0,20	0,88	0,69
анализ									
				<u> </u>	<u> </u>				·

Основной	0,40	0,15	0,49	0,36	0,25	0,13	0,33	0,55	0,18
анализ									

Контрольный	0,46	0,25	0,69	0,61	0,42	0,17	0,49	0,59	0,22
анализ									
Основной	0,	20	0,27	0,50	0,38	0,2	27 (),15	0,34
анализ									
Контрольный	0,	16	0,33	0,69	0,54	0,3	38 (),30	0,48

Проверить гипотезу об однородности основной и контрольной выборки.

13.5. Физическая подготовка 9 спортсменов была проверена при поступлении в спортивную школу, а затем после недели тренировок. Итоги проверки в баллах оказались следующими.

Число баллов	76	71	57	49	70	69	26	65	59
полученных при									
поступлении в школу									
Число баллов	81	85	52	52	70	63	33	83	62
полученных после									
недели обучения									

Проверить значимо или незначимо изменилась физическая подготовка спортсменов.

Тема 14. Ошибки первого и второго рода. Мощность критерия

С проверкой статистических гипотез связывают ошибки двух типов.

Определение. *Ошибкой первого рода* называют событие, когда верная проверяемая гипотеза отвергается критерием.

Определение. *Ошибкой второго рода* называют событие, когда неверная проверяемая гипотеза принимается критерием.

Вероятности ошибок первого и второго рода обозначают α и β , соответственно. Вероятность ошибки второго рода зависит от выдвигаемой конкурирующей гипотезы.

Определение. Вероятность отклонения ложной проверяемой гипотезы, т.е. принятия правильного решения в пользу конкурирующей, называется *мощностью*, и она равна $1-\beta$.

Определение. Вероятность ошибки первого рода также называют *уровнем значимости* критерия.

Задачи:

14.1. На рисунке 14.1 даны распределения статистики критерия при верной нулевой гипотезе $G(S \big| H_0)$, и при верной альтернативной гипотезе — $G(S \big| H_1)$.

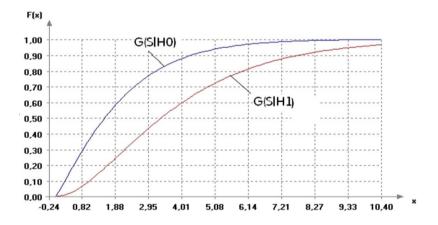


Рис. 14.1 Распределения статистики при верности H_0 и при верности H_1

- а) Вычислите ошибки первого и второго рода, если гипотеза H_0 отвергается при S>6,14. Является ли критерий несмещенным?
- б) Вычислите вероятности ошибок первого и второго рода, если известно, что гипотеза H_0 отклоняется при значениях статистики меньше 1,88 или больше 6,14. Является ли критерий несмещенным?
- в) Вычислите мощность критерия при заданном уровне значимости критерия 0,1, если известно, что гипотеза H_0 отклоняется при больших значениях статистики. Является ли критерий несмещенным?
- г) Вычислите мощность критерия при заданном уровне значимости критерия 0,3, если известно, что гипотеза H_0 отклоняется при значениях статистики близких к нулю. Является ли данный критерий несмещенным?

14.2. На рисунке 14.2 даны распределения статистики критерия при верной нулевой гипотезе $G(S|H_0)$, и при верной альтернативной гипотезе $-G(S|H_1)$.

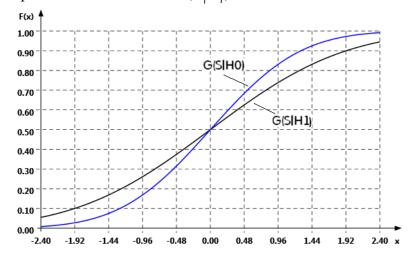


Рис. 14.2 Распределения статистики при верности $oldsymbol{H}_0$ и при верности $oldsymbol{H}_1$

- а) Вычислите ошибки первого и второго рода, если гипотеза H0 отвергается при S>1,44 или S<-1,44. Является ли критерий несмещенным?
- б) Вычислите мощность критерия при заданном уровне значимости критерия 0,1, если известно, что гипотеза отклоняется при больших значениях статистики. Является ли критерий несмещенным?
- в) Вычислите мощность критерия при заданном уровне значимости критерия 0,1, если известно, что гипотеза отклоняется при больших по абсолютной величине значениях статистики. Является ли данный критерий несмещенным?

Тема 15. Построение наиболее мощного критерия

Пусть в эксперименте требуется выбрать одну из двух гипотез:

$$H_0: f_{\xi}(x) = f_0(x)$$
,

$$H_1: f_{\varepsilon}(x) = f_1(x),$$

где f(x)- функция плотности, если случайная величина ξ является непрерывной, и вероятность события $P\{\xi=x\}$, если ξ - дискретная случайная величина.

Критерий отношения правдоподобия строится на основании *статистики отношения правдоподобия*

$$\Lambda_n = \prod_{i=1}^n \frac{f_1(x_i)}{f_0(x_i)},$$

или ее логарифма

$$\lambda_n = \ln \Lambda_n = \ln \prod_{i=1}^n \frac{f_1(x_i)}{f_0(x_i)} = \sum_{i=1}^n \ln \frac{f_1(x_i)}{f_0(x_i)}.$$

 Гипотеза H_0 отвергается по критерию отношения правдоподобия, если $\lambda_n > t_\alpha$, где критическое значение t_α определяется из условия $P\left\{\lambda_n > t_\alpha \left| H_0 \right.\right\} = \alpha$.

Отсюда, в случае непрерывности случайной величины ξ

$$t_{\alpha} = F_{\lambda_n|H_0}^{-1} \left(1 - \alpha\right) ,$$

где $F_{\lambda_n|H_0}^{-1}(y)$ - обратная функция распределения случайной величины λ_n при верной гипотезе H_0 . Мощность критерия отношения правдоподобия определяется соотношением

$$1 - \beta = P\left\{\lambda_n > t_\alpha \left| H_1 \right\} = F_{\lambda_n \mid H_1} \left(t_\alpha \right) \ .$$

Лемма Неймана-Пирсона. Среди всех критериев заданного уровня значимости α , проверяющих две простые

гипотезы H_0 и H_1 , критерий отношения правдоподобия является наиболее мощным.

Задачи:

Вычислить мощность.

Построить наиболее мощный критерий проверки гипотезы $H_0: f(x) = \begin{cases} 2x, & x \in [0,1] \\ 0 & x \notin [0,1] \end{cases}$ против альтернативы $H_1: f(x) = \begin{cases} 2(1-x), & x \in [0,1] \\ 0 & x \notin [0,1] \end{cases}$ по выборке из одного

наблюдения. Вычислить мощность.

15.2. Построить наиболее мощный критерий проверки гипотезы $H_0: f(x) = \begin{cases} 1, & x \in [0,1] \\ 0 & x \not\in [0,1] \end{cases}$ против альтернативы $H_1: f(x) = e^{-x}, x \geq 0$ по выборке из одного наблюдения.

15.3. Пусть X_1 — выборка объема 1. Гипотеза состоит в том, что X_1 имеет показательное распределение с параметром $\alpha=2$. Альтернатива состоит в том, что X_1 имеет плотность

$$f_2 = \begin{cases} 1/2, y \in [0,1], \\ 1, y \in [3/2,2], \\ 0, y \notin [0,1] \cup [3/2,2]. \end{cases}$$

Построить наиболее мощный критерий размера 1/3.

15.4. Построить наиболее мощный критерий проверки гипотезы $H_0: \theta = \theta_0$ против простой альтернативы $H_1: \theta = \theta_1 > \theta_0$ о параметре θ распределения с плотностью $f(x,\theta) = \frac{1}{\theta} e^{-x/\theta}, x \ge 0$ по выборке из одного наблюдения. Вычислить мощность.

- 15.5. Пусть $X_1,...,X_n$ выборка из нормального распределения со средним a и известной дисперсией σ^2 . Построить наиболее мощный критерий размера ε для проверки гипотезы $H_1 = \{a = a_1\}$ против альтернативы $H_2 = \{a = a_2\}$, где $a_1 < a_2$. Будет ли этот критерий состоятельным?
- 15.6. При выборке объема n при заданной вероятности ошибки первого рода построить наиболее мощный критерий для различения двух простых гипотез относительно неизвестной дисперсии нормального распределения, если математическое ожидание известно и равно нулю.

Литература

- 1. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1983. 416 с.
- 2. Губарев В.В. Вероятностные модели: Справочник. В 2-х ч. /Новосиб.электротехн. ин-т. Новосибирск, 1992. Ч.1 198 с. Ч.2 188 с.
- 3. Гланц С. Медико-биологическая статистика. М.: Практика, 1998.-459 с.
- 4. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М.: Высшая школа, 1979— 400 с.
- 5. Ивченко Г.И., Медведев Ю.А. Математическая статистика: Учеб. пособие для втузов. М.: Высшая школа, 1994. 248 с.
- 6. Ивченко Г.И., Медведев Ю.А., Чистяков А.В. Сборник задач по математической статистике. М.: Высшая школа, 1989.-255 с.
- 7. Коршунов Д.А., Чернова Н.И. Сборник задач и упражнений по математической статистике: учебное пособие. Новосибирск: Изд-во Института математики, 2001. 120 с.
- 8. Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. Новосибирск: изд-во $H\Gamma TY$, 2004. 120 с.
- 9. Лемешко Б.Ю. Критерии проверки гипотез об однородности. Руководство по применению : монография / Б.Ю. Лемешко. М. : ИНФРА-М, 2017. 208 с. (Научная мысль). ISBN: 978-5-16-012557-2 (print), 978-5-16-105463-5(online) DOI: 10.12737/22368
- 10. Лемешко Б.Ю., Блинов П.Ю. Критерии проверки отклонения распределения от равномерного закона. Руководство по применению: Монография / Б.Ю. Лемешко,

- П.Ю. Блинов. М.: ИНФРА-М, 2015. 183 с. (Научная мысль). DOI: 10.12737/11304
- 11. Лемешко Б.Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению: Монография / Б.Ю. Лемешко. М.: ИНФРА-М, 2015. 160 с. (Научная мысль). DOI: 10.12737/6086
- 12. Лемешко Б.Ю. Непараметрические критерии согласия: Руководство по применению: Монография / Б.Ю. Лемешко.— М.: ИНФРА-М, 2014.-163 с. DOI: 10.12737/11873
- 13. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: Монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. Новосибирск: Изд-во НГТУ, 2011. 888 с. (серия «Монографии НГТУ»).

Приложение 1 Основные законы распределения случайных величин

No	Распределение,	Плотность (вероятность) распределения,
	параметры	область значений случайной величины
		Дискретные
1.	Биномиальное,	$P\{\xi=k\}=C_m^k p^k (1-p)^{m-k},$
	$p \in [0,1]$	$k=0, \ldots, m$
2.	Отрицательное биномиальное,	$P\{\xi=k\}=C_{m+k-1}^{k}(1-p)^{k}p^{m},$
	$p \in [0,1]$	$k = 0, 1, \dots$
3.	Геометрическое, $p \in [0,1]$	$P\{\xi=k\}=(1-p)^k p, k=0, 1,$
4.	Пуассона, $p \in [0,1]$	$P\{\xi=k\} = \frac{p^k}{k!}e^{-p}, k=0, 1, \dots$
5.	Паскаля, $p \in [0,1]$	$P\{\xi=k\} = \frac{p^k}{(1+p)^{k+1}}, k=0, 1, \dots$
	Нег	прерывные
6.	Равномерное, $\lambda > 0$	$f(x) = \frac{1}{\lambda}, \ x \in [0, \lambda]$
7.	Бета-распределение, $\lambda > 0$, $\alpha > 0$, $\beta > 0$	$f(x) = \frac{\Gamma(\alpha + \beta)}{\lambda^{\alpha + \beta - 1} \Gamma(\alpha) \Gamma(\beta)} x^{\alpha - 1} (\lambda - x)^{\beta - 1},$
		$x \in [0, \lambda]$
8.	Нормальное, $a \in R$, $\lambda > 0$	$f(x) = \frac{1}{\sqrt{2\pi}\lambda} e^{-\frac{(x-a)^2}{2\lambda^2}}, x \in R$
9.	Лапласа, $a \in R$, $\lambda > 0$	$f(x) = \frac{1}{2\lambda} e^{-\frac{ x-a }{\lambda}}, x \in R$
10.	Двустороннее экспоненциальное, $a \in R$, $\lambda > 0$, $\alpha > 0$	$f(x) = \frac{\alpha}{2\lambda\Gamma(1/\alpha)}e^{-\left \frac{x-a}{\lambda}\right ^{\alpha}}, x \in R$

11.	Экспоненциальное, $a \in R$, $\lambda > 0$	$f(x) = \frac{1}{\lambda} e^{-(x-a)/\lambda} , x \ge a$
12.	Полунормальное, $\lambda > 0$	$f(x) = \frac{2}{\lambda\sqrt{2\pi}}e^{-\frac{x^2}{2\lambda^2}}, \ x \ge 0$
13.	Рэлея, $\lambda > 0$	$f(x) = \frac{x}{\lambda^2} e^{-\frac{x^2}{2\lambda^2}}, x \ge 0$
14.	Максвелла, $\lambda > 0$	$f(x) = \frac{2x^2}{\lambda^3 \sqrt{2\pi}} e^{-\frac{x^2}{2\lambda^2}}, \ x \ge 0$
15.	Гамма, $\lambda > 0$, $\alpha > 0$	$f(x) = \frac{1}{\lambda^{\alpha} \Gamma(\alpha)} x^{\alpha - 1} e^{-x/\lambda}, \ x \ge 0$
16.	Вейбулла-Гнеденко, $\lambda > 0$, $\alpha > 0$	$f(x,\theta) = \frac{\alpha x^{\alpha-1}}{\lambda^{\alpha}} \exp\left\{-\left(\frac{x}{\lambda}\right)^{\alpha}\right\}, \ x \ge 0$
17.	Логнормальное, $\alpha \in R$, $\beta > 0$	$f(x) = \frac{1}{x\beta\sqrt{2\pi}}e^{-(\ln x - \alpha)^2/2\beta^2}, \ x \ge 0$
18.	Парето $\beta > 0, \ \theta > 0$	$f(x) = \beta \theta^{\beta} x^{-(\beta+1)}, \ x \ge \beta$

Приложение 2

Таблица стандартного нормального распределения

В таблице показаны значения функции распределения стандартного нормального закона, например, значение функции в точке 2,57 находится в строке "2,5" и колонке "0,07" и равно 0,9949.

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,504	0,504	0,508	0,512	0,516	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,591	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,648	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,67	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,695	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,719	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,758	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,791	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,834	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,877	0,879	0,881	0,883
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,898	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,937	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,975	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,983	0,9834	0,9838	0,9842	0,9846	0,985	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,989

2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,992	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,994	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,996	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,997	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,998	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,999	0,999
3,1	0,999	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997

Приложение 3

Таблица распределения Стьюдента

В таблице показаны квантили $t_{p,\gamma}$ функции распределения Стьюдента при разных степенях свободы n.

$$p = \frac{\gamma + 1}{2} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\Gamma(n/2)} \int_{-\infty}^{t_{p,n}} \frac{dx}{(x^2/n+1)^{(n+1)/2}}$$

	p = 0.95	p = 0.975	p = 0.99	p = 0.995
n	$\gamma = 0.9$	$\gamma = 0.95$	$\gamma = 0.98$	$\gamma = 0.99$
1	6,3138	12,7062	31,8205	63,6567
2	2,9200	4,3027	6,9646	9,9248
3	2,3534	3,1825	4,5407	5,8409
4	2,1318	2,7765	3,7470	4,6041
5	2,0150	2,5706	3,3649	4,0321
6	1,9432	2,4469	3,1427	3,7074
7	1,8946	2,3646	2,9980	3,4995
8	1,8595	2,3060	2,8965	3,3554
9	1,8331	2,2622	2,8214	3,2498
10	1,8125	2,2281	2,7638	3,1693
11	1,7959	2,2010	2,7181	3,1058
12	1,7823	2,1788	2,6810	3,0545
13	1,7709	2,1604	2,6503	3,0123
14	1,7613	2,1448	2,6245	2,9768
15	1,7531	2,1315	2,6025	2,9467
16	1,7459	2,1199	2,5835	2,9208
17	1,7396	2,1098	2,5669	2,8982
18	1,7341	2,1009	2,5524	2,8784
19	1,7291	2,0930	2,5395	2,8609
		C 0		

20	1,7247	2,0860	2,5280	2,8453
21	1,7207	2,0796	2,5177	2,8314
22	1,7171	2,0739	2,5083	2,8188
23	1,7139	2,0687	2,4999	2,8073
24	1,7109	2,0639	2,4922	2,7969
25	1,7081	2,0595	2,4851	2,7874
26	1,7056	2,0555	2,4786	2,7787
27	1,7033	2,0518	2,4727	2,7707
28	1,7011	2,0484	2,4671	2,7633
29	1,6991	2,0452	2,4620	2,7564
30	1,6973	2,0423	2,4573	2,7500
∞	1,6449	1,9600	2,3264	2,5758

Приложение 3

Таблица распределения χ^2

В таблице показаны квантили $t_{\mathrm{l-}\alpha}$ функции распределения χ^2 при разных степенях свободы n.

$$1 - \alpha = \frac{1}{2^{n/2} \Gamma(n/2)} \int_{0}^{t_{1-\alpha}} x^{n/2-1} e^{-x/2} dx$$

n	$\alpha = 0,1$	$\alpha = 0.05$	$\alpha = 0.01$
1	2,71	3,84	6,64
2	4,61	5,99	9,21
3	6,25	7,81	11,3
4	7,78	9,49	13,3
5	9,24	11,1	15,1
6	10,6	12,6	16,8
7	12,0	14,1	18,5
8	13,4	15,5	20,1
9	14,7	16,9	21,7
10	16,0	18,3	23,2
11	17,3	19,7	24,7
12	18,5	21,0	26,2
13	19,8	22,4	27,7
14	21,1	23,7	29,1
15	22,3	25,0	30,6
16	23,5	26,3	32,0
17	24,8	27,6	33,4
18	26,0	28,9	34,8
19	27,2	30,1	36,2
20	28,4	31,4	37,6

Приложение 4 Таблица распределения Колмогорова

В таблице показаны квантили $t_{1-\alpha} = K^{-1}(1-\alpha)$ функции распределения Колмогорова.

	$\alpha = 0,1$	$\alpha = 0.05$	$\alpha = 0.01$
t_{1-lpha}	1,2238	1,3581	1,6276

Приложение 5 Таблица распределения статистики Колмогорова при проверке сложных гипотез

В таблице показаны квантили $t_{1-\alpha} = G^{-1}(1-\alpha)$ функции распределения статистики критерия Колмогорова при проверке сложной гипотезы, когда параметры распределения при верной гипотезе H_0 оцениваются по методу максимального правдоподобия.

№	Распределение	$\alpha = 0,1$	$\alpha = 0.05$	$\alpha = 0.01$
1	Экспоненциальное	0,9841	1,0794	1,2838
2	Лапласа	0,8710	0,9497	1,1206
3	Нормальное	0,8333	0,9042	1,0599
4	Логистическое	0,7451	0,8036	0,9261

Оглавление

Тема 1. Выборки и статистики	3
Тема 2. Порядковые статистики	6
Тема 3. Непараметрическое оценивание функции распределения	8
3.1. Эмпирическая функция распределения	
3.2. Цензурированная выборка. Оценка Каплана-Мейера	11
Тема 4. Непараметрическое оценивание функции плотности	14
4.1. Гистограмма	14
4.2. Ядерная оценка плотности	
Тема 5. Свойства оценок. Состоятельность и несмещенность	18
Тема 6. Метод моментов	
Тема 7. Метод максимального правдоподобия	
Тема 8. Информационное количество Фишера, неравенство Рао-Кр	амера и
критерий эффективности	
Тема 9. Доверительные интервалы	33
9.1. Построение доверительного интервала с использованием	
распределения точечной оценки параметров	33
9.2. Построение доверительного интервала с использованием	
центральной статистики	34
Тема 10. Асимптотические доверительные интервалы	38
Тема 11. Проверка гипотезы о виде распределения по критерию хи-	-
Тема 12. Проверка гипотезы о виде распределения по критерию	41
Колмогорова	46
Тема 13. Проверка гипотезы независимости, проверка гипотезы	
однородности	49
13.1. Гипотеза независимости	
13.2. Гипотеза однородности	
13.3. Критерий однородности χ^2 Пирсона	
13.4. Критерий Смирнова	
Тема 14. Ошибки первого и второго рода. Мощность критерия	
Тема 15. Построение наиболее мощного критерия	
Литература	
Приложение 1	
Основные законы распределения случайных величин	
Приложение 2	
Таблица стандартного нормального распределения	
Приложение 3	

Таблица распределения Стьюдента	68
Приложение 3	
Таблица распределения χ^2	70
Приложение 4	
Таблица распределения Колмогорова	
Приложение 5	
Таблица распределения статистики Колмогорова при проверке слож	жных
гипотез	71