

Министерство образования и науки Российской Федерации
САНКТ-ПЕТЕРБУРГСКИЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО

Б. И. Положинцев

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА
ВВЕДЕНИЕ В МАТЕМАТИЧЕСКУЮ
СТАТИСТИКУ

Учебное пособие

2-е издание, переработанное и дополненное



Санкт-Петербург
2018

УДК 519.2(075.8)
П52

Положинцев Б. И. **Теория вероятностей и математическая статистика. Введение в математическую статистику** : учеб. пособие / Б. И. Положинцев. – 2-е изд., перераб. и доп. – СПб. : Изд-во Политехн. ун-та, 2018. – 96 с.

В пособии изложены основные понятия и методы статистической обработки экспериментальных данных: точечное и интервальное оценивание параметров распределений, проверка статистических гипотез о параметрах распределений (критерии значимости), а также о виде распределений (критерии согласия).

Предназначено для студентов, обучающихся по направлениям 11.03.01 “Радиотехника” и 11.03.02 “Инфокоммуникационные технологии и системы связи” Института физики, нанотехнологий и телекоммуникаций.

Печатается по решению
Совета по издательской деятельности Ученого совета
Санкт-Петербургского политехнического университета Петра Великого.

ISBN 978-5-7422-6083-7

© Положинцев Б. И., 2018
© Санкт-Петербургский политехнический
университет Петра Великого, 2018

СОДЕРЖАНИЕ

1. Определение случайной выборки.....	5
2. Закон распределения порядковых статистик.....	6
3. Эмпирическая функция распределения.....	8
4. Группирование выборочных данных, гистограмма.....	11
5. Определение и свойства точечных оценок параметров распределения: состоятельность, несмещенность, эффективность.....	13
6. Оценки основных числовых характеристик распределения и их свойства.....	16
7. Выборочные квантили.....	21
8. Нахождение оценок параметров распределений методом максимального правдоподобия.....	23
9. Примеры нахождения оценок максимального правдоподобия (МП- оценок) параметров распределений.....	25
10. Понятие доверительного интервала.....	28
11. Основные этапы процедуры построения доверительных интервалов.....	29
12. Доверительный интервал для математического ожидания нормального распределения при известной дисперсии. Пример 1_ди.....	31
13. Распределение χ^2 , распределение Стьюдента, лемма Фишера..	34
14. Доверительный интервал для математического ожидания нормального распределения при неизвестной дисперсии Пример 2_ди и Пример 3_ди.....	36
15. Доверительный интервал для дисперсии нормального распределения.....	40
16. Приближенная интервальная оценка для математического ожидания произвольного распределения по выборке большого объема.....	42
17. Приближенная интервальная оценка вероятности p в схеме Бернулли по выборке большого объема. Пример 4_ди.....	43
18. Постановка задачи проверки статистических гипотез. Пример 1_кз.....	46
19. Критерии значимости: гипотезы, критическая область, решения, ошибки.....	48

20. Основные этапы процедуры проверки статистических гипотез.....	53
21. Подход к проверке статистических гипотез о параметрах распределений, основанный на доверительных интервалах Пример 1_кди.....	53
22. Примеры проверки гипотез о параметрах распределений: Пример 2_кз_кди; Пример 3_кз_кди (левосторонний критерий); Пример 4_кз (правосторонний критерий); Пример 5_кз_кди (двусторонний критерий).....	55
23. Распределение Фишера, свойство квантилей.....	66
24. Проверка гипотезы о равенстве дисперсий двух нормальных распределений (критерий Фишера). Пример 6_кз.....	67
25. Проверка гипотезы о равенстве математических ожиданий двух нормальных распределений (критерий Стьюдента). Пример 7_кз.....	68
26. Теорема Пирсона, проверка простой статистической гипотезы.....	72
27. Проверка гипотезы о виде распределения – метод χ^2 для простой гипотезы.....	75
28. Проверка гипотезы о виде распределения – метод χ^2 для сложной гипотезы.....	77
29. Пример 1_кс (нормальное распределение).....	79
30. Пример 2_кс (распределение Пуассона).....	81
31. Проверка гипотезы о равенстве параметров p_1 и p_2 (вероятностей) двух биномиальных распределений по выборкам большого объема.....	85
32. Понятие p – значения.....	87
<i>Приложение 1. Предельные теоремы теории вероятностей.....</i>	<i>87</i>
<i>Приложение 2. Получение выборки из заданного распределения..</i>	<i>93</i>
Литература.....	95

В пособии рассматриваются основные понятия и методы анализа данных, полученных в результате опыта – наблюдений (измерений, регистраций) величин, случайных по своей природе. При этом принципиально осуществимой предполагается возможность неограниченного числа таких наблюдений в одних и тех же условиях.

В большинстве случаев далее считается, что статистические данные представляют собой результат серии независимых опытов, в каждом из которых зарегистрировано значение исследуемой одномерной случайной величины.

1. Определение случайной выборки

Пусть X – исследуемая случайная величина, $F_X(x) = P(X < x)$ – ее функция распределения (вообще говоря, неизвестная). В ряде случаев может быть известен вид распределения случайной величины, а неизвестными являются один или несколько параметров, от которых зависит функция распределения. Ради краткости в записи $F_X(x)$ индекс может в дальнейшем опускаться. Условимся также указывать, непрерывной или дискретной является исследуемая случайная величина.

Пусть проводится серия n независимых наблюдений (измерений) случайной величины X в одних и тех же условиях (эксперимент). В результате эксперимента получают n чисел – значений x_1, x_2, \dots, x_n , которые случайная величина X последовательно принимала в данной серии наблюдений. Эти числа будем считать значениями n одинаково распределенных независимых случайных величин X_1, \dots, X_n , каждая из которых имеет функцию распределения $F_X(x)$ – ту же, что исследуемая случайная величина X .

Конечную последовательность n независимых одинаково распределенных случайных величин будем называть *случайной выборкой* X_1, \dots, X_n (короче – *выборкой*) из распределения $F_X(x)$, а указанные числа x_1, x_2, \dots, x_n , полученные в данном эксперименте – *реализацией выборки*. Отметим, что множество всех возможных

значений исследуемой случайной величины называют генеральной совокупностью.

На основе выборок строят *оценки* параметров распределения исследуемой случайной величины X , таких как математическое ожидание, стандартное отклонение и других, а также судят о виде функции распределения $F_X(x)$.

Понятно, что числа x_1, x_2, \dots, x_n можно также рассматривать как значение n -мерной случайной величины (X_1, \dots, X_n) , компоненты которой X_1, \dots, X_n независимы и одинаково распределены.

Всякую функцию выборки $\varphi(X_1, \dots, X_n)$ называют статистикой. Статистика $\varphi(X_1, \dots, X_n)$ – случайная величина, распределение которой зависит от распределения $F_X(x)$, из которого извлечена выборка, и от объема выборки n .

2. Закон распределения порядковых статистик

Пусть X_1, \dots, X_n – выборка объема n из распределения $F_X(x)$; x_1, x_2, \dots, x_n – некоторая ее реализация.

Упорядочим числа x_1, x_2, \dots, x_n по возрастанию и обозначим их следующим образом: $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, где $x_{(1)} = \min(x_1, x_2, \dots, x_n)$, $x_{(n)} = \max(x_1, x_2, \dots, x_n)$, $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

Представим, что упорядочены все возможные реализации выборки X_1, \dots, X_n и введем новую случайную величину $X_{(k)}$ – порядковую статистику порядка k ($k = 1, 2, \dots, n$).

Множество возможных значений случайной величины $X_{(k)}$ определим так: оно состоит из тех и только тех чисел $x_{(k)}^i$, которые оказываются на k -м месте при упорядочении любой реализации x_1, x_2, \dots, x_n выборки X_1, \dots, X_n (индекс $i = 1, 2, \dots$ – номер реализации).

Таким образом, по выборке X_1, \dots, X_n построена последовательность $X_{(1)}, \dots, X_{(k)}, \dots, X_{(n)}$, называемая вариационным

рядом. Элементы вариационного ряда – порядковые статистики удовлетворяют соотношениям: $X_{(1)} \leq \dots \leq X_{(k)} \leq \dots \leq X_{(n)}$, при этом в любой реализации вариационного ряда числа $x^i_{(1)}, \dots, x^i_{(k)}, \dots, x^i_{(n)}$ связаны неравенствами $x^i_{(1)} \leq \dots \leq x^i_{(k)} \leq \dots \leq x^i_{(n)}$ (верхний индекс i – номер реализации, $i = 1, 2, \dots$).

$$\begin{array}{cccc}
 x^1_{(1)} & \dots & x^1_{(k)} & \dots & x^1_{(n)} \\
 \cdot & & \cdot & & \cdot \\
 \cdot & & \cdot & & \cdot \\
 \cdot & & \cdot & & \cdot \\
 x^i_{(1)} & \dots & x^i_{(k)} & \dots & x^i_{(n)} \\
 \cdot & & \cdot & & \cdot \\
 \cdot & & \cdot & & \cdot \\
 \cdot & & \cdot & & \cdot \\
 \hline
 X_{(1)} & \dots & X_{(k)} & \dots & X_{(n)}
 \end{array}$$

Найдем функцию распределения k -й порядковой статистики

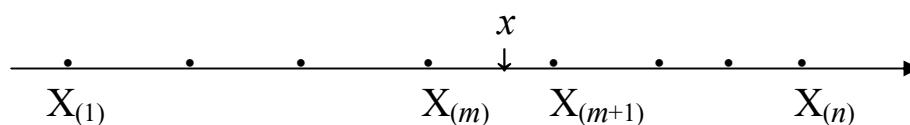
$$X_{(k)}: F_{X_{(k)}}(x) = P(X_{(k)} < x) \quad (k = 1, 2, \dots, n).$$

Эмпирической частотой $N_n(x)$ назовем случайную величину, равную числу элементов выборки X_1, \dots, X_n , меньших x (иначе – числу элементов вариационного ряда $X_{(1)}, \dots, X_{(n)}$, меньших x). Ясно, что возможные значения эмпирической частоты $N_n(x)$ – число осуществлений события $(X < x)$ на выборке X_1, \dots, X_n объема n – это числа $m = 0, 1, \dots, n$. Действительно,

$$(N_n(x) = 0) = (x \leq X_{(1)});$$

$$(N_n(x) = m) = (X_{(m)} < x \leq X_{(m+1)}) \quad \forall m = 1, \dots, n-1;$$

$$(N_n(x) = n) = (x > X_{(n)}).$$

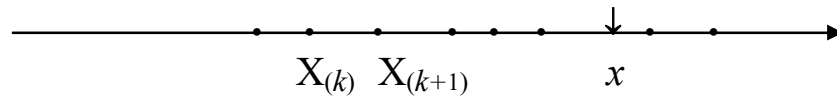


Извлечение выборки из распределения $F_X(x)$ представляет собой серию n независимых испытаний – n наблюдений (регистраций значений) исследуемой случайной величины X . Для каждого из указанных испытаний вероятность события $(X < x)$ равна $P(X < x) = F_X(x)$.

Отсюда следует, что случайная величина $N_n(x)$ подчиняется биномиальному распределению:

$$P(N_n(x) = m) = C_n^m (F_X(x))^m (1 - F_X(x))^{n-m} \quad (m = 0, 1, \dots, n).$$

Заметим, что события $(X_{(k)} < x)$ и $(N_n(x) \geq k)$ равносильны,



то есть $(X_{(k)} < x) = (N_n(x) \geq k) = \sum_{m=k}^n (N_n(x) = m)$.

Таким образом, получаем:

$$P(X_{(k)} < x) = F_{X_{(k)}}(x) = \sum_{m=k}^n C_n^m (F_X(x))^m (1 - F_X(x))^{n-m}, \quad \forall k = 1, \dots, n -$$

– закон распределения порядковых статистик.

При $k=1$ и $k=n$ имеем распределения *экстремальных порядковых статистик*:

минимальной $X_{(1)}$: $F_{X_{(1)}}(x) = 1 - (1 - F_X(x))^n$ и

максимальной $X_{(n)}$: $F_{X_{(n)}}(x) = (F_X(x))^n$.

3. Эмпирическая функция распределения

Пусть X_1, \dots, X_n – выборка из распределения $F_X(x)$, $X_{(1)}, \dots, X_{(k)}, \dots, X_{(n)}$ – вариационный ряд, $N_n(x)$ – эмпирическая частота. Случайная величина $F_n(x) = N_n(x)/n$, называемая эмпирической

функцией распределения – относительная частота числа элементов выборки X_1, \dots, X_n , удовлетворяющих условию $X_i < x$.

Ясно, что множество возможных значений эмпирической функции распределения есть: $0, 1/n, \dots, m/n, \dots, n/n$.

События $(F_n(x) = m/n)$ и $(N_n(x) = m)$ – равносильны, эмпирическая частота $N_n(x)$ распределена по биномиальному закону, поэтому

$$P(F_n(x) = m/n) = C_n^m (F_X(x))^m (1 - F_X(x))^{n-m} \quad (m=0,1,\dots,n) -$$

– закон распределения эмпирической функции распределения $F_n(x)$.

С помощью функции единичного скачка (функции Хевисайда) эмпирическая функция распределения может быть записана в виде:

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n e(x - X_{(k)}),$$

$$\text{где } e(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases} - \text{ функция Хевисайда.}$$

Заметим, что для каждой реализации выборки эмпирическая функция распределения обладает всеми свойствами функции распределения. Действительно, пусть x_1, x_2, \dots, x_n – некоторая реализация выборки, $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ – соответствующая реализация вариационного ряда, где $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

Среди чисел x_1, x_2, \dots, x_n выберем только различные, упорядочим их и обозначим через $x_{i_1}, x_{i_2}, \dots, x_{i_k}$, тогда

$$x_{(1)} = x_{i_1} < \dots < x_{i_k} = x_{(n)} \quad (k \leq n)$$

x_{i_1}	\dots	x_{i_m}	\dots	x_{i_k}
n_1	\dots	n_m	\dots	n_k
n_1/n	\dots	n_m/n	\dots	n_k/n

Здесь n_m – абсолютная, n_m/n – относительная частота элемента x_{i_m} , при этом, очевидно: $\sum_{m=1}^k n_m = n$; $\sum_{m=1}^k \frac{n_m}{n} = 1$.

Введем случайную величину X^* , заданную рядом распределения

X^*	x_{i_1}	\dots	x_{i_m}	\dots	x_{i_k}
$P(X^* = x_{i_m})$	n_1/n	\dots	n_m/n	\dots	n_k/n

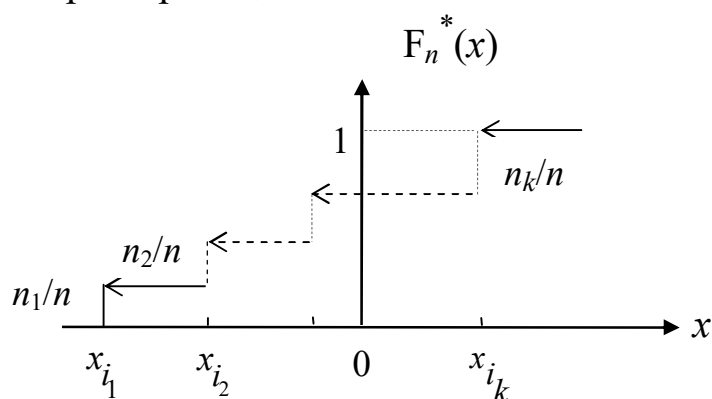
Заметим, что таким образом каждому элементу реализации выборки x_1, x_2, \dots, x_n приписана вероятность $1/n$.

Обозначим через $F_n^*(x)$ реализацию случайной величины $F_n(x)$, отвечающую данной реализации выборки, тогда

$$F_n^*(x) = P(X^* < x) = \sum_{m: x_{i_m} < x} \frac{n_m}{n}.$$

$F_n^*(x)$ – кусочно-постоянная (ступенчатая) функция, принимающая свои значения на отрезке $[0; 1]$. В каждой точке x , кроме точек x_{i_m} , функция $F_n^*(x)$ непрерывна; в точках x_{i_m} – она непрерывна слева, величина скачка справа равна n_m/n ($m = 1, 2, \dots, k$).

График эмпирической функции распределения для некоторой реализации выборки приведен ниже:



Эмпирическая функция распределения $F_n(x)$ как относительная частота числа осуществлений на выборке события $(X < x)$ при любом x сходится по вероятности к вероятности этого события

$P(X < x) = F_X(x)$, – к теоретической функции распределения, вообще говоря, неизвестной: $\forall x F_n(x) \xrightarrow[n \rightarrow +\infty]{P} F_X(x)$, поэтому, если объем выборки n достаточно велик, то значение эмпирической функции распределения $F_n(x)$ в каждой точке x оказывается близким к соответствующему значению теоретической функции распределения $F_X(x)$.

Доказано (*теорема Гливленко*), что отклонение эмпирической функции распределения $F_n(x)$ – случайной величины – от теоретической функции распределения $F_X(x)$ с вероятностью 1 сколь угодно мало при достаточно большом объеме выборки n :

$$\forall \varepsilon > 0 \quad P \left(\sup_x |F_n(x) - F_X(x)| \leq \varepsilon \right) \xrightarrow[n \rightarrow +\infty]{} 1,$$

при этом $F_n(x)$ служит равномерным приближением $F_X(x)$ на всей числовой оси. Заметим, что разность $(F_n(x) - F_X(x))$ асимптотически нормальна с нулевым математическим ожиданием.

4. Группирование выборочных данных, гистограмма

Эмпирическая функция распределения является характеристикой выборки, позволяющей наглядно представлять статистические данные и выдвигать предположения о виде неизвестной функции распределения исследуемой (наблюдаемой) случайной величины.

Другой способ представления статистического материала – это построение группированного статистического ряда и гистограммы.

Пусть исследуемая случайная величина X – непрерывна. Если выборка достаточно большая (обычно в статистике большими считают выборки объемом $n \geq 100$), то ее реализацию (x_1, x_2, \dots, x_n) подвергают группировке следующим образом.

Отрезок $[x_{(1)}; x_{(n)}]$, где $x_{(1)} = \min(x_1, x_2, \dots, x_n)$, $x_{(n)} = \max(x_1, x_2, \dots, x_n)$, содержащий все элементы выборки, разбивают на k равных интервалов Δ_i (обычно $5 \leq k \leq 15$):

Относительная частота события по вероятности сходится к вероятности этого события, поэтому если длина интервалов разбиения h достаточно мала, то $\frac{n_i}{n} \cong f_X(x)h \forall x \in \Delta_i$. При больших n верхний контур гистограммы (ступенчатый график) служит приближением графика плотности вероятности $f_X(x)$ (вообще говоря, неизвестной). Таким образом, разумно построенная гистограмма позволяет выдвинуть гипотезу о виде распределения исследуемой случайной величины X . Заметим, что слишком малое или слишком большое число интервалов разбиения k при построении гистограммы может привести к ее недостаточной информативности.

Число интервалов k при разбиении отрезка $[x_{(1)}; x_{(n)}]$ обычно определяют по формуле $k = 1 + 3,32 \lg n$ (формула Старджесса), либо по формуле $k = 1,72 n^{1/3}$.

5. Определение и свойства точечных оценок параметров распределения: состоятельность, несмещенность, эффективность

Пусть θ – некоторый параметр распределения $F_X(x, \theta)$. Информация, необходимая для нахождения оценки $\hat{\theta}$ неизвестного параметра θ , содержится в выборке X_1, \dots, X_n из данного распределения. Таким образом, возникает задача построения оценки $\hat{\theta}$ параметра распределения как функции случайной выборки: $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$.

Заметим, что оценка параметра распределения является случайной величиной (статистикой). В результате проведения эксперимента (серии n независимых наблюдений) получают реализацию выборки – числа x_1, x_2, \dots, x_n . При этом оценка $\hat{\theta}$ принимает соответствующее числовое значение $\hat{\theta}_e = \hat{\theta}(x_1, x_2, \dots, x_n)$, которое является приближенным значением неизвестного параметра θ . Оценки указанного типа называют точечными, их применение целесообразно

при достаточно больших выборках. При малых объемах выборок используют интервальные оценки, которые будут рассмотрены далее.

Ниже определяются свойства точечных оценок: *состоятельность*, *несмещенность*, *эффективность*, каждое из которых определенным образом характеризует меру близости оценки $\hat{\theta}$ (случайной величины) к истинному значению (неслучайной величине) неизвестного параметра θ распределения $F_X(x, \theta)$. Понятно, что вопрос обладает ли данная оценка каким-либо (или всеми) из указанных свойств требует специального рассмотрения.

Состоятельность. Оценка $\hat{\theta}$ называется состоятельной оценкой параметра θ , если оценка $\hat{\theta}$ по вероятности сходится к оцениваемому параметру: $\forall \varepsilon > 0 \quad P(|\hat{\theta} - \theta| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$ (символически это принято записывать так: $\hat{\theta} \xrightarrow[n \rightarrow +\infty]{P} \theta$). Иными словами, состоятельная оценка

обладает свойством: с увеличением объема выборки n уменьшается вероятность того, что абсолютная величина отклонения оценки от оцениваемого параметра θ превзойдет любое наперед заданное $\varepsilon > 0$.

Несмещенность. Оценка $\hat{\theta}$ называется несмещенной оценкой параметра θ , если ее математическое ожидание равно оцениваемому параметру: $M\hat{\theta} = \theta$.

Если $M\hat{\theta} \neq \theta$, то имеет место систематическая ошибка, величину $|M\hat{\theta} - \theta|$ называют смещением.

Оценка $\hat{\theta}$ называется асимптотически несмещенной, если

$$M\hat{\theta} \xrightarrow[n \rightarrow +\infty]{} \theta$$

В качестве упражнения докажем следующее утверждение:

Пусть $\hat{\theta}$ – несмещенная оценка параметра распределения θ , причем $D[\hat{\theta}] \xrightarrow[n \rightarrow +\infty]{} 0$, тогда $\hat{\theta}$ – состоятельна.

Запишем неравенство Чебышева: $\forall \varepsilon > 0 \quad 0 \leq P(|\hat{\theta} - M\hat{\theta}| \geq \varepsilon) \leq \frac{D[\hat{\theta}]}{\varepsilon^2}$.

Учтем, что по условию $\hat{\theta}$ – несмещенная оценка ($M\hat{\theta} = \theta$), тогда, переходя к пределу при $n \rightarrow +\infty$, по теореме о сжатой переменной имеем:

$\forall \varepsilon > 0 \quad P(|\hat{\theta} - \theta| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$, то есть $\hat{\theta}$ – состоятельна: $\hat{\theta} \xrightarrow[n \rightarrow +\infty]{P} \theta$.

Эффективность. Оценку $\hat{\theta}$ называют эффективной оценкой параметра θ в классе оценок, имеющих дисперсию, если дисперсия $\hat{\theta}$ минимальна. Из двух оценок, принадлежащих данному классу, более эффективна та, которая имеет меньшую дисперсию.

Рассмотрим в качестве примера свойства относительной частоты как оценки вероятности в схеме Бернулли.

Пусть случайная величина X подчиняется биномиальному распределению с параметрами n и p .

Относительная частота $\frac{X}{n} = \hat{p}$ как оценка вероятности p обладает следующими свойствами:

1) \hat{p} – состоятельная, что следует из закона больших чисел:

$$\hat{p} = \frac{X}{n} \xrightarrow[n \rightarrow +\infty]{P} p.$$

2) \hat{p} – несмещенная, так как:

$$M[\hat{p}] = M\left[\frac{1}{n} X\right] = \frac{1}{n} M[X] = \frac{1}{n} np = p$$

3) Дисперсия \hat{p} – бесконечно малая при $n \rightarrow +\infty$:

$$D[\hat{p}] = D\left[\frac{X}{n}\right] = \frac{1}{n^2} npq = (1/n) pq \xrightarrow[n \rightarrow +\infty]{} 0.$$

4) Согласно теореме Муара-Лапласа эта оценка является асимптотически нормальной: $\hat{p} = \frac{X}{n} \xrightarrow[n \rightarrow +\infty]{\sim} N\left(p; \sqrt{\frac{pq}{n}}\right)$.

6. Оценки основных числовых характеристик распределения и их свойства

Функция распределения $F_X(x)$, полностью определяющая свойства случайной величины X , в большинстве случаев неизвестна. Однако информацию о распределении, достаточную во многих случаях для практических целей, могут дать оценки основных числовых характеристик распределения.

Итак, пусть (X_1, \dots, X_n) – выборка из распределения $F_X(x)$ исследуемой одномерной случайной величины, (x_1, x_2, \dots, x_n) – реализация выборки.

Числовая характеристика случайной величины X	Оценка числовой характеристики – функция выборки (X_1, \dots, X_n)	Реализация оценки – значение оценки в точке (x_1, x_2, \dots, x_n)
$\alpha_k = M[X^k]$ – начальный момент k -го порядка	$\hat{\alpha}_k = \frac{1}{n} \sum_{i=1}^n X_i^k$ – выборочный начальный момент k -го порядка	$\hat{\alpha}_{ke} = \frac{1}{n} \sum_{i=1}^n x_i^k = a_k$ – выборочный начальный момент k -го порядка
$MX = m_X = m = \alpha_1$ – математическое ожидание	$\bar{X} = \hat{m} = \frac{1}{n} \sum_{i=1}^n X_i = \hat{\alpha}_1$ – выборочное среднее	$\bar{x} = \hat{m}_e = \frac{1}{n} \sum_{i=1}^n x_i = \hat{\alpha}_{1e}$ – выборочное среднее
$\mu_k = M[(X - MX)^k]$ – центральный момент k -го порядка	$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$ – выборочный центральный момент k -го порядка	$\hat{\mu}_{ke} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k = m_k$ – выборочный центральный момент k -го порядка
$DX = M[(X - MX)^2] = \sigma^2 = \mu_2$ – дисперсия $\sigma_X = \sigma$ – стандартное отклонение	$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \hat{\mu}_2 = \hat{\sigma}^2$ – выборочная дисперсия $S^{*2} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ – исправленная (несмещенная) выборочная дисперсия	$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \hat{\mu}_{2e} = \hat{\sigma}_e^2$ – выборочная дисперсия $s^{*2} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ – исправленная (несмещенная) выборочная дисперсия

Подчеркнем еще раз, что числовая характеристика случайной величины – неслучайная величина, ее оценка (функция выборки) – случайная величина, реализация оценки (значение оценки, которое она принимает в точке (x_1, x_2, \dots, x_n)) – неслучайная величина (число).

Следующие теоремы раскрывают свойства указанных оценок.

Теорема Выборочный начальный момент k -го порядка $\hat{\alpha}_k$ ($\hat{\alpha}_k = \frac{1}{n} \sum_{i=1}^n X_i^k$) является несмещенной, состоятельной, асимптотически нормальной оценкой начального момента k -го порядка α_k ($\alpha_k = M[X^k]$) случайной величины X (при условии, что существуют конечные α_k и α_{2k}).

Доказательство Поскольку $MX_i = MX$ и $MX_i^k = MX^k$,

имеем: $M[\hat{\alpha}_k] = M\left[\frac{1}{n} \sum_{i=1}^n X_i^k\right] = \frac{1}{n} (nM[X^k]) = \alpha_k$ – несмещенность.

$$\begin{aligned} \text{Далее, } D[\hat{\alpha}_k] &= D\left[\frac{1}{n} \sum_{i=1}^n X_i^k\right] = \left(\frac{1}{n}\right)^2 \sum_{i=1}^n D[X_i^k] = \left(\frac{1}{n}\right)^2 nD[X^k] = \\ &= \frac{1}{n} (M[X^{2k}] - (M[X^k])^2) = \frac{\alpha_{2k} - \alpha_k^2}{n} \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

Таким образом $\hat{\alpha}_k$ – несмещенная оценка, ее дисперсия – бесконечно малая при $n \rightarrow +\infty$, откуда следует состоятельность $\hat{\alpha}_k$ (см. утверждение о несмещенной оценке, дисперсия которой – бесконечно малая при $n \rightarrow +\infty$, доказанное в п. 5):

$$\hat{\alpha}_k \xrightarrow[n \rightarrow +\infty]{P} \alpha_k.$$

Заметим, что выборочные центральные моменты $\hat{\mu}_k$ – также являются состоятельными оценками центральных моментов μ_k случайной величины X .

Докажем теперь, что $\hat{\alpha}_k$ – асимптотически нормальная оценка начального момента α_k .

Действительно, выборочный начальный момент $\hat{\alpha}_k$ есть сумма одинаково распределенных независимых случайных величин, имеющих конечное математическое ожидание и дисперсию.

Согласно центральной предельной теореме:

$$\frac{\hat{\alpha}_k - M[\hat{\alpha}_k]}{\sqrt{D[\hat{\alpha}_k]}} = \frac{\hat{\alpha}_k - \alpha_k}{\sqrt{\frac{\alpha_{2k} - \alpha_k^2}{n}}} \underset{n \rightarrow +\infty}{\sim} N(0;1).$$

Таким образом, $\hat{\alpha}_k$ – асимптотически нормальная случайная величина с математическим ожиданием и стандартным отклонением, равными, соответственно: α_k и $\sqrt{\frac{\alpha_{2k} - \alpha_k^2}{n}}$. Символически это можно

записать так: $\hat{\alpha}_k \underset{n \rightarrow +\infty}{\sim} N\left(\alpha_k; \sqrt{\frac{\alpha_{2k} - \alpha_k^2}{n}}\right)$.

Следствие Выборочное среднее $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \hat{\alpha}_1$ является несмещенной, состоятельной, асимптотически нормальной оценкой математического ожидания MX случайной величины X .

Теорема. Выборочная дисперсия $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ – состоятельная, асимптотически несмещенная оценка дисперсии DX случайной величины X .

Доказательство

$$\begin{aligned} \text{Запишем следующее очевидное равенство: } S^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \\ &= \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2) = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 = \hat{\alpha}_2 - \hat{\alpha}_1^2. \end{aligned}$$

В предыдущей теореме доказана состоятельность оценки $\hat{\alpha}_k$ k -го начального момента α_k :

$$\hat{\alpha}_k \xrightarrow[n \rightarrow +\infty]{P} \alpha_k.$$

Известные теоремы о пределе суммы и произведения функций справедливы и для сходимости по вероятности, поэтому имеем:

$$S^2 = (\hat{\alpha}_2 - \hat{\alpha}_1^2) \xrightarrow[n \rightarrow +\infty]{P} (\alpha_2 - \alpha_1^2) = M[X^2] - (MX)^2 = DX.$$

Таким образом, S^2 – состоятельная оценка дисперсии DX :

$$S^2 \xrightarrow[n \rightarrow +\infty]{P} DX = \sigma_X^2.$$

Докажем теперь, что выборочная дисперсия S^2 является *асимптотически несмещенной* оценкой дисперсии DX .

Поскольку $S^2 = \hat{\alpha}_2 - \hat{\alpha}_1^2$, имеем: $M[S^2] = M\hat{\alpha}_2 - M(\hat{\alpha}_1^2)$.

Как было доказано, выборочные начальные моменты $\hat{\alpha}_k$ – несмещенные оценки соответствующих начальных моментов α_k , поэтому $M[\hat{\alpha}_2] = \alpha_2 = M(X^2)$.

Далее запишем:

$$M(\hat{\alpha}_1^2) = M(\bar{X}^2) = D\bar{X} + (M\bar{X})^2 = \frac{1}{n}DX + (M\bar{X})^2 = \frac{1}{n}DX + (MX)^2$$

(здесь использованы очевидные равенства $M\bar{X} = MX$, $D\bar{X} = \frac{1}{n}DX$).

В итоге получаем:

$$MS^2 = M\hat{\alpha}_2 - M(\hat{\alpha}_1^2) = M(X^2) - \frac{1}{n}DX - (MX)^2 = DX - \frac{1}{n}DX = \frac{n-1}{n}DX,$$

выборочная дисперсия S^2 не является несмещенной; однако, эта оценка – *асимптотически несмещенная*, поскольку

$$MS^2 = \frac{n-1}{n}DX \xrightarrow[n \rightarrow +\infty]{} DX.$$

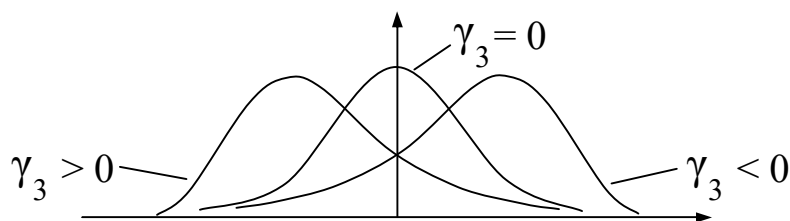
Наряду с выборочной дисперсией S^2 в качестве оценки дисперсии DX используют также *исправленную* выборочную дисперсию $S^{*2} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n}{n-1} S^2$ – несмещенную состоятельную оценку дисперсии DX .

Действительно: $MS^{*2} = \frac{n}{n-1} (\frac{n-1}{n}DX) = DX$ и $S^{*2} = (\frac{n}{n-1} S^2) \xrightarrow[n \rightarrow +\infty]{P} DX$.

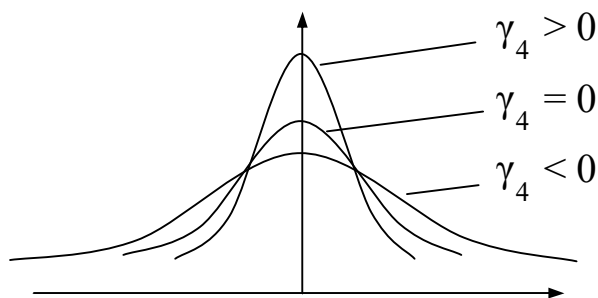
Замечание

Коэффициент асимметрии (асимметрия) $\gamma_3 = \mu_3 / \sigma^3$ и коэффициент эксцесса (эксцесс) $\gamma_4 = \mu_4 / \sigma^4 - 3$ характеризуют асимметрию и “островершинность” распределения, соответственно. Используют также обозначение *sk* (skewness) для асимметрии и *ku* (kurtosis) – для эксцесса. Указанные числовые характеристики определены для распределений, у которых существуют конечные центральные моменты до четвертого включительно.

Для симметричных распределений (в частности – для нормального распределения) коэффициент асимметрии $\gamma_3 = 0$.



Нормальное распределение имеет эксцесс равный нулю, $\gamma_4 = 0$.



Оценками указанных числовых характеристик служат *выборочная асимметрия* $\hat{\gamma}_3 = \hat{\mu}_3 / S^3$ и *выборочный эксцесс* $\hat{\gamma}_4 = \hat{\mu}_4 / S^4 - 3$.

Реализации этих оценок: $\hat{\gamma}_{3e} = \hat{\mu}_{3e} / s^3$, $\hat{\gamma}_{4e} = \hat{\mu}_{4e} / s^4 - 3$

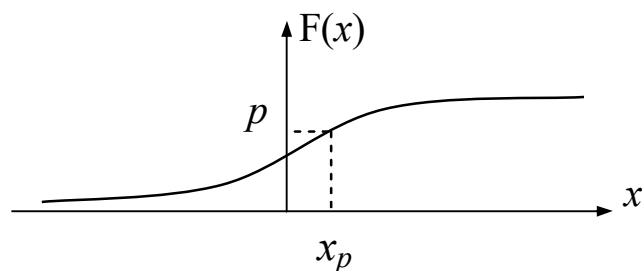
Выборочные коэффициенты эксцесса и асимметрии можно использовать для грубой проверки выборки “на нормальность” (а именно – отклонения гипотезы о нормальности распределения): если

отличие от нуля значения эксцесса ($\hat{\gamma}_{3e}$) или асимметрия ($\hat{\gamma}_{4e}$) оказывается существенным, то гипотезу о нормальности распределения следует отвергнуть.

Заметим, что в ряде статистических модулей прикладных программ (в частности, в Excel) реализованы *несмещенные* выборочные оценки числовых характеристик распределений, в том числе, асимметрии и эксцесса.

7. Выборочные квантили

Напомним: *квантилью порядка p* одномерного распределения называется корень x_p уравнения $F(x) = p$, где $F(x)$ – функция распределения.



Вообще функция распределения $F(x)$ – неубывающая. Если она строго монотонна, то уравнение $F(x) = p$ имеет единственный корень x_p ; в противном случае при некоторых значениях p это уравнение может иметь более одного решения, тогда в качестве квантили x_p берут наименьшее из них. Заметим, что порядок p квантили x_p – это вероятность того, что случайная величина X примет значение *левее* точки x_p : $P(X < x_p) = p$.

Выборочной квантилью $Z_{n,p}$ порядка p называется следующая статистика:

$$Z_{n,p} = \begin{cases} X_{([np]+1)}, & \text{если } np - \text{ дробное} \\ X_{(np)}, & \text{если } np - \text{ целое} \end{cases} \quad \text{здесь } [np] - \text{ целая часть } np.$$

Напомним: целой частью данного числа называют наибольшее целое число, не превосходящее данное; $X_{(k)}$ – k -я порядковая статистика (элемент вариационного ряда).

Из определения следует, что $Z_{n,p}$ – это максимальная из порядковых статистик, обладающая свойством: *левее* нее располагаются члены вариационного ряда, доля которых $\frac{[np]}{n}$ *не превосходит* p .

Таким образом, выборочная квантиль является статистическим аналогом квантили x_p исследуемой случайной величины X .

Частные случаи.

Значению $p = 1/2$ отвечает выборочная медиана

$$Z_{n,1/2} = Med = \begin{cases} X_{([n/2]+1)}, & \text{если } n/2 \text{ - дробное} \\ X_{(n/2)}, & \text{если } n/2 \text{ - целое} \end{cases}.$$

Выборочная медиана Med – оценка медианы MeX распределения случайной величины X (напомним: $P(X < MeX) = P(X > MeX) = 1/2$). Реализацию med выборочной медианы Med вычисляют по реализации вариационного ряда $x_{(1)}, x_{(2)}, \dots, x_{(n)}$.

б) Значениям $p = 1/4$ и $p = 3/4$ отвечают выборочные *квартили* $Z_{n,1/4}$ и $Z_{n,3/4}$ (оценки нижней и верхней квартилей $x_{1/4}$ и $x_{3/4}$), их реализации обозначают $z_{n,1/4}$ и $z_{n,3/4}$, соответственно.

Замечание При наличии выбросов при измерениях или в случае “зашумленных” выборок в качестве оценок математического ожидания MX и дисперсии DX симметричных распределений, целесообразным оказывается использование также оценок, перечисленных ниже.

Оценки MX (положение центра распределения):

Med – выборочная медиана,

$\hat{\theta}_R = \frac{1}{2} (X_{(1)} + X_{(n)})$ – среднее арифметическое экстремальных статистик.

$\hat{\theta}_Q = \frac{1}{2} (Z_{n,1/4} + Z_{n,3/4})$ – среднее арифметическое выборочных квартилей.

Оценки DX (мера рассеяния распределения):

$D = \frac{1}{n} \sum_{i=1}^n |X_i - Med|$ – среднее абсолютное отклонение,

$R = X_{(n)} - X_{(1)}$ – размах,

$Q = Z_{n,3/4} - Z_{n,1/4}$ – интерквартильная широта.

8. Нахождение оценок параметров распределений методом максимального правдоподобия

Пусть X_1, \dots, X_n – выборка из распределения $F_X(x, \theta)$, зависящего от одного неизвестного параметра θ и стоит задача построить оценку этого параметра. Один из методов нахождения оценок параметров непрерывных и дискретных распределений – *метод максимального правдоподобия*.

а) пусть X – непрерывная исследуемая случайная величина, X_1, \dots, X_n – выборка из распределения с плотностью вероятности $f_X(x, \theta)$, зависящей от неизвестного параметра, причем вид функции f известен, и x_1, x_2, \dots, x_n – некоторая реализация выборки.

Функция $L(x_1, x_2, \dots, x_n, \theta) = f_X(x_1, \theta) \dots f_X(x_n, \theta) = \prod_{i=1}^n f_X(x_i, \theta)$,

рассматриваемая в фиксированной точке (x_1, x_2, \dots, x_n) как функция параметра θ , называется функцией правдоподобия.

Вероятностный смысл этой функции – значение плотности вероятности n -мерной случайной величины X_1, \dots, X_n , вычисленное в данной точке x_1, x_2, \dots, x_n и зависящее от параметра θ (говорят – *апостериорное* значение плотности вероятности).

б) пусть теперь x_1, x_2, \dots, x_n – некоторая реализация выборки X_1, \dots, X_n из распределения дискретной случайной величины,

множество возможных значений которой $\{x_i\} \quad i=1,2,\dots$, причем распределение $P(X=x_i)=p_i(\theta)$ зависит от параметра θ .

Пусть в данной реализации x_1, x_2, \dots, x_n значение x_m встречается n_m раз (здесь $m=1, 2, \dots, k$; причем $n_1 + \dots + n_k = n$).

В случае дискретного распределения функцию правдоподобия определяют так:

$$L(x_1, x_2, \dots, x_n, \theta) = p_1^{n_1}(\theta) \dots p_k^{n_k}(\theta) = \prod_{m=1}^k p_m^{n_m}(\theta)$$

Вероятностный смысл функции правдоподобия для случая дискретного распределения состоит в следующем: это вероятность того, что случайная выборка X_1, \dots, X_n примет значение, равное именно данной реализации выборки x_1, x_2, \dots, x_n .

Понятно, что чем ближе значение переменной θ к истинному (неизвестному) значению параметра распределения $F_X(x, \theta)$, тем выше вероятность при проведении эксперимента получить данную реализацию выборки x_1, x_2, \dots, x_n .

Оценкой максимального правдоподобия $\hat{\theta}_{МП}$ неизвестного параметра θ (точнее – значением оценки, отвечающим данной конкретной реализации выборки x_1, x_2, \dots, x_n) называется такое значение переменной θ , которое доставляет максимум функции правдоподобия $L(x_1, x_2, \dots, x_n, \theta)$.

Функция правдоподобия L , определенная выше, представляет собой произведение ряда сомножителей, поэтому при поиске точки максимума L целесообразно перейти к $\ln L$ (очевидно, что $\ln L$ и L имеют максимум в одной и той же точке) и оценку максимального правдоподобия $\hat{\theta}_{МП}$ параметра θ находить из уравнения правдоподобия $\frac{\partial}{\partial \theta}(\ln L) = 0$.

В случае, когда неизвестными являются m параметров $\theta_1, \dots, \theta_m$, оценки $\hat{\theta}_{1МП}, \dots, \hat{\theta}_{mМП}$ находят из соответствующей системы m уравнений.

Заметим, что метод максимального правдоподобия всегда приводит к состоятельным оценкам, распределенным асимптотически нормально, имеющим наименьшую возможную дисперсию среди других асимптотически нормальных оценок. Однако на практике он может осложняться трудностями, связанными с решением систем уравнений правдоподобия.

9. Примеры нахождения оценок максимального правдоподобия (МП-оценок) параметров распределений

Пример 1_мп (нормальное распределение).

Пусть имеется выборка X_1, \dots, X_n из нормального распределения $N(m; \sigma)$ и x_1, x_2, \dots, x_n – некоторая реализация выборки. Найдем оценки максимального правдоподобия $\hat{m}_{МП}$ и $\hat{\sigma}_{МП}$ параметров распределения m и σ .

Функция правдоподобия в точке (x_1, x_2, \dots, x_n) равна

$$L(x_1, x_2, \dots, x_n, m, \sigma) = \prod_{i=1}^n f_X(x_i, m, \sigma) = \prod_{i=1}^n \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-m)^2}{2\sigma^2}} \right),$$

$$\text{ее логарифм } \ln L = \left(-\frac{n}{2}\right) \ln 2\pi - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2.$$

Решая систему уравнений правдоподобия

$$\frac{\partial \ln L}{\partial m} = 0, \quad \frac{\partial \ln L}{\partial \sigma} = 0 \text{ относительно неизвестных } m \text{ и } \sigma,$$

$$\text{получаем: } m = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}; \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 = s^2.$$

Полученное решение – значение реализаций оценок параметров m и σ , соответствующее данной реализации выборки, однако все

приведенные рассуждения справедливы для любой реализации выборки, поэтому искомые оценки равны, соответственно:

$$\hat{m}_{\text{МП}} = \bar{X}; \hat{\sigma}_{\text{МП}} = S^2.$$

Пример 2_мп (распределение Пуассона)

Найдем оценку максимального правдоподобия $\hat{a}_{\text{МП}}$ параметра a распределения Пуассона $P(X = i) = p_i = \frac{a^i}{i!} e^{-a}$ ($i = 0, 1, 2, \dots$).

Пусть x_1, x_2, \dots, x_n – некоторая реализация выборки X_1, \dots, X_n из распределения Пуассона, так что числа x_1, x_2, \dots, x_n – целые неотрицательные. Обозначим через k наибольшее из них и подсчитаем число раз, которое каждое из чисел $0, 1, \dots, k$ встретилось в данной реализации выборки:

$0 - n_0$ раз, $1 - n_1$ раз, ..., $m - n_m$ раз, ..., $k - n_k$ раз,

$$\text{при этом } \sum_{m=0}^k n_m = n, \quad \sum_{m=0}^k m n_m = \sum_{i=1}^n x_i.$$

$$\text{Далее, } L(x_1, x_2, \dots, x_n, a) = \prod_{m=0}^k p_m^{n_m}(a), \quad \ln L = \sum_{m=0}^k n_m \ln p_m(a),$$

$$\frac{\partial \ln L}{\partial a} = \sum_{m=0}^k n_m \left(\frac{m}{a} - 1\right) = 0, \quad \frac{1}{a} \sum_{m=0}^k m n_m - \sum_{m=0}^k n_m = 0, \quad \frac{1}{a} \sum_{i=1}^n x_i = n,$$

$$\text{откуда } a = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$

Таким образом, оценка максимального правдоподобия параметра a распределения Пуассона равна $\hat{a}_{\text{МП}} = \bar{X}$. Заметим, что эта оценка несмещенная состоятельная асимптотически нормальная.

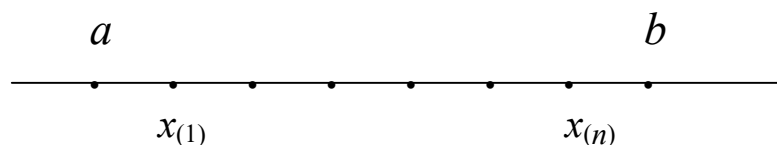
Пример 3_мп (равномерное распределение).

Пусть x_1, x_2, \dots, x_n – некоторая реализация выборки X_1, \dots, X_n из равномерного распределения с параметрами a и b , а $x_{(1)} \leq \dots \leq x_{(n)}$ –

соответствующая реализация вариационного ряда. Найдем оценки максимального правдоподобия параметров a и b .

Функция правдоподобия в точке (x_1, x_2, \dots, x_n) равна:

$$- L(x_1, x_2, \dots, x_n; a, b) = \left(\frac{1}{b-a}\right)^n, \text{ если } \forall i x_i \in [a; b]$$



- $L(x_1, x_2, \dots, x_n; a, b) = 0$, если хотя бы одно число из x_1, x_2, \dots, x_n лежит вне $[a; b]$.

Ясно, что функция правдоподобия $L(x_1, x_2, \dots, x_n; a, b) = \left(\frac{1}{b-a}\right)^n$ максимальна (как функция параметров a и b) при условии, что величина разности $b-a$ минимальна. Таким образом, поскольку $a \leq x_{(1)}$, $b \geq x_{(n)}$, то значения параметров, доставляющие максимум функции правдоподобия $a = x_{(1)}$, $b = x_{(n)}$, а искомые оценки $\hat{a}_{МП} = X_{(1)}$, $\hat{b}_{МП} = X_{(n)}$.

Пример 4_мп (распределение Бернулли)

Найдем оценку максимального правдоподобия вероятности p (вероятности успеха) в каждом испытании при проведении n независимых испытаний по схеме Бернулли. Индикатор X_i появления успеха в i -м испытании – случайная величина, принимающая два возможных значения 1 или 0, а именно, $P(X_i=1) = p$, если в результате i -го испытания осуществился успех и $P(X_i=0) = 1-p = q$ если результат i -го испытания – неуспех (распределение Бернулли):

X_i	0	1
P	q	p

Пусть в результате данной серии n испытаний получена реализация выборки из распределения Бернулли, в которой значение 1 встретилось точно m раз, а значение 0 соответственно $n-m$ раз (ровно m успехов в n испытаниях). Функция правдоподобия имеет вид: $L(x_1, x_2, \dots, x_n, p) = p^m q^{n-m} = p^m (1-p)^{n-m}$.

$$\frac{\partial \ln L}{\partial p} = \frac{m}{p} - \frac{n-m}{1-p} = 0, \text{ откуда } p = \frac{m}{n}.$$

Таким образом, искомой оценкой максимального правдоподобия вероятности p является относительная частота $\hat{p}_{МП} = \frac{X}{n}$ числа успехов при проведении n независимых испытаний.

10. Понятие доверительного интервала

Пусть θ – некоторый неизвестный параметр распределения. По выборке X_1, \dots, X_n из данного распределения построим интервальную оценку параметра θ распределения, то есть найдем такой интервал, внутри которого с заданной (высокой) вероятностью $1-\alpha$ находится истинное значение неизвестного параметра θ . Указанную вероятность $1-\alpha$ называют доверительной вероятностью, а величину α – уровнем значимости.

В качестве значений доверительной вероятности обычно выбирают величины 0,9, 0,95, 0,99, достаточно близкие к 1. В каждом конкретном случае выбор величины доверительной вероятности определяется спецификой решаемой практической задачи.

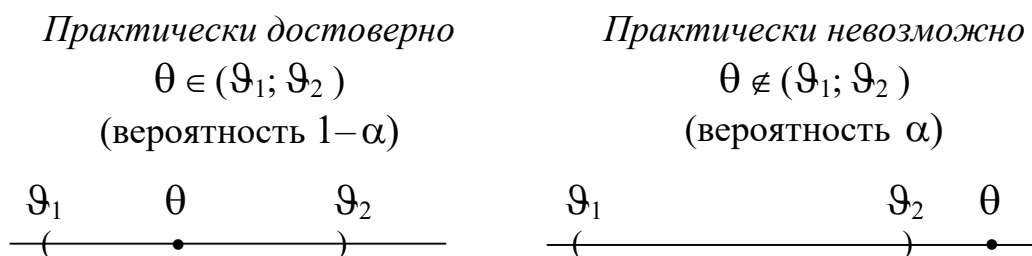
Итак, пусть X_1, \dots, X_n – выборка из данного распределения и задана величина доверительной вероятности $1-\alpha$. Интервал $(\vartheta_1; \vartheta_2)$ называют доверительным интервалом для параметра θ , отвечающим доверительной вероятности $1-\alpha$, если его границы $\vartheta_1 = \vartheta_1(X_1, \dots, X_n)$

и $\vartheta_2 = \vartheta_2(X_1, \dots, X_n)$ – две статистики такие, что верно равенство:

$$P(\vartheta_1 < \theta < \vartheta_2) = 1 - \alpha.$$

Заметим, что границы доверительного интервала ϑ_1 и ϑ_2 – случайные величины (функции выборки X_1, \dots, X_n), параметр θ – неслучайная величина, так что интервал $(\vartheta_1; \vartheta_2)$ “накрывает” величину θ с вероятностью $1 - \alpha$ (соответственно, “не накрывает” с вероятностью α).

Длина интервала $\vartheta_2 - \vartheta_1$ характеризует точность, а доверительная вероятность $1 - \alpha$ – надежность интервальной оценки. Очевидно, что точность и надежность взаимосвязаны: увеличение надежности приводит к уменьшению точности – увеличению длины интервала $(\vartheta_2 - \vartheta_1)$. Выбирая величину доверительной вероятности $1 - \alpha$, принимают соглашение: считать события, вероятность которых $P \geq 1 - \alpha$, – практически достоверными, а события, вероятность которых $P \leq \alpha$ – практически невозможными.



11. Основные этапы процедуры построения доверительных интервалов

Напомним, что границы доверительного интервала $\vartheta_1(X_1, \dots, X_n)$ и $\vartheta_2(X_1, \dots, X_n)$ – случайные величины (функции выборки X_1, \dots, X_n). Результатом эксперимента (серии n независимых измерений данной случайной величины) является реализация выборки x_1, x_2, \dots, x_n . Соответственно, значения статистик ϑ_1 и ϑ_2 в точке (x_1, x_2, \dots, x_n) – это числа $\vartheta_{1e} = \vartheta_1(x_1, x_2, \dots, x_n)$ и $\vartheta_{2e} = \vartheta_2(x_1, x_2, \dots, x_n)$.

Таким образом, будем различать, идет ли речь о *доверительном интервале*, границы которого по смыслу – случайные величины, или о *реализации доверительного интервала*, границами которого являются конкретные числа.

Для построения реализации доверительного интервала на основе данной реализации выборки x_1, x_2, \dots, x_n выполняют следующие действия:

(1) формулируют предположения о распределении и о выборке X_1, \dots, X_n (допущения, принимаемые при построении априорной теоретической модели).

(2) строят доверительный интервал, для чего:

а) выбирают значение доверительной вероятности $1-\alpha$ (или уровня значимости α), то есть принимают соглашение:

– *вероятность практически достоверного события* = $1-\alpha$;

– *вероятность практически невозможного события* = α .

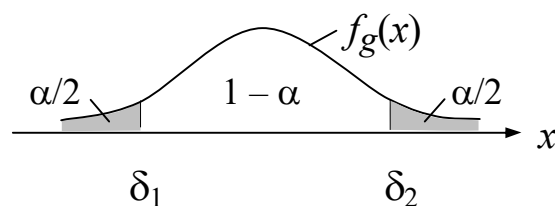
б) записывают вероятностное равенство:

$$P(\delta_1 < g(\theta; \hat{\theta}) < \delta_2) = \int_{\delta_1}^{\delta_2} f_g(x) dx = 1-\alpha,$$

где статистика g имеет известную (табулированную) плотность вероятности $f_g(x)$, θ – оцениваемый параметр, $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ – некоторая его оценка.

Существует бесконечное множество значений величин δ_1 и δ_2 , обеспечивающих справедливость указанного равенства, однако, если использовать дополнительное условие

$P(g(\theta; \hat{\theta}) < \delta_1) = P(g(\theta; \hat{\theta}) > \delta_2) = \alpha/2$, то решение $(\delta_1; \delta_2)$ будет единственным:



в) преобразуют неравенство $\delta_1 < g(\theta; \hat{\theta}) < \delta_2$ к виду $\vartheta_1 < \theta < \vartheta_2$, где $\vartheta_1 = \vartheta_1(X_1, \dots, X_n)$, $\vartheta_2 = \vartheta_2(X_1, \dots, X_n)$ – функции выборки, тогда

$$P(\vartheta_1 < \theta < \vartheta_2) \equiv 1 - \alpha.$$

(3) проводят эксперимент – получают конкретную реализацию выборки x_1, x_2, \dots, x_n .

(4) вычисляют значения $\vartheta_{1e} = \vartheta_1(x_1, x_2, \dots, x_n)$ и $\vartheta_{2e} = \vartheta_2(x_1, x_2, \dots, x_n)$.

В результате перечисленных действий (1) – (4) получают реализацию доверительного интервала – числовой интервал $(\vartheta_{1e}; \vartheta_{2e})$.

Степень уверенности в том, что полученный интервал $(\vartheta_{1e}; \vartheta_{2e})$ в действительности содержит неизвестный параметр θ выражается *выбранной априори* величиной доверительной вероятности $1 - \alpha$ (вероятностью практически достоверного события). Иными словами, априори допускается, что утверждение “данная реализация доверительного интервала $(\vartheta_{1e}; \vartheta_{2e})$ содержит оцениваемый параметр θ ” может оказаться ошибочным, однако, число таких случаев мало и может наблюдаться лишь в $\alpha \cdot 100\%$ общего числа случаев; при этом приемлемая доля указанных ошибок выражается уровнем значимости α – вероятностью практически невозможного события.

12. Доверительный интервал для математического ожидания нормального распределения при известной дисперсии

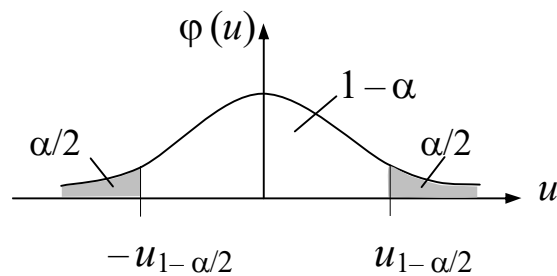
Пример 1_ди

Пусть X_1, \dots, X_n – выборка из нормального распределения $N(m; \sigma)$. Предполагается, что параметр σ известен: $\sigma = \sigma_0$ (например, этот параметр определен в результате специальных многократных измерений).

При указанных предположениях справедливо: $\bar{X} \sim N(m; \sigma_0 / \sqrt{n})$

или иначе, $\frac{\bar{X} - m}{\sigma_0 / \sqrt{n}} \sim N(0; 1)$.

Построим доверительный интервал для математического ожидания m . Зададим величину доверительной вероятности $1 - \alpha$ и запишем вероятностное равенство $P\left(\left|\frac{\bar{X} - m}{\sigma_0 / \sqrt{n}}\right| < u\right) = 1 - \alpha$, откуда имеем: $u = u_{1-\alpha/2}$ – квантиль порядка $1 - \alpha/2$ (см. рисунок ниже, где $\varphi(u)$ – плотность вероятности стандартного нормального распределения).



Далее, $P\left(\left|\frac{\bar{X} - m}{\sigma_0 / \sqrt{n}}\right| < u_{1-\alpha/2}\right) = 1 - \alpha$, откуда

$$P\left(\bar{X} - \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2} < m < \bar{X} + \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2}\right) = 1 - \alpha.$$

Таким образом, искомый доверительный интервал, отвечающий доверительной вероятности $1 - \alpha$ имеет вид:

$$\left(\bar{X} - \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2}; \bar{X} + \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2}\right)$$

Заметим, что для данной величины $1 - \alpha$ доверительной вероятности длина этого интервала равна $d = 2 \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2}$.

При данном объеме выборки n длина d постоянна, от выборки к выборке меняется только положение центра интервала \bar{X} .

Замечание

Проиллюстрируем связь точности и надежности интервальной оценки при фиксированных значениях n и σ_0 :

$$1 - \alpha = 0,9 \quad u_{1-\alpha/2} = u_{0,95} = 1,65 \quad d = 3,3 \frac{\sigma_0}{\sqrt{n}};$$

$$1 - \alpha = 0,95 \quad u_{1-\alpha/2} = u_{0,975} = 1,96 \quad d = 3,92 \frac{\sigma_0}{\sqrt{n}};$$

$$1 - \alpha = 0,99 \quad u_{1-\alpha/2} = u_{0,995} = 2,58 \quad d = 5,16 \frac{\sigma_0}{\sqrt{n}}.$$

Видим, что при данном объеме выборки n с ростом надежности оценки $1 - \alpha$ ее точность убывает (длина d соответствующего интервала растет). Таким образом, как уже упоминалось, плата за повышение надежности – уменьшение точности интервальной оценки.

Пусть заданы $\varepsilon > 0$ и $1 - \alpha$ – величины, характеризующие соответственно точность и надежность оценки. Найдем объем выборки, достаточный для обеспечения одновременно заданных значений точности и надежности оценки. Из условия $d \leq 2\varepsilon$ получаем:

$$n \geq \left(\frac{\sigma_0}{\varepsilon} u_{1-\alpha/2} \right)^2.$$

Пример 1_ди (доверительный интервал для m при известном σ , $1 - \alpha = 0,95$)

Точность прибора известна (в паспорте прибора указано $\sigma_0 = 0,02$). С помощью этого прибора проведено n независимых повторных измерений ($n = 25$) некоторой физической величины. По результатам измерений x_1, x_2, \dots, x_{25} вычислено среднее выборочное, оказавшееся равным $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 2,42$. Считая числа x_1, x_2, \dots, x_{25} реализацией выборки из нормального распределения $N(m; \sigma_0)$, где $\sigma_0 = 0,02$, найти

реализацию доверительного интервала для математического ожидания m ; доверительная вероятность $1 - \alpha = 0,95$.

Решение

Выражение для реализации доверительного интервала для математического ожидания m нормального распределения (при $\sigma = \sigma_0$), отвечающего доверительной вероятности $1 - \alpha$, имеет вид:

$$\left(\bar{x} - \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2}; \bar{x} + \frac{\sigma_0}{\sqrt{n}} u_{1-\alpha/2} \right).$$

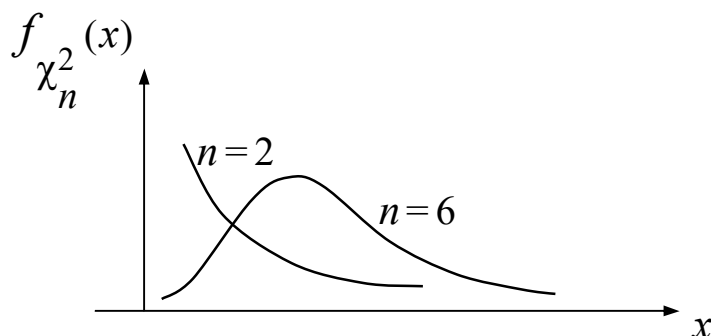
По таблице квантилей стандартного нормального распределения находим квантиль порядка $1 - \alpha/2$: $u_{1-\alpha/2} = u_{0,975} = 1,96$. Подставляя $\bar{x} = 2,42$, $\sigma_0 = 0,02$, $n = 25$ в выражение для реализации доверительного интервала, получаем:

$$(2,420 - 0,008; 2,420 + 0,008) = (2,412; 2,428).$$

13. Распределение χ^2 , распределение Стьюдента, лемма Фишера

Распределение χ^2

Сумма квадратов n независимых стандартных нормальных случайных величин $\sum_{i=1}^n X_i^2 = \chi_n^2$ называется случайной величиной χ_n^2 (с n степенями свободы). Плотность вероятности распределения χ_n^2 табулирована, ее график имеет вид, представленный на рисунке:



Для распределения χ_n^2 имеют место следующие соотношения: $M[\chi_n^2] = n$, $D[\chi_n^2] = 2n$, $Mo[\chi_n^2] = n - 2$. Заметим, что с ростом n кривая

$f_{\chi_n^2}(x)$ становится более симметричной, а ее максимум смещается вправо.

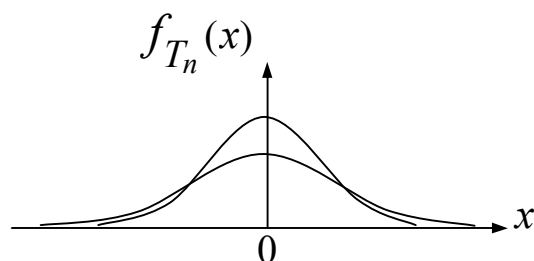
Заметим также, что в силу центральной предельной теоремы центрированная и нормированная случайная величина $\frac{\chi_n^2 - n}{\sqrt{2n}} \underset{n \rightarrow +\infty}{\sim} N(0;1)$ асимптотически нормальна, поэтому в таблицах для распределения χ_n^2 приводятся квантили только для $n \leq 30$.

Распределение Стьюдента

Пусть случайные величины $X \sim N(0; 1)$ и χ_n^2 – независимы.

Случайная величина $T_n = \frac{X}{\sqrt{\frac{\chi_n^2}{n}}}$ называется отношением Стьюдента

(t -отношением). Плотность вероятности распределения Стьюдента табулирована, ее график имеет вид, представленный на рисунке:



Распределение $f_{T_n}(x)$ симметрично, $M[T_n]=0$ и имеет место асимптотическое свойство: $T_n \underset{n \rightarrow +\infty}{\sim} N(0; 1)$.

При малых значениях n распределение Стьюдента заметно отличается от стандартного нормального распределения, однако при $n > 30$ эти распределения близки.

Лемма Фишера (о совместном распределении \bar{X} и S^2 для выборки из нормального распределения)

Пусть X_1, \dots, X_n – выборка из нормального распределения $N(m; \sigma)$ тогда:

- выборочное среднее \bar{X} и выборочная дисперсия S^2 (или исправленная выборочная дисперсия S^{*2}) – взаимно независимы;
- выборочное среднее \bar{X} подчиняется нормальному распределению: $\bar{X} \sim N(m; \sigma / \sqrt{n})$;
- случайная величина $\frac{nS^2}{\sigma^2}$ (или $\frac{(n-1)S^{*2}}{\sigma^2}$) распределена по закону χ_{n-1}^2 (с $n-1$ степенью свободы).

Заметим, что из леммы Фишера следует независимость \bar{X} и $\frac{nS^2}{\sigma^2}$, а также \bar{X} и $\frac{(n-1)S^{*2}}{\sigma^2}$.

14. Доверительный интервал для математического ожидания нормального распределения при неизвестной дисперсии **Примеры 2_ди и 3_ди**

Пусть X_1, \dots, X_n – выборка из нормального распределения $N(m; \sigma)$, параметры которого m и σ неизвестны. Найдем интервальную оценку параметра m , отвечающую заданной величине доверительной вероятности $1 - \alpha$.

Согласно лемме Фишера статистики $U = \frac{\bar{X} - m}{\sigma / \sqrt{n}} \sim N(0; 1)$ и

$\frac{(n-1)S^{*2}}{\sigma^2} = \chi_{n-1}^2$ – независимы. Составим отношение

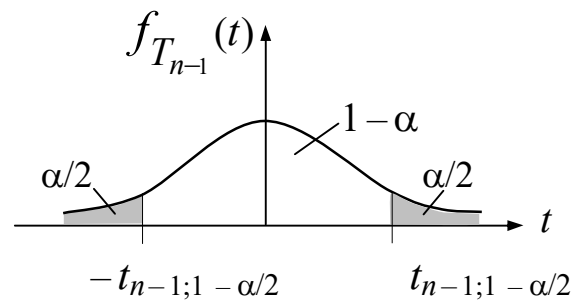
$$T_{n-1} = \frac{U}{\sqrt{\frac{\chi_{n-1}^2}{n-1}}} = \frac{\bar{X} - m}{S^* / \sqrt{n}} \text{ – отношение Стьюдента.}$$

Из вероятностного равенства $P(|T_{n-1}| < t) = 1 - \alpha$ имеем:

$t = t_{n-1; 1-\alpha/2}$ – квантиль порядка $1 - \alpha/2$ распределения Стьюдента.

Далее, $P\left(\left|\frac{\bar{X} - m}{S^* / \sqrt{n}}\right| < t_{n-1; 1-\alpha/2}\right) \equiv 1 - \alpha$, откуда

$$P\left(\bar{X} - \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2} < m < \bar{X} + \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}\right) \equiv 1 - \alpha.$$



$$P\left(\bar{X} - \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2} < m < \bar{X} + \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}\right) \equiv 1 - \alpha.$$

Таким образом, искомый доверительный интервал, отвечающий доверительной вероятности $1 - \alpha$ имеет вид:

$$\left(\bar{X} - \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}; \bar{X} + \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}\right).$$

Заметим, что положение центра этого интервала, как и его длина – случайные величины.

Пример 2_ди (доверительный интервал для m при неизвестном σ , $1 - \alpha = 0,6827$)

По результатам пяти измерений вычислено выборочное среднее \bar{x} и исправленная выборочная дисперсия S^{*2} , соответственно:

$$\bar{x} = \frac{1}{5} \sum_{i=1}^5 x_i = 4 \text{ и } S^{*2} = \frac{1}{4} \sum_{i=1}^5 (x_i - \bar{x})^2 = 10/4.$$

Считая числа x_1, x_2, \dots, x_5 реализацией выборки из нормального распределения $N(m; \sigma)$, где σ неизвестно, записать реализацию доверительного интервала для математического ожидания m , приняв величину доверительной вероятности равной $1 - \alpha = 0,6827$.

Решение

Реализация доверительного интервала, отвечающего доверительной вероятности $1 - \alpha$, для математического ожидания m нормального распределения имеет вид:

$$\left(\bar{x} - \frac{s^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}; \bar{x} + \frac{s^*}{\sqrt{n}} t_{n-1; 1-\alpha/2} \right)$$

Здесь $\frac{s^*}{\sqrt{n}}$ – стандартная ошибка, $\frac{s^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}$ – погрешность.

Правило записи реализации доверительного интервала (для $1 - \alpha = 0,6827$):

- 1) округление значения погрешности – округляемая цифра отбрасывается, если она < 5 , если она ≥ 5 то в предыдущий разряд добавляется единица;
- 2) в погрешности оставляют одну значащую цифру, если она ≥ 4 и две, если первая из них < 4 ;
- 3) в значении \bar{x} оставляют последнюю значащую цифру в том же разряде, что и в погрешности;
- 4) общий множитель вида 10^k выносят за скобки.

Находим (по таблицам) значение квантили $t_{n-1; 1-\alpha/2} = t_{4; 0,84135} = 1,22$, вычисляем значение погрешности $\frac{s^*}{\sqrt{5}} t_{4; 0,84135} = 0,9$ и получаем искомую реализацию доверительного интервала: $(4,0 - 0,9; 4,0 + 0,9) = (3,1; 4,9)$.

Пример 3_ди (доверительный интервал для m при неизвестном σ , $1 - \alpha = 0,95$)

Произведено $n=19$ измерений x_1, x_2, \dots, x_{19} некоторой физической величины с помощью прибора, измеряющего с точностью до 0,1.

Вычислено выборочное среднее $\bar{x} = \frac{1}{19} \sum_{i=1}^{19} x_i = 7,07$ и исправленная

выборочная дисперсия $S^{*2} = \frac{1}{18} \sum_{i=1}^{19} (x_i - \bar{x})^2 = 0,041$.

Считая, что x_1, x_2, \dots, x_{19} – реализация выборки из нормального распределения $N(m; \sigma)$, где σ неизвестно, записать реализацию доверительного интервала для математического ожидания m при доверительной вероятности $1 - \alpha = 0,95$.

Решение

Правило записи реализации доверительного интервала для $1 - \alpha = 0,95$:

- 1) выборочное среднее \bar{x} вычисляется с точностью на порядок большей, чем точность измерений;
- 2) выборочное стандартное отклонение S^* (исправленное) – с точностью, на порядок большей, чем точность вычисления среднего;
- 3) правило округления: если округляемая цифра < 5 , то она отбрасывается, если она ≥ 5 – в предыдущий разряд добавляется единица.

Реализация доверительного интервала, отвечающего доверительной вероятности $1 - \alpha$, для математического ожидания m при неизвестном σ имеет вид:

$$\left(\bar{x} - \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2}; \bar{x} + \frac{S^*}{\sqrt{n}} t_{n-1; 1-\alpha/2} \right)$$

Находим по таблицам значение квантили $t_{n-1; 1-\alpha/2} = t_{18; 0,975} = 2,10$, вычисляем значение погрешности

$\frac{s^*}{\sqrt{19}} t_{18; 0,975} = 0,098$ и получаем искомую реализацию доверительного интервала для m :

$$(7,07 - 0,098; 7,07 + 0,098) = (6,972; 7,168).$$

15. Доверительный интервал для дисперсии нормального распределения

Пусть X_1, \dots, X_n – выборка из нормального распределения $N(m; \sigma)$. Параметры распределения m и σ неизвестны. Найдем интервальную оценку дисперсии распределения σ^2 , отвечающую заданной величине доверительной вероятности $1 - \alpha$.

Известно (п. 13), что статистика $\frac{(n-1) S^{*2}}{\sigma^2}$ распределена по закону χ_{n-1}^2 (хи-квадрат с $n-1$ степенью свободы):

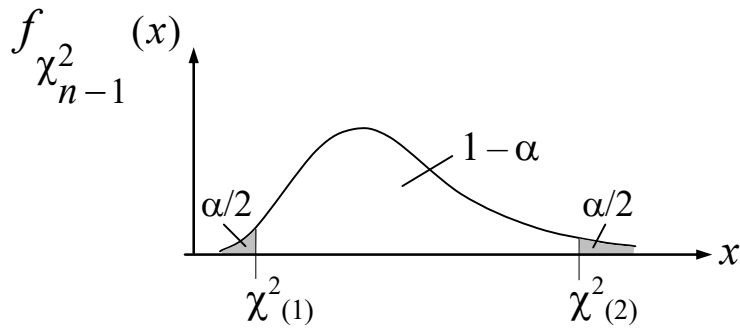
$$\frac{(n-1) S^{*2}}{\sigma^2} = \chi_{n-1}^2$$

Запишем равенство:

$$P(\chi_{(1)}^2 < \frac{(n-1) S^{*2}}{\sigma^2} < \chi_{(2)}^2) = 1 - \alpha,$$

здесь $\chi_{(1)}^2$ и $\chi_{(2)}^2$ – границы доверительного интервала, подлежащие определению. С учетом дополнительного условия $P(\chi_{n-1}^2 < \chi_{(1)}^2) = P(\chi_{n-1}^2 > \chi_{(2)}^2) = \alpha/2$ эти границы определяются единственным образом. Действительно $\chi_{(1)}^2 = \chi_{n-1; \alpha/2}^2$ и $\chi_{(2)}^2 = \chi_{n-1; 1-\alpha/2}^2$ – квантили распределения случайной величины χ_{n-1}^2 порядка $\alpha/2$ и $1 - \alpha/2$, соответственно. В итоге получаем:

$$P\left(\frac{(n-1) S^{*2}}{\chi_{(2)}^2} < \sigma^2 < \frac{(n-1) S^{*2}}{\chi_{(1)}^2}\right) \equiv 1 - \alpha.$$



Замечание

В случае выборки большого объема из произвольного распределения может быть построен приближенный доверительный интервал для стандартного отклонения σ . При этом предполагается, что у распределения, из которого извлечена выборка, существуют конечные первые четыре момента.

В выборочной дисперсии $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ не все слагаемые являются независимыми, так как $\sum_{i=1}^n (X_i - \bar{X}) = 0$. Если объем выборки n достаточно велик, то этой связью можно пренебречь и считать, что все слагаемые независимы. При этом, согласно центральной предельной теореме, центрированная и нормированная случайная величина $\frac{S^2 - MS^2}{DS^2}$ будет распределена асимптотически нормально.

Таким образом примем, что при данном объеме выборки справедливо:

$$\frac{S^2 - MS^2}{\sqrt{D(S^2)}} \sim N(0; 1).$$

Полагая, что выполнены все указанные условия, приведем без доказательства формулу для асимптотического доверительного интервала:

$$S(1 + u_{1-\alpha/2} \sqrt{(\hat{\gamma}_4 + 2)/n})^{-1/2} < \sigma < S(1 - u_{1-\alpha/2} \sqrt{(\hat{\gamma}_4 + 2)/n})^{-1/2}.$$

Здесь $u = u_{1-\alpha/2}$ – квантиль порядка $1-\alpha/2$ стандартного нормального распределения, $\hat{\gamma}_4 = \hat{\mu}_4 / S^4 - 3$ – выборочный эксцесс.

16. Приближенная интервальная оценка для математического ожидания произвольного распределения по выборке большого объема

Пусть X_1, \dots, X_n – выборка из некоторого (произвольного) распределения, причем объем выборки n достаточно велик.

Введем обозначения: $MX_i = m$, $DX_i = \sigma^2$ ($i = 1, \dots, n$), тогда $M\bar{X} = m$, $D\bar{X} = (\sigma/\sqrt{n})^2$.

Согласно центральной предельной теореме имеем: $\frac{\bar{X} - m}{\sigma/\sqrt{n}} \sim N(0;1)$ при $n \rightarrow +\infty$. По условию n – велико, поэтому примем допущение, что

указанная нормальность $\frac{\bar{X} - m}{\sigma/\sqrt{n}} \sim N(0;1)$ имеет место при данном объеме выборки n .

Зададимся величиной доверительной вероятности $1-\alpha$ и запишем: $P\left(\left|\frac{\bar{X} - m}{\sigma/\sqrt{n}}\right| < u\right) = 1-\alpha$, откуда $u = u_{1-\alpha/2}$ – квантиль порядка $1-\alpha/2$ стандартного нормального распределения.

Таким образом, $P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2} < m < \bar{X} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right) \equiv 1-\alpha$.

Заменяя в последнем равенстве неизвестную величину σ ее несмещенной состоятельной оценкой S^* , получаем *приближенную* интервальную оценку для математического ожидания m , отвечающую доверительной вероятности $1-\alpha$ при большом объеме выборки:

$$\left(\bar{X} - \frac{S^*}{\sqrt{n}} u_{1-\alpha/2} < m < \bar{X} + \frac{S^*}{\sqrt{n}} u_{1-\alpha/2}\right).$$

17. Приближенная интервальная оценка вероятности p в схеме Бернулли по выборке большого объема. Пример 4_ди

Пусть проводится n независимых испытаний, p – вероятность успеха в каждом испытании и X_1, \dots, X_n – выборка из соответствующего распределения Бернулли, причем объем выборки n достаточно велик. Каждый элемент выборки – индикатор X_i появления успеха в i -м испытании. Случайная величина $X = \sum_{i=1}^n X_i$ – число успехов – подчиняется биномиальному распределению:

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

Согласно теореме Муавра-Лапласа, центрированная и нормированная случайная величина $\frac{X - np}{\sqrt{npq}} = \frac{X/n - p}{\sqrt{pq/n}}$ является

асимптотически нормальной, что можно символически записать так:

$$\frac{\hat{p} - p}{\sqrt{pq/n}} \sim N(0; 1) \quad \text{при } n \rightarrow +\infty. \quad \text{Здесь относительная частота } \hat{p} = \frac{X}{n}$$

оценка максимального правдоподобия вероятности p (несмещенная состоятельная асимптотически эффективная и асимптотически нормальная точечная оценка), $q = 1 - p$.

Это означает, что при достаточно больших значениях n распределение статистики $\frac{\hat{p} - p}{\sqrt{pq/n}}$ близко к стандартному нормальному. Будем считать, что нормальность распределения имеет место при данном (достаточно большом) объеме выборки n .

Зададимся величиной доверительной вероятности $1 - \alpha$ и запишем: $P\left(\left|\frac{\hat{p} - p}{\sqrt{pq/n}}\right| < u\right) = 1 - \alpha$, откуда $u = u_{1-\alpha/2}$ – квантиль порядка $1 - \alpha/2$ стандартного нормального распределения.

Таким образом, $P\left(\left|\frac{\hat{p}-p}{\sqrt{pq/n}}\right| < u_{1-\alpha/2}\right) \equiv 1-\alpha$,

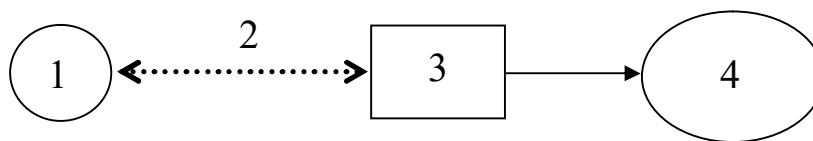
$$P\left(\hat{p}-u_{1-\alpha/2}\sqrt{p(1-p)/n} < p < \hat{p}+u_{1-\alpha/2}\sqrt{p(1-p)/n}\right) \equiv 1-\alpha.$$

Заменяя в последнем тождестве неизвестную величину p ее оценкой \hat{p} , получим выражение для приближенного (асимптотического) доверительного интервала, отвечающего доверительной вероятности $1-\alpha$:

$$\left(\hat{p}-u_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}; \hat{p}+u_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right)$$

Отметим еще раз, что полученное выражение используют для нахождения *приближенной* интервальной оценки вероятности p при достаточно больших значениях n .

Пример 4_ди



Приемное устройство (3) охранной системы находится в режиме ожидания (дежурном режиме) и может получать сообщения (пакеты) фиксированной длительности $\Delta\tau$ от датчика (1) по радиоканалу (2). Если в каком-либо интервале времени $\Delta\tau$ от датчика поступает сообщение, то на исполнительное устройство (4) передается извещение “тревога”.

Под *ложной тревогой* понимают ошибочное формирование извещения “тревога” при условии, что сообщение от датчика отсутствует; ложная тревога обусловлена только наличием собственного шума в системе, в частности – в радиоканале.

Для оценки вероятности ложной тревоги осуществили наблюдение за работой системы в течение интервала времени, равного $10^4\Delta\tau$. Подсчитали частоту события “ложная тревога” (число интервалов $\Delta\tau$,

таких, в которых было ошибочно сформировано извещение “тревога”) и его относительную частоту, оказавшуюся равной 0,006.

Найдем приближенную интервальную оценку вероятности ложной тревоги при доверительной вероятности $1 - \alpha = 0,95$.

Указание: Принять допущение, что вероятностной моделью эксперимента по оценке вероятности p ложной тревоги служит последовательность независимых испытаний (схема Бернулли), то есть, что выборка объема $n = 10^4$ извлечена из распределения Бернулли с параметром p .

Решение

Выражение для реализации приближенного доверительного интервала для неизвестного параметра p , отвечающего доверительной вероятности $1 - \alpha$, имеет вид:

$$(\hat{p}_e - u_{1-\alpha/2} \sqrt{\hat{p}_e(1-\hat{p}_e)/n}; \hat{p}_e + u_{1-\alpha/2} \sqrt{\hat{p}_e(1-\hat{p}_e)/n}),$$

где \hat{p}_e – значение оценки $\hat{p} = X/n$, вычисленное по реализации выборки (по результатам эксперимента); по условию задачи $\hat{p}_e = 0,006$.

По таблицам квантилей стандартного нормального распределения находим квантиль $u_{1-\alpha/2} = u_{0,975} = 1,96$. Вычисляем

$$u_{1-\alpha/2} \sqrt{\hat{p}_e(1-\hat{p}_e)/n} = 1,96 \sqrt{0,006(1-0,006)/10000} = 0,0015,$$

$$\hat{p}_e - u_{1-\alpha/2} \sqrt{\hat{p}_e(1-\hat{p}_e)/n} = 0,006 - 0,0015 = 0,0045,$$

$$\hat{p}_e + u_{1-\alpha/2} \sqrt{\hat{p}_e(1-\hat{p}_e)/n} = 0,006 + 0,0015 = 0,0075.$$

В итоге получаем искомую реализацию приближенного доверительного интервала: $(0,0045; 0,0075)$, соответствующего доверительной вероятности $1 - \alpha = 0,95$.

18. Постановка задачи проверки статистических гипотез

Пример 1_кз

Задачу проверки статистических гипотез рассмотрим на примере.

Пример 1_кз (двусторонний критерий).

В результате многократных измерений некоторого параметра эталонного образца получено значение 2,40 (условных единиц). Точность прибора по паспорту $\sigma_0 = 0,02$. Периодически настройку прибора проверяют и при необходимости корректируют. Прибором некоторое время пользовались, а затем произвели $n = 25$ контрольных независимых измерений x_1, x_2, \dots, x_{25} указанного параметра того же эталонного образца и получили $\bar{x} = 2,42$.

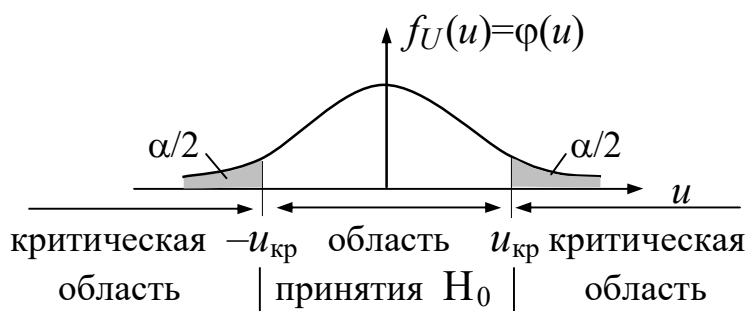
Допустим, что полученные числа x_1, x_2, \dots, x_{25} – реализация выборки X_1, \dots, X_{25} из нормального распределения $N(m; \sigma)$, один из параметров которого известен: $\sigma = \sigma_0 = 0,02$ и проверим, является ли обоснованным предположение: контрольная выборка извлечена из нормального распределения с математическим ожиданием $m = m_0 = 2,40$, а отличие выборочного среднего от математического ожидания естественно объясняется случайностью выборки?

Иными словами, проверим гипотезу $H_0: m = m_0 = 2,40$ о значении математического ожидания нормального распределения $N(m; \sigma_0)$, из которого извлечена выборка. Гипотезу $H_1: m \neq m_0$ рассмотрим как альтернативную к гипотезе H_0 .

При справедливости гипотезы H_0 имеем:

$$\frac{\bar{X} - m_0}{\frac{\sigma_0}{\sqrt{n}}} = U \sim N(0;1).$$

Статистику U назовем *статистикой критерия*. Зададим уровень значимости $\alpha = 0,05$. Всякое событие, вероятность которого $P \leq \alpha$ будем считать практически невозможным при справедливости гипотезы H_0 , а событие, вероятность которого $P \geq 1 - \alpha$ – практически достоверным при справедливости гипотезы H_0 .



При выбранном уровне значимости α так называемое критическое число $u_{кр}$ определяется условиями $P(|U| \geq u_{кр}) = \alpha$, $P(|U| < u_{кр}) = 1 - \alpha$, так что при справедливости гипотезы H_0 событие $(|U| \geq u_{кр})$ – *практически невозможное*, а событие $(|U| < u_{кр})$ – *практически достоверное*. Таким образом, множество значений статистики критерия U разбивается на интервалы, соответствующие практически невозможному событию (*критическая область*), либо практически достоверному событию (*область принятия H_0*). Указанное разбиение позволяет сформулировать следующее правило принятия решения об отклонении или принятии гипотезы H_0 :

- при $(|U| \geq u_{кр})$ гипотеза H_0 отвергается,
- при $(|U| < u_{кр})$ гипотеза H_0 принимается.

При $\alpha = 0,05$ имеем $u_{кр} = u_{1 - \alpha/2} = u_{0,975} = 1,96$.

Обозначим через u_e значение, которое статистика U приняла в результате контрольных измерений x_1, x_2, \dots, x_{25} , тогда

$$u_e = U(x_1, x_2, \dots, x_{25}) = \frac{\bar{x} - m_0}{\sigma_0 / \sqrt{n}} = \frac{2,42 - 2,40}{0,02 / \sqrt{25}} = 5 > u_{кр} = 1,96.$$

Попадание значения u_e в критическую область означает, что произошло событие, практически невозможное при справедливости гипотезы H_0 . Данные измерений x_1, x_2, \dots, x_{25} не согласуются с гипотезой H_0 , а напротив, опровергают ее, поэтому гипотезу $H_0: m = m_0$ следует отклонить (соответственно – принять альтернативную к ней гипотезу $H_1: m \neq m_0$). Практически это означает, что прибор следует настроить заново.

19. Критерии значимости: гипотезы, критическая область, решения, ошибки

Пусть X – исследуемая случайная величина, $F_X(x) = P(X < x)$ – ее функция распределения, зависящая от одного или нескольких параметров, и пусть о параметре распределения или о виде распределения выдвинута и подлежит проверке некоторая гипотеза H_0 (нуль-гипотеза) и указана альтернативная к ней гипотеза H_1 .

Поставим задачу: на основе опытных данных либо отвергнуть гипотезу H_0 , если опытные данные и гипотеза противоречат друг другу, либо принять H_0 , то есть сделать вывод о том, что эта гипотеза согласуется с опытными данными. Таким образом, решение об отклонении гипотезы H_0 или ее принятии будет строиться на основе выборки X_1, \dots, X_n из распределения $F_X(x)$.

Гипотезу называют простой, если она полностью определяет распределение и сложной – в противном случае.

В рассмотренном в п.18 примере вид распределения предполагался известным (нормальное распределение $N(m; \sigma)$), один из параметров которого $\sigma = \sigma_0$ – известен, поэтому нуль-гипотеза $H_0: m = m_0$ о значении другого параметра распределения – *простая*, так как она полностью определяет распределение, в то время как альтернативная к ней гипотеза $H_1: m \neq m_0$ – *сложная*.

Правило принятия решения, согласно которому принимается или отвергается гипотеза H_0 , называют *статистическим критерием*.

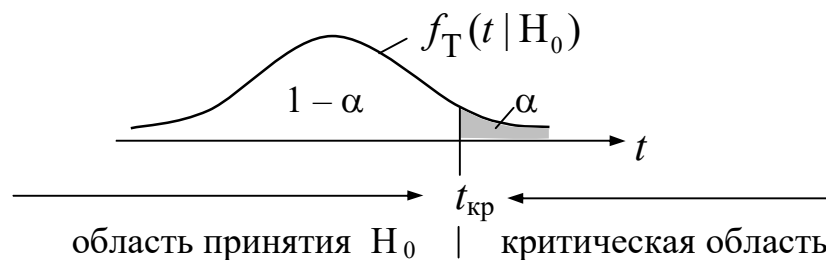
Пусть $\mathcal{X} = \{(X_1, \dots, X_n)\}$ – множество всех возможных значений случайной выборки и задана вероятность α практически невозможного события при справедливости гипотезы H_0 (эту вероятность называют *уровнем значимости* α). При этом множество \mathcal{X} разбивается на два подмножества $\mathcal{X}_{кр}$ – *критическая область* и

$\mathcal{X} \setminus \mathcal{X}_{\text{кр}}$ – область принятия гипотезы H_0 следующим образом:
 $\mathcal{X}_{\text{кр}}$ – подмножество \mathcal{X} такое, что любое событие $((X_1, \dots, X_n) \in \mathcal{X}_{\text{кр}})$,
 вероятность которого $P((X_1, \dots, X_n) \in \mathcal{X}_{\text{кр}} | H_0) = \alpha$ – это практически
 невозможное события при справедливости гипотезы H_0 .
 Соответственно, $P((X_1, \dots, X_n) \in \mathcal{X} \setminus \mathcal{X}_{\text{кр}} | H_0) = 1 - \alpha$ – вероятность
 практически достоверного события. Если в результате опыта
 происходит событие $((X_1, \dots, X_n) \in \mathcal{X}_{\text{кр}})$, то гипотезу H_0 отвергают;
 если же происходит событие $((X_1, \dots, X_n) \in \mathcal{X} \setminus \mathcal{X}_{\text{кр}})$, то H_0 принимают.

Критерий (правило принятия решения об отклонении или
 принятии гипотезы H_0) строится на основе соответствующей
 случайной величины – *статистики критерия* $T(X_1, \dots, X_n)$. При этом
 предполагается, что распределение статистики критерия
 $f_T(t | H_0)$ известно при справедливости гипотезы H_0 . Статистика
 критерия отображает множество $\mathcal{X} = \{(X_1, \dots, X_n)\}$ на числовую
 прямую. По распределению статистики критерия $f_T(t | H_0)$ и
 заданному уровню значимости α находят так называемые
критические числа, которые разбивают все множество значений
 статистики критерия на интервалы, соответствующие либо принятию
 гипотезы H_0 , либо ее отклонению. Понятно, что отклонение H_0
 означает принятие альтернативной гипотезы H_1 при данном уровне
 значимости α .

Рассмотрим примеры.

Односторонний (правосторонний) критерий



В случае правостороннего критерия критическое число $t_{кр}$ определяется соотношениями:

$$P(T(X_1, \dots, X_n) \geq t_{кр} | H_0) = \alpha,$$

$$P(T(X_1, \dots, X_n) < t_{кр} | H_0) = 1 - \alpha.$$

Пусть x_1, x_2, \dots, x_n – реализация выборки X_1, \dots, X_n . Обозначим через t_e значение статистики критерия в точке (x_1, x_2, \dots, x_n) : $t_e = T(x_1, x_2, \dots, x_n)$.

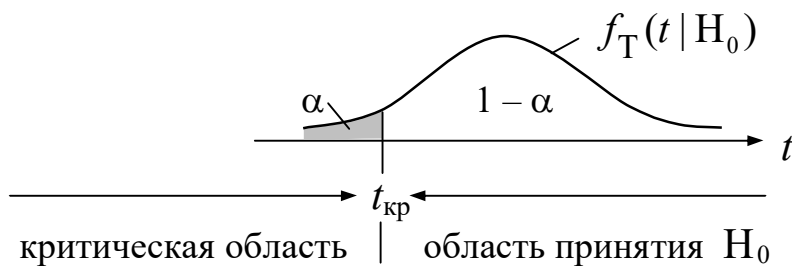
Правило принятия решения в случае правостороннего критерия формулируют так:

если $t_e < t_{кр}$, то гипотезу H_0 принимают,

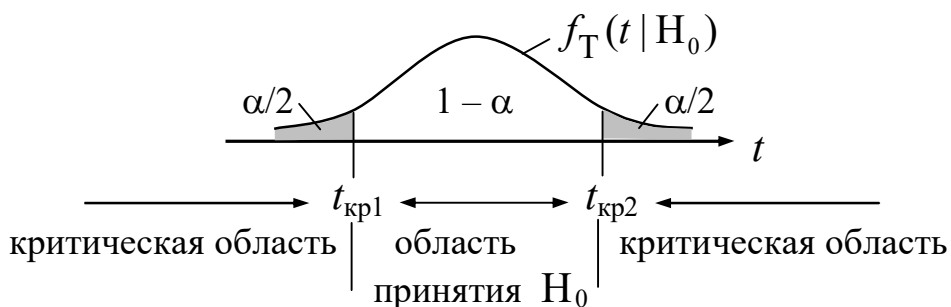
если $t_e \geq t_{кр}$, то гипотезу H_0 отвергают.

Аналогично строятся левосторонний и двусторонний критерии.

Односторонний (левосторонний) критерий



Двусторонний критерий



Описанные критерии называют *критериями значимости*. Если в результате опыта наблюдается значение статистики критерия, попадающее в критическую область, то такой результат можно

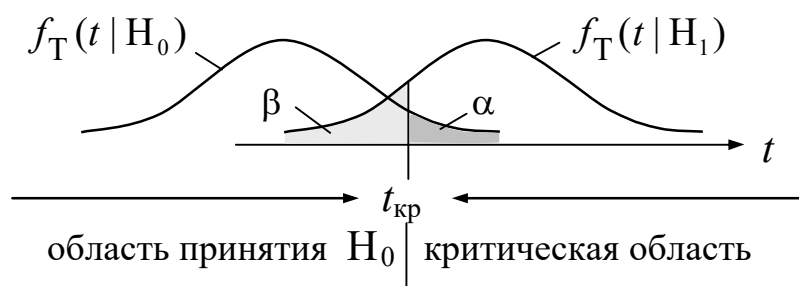
рассматривать как *значимое опровержение* гипотетического согласия между результатами наблюдений и проверяемой гипотезой.

Замечания

1. Из гипотезы H_0 логически следует, что при проведении эксперимента (извлечении реализации выборки x_1, x_2, \dots, x_n) значение t_e *практически неизбежно* должно попадать в область принятия H_0 . Если же в результате эксперимента произошло событие, *практически невозможное при справедливости H_0* (значение t_e попало в критическую область), то гипотеза H_0 должна быть отвергнута.
2. Таким образом, проверка гипотезы H_0 осуществляется косвенно. Доказать справедливость гипотезы H_0 косвенным образом нельзя, так как правильное заключение может следовать и из неверной посылки, однако отвергнуть H_0 – можно.
3. Принятие H_0 не означает, что гипотеза H_0 – единственно верное утверждение, это означает лишь, что гипотеза H_0 не противоречит имеющимся экспериментальным данным.

Решения и ошибки (случай простых гипотез H_0 и H_1)

Пусть гипотеза H_0 (нуль-гипотеза) и альтернативная к ней гипотеза H_1 – обе простые и рассматривается правосторонний критерий; при этом $f_T(t|H_0)$ и $f_T(t|H_1)$ – распределения статистики критерия T при справедливости гипотез H_0 и H_1 , соответственно:



Ошибки, допускаемые при этом, классифицируют следующим образом:

Ошибка 1-го рода – отклонить H_0 , когда верна H_0 .

Ошибка 2-го рода – принять H_0 , когда верна H_1 .

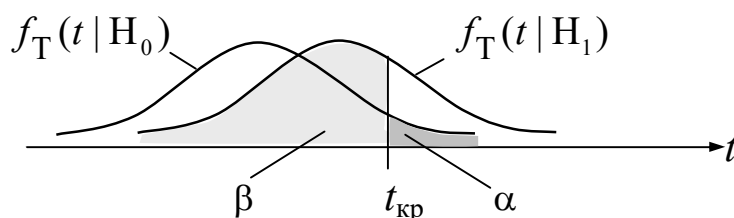
Вероятность ошибки 1-го рода равна уровню значимости:

$$P(T(X_1, \dots, X_n) \geq t_{кр} | H_0) = \alpha$$

Вероятность ошибки 2-го рода обозначают через β , она равна

$$P(T(X_1, \dots, X_n) < t_{кр} | H_1) = \beta.$$

Заметим, что при фиксированном α величина β зависит от гипотезы H_1 и от вида критической области:



Обычно критерий стараются выбрать так, чтобы при данном уровне значимости α величина ошибки второго рода β была минимальной (чтобы максимальной была так называемая мощность критерия $1 - \beta$).

20. Основные этапы процедуры проверки статистических гипотез

Процедура проверки состоит из двух этапов

На первом этапе

- формулируют предположения о распределении и о выборке X_1, \dots, X_n из этого распределения;
- выдвигают нуль-гипотезу H_0 и альтернативную гипотезу H_1 ;
- задают уровень значимости α ;
- выбирают статистику критерия $T(X_1, \dots, X_n)$;
- находят критические числа (границы критической области и области принятия гипотезы H_0);
- формулируют правило принятия решения.

На втором этапе

- проводят эксперимент – получают реализацию выборки x_1, x_2, \dots, x_n ;
- вычисляют $T(x_1, x_2, \dots, x_n) = t_e$;
- сопоставляют t_e с критическими числами;
- формулируют вывод об отклонении гипотезы H_0 на уровне значимости α или о ее принятии.

21. Подход к проверке статистических гипотез о параметрах распределений, основанный на доверительных интервалах

Пример 1_кди

Пусть проверяется гипотеза $H_0: \theta = \theta_0$ о параметре распределения против альтернативы $H_1: \theta \neq \theta_0$ (или $H_1: \theta < \theta_0$, или $H_1: \theta > \theta_0$) при заданном уровне значимости α .

Для проверки гипотезы указанного типа (параметрической гипотезы) можно воспользоваться подходом, основанном на доверительном интервале для параметра θ .

Пусть $(\vartheta_1; \vartheta_2)$ – доверительный интервал для параметра θ , соответствующий данному уровню значимости α : $P(\vartheta_1 < \theta < \vartheta_2) \equiv 1 - \alpha$, где $\vartheta_1 = \vartheta_1(X_1, \dots, X_n)$, $\vartheta_2 = \vartheta_2(X_1, \dots, X_n)$ – функции выборки. Правило принятия решения формулируют так: если интервал $(\vartheta_1 < \theta < \vartheta_2)$ накрывает θ_0 , то гипотезу H_0 принимают при данном уровне значимости α , в противном случае – отвергают.

По реализации выборки x_1, x_2, \dots, x_n вычисляют границы реализации доверительного интервала $\vartheta_{1e} = \vartheta_1(x_1, x_2, \dots, x_n)$, $\vartheta_{2e} = \vartheta_2(x_1, x_2, \dots, x_n)$. Если $\theta_0 \in (\vartheta_{1e}; \vartheta_{2e})$, то гипотезу H_0 принимают, в противном случае $\theta_0 \notin (\vartheta_{1e}; \vartheta_{2e})$ – гипотезу H_0 отвергают (соответственно, принимают H_1) на заданном уровне значимости α .

Замечание

Отыскание критической области и доверительного интервала приводит к одинаковым результатам, однако их истолкование различно:

- критическая область определяет критические числа, между которыми заключено $(1-\alpha)\%$ числа наблюдаемых значений *статистики критерия*;
- доверительный интервал определяет границы (концы доверительного интервала), между которыми в $(1-\alpha)\%$ опытов заключено *истинное значение* оцениваемого параметра θ распределения.

Пример 1_кди

Проиллюстрируем описанный подход на примере проверки гипотезы о математическом ожидании нормального распределения в случае, когда стандартное отклонение известно. Конкретное содержание задачи приведено в п.18 и п.12 (Пример 1_кз и Пример 1_ди).

Допустим, что выборка X_1, \dots, X_{25} извлечена из нормального распределения $N(m; \sigma)$, стандартное отклонение которого известно: $\sigma = \sigma_0 = 0,02$. Проверке подлежит гипотеза о математическом ожидании m этого распределения: $H_0: m = m_0 = 2,40$ против альтернативы $H_1: m \neq m_0$ на уровне значимости $\alpha = 0,05$.

В п.12 по реализации выборки x_1, x_2, \dots, x_{25} получена реализация доверительного интервала $(2,412; 2,428)$ для m , отвечающего доверительной вероятности $1 - \alpha = 0,95$.

Поскольку $m_0 = 2,40 \notin (2,412; 2,428)$, делаем вывод, что статистические данные не согласуются с гипотезой H_0 , эта гипотеза должна быть отклонена на уровне значимости $\alpha = 0,05$.

22. Примеры проверки гипотез о параметрах распределений

Пример 2_кз_кди

По результатам многократных проверок установлено, что вероятность ложной тревоги для охранной системы, функционирующей в штатном режиме, составляет 0,005 (описание системы приведено в п.17). При очередной плановой проверке исправности системы относительная частота события “ложная тревога”, вычисленная по реализации контрольной выборки объема $n = 10^4$, оказалась равной 0,006.

Необходимо выяснить, позволяет ли результат проверки считать, что система работает исправно, а отличие относительной частоты $\hat{p}_e = 0,006$ от вероятности $p = 0,005$ объясняется случайными факторами. Уровень значимости принять равным $\alpha = 0,05$.

Решение

а) двусторонний критерий значимости

Допустим, что выполнены условия применимости схемы Бернулли: проведено n независимых испытаний, p – вероятность успеха в каждом испытании. Индикатор “успеха” в i -м испытании; X_i подчиняется распределению Бернулли, а случайная величина $X = X_1 + \dots + X_n$ – биномиальному распределению (см. п. 17)

Проверке подлежит гипотеза $H_0: p = p_0 = 0,005$ о вероятности p в схеме Бернулли против альтернативы $H_1: p \neq p_0$ при заданном уровне значимости α .

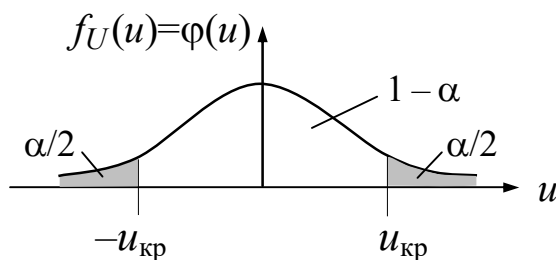
Относительная частота числа успехов в n испытаниях:

$$\hat{p} = \frac{X}{n} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Согласно теореме Муавра-Лапласа, центрированная и нормированная случайная величина $U = \frac{\hat{p} - p}{\sqrt{pq/n}}$ распределена асимптотически нормально при $n \rightarrow +\infty$. Объем выборки $n = 10^4$,

поэтому примем допущение, что статистика U при данном (большом) объеме выборки и при справедливости гипотезы H_0 подчиняется стандартному нормальному распределению:

$$U = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} \sim N(0;1) - \text{статистика критерия.}$$



Правило принятия решения имеет вид:

при $(|U| \geq u_{кр})$ гипотеза H_0 отвергается,
 при $(|U| < u_{кр})$ гипотеза H_0 принимается.

При $\alpha = 0,05$ имеем: $u_{кр} = u_{1-\alpha/2} = u_{0,975} = 1,96$.

Обозначим через u_e значение, которое статистика U приняла в результате $n = 10^4$ указанных контрольных измерений, тогда:

$$|u_e| = \left| \frac{\hat{p}_e - p_0}{\sqrt{p_0 q_0 / n}} \right| = \left| \frac{0,006 - 0,005}{\sqrt{0,005(1-0,005)/10000}} \right| = 1,42 < u_{кр} = 1,96.$$

Статистические данные не противоречат проверяемой гипотезе о параметре p биномиального распределения, а именно, $H_0: p = p_0 = 0,005$. Таким образом, эта гипотеза принимается на уровне значимости $\alpha = 0,05$.

Практический вывод – результат проверки не дает оснований считать, что *вероятность* ложной тревоги отличается от значения $p = 0,005$, характерного для функционирования системы в штатном режиме.

б) критерий, основанный на доверительном интервале

В п. 17 для вероятности p ложной тревоги при относительной частоте $\hat{p}_e = 0,006$ и объеме выборки $n = 10^4$ была найдена реализация

доверительного интервала $(0,0045; 0,0075)$, соответствующего доверительной вероятности $1 - \alpha = 0,95$.

Проверим гипотезу о параметре p биномиального распределения $H_0: p = p_0 = 0,005$ против альтернативы $H_1: p \neq p_0$ при заданном уровне значимости $\alpha = 0,05$ с помощью указанного доверительного интервала.

Поскольку $p = p_0 = 0,005 \in (0,0045; 0,0075)$, гипотеза H_0 принимается на уровне значимости $\alpha = 0,05$.

Пример 3_кз_кди (левосторонний критерий)

Стандартное содержание нежелательной примеси в выпускаемом продукте равно 2%. Контроль осуществляется с помощью прибора, точность которого указана в его паспорте и характеризуется параметром $\sigma_0 = 0,4\%$. После усовершенствования технологии тем же прибором провели $n = 16$ повторных независимых измерений содержания примеси x_1, x_2, \dots, x_{16} и получили $\bar{x} = 1,8\%$. Необходимо выяснить, объясняется ли уменьшение содержания примеси случайными факторами или физическими причинами – новой технологией. Уровень значимости принять равным $\alpha = 0,05$. *Указание:* считать, что контрольная выборка извлечена из нормального распределения.

Какой вывод об успешности новой технологии при величине $\bar{x} = 1,8\%$, полученной по серии контрольных измерений, можно сделать в случаях:

(1) при уровне значимости $\alpha = 0,01$ при том же объеме контрольной выборки $n = 16$; (2) при $\alpha = 0,01$ и объеме выборки $n = 25$.

Решение

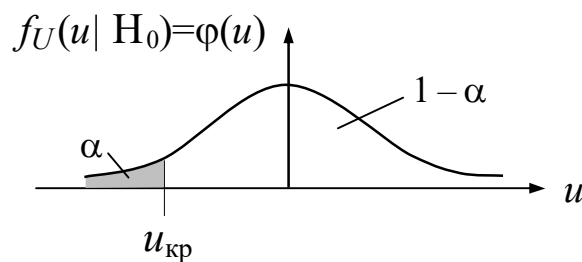
Допустим, что x_1, x_2, \dots, x_{16} – реализация выборки X_1, \dots, X_{16} – из нормального распределения $N(m; \sigma)$, один параметр которого σ известен: $\sigma = \sigma_0$. Проверке подлежит гипотеза $H_0: m = m_0 = 2\%$ о

математическом ожидании распределения, из которого (после изменения технологии) извлечена выборка, против альтернативы $H_1: m < m_0$ (левосторонняя альтернатива) при уровне значимости $\alpha = 0,05$.

(а) *односторонний (левосторонний) критерий значимости*

При указанных предположениях о распределении, из которого извлечена выборка, и справедливости гипотезы H_0 статистика

критерия $\frac{\bar{X} - m_0}{\sigma_0 / \sqrt{n}} = U$ подчиняется стандартному нормальному распределению: $U \sim N(0; 1)$.



Правило принятия решения, основанное на данной статистике критерия U :

- при $(U \leq u_{кр})$ гипотеза H_0 отвергается,
- при $(U > u_{кр})$ гипотеза H_0 принимается.

Критическое число $u_{кр}$ является квантилью порядка α стандартного нормального распределения:

$$u_{кр} = u_{\alpha} = -u_{1-\alpha} = u_{0,05} = -u_{0,95} = -1,65.$$

Вычислим значение u_e статистики критерия и сравним его с критическим числом $u_{кр}$:

$$u_e = \frac{\bar{x} - m_0}{\sigma_0} \sqrt{n} = \frac{0,018 - 0,020}{0,004} \sqrt{16} = -\frac{0,002}{0,004} \cdot 4 = -2 < -1,65 = u_{кр}.$$

Таким образом, гипотеза $H_0: m = m_0 = 2\%$ противоречит результату контрольных измерений x_1, x_2, \dots, x_{16} , поэтому ее следует отвергнуть и принять альтернативную гипотезу $H_1: m < m_0$.

Практический вывод: можно считать, что изменение технологии привело к цели – уменьшению содержания нежелательной примеси в

выпускаемом продукте. Заметим, что этот вывод справедлив при уровне значимости $\alpha = 0,05$.

Ответим теперь на поставленные в задаче дополнительные вопросы (1) и (2).

(1) При $\alpha = 0,01$ и объеме выборки $n = 16$ имеем:

$$u_{кр} = u_{\alpha} = -u_{1-\alpha} = u_{0,01} = -u_{0,99} = -2,33 < -2 = u_e .$$

В соответствие с указанным правилом принятия решения обнаруживаем, что нулевая гипотеза $H_0: m = m_0 = 2\%$ согласуется со статистическими данными при уровне значимости $\alpha = 0,01$.

Практический вывод в этом случае: контрольные измерения не дают оснований считать, что содержание примеси в выпускаемом продукте уменьшилось.

(2) Заметим, что величина $|u_e|$ растет с увеличением объема

выборки n :
$$u_e = \frac{\bar{x} - m_0}{\sigma_0} \sqrt{n} = \frac{0,018 - 0,020}{0,004} \sqrt{n} = -\frac{\sqrt{n}}{2} .$$

Отклонению нулевой гипотезы H_0 соответствует выполнение условия $u_e \leq u_{кр} : -\frac{\sqrt{n}}{2} \leq -2,33$, откуда $n \geq 22$.

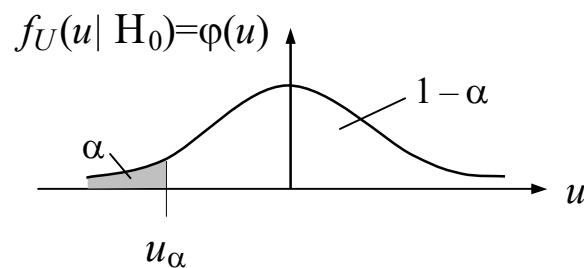
Если бы результат контрольных измерений $\bar{x} = 1,8\%$ был получен при $n \geq 22$, то гипотезу $H_0: m = m_0 = 2\%$ при уровне значимости $\alpha = 0,01$ следовало бы отвергнуть, а альтернативную гипотезу $H_1: m < m_0$ – принять. Таким образом, результат $\bar{x} = 1,8\%$ при $n = 25$ свидетельствует против гипотезы H_0 .

Обычно, для обоснованного отклонения нулевой гипотезы H_0 выбирают $\alpha = 0,01$. Если же речь идет о принятии H_0 , то уровень значимости α выбирают равным $0,05$.

(б) критерий, основанный на одностороннем (левостороннем) доверительном интервале

Как и в предыдущем пункте (а), допустим, что x_1, x_2, \dots, x_{16} – реализация выборки X_1, \dots, X_{16} – из нормального распределения $N(m; \sigma)$, один параметр которого σ известен $\sigma = \sigma_0$, таким образом, полагаем $X_i \sim N(m; \sigma_0) \quad i=1, \dots, 16$. При указанных предположениях и справедливости гипотезы H_0 статистика $U = \frac{\bar{X} - m}{\sigma_0 / \sqrt{n}}$ подчиняется стандартному нормальному распределению $N(0; 1)$.

Построим левосторонний доверительный интервал для математического ожидания m . Зададим величину доверительной вероятности $1 - \alpha$ и запишем равенство $P\left(\frac{\bar{X} - m}{\sigma_0 / \sqrt{n}} > u\right) = 1 - \alpha$, откуда $u = u_\alpha$ – квантиль порядка α (см. рисунок ниже, где $\varphi(u)$ – плотность вероятности стандартного нормального распределения).



$$P\left(-\infty < m < \bar{X} - u_\alpha \frac{\sigma_0}{\sqrt{n}}\right) = 1 - \alpha.$$

Правую границу этого интервала $\bar{X} - u_\alpha \frac{\sigma_0}{\sqrt{n}}$ – функцию выборки обозначим через $\tilde{m}_{cp}(X_1, \dots, X_n)$.

Правило принятия решения при проверке гипотезы $H_0: m = m_0$:
если $m_0 \in (-\infty; \tilde{m}_{cp})$, то гипотезу H_0 принимают;

если $m_0 \notin (-\infty; \tilde{m}_{zp})$ – гипотезу H_0 отвергают (соответственно, принимают альтернативную гипотезу H_1) на заданном уровне значимости α .

По реализации выборки $(x_1, x_2, \dots, x_{16})$ построим реализацию левостороннего доверительного интервала.

При $\alpha = 0,05$ $u_\alpha = -u_{1-\alpha} = u_{0,05} = -u_{0,95} = -1,65$.

Вычислим значение правой границы интервала \tilde{m}_{zpe} в точке

$$(x_1, x_2, \dots, x_{16}): \tilde{m}_{zpe} = \bar{x} - u_{0,05} \frac{\sigma_0}{\sqrt{n}} = 0,018 + 1,65 \frac{0,004}{4} = 0,01965$$

Видим, что $m_0 = 0,02 \notin (-\infty; 0,01965)$, поэтому гипотезу H_0 следует отвергнуть при $\alpha = 0,05$.

При $\alpha = 0,01$ и объеме выборки $n = 16$ имеем: $u_\alpha = -u_{1-\alpha} = u_{0,01} =$

$$= -u_{0,99} = -2,33; \text{ тогда } \tilde{m}_{zpe} = 0,018 + 2,33 \frac{0,004}{4} = 0,02033.$$

Гипотеза H_0 не противоречит статистическим данным, полученным в результате контрольных измерений, поскольку $m_0 = 0,02 \in (-\infty; 0,02033)$. Таким образом, гипотеза H_0 принимается на уровне значимости $\alpha = 0,01$.

Пример 4_кз (правосторонний критерий)

Средний срок службы до первого отказа для приборов, выпускаемых по стандартной технологии, равен 1000 час. Для повышения надежности выпускаемой продукции технологию усовершенствовали. С целью контроля эффективности новой технологии отобрали и испытали опытную партию из $n = 10$ приборов. По этой выборке вычислили выборочное среднее и выборочное стандартное отклонение (исправленное) срока службы, соответственно, $\bar{x} = 1100$ час. и $s^* = 100$ час. Можно ли считать, что новая технология увеличила срок службы приборов? Принять уровень

значимости равным $\alpha=0,01$. *Указание:* считать, что контрольная выборка извлечена из нормального распределения.

Решение

Положим, что контрольная выборка x_1, x_2, \dots, x_{10} – реализация выборки X_1, \dots, X_{10} из нормального распределения $X \sim N(m; \sigma)$ с неизвестными параметрами m и σ .

Проверим гипотезу $H_0: m = m_0 = 1000$ о математическом ожидании распределения, из которого (после изменения технологии) извлечена выборка, против альтернативы $H_1: m > m_0$ (правосторонняя альтернатива) при уровне значимости $\alpha=0,01$. Согласно лемме

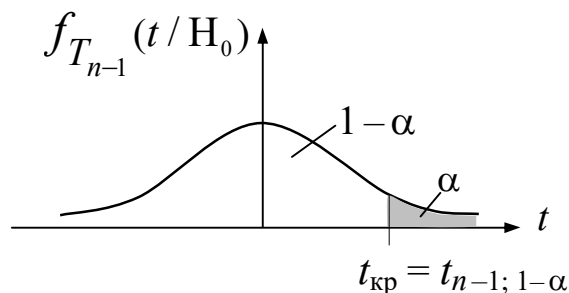
Фишера (п. 13) $U = \frac{\bar{X} - m_0}{\sigma / \sqrt{n}} \sim N(0; 1)$ и $\frac{(n-1) S^{*2}}{\sigma^2} = \chi_{n-1}^2$ –

независимые статистики.

В качестве статистики критерия проверки гипотезы H_0 возьмем отношение $T_{n-1} = \frac{U}{\sqrt{\frac{\chi_{n-1}^2}{n-1}}} = \frac{\bar{X} - m_0}{\frac{S^*}{\sqrt{n}}}$, подчиняющееся распределению

Стьюдента при справедливости H_0 .

Сформулируем правило принятия решения, основанное на распределении статистики критерия T_{n-1} при справедливости H_0 :



при $(T_{n-1} \geq t_{кр})$ гипотеза H_0 отвергается,

при $(T_{n-1} < t_{кр})$ гипотеза H_0 принимается.

Критическое число $t_{кр}$ – это квантиль порядка $1-\alpha$ распределения Стьюдента с $n-1$ степенью свободы: $t_{кр} = t_{n-1; 1-\alpha} = t_{9; 0,99} = 2,82$ при $\alpha = 0,01$. Вычислим значение статистики критерия t_e , отвечающее контрольной выборке x_1, x_2, \dots, x_{10} и сравним с значением $t_{кр}$:

$$t_e = \frac{\bar{x} - m_0}{s^*} \sqrt{n} = \frac{1100 - 1000}{100} \sqrt{10} = 3,16 > 2,82 = t_{кр}$$

Таким образом, гипотеза H_0 не согласуется с экспериментальными данными (значение t_e попадает в критическую область), поэтому гипотезу H_0 следует отвергнуть на уровне значимости $\alpha = 0,01$ и принять альтернативную гипотезу $H_1: m > m_0 = 1000$ (правостороннюю).

Практический вывод, который можно сделать в результате исследования опытной партии – применение новой технологии привело к увеличению срока службы выпускаемых приборов.

Пример 5_кз_кди (двусторонний критерий)

По результатам предварительных исследований установлено, что разброс значений контролируемого параметра изделий, произведенных на автоматической линии, характеризуется стандартным отклонением, равным $\sigma_0 = 20$ (условных единиц). Для проверки стабильности работы линии извлекли контрольную выборку из $n = 25$ изделий и получили выборочное стандартное отклонение (исправленное) $s^* = 24,3$ единицы. Считая, что контрольная выборка извлечена из нормального распределения, проверить обоснованность предположения о стабильности работы линии. Уровень значимости принять равным $\alpha = 0,05$.

Решение

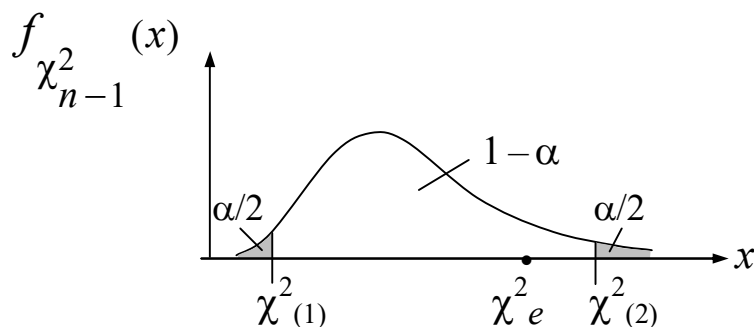
По условию выборочное стандартное отклонение $s^* = 24,3$ вычислено по реализации выборки из нормального распределения $N(m; \sigma)$.

Необходимо проверить гипотезу $H_0: \sigma = \sigma_0 = 20$ о стандартном отклонении распределения, из которого извлечена выборка, против альтернативы $H_1: \sigma \neq \sigma_0$ (двусторонний критерий) на уровне значимости $\alpha = 0,05$.

(а) двусторонний критерий значимости

При справедливости гипотезы H_0 имеем: $\frac{(n-1) S^{*2}}{\sigma_0^2} = \chi_{n-1}^2$ (п. 13

– лемма Фишера).



Находим критические числа $\chi_{(1)}^2 = \chi_{n-1; \alpha/2}^2 = \chi_{24; 0,025}^2 = 12,40$

и $\chi_{(2)}^2 = \chi_{n-1; 1-\alpha/2}^2 = \chi_{24; 0,975}^2 = 39,36$

Значение статистики критерия (вычисленное по контрольной выборке) $\chi_e^2 = \frac{(n-1) s^{*2}}{\sigma_0^2} = \frac{24 \cdot (24,3)^2}{(20)^2} = 35,43$ принадлежит области принятия гипотезы. Вывод: гипотеза H_0 не противоречит результату контрольных измерений и может быть принята на уровне значимости $\alpha = 0,05$. Таким образом, предположение о стабильности работы линии можно считать обоснованным.

(б) критерий, основанный на доверительном интервале

Доверительный интервал для квадрата стандартного отклонения σ^2 нормального распределения, отвечающий доверительной вероятности $1 - \alpha$, имеет вид (п. 15):

$$\left(\frac{(n-1) S^{*2}}{\chi^2_{(2)}}; \frac{(n-1) S^{*2}}{\chi^2_{(1)}} \right) = (\vartheta_1; \vartheta_2).$$

Для проверки гипотезы H_0 необходимо установить, покрывает ли реализация этого интервала $(\vartheta_{1e}; \vartheta_{2e})$ значение $\sigma_0^2 = 400$.

Вычислим границы этой реализации

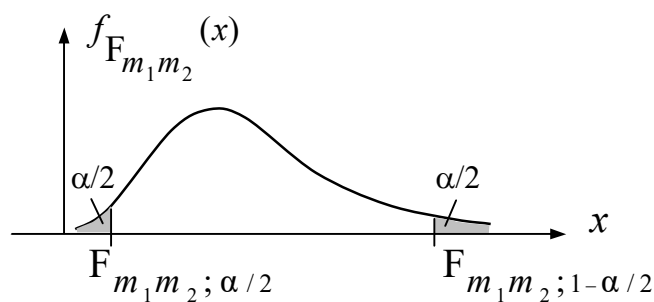
$$\vartheta_{1e} = \vartheta_1(x_1, x_2, \dots, x_{20}) = \frac{(n-1) S^{*2}}{\chi^2_{(2)}} = \frac{24 \cdot 590,49}{39,36} \cong 360;$$

$$\vartheta_{2e} = \vartheta_2(x_1, x_2, \dots, x_{20}) = \frac{(n-1) S^{*2}}{\chi^2_{(1)}} = \frac{24 \cdot 590,49}{12,40} \cong 1143.$$

Поскольку $\sigma_0^2 = 400 \in (360; 1143)$, гипотезу H_0 принимаем на уровне значимости $\alpha = 0,05$.

23. Распределение Фишера, свойство квантилей

Пусть случайные величины $\chi^2_{m_1}$ и $\chi^2_{m_2}$ – независимы. Случайная величина $F_{m_1 m_2} = \frac{\chi^2_{m_1}/m_1}{\chi^2_{m_2}/m_2}$ – отношение Фишера (F– отношение) подчиняется распределению Фишера с m_1 и m_2 степенями свободы; плотность вероятности $f_{F_{m_1 m_2}}(x)$ – известна (табулирована).



Докажем следующее свойство квантилей распределения Фишера:

$$F_{m_2 m_1; \alpha/2} = \frac{1}{F_{m_1 m_2; 1-\alpha/2}}$$

Учтем очевидное соотношение $\frac{1}{F_{m_1 m_2}} = F_{m_2 m_1}$ и определение квантили соответствующего порядка и запишем:

$$\begin{aligned} P(F_{m_1 m_2} > F_{m_1 m_2; 1-\alpha/2}) &= P\left(\frac{1}{F_{m_1 m_2}} < \frac{1}{F_{m_1 m_2; 1-\alpha/2}}\right) = \\ &= P\left(F_{m_2 m_1} < \frac{1}{F_{m_1 m_2; 1-\alpha/2}}\right) = \alpha/2 \Rightarrow \frac{1}{F_{m_1 m_2; 1-\alpha/2}} = F_{m_2 m_1; \alpha/2} \cdot \end{aligned}$$

В таблицах квантилей распределения Фишера приведены только “правые” квантили, так как “левые” могут быть легко получены из указанного соотношения. Заметим также, что правые квантили обладают свойством: $F_{m_1 m_2; 1-\alpha/2} > 1$.

24. Проверка гипотезы о равенстве дисперсий двух нормальных распределений (критерий Фишера). Пример 6_кз

Имеем две независимые выборки X_1, \dots, X_{n_1} и Y_1, \dots, Y_{n_2} объема n_1 и n_2 из нормальных распределений $X \sim N(m_X; \sigma_X)$, $Y \sim N(m_Y; \sigma_Y)$ (элементы внутри каждой из выборок – взаимно независимы по определению случайной выборки).

Проверяется гипотеза $H_0: \sigma_X = \sigma_Y$ против альтернативной гипотезы $H_1: \sigma_X \neq \sigma_Y$ при уровне значимости α .

Статистика критерия:

Исправленные выборочные дисперсии для указанных независимых выборок S_X^{*2} и S_Y^{*2} – независимы, кроме того, известно:

$$(n_1 - 1) S_X^{*2} / \sigma_X^2 = \chi_{n_1-1}^2, \quad (n_2 - 1) S_Y^{*2} / \sigma_Y^2 = \chi_{n_2-1}^2 \quad (\text{п. 13}), \text{ поэтому}$$

$$\text{статистика} \quad \frac{S_X^{*2} / \sigma_X^2}{S_Y^{*2} / \sigma_Y^2} = \frac{\chi_{n_1-1}^2 / (n_1 - 1)}{\chi_{n_2-1}^2 / (n_2 - 1)} \quad \text{подчиняется распределению}$$

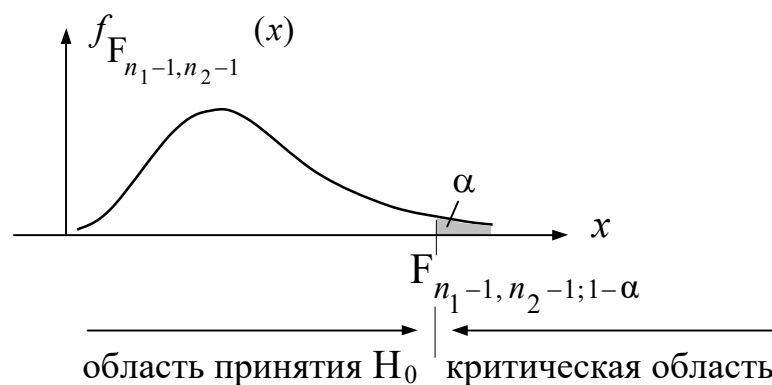
Фишера с $n_1 - 1$ и $n_2 - 1$ степенями свободы.

При справедливости гипотезы H_0 имеем $\sigma_X = \sigma_Y$, откуда отношение $\frac{S_X^{*2}}{S_Y^{*2}} = F_{n_1-1, n_2-1}$ – статистика критерия $\frac{S_X^{*2}}{S_Y^{*2}}$ подчиняется распределению Фишера.

Примем соглашение: обозначать бóльшую из выборочных дисперсий через S_X^{*2} , меньшую – через S_Y^{*2} , тогда $\frac{S_X^{*2}}{S_Y^{*2}} > 1$.

Таким образом, приходим к правостороннему критерию: проверке подлежит гипотеза $H_0: \sigma_X = \sigma_Y$ против альтернативы $H_1: \sigma_X > \sigma_Y$.

Решение об отклонении или принятии гипотезы H_0 принимают, сопоставляя значение статистики критерия $\frac{S_X^{*2}}{S_Y^{*2}} = F_e$, полученное в эксперименте, с критическим числом – квантилью порядка $1-\alpha$ распределения Фишера $F_{n_1-1, n_2-1; 1-\alpha}$.



Пример 6_кз (правосторонний критерий)

Выяснить, значимо ли варьирует от одного дня к другому величина контролируемого параметра – изменчивости температуры в термостатируемом помещении, если в первый день по выборке объема $n_1=16$ получена выборочная дисперсия $s_1^{*2} = 1,23$, а во второй – по выборке объема $n_2=20$ получена $s_2^{*2} = 0,97$, соответственно.

Принять уровень значимости $\alpha = 0,05$. *Указание:* принять допущение, что выборки извлечены из нормальных распределений.

Решение

Следуя описанной выше процедуре, проверим гипотезу о равенстве дисперсий двух нормальных распределений, из которых извлечены выборки:

$H_0: \sigma_X = \sigma_Y$ против альтернативы $H_1: \sigma_X > \sigma_Y$.

Найдем критическое число $F_{n_1-1, n_2-1; 1-\alpha} = F_{15; 19; 0,95} = 2,23$

(по таблице квантилей распределения Фишера) и сравним его с

значением статистики критерия $F_e = \frac{s_X^{*2}}{s_Y^{*2}} = \frac{1,23}{0,97} = 1,268 < 2,23$

Вывод: на уровне значимости $\alpha = 0,05$ гипотеза $H_0: \sigma_X = \sigma_Y$ и данные опыта не противоречат друг другу. Это означает, что отличие значений контролируемого параметра – изменчивости температуры в термостатируемом помещении, измеренных в первый и второй день, может быть объяснено случайными факторами (незначимо на уровне значимости $\alpha = 0,05$).

25. Проверка гипотезы о равенстве математических ожиданий двух нормальных распределений (критерий Стьюдента)

Пример 7_кз

Пусть X_1, \dots, X_{n_1} и Y_1, \dots, Y_{n_2} – две независимые выборки объема n_1 и n_2 , соответственно, из нормальных распределений $N(m_X; \sigma_X)$ и $N(m_Y; \sigma_Y)$. Элементы внутри каждой из выборок – взаимно независимы по определению случайной выборки.

Проверяется гипотеза $H_0: m_X = m_Y$ против альтернативной гипотезы $H_1: m_X \neq m_Y$ при уровне значимости α .

Статистика критерия:

а) Если стандартные отклонения σ_X и σ_Y известны (вообще говоря, $\sigma_X \neq \sigma_Y$), то при сделанных предположениях имеем:

$$\bar{X} \sim N(m_X; \sigma_X / \sqrt{n_1}), \quad \bar{Y} \sim N(m_Y; \sigma_Y / \sqrt{n_2}).$$

Статистики \bar{X} и \bar{Y} независимы, как функции независимых выборок, а их разность $\bar{X} - \bar{Y}$ (композиция нормальных распределений) распределена по нормальному закону с параметрами

$$M(\bar{X} - \bar{Y}) = m_X - m_Y \quad \text{и} \quad D(\bar{X} - \bar{Y}) = \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}.$$

При справедливости гипотезы $H_0: m_X - m_Y = 0$ имеем:

$$U = (\bar{X} - \bar{Y}) / \sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}} \sim N(0; 1).$$

Правило принятия решения относительно гипотезы H_0 имеет вид:

при $|U| < u_{кр} = u_{1-\alpha/2}$ гипотеза H_0 принимается;

при $|U| \geq u_{кр} = u_{1-\alpha/2}$ гипотеза H_0 отклоняется.

б) Рассмотрим теперь задачу проверки гипотезы H_0 при неизвестных σ_X и σ_Y , причем $\sigma_X = \sigma_Y = \sigma$ (допущение о равенстве дисперсий может быть проверено с помощью критерия Фишера).

Имеем: $(n_1 - 1) S_X^{*2} / \sigma^2 = \chi_{n_1 - 1}^2$, $(n_2 - 1) S_Y^{*2} / \sigma^2 = \chi_{n_2 - 1}^2$ (лемма Фишера п. 13), при этом $\chi_{n_1 - 1}^2$ и $\chi_{n_2 - 1}^2$ независимы, так как выборочные дисперсии S_X^{*2} и S_Y^{*2} – функции независимых выборок.

Отсюда, по определению случайной величины χ_n^2 , получаем:

$$\chi_{n_1 - 1}^2 + \chi_{n_2 - 1}^2 = \chi_{n_1 + n_2 - 2}^2 = \frac{(n_1 - 1) S_X^{*2} + (n_2 - 1) S_Y^{*2}}{\sigma^2}.$$

Далее, при справедливости гипотезы $H_0: m_X = m_Y$, имеем:

$U \sim N(0; 1)$ – стандартное нормальное распределение.

При $\sigma_X = \sigma_Y = \sigma$ и справедливости гипотезы $H_0 : m_X = m_Y$ статистики $U = (\bar{X} - \bar{Y}) / \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}$ и $\chi^2_{n_1+n_2-2}$ – независимы,

поэтому $\frac{U}{\sqrt{\frac{\chi^2_{n_1+n_2-2}}{n_1+n_2-2}}} = T_{n_1+n_2-2}$ – отношение Стьюдента.

Таким образом, получаем выражение для статистики критерия:

$$T_{n_1+n_2-2} = (\bar{X} - \bar{Y}) \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \sqrt{\frac{n_1 + n_2 - 2}{(n_1 - 1)S_X^{*2} + (n_2 - 1)S_Y^{*2}}}.$$

Решения о принятии или отклонении гипотезы H_0 принимают на основе указанной статистики критерия (критерий Стьюдента):

при $|T_{n_1+n_2-2}| < t_{кр} = t_{n_1+n_2-2; 1-\alpha/2}$ гипотеза H_0 принимается;

при $|T_{n_1+n_2-2}| \geq t_{кр}$ гипотеза H_0 отвергается.

При $n_1 = n_2 = n$ выражение для статистики критерия имеет вид:

$$T_{2(n-1)} = \frac{\bar{X} - \bar{Y}}{\sqrt{S_X^{*2} + S_Y^{*2}}} \sqrt{n}.$$

Замечание

В случае, когда дисперсии неизвестны и их равенство не предполагается, используют статистику критерия, аналогичную рассмотренной в настоящем параграфе для случая, когда дисперсии считались известными, но не равными.

Пример 7_кз (двусторонний критерий)

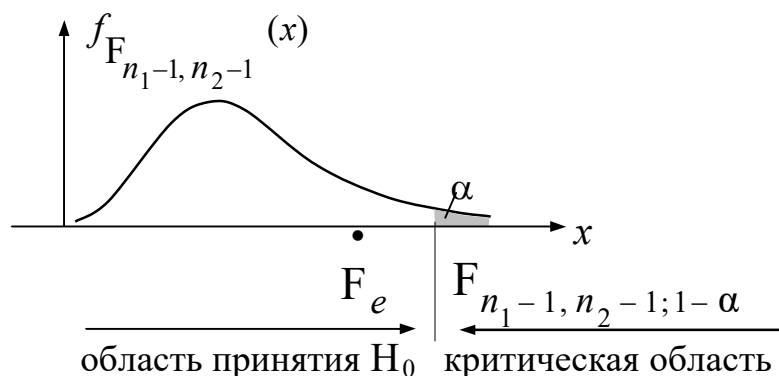
Проверить гипотезу о равенстве математических ожиданий двух нормальных распределений $N(m_1; \sigma_1)$ и $N(m_2; \sigma_2)$ на основе двух независимых выборок одинакового объема $n_1 = n_2 = n = 15$ из этих распределений. При обработке реализации первой выборки получено: $\bar{x}_1 = 18,80$, $s_1^{*2} = 24,03$ (выборочное среднее и исправленная

выборочная дисперсия); второй: $\bar{x}_2 = 16,13$; $s_2^{*2} = 15,43$. Уровень значимости принять равным $\alpha = 0,05$.

Указание: Предварительно с помощью критерия Фишера (см. п.24) проверить на уровне значимости $\alpha = 0,05$ справедливость гипотезы о равенстве стандартных отклонений $\sigma_1 = \sigma_2 = \sigma$ распределений, из которых извлечены выборки.

Решение

1) Применим критерий Фишера для проверки гипотезы $H_0: \sigma_1 = \sigma_2$ против альтернативы $H_1: \sigma_1 > \sigma_2$ при уровне значимости $\alpha = 0,05$. Найдем критическое число $F_{n_1-1, n_2-1; 1-\alpha} = F_{14; 14; 0,95} = 2,40$ (по таблице квантилей распределения Фишера) и сравним его с значением статистики критерия $F_e = \frac{s_1^{*2}}{s_2^{*2}} = \frac{24,03}{15,43} = 1,56$.



Значение статистики критерия F_e принадлежит области принятия гипотезы, на уровне значимости $\alpha = 0,05$ гипотеза $H_0: \sigma_1 = \sigma_2 = \sigma$ принимается, поэтому будем считать, что стандартные отклонения распределений, из которых извлечены выборки, равны.

2) Полагая $\sigma_1 = \sigma_2 = \sigma$ (σ — неизвестно), с помощью критерия Стьюдента проверим теперь гипотезу о равенстве математических ожиданий указанных распределений $H_0: m_1 = m_2$ против

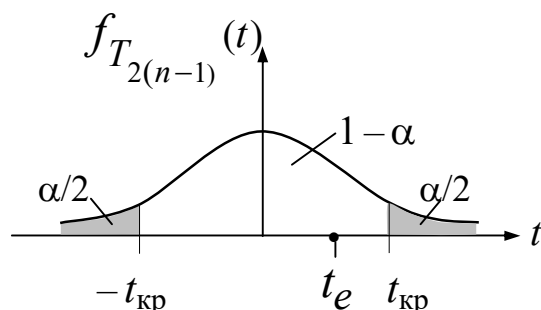
альтернативной гипотезы $H_1: m_1 \neq m_2$ на уровне значимости $\alpha = 0,05$.

Вычислим значение статистики критерия

$$T_{2(n-1)} = \frac{\bar{X} - \bar{Y}}{\sqrt{S_X^{*2} + S_Y^{*2}}} \sqrt{n}, \text{ отвечающее данным выборкам:}$$

$$t_e = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^{*2} + s_2^{*2}}} \sqrt{n} = \frac{18,80 - 16,13}{\sqrt{24,03 + 15,43}} \sqrt{15} = 1,646.$$

По таблицам квантилей распределения Стьюдента найдем критическое число $t_{кр} = t_{2(n-1); 1-\alpha/2} = t_{28; 0,975} = 2,048$.



Таким образом, значение t_e статистики критерия принадлежит области принятия гипотезы; гипотеза $H_0: m_1 = m_2$ о равенстве математических ожиданий распределений не противоречит опытным данным и может быть принята на уровне значимости $\alpha = 0,05$.

26. Теорема Пирсона, проверка гипотезы о вероятностях в обобщенной схеме Бернулли

1. Рассмотрим последовательность n независимых испытаний с числом k ($k \geq 2$) исходов в каждом испытании (обобщенная схема Бернулли).

Теорема К. Пирсона. Пусть n – число независимых испытаний, результатом каждого испытания является один из k исходов A_1, \dots, A_k ($k \geq 2$). Вероятности исходов A_1, \dots, A_k равны, соответственно: p_1, \dots, p_k и не зависят от номера испытания; все $p_i \neq 0$ и $p_1 + \dots + p_k = 1$.

Пусть в результате проведения n испытаний исход A_1 наблюдался N_1 раз; ... $A_i - N_i$ раз; ... $A_k - N_k$ раз, при этом $N_1 + \dots + N_k = n$.

Заметим, что $N_i (i = 1, \dots, k)$ – случайные величины, подчиняющиеся биномиальному распределению с параметрами n и p_i , при этом $M(N_i) = np_i$, $D(N_i) = np_i(1 - p_i)$.

Примем без доказательства утверждение (*теорема Пирсона*): случайная величина (хи-квадрат) $\mathbb{X}^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i}$ при $n \rightarrow +\infty$ распределена как χ_{k-1}^2 (хи-квадрат с $k-1$ степенью свободы):

$$\forall x \quad P(\mathbb{X}^2 < x) \xrightarrow{n \rightarrow +\infty} P(\chi_{k-1}^2 < x).$$

Величины N_i называют *наблюдаемыми* частотами, а np_i – *ожидаемыми* частотами.

2. Пусть относительно вероятностей p_1, \dots, p_k выдвинута *простая* гипотеза о вероятностях $H_0: p_1 = p_1^0, \dots, p_k = p_k^0$ (альтернативная гипотеза $H_1: \exists i (i = 1, \dots, k) p_i \neq p_i^0$ и задан уровень значимости α).

В качестве статистики критерия для проверки гипотезы H_0 возьмем $\mathbb{X}^2 = \sum_{i=1}^k \frac{(N_i - np_i^0)^2}{np_i^0} = n \left(\sum_{i=1}^k \frac{(\frac{N_i}{n} - p_i^0)^2}{p_i^0} \right)$.

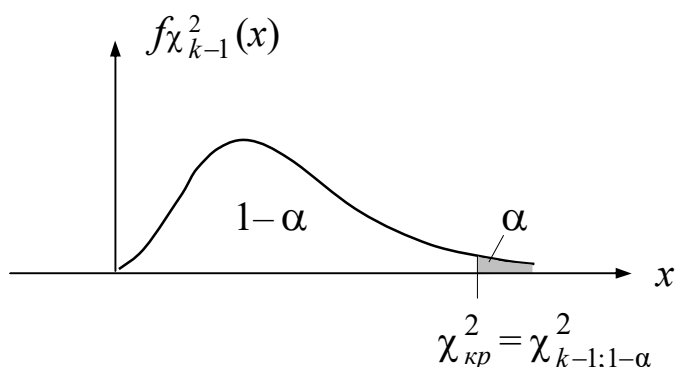
Если гипотеза H_0 верна, то согласно теореме Пирсона имеем:

$$\mathbb{X}^2 \xrightarrow{n \rightarrow +\infty} \sim \chi_{k-1}^2.$$

Если H_0 – неверна, то хотя бы одна из относительных частот $\frac{N_i}{n}$ сходится по вероятности к величине p_i , отличной от p_i^0 :

$$\frac{N_i}{n} \xrightarrow{P} p_i \neq p_i^0, \text{ поэтому } \mathbb{X}^2 = n \left(\sum_{i=1}^k \frac{(\frac{N_i}{n} - p_i^0)^2}{p_i^0} \right) \xrightarrow{n \rightarrow +\infty} \infty.$$

Отсюда следует, что гипотеза H_0 должна быть отвергнута, если полученное в опыте значение X^2 велико:



Таким образом, приходим к правостороннему критерию:

при $X^2 \geq \chi^2_{кр}$ – гипотезу H_0 отклоняют;

при $X^2 < \chi^2_{кр}$ – H_0 принимают.

Пусть в данном эксперименте частоты N_i (случайные величины) приняли конкретные значения n_i , соответственно. Вычисляют

$\chi_e^2 = \sum_{i=1}^k \frac{(n_i - np_i^0)^2}{np_i^0}$, и решение об отклонении или принятии гипотезы

H_0 принимают, сопоставляя значение χ_e^2 с критическим числом

$$\chi_{кр}^2 = \chi_{k-1; 1-\alpha}^2.$$

Не следует считать, что при справедливости гипотезы H_0 величина X^2 должна быть близкой к нулю, поскольку результаты наблюдений (измерений) – это реализация *случайной* выборки.

Также необходимо учесть, что применение непрерывного распределения χ_{k-1}^2 в качестве аппроксимации распределения дискретной случайной величины X^2 порождает ряд ограничений. В частности требуется, чтобы n было достаточно велико ($n \geq 50$), “ожидаемые” частоты np_i^0 , а также значения n_i не должны быть

малыми. Более детально практические рекомендации будут обсуждены далее.

27. Проверка гипотезы о виде распределения – метод χ^2 для простой гипотезы

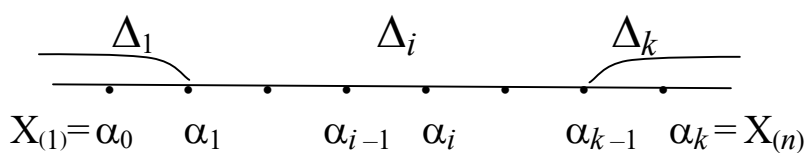
Пусть X_1, \dots, X_n – выборка из распределения $F_X(x)$ непрерывной случайной величины X , объем выборки n – достаточно велик, элементы выборки независимы и $F_{X_i}(x) = F_X(x)$, для всех $i=1, \dots, n$.

Пусть относительно распределения $F_X(x)$ проверке подлежит гипотеза $H_0: F_X(x) = F_0(x)$, альтернативная гипотеза $H_1: F_X(x) \neq F_0(x)$ и задан уровень значимости α .

Статистика критерия

По вариационному ряду $X_{(1)}, \dots, X_{(n)}$ построим k промежутков аналогично тому, как это делалось при построении гистограммы, с тем отличием, что в качестве крайних промежутков возьмем полубесконечные: $\Delta_1 = (-\infty; \alpha_1]$, ..., $\Delta_i = (\alpha_{i-1}; \alpha_i]$, ..., $\Delta_k = (\alpha_{k-1}; +\infty)$.

Число интервалов разбиения k обычно берут таким же, как при построении гистограммы, а именно, применяют либо формулу Старджесса: $k = 1 + 3,32 \lg n$, либо формулу: $k = 1,72 n^{1/3}$, а сами промежутки полагают равными (за исключением крайних – полубесконечных):



Наблюдаемые частоты N_i (случайные величины) – число элементов выборки, попавших в i -й промежуток разбиения $\Delta_i = (\alpha_{i-1}; \alpha_i]$.

Обозначим через p_i вероятность $P(X \in \Delta_i)$ для случайной величины X принять значение в промежутке Δ_i .

При справедливости гипотезы H_0 имеем:

$$P(X \in \Delta_1 | H_0) = F_0(\alpha_1) = p_1^0;$$

$$P(X \in \Delta_i | H_0) = F_0(\alpha_i) - F_0(\alpha_{i-1}) = p_i^0 \quad (i=2, \dots, k-1);$$

$$P(X \in \Delta_k | H_0) = 1 - F_0(\alpha_{k-1}) = p_k^0.$$

Проверяемая гипотеза о распределении $H_0: F_X(x) = F_0(x)$ равносильна гипотезе о том, что упомянутые вероятности p_i приняли определенные значения p_i^0 . Таким образом, приходим к задаче о проверке *простой гипотезы о вероятностях в обобщенной схеме Бернулли*, рассмотренной в п. 26: $H_0: p_i = p_i^0 \quad (i=1, \dots, k)$.

Правило принятия решения об отклонении (принятии) проверяемой гипотезы H_0 о виде распределения строится на основе приближения распределения статистики критерия Пирсона $\chi^2 = \sum_{i=1}^k \frac{(N_i - np_i^0)^2}{np_i^0}$ распределением χ_{k-1}^2 при больших объемах выборки ($n \geq 50$).

Практически по реализации выборки (x_1, x_2, \dots, x_n) получают реализацию вариационного ряда и отрезок $[x_{(1)}; x_{(n)}]$, где $x_{(1)} = \min(x_1, x_2, \dots, x_n) = \alpha_{(0)}$, $x_{(n)} = \max(x_1, x_2, \dots, x_n) = \alpha_{(k)}$, содержащий все элементы выборки, разбивают на k равных интервалов. Подсчитывают $\chi_e^2 = \sum_{i=1}^k \frac{(n_i - np_{ie}^0)^2}{np_{ie}^0}$, где n_i – наблюдаемые частоты (значения случайных величин N_i) – число элементов реализации выборки, фактически попавших в соответствующий интервал Δ_i , np_{ie}^0 – вычисленные *ожидаемые* частоты.

Решение об отклонении или принятии гипотезы H_0 принимают, сопоставляя значение χ_e^2 , соответствующее данной реализации выборки, с критическим числом $\chi_{кр}^2 = \chi_{k-1; 1-\alpha}^2$.

28. Проверка гипотезы о виде распределения – метод χ^2 для сложной гипотезы

Пусть X_1, \dots, X_n – выборка из распределения непрерывной случайной величины X , функция распределения которой зависит от r неизвестных параметров $F_X(x, \theta_1, \theta_2, \dots, \theta_r)$. В этом случае гипотеза о виде распределения $H_0: F_X(x, \theta_1, \dots, \theta_r) = F_0(x, \theta_1, \dots, \theta_r)$ – сложная.

В функции распределения F_0 неизвестные параметры заменим оценками максимального правдоподобия $\hat{\theta}_{1МП}, \dots, \hat{\theta}_{rМП}$ и, действуя аналогично процедуре проверки простой гипотезы, рассмотренной в п. 27, вычислим вероятности $\hat{p}_i^0 = P(X \in \Delta_i | H_0)$:

$$\hat{p}_i^0 = F_0(\alpha_i, \hat{\theta}_{1МП}, \dots, \hat{\theta}_{rМП}) - F_0(\alpha_{i-1}, \hat{\theta}_{1МП}, \dots, \hat{\theta}_{rМП}), \quad (i=2, \dots, k-1).$$

$$\text{для } \Delta_1: \hat{p}_1^0 = F_0(\alpha_1; \hat{\theta}_{1МП}, \dots, \hat{\theta}_{rМП}),$$

$$\text{для } \Delta_k: \hat{p}_k^0 = 1 - F_0(\alpha_{k-1}; \hat{\theta}_{1МП}, \dots, \hat{\theta}_{rМП}).$$

Гипотезы H_0 и H_1 при этом формулируются следующим образом:

$$H_0: p_i = \hat{p}_i^0 \quad (i=1, \dots, k); \quad H_1: \exists i \quad (i=1, \dots, k) \quad p_i \neq \hat{p}_i^0.$$

Доказано (теорема Фишера) что распределение статистики $\chi^2 = \sum_{i=1}^k \frac{(N_i - n\hat{p}_i^0)^2}{n\hat{p}_i^0}$ при справедливости гипотезы H_0 при $n \rightarrow +\infty$

стремится к распределению случайной величины χ_{k-r-1}^2 (с $k-r-1$

степенью свободы): $\forall x \quad P\left(\sum_{i=1}^k \frac{(N_i - n\hat{p}_i^0)^2}{n\hat{p}_i^0} < x\right) \xrightarrow{n \rightarrow +\infty} P(\chi_{k-r-1}^2 < x)$,

где r – число параметров, оцениваемых по выборке.

В остальном проверка гипотезы H_0 совпадает с рассмотренной в п. 27 процедурой проверки для случая простой гипотезы:

по реализации выборки (x_1, x_2, \dots, x_n) вычисляют значение

$$\chi_e^2 = \sum_{i=1}^k \frac{(n_i - n\hat{p}_{ie}^0)^2}{n\hat{p}_{ie}^0},$$

где n_i – значения *наблюдаемых* частот, фактически полученные в эксперименте; $n\hat{p}_{ie}^0$ – вычисленные *ожидаемые* частоты. Сопоставляя значение χ_e^2 с $\chi_{кр}^2 = \chi_{k-r-1; 1-\alpha}^2$, принимают решения: на уровне значимости α гипотезу H_0 отвергнуть, если $\chi_e^2 \geq \chi_{кр}^2$ или гипотезу H_0 принять, если $\chi_e^2 < \chi_{кр}^2$.

Замечания

Применимость аппроксимации непрерывным распределением χ_{k-1}^2 (в случае простой гипотезы) и χ_{k-r-1}^2 (в случае сложной гипотезы) к соответствующей статистике X^2 , распределение которой дискретно, накладывает определенные ограничения на построение упомянутого разбиения. Промежутки разбиения Δ_i следует строить так, чтобы выполнялось условие: “ожидаемое” $n\hat{p}_{ie}^0 \geq 5$. При этом длины промежутков не должны быть обязательно равными. Если для какого-либо промежутка $n\hat{p}_{ie}^0 < 5$, или наблюдаемая частота $n_i < 5$, то такой промежуток объединяют с соседним.

Число промежутков k таким образом может сократиться по сравнению с первоначальным и решение об отклонении (принятии) гипотезы H_0 принимают, сравнивая χ_e^2 с $\chi_{кр}^2 = \chi_{k^*-r-1; 1-\alpha}^2$, где k^* – окончательное число промежутков после объединения.

В то же время, условие “минимальное ожидаемое $n\hat{p}_{ie}^0 \geq 5$ ” может оказаться слишком жестким – допустимый минимум зависит от числа степеней свободы k .

В специальной литературе по прикладной статистике приводятся также следующие рекомендации. Общее количество промежутков разбиения k должно быть не меньше 8. В каждый интервал разбиения должно попасть не менее 7–10 элементов реализации выборки, причем желательно, чтобы в разные интервалы попало примерно одинаковое число точек.

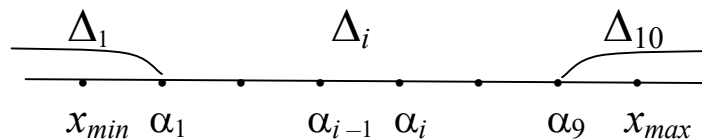
29. Пример 1_кс. Проверка гипотезы о нормальном распределении

Пусть имеется реализация выборки x_1, x_2, \dots, x_{500} из некоторого распределения, объем выборки $n = 500$. Вычислены выборочное среднее и исправленное стандартное отклонение, равные соответственно $\bar{x} = 0,044289$ и $s^* = 1,015813$, определены минимальный и максимальный элементы выборки $x_{(1)} = x_{min} = -2,769448$, $x_{(500)} = x_{max} = 2,835368$.

На уровне значимости $\alpha = 0,05$ необходимо проверить гипотезу H_0 о нормальности распределения, для которого экспериментально получена данная реализация выборки.

Решение

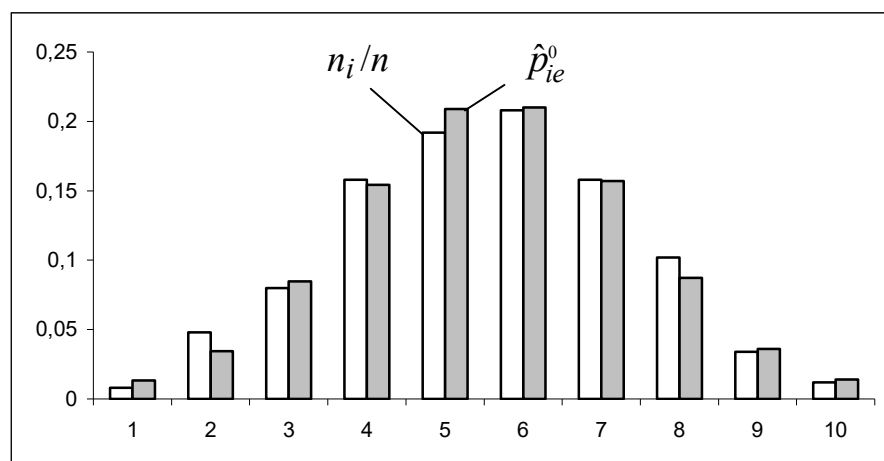
Построим группированный статистический ряд. Число интервалов разбиения примем равным $k = 10$, крайний левый и крайний правый интервалы равными, соответственно, $\Delta_1 = (-\infty; \alpha_1]$, $\Delta_{10} = (\alpha_9; +\infty)$.



Неизвестные параметры нормального распределения m и σ заменим значениями их оценок, вычисленными по данной реализации выборки, а именно $m = \hat{m}_e = \bar{x}$, $\sigma = \hat{\sigma}_e = s^*$. Для случайной величины, подчиняющейся нормальному распределению $N(\bar{x}; s^*)$, вычислим вероятности принять значение внутри соответствующих интервалов разбиения: $\hat{p}_{1e}^0 = \Phi((\alpha_1 - \bar{x})/s^*)$; $\hat{p}_{10e}^0 = 1 - \Phi((\alpha_9 - \bar{x})/s^*)$; $\hat{p}_{ie}^0 = \Phi((\alpha_i - \bar{x})/s^*) - \Phi((\alpha_{i-1} - \bar{x})/s^*)$, $i = 2, \dots, 9$, ($\Phi(x)$ – функция Лапласа).

Группированные числовые данные и результаты расчетов приведены в таблице и представлены на графике ниже.

i	Интервал $\Delta_i = (\alpha_{i-1}; \alpha_i]$		“Наблюдаемые” частоты n_i	“Ожидаемые” значения $n\hat{p}_{ie}^0$	$\frac{(n_i - n\hat{p}_{ie}^0)^2}{n\hat{p}_{ie}^0}$
	1	$-\infty$	-2,209	4	6,635
2	-2,209	-1,648	24	17,272	2,621
3	-1,648	-1,088	40	42,341	0,129
4	-1,088	-0,527	79	77,126	0,046
5	-0,528	0,033	96	104,401	0,676
6	0,0330	0,593	104	105,029	0,010
7	0,593	1,154	79	78,526	0,002
8	1,154	1,714	51	43,631	1,245
9	1,714	2,275	17	18,013	0,057
10	2,275	$+\infty$	6	7,025	0,149
					$\chi_e^2 = 5,983$



На рисунке выше представлены экспериментальная и теоретическая гистограммы: белые прямоугольники соответствуют наблюдаемым относительным частотам n_i/n , серые – вероятностям \hat{p}_{ie}^0 попадания значений выборки в соответствующий интервал.

Найдем (по таблице) критическое число $\chi_{кр}^2 = \chi_{k-r-1; 1-\alpha}^2$ – квантиль порядка $1-\alpha = 0,95$ распределения хи-квадрат с числом степеней свободы $k-r-1 = 10-2-1 = 7$: $\chi_{k-r-1; 1-\alpha}^2 = \chi_{7; 0,95}^2 = 14,067$.

Имеем:
$$\chi_e^2 = \sum_{i=1}^{10} \frac{(n_i - n\hat{p}_{ie}^0)^2}{n\hat{p}_{ie}^0} = 5,983 < \chi_{7; 0,95}^2 = 14,067.$$

Таким образом, гипотеза H_0 о нормальности распределения, из которого получена реализация выборки x_1, x_2, \dots, x_{500} на уровне значимости $\alpha = 0,05$ не противоречит экспериментальным данным и может быть принята.

Заметим, что по существу проверялась гипотеза о нормальности распределения с параметрами $m = \hat{m}_e = \bar{x}$, $\sigma = \hat{\sigma}_e = s^*$.

30. Пример 2_кс: проверка гипотезы о распределении Пуассона

Для статистического анализа процесса возникновения метеорных следов в определенной области атмосферы полное время наблюдения разбили на 2550 равных промежутков длительностью Δt , в каждом из которых регистрировалось число обнаруженных метеорных следов.

Результаты приведены в таблице, где наблюдаемые частоты n_i – число промежутков из 2550, в которых было зарегистрировано соответствующее число i ($i=0,1,\dots,13$) метеорных следов:

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13
n_i	44	189	375	475	491	430	265	143	83	31	15	7	1	1

На уровне значимости $\alpha = 0,05$ необходимо проверить гипотезу о том, что случайная величина X – число зарегистрированных метеорных следов за временной промежуток длительностью Δt , подчиняется распределению Пуассона:

$$P(X=i) = p_i = \frac{a^i}{i!} e^{-a} \quad (i=0,1,2,\dots).$$

Решение

Поскольку значение n_i в последних двух столбцах исходной таблицы меньше 5, объединим три последних столбца, получим таблицу для расчетов:

i	0	1	2	3	4	5	6	7	8	9	10	11
n_i	44	189	375	475	491	430	265	143	83	31	15	9

Используя данные *исходной* таблицы, вычислим значение \hat{a}_e оценки неизвестного параметра a распределения Пуассона ($\hat{a}_e = \bar{x}$, где \bar{x} – выборочное среднее):

$$\bar{x} = \frac{1}{2550} \sum_{i=1}^{13} i n_i = 10264/2550 = 4,025098 = \hat{a}_e.$$

Заменим в гипотетическом распределении Пуассона $p_i = \frac{a^i}{i!} e^{-a}$ неизвестный параметр a значением его оценки \hat{a}_e , вычисленным по экспериментальной выборке $\hat{a}_e = 4,025098$.

Таким образом, проверке подлежит гипотеза $H_0: p_i = \hat{p}_{ie}^0 = \frac{\hat{a}_e^i}{i!} e^{-\hat{a}_e}$ $i = 0, 1, \dots, 11$ (против альтернативы $H_1: \exists i (i = 1, \dots, k) p_i \neq \hat{p}_{ie}^0$), на уровне значимости $\alpha = 0,05$. Заметим, что проверяемая гипотеза H_0 – сложная, так как распределение содержит неизвестный параметр a , значение которого заменено значением его оценки $\hat{a}_e = \bar{x}$.

В таблице ниже приведен расчет величины $\chi_e^2 = \sum_{i=1}^{11} \frac{(n_i - n\hat{p}_{ie}^0)^2}{n\hat{p}_{ie}^0}$.

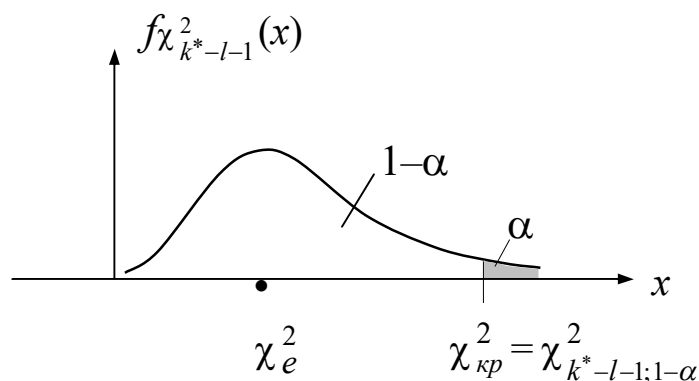
i	n_i	$i n_i$	\hat{p}_{ie}^0	$n\hat{p}_{ie}^0$	$\frac{(n_i - n\hat{p}_{ie}^0)^2}{n\hat{p}_{ie}^0}$
0	44	0	0,017883	45,54727	0,052561
1	189	189	0,071959	183,3322	0,175222
2	375	750	0,144778	368,9651	0,09871
3	475	1425	0,194192	495,0402	0,811265
4	491	1964	0,195353	498,1463	0,10252
5	430	2150	0,157217	401,0176	2,094627
6	265	1590	0,105438	269,0225	0,060145
7	143	1001	0,060611	154,6917	0,883667
8	83	664	0,030487	77,83116	0,343268
9	31	279	0,013631	34,80867	0,416735
10	15	150	0,005485	14,01083	0,069836
≥ 11	9	99	0,002969	7,586556	0,263337
	2550	10261	1,0	2550	$\chi_e^2 = 5,371893$

Заметим, что при расчете “теоретических” вероятностей в случае *целочисленной* случайной величины для крайних значений выборки следует поступать так же, как это делалось для непрерывного распределения (п. 28), когда крайние интервалы считались полубесконечными. В данном случае вероятность в строке таблицы, обозначенной ≥ 11 , равна сумме вероятностей всех значений $i \geq 11$ случайной величины, подчиняющейся распределению Пуассона:

$$\hat{p}_{\geq 11}^0 = 1 - \sum_{i=1}^{10} \hat{p}_{ie}^0.$$

Для случая сложной гипотезы статистикой критерия служит χ_{k-l-1}^2 , где l – число параметров, оцениваемых по выборке, k – общее число *различных* значений случайной величины X , зарегистрированных в данном эксперименте (аналог числа интервалов разбиения, используемого в методе χ -квадрат при проверке гипотезы о непрерывном распределении). В рассматриваемом примере $k = 14$.

Учтем, что три последних столбца исходной таблицы были объединены в один, поэтому $k^* = k - 2 = 12$, а также то, что неизвестное значение параметра распределения a было заменено значением его оценки \hat{a}_e , поэтому число степеней свободы для χ -квадрат равно окончательно $k^* - l - 1 = 12 - 1 - 1 = 10$.



Квантиль порядка 0,95 распределения χ^2 с числом степеней свободы равным 10, равна: $\chi_{кр}^2 = \chi_{10; 0,95}^2 = 18,30704$.

Поскольку $\chi_e^2 = 5,371893 < 18,30704 = \chi_{кр}^2$, – нет оснований для отклонения гипотезы H_0 .

Таким образом, гипотеза H_0 о том, что случайная величина – число метеорных следов, возникающих в выделенной области атмосферы за промежуток времени $\Delta\tau$, подчиняется распределению Пуассона с параметром, равным $\hat{a}_e = 4,025098$, не противоречит результатам наблюдений и может быть принята на уровне значимости $\alpha = 0,05$.

Замечание

Другой подход к проверке гипотезы H_0 о распределении Пуассона основан на известном свойстве этого распределения: $MX = DX = a$. Если X_1, \dots, X_n – выборка из распределения Пуассона, то $S^{*2} \xrightarrow[n \rightarrow +\infty]{P} DX = a$ и $\bar{X} \xrightarrow[n \rightarrow +\infty]{P} MX = a$, поэтому отношение S^{*2}/\bar{X} должно быть близким к 1 при достаточно больших значениях n . Доказано, что это отношение распределено асимптотически нормально: $S^{*2}/\bar{X} \sim N(1; \sqrt{2a^2/n})$ при $n \rightarrow +\infty$.

Если объем выборки n достаточно велик, то распределение центрированной и нормированной случайной величины $((S^{*2}/\bar{X}) - 1)/\sqrt{2a^2/n}$ будет близким к стандартному нормальному распределению $N(0; 1)$.

Заменим в последнем выражении неизвестный параметр a его несмещенной состоятельной асимптотически нормальной оценкой $\hat{a} = \bar{X}$, тогда статистика критерия $U = \sqrt{n}(S^{*2} - \bar{X})/(\bar{X}^2\sqrt{2})$ будет распределена приблизительно нормально при достаточно больших n . Возьмем U в качестве статистики критерия; правило принятия решения при уровне значимости α имеет вид:

при $(|U| \geq u_{1-\alpha/2})$ гипотеза H_0 отвергается,

при $(|U| < u_{1-\alpha/2})$ гипотеза H_0 принимается.

31. Проверка гипотезы о равенстве параметров p_1 и p_2 двух биномиальных распределений по выборкам большого объема из соответствующих распределений Бернулли

Пусть X_1, \dots, X_{n_1} и Y_1, \dots, Y_{n_2} – две независимые выборки объема n_1 и n_2 из распределений Бернулли с параметрами p_1 и p_2 , соответственно. Таким образом, проведено n_1 независимых испытаний с вероятностью успеха p_1 в каждом испытании и n_2 испытаний с вероятностью успеха p_2 , причем величины n_1 и n_2 (объемы выборок) достаточно велики. Как обычно, будем обозначать $1 - p_1 = q_1$ и $1 - p_2 = q_2$.

Необходимо проверить гипотезу $H_0: p_1 = p_2$ против альтернативной гипотезы $H_1: p_1 \neq p_2$ при уровне значимости α .

Статистика критерия:

Обозначим через X число успехов при проведении n_1 испытаний по схеме Бернулли с вероятностью успеха p_1 в каждом испытании, через Y – число успехов при проведении n_2 испытаний с вероятностью p_2 . Относительные частоты $\hat{p}_1 = \frac{X}{n_1}$ и $\hat{p}_2 = \frac{Y}{n_2}$ являются асимптотически нормальными случайными величинами (см. п. 17):

$$\hat{p}_1 \sim N\left(p_1; \sqrt{\frac{p_1 q_1}{n_1}}\right) \text{ при } n_1 \rightarrow +\infty; \quad \hat{p}_2 \sim N\left(p_2; \sqrt{\frac{p_2 q_2}{n_2}}\right) \text{ при } n_2 \rightarrow +\infty.$$

Будем считать, что нормальность распределений имеет место при данных (больших) объемах выборок n_1 и n_2 . По условию \hat{p}_1 и \hat{p}_2 – независимы, поэтому их разность также подчиняется нормальному распределению (как композиция нормальных распределений):

$$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2; \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}\right).$$

Заменив в последнем отношении неизвестные параметры p_1 и p_2 их оценками \hat{p}_1 и \hat{p}_2 , получим, что при справедливости гипотезы

$$H_0: p_1 = p_2, \quad \text{статистика} \quad U = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}} \sim N(0;1) \quad \text{подчиняется}$$

(приближенно) стандартному нормальному распределению. Возьмем U в качестве статистики критерия проверки гипотезы $H_0: p_1 = p_2$, тогда правило принятия решения имеет вид:

- при $(|U| \geq u_{1-\alpha/2})$ гипотеза H_0 отвергается,
- при $(|U| < u_{1-\alpha/2})$ гипотеза H_0 принимается.

Практически процедура проверки гипотезы сводится к сравнению значения статистики критерия, вычисленного по результатам конкретного эксперимента, $u_e = U(\hat{p}_{1e}; \hat{p}_{2e})$ с квантилью $u_{1-\alpha/2}$ порядка $1 - \alpha/2$ стандартного нормального распределения.

Замечание

При данных (больших) объемах выборок n_1 и n_2 будем считать справедливым утверждение:

$$\frac{p_1 - p_2 - (\hat{p}_1 - \hat{p}_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0;1). \quad \text{Заменим в знаменателе этого отношения}$$

неизвестные параметры распределений их оценками \hat{p}_1 , $1 - \hat{p}_1 = \hat{q}_1$ и \hat{p}_2 , $1 - \hat{p}_2 = \hat{q}_2$.

Действуя аналогично п. 17., построим приближенный доверительный интервал для разности $p_1 - p_2$, соответствующий доверительной вероятности $1 - \alpha$:

$$\hat{p}_1 - \hat{p}_2 - u_{1-\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + u_{1-\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}.$$

Сформулируем правило принятия решения при проверке гипотезы $H_0: p_1 - p_2 = 0$, основанное на доверительном интервале:

если $0 \notin (\hat{p}_1 - \hat{p}_2 - u_{1-\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}; \hat{p}_1 - \hat{p}_2 + u_{1-\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}})$,

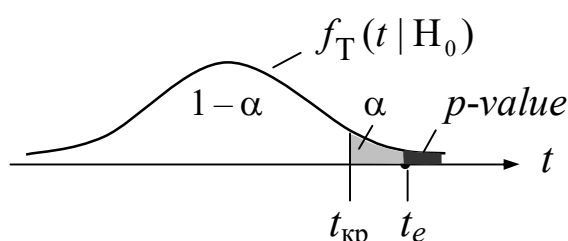
то гипотеза H_0 отвергается;

если $0 \in (\hat{p}_1 - \hat{p}_2 - u_{1-\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}; \hat{p}_1 - \hat{p}_2 + u_{1-\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}})$,

то гипотеза H_0 принимается на данном уровне значимости α .

32. Понятие p -значения

Рассмотрим односторонний (правосторонний) критерий).



При заданном уровне значимости α критическое число $t_{кр}$ определяется соотношением $P(T(X_1, \dots, X_n) \geq t_{кр} | H_0) = \alpha$, а p -значение (p -value) – соотношением $P(T(X_1, \dots, X_n) \geq t_e | H_0) = p$ -value.

В русскоязычной литературе p -значение (p -value) называют также *достигаемым уровнем значимости* (пи-величиной). Чем меньше оказывается значение p -value, тем сильнее свидетельствует совокупность наблюдений x_1, x_2, \dots, x_n против нулевой гипотезы.

Приложение 1. Предельные теоремы теории вероятностей **Теорема (неравенство Чебышева)**

Пусть непрерывная случайная величина X имеет второй начальный момент MX^2 , тогда:

$$\forall \varepsilon > 0 \quad P(|X| \geq \varepsilon) \leq MX^2 / \varepsilon^2.$$

Следствие Если случайная величина X имеет конечные математическое ожидание MX и дисперсию $DX = \sigma^2$, тогда

$$\forall \varepsilon > 0 \quad P(|X - MX| \geq \varepsilon) \leq DX / \varepsilon^2.$$

При $\varepsilon = 3\sigma$ из последнего неравенства получаем “3 σ -оценку” вероятности абсолютного отклонения случайной величины от ее математического ожидания на величину, не меньшую 3σ :

$$P(|X - MX| \geq 3\sigma) \leq 1/9.$$

Заметим, что эта оценка универсальна – она справедлива для любого распределения, однако является довольно грубой: для нормального распределения имеем: $P(|X - MX| \geq 3\sigma) = 0,027 \leq 1/9$; для равномерного распределения: $P(|X - MX| \geq 3\sigma) = 0 \leq 1/9$.

Закон больших чисел (теорема Чебышева)

Пусть $\{X_i\}$ – последовательность независимых случайных величин и пусть каждый член последовательности X_i имеет конечные математическое ожидание $MX_i = m_i$ и дисперсию $DX_i = D_i$, причем дисперсии ограничены в совокупности ($\exists B \quad 0 < B < +\infty \quad \forall i \quad D_i \leq B < +\infty$), тогда

$$\forall \varepsilon > 0 \quad P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n m_i\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0.$$

Приведем определение сходимости по вероятности. Последовательность случайных величин $\{S_n\}$ называют сходящейся по вероятности к случайной величине S , если

$$\forall \varepsilon > 0 \quad P(|S_n - S| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$$

Для сходимости по вероятности принято обозначение: $S_n \xrightarrow[n \rightarrow +\infty]{P} S$.

Поскольку любую неслучайную величину C ($C = \text{const}$) можно рассматривать как случайную величину X , принимающую свое единственное значение с вероятностью единица $P(X = C) = 1$, то можно говорить и о сходимости по вероятности последовательности случайных величин к неслучайной величине.

Следствие 1 из закона больших чисел – о сходимости по вероятности среднего арифметического одинаково распределенных случайных величин к их математическому ожиданию.

Пусть $\{X_i\}$ – последовательность одинаково распределенных независимых случайных величин, имеющих конечные математическое ожидание $MX_i = m$ и дисперсию $DX_i = \sigma^2$ (m и σ^2 – одни и те же для всех X_i), тогда: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow +\infty]{P} m$.

Следствие 2 из закона больших чисел – теорема Бернулли о сходимости по вероятности относительной частоты к вероятности.

Пусть случайная величина X – число наступлений события A (успеха) при проведении n независимых испытаний по схеме Бернулли, $p = P(A)$ – вероятность успеха в каждом испытании, $\hat{p} = \frac{X}{n}$ – относительная частота числа успехов в n испытаниях, тогда

$$\hat{p} \xrightarrow[n \rightarrow +\infty]{P} p.$$

Центральная предельная теорема

Рассмотрим последовательность $\{X_i\}$ независимых случайных величин ($\forall n X_1, \dots, X_n$ – независимы в совокупности) с математическими ожиданиями $\{m_i\}$ и дисперсиями $\{D_i\}$. При любом фиксированном n имеем:

$$M\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n MX_i = \sum_{i=1}^n m_i,$$

$$D\left(\sum_{i=1}^n X_i\right) = [\text{в силу независимости } X_i] = \sum_{i=1}^n DX_i = \sum_{i=1}^n D_i.$$

Случайная величина $Y_n = \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n m_i}{\sqrt{\sum_{i=1}^n D_i}}$ является центрированной и

нормированной по определению (ее математическое ожидание $MY_n = 0$, а дисперсия $DY_n = 1$).

Термин *центральная предельная теорема* (ЦПТ) вообще означает ряд теорем, утверждающих (при каких-либо условиях) справедливость предельного равенства вида:

$$\lim_{n \rightarrow \infty} P(Y_n < x) = \lim_{n \rightarrow \infty} P\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n m_i}{\sqrt{\sum_{i=1}^n D_i}} < x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

Если указанное предельное равенство имеет место, то при достаточно больших значениях n на его основе строится следующее приближение:

$$P\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n m_i}{\sqrt{\sum_{i=1}^n D_i}} < x\right) \approx \Phi(x).$$

В статистике принято считать, что такое приближение приемлемо уже при $n > 30$.

Теорема (ЦПТ для независимых одинаково распределенных величин)

Пусть $\{X_i\}$ – последовательность независимых одинаково распределенных случайных величин, имеющих конечные математическое ожидание m и дисперсию σ^2 , тогда

$$\forall x \lim_{n \rightarrow \infty} P\left(\frac{\sum_{i=1}^n X_i - nm}{\sigma\sqrt{n}} < x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

Символически утверждение ЦПТ записывают так:

$\sum_{i=1}^n X_i \underset{n \rightarrow \infty}{\sim} N(nm; \sigma\sqrt{n})$. Отсюда $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \underset{n \rightarrow \infty}{\sim} N(m; \sigma/\sqrt{n})$ – среднее

арифметическое любых независимых одинаково распределенных случайных величин (имеющих математическое ожидание и дисперсию) асимптотически нормально.

Заметим, что две последние символические записи удобны для применения ЦПТ в “допредельной” форме – при конечных достаточно больших значениях n .

Предельная теорема Муавра-Лапласа

Пусть X – число успехов при проведении n независимых испытаний (по схеме Бернулли) с вероятностью p успеха в каждом испытании, при этом $MX = np$, $DX = npq$, тогда для центрированной и нормированной случайной величины $\frac{X - np}{\sqrt{npq}}$ справедливо предельное равенство:

$$\forall x \lim_{n \rightarrow \infty} P\left(\frac{X - np}{\sqrt{npq}} < x\right) = \Phi(x).$$

Иными словами, теорема утверждает, что число успехов X в схеме Бернулли – асимптотически нормально: $X \underset{n \rightarrow +\infty}{\sim} N(np; \sqrt{npq})$.

Действительно, пусть X_i – индикатор появления успеха в i -м испытании (X_i – случайная величина, принимающая значение 1, если результатом i -го испытания является успех и значение 0 в противном случае):

X_i	0	1
P	q	p

$MX_i = p$, $DX_i = pq$, тогда $X = \sum_{i=1}^n X_i$ – сумма одинаково

распределенных независимых случайных величин (индикаторов X_i) удовлетворяет требованиям центральной предельной теоремы, откуда и следует утверждение теоремы Муавра-Лапласа.

Замечание 1

Как известно, случайная величина X подчиняется биномиальному распределению:

$$P(X = k) = b_k(n; p) = C_n^k p^k (1 - p)^{n-k} \quad (k = 0, 1, \dots, n).$$

При больших n и малых p непосредственные вычисления по этой формуле затруднительны, поэтому при $n > 40$, $p < 1/10$

используют приближение, основанное на теореме Пуассона:

$$b_k(n;p) \approx \frac{a^k}{k!} e^{-a}, \text{ где } a = np.$$

Замечание 2

На основании теоремы Муавра-Лапласа при достаточно больших значениях n имеем приближенно:

$$P(k_1 \leq X \leq k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right)$$

Это приближение применяют при значениях p не близких к нулю, величинах k_1 и k_2 – порядка несколько десятков, а $\frac{k_1 - np}{\sqrt{npq}}$ и $\frac{k_2 - np}{\sqrt{npq}}$ – порядка нескольких единиц.

Упражнение

(а) Вычислить приближенно $P\left(\left|\frac{X}{n} - p\right| \leq \varepsilon\right)$ – вероятность отклонения относительной частоты $\hat{p} = \frac{X}{n}$ числа успехов от вероятности p на величину, не превосходящую по абсолютной величине ε (величины n , p и $\varepsilon > 0$ заданы) при проведении n независимых испытаний.

(б) найти наименьшее значение числа испытаний n , при котором выполняется условие $P\left(\left|\frac{X}{n} - p\right| \leq \varepsilon\right) = 0,95$ при $\varepsilon = 0,01$ и $p = 0,5$.

Решение

(а) Если случайная величина Y распределена по нормальному закону $N(m; \sigma)$, то верно равенство $\forall t \ P(|Y - m| \leq \sigma t) = 2\Phi_0(t)$.

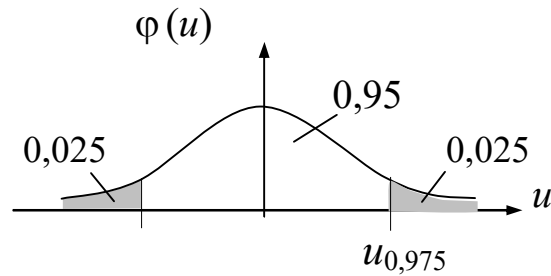
При достаточно больших n имеем приближенно:

$$\frac{X}{n} \sim N\left(p; \frac{\sqrt{pq}}{\sqrt{n}}\right), \text{ откуда } P\left(\left|\frac{X}{n} - p\right| \leq t \frac{\sqrt{pq}}{\sqrt{n}}\right) \approx 2\Phi_0(t).$$

Полагая $\varepsilon = t \frac{\sqrt{pq}}{\sqrt{n}}$, $t = \varepsilon \frac{\sqrt{n}}{\sqrt{pq}}$, получаем:

$$P\left(\left|\frac{X}{n} - p\right| \leq \varepsilon\right) \approx 2 \Phi_0\left(\varepsilon \frac{\sqrt{n}}{\sqrt{pq}}\right).$$

(б) По условию $P\left(\left|\frac{X}{n} - p\right| \leq \varepsilon\right) = 0,95$, $2 \Phi_0(u) = 0,95$, отсюда



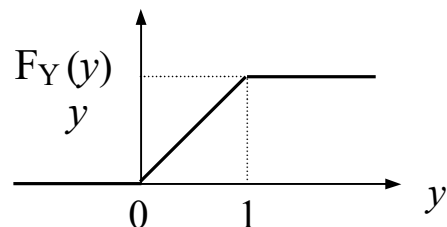
$u = u_{0,975} = 1,96$ и $\varepsilon \frac{\sqrt{n}}{\sqrt{pq}} = 1,96$. При $\varepsilon = 0,01$, $p = 0,5$ имеем

$1,96 = 0,01 \frac{\sqrt{n}}{\sqrt{0,25}}$, поэтому условие (б) выполняется при $n \geq (98)^2$.

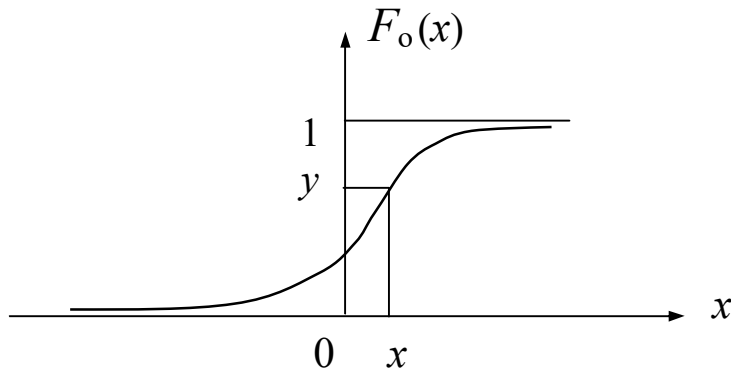
Приложение 2. Получение выборки из заданного распределения

Пусть заданная функция распределения $F_0(x)$ непрерывна и строго монотонна (так что существует обратная к ней функция F_0^{-1}), а случайная величина Y равномерно распределена на отрезке $[0;1]$:

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ y, & y \in [0;1] \\ 1, & y > 1 \end{cases}$$



Докажем, что при этом функция распределения $F_X(x) = P(X < x)$ случайной величины X , полученной преобразованием $X = F_0^{-1}(Y)$, будет равна заданной функции распределения $F_0(x)$.



В силу монотонности $F_0(x)$ события $(X < x)$ и $(0 < Y < y)$ равносильны, $y = F_0(x)$ и $x = F_0^{-1}(y)$, поэтому:

$$F_X(x) = P(X < x) = P(0 < Y < y) = F_Y(y) - F_Y(0) = y = F_0(x).$$

Для любого значения y случайной величины Y , *равномерно распределенной на $[0;1]$* , число x , полученное путем преобразования $x = F_0^{-1}(y)$, будет *соответствующим* значением (реализацией) случайной величины X , подчиняющейся наперед заданному распределению $F_0(x)$. Таким образом, реализация выборки x_1, x_2, \dots, x_n из распределения $F_0(x)$, может быть получена из реализации выборки y_1, y_2, \dots, y_n из указанного равномерного распределения с помощью преобразования $x_i = F_0^{-1}(y_i)$.

Заметим, что верно также следующее утверждение.

Пусть случайная величина X имеет строго монотонную функцию распределения $F_X(x)$ и $Y = F_X(X)$, тогда Y – равномерно распределена на отрезке $[0;1]$.

Действительно, поскольку $0 \leq Y \leq 1$ и $x = F_X^{-1}(y)$ (F_X – строго монотонна по условию, поэтому обратная функция F_X^{-1} существует), имеем:

$$F_Y(y) = P(0 < Y < y) = P(X < x) = F_X(x) = F_X(F_X^{-1}(y)) = y.$$

$F_Y(y) = y, y \in [0;1]$: Y – равномерно распределена на отрезке $[0;1]$.

Литература

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Основы моделирования и первичная обработка данных. – М.: Финансы и статистика, 1983, 471 с.
2. Браунли К.А. Статистическая теория и методология в науке и технике.– М.: Наука, 1977. – 408 с., ил.
3. Вероятностные разделы математики. Учебник для бакалавров технических направлений.//Под ред. Максимова Ю.Д.– СПб.: «Иван Федоров», 2001.– 592 с., ил.
4. Ивченко Г.И., Медведев Ю.И. Математическая статистика: Учеб. пособие для вузов.– 2-е изд. доп. М.: Высш. шк., 1992. – 304 с., ил.
5. Калинин В.М., Тихомиров С.Р. Лекции по теории вероятностей и математической статистике. СПб., 2002. 89 с.
6. Крамер Г. Математические методы статистики. – М.: Мир, 1975, 648 с.
7. Положинцев Б.И. Теория вероятностей и математическая статистика. Введение в математическую статистику : учеб. пособие .– СПб., Изд-во Политехн. ун-та 2010 .– 96с.
8. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере/ под ред. Фигурнова В.Э. –М.: ИНФРА – М. 1998. –528 с., ил.

Положинцев Борис Иосифович

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА
ВВЕДЕНИЕ В МАТЕМАТИЧЕСКУЮ
СТАТИСТИКУ**

Учебное пособие

Налоговая льгота – Общероссийский классификатор продукции
ОК 005-93, т. 2; 95 3005 – учебная литература

Подписано в печать 16.02.2018. Формат 60×84/16.
Усл. печ. л. 6,0. Тираж 120. Заказ 16.

Отпечатано с готового оригинал-макета, предоставленного автором,
в Издательско-полиграфическом центре Политехнического университета.
195251, Санкт-Петербург, Политехническая ул., 29.
Тел.: (812) 552-77-17; 550-40-14.