Министерство науки
и высшего образования Российской Федерации
Алтайский государственный университет

# Сборник задач по статистике, случайным числам и процессам

*Задачник*

Барнаул
Издательство
Алтайского государственного
университета
2018

Составитель
к.ф.-м.н., доцент ***С.В. Дронов***

Рецензенты
д.ф.-м.н., профессор ***Е.Д. Родионов (АлтГУ)***
к.ф.-м.н., доцент ***К.О. Кизбикенов (АлтГПУ)***

Задачник соответствует программам учебных курсов «Математическая статистика и теория случайных процессов», «Математическая статистика и теория случайных чисел»второго курса математических направлений. Приводится большое количество задач и разнообразного статистического материала для практических занятий по курсам, разобраны типичные примеры.

Подписано в печать 17.07.2018. Формат 60x84/16
Усл.-печ. л. 2,56. Тираж 100 экз. Заказ № 362
Типография Алтайского государственного университета:
656099, Барнаул, ул. Димитрова, 66

# 1. Знакомство с основными понятиями математической статистики

Основным объектом, с которым мы будем иметь дело, будет выборка из распределения некоей случайной величины $\xi$, т.е. набор $X = (x_1, ..., x_n)$ из $n$ независимых наблюдений над $\xi$. Число $n$ называют объемом выборки. В процессе теоретических исследований принято смотреть на $x_1, ..., x_n$ как на независимые случайные величины, каждая из которых одинаково распределена с наблюдаемой случайной величиной $\xi$. Дискретная случайная величина $\xi^*$ с рядом распределения

| $\xi^*$ | $x_1$ | ... | $x_n$ |
|---|---|---|---|
| | $1/n$ | ... | $1/n$ |

называется эмпирическим аналогом $\xi$.

Если выражаться более точно, при построении эмпирического аналога за каждым из выборочных значений $x_i$ закрепляется вероятность $1/n$. Поэтому в приведенной выше таблице в первой строке каждое значение $x_i$ встречается лишь один раз, а вероятность ему соответствует $n_i/n$, где $n_i$ – повторность, количество раз, которое это значение встречается в выборке.

Все характеристики эмпирического аналога носят название эмпирических или выборочных.Например,

$$\mathbf{M}\xi^* = \frac{1}{n}\sum_{j=1}^{n} x_j \equiv \bar{X} \quad -$$

выборочное математическое ожидание (среднее),

$$\mathbf{D}\xi^* = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{X})^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 - (\bar{X})^2 \equiv S^2 -$$

выборочная дисперсия и т.п.

Функция распределения эмпирического аналога носит название эмпирической функции распределения и обозначается $F_n^*$. Любая измеримая функция выборки называется статистикой. Если мы упорядочим элементы выборки $X$, получим так называемый вариационный ряд:

$$X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}.$$

$X_{(k)}$ – $k$-й по порядку выборочный элемент. Его часто называют $k$-й порядковой статистикой. Название оправдывается тем, что порядковая статистика есть функция выборки в целом.

Наиболее вероятное (чаще всего встречающееся) значение $x_j$ называют модой, а то числовое значение, которое делит вариационный ряд на две равные по длине половины, медианой выборки. При этом, если выборка имеет нечетный объем, то медиана совпадает с одним из выборочных значений, а если четный, то за ее значение принимают середину интервала между соответствующими выборочными значениями. Заметим, наконец, что любая статистика (и медиана, в частности) является случайной величиной.

**Пример 1.** *Найдите математическое ожидание и дисперсию эмпирической функции распределения, считая теоретическую функцию распределения $F$ известной.*

**Решение**. Для начала припомним формулу, связывающую функцию распределения и ряд распределения дискретной величины, имеющую в нашем случае вид

$$F_n^*(x) \;=\; \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{(-\infty,x)}(x_j),$$

откуда

$$\mathbf{M}F_n^*(x) \;=\; \tfrac{1}{n}\sum_{j=1}^{n} P(x_j \in (-\infty, x)) \;=$$
$$=\; \tfrac{1}{n}\sum_{j=1}^{n} P(\xi < x) \;=\; F(x).$$

Вновь используя то же представление, запишем

$$\mathbf{M}(F_n^*(x))^2 \;=\; \tfrac{1}{n^2}\mathbf{M}\left(\sum_{j=1}^{n} \mathbf{1}_{(-\infty,x)}(x_j)\right)^2 \;=$$
$$= \tfrac{1}{n^2}\sum_{j=1}^{n} \textstyle\sum_{i=1}^{n} \mathbf{M}\mathbf{1}_{\{x_j<x\}}\mathbf{1}_{\{x_i<x\}}.$$

Отметим далее, что

$$(i \neq j) \;\Rightarrow\; \mathbf{M}\mathbf{1}_{\{x_j<x\}}\mathbf{1}_{\{x_i<x\}} = \mathbf{M}\mathbf{1}_{\{x_j<x\}}\mathbf{M}\mathbf{1}_{\{x_i<x\}} \;=\; F^2(x),$$

если $i = j$, то такое математическое ожидание равно $F(x)$. Таким образом, выписанная двойная сумма состоит из $n$ слагаемых $F(x)$ и $n(n-1)$ слагаемых $F^2(x)$. Окончательно, получаем

$$\mathbf{M}(F_n^*(x))^2 \;=\; \frac{1}{n}F(x) + \frac{n-1}{n}F^2(x).$$

Осталось записать

$$\mathbf{D}F_n^*(x) \;=\; \mathbf{M}(F_n^*(x))^2 - (\mathbf{M}F_n^*(x))^2 \;=\; \frac{F(x)(1-F(x))}{n}.$$

**Пример 2.** *Найдите распределение $X_{(1)}$ для выборки из распределения с функцией*

$$F(x) \;=\; \begin{cases} 1-e^{-x}, & x > 0, \\ 0 & , x \le 0. \end{cases}$$

**Решение**. Заметим, что

$$P(X_{(1)} \ge t) \;=\; P(x_1 \ge t, x_2 \ge t, ..., x_n \ge t),$$

откуда, в силу независимости $x_1, ..., x_n$ и их одинаковой распределенности с наблюдаемой случайной величиной $\xi$ следует

$$P(X_{(1)} \geq t) = P^n(x_1 \geq t) = (1 - F(x))^n = e^{-nx}$$

при $x > 0$. Таким образом, функция распределения $X_{(1)}$ имеет вид

$$G(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-nx}, & x > 0. \end{cases}$$

## ЗАДАЧИ К ЗАНЯТИЮ

1. В отделе технического контроля завода «Фрезер» была измерена глубина паза 120 плашек (в мм), и результаты приведены в таблице

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2,4 | 2,4 | 2,5 | 2,4 | 2,7 | 2,7 | 2,2 | 2,6 | 2,4 | 2,6 | 2,7 | 2,6 |
| 2,3 | 2,6 | 2,5 | 2,3 | 2,5 | 2,4 | 2,7 | 2,5 | 2,4 | 2,5 | 2,5 | 2,4 |
| 2,4 | 2,5 | 2,4 | 2,7 | 2,4 | 2,4 | 2,5 | 2,4 | 2,6 | 2,3 | 2,5 | 2,6 |
| 2,6 | 2,6 | 2,4 | 2,4 | 2,4 | 2,4 | 2,5 | 2,2 | 2,4 | 2,8 | 2,8 | 2,6 |
| 2,4 | 2,4 | 2,4 | 2,3 | 2,2 | 2,5 | 2,4 | 2,4 | 2,5 | 2,4 | 2,5 | 2,2 |
| 2,3 | 2,6 | 2,6 | 2,5 | 2,5 | 2,5 | 2,7 | 2,5 | 2,7 | 2,1 | 2,5 | 2,6 |
| 2,4 | 2,4 | 2,5 | 2,6 | 2,6 | 2,7 | 2,7 | 2,6 | 2,5 | 2,3 | 2,5 | 2,4 |
| 2,6 | 2,4 | 2,3 | 2,5 | 2,2 | 2,6 | 2,5 | 2,5 | 2,5 | 2,3 | 2,8 | 2,7 |
| 2,4 | 2,6 | 2,4 | 2,4 | 2,5 | 2,1 | 2,6 | 2,7 | 2,7 | 2,6 | 2,4 | 2,4 |
| 2,2 | 2,5 | 2,5 | 2,3 | 2,5 | 2,5 | 2,3 | 2,4 | 2,7 | 2,5 | 2,3 | 2,5 |

Постройте гистограмму, полигон, эмпирическую функцию распределения. Найдите моду, медиану, выборочные математическое ожидание и дисперсию глубины паза.

2. В таблице приведены данные по количеству пациентов, обратившихся в городской травмопункт в каждый из дней недели.

| $X$ | пн | вт | ср | чт | пт | сб | вс |
|---|---|---|---|---|---|---|---|
| | 37 | 53 | 35 | 27 | 30 | 44 | 28 |

Постройте гистограмму, полигон, эмпирическую функцию распределения. Найдите моду, медиану, выборочные математическое ожидание и дисперсию а) считая, что с каждым пациентом связано наблюдение, результатом которого является день его посещения; б) считая наблюдаемой величиной количество пациентов, а дни недели – повторениями эксперимента. Какая проблема возникает в задаче а )? Предложите несколько вариантов ее решения. Подумайте о степени полезности гистограммы и полигона в задаче б).

3. Функция $F(x)$ задана соотношением

$$F(x) = \begin{cases} 0, & x \le 0, \\ 0{,}5, & 0 < x \le 1, \\ 1, & x > 1. \end{cases}$$

Постройте хотя бы два примера выборок разного объема, для которых эта функция является эмпирической функцией распределения. Если объем выборки $n$ равен 4, то сколько вариантов таких выборок может быть построено? А если $n = 5$?

4. Функция $F(x)$ задана соотношением

$$F(x) = \begin{cases} 0, & x \le 0; \\ 0{,}3, & 0 < x \le 1; \\ 0{,}9, & 2 < x \le 3; \\ 1, & x > 3. \end{cases}$$

Постройте хотя бы два примера выборок разного объема, для которых эта функция является эмпирической функцией распределения. Какого объема должна быть выборка, для которой такая функция является эмпирической функцией распределения? Сколько можно предложить вариантов такой выборки объема 10?

5. Какова вероятность того, что эмпирическая функция выборки объема $n$ из абсолютно непрерывного распределения будет иметь скачок величины не меньшей, чем $2/n$?

6. Как может выглядеть график эмпирической функции распределения для выборки из биномиального распределения $B_{n,p}$ ?

7. Пусть $F_n^*(x)$ – эмпирическая функция распределения, $a > 0$. Будет ли эмпирической функцией распределения для некоторой выборки функция $G(x) = F_n^*(ax)$? $H(x) = aF_n^*(x)$? Если возможен ответ да, в обоих случаях укажите все допустимые для этого $a$.

8. Пусть заданы числа $a, b$ и эмпирическая функция распределения $F_n^*$ строится по выборке с известной теоретической функцией распределения $F$. Чему тогда равна вероятность $P(F_n^*(a) < F_n^*(b))$?

9. Пусть выборка построена из наблюдений над случайной величиной с функцией распределения $F$. Докажите, что

$$P\left(F_n^*(x) = \frac{k}{n}\right) \;=\; C_n^k F^k(x)(1 - F(x))^{n-k},\; k = 0, ..., n.$$

10. В условиях предыдущей задачи найдите вероятности

$$P(X_{(k)} < y), \quad P(X_{(k)} < y, X_{(k+1)} \geq y).$$

11. Пусть $X$ – выборка объема $n$ из распределения с непрерывной функцией $F$. Какому распределению соответствует выборка $(y_1, ..., y_n)$, где $y_i = F(x_i),\; i = 1, ..., n$?

12. Пусть $X$ имеет равномерное распределение на интервале $[0, \theta]$. Найдите плотность распределения $k$-й порядковой статистики.

13. Пусть задано натуральное число $N$, и $P(\xi = k) = p_k,\; k = 0, ..., N$. Считая, что $p_0 + ... + p_N = 1$, найдите распределение $k$-й порядковой статистики для выборки из этого распределения.

14. Пусть выборка $X$ объема 3 взята из равномерного распределения на интервале $[0, 1]$. Оцените величину

$$Q \;=\; \sup_t |F_3^*(t) - F(t)|$$

неслучайной константой. Вычислите $\mathbf{M}Q$.

## 2. Оценки и их простейшие свойства

Основной задачей математической статистики является восстановление распределения случайной величины по результатам наблюдений. Если никакой априорной информации при этом нет, то первым шагом в этом направлении является построение эмпирической функции распределения. Но иногда относительно наблюдаемой случайной величины $\xi$ заранее известно, что ее распределение относится к одному из известных параметрических семейств распределений, например, является нормальным. В этом случае бывает достаточно оценить параметр этого распределения – одномерный или многомерный.

Пусть $X$ – выборка из одного из распределений $P_\theta$, где $\theta$ – какой-то элемент множества $\Theta$, и вместо него мы используем некоторую статистику $S(X)$. Тогда $\theta^* = S(X)$ называется оценкой неизвестного параметра $\theta$. Если

$$\mathbf{M}\theta^* \ = \ \theta,$$

то $\theta^*$ называется несмещенной оценкой $\theta$. Если это не так, то мы говорим о смещенной оценке. Наиболее известной смещенной оценкой является оценка дисперсии $S^2$: для выборки объема $n$

$$\mathbf{M}S^2 \ = \ \frac{n-1}{n}\mathbf{D}\xi \ < \ \mathbf{D}\xi.$$

Если $\theta^*$ сходится к истинному значению $\theta$ в каком-либо смысле при $n \to \infty$, то $\theta^*$ называется состоятельной.

Одним из простейших методов получения оценок является следующий. Пусть $\theta = G(F)$, где $G$ – функционал, задающий последовательность операций над функцией распределения $F$ для извлечения параметра $\theta$. Подставим в него вместо $F$ эмпирическую функцию распределения. Полученный в результате этой процедуры результат

$$\theta^* \;=\; G(F_n^*)$$

называют оценкой подстановки для $\theta$.

Еще один метод получения оценок состоит в том, что можно приравнять соответствующие друг другу эмпирические и теоретические характеристики $\xi$. Полученные уравнения решаются относительно неизвестных значений параметров. Такой метод носит название метода моментов, поскольку в качестве рассматриваемых характеристик чаще всего выступают моменты распределения $\xi$ (например, математическое ожидание или дисперсия, $\mathbf{M}\xi^2$, $\mathbf{M}\xi^3$ и т.д.).

Пусть задано некоторое расстояние $d$ между функциями распределения. Оценкой $\theta$ по методу минимума расстояния называют такое значение $\theta^*$, что

$$d(F_{\theta^*},\, F_n^*) \;=\; \min\{d(F_\theta,\, F_n^*),\ \theta \in \Theta\}.$$

**Пример 3.** *Пусть выборка $X$ берется из биномиального распределения $B_{n,p}$. Найдите оценки для $n$ и $p$ по методу моментов. Проверьте их состоятельность.*

**Решение**. Известно, что

$$\mathbf{M}x_1 \;=\; np, \quad \mathbf{D}x_1 \;=\; np(1-p).$$

Поскольку параметров два, построим систему метода моментов, используя математическое ожидание и дисперсию.

$$\begin{cases} \bar{X} \;=\; np; \\ S^2 \;=\; np(1-p), \end{cases}$$

откуда

$$p^* \;=\; 1 - \frac{S^2}{\bar{X}}, \quad n^* \;=\; \frac{(\bar{X})^2}{\bar{X} - S^2} -$$

нужные нам оценки по методу моментов. В силу усиленного закона больших чисел и теоремы непрерывности

$$\bar{X} \xrightarrow{\text{п.н.}} np, \quad S \xrightarrow{\text{п.н.}} npq.$$

Отсюда

$$\frac{S^2}{\bar{X}} \xrightarrow{\text{п.н.}} q \text{ и } p^* \xrightarrow{\text{п.н.}} 1 - q \ = \ p,$$

т.е. $p^*$ сильно состоятельна. Аналогично доказывается и сильная состоятельность оценки $n^*$.

**Пример 4.** *Пусть для параметр $\theta$ равномерного на $[0, \theta]$ распределения выбрана оценка $\theta^* = X_{(n)}$. Исследуйте ее свойства.*

**Решение**. Найдем распределение $\theta^*$.

$$P(\theta^* < x) = P_\theta(x_1 < x, ..., x_n < x) = \\ = \prod_{j=1}^n P(x_j < x) = P^n(x_1 < x).$$

Поскольку распределение $x_1$ равномерно, то при $x \in [0, \theta]$ $P(x_1 < x) = x/\theta$. Тем самым,

$$P(\theta^* < x) = (x/\theta)^n \quad (0 \leq x \leq \theta),$$

и плотность распределения $\theta^*$ равна

$$p_{\theta^*}(x) \ = \ \frac{nx^{n-1}}{\theta^n} \quad (0 \leq x \leq \theta).$$

Переходя к пределу при $n \to \infty$, видим, что функция распределения $\theta^*$ поточечно сходится к индикатору интервала $(\theta, +\infty)$, который, как известно, является функцией распределения константы, равной $\theta$. Таким образом, имеется по меньшей мере, слабая сходимость $\theta^*$ к $\theta$, и оценка состоятельна.

Далее,

$$\mathbf{M}\theta^* = \int\limits_0^\theta \frac{nx^n}{\theta^n}\, dx = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n\theta}{n+1},$$

т.е. оценка смещенная. Для того, чтобы получить несмещенную оценку, надо взять

$$\theta^{**} = (n+1)\theta/n = (1+1/n)X_{(n)}.$$

**Пример 5.** *Постройте оценку параметра p распределения Бернулли по минимуму расстояния*

$$d(F,G) \;=\; \sup_x |F(x) - G(x)|.$$

*Проверьте ее свойства.*

**Решение**. Легко увидеть, что теоретическая и эмпирическая функции распределения в этом случае имеют, соответственно, вид

$$F(x) = \begin{cases} 0, & x \le 0; \\ 1-p, & 0 < x \le 1; \\ 1, & x > 1, \end{cases} \qquad F_n^*(x) = \begin{cases} 0, & x \le 0; \\ 1-\bar{X}, & 0 < x \le 1; \\ 1, & x > 1, \end{cases}$$

откуда $d(F, F_n^*) = |p - \bar{X}|$, и искомая оценка $p^* = \bar{X}$. Из курса лекций известно, что эта оценка является несмещенной и сильно состоятельной.

## ЗАДАЧИ К ЗАНЯТИЮ

15. Пусть наблюдаемая случайная величина абсолютно непрерывна и $\theta$ – точка максимума ее плотности. Предложите оценку этого параметра методом подстановки. В каком случае можно заведомо утверждать, что полученная оценка будет состоятельной?

16. Квантилью распределения $F$ уровня $\alpha$ называется решение неравенства $F(x) \le \alpha \le F(x+0)$. Найдите оценки подстановки квантили уровня $1/2$ и уровня $1/3$. Вычислите значения этих оценок по данным задач 1 и 2.

17. Постройте теоретическую характеристику, для которой медиана выборки является ее оценкой подстановки. Как следует назвать эту характеристику?

18. Постройте оценки по методу моментов для параметров следующих распределений а) $N(a,1)$; б) $N(0,\sigma^2)$; в) $U_{[0,\theta]}$; г) распределения Бернулли; д) экспоненциального распределения; е)$U_{[\theta-1,\theta+1]}$; ж) $U_{[-\theta,\,\theta]}$.

19. Оцените с помощью метода моментов параметр распределения с плотностью

$$f(x) \;=\; \frac{\lambda}{2}e^{-\lambda|x|}.$$

Проверьте состоятельность и несмещенность этой оценки.

20. Постройте оценки по методу моментов для векторных параметров распределений

$$\text{а) } U_{[a,b]}; \qquad \text{б) } \Gamma_{\alpha,\lambda}.$$

Исследуйте (по возможности) свойства полученных оценок. Если какое-то исследование не получается произвести строго, выскажите аргументированную гипотезу о его возможном результате.

21. Оцените параметры дискретного распределения

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

по методу моментов.

22. Пусть дана выборка из распределения с параметром $\theta$, смысл которого в том, что теоретическая функция распределения имеет вид

$$F_\theta(x) \;=\; \theta F_1(x)+(1-\theta)F_2(x),$$

$0<\theta<1$, функции распределения $F_1, F_2$ известны (смесь двух известных распределений с неизвестными весами). Постройте оценку неизвестного параметра с помощью метода моментов.

23. Можно ли построить оценку неизвестного математического ожидания нормального распределения по ме-

тоду моментов, используя равенство вторых теоретического и эмпирического моментов, если дополнительно известно, что это математическое ожидание положительно? А если ничего не известно о его знаке?

24. Используя равенство $k$-х теоретического и эмпирического моментов, постройте счетные семейства оценок параметров следующих распределений

а) $N(a,1),\ a \geq 0$; б) $U_{[0,\theta]}$; в) $E_\alpha$.

25. Для оценивания параметра $\alpha$ распределения с плотностью

$$f(x) \ = \ e^{\alpha - x}, \quad (x \geq \alpha)$$

предлагается оценка $\alpha^* = X_{(1)}$. Проверьте ее состоятельность и несмещенность.

26. Пусть $X$ – выборка из $U_{[0,\theta]}$. Являются ли несмещенными оценки

$$\theta_1^* = 2\bar{X}, \quad \theta_2^* = X_{(1)} + X_{(n)}?$$

27. Установите состоятельность оценки $\theta^* = X_{(n)}$ для параметра $U_{[0,\theta]}$ с помощью неравенств Чебышева.

28. Пусть $X$ – выборка из распределения $B_{1,\sqrt{p}}$. Является ли оценка $p^* = (\bar{X})^2$ несмещенной? Состоятельной?

29. Пусть требуется оценить функцию $\tau(p) = 1/p$ по выборке из распределения Бернулли. Покажите, что в данном случае не существует несмещенной оценки.

30. В условиях задачи 22 постройте оценку параметра, минимизирующую расстояние

$$d(F,G) \ = \ \int\limits_{-\infty}^{+\infty} (F(x) - G(x))^2 dF(x),$$

выбирая в качестве $F$ эмпирическую функцию распределения.

31. Будет ли оценка $\theta^* = \bar{X}$ состоятельной оценкой для параметра распределения Коши с плотностью

$$p_\theta(x) \;=\; \frac{1}{\pi(1+(x-\theta)^2)}?$$

## 3. Метод максимального правдоподобия

Пусть распределение $P_\theta$, из которого берется выборка, либо абсолютно непрерывно, либо дискретно. В первом случае через $f_\theta$ будем обозначать плотность этого распределения, во втором – вероятности отдельных значений (т.е. его плотность относительно считающей меры). Функцией правдоподобия называется

$$L(X,\theta) \;=\; \prod_{j=1}^{n} f_\theta(x_j).$$

Оценка по методу максимума правдоподобия $\hat{\theta}$ – это такое значение $\theta$, при котором функция правдоподобия достигает своего максимума. При этом вместо поиск максимума $L$ нередко ищут максимум

$$l(X,\theta) \;=\; \ln L(X,\theta) \;=\; \sum_{j=1}^{n} \ln f_\theta(x_j).$$

Последняя функция называется логарифмической функцией правдоподобия.

Метод максимального правдоподобия, как правило, дает наилучшие из возможных оценок.

### ЗАДАЧИ К ЗАНЯТИЮ

32. Пусть $f_\theta$ дифференцируема по $\theta$ в каждой точке два раза. Достаточно ли для достижения максимума функции правдоподобия в точке потребовать, чтобы в этой точке ее производная равнялась бы 0?

33. Пусть распределение наблюдаемой случайной величины имеет плотность $p_\theta(x) = f(x-\theta)$, где известная функция $f$ имеет единственный максимум в точке 0. Оцените параметр $\theta$ методом максимального правдоподобия по одному наблюдению.

34. В условиях предыдущей задачи покажите, что оценка для $\theta$, построенная методом максимального правдоподобия по $n$ наблюдениям, лежит в интервале $[X_{(1)},\ X_{(n)}]$.

35. Найдите оценки максимального правдоподобия для параметров следующих распределений:

а) $N(a,1)$; б) $N(0,\sigma^2)$; в) параметра $p$ биномиального распределения (считая второй параметр $m$ известным); г) $\Pi_\lambda$; д) экспоненциального распределения; е) $U_{[-\theta,0]}$; ж) $U_{[-\theta,\theta]}$.

36. Постройте методом максимального правдоподобия оценки векторных параметров $U_{[a,b]}$; $\Gamma_{\alpha,\lambda}$.

37. Распределение Кэптейна обладает плотностью

$$p_\theta(x) \ = \ \frac{g'(x)}{\sqrt{2\pi}}\exp\{-\frac{1}{2}(g(x)-\theta)^2\},$$

где $g(x)$ – известная дифференцируемая функция. Оцените параметр $\theta$ методом максимального правдоподобия.

38. Найдите оценку максимального правдоподобия распределения Лапласа с плотностью

$$p_\mu(x) \ = \ \frac{1}{2}e^{-|x-\mu|}.$$

39. Найдите оценки максимального правдоподобия для распределений со следующими плотностями:

$$p_\theta(x) = \begin{cases} \theta x^\theta, & x\in[0,\ 1], \\ 0, & x\notin[0,\ 1]; \end{cases} \qquad p_\theta(x) = \begin{cases} \frac{2x}{\theta^2}, & x\in[0,\ \theta], \\ 0, & x\notin[0,\ \theta]; \end{cases}$$

$$p_\theta(x) = \begin{cases} \frac{\theta}{\sqrt{2\pi x^3}}e^{-\theta^2/(2x)}, & x>0, \\ 0, & x\le 0; \end{cases}$$

$$p_\theta(x) = \begin{cases} e^{\theta - x}, & x \geq \theta, \\ 0, & x < \theta; \end{cases} \qquad p_{\alpha,\mu}(x) = \frac{1}{2\alpha} \exp\left\{ -\frac{|x-\mu|}{\alpha} \right\}.$$

# 4. Сравнение оценок. Асимптотическая нормальность

Оценка $\theta^*$ для $\theta$ называется асимптотически нормальной с коэффициентом рассеивания $\sigma^2$, если для некоторой $\eta$, имеющей распределение $N(0, \sigma^2)$,

$$(\theta^* - \theta)\sqrt{n} \xrightarrow{d} \eta, \ n \to \infty.$$

Это определение можно также переписать в виде

$$\frac{\theta^* - \theta}{\sigma} \cdot \sqrt{n} \xrightarrow{d} \eta_1, \ n \to \infty,$$

где $\eta_1$ является стандартной нормальной случайной величиной. Для определения коэффициента рассеивания $\sigma^2$ обычно используется следующая методика. Пусть

$$\theta^* = h\left( \frac{1}{n} \sum_{j=1}^{n} g(x_j) \right),$$

тогда при выполнении условий $\mathbf{M}g^2(\xi) < \infty$, $h'(a) \neq 0$, где $a = \mathbf{M}g(x_1)$, эта оценка для $\theta = h(g(a))$ асимптотически нормальна, и коэффициент ее рассеивания равен

$$\sigma^2 = (h'(a))^2 \int_{-\infty}^{+\infty} (g(x) - a)^2 dF_\theta(x) = (h'(a))^2 \mathbf{D}g(\xi), \quad (1)$$

Для сравнения между собой двух оценок одного и того же параметра существует два основных подхода.

1). Среднеквадратический. Та из двух оценок лучше, для которой $d(\theta) = \mathbf{M}(\theta^* - \theta)^2$ меньше.

2). Асимптотический. Та из двух асимптотически нормальных оценок лучше, у которой коэффициент рассеивания меньше.

Оценка называется эффективной в классе, если она не хуже всех остальных оценок этого класса. Обычно в качестве классов здесь выбирают множества оценок с фиксированным смещением, имеющих конечные вторые моменты:

$$K_b \;=\; \{\theta^* |\, \mathbf{M}(\theta^*)^2 < \infty,\; \mathbf{M}\theta^* = \theta + b\}.$$

Оказывается, если оценка максимального правдоподобия не смещена, то она, как правило, является эффективной в классе $K_0$. Такую оценку называют эффективной. Для корректной проверки эффективности оценки максимального правдоподобия достаточно убедится, что она не смещена, и для нее в неравенстве Рао - Крамера достигается равенство:

$$\mathbf{D}\theta^* \;=\; \frac{1}{nI(\theta)},$$

где в знаменателе дроби фигурирует информация Фишера, которая вычисляется по формуле

$$I(\theta) \;=\; \mathbf{M}\left(ln'(f_\theta(\xi)\right)^2.$$

Вместо проверки равенства в неравенстве Рао-Крамера можно проверить условие экспоненциальности семейства распределений относительно оценки $\theta^*$ (см. задачу 52).

**Пример 6.** *Пусть $X$ – выборка из $U_{[0,\theta]}$. Какая из двух оценок: $\theta_1^* = 2\bar{X}$ или $\theta_2^* = (1 + 1/n)\, X_{(n)}$ лучше в среднеквадратическом смысле?*

**Решение**. Как известно, обе эти оценки не смещены, откуда немедленно следует, что

$$\mathbf{M}(\theta_i^* - \theta)^2 \;=\; \mathbf{D}\theta_i^*,\; i = 1, 2.$$

Учитывая это обстоятельство, получаем

$$d_1(\theta) = \mathbf{D}\theta_1^* = 4\mathbf{D}\bar{X} = \tfrac{1}{n^2}(4\sum_{j=1}^{n}\mathbf{D}x_j) =$$
$$= (4n\mathbf{D}x_1)/n^2 = \theta^2/3n.$$

С другой стороны,

$$d_2(\theta) = \mathbf{D}\theta_2^* = (1+\tfrac{1}{n})^2\mathbf{M}X_{(n)}^2 - ((1+\tfrac{1}{n})\mathbf{M}X_{(n)})^2 =$$
$$= (1+\tfrac{1}{n})^2\mathbf{M}X_{(n)}^2 - \theta^2.$$

Но плотность распределения $X_{(n)}$ была уже вычислена в примере 4, поэтому можно записать

$$\mathbf{M}X_{(n)}^2 = \int\limits_0^\theta x^2\frac{nx^{n-1}}{\theta^n}\,dx = \frac{n}{\theta^n}\frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2,$$

и, окончательно,

$$d_2(\theta) \ = \ \frac{\theta^2}{n(n+2)}.$$

Сравнив $d_1(\theta)$ и $d_2(\theta)$, видим, что $\theta_1^*$ может быть не хуже лишь при $n=1$. Поэтому предпочесть следует $\theta_2^*$.

**Пример 7.** *Пусть $X$ имеет пуассоновское распределение с параметром $\lambda$. Рассмотрим две оценки для $\lambda$*

$$\lambda_1^* = \bar{X}, \quad \lambda_2^* = \sqrt{\overline{X^2} + \frac{1}{4}} - \frac{1}{2}.$$

*Какая из них лучше в асимптотическом смысле?*

**Решение**. В формуле (1) для $\lambda_1^*$ следует взять

$$g(x) = h(x) = x, \quad a = Mx_1 = \lambda, \quad h'(a) = 1,$$

откуда $\sigma_1^2 = Dx_1 = \lambda$. Аналогично, для $\lambda_2^*$ положим

$$g(x) = x^2, \quad h(x) = \sqrt{x+1/4} - 1/2.$$

Здесь
$$a = \mathbf{M}x_1^2 = \lambda^2 + \lambda, \quad h'(a) = \frac{1}{2\lambda+1},$$
и, таким образом,
$$\sigma_2^2 \;=\; \frac{\mathbf{D}x_1^2}{(2\lambda+1)^2} = \frac{\mathbf{M}x_1^4 - (\mathbf{M}x_1^2)^2}{(2\lambda+1)^2}.$$

В нашей задаче $x_1$ имеет распределение $\Pi_\lambda$, поэтому
$$\mathbf{M}x_1^4 = \lambda^4 + 6\lambda^3 + 6\lambda^2 + \lambda;\ (\mathbf{M}x_1^2)^2 = \left(\lambda^2+\lambda\right)^2,$$
и, окончательно,
$$\sigma_2^2 = \frac{4\lambda^3 + 5\lambda^2 + \lambda}{(2\lambda+1)^2} = \lambda\left(1 + \frac{\lambda}{(2\lambda+1)^2}\right).$$
Итак, $\sigma_1^2 < \sigma_2^2$, и оценку $\lambda_1^*$ следует предпочесть.

## ЗАДАЧИ К ЗАНЯТИЮ

40. Докажите, что любая асимптотически нормальная оценка является состоятельной.

41. Пусть $\theta^*$ – асимптотически нормальная оценка $\theta$. Будет ли $|\theta^*|$ асимптотически нормальной оценкой $|\theta|$?

42. Пусть параметр $\lambda$ распределения $\Gamma_{\alpha,\lambda}$ задан. Найдите оценку для $\alpha$ по методу моментов. Докажите ее асимптотическую нормальность и найдите коэффициент рассеивания.

43. Пусть $X$ – выборка из экспоненциального распределения с параметром $\alpha$. Докажите, что при любом натуральном $k$ оценка
$$\alpha_k^* \;=\; \left(\frac{k!}{\overline{X^k}}\right)^{1/k}$$
является асимптотически нормальной. Вычислите ее коэффициент рассеивания. Какую из этих оценок следует признать лучшей в асимптотическом смысле?

44. Выясните, которая из оценок

$$\theta_k^* = \left((k+1)\overline{X^k}\right)^{1/k}, \; k \in N$$

для параметра равномерного распределения $U_{[0,\theta]}$ является лучшей в асимптотическом смысле. Найдите $\lim_{k\to\infty} \theta_k^*$.

45. Сравните оценки

$$\theta_1^* = X_{(1)} \text{ и } \theta_2^* = 2\overline{X} - 1$$

параметра распределения $U_{[\theta,1]}$ в среднеквадратическом смысле.

46. Пусть $X$ – выборка из $N(\alpha, 1)$, $\alpha > 0$. Сравните

$$\alpha_1^* = \bar{X} \text{ и } \alpha_2^* = \max\{0, \bar{X}\}$$

в среднеквадратическом смысле.

47. Пусть параметр $\sigma > 0$ распределения с плотностью

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{(\ln x - \theta)^2}{2\sigma^2}\right\}, \; x > 0$$

задан. Покажите, что оценка $\theta^* = \overline{\ln X}$ является эффективной.

48. Проверьте эффективность оценки

$$\alpha^* = \frac{n-1}{n\bar{X}}$$

для выборки из распределения $\Gamma_{\alpha,1}$.

49. Проверьте эффективность в классе оценок с фиксированным смещением оценки $\theta^* = (1 + 1/n)X_{(n)}$ для параметра равномерного на $[0, \theta]$ распределения.

50. По выборке объема 1 из распределения $U_{[\theta,\theta+1]}$ строится оценка $\theta^* = [x_1]$ (целая часть имеющегося наблюдения). Вычислите дисперсию и смещение этой оценки. В каком классе такая оценка будет эффективной?

51. Пусть $\theta^*_\alpha$ – эффективная оценка в классе оценок со смещением $\alpha\theta$. Постройте эффективную (в частности, несмещенную) оценку.

52. Параметрическое семейство распределений называется экспоненциальным относительно статистики $T(X)$, если для функции правдоподобия имеет место представление

$$L(X,\theta) \ = \ h(X)\exp\{A(\theta)T(X)+B(\theta)\},$$

где измеримые функции $h, A, T, B$ зависят лишь от своих аргументов и, возможно, от объема выборки. Будут ли экспоненциальными семейства распределений а) $N(a,1)$; б) пуассоновские распределения; в) распределения Бернулли; в) $N(a,\sigma^2)$; г) $U_{[a,b]}$; д) $\Gamma_{\alpha,\lambda}$? (в трех последних вопросах параметр двумерный).

## 5. Доверительные интервалы

Понятие доверительного интервала оказывается полезным тогда, когда требуется верно оценить неизвестный параметр с высокой вероятностью. Пусть $\theta^\pm$ – статистики, $\varepsilon > 0$ – малое число. Интервал $[\theta^-, \theta^+]$ называют (точным) доверительным интервалом для $\theta$ уровня доверия $1-\varepsilon$, если

$$P(\theta^- \ \le \ \theta \ \le \ \theta^+) \ = \ 1-\varepsilon.$$

Соответствующий интервал называют асимптотическим доверительным, если

$$\lim_{n\to\infty} P(\theta^- \le \theta \le \theta^+) \ \ge \ 1-\varepsilon.$$

Задача построения точных доверительных интервалов, как правило, требует в каждом случае поиска распределений специально подобранных статистик. Часто при решении подобных задач помогают теоремы Фишера о распределениях некоторых статистик в нормальной совокупности. Пусть $X$ – выборка объема $n$ из $N(a, \sigma^2)$.

Теорема 1. $(\overline{X} - a)/\sigma \cdot \sqrt{n}$ *имеет стандартное нормальное распределение.*

Теорема 2. $\frac{1}{\sigma^2} \sum_{j=1}^{n} (x_j - a)^2$ *имеет распределение* $\chi^2_n$.

Теорема 3. $nS^2/\sigma^2$ *имеет распределение* $\chi^2_{n-1}$ *и не зависит от* $\overline{X}$.

Теорема 4. $(\overline{X} - a)/S \cdot \sqrt{n-1}$ *имеет распределение Стьюдента с* $n - 1$ *степенью свободы.*

Построение асимптотических доверительных интервалов часто может быть реализовано с помощью процедуры, описанной в следующем примере.

**Пример 8.** *Пусть* $\theta^*$ *– асимптотически нормальная оценка* $\theta$ *с известным коэффициентом рассеивания* $\sigma^2$. *Постройте асимптотический доверительный интервал для* $\theta$ *заданного уровня* $1 - \varepsilon$.

**Решение**. В силу определения сходимости по распределению, того, что функция стандартного нормального распределения $\Phi$ является непрерывной и свойств нормальных распределений для произвольного $x$ имеет место соотношение

$$P(|(\theta^* - \theta)|\sqrt{n}/\sigma \le x) \ \to P(|\eta_1| < x) \ = \\ = \ \Phi(x) - \Phi(-x) \ = \ 2\Phi(x) - 1, \ n \to \infty,$$

где $\eta_1$, как и выше, стандартная нормальная величина. Теперь, выбирая при заданном $\varepsilon$ число $x$ так, чтобы

$$2\Phi(x) - 1 \ = \ 1 - \varepsilon,$$

или, иначе,

$$\Phi(x) \ = 1 - \frac{\varepsilon}{2},$$

мы получим, что $|\theta^* - \theta|\sqrt{n} \le x\sigma$ с нужной вероятностью. Осталось решить это неравенство относительно $\theta$, имея, впрочем, в виду, что $\sigma$ также может зависеть от $\theta$.

## ЗАДАЧИ К ЗАНЯТИЮ

53. Найдите асимптотически нормальные оценки параметров, использующие статистику $\bar{X}$. На их основе постройте асимптотические доверительные интервалы для параметров следующих распределений а) биномиального с известным числом испытаний $m$; б) экспоненциального; в) пуассоновского.

54. Используя поочередно асимптотически нормальные оценки, основанные на статистиках $\bar{X}$, $\overline{X^2}$, постройте два асимптотических доверительных интервала для параметра распределения $U_{[0,\theta]}$ и выясните, который из них имеет меньшую длину.

55. С помощью асимптотически нормальных оценок

$$\alpha_1^* = \frac{1}{\bar{X}}, \quad \alpha_2^* = \sqrt{\frac{2}{\overline{X^2}}}$$

параметра экспоненциального распределения постройте два асимптотических доверительных интервала. Какой из них короче?

56. По выборке из нормального распределения $N(a, \sigma^2)$ с известным средним $a$ постройте асимптотический доверительный интервал для $\sigma$, используя ее асимптотически нормальную оценку $\sqrt{(\pi/2)} \cdot \overline{|X - a|}$. Сравните его со стандартным точным доверительным интервалом того же уровня.

57. Пусть дана выборка из равномерного распределения на $[0, \theta]$. Постройте асимптотический доверительный интервал для параметра а) при помощи асимптотически нормальной оценки $\theta^* = 2\bar{X}$; б) при помощи оценки максимального правдоподобия $\hat{\theta} = X_{(n)}$. *Указание.* В задаче

б) используйте то, что в этом случае

$$\lim_{n\to\infty} P\left(\frac{nX_{(k)}}{\theta} < y\right) \;\to\; \Gamma_{1,k}(y)$$

при любых фиксированных $k, y$.

58. Пусть дана выборка из распределения, у которого неизвестный параметр $\theta$ – математическое ожидание, причем дисперсия является известной непрерывной функцией $\sigma^2(\theta)$ от параметра. Постройте асимптотический доверительный интервал для $\theta$.

59. При наблюдении выборки из $N(a, \sigma^2)$ параметр $a$ считается точно известным. Найдя распределение статистики $|\bar{X} - a|$, постройте точный доверительный интервал для дисперсии. Сравните его со стандартным доверительным интервалом.

60. Пусть $X$ – выборка из $N(\theta, \theta^2)$. Постройте точный доверительный интервал для $\theta$.

61. Для параметра $\theta$ распределения $U_{[0,\theta]}$ постройте интервалы, в которых параметр содержится с вероятностью, не меньшей $1-\varepsilon$, поочередно применяя неравенство Чебышева к статистикам $\bar{X},\ X_{(n)}$. Сравните эти интервалы.

62. Дана выборка из $U_{[0,\theta]}$. Покажите, что в качестве интервала, в котором $\theta$ содержится с вероятностью, не меньшей, чем $1 - \varepsilon$, можно взять $[X_{(n)},\ X_{(n)}/\Psi]$, где $\Psi$ – решение уравнения

$$\Psi^{n-1}(n - \ (n-1)\Psi) \;=\; \varepsilon.$$

63. Путем изучения распределения статистики $X_{(1)}$ постройте точный доверительный интервал для параметра а) $U_{[\theta,\theta+1]}$; б) $U_{[\theta,\, 2\theta]}$.

# 6. Проверка статистических гипотез

Пусть имеются две несовместные гипотезы: $H_0$ – основная и $H_1$ – альтернативная. Статистический критерий это измеримое отображение множества всех выборок на двухэлементное множество $\delta : \mathcal{X} \to \{0, 1\}$. Значение критерия на имеющейся выборке показывает номер принимаемой на данной выборке гипотезы. Поскольку гипотезы всего две, то достаточно задать только множество выборок, на котором принимается одна из них. Поэтому критерий обычно отождествляют с его так называемым критическим множеством

$$S \ = \ \{X \in \mathcal{X} :\ \delta(X) = 1\},$$

на котором основная гипотеза отвергается. Вероятностью ошибки первого рода называется $\alpha(\delta) = P(S/H_0)$, вероятностью ошибки второго рода

$$\alpha'(\delta) \ = \ P(\delta(X) = 0/H_1) \ = \ 1 - P(S/H_1).$$

Близкие к этим термины: $\alpha(\delta)$ называют размером критерия, $\beta(\delta) = P(S/H_1)$ – мощностью критерия $\delta$.

При вычислении вероятностей ошибок условные вероятности часто упрощенно записывают с помощью индексов. Например, вместо $P(A/H_0)$ пишут $P_0(A)$.

Обычная схема статистических критериев использует сравнение ожидаемой с точки зрения справедливости основной гипотезы и наблюдаемой картин (выборок). Одним из самых распространенных критериев такого типа является критерий Пирсона $\chi^2$. По наблюдаемым $x_j,\ j = 1, ...n$ и ожидаемым (идеальным) значениям $z_j,\ j = 1...n$ вычисляют

$$\chi^2 \ = \ \sum_{j=1}^{n} \frac{(x_j - z_j)^2}{z_j}$$

и сравнивают с критической точкой распределения $\chi^2_{n-1}$. Если значение $\chi^2$ не превосходит критической точки, то гипотезу можно принять.

В случае, когда для вычисления ожидаемых значений нужно знание $m$ некоторых неизвестных параметров, рекомендуется оценить их методом максимального правдоподобия. Тогда решение принимается путем сравнения рассчитанного $\chi^2$ с критической точкой распределения $\chi^2_{n-m-1}$ (число степеней свободы уменьшается на количество оцениваемых параметров).

Гипотеза называется простой, если ей удовлетворяет единственное распределение. В случае, когда основная и альтернативная гипотезы являются простыми при выполнении некоторых достаточно широких условий регулярности наиболее мощный критерий задается с помощью критического множества

$$S = \left\{ X \in \mathcal{X} : \frac{L(X,1)}{L(X,0)} > t \right\},$$

где $t$ подбирается из условия $P_0(S) = \alpha$. Этот критерий носит название критерия Неймана - Пирсона. Если построенный наиболее мощный критерий для проверки фиксированной (простой) альтернативы оказывается не зависящим от вида этой альтернативы, то такой критерий называют равномерно наиболее мощным (РНМ).

## ЗАДАЧИ К ЗАНЯТИЮ

64. Постройте критерий, обладающий нулевыми вероятностями ошибок и первого, и второго рода для проверки гипотезы $H_0 = \{N(0,1)\}$ против простой альтернативы $H_1 = \{\Pi_2\}$.

65. При каких условиях можно добиться того, чтобы обе ошибки критерия имели нулевую вероятность?

66. Проверяется гипотез $H_0 = \{N(0,1)\}$ против альтернативы $H_1 = \{P(0 \leq \xi \leq 1) = 1\}$. Постройте критерий с нулевой вероятностью ошибки второго рода. Какой наименьшей вероятностью ошибки первого рода может обладать такой критерий?

67. Пусть выборка берется из $N(\alpha, 1)$. Для проверки гипотезы $H_0 = \{\alpha = 0\}$ против $H_1 = \{\alpha = 1\}$ используется критерий

$$S = \{X : X_{(n)} \geq 3\}.$$

Найдите вероятности обеих ошибок. Сформулируйте ограничения для возможности применения этого критерия, ориентируясь на найденные вероятности.

68. Для проверки $H_0 = \{U_{[0,1]}\}$ при альтернативе $H_1$, состоящей в том, что наблюдаемая величина имеет любое иное распределение, по выборке из трех наблюдений используется критерий Колмогорова: основная гипотеза отвергается, если

$$\sup_{t \in [0,1]} |F_n^*(t) - t| > \frac{1}{3}.$$

Опишите критическое множество в терминах порядковых статистик. Чему равен размер этого критерия?

69. Изучаются две выборки одинакового объема $n$ из распределений с абсолютно непрерывными функциями $F$ и $G$. Для проверки гипотезы $H_0 = \{F = G\}$ применяется следующий критерий знаков: гипотеза отклоняется, если число попарных разностей элементов выборок с одинаковыми номерами отличается от $n/2$ более, чем на заранее выбранное число $\delta > 0$. Найдите вероятность ошибки первого рода точно и оцените ее значение при больших $n$, используя нормальное приближение. Каким следует выбрать число $\delta$, чтобы вероятность ошибки первого рода не превосходила бы заданного $\varepsilon$?

70. Используя конструкцию доверительного интервала для соответствующего параметра, постройте критерий размера $\alpha$ для проверки простой гипотезы $H_0 = \{\theta = 1\}$, если выборка берется из а) $N(\theta, 1)$; б) $N(1, \theta)$; в) экспоненциального распределения $E_\theta$ ; г) распределения Бернулли с параметром $p = \theta/2$; д) $\Pi_\theta$.

71. По выборке из $N(0, \sigma^2)$ постройте РНМ-критерий для проверки $H_0 = \{\sigma^2 = \sigma_1^2\}$ при альтернативе $H_1 = \{\sigma^2 < \sigma_1^2\}$.

71. Постройте критерии Неймана -Пирсона для выписанных пар простых гипотез. Вычислите мощности этих критериев при больших объемах выборок. Будут ли построенные критерии равномерно наиболее мощными для каких-либо видов альтернатив?

а) $H_i = \{E_{\alpha_i}\},\ i = 0, 1$; б) $H_i = \{\Pi_{\lambda_i}\},\ i = 0, 1$; в) гипотезы о биномиальных распределениях с известным числом испытаний $m$ и параметрами $p_i, i = 1, 2$.

72. Постройте критерий Неймана -Пирсона для проверки гипотезы $\{N(a, 1)\}$ против альтернативы $\{N(a, 2)\}$, где параметр $a$ известен. Найдите его мощность.

73. Постройте критерий Неймана -Пирсона размера $\varepsilon$ для проверки $H_0 = \{U_{[0,1]}\}$ против $H_1 = \{E_1\}$. Чему равна вероятность ошибки второго рода этого критерия?

74. Постройте критерий Неймана -Пирсона для проверки гипотезы $\{N(a, 1)\}$ против альтернативы $\{N(b, 2)\}$ $(a \neq b)$.

75. Пусть $a, \sigma, b, \tau$ известны. Какова мощность наиболее мощного критерия для проверки простой гипотезы $H_0 = \{N(a, \sigma)\}$ против $H_1 = \{N(b, \tau)\}$ ?

76. По выборке объема 1 проверяется гипотеза $H_0 = \{U_{[0,1]}\}$ против альтернативы, в роли которой выступает

распределение с плотностью

$$f(t) = \begin{cases} 3/2, \ t \in [0, 1/2], \\ 1/2, \ t \in [1/2, 1], \\ 0, \quad t \notin [0, 1]. \end{cases}$$

Постройте наиболее мощный критерий размера 0,25.

77. Гипотеза о наличии телепатических способностей у данного человека принимается, если в 100 независимых однотипных экспериментах по угадыванию мыслей не менее 70 заканчиваются успехом. Чему равна вероятность принять за телепата человека без телепатических способностей, если в этом случае вероятность угадать в единичном эксперименте равна $1/2$? Переведите условие и ответ этой задачи на строгий язык задачи проверки гипотез.

78. При 4040 бросаниях монеты Бюффон получил 2048 выпадений герба и 1992 выпадения решки. Согласуется ли это с гипотезой о симметричности монеты?

79. События $A, B, C$ составляют полную группу. При 4000 независимых экспериментах они наблюдались 1905, 1015 и 1080 раз соответственно. Согласуется ли это с гипотезой о том, что $P(A) = 0,5, \ P(B) = P(C) = 0,25$?

80. Цифры 0,1, ..., 9 появляются среди первых 800 знаков числа $\pi$ 74, 92, 83, 79, 80, 73, 77, 75, 76, 91 раз соответственно. Проверьте согласованность этих данных с предположением о равных вероятностях их появления.

81. Ниже приведены данные Уэлдона о 4096 опытах по одновременному подбрасыванию 12 костей. В каждом из опытов подсчитывалось число костей, выпавших шестерками. Проверьте гипотезу о правильности подбрасывавшихся Уэлдоном костей.

| Шестерок | 0 | 1 | 2 | 3 | 4 | 5 | 6 | больше 6 |
|---|---|---|---|---|---|---|---|---|
| Случаев | 447 | 1145 | 1181 | 796 | 380 | 115 | 24 | 8 |

82. В Швеции в 1935 году всего родилось 88273 ребенка. По месяцам имелось следующее распределение числа

родившихся детей. Январь – 7280, февраль – 6957, март – 7883, апрель – 7884, май – 7892, июнь – 7609, июль – 7585, август – 7393, сентябрь – 7203, октябрь – 6903, ноябрь – 6552, декабрь – 7132. По этим данным проверьте гипотезу о равных вероятностях рождения детей в любой из 365 дней года.

83. Во время переписи населения Англии и Уэлльса в 1901 году было зарегистрировано 15 729 037 мужчин и 16 799 234 женщины. При этом 3497 мужчин и 3072 женщины были зарегистрированы, как глухонемые от рождения. Предложите способ проверки гипотезы о том, что глухонемота связана с полом. Примените его к приведенным данным.

83. Давление в камере измерялось одновременно двумя манометрами. Всего было произведено 10 экспериментов. Результаты первого манометра дали $\overline{X} = 1573$, $S_X^2 = 0,72$, результаты второго – $\overline{Y} = 1671$, $S_Y^2 = 0,75$. Можно ли считать, что манометры откалиброваны одинаково, если ошибки измерения распределены по нормальному закону?

84. При 1000 наблюдениях над целочисленной случайной величиной значение 0 встретилось 343 раза, 1 – 372, значение 2 – 201 раз, значение 3 – 68 раз, а в остальных случаях появлялось значение, не меньшее 4. Проверьте по этим данным гипотезу о том, что наблюдаемая случайная величина имеет распределение Пуассона.

# 7. Индивидуальное расчетное задание

Для получения зачета по математической статистике предлагается выполнить типовой расчет по индивидуально полученным данным. Расчеты производятся с двумя цифрами после запятой.

Даны две выборки.

1. По первой выборке постройте доверительные интервалы для математического ожидания, считая дисперсию
   (a) неизвестной;
   (b) равной единице.
2. По второй выборке постройте доверительные интервалы для дисперсии, считая математическое ожидание
   (a) неизвестным;
   (b) равным 10.
3. Проверьте гипотезы:
   (a) о нормальности первой выборки;
   (b) о независимости выборок;
   (c) об однородности выборок;
   (d) о равенстве средних значений выборок.

# 8. Случайные процессы. Их характеристики

Случайным процессом формально называют действительнозначную функцию двух аргументов, у которой первый аргумент – результат какого-то случайного эксперимента, а второй интерпретируется как время. При этом фиксация второго превращает случайный процесс в случайную величину, а фиксация первого – в некую неслучайную функцию, которая называется реализацией или траекторий случайного процесса. Таким образом, если область определения процесса по второму аргументу (так называемое индексное множество $T$) представляет собой отрезок числовой прямой, то можно представлять себе случайный процесс $\xi(t),\ t \in T$ как семейство несчетного количества случайных величин, или как бесконечномерный случайный вектор с несчетным числом координат. Распределение любого конечномерного подвектора этого вектора называют конечномерным распределением процесса.

Например, произвольное двумерное распределение $\xi(t)$ строится так. Выберем $t_1, t_2 \in T$ и рассмотрим двумерный вектор $(\xi(t_1),\ \xi(t_2))$. Построение двумерных распределений можно считать законченным, если при произвольном выборе $t_1, t_2$ нам удалось выписать функцию или плотность распределения этого вектора.

Кроме конечномерных распределений к основным характеристикам любого случайного процесса относятся его математическое ожидание $m(t) = \mathbf{M}\xi(t)$ и ковариационная функция:

$$K(t,s) = \operatorname{cov}(\xi(t),\ \xi(s)) = \mathbf{M}\xi(t)\xi(s) - \mathbf{M}\xi(t) \cdot \mathbf{M}\xi(s).$$

Запись случайного процесса в виде

$$\xi(t) \;=\; f_0(t) + \sum_{j=1}^{\infty} f_j(t)\xi_j,$$

где $\xi_j,\; j = 1, 2, ...$ – случайные величины с условиями

$$\mathbf{M}\xi_j \;=\; 0, \quad (i \neq j) \;\Rightarrow\; \mathbf{M}\xi_i\xi_j \;=\; 0,$$

а $f_0(t), f_1(t), ...$ – неслучайные функции, называют каноническим разложением $\xi(t)$. Такая запись, в частности, облегчает поиск математического ожидания и ковариационной функции процесса

$$m(t) \;=\; f_0(t), \quad K(t, s) \;=\; \sum_{j=1}^{\infty} f_j(t)f_j(s)\mathbf{D}\xi_j.$$

## ЗАДАЧИ К ЗАНЯТИЮ

85. Пусть $\eta$ – случайная величина с функцией распределения $F(x)$. Найдите все конечномерные распределения случайного процесса $\xi(t) = \eta + t$.

86. Пусть случайный процесс $\xi(t)$ задан на вероятностном пространстве $< \Omega, \mathcal{F}, P >$, где $\Omega = \{1, 2\}$ , $\mathcal{F}$ — множество всех подмножеств $\Omega$, а $P$ приписывает вероятности 0,5 обоим множествам $\{1\}$ и $\{2\}$. Пусть $t$ пробегает отрезок $[0,\ 1]$ и $\xi(t) = \omega t$. Найдите класс всех реализаций процесса, а также одно, двух и трехмерные его распределения.

87. Пусть случайный процесс $\xi(t)$ задан на вероятностном пространстве $< \Omega, \mathcal{F}, P >$, где $\Omega = [0, 1]$ , класс $\mathcal{F}$ — $\sigma$–алгебра борелевских подмножеств $\Omega$, $P$ – мера Лебега. Пусть $t$ пробегает отрезок $[0,\ 1]$ и

$$\xi(t) \;=\; \mathbf{1}_{\{t \leq \omega\}}.$$

Найдите класс всех реализаций и двумерные распределения процесса.

88. Пусть $\eta_1, \eta_2$ – независимые случайные величины, обе имеющие равномерное распределение на $[0,\ 1]$, в качестве времени рассматривается двумерный параметр $\vec{t} = (t_1, t_2) \in \mathbf{R}^2$. Найдите все такие значения $a$, что почти все реализации случайного процесса

$$X(\vec{t}\,) \ = \ t_1(\eta_1 + t_2(\eta_2 + 2a))$$

монотонно возрастают по $t_1$ при $t_2 = a$.

89. Пусть $\eta$ – случайная величина, подчиняющаяся закону распределения $N(m, \sigma^2)$, $b$ – действительное число. Найдите математическое ожидание и ковариационную функцию случайного процесса

$$\xi(t) \ = \eta t + bt^2, \quad t \geq 0.$$

90. Случайный процесс задан соотношением $Z(t) = 2U \sin 2t + 3Vt^2 + 5$, где $U, V$ – случайные величины,

$$\mathbf{M}U = 5,\ \mathbf{M}V = 5,\ \mathbf{D}U = 1,\ \mathbf{D}V = 0,05,\ \rho(U, V) = -0,2.$$

Найдите математическое ожидание, дисперсию и ковариационную функцию процесса.

91. Дан случайный процесс $\xi(t) = xt + y \sin t$, где среднее случайного вектора $(x, y)$ равно $(1, -1)$, а его ковариационная матрица

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

Для процесса $\xi(t)$ найдите математическое ожидание и ковариационную функцию. Постройте каноническое разложение этого процесса.

92. В представлении случайного процесса

$$W(t) \ = \ xf_1(t) + yf_2(t) + zf_3(t)$$

заданы $\mathbf{M}x = 0$, $\mathbf{M}y = 1$, $\mathbf{M}z = 2$, и ковариационная матрица вектора $(x, y, z)^t$ имеет вид

$$\begin{pmatrix} 2 & 1 & 0 \\ & 3 & 2 \\ & & 4 \end{pmatrix}.$$

Найдите математическое ожидание, дисперсию и ковариационную функцию этого процесса, если $f_i(t), i = 1, 2, 3$ – неслучайные функции. Как построить каноническое разложение этого процесса?

93. Пусть $\xi_1(t), \ldots, \xi_n(t)$ – независимые случайные процессы с нулевыми математическими ожиданиями и известными ковариационными функциями

$$K_1(s, t), \ldots, K_n(s, t).$$

Найдите ковариационную функцию $\xi_1(t) + ... + \xi_n(t)$.

94. Пусть $\xi_1(t), \xi_2(t)$ – два независимых случайных процесса с известными ковариационными функциями $K_1(s, t)$, $K_2(s, t)$ и математическими ожиданиями $m_1(t), m_2(t)$ соответственно. Найдите ковариационную функцию случайного процесса $\xi(t) = \xi_1(t) \cdot \xi_2(t)$.

95. Пусть $f_1(t), \ldots, f_n(t)$ – неслучайные функции, принимающие лишь действительные значения, $c_1, \ldots, c_n$ – положительные числа. Докажите, что функция

$$K(t, s) \;=\; \sum_{j=1}^{n} c_j f_j(s) f_j(t)$$

является ковариационной функцией некоторого случайного процесса.

96*. Докажите, что функция

$$K(t, s) = \begin{cases} |t - s|, & |t - s| < 1, \\ 0 & |t - s| \geq 1 \end{cases}$$

не является ковариационной функцией никакого случайного процесса.

# 9. Интегрирование и дифференцирование случайных процессов

На семействе случайных процессов, интегрируемых с квадратом, может быть определен так называемый предел в среднем квадратическом (L.i.m.):

$$\xi = \text{L.i.m.}_{t\to a}\xi(t) \;\Leftrightarrow\; \lim_{t\to a}\mathbf{M}|\xi(t)-\xi|^2 = 0.$$

Введем обозначение

$$\Delta_h\xi(a) \;=\; \frac{\xi(a+h)-\xi(a)}{h}.$$

Тогда

$$\frac{d\xi}{dt}(a) \;=\; \text{L.i.m.}_{h\to 0}\Delta_h\xi(a)$$

называют производной процесса $\xi(t)$ в точке $a$ в среднем квадратическом. Условием существования такой производной является дифференцируемость в этой точке математического ожидания исходного процесса и существование конечной смешанной производной второго порядка функции $K(t,s)$ при $t=s=a$.

Характеристики производной случайного процесса вычисляются исходя из следующих трех утверждений.

1. $\mathbf{M}\left(\frac{d\xi}{dt}(a)\right) \;=\; m'(a)$;
2. $K_{d\xi/dt}(t,s) \;=\frac{\partial^2 K(t,s)}{\partial t\,\partial s}$;
3. $\operatorname{cov}\left(\xi(t),\, \frac{d\xi}{dt}(s)\right) \;=\frac{\partial K(t,s)}{\partial s}$.

Пусть случайный процесс $\xi(t)$ задан на ограниченном отрезке $[a,\,b]$. Рассмотрим при натуральном $n$ произвольное разбиение этого отрезка на части

$$a=t_0<t_1<...<t_n=b,\;\; \Delta t_j = t_j - t_{j-1},\;\; j=1,...,n.$$

Определим $\Delta = \max\{\Delta t_j\}$ и выберем на каждом из полученных отрезков произвольным образом точку $\theta_j$, $j = 1, ..., n$. Выражение

$$S_n = \sum_{j=1}^{n} \xi(\theta_j) \Delta t_j$$

называют интегральной суммой. Если существует конечный $\mathrm{L.i.m.}_{\Delta \to 0} S_n$, то этот предел называют (определенным) интегралом от процесса по интервалу $[a, b]$ в среднем квадратическом и обозначают

$$\int_a^b \xi(t)\, dt.$$

Разумеется, интеграл уже не будет случайным процессом. Его можно превратить в случайный процесс, рассмотрев верхний предел интеграла, как переменную величину.

Условия интегрируемости процесса по интервалу: функция математического ожидания исходного процесса должна быть интегрируемой по этому интервалу в обычном смысле, а также должен существовать двойной интеграл

$$\int_a^b \int_a^b K(t, s)\, dt ds,$$

где $K(t, s)$ – ковариационная функция процесса $\xi(t)$.

Характеристики интеграла в среднем квадратическом и взаимные характеристики интеграла и процесса задаются следующими формулами.

1. $\mathbf{M} \left( \int_a^b \xi(t)\, dt \right) = \int_a^b m(t)\, dt;$
2. $\mathrm{cov} \left( \int_a^t \xi(u)\, du; \int_a^s \xi(u)\, du \right) = \int_a^t \int_a^s K(u, v) du dv;$
3. $\mathrm{cov} \left( \int_a^t \xi(u)\, du; \xi(s) \right) = \int_a^t K(u, s)\, du.$

Особую важность производные и интегралы в среднем квадратическом имеют в так называемой теории стационарных процессов. Это такие случайные процессы, у которых распределения их приращений на отрезках равной длины одинаковы независимо от места расположения этих отрезков на числовой прямой. В частности, для таких процессов математическое ожидание является константой (не зависит от времени), а ковариационная функция $K(t,s)$ зависит лишь от разности $u = t - s$, и записывается в виде функции одного аргумента $u$.

## ЗАДАЧИ К ЗАНЯТИЮ

97. Известны характеристики случайного процесса:

$$\mathbf{M}X(t) = 2t + 1, \;\; K(t,s) = \exp\left\{-(t-s)^2\right\}.$$

Проверьте существование производной этого процесса в среднем квадратическом. Найдите математическое ожидание, ковариационную функцию и дисперсию его производной.

98. Случайный процесс задан каноническим разложением:

$$X(t) \;=\; 1 + t + Ut + Vt^2 + Wt^3,$$

где $\mathbf{D}U = 2, \;\; \mathbf{D}V = 1, \;\; \mathbf{D}W = 0,5$. Найдите математическое ожидание и ковариационную функцию его производной в среднем квадратическом.

99. Будет ли пуассоновский процесс дифференцируем в среднем квадратическом? А винеровский процесс?

100. На вход интегрирующего устройства поступает случайный процесс, заданный своим каноническим разложением $X(t) = 1 + Ut + Vt^2$, $\mathbf{D}U = 5, \; \mathbf{D}V = 2$. Найдите математическое ожидание и дисперсию процесса на выходе устройства. Вычислите ковариацию между исходным процессом и процессом на выходе интегратора.

101. На вход интегрирующего устройства поступает случайный процесс $X(t)$. Известно, что

$$\mathbf{M}X(t) \;=\; 0,2\cos^2 t, \quad K(t,s) = 0,4\cos t\cos s.$$

Найдите математическое ожидание и дисперсию процесса на выходе устройства.

103. Проверьте, что операции неопределенного интегрирования и дифференцирования в среднем квадратическом взаимно обратны, то есть если траектории процесса $X(t)$ непрерывны с вероятностью 1, и

$$Y(t) = \int\limits_0^t X(s)ds,$$

то $X(t)$ – производная процесса $Y(t)$ в среднем квадратическом.

104. Выведите формулу, связывающую ковариационную функцию стационарного случайного процесса и ковариационную функцию его производной в среднем квадратическом. Сформулируйте условия его дифференцируемости.

105. Стационарный процесс $X(t)$ имеет математическое ожидание, равное 5, и ковариационную функцию

$$K(t) = e^{-2|t|}(\cos 2t + \sin 2|t|).$$

При этом известно, что все его конечномерные приращения имеют нормальные распределения в пространствах соответствующей размерности (гаусовский процесс). Найдите одномерную плотность его производной, а также вероятность того, что процесс $X(t)$ по абсолютной величине не превзойдет 7.

106. Пусть $X(t)$ – стационарный случайный процесс с

ковариационной функцией $K(t)$,

$$Y = \int\limits_a^b X(t)dt.$$

Докажите, что

$$\mathbf{D}Y = \int\limits_a^b dt \int\limits_{t-b}^{t-a} K(s)ds.$$

107. Стационарный центрированный случайный процесс $X(t)$ имеет ковариационную функцию следующего вида

$$K(t) = Ae^{-2|t|}(1 + 2|t|).$$

Найдите отношение дисперсии процесса к дисперсии случайной величины

$$Y = \int_0^1 X(t)\, dt$$

.

108. Известен вид ковариационной функции случайного процесса. Найдите ковариацию между исходным процессом и его второй производной в среднем квадратическом. Рассмотрите отдельно случай стационарного процесса.

## Библиографический список

1. Дронов С.В. Математическая статистика. АлтГУ, ФМиИТ. – Барнаул : Изд-во АлтГУ, 2016. – 1 эл. опт. диск (CD-ROM). – № гос. регистрации 0321602907. Режим доступа: http://elibrary.asu.ru/xmlui/handle/asu/2845.

2. Боровков А.А. Математическая статистика. – М.: Лань, 2010.– 704 с.

3. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.:Наука, 1983.– 416 с.

4. Дронов С.В. Конспект лекций по теории случайных процессов. Барнаул: Изд-во Алт. ун-та. – 2014. – 99 с.

# Приложение

## Некоторые параметрические семейства распределений

В этом приложении перечислены основные (часто встречающиеся в нашем курсе) параметрические семейства распределений, указаны их числовые характеристики и некоторые полезные свойства (без доказательств). Это приложение должно использоваться как краткий справочный материал.

## A Дискретные распределения

1. *Вырожденное* $I_c$. Ряд распределения : $P(\xi = c) = 1$. Числовые характеристики: $\mathbf{M}\xi = c$, $\mathbf{D}\xi = 0$, характеристическая функция $\varphi(t) = \exp\{itc\}$.

2. *Бернуллиевское* $B_{1,p}$ $(0 < p < 1)$. Ряд распределения: $P(\xi = 1) = p$, $P(\xi = 0) = q = 1 - p$. Здесь $\mathbf{M}\xi = p$, $\mathbf{D}\xi = pq$, характеристическая функция $\varphi(t) = pe^{it} + q$.

3. *Биномиальное* $B_{n,p}$ $(0 < p < 1, n \in \mathbf{N})$. Ряд распределения: $P(\xi = k) = C_n^k p^k q^{n-k}$, $0 \le k \le n$. Биномиальная случайная величина представляется в виде суммы $n$ независимых бернуллиевских. Эта величина равна количеству успехов при проведении $n$ испытаний Бернулли с вероятностью успеха $p$ в каждом испытании.

$$\mathbf{M}\xi = np,\ \mathbf{D}\xi = npq,\ \varphi(t) = \left(pe^{it} + q\right)^n.$$

4. *Пуассоновское* $\Pi_\lambda$ $(\lambda > 0)$. Ряд распределения:

$$P(\xi = k) = \frac{\lambda^k}{k!}\, e^{-\lambda}, \quad k = 0, 1, \ldots$$

Среднее, дисперсия и характеристическая функция:

$$\mathbf{M}\xi = \lambda,\ \mathbf{D}\xi = \lambda,\ \varphi(t) = \exp\left\{\lambda(e^{it}-1)\right\}.$$

Пуассоновское распределение есть слабый предел биномиальных $B_{n,p}$ в случае, когда $np \to \lambda$. Класс пуассоновских распределений замкнут относительно суммирований независимых случайных величин. Параметры в этом случае складываются. Любое количество случайных событий на единичном интервале времени, обладающих свойствами относительной редкости и независимости, является пуассоновским.

## B Абсолютно непрерывные распределения

.

1. *Равномерное* $U_{[a,b]}$ $(b > a)$. Плотность: $f(x) = \frac{1}{b-a}$ при $x \in [a,\ b]$, 0 вне этого интервала. Характеристики

   $$\mathbf{M}\xi = \frac{a+b}{2},\ \mathbf{D}\xi = \frac{(b-a)^2}{12},\ \varphi(t) = \frac{e^{itb}-e^{ita}}{it(b-a)}.$$

   Значения $\xi$ сосредоточены на $[a, b]$. Если $\xi$ имеет распределение $U_{[0,1]}$, $F$ – любая функция распределения, то $F(\xi)$ имеет функцию распределения $F$.

2. *Нормальное* $N(a, \sigma^2)$ $(a \in \mathbf{R}, \sigma > 0)$. Плотность:

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

   Характеристики

   $$\mathbf{M}\xi = a,\ \mathbf{D}\xi = \sigma^2,\ \varphi(t) = \exp\left\{ita - \frac{t^2\sigma^2}{2}\right\}.$$

Класс нормальных распределений замкнут относительно линейных комбинаций, а также относительно суммирования произвольным образом зависимых случайных величин, имеющих совместное нормальное распределение. При $a = 0$, $\sigma = 1$ – стандартное нормальное распределение. Случайная величина $(\xi - a)/\sigma$ имеет стандартное нормальное распределение («следствие за четыреста рублей»).

3. *Гамма-распределение* $\Gamma_{\alpha,\lambda}$ $(\alpha, \lambda > 0)$. Плотность:

$$f(x) = \frac{\alpha^{\lambda}}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x} \quad (x \geq 0).$$

Характеристики

$$\mathbf{M}\xi = \frac{\lambda}{\alpha}\,, \quad \mathbf{D}\xi = \frac{\lambda}{\alpha^2}\,, \quad \varphi(t) = \left(1 - \frac{it}{\alpha}\right)^{-\lambda}.$$

При $\lambda = 1$ это распределение называют *экспоненциальным* или *показательным* и обозначают $E_{\alpha}$.

4. *Хи-квадрат распределение с $n$ степенями свободы* $\chi^2_n$. Это распределение является частным случаем гамма-распределения при $\alpha = \frac{1}{2}$, $\lambda = \frac{n}{2}$. Оно особенно важно в силу того, что случайная величина, имеющая такое распределение, представима в виде суммы квадратов $n$ независимых стандартных нормальных случайных величин. Характеристики:

$$\mathbf{M}\xi = n, \quad \mathbf{D}\xi = 2n, \quad \varphi(t) = (1 - 2it)^{n/2}.$$

5. *Распределение Стьюдента с $n$ степенями свободы* $T_n$. Плотность:

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

Распределение $T_n$ при $n \to \infty$ аппроксимируется стандартным нормальным. Имеет место представление

$$T_n = \frac{N(0,1)}{\sqrt{\frac{1}{n}\chi_n^2}} \ ,$$

где числитель и знаменатель независимы. При $n = 1$ называется *распределением Коши* и не имеет математического ожидания. Плотность распределения Коши имеет вид

$$p(x) \ = \ \frac{1}{\pi} \, \frac{1}{1+x^2} \ .$$

При $n = 2$ имеет нулевое среднее и не имеет дисперсии. Распределение Стьюдента симметрично, а следовательно,

$$(\forall x) \ T_n(x) + T_n(-x) = 1,$$

где $T_n(x)$ обозначена функция соответствующего распределения.

Это распределение в англоязычной литературе и некоторых статистических пакетах называют $T$-распределением. Именно поэтому статистики, имеющие такое распределение, часто обозначаются буквой $T$.

6. *Распределение Фишера ($F$-распределение) с $n, m$ степенями свободы $F_{n,m}$.* Так называется отношение двух независимых распределений хи-квадрат, нормированных соответственно своими степенями свободы:

$$F_{n,m} \ = \ \frac{m\chi_n^2}{n\chi_m^2}.$$

Его функция распределения задается формулой

$$F(x, n, m) =$$

$$= \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \, n^{n/2} m^{m/2} \int\limits_0^x z^{\frac{n}{2}-1} (m + nz)^{-\frac{n+m}{2}} \, dz$$

при $x > 0$.