## Тема 2. Понятие о методах статистических выводов

Проблемы статистических выводов традиционно делятся на проблемы оценивания и проверку гипотез. Главное различие между этими двумя проблемами состоит в том, что при оценивании мы должны определить величину параметра или нескольких параметров. В то время как при проверке гипотез мы должны решить: принять или отвергнуть специфическую величину (или ряд специфических величин) параметра или нескольких параметров.

### 2.2.1. Понятие оценки параметров

В общем виде задача оценки параметров формулируется следующим образом.

Пусть распределение признака $X$ – генеральной совокупности – задается функцией вероятности $f(x,\theta)=P(X=x_i)$ для дискретной случайной величины или плотностью вероятностей $f(x,\theta)$ для непрерывной случайной величины, которая содержит неизвестный параметр $\theta$.

Для вычисления параметра $\theta$ используют выборку $x_1,x_2,\ldots,x_n$, каждая из которых имеет один и тот же закон распределения, что и признак $X$.

Оценкой $\theta_n$ параметра $\theta$ называют всякую функцию результатов наблюдений (иначе – статистику), с помощью которой делают вывод о значении параметра $\theta$:

$$\theta_n=\theta_n(x_1,x_2,\ldots,x_n).$$

Так как $x_1,x_2,\ldots,x_n$ – случайные величины, то и оценка $\theta_n$ является случайной величиной, которая зависит от закона распределения и объема выборки $n$. Оцениваемый параметр $\theta$ является постоянной величиной.

Всегда существует множество функций от результатов наблюдений $x_1,x_2,\ldots,x_n$, которые можно предложить в качестве оценки параметра $\theta$. Например, для математического ожидания в качестве оценки $\theta_n$ по выборке можно взять среднюю арифметическую результатов наблюдений $\overline{X}$, моду $M_0$, медиану $M_e$ и т. д.

Какими свойствами должна обладать оценка $\theta_n$?

Так как $\theta_n$ – случайная величина, то невозможно предсказать индивидуальное значение оценки в данном частном случае. Поэтому о качестве оценки следует судить не по ее индивидуальным значениям, а по распределению ее значений при достаточно большом числе испытаний, т. е. по выборочному распределению оценки.

### 2.2.2. Наиболее важные свойства оценок

Оценка $\theta_n$ параметра θ называется *несмещенной*, если ее математическое ожидание равно оцениваемому параметру, т. е.

$$M(\theta_n) = \theta .$$

В противном случае оценка называется *смещенной*. Если это равенство не выполняется, то оценка $\theta_n$, полученная по разным выборкам, будет либо завышать θ, если $M(\theta_n) > \theta$, либо занижать его, если $M(\theta_n) < \theta$. Таким образом, требование несмещенности гарантирует отсутствие систематических ошибок при оценивании.

Являются ли оценки математического ожидания μ и дисперсии $\sigma^2$ несмещенными?

Для матожидания оценкой, полученной по выборке, является среднее арифметическое $\overline{X}$:

$$M[\overline{X}] = M\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} Mx_i = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu ,$$

что говорит о том, что $\overline{X}$ – несмещенная оценка для матожидания.

Для дисперсии $\sigma^2$ оценкой, полученной по выборке, является $S^2$:

$$M[S^2] = \frac{1}{n} M\left[\sum_{i=1}^{n}((x_i - \mu) - (\overline{X} - \mu))^2\right] =$$
$$= \frac{1}{n} M\left[\sum_{i=1}^{n}(x_i - \mu)^2 - 2(\overline{X} - \mu)\sum_{i=1}^{n}(x_i - \mu) + \sum_{i=1}^{n}(\overline{X} - \mu)^2\right] =$$
$$= \frac{1}{n} M\left[\sum_{i=1}^{n}(x_i - \mu)^2 - 2(\overline{X} - \mu)(n\overline{X} - n\mu) + n(\overline{X} - \mu)^2\right] =$$
$$= \frac{1}{n} M\left[\sum_{i=1}^{n}(x_i - \mu)^2 - 2n(\overline{X} - \mu)^2 + n(\overline{X} - \mu)^2\right] =$$
$$= \frac{1}{n} M\left[\sum_{i=1}^{n}(x_i - \mu)^2 - n(\overline{X} - \mu)^2\right] =$$
$$= \frac{1}{n}\sum_{i=1}^{n} M(x_i - \mu)^2 - nM(\overline{X} - \mu)^2 =$$
$$= \frac{1}{n}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \frac{1}{n}(n\sigma^2 - \sigma^2) =$$
$$= \sigma^2\left(\frac{n-1}{n}\right),$$

так как в случае независимых и одинаково распределенных с дисперсией $\sigma^2$ наблюдений $x_1, x_2, \ldots, x_n$ имеем

$$D\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} Dx_i = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

(см. свойства дисперсии, с. 85).

Для устранения смещения в оценке дисперсии достаточно оценку $S^2$ домножить на $\frac{n}{n-1}$, тогда несмещенной оценкой генеральной дисперсии будет выборочная дисперсия

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2 .$$

Коэффициент $\frac{1}{n-1}$ особенно важен для выборок малого объема.

Для оценок важными являются еще свойства *состоятельности* и *эффективности*, определения которых приводятся ниже.

Оценка $\theta_n$ параметра $\theta$ называется *состоятельной*, если она удовлетворяет закону больших чисел, т. е. сходится по вероятности к оцениваемому параметру

$$\lim_{n\to\infty} P(|\theta_n - \theta| < \varepsilon) = 1$$

или

$$\theta_n \xrightarrow[n\to\infty]{P} \theta.$$

Если оценка состоятельна, то практически достоверно, что при достаточно большом $n$ $\theta_n \approx \theta$.

*Несмещенная* оценка $\theta_n$ параметра $\theta$ является *эффективной*, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра $\theta$, вычисленных по выборкам одного и того же объема $n$. Так как для несмещенной оценки $M(\theta_n - \theta)^2$ есть дисперсия $\sigma^2_{\theta_n}$, то эффективность является решающим свойством, определяющим качество оценки.

В качестве статистических оценок параметров генеральной совокупности желательно использовать оценки, удовлетворяющие одновременно требованиям *несмещенности*, *состоятельности* и *эффективности*.

Для случайной величины $X$, распределенной по нормальному закону, среднее арифметическое $\overline{X}$ является несмещенной, состоятельной и эффективной оценкой для математического ожидания.

Однако на практике не всегда оценки удовлетворяют всем трем требованиям. Может оказаться, что даже если эффективная оценка существует, то формулы для ее вычисления оказываются слишком сложными, и тогда используют оценку, дисперсия которой несколько больше. Иногда, в интересах простоты расчетов, применяются незначительно смещенные оценки. Выбору оценки всегда должно предшествовать ее критическое рассмотрение.

### 2.2.3. Понятие о точечной оценке

Выборочная характеристика, используемая в качестве приближенного значения неизвестной генеральной характеристики, называется ее *точечной статистической оценкой*.

«Точечная» означает, что оценка представляет собой число или точку на числовой оси.

Точечные оценки могут быть получены с использованием метода моментов, метода максимального правдоподобия и метода наименьших квадратов.

**Метод моментов** состоит в том, что выборочные моменты приравниваются к теоретическим моментам распределения.

*Метод максимального правдоподобия*

Основу метода составляет функция правдоподобия, выражающая плотность вероятности совместного появления результатов выборки $x_1, x_2, \ldots, x_n$,

$$L(x_1, x_2, \ldots, x_n) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_n, \theta).$$

Согласно методу максимального правдоподобия в качестве оценки неизвестного параметра θ принимается такое значение $\theta_n$, которое максимизирует функцию $L$.

Нахождение оценки $\theta_n$ упрощается, если максимизировать не саму функцию $L$, а $\ln L$, так как максимум обеих функций достигается при одном значении θ.

Для отыскания оценки параметра θ необходимо решить систему уравнений правдоподобия, получаемую приравниванием производных по параметру нулю:

$$\frac{d \ln L}{d\theta} = 0,$$

а затем отобрать то решение, которое обращает функцию $\ln L$ в максимум.

*Метод наименьших квадратов*

Метод наименьших квадратов является частным случаем метода максимального правдоподобия и заключается в том, что оценка определяется из условия минимизации суммы квадратов отклонений выборочных данных от определяемой оценки.

Оценка $\theta_n$ определяется из условия минимизации суммы

$$u = \sum_{i=1}^{n} (x_i - \theta_n)^2 \to \min.$$

Метод наименьших квадратов получил широкое применение в практике, так как хорошо разработан в плане вычислительной реализации.

### 2.2.4. Понятие об интервальных оценках

Вычисляя на основании результатов наблюдений точечную оценку

$\theta$ неизвестного генерального параметра $\theta$, мы понимаем, что оценка $\theta$ является приближенным значением $\theta$. Если для большого объема выборки точность приближения бывает достаточной, то для выборок малого объема вопрос о точности оценок очень важен. В математической статистике он решается следующим образом.

По выборке находится точечная оценка $\theta_n$ неизвестного $\theta$. Затем задаются вероятностью $P = 1 - \alpha$ и по определенным правилам находят число $\varepsilon > 0$, чтобы выполнялось соотношение $P(|\theta_n - \theta| < \varepsilon) = 1 - \alpha$ или

$$P(\theta_n - \varepsilon < \theta < \theta_n + \varepsilon) = 1 - \alpha. \qquad (2.32)$$

Из приводимых соотношений видно, что абсолютная погрешность оценки $\theta_n$ не превосходит числа $\varepsilon$. Это утверждение верно с вероятностью $P = 1 - \alpha$. Число $\varepsilon$ называется *точностью оценки* $\theta$.

Числа $\theta_n - \varepsilon$, $\theta_n + \varepsilon$ называются *доверительными границами*, интервал $(\theta_n - \varepsilon,\ \theta_n + \varepsilon)$ – *доверительным интервалом*.

Вероятность $P = 1 - \alpha$ называется *доверительной вероятностью*, или *надежностью интервальной оценки*. Величина $\alpha$ называется *уровнем значимости*. Доверительные границы могут изменяться при изменении объема выборки, кроме того, они могут изменяться при изменении вероятности $P = 1 - \alpha$. При этом чем шире интервал, тем точность оценивания хуже. Генеральная характеристика $\theta$ – постоянная величина.

Соотношение (2.32) следует читать так: вероятность того, что $\theta_n - \varepsilon$, $\theta_n + \varepsilon$ накроет характеристику $\theta$, равна $P = 1 - \alpha$.

На рис. 2.9 друг над другом изображены доверительные интервалы для параметра $\theta$, построенные для разных выборок; центры интервалов – это выборочные значения оценки $\theta_n$.

Надежность принято выбирать равной 0.95, 0.99, 0.999, соответственно уровень значимости $\alpha$ = 0.05, 0.01, 0.001. В приведенном соотношении (2.32) доверительные границы симметричны относительно точечной оценки $\theta_n$. Рассмотренные доверительные интервалы являются двусторонними. На практике не всегда доверительные интервалы являются симметричными, кроме того, не всегда являются *двусторонними*. В этом случае они называются *односторонними*.

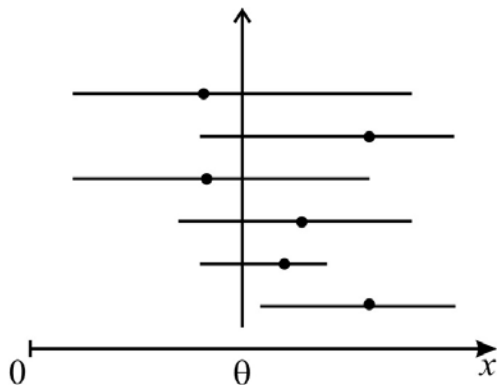


Рис. 2.9

### 2.2.5. Построение доверительного интервала для генерального среднего

Пусть $X$ – случайная величина, имеющая нормальный закон распределения с параметрами μ, σ, т. е. $X = N(\mu,\sigma)$. Будем предполагать, что наблюдения над этой величиной независимы и проводятся в одинаковых условиях, т. е. возможные результаты $X_1, X_2, \ldots, X_n$ этих наблюдений обладают следующими свойствами:

- $X_1, X_2, \ldots, X_n$ – независимые случайные величины;
- закон распределения любой из величин $X_1, X_2, \ldots, X_n$ совпадает с законом распределения величины $X$, т. е.

$$X_1 = N(\mu,\sigma),\ X_2 = N(\mu,\sigma), \ldots,\ X_n = N(\mu,\sigma).$$

### 2.2.6. Интервальная оценка математического ожидания нормального распределения при известной дисперсии

Как результат центральной предельной теоремы следующая $Z$-формула используется в данном случае

$$Z = \frac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}}, \tag{2.33}$$

откуда имеем

$$\mu = \overline{X} - Z\frac{\sigma}{\sqrt{n}}. \tag{2.34}$$

Так как выборочное среднее может быть больше или меньше, чем генеральный параметр, то предыдущее выражение берется в следующей форме:

$$\overline{X} \pm Z\frac{\sigma}{\sqrt{n}}. \tag{2.35}$$

Отсюда получаем доверительный интервал в виде

$$\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \tag{2.36}$$

где $\alpha$ – уровень значимости, изображаемый площадью под кривой нормального распределения вне площади, соответствующей доверительной вероятности;

$\alpha/2$ – площадь под кривой нормального распределения на правом и на левом хвостах распределения (рис. 2.10).

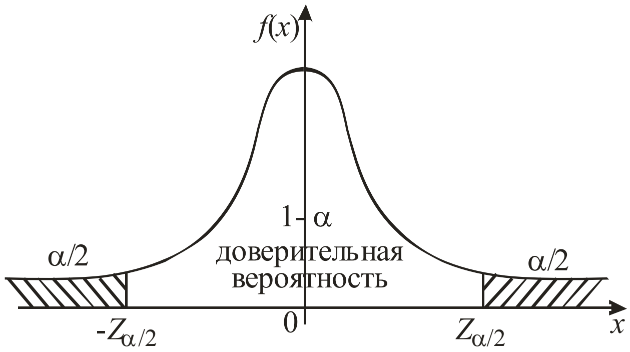


Рис. 2.10

Уровень значимости $\alpha$ используется, чтобы определить положение $Z$, значение которого определяется из таблицы функции Лапласа (см. Приложение, табл. I).

Если мы хотим определить 95 %-й доверительный интервал для $\mu$, то это означает, что из ста интервалов, построенных по случайным выборкам, взятым из генеральной совокупности, 95 интервалов будут накрывать генеральный параметр $\mu$, а 5 интервалов – нет.

В данном случае формула в виде вероятностного утверждения имеет вид

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha = 0.95. \tag{2.37}$$

Откуда $\alpha = 0.05$; $\alpha/2 = 0.025$.

Распределение выборочного среднего для вероятности 0,95 имеет вид (рис. 2.11)

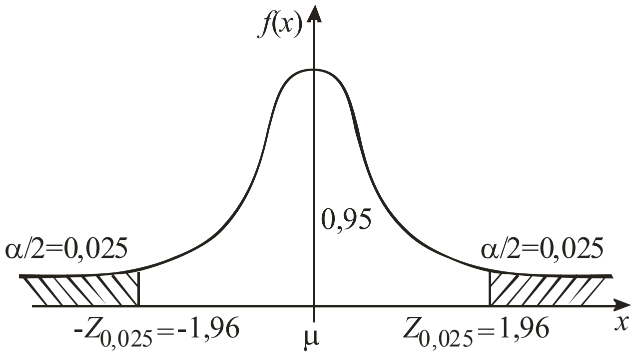


Рис. 2.11

**Пример.** Случайная величина $X$ распределена по $N(\mu, 1.1)$. По случайной выборке определено среднее арифметическое $\overline{X} = 4.26$, $n = 60$. Определить 95 %-й доверительный интервал для μ и точность оценивания μ.

*Решение*

$$P\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha = 0.95;$$

$$\alpha = 0.05, \ \alpha/2 = 0.025;$$

$$Z_{1-\alpha/2} = Z_{0.975} = 1.96; \qquad Z_{0.025} = -1.96.$$

Доверительный интервал имеет вид

$$4.26 - 1.96\frac{1.1}{\sqrt{60}} < \mu < 4.26 + 1.96\frac{1.1}{\sqrt{60}}.$$

Точность оценивания $1.96 \cdot \dfrac{1.1}{\sqrt{60}} = 0.28$.

Окончательно имеем

$$4.26 - 0.28 \le \mu \le 4.26 + 0.28;$$

$$3.98 \le \mu \le 4.54.$$

*Примечание.* Приводимая формула $Z$ для выборочных средних может быть использована для выборок большого размера, несмотря на форму генерального распределения.

Эта же формула используется для выборок малого размера, если генеральная совокупность нормально распределена и генеральное σ известно.

### 2.2.7. Доверительный интервал для оценки генерального среднего, если неизвестно σ

Если генеральное среднеквадратическое отклонение неизвестно, то вместо него используют выборочное среднеквадратическое отклонение $S$.

В случае, если объем выборки большой ($n > 30$), выборочное стандартное отклонение $S$ является хорошей оценкой генерального стандартного отклонения σ, и доверительный интервал для оценки математического ожидания μ, если σ не известна и объем выборки большой, имеет вид

$$\bar{X} - Z_{\alpha/2}\frac{S}{\sqrt{n}} \le \mu \le \bar{X} + Z_{\alpha/2}\frac{S}{\sqrt{n}}. \tag{2.38}$$

Если имеем выборку малого объема, то, несмотря на то, что генеральная совокупность распределена по нормальному закону, приведенную выше формулу использовать нельзя. В этом случае выражение

$$Z = \frac{\bar{X} - \mu}{\dfrac{S}{\sqrt{n}}}$$

не дает нормального распределения для выборок малого объема, если даже генеральная совокупность распределена по нормальному закону, потому что генеральное стандартное отклонение неизвестно.

Вводится $t$-статистика для выборочного среднего:

$$t = \frac{\bar{X} - \mu}{\dfrac{S}{\sqrt{n}}}. \tag{2.39}$$

Когда генеральное стандартное отклонение неизвестно, выборка малого объема и генеральное распределение является нормальным, статистика $t$ имеет распределение Стьюдента.

Распределение Стьюдента, или $t$-распределение, является на самом деле серией распределений, так как каждый объем выборки дает свое распределение. На рис. 2.12 представлены два $t$-распределения для разных объемов выборки и одно нормальное распределение.

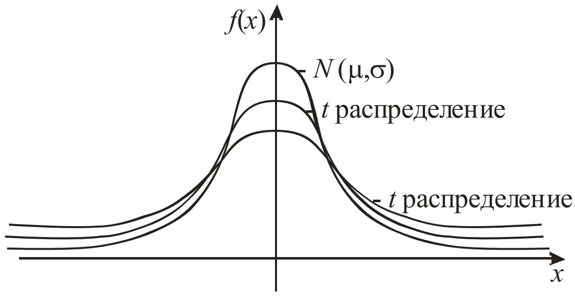


Рис. 2.12

Для определения доверительного интервала для генерального среднего, когда σ неизвестно и генеральная совокупность распределена по нормальному закону для выборок малого объема ($n$ < 30), используется $t$-статистика.

Доверительный интервал в этом случае имеет вид

$$\overline{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \qquad (2.40)$$

где $\alpha$ – уровень значимости;

$n-1$ – число степеней свободы.

Под числом степеней свободы понимается количество значений варьирования наблюдений, которые могут принимать «произвольные» значения, не изменяя общего уровня, около которого эти значения варьируют.

**Пример.** Необходимо вычислить среднее арифметическое по следующим наблюдениям:

0,24 0,02 0,01 0,22 0,04 0,14 0,18;

$$\overline{X} = \frac{0{,}95}{7} = 0{,}136.$$

Если перед нами поставлена задача произвольно отобрать еще одну совокупность таких же 7-ми значений, не изменяя при этом вычисленного уровня (т. е. $\overline{X} = 0.136$), то свободно варьирующих значений было бы не 7, а 6. Седьмое значение должно быть таким, чтобы общая сумма всех значений оказалась бы неизменной, т. е. $\sum x_i = 0.95$. Число степеней свободы в данном случае равно $7-1=6$, или объем выборки $n$ минус 1.

Значение $t_{\alpha/2,n-1}$ определяется по таблице $t$-распределения Стьюдента. Если должен быть вычислен, например, 90 %-й доверительный интервал, то общая площадь на двух концах распределения составляет 10 %. Таким образом, уровень значимости $\alpha = 0.10$; $\alpha/2 = 0.05$ (рис. 2.13).

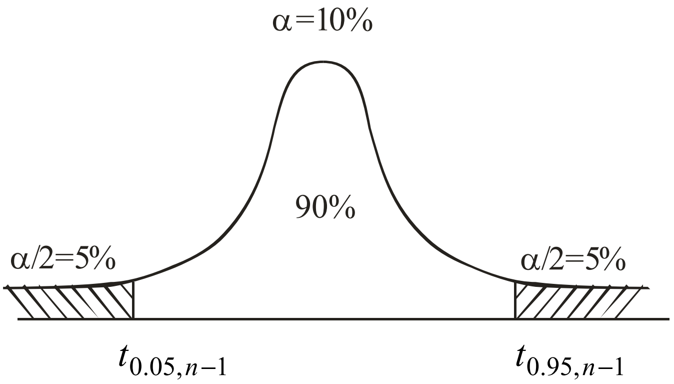


Рис. 2.13

Таблица $t$-распределения Стьюдента представляет собой значения переменной $t$ для соответствующего числа степеней свободы и уровня значимости $\alpha/2\,(0.1;\ 0.05;\ 0.025;\ 0.01;\ 0.005)$ (см. табл. III Приложения).

Пример. Значение $t$ находится на пересечении соответствующего числа степеней свободы и уровня значимости. Если определяется 90 %-й доверительный интервал по выборке объема $n = 25$, то $\alpha = 1 - 0.1$, $\alpha/2 = 0.05$; число степеней свободы $n - 1 = 25 - 1 = 24$, то $t_{0.05,24} = 1.711$ (табл. 2.7).

Таблица 2.7

| Число степеней свободы | $t_{0.1}$ | $t_{0.05}$ | $t_{0.025}$ | $t_{0.01}$ | $t_{0.005}$ |
|---|---|---|---|---|---|
| . | | . | | | |
| . | | . | | | |
| . | | . | | | |
| 23 | | . | | | |
| 24 | | 1.711 | | | |
| 25 | | . | | | |
| . | | . | | | |
| . | | . | | | |

*Примечание*. Данная таблица составлена для двусторонней критической области, что будет соответствовать доверительной вероятности $P = 1 - \alpha$.

**Пример.** Определить 90 %-й доверительный интервал для генерального среднего по выборке объема $n = 18$ из нормально распределенной генеральной совокупности для $\overline{X} = 13.56;\quad S = 7.8$.

**Решение**

Значение $t_{0.05,17} = 1.740$;

$$13.56 - 1.740 \cdot \frac{7.8}{\sqrt{18}} \le \mu \le 13.56 + 1.740 \cdot \frac{7.8}{\sqrt{18}};$$

$$10.36 \le \mu \le 16.76;$$

$$P(10.36 \le \mu \le 16.76) = 0.9.$$

*Примечание*. Кроме указанной выше таблицы, существуют таблицы для односторонней критической области, для которой $t_{\alpha,k}$ является решением уравнения

$$P = (t > t_{\alpha,k}) = \int_{t_{\alpha,k}}^{\infty} f_t(x)dx = \alpha.$$

При интервальном оценивании генерального среднего μ возможно решение следующих проблем:

– определение точности оценивания μ;

– определение объема выборки, необходимого для достижения необходимой точности;

– определение доверительного интервала μ.

Точность оценивания определяется в виде:

- $\varepsilon = Z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}$ – для выборок из нормального распределения генеральной совокупности, несмотря на объем выборки, если известно σ;
- $\varepsilon = Z_{\alpha/2}\cdot\frac{S}{\sqrt{n}}$ – для выборок большого объема, если неизвестно σ (выборки могут быть необязательно из нормально распределенной совокупности);
- $\Delta = t_{\alpha/2,n-1}\cdot\frac{S}{\sqrt{n}}$ – для выборок из нормально распределенной совокупности, выборки малого объема и σ неизвестно. Соответственно объем выборки определяется по формулам:

$$n = \frac{Z_{\alpha/2}\cdot\sigma^2}{\varepsilon^2}; \tag{2.41}$$

$$n = \frac{Z^2_{\alpha/2}\cdot S^2}{\varepsilon^2}; \tag{2.42}$$

$$n = \frac{t^2_{\alpha/2,n-1}\cdot S^2}{\varepsilon^2}. \tag{2.43}$$

### 2.2.8. Частость как точечная оценка вероятности события

Обозначим через $p$ неизвестную вероятность появления случайного события А в единичном испытании.

Приближенное значение $\hat{p}$ вероятности $p$ определяется в виде

$$\hat{p} = \frac{m}{n}, \tag{2.44}$$

где $\hat{p}$ – частость появления события $A$ в $n$ испытаниях;

$m$ – число появления события $A$ в $n$ испытаниях.

Серия независимых испытаний, в каждом из которых событие $A$ происходит с вероятностью $q = 1 - p$, является последовательностью испытаний Бернулли.

***Теорема.*** *Пусть m – число наступлений события A в n независимых испытаниях, p – вероятность наступления события A в каждом из испытаний. Тогда* $\hat{p} = \frac{m}{n}$ *– состоятельная, несмещенная и эффективная оценка вероятности p.*

### 2.2.9. Интервальная оценка вероятности события

«Хорошей» точечной оценкой вероятности $p$ события $A$ является частость $\hat{p} = \frac{m}{n}$, где $n$ – общее число испытаний, а событие $A$ может произойти с вероятностью $p$ или не произойти с вероятностью $q = 1 - p$ (последовательность испытаний Бернулли).

Интервальная оценка для $p$ задается в виде

$$P(p_1 < p < p_2) = 1 - \alpha,$$

где $(p_1, p_2)$ – границы интервала для вероятности $p$, отвечающие надежности $1 - \alpha$, $\alpha$ – уровень значимости.

Интервальная оценка зависит от объема выборки $n$.

### 2.2.10. Интервальная оценка вероятности для большого числа испытаний Бернулли

Так как $A$ – случайное событие, то $m$ – число появлений $A$ в $n$ испытаниях – тоже случайно.

При выполнении условий (грубых), что $n$ порядка нескольких десятков, а $np > 10$, распределение величины $m$ в силу локальной и интегральной теорем Муавра-Лапласа близко к нормальному распределению с математическим ожиданием, равным $np$, и дисперсией, равной $npq$, т. е.

$$m \approx N(np, \sqrt{npq}).$$

При делении $m$ на $n$ его распределение не изменяется, изменяются только его параметры. Поэтому при большом $n$ распределение частости $\hat{p} = \frac{m}{n}$, так же как и распределение частоты $m$, близко к нормальному, но с математическим ожиданием

$$Mp = M\left(\frac{m}{n}\right) = \frac{1}{n}M(m) = \frac{1}{n}np = p \qquad (2.45)$$

и дисперсией

$$Dp = D\left(\frac{m}{n}\right) = \frac{1}{n^2}D(m) = \frac{1}{n^2}npq = \frac{pq}{n}. \qquad (2.46)$$

Таким образом, при большом числе $n$ испытаний Бернулли $p \approx N\left(p, \sqrt{p\frac{q}{n}}\right)$ и распределение величины $\dfrac{p - \hat{p}}{\sqrt{pq/n}}$ близко к нормальному с нулевым математическим ожиданием и нулевой дисперсией.

Из условия

$$P\left(\left|\frac{p - \hat{p}}{\sqrt{pq/n}}\right| < Z_\alpha\right) = 1 - \alpha \qquad (2.47)$$

определяются границы в виде

$$p_1 = \hat{p} - Z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}; \qquad (2.48)$$

$$p_2 = \hat{p} + Z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}. \qquad (2.49)$$

Отсюда получается, что частость $\hat{p} = m/n$ определяет значение неизвестной вероятности с точностью $\varepsilon = Z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}$.

**Пример.** Событие $A$ в серии из $n = 100$ испытаний Бернулли произошло $m = 78$ раз. Найти интервальную оценку для вероятности $p$ события $A$ с надежностью 0.9.

**Решение.** Точечная оценка $\hat{p} = 78/100 = 0.78$ для $P = 0.9$; $Z_{\alpha/2} = 1.64$;

$\varepsilon = 1.64\sqrt{0.78 \cdot 0.22/100} = 0.068$;

$p_1 = 0.78 - 0.068 = 0.71$; $p_2 = 0.78 + 0.068 = 0.848$.

### 2.2.11. Интервальная оценка вероятности при малом числе *n* испытаний Бернулли

Законом распределения числа $X$ события $A$ в $n$ испытаниях в данном случае является биномиальный закон распределения, т. е. вероятность

$$P(X=m)=C_n^m p^m (1-p)^{n-m}; \quad m=0,1,2,\ldots,n.$$

При доверительной вероятности $\gamma = 1-\alpha$ вероятность

$$P=(p_1<p<p_2)=\gamma. \tag{2.50}$$

Так как при $p \neq 0{,}5$ и $n$ малом биномиальное распределение несимметрично, то в качестве доверительного интервала для $p$ берут такой интервал $(p_1, p_2)$, что вероятность попадания левее $p_1$ и правее $p_2$ одна и та же и равна $\frac{1-\gamma}{2}$:

$$\sum_{m=\tilde{m}} C_n^m p_1^m (1-p_1)^{n-m} = \frac{1-\gamma}{2}; \qquad \sum_{m_0=0}^{\tilde{m}} C_n^m p_2^m (1-p_2)^{n-m} = \frac{1-\gamma}{2}, \tag{2.51}$$

где $\tilde{m}$ – фактическое число элементов выборки, обладающих интересующим признаком.

Решение таких уравнений упрощается, если использовать таблицы для нахождения $p_1$ и $p_2$ по заданным $n$, $n-m$, $\gamma$. Фрагмент таблицы приведен ниже (табл. 2.8).

Таблица 2.8

Доверительные границы $p_1$ и $p_2$ для параметра $p$ при $\gamma = 0.95$

| n-m | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0,975 | 842 | 708 | 602 | 522 | 459 | 410 | 369 | 336 | 308 |
| | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |
| 1 | 987 | 906 | 806 | 716 | 641 | 579 | 527 | 483 | 445 | 413 |
| | 013 | 008 | 006 | 005 | 004 | 004 | 003 | 003 | 003 | 002 |
| 2 | 992 | 932 | 853 | 777 | 710 | 651 | 600 | 556 | 518 | 484 |
| | 094 | 068 | 053 | 043 | 037 | 032 | 028 | 025 | 023 | 021 |
| 3 | 994 | 947 | 882 | 816 | 755 | 701 | 652 | 610 | 572 | 538 |
| | 194 | 147 | 118 | 099 | 085 | 075 | 067 | 060 | 055 | 050 |
| 4 | 995 | 957 | 901 | 843 | 788 | 738 | 692 | 651 | 614 | 581 |
| | 284 | 223 | 184 | 157 | 137 | 122 | 109 | 099 | 091 | 084 |

**Пример.** В пяти испытаниях Бернулли событие $A$ произошло три раза. Найти с надежностью 0.95 интервальную оценку вероятности $p$ появления события $A$ в одном испытании.

**Решение.** Имеем $n=5$, $m=3$, $\gamma=0{,}95$. По таблице находим $p_1=0{,}147$; $p_2=0{,}947$. Таким образом, интервальная оценка вероятности $p$ для $\gamma=0.95$: (0.147; 0.947).

### 2.2.12. Задание для самостоятельной работы

1. Определить 99 %-й доверительный интервал для выборки объема $n = 110$ со средним арифметическим, равным 85.5, и $S = 19.3$.

Ответ: (80.75; 90.25).

2. Предположим, что $X$ нормально распределена. Используйте следующую информацию, чтобы определить 90 %-й доверительный интервал для математического ожидания:

313 320 319 340 325 310 321 329 317 311 307 318.

Определить ошибку оценивания.

Ответ: (313.376; 324.956); 5.79.

Определить 99 %-й доверительный интервал.

Ответ: (309.9676; 328.3644).

Определить ошибку оценивания.

Ответ: 9.1984.

Сравнить ширину 90 %-го интервала и 99 %-го интервала.

Сделать вывод, какая из оценок хуже.

3. Может ли исследователь использовать 100 %-й доверительный интервал? Объясните.

4. Исследователь по маркетингу утверждает, что имеет 95 %-й доверительный интервал средней стоимости ежемесячной торговли продуктами ($ 70.000, $ 200.000). Объясните, что означает это утверждение.

5. Предположим, что менеджер по поставке краски хотел оценить точное количество краски, содержащейся в one-gallon канистре. Из технических условий производителя известно, что стандартное отклонение количества краски равняется 0.02 gallon.

По выборке из 50-ти канистр определено среднее количество краски на каждую one-gallon канистру, равное 0.995 gallon.

Построить 99 %-й доверительный интервал для оценки генерального среднего количества краски, находящегося в one-gallon канистре.

Ответ: (0.987, 1.0022).

Основываясь на Ваших результатах, думаете ли Вы, что владелец склада имеет право жаловаться на изготовителя? Объясните.

6. Определить необходимый объем выборки для интервального оценивания математического ожидания, если известно, что генеральное среднеквадратическое отклонение равно 1.9, ошибка оценивания равна 0.7, доверительная вероятность равна 0.95.

Ответ: 28.

7. Какие факторы влияют на повышение точности интервального

оценивания? Объясните.

### 2.2.13. Проверка гипотез

Одним из наиболее важных статистических приемов при принятии решений является проверка гипотез. Она используется, когда необходим обоснованный вывод о преимуществах того или иного способа инвестиций, измерений, технологического процесса, эффективности нового метода обучения, лечения, управления и т. д.

Статистической гипотезой называется любое предположение о виде или параметрах неизвестного закона распределения.

Например, задачи на проверку гипотез могут формулироваться следующим образом:

– Инженер должен решить на основании выборочных данных, является ли генеральное среднее продолжительности службы определенного вида шин равным 22000 км.

– Производитель лекарств должен решить на основании выборки, выздоравливают ли 90 % пациентов, если они принимают новое лекарство от данной болезни.

– Агроном должен решить на основании эксперимента, дает ли данное удобрение более высокий урожай по сравнению с другими.

Эти проблемы могут быть решены, если они будут переведены на язык проверки гипотез.

Большинство статистических гипотез касается параметров распределения. Так, в первом случае инженер должен проверить гипотезу о том, что параметр $\theta$ экспоненциального распределения по крайней мере 22000 км.

Во втором примере производитель лекарств должен решить, является ли параметр $\theta$ биномиального распределения равным 0.9.

В третьем случае агроном должен решить, что $\mu_1 > \mu_2$, где $\mu_1$ и $\mu_2$ – генеральные средние двух нормально распределенных совокупностей.

В каждом случае должно предполагаться, что указанные распределения правильно описывают экспериментальные данные.

Кроме проверки гипотез о параметрах, могут проверяться гипотезы о виде распределения.

Так, в первом примере инженер может решить вопрос о том, действительно ли выборка, по которой он хочет сделать вывод, является выборкой из экспоненциально распределенной генеральной совокупности.

Проверка гипотез предполагает выполнение определенных шагов.

Первый шаг – это установление нулевой гипотезы $H_0$ и альтернативной (конкурирующей) $H_1$. Нулевая $H_0$ и альтернативная гипотезы устанавливаются в противоположность друг другу.

Нулевая гипотеза формулируется в отрицании различий. Так, в ранее приведенных примерах

– в первом случае $H_0: \ \theta = 22000$;

– во втором случае $H_0: \ \theta = 0.9$;

– в третьем случае $H_0: \ \mu_1 = \mu_2$.

Альтернативная гипотеза

– для первого случая может быть

$H_1: \ \theta < 22000$, или $H_1: \ \theta \geq 22000$, или $H_1: \ \theta \neq 22000$;

– для второго случая

$H_1: \ \mu_1 > \mu_2$, или $H_1: \ \mu_1 \leq \mu_2$, или $H_1: \ \mu_1 \neq \mu_2$;

– для третьего случая

$H_1: \ \theta < 0.9$, или $H_1: \ \theta \geq 0.9$, или $H_1: \ \theta \neq 0.9$.

Если гипотезы $H_0$ или $H_1$ полностью определяют значение параметра равным $\theta_0$ или $\theta_1$, то они называются *простыми*.

В этом случае они формулируются следующим образом: $H_0 : \theta = \theta_0$; $H_1 : \theta = \theta_1$.

Если гипотезы не полностью определяют параметры распределения, то они называются *сложными* (см. гипотезы $H_1$ в приведенных выше случаях).

Проверка статистических гипотез – это использование ряда правил для решения: принять нулевую гипотезу или отвергнуть в пользу альтернативной гипотезы $H_1$.

Проверка гипотез включает следующие шаги.

1. Располагая выборочными данными $x_1, x_2, \ldots, x_n$ и руководствуясь конкретными условиями рассматриваемой задачи, формулируют гипотезу $H_0$ и конкурирующую гипотезу $H_1$.

Например, если нулевая гипотеза такова:

$H_0$: математическое ожидание $\mu = 5$, $H_1$ может быть сформулирована следующим образом:

$H_1$: а) $\mu < 5$ или $\mu > 5$; б) $\mu \neq 5$.

2. Задаются вероятностью $\alpha$, которую называют *уровнем значимости*. Решение о том, можно ли считать высказывание $H_0$ справедливым для генеральной совокупности, принимается по выборочным данным,

т. е. по ограниченному ряду наблюдений, следовательно, это решение может быть ошибочным. При этом может иметь место ошибка двух родов (табл. 2.9):

– отвергают гипотезу $H_0$ и принимают гипотезу $H_1$, тогда как на самом деле гипотеза $H_0$ верна; это ошибка первого рода;

– принимают гипотезу $H_0$, тогда как верна гипотеза $H_1$; это ошибка второго рода.

Уровень значимости $\alpha$ – это вероятность ошибки 1-го рода. Вероятность $\alpha$ задается заранее некоторыми стандартными значениями: 0.05; 0.01; 0.005; 0.001.

Вероятность ошибки 2-го рода обозначается $\beta$. Зная $\alpha$, можно определить $\beta$.

Таблица 2.9

| Решение, принимаемое об $H_0$ по выборке | Гипотеза $H_0$ отвергается, принимается $H_1$ | Гипотеза $H_0$ принимается |
|---|---|---|
| Гипотеза $H_0$ верна | Ошибка 1-го рода, ее вероятность $\alpha$ | Правильное решение, его вероятность $1-\alpha$ |
| Гипотеза $H_0$ неверна, верна $H_1$ | Правильное решение, его вероятность $1-\beta$ | Ошибка 2-го рода, его вероятность $\beta$ |

3. Установление решающего правила.

Для этого используется статистика $T(x_1,\ldots,x_n)$, полученная по выборке, точное или приближенное распределение которой известно.

Устанавливается правило, по которому гипотеза $H_0$ отвергается или принимается. Это правило называется *статистическим критерием*.

Построение критерия начинается с выбора такого множества на действительной прямой, что если значение статистики принадлежит этому множеству, то принимается гипотеза $H_0$ и тем самым отвергается $H_1$. Такое множество называется *множеством принятия гипотезы* $H_0$.

Дополнительное множество к множеству принятия нулевой гипотезы называется *множеством отклонения нулевой гипотезы, или критическим множеством*.

Уровень значимости $\alpha$ определяет «размер» критической области. Размер критической области определяет также и ошибку 2-го рода $\beta$. Очевидно желание сделать как угодно малыми и $\alpha$, и $\beta$. Однако, при фиксированном объеме выборки можно сделать как угодно малой только одну из величин – или α, или β, уменьшение одной сопряжено с неизбежным увеличением другой. Положение критической области на действительной прямой зависит от формулировки гипотезы $H_1$.

На рис. 2.14 показано расположение критической области (под заштрихованной фигурой) для различных альтернативных гипотез.

Здесь $f_T(X/H_0)$ – плотность распределения статистики критерия $T$ при условии, что верна гипотеза $H_0$.

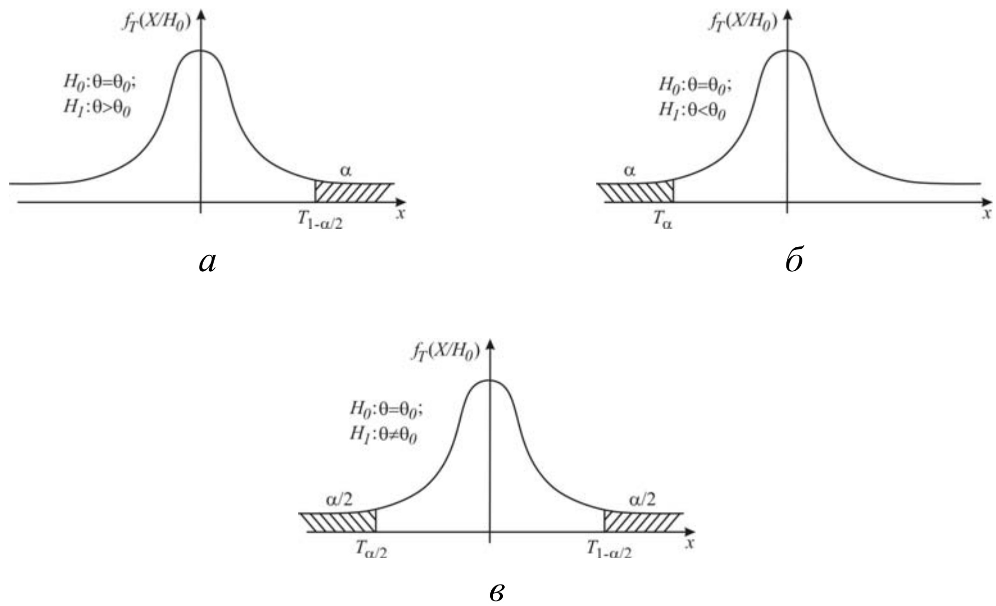


Рис. 2.14

В случаях *а* и *б* критерий называется *односторонним*. В случае *в* двусторонним. Хотя по своему назначению и по характеру решаемых задач статистические критерии очень разнообразны, их объединяет общность логической схемы, по которой они строятся и применяются.

Кратко эту схему можно описать следующим образом:

1. Выдвигаются гипотезы $H_0$ и $H_1$.

2. Задается уровень значимости α.

3. Выбирается статистика критерия $T$ для проверки гипотезы $H_0$ против $H_1$.

4. Определяется распределение статистики $T$ при условии, что верна $H_0$.

5. Исходя из распределения статистики $T$ и формулировки гипотезы $H_1$, определяется критическая область, обычно она определяется од-

ним из неравенств $T > T_{1-\alpha}$ или $T < T_{\alpha}$ или совокупностью неравенств $T < T_{1-\alpha/2}$ и $T < T_{\alpha/2}$.

6. По имеющимся данным наблюдений вычисляется значение $T^*$ статистики $T$.

7. Принимается статистическое решение с использованием пункта 5.

Если $T^*$ принадлежит критической области, то $H_0$ отвергается.

Если $T^*$ не принадлежит критической области, то $H_0$ принимается.

Одной из наиболее часто встречающихся задач проверки гипотез является задача о проверке гипотезы относительно генерального среднего.

Задача формируется, например, следующим образом. Исследователь заинтересован в определении того, что установленное (генеральное) среднее $\mu$ для выпускаемой продукции все еще является таковым. По имеющимся выборочным данным он получает среднее арифметическое $\overline{X}$. Ему необходимо проверить гипотезу $H_0 : \overline{X} = \mu$. При этом альтернативная гипотеза $H_1$ может быть сформулирована в виде

$H_1 : \overline{X} \neq \mu$ – двусторонняя гипотеза;

$H_1 : \overline{X} < \mu$ – левосторонняя гипотеза;

$H_1 : \overline{X} > \mu$ – правосторонняя гипотеза.

Критерии значимости для проверки гипотез о генеральном среднем приведены в табл. 2.10.

Таблица 2.10

Критерии значимости для проверки гипотез о генеральном среднем

| Предположения | Статистика критерия | Область принятия гипотезы $H_0$ | | |
|---|---|---|---|---|
| | | для двусторонней $H_1$ | для правосторонней $H_1$ | для левосторонней $H_1$ |
| 1. Выборка из нормальной генеральной совокупности с известной $\sigma^2$ | $Z = \frac{(\overline{X} - \mu)\sqrt{n}}{\sigma}$ | $\lvert Z \rvert < Z_{1-\alpha/2}$ | $Z < Z_{1-\alpha}$ | $Z > Z_{\alpha}$ |

| | | | | |
|---|---|---|---|---|
| 2. Выборка из нормальной генеральной совокупности с неизвестной $\sigma^2$ | $t = \dfrac{(\overline{X} - \mu)\sqrt{n}}{S}$ | $\lvert t \rvert < t_{1-\alpha/2, n-1}$ | $t < t_{1-\alpha, n-1}$ | $t > t_{\alpha, n-1}$ |
| 3. Выборка из произвольной совокупности большого объема | $Z = \dfrac{(\overline{X} - \mu)\sqrt{n}}{S}$ | $\lvert Z \rvert < Z_{1-\alpha/2}$ | $Z < Z_{1-\alpha}$ | $Z > Z_{\alpha}$ |

Здесь $n$ – объем выборки;

$\alpha$ – уровень значимости;

$Z_{\alpha}$ – квантиль порядка $\alpha$ стандартного нормального распределения (см. табл. II Приложения);

$t_{\alpha, n-1}$ – квантиль порядка $\alpha$ стандартного распределения Стьюдента с $n-1$ степенями свободы (см. табл. III Приложения).

**Пример.** Пусть необходимо проверить нулевую гипотезу $H_0 : \mu = 108$ против альтернативной гипотезы $H_1 : \mu \neq 108$. Распределение исследуемого признака можно считать нормальным; генеральная дисперсия $\sigma^2$ неизвестна. По полученной выборке имеем $\overline{X} = 108.85$; $S = 12.65$; $n = 117$.

**Решение.** Область принятия $H_0$ при двусторонней альтернативе определяется неравенством (см. табл. 2.11)

$$|Z| < Z_{1-\alpha/2},$$

где статистика критерия

$$Z = \frac{(\overline{X} - \mu)\sqrt{n}}{S};$$

$$Z = (108.85 - 108) \cdot \frac{\sqrt{117}}{12.65} = 0.73; \qquad Z_{1-\alpha/2} = Z_{0.975} = 1.96.$$

Так как $0.73 < 1.96$, гипотеза $H_0$ не противоречит имеющимся данным.

### 2.2.14. Гипотеза о равенстве двух средних

Предполагается, что случайные величины $X$ и $Y$ имеют нормальные законы распределения. При этом $X \sim N(\mu_1, \sigma_1)$, $Y \sim N(\mu_2, \sigma_2)$.

$\mu_1$ и $\mu_2$ неизвестны, относительно $\sigma_1$ и $\sigma_2$ возможны две ситуации:

$\sigma_1$ и $\sigma_2$ – известны;

$\sigma_1$ и $\sigma_2$ – неизвестны.

Критерии проверки гипотез о равенстве двух средних с учетом этих условий приводятся в табл. 2.11.

Таблица 2.11

Критерии проверки гипотез о равенстве двух средних

| $H_0$ | Предположения | Статистика критерия | $H_1$ | Область принятия $H_0$ |
|---|---|---|---|---|
| $\mu_1 = \mu_2$ | $\sigma_1$ и $\sigma_2$ – известны | $Z = \dfrac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_x} + \dfrac{\sigma_2^2}{n_y}}}$ | $\mu_1 > \mu_2$<br>$\mu_1 < \mu_2$<br>$\mu_1 \neq \mu_2$ | $Z < Z_{1-\alpha}$;<br>$Z > Z_{\alpha}$;<br>$\lvert Z \rvert < Z_{1-\alpha/2}$ |
| | $\sigma_1$ и $\sigma_2$ – неизвестны | $t = \dfrac{\overline{X} - \overline{Y}}{S\sqrt{\dfrac{1}{n_x} + \dfrac{1}{n_y}}}$, где<br>$S^2 = \dfrac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}$ | $\mu_1 > \mu_2$<br>$\mu_1 < \mu_2$<br>$\mu_1 \neq \mu_2$ | $t < t_{\alpha, n_x + n_y - 2}$;<br>$t > t_{\alpha, n_x + n_y - 2}$;<br>$\lvert t \rvert < t_{\alpha/2, n_x + n_y - 2}$ |

**Пример.** Произведены две выборки урожая пшеницы: при своевременной уборке и уборке с некоторым опозданием. В первом случае, при наблюдении восьми участков, выборочная средняя урожайность составила 16.2 ц/га, $S_x = 3.2$ ц/га. Во втором случае, при наблюдении девяти участков, $\overline{Y} = 13.9$ ц/га, $S_Y = 2.1$ ц/га. На уровне значимости $\alpha = 0.05$ выяснить влияние своевременности уборки урожая на среднее значение урожайности.

**Решение**

$H_0 : \overline{X} = \overline{Y}$;

$H_1 : \overline{X} > \overline{Y}$.

Принятие гипотезы $H_1$ будет означать существенное влияние на урожайность сроков уборки:

$$t = \frac{16.2 - 13.9}{\sqrt{\frac{7 \cdot 3.2^2 + 8 \cdot 2.1^2}{8 + 9 - 2}\left(\frac{1}{8} + \frac{1}{9}\right)}} = 1.62;$$

$$t_{0.05,9+8-2} = t_{0.05,15} = 1.75.$$

Так как $1.62 < 1.75$, то $H_0$ принимается. Это означает, что для данных выборок сроки уборки не влияют на урожайность, но этот вывод может измениться для выборок большего объема.

Если гипотезу $H_0$ принимают, то говорят, что различие выборочных средних $X$ и $Y$ незначимо и оценка математического ожидания определяется в виде

$$(\overline{X} \cdot n_x + \overline{Y} \cdot n_y)/(n_x + n_y).$$

### 2.2.15. Исключение резко выделяющихся наблюдений

Пусть $x^*, x_1, x_2, \ldots, x_n$ – выборка, где $x^*$ – резко выделяющееся наблюдение.

Необходимо решить вопрос о принадлежности резко выделяющегося наблюдения $x^*$ к остальным наблюдениям.

Для ряда наблюдения $x_1, x_2, \ldots, x_n$ определяют $\overline{X}$ и $S$.

Выдвигается гипотеза $H_0 : \overline{X} = x^*$.

Альтернативная гипотеза $H_1 : \overline{X} > x^*$ или

$H_1 : \overline{X} < x^*$.

Определяется статистика
$$t = \frac{\overline{X} - x^*}{S}. \qquad (2.52)$$

Статистика $t$ имеет распределение Стьюдента с $k = n - 1$ степенями свободы. Гипотеза $H_0$ отвергается, если $|t| > t_{1-2\alpha,n-1}$, и принимается, если $|t| < t_{1-2\alpha,n-1}$.

**Пример.** Имеются следующие данные об урожайности пшеницы на 8-ми опытных участках одинакового размера:

26.5; 26.2; 35.9; 30.1; 32.3; 29.3; 26.1; 25.0.

Есть основание предполагать, что значение урожайности третьего участка ($x^* = 35.9$) зарегистрировано неверно. Является ли это значение резко выделяющимся для $\alpha = 0.05$?

**Решение.** Исключив $x^* = 35.9$, определяем

$$\overline{X} = \frac{1}{7}(26.5 + 26.2 + 30.1 + 32.3 + 29.3 + 26.1 + 25.0) = 27.93;$$

$$H_0 : \overline{X} = x^*;$$

$$H_1 : \overline{X} < x^*;$$

$$S = 2.67; \; t = \frac{27.93 - 35.9}{2.67} = -2.98;$$

$$t_{1-2\alpha, n-1} = t_{1-0.1,6} = t_{0.9,6} = 1.94; \qquad |t| = 2.98 > 1.94.$$

$H_0$ отвергается, т. е. наблюдение $x^* = 35.9$ является резко выделяющимся и его следует отбросить.

### 2.2.16. Проверка гипотезы о равенстве долей признака в двух совокупностях

Сравнение долей признака в двух совокупностях – часто встречающаяся на практике задача.

Например, если выборочная доля признака в одной совокупности отличается от такой же доли в другой совокупности, то указывает ли это на то, что наличие признака в одной совокупности действительно вероятнее, или полученное расхождение долей является случайным?

Задача ставится следующим образом. Имеется две совокупности, генеральные доли признака, в которых равны соответственно $p_1$ и $p_2$. Необходимо проверить нулевую гипотезу о равенстве генеральных долей, т. е. $H_0 : p_1 = p_2$.

Гипотеза $H_1$ формулируется так же, как при проверке средних, т. е.

$$H_1 : p_1 < p_2; \qquad H_1 : p_1 > p_2; \qquad H_1 : p_1 \neq p_2.$$

Для проверки гипотезы $H_0$ из двух генеральных совокупностей взяты две независимые выборки достаточно большого объема ($n_1$ и $n_2$). Выборочные доли (частость) соответственно равны

$$W_1 = \frac{m_1}{n_1} \text{ и } W_2 = \frac{m_2}{n_2}, \tag{2.53}$$

где $m_1$ и $m_2$ – число элементов в первой и второй выборках, обладающих данным признаком.

При достаточно больших $n_1$ и $n_2$ выборочные доли $W_1$ и $W_2$ имеют приближенно нормальный закон распределения с $\mu_1 = p_1$, $\mu_2 = p_2$ и

$$\sigma_1^2 = \frac{p_1 \cdot (1 - p_1)}{n_1}, \qquad \sigma_2^2 = \frac{p_2 \cdot (1 - p_2)}{n_2}. \tag{2.54}$$

При справедливости гипотезы $H_0 : p_1 = p_2$ разность $W_1 - W_2$ имеет нормальный закон распределения с матожиданием

$M[W_1 - W_2] = p - p = 0$ и дисперсией $\sigma^2 = \sigma_1^2 + \sigma_2^2 = p \cdot (1 - p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$,

поэтому вводится статистика $Z = \frac{W_1 - W_2}{\sigma^2_{W_1 - W_2}} = \frac{W_1 - W_2}{\sqrt{\hat{p} \cdot \left(1 - \hat{p}\right) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$, которая имеет стандартное нормальное распределение $N(0, 1)$. В качестве $\hat{p}$ выбирается значение

$$\hat{p} = \frac{m_1 + m_2}{n_1 + n_2}. \tag{2.55}$$

**Пример.** Контрольную работу по высшей математике по индивидуальным вариантам выполняли студенты двух групп первого курса. В 1-й группе было предложено 105 задач, из которых верно решено 60, во второй группе из 140 предложенных задач верно решено 69. На уровне значимости 0.02 проверить гипотезу об отсутствии существенных различий в усвоении учебного материала студентами обеих групп.

**Решение**

$H_0 : p_1 = p_2 = p$;

$H_1 : p_1 \neq p_2$.

Оценкой $p$ является $\hat{p} = \frac{60 + 69}{105 + 140} = \frac{129}{245} = 0.527$. Выборочные доли для каждой группы

$$W_1 = \frac{m_1}{n_1} = \frac{60}{105} = 0.571; \qquad W_2 = \frac{m_2}{n_2} = \frac{69}{140} = 0.493;$$

$$Z = \frac{0.571 - 0.493}{\sqrt{0.527(1 - 0.527)\left(\frac{1}{105} + \frac{1}{140}\right)}} = 1.21; \qquad Z_{1-\alpha/2} = Z_{0.99}.$$

По таблице Приложений $Z_{0.99} = 2.33$, $1.21 < 2.33$; $H_0$ принимается.

### 2.2.17. *p* – значение

Иногда более удобно производить другую процедуру, являющуюся в некотором смысле обратной к только что приведенной. А именно: вместо того, чтобы работать с фиксированным уровнем значимости α, можно найти α, соответствующее реализации случайной величины *X*. Число α, полученное таким образом, называется *p-значением*, которое определяет наименьший уровень значимости отверженной гипотезы $H_0$.

**Пример.** Предположим, что нормальная случайная величина *X*, дисперсия которой равна 10, приняла значение 25. Мы хотим проверить гипотезу $H_0 : \mu = 100$ против $H_1 : \mu \neq 100$ при уровне значимости $\alpha = 0.05$, объем выборки $n = 25$.

Так как $\dfrac{|25-10|\sqrt{25}}{\sqrt{100}} = 2.5 > Z_{0.975} = 1.96$, то критерий уровня $\alpha = 0.05$ отвергает нулевую гипотезу.

С другой стороны, решая относительно *p* уравнение

$$\frac{|25-10|\sqrt{25}}{\sqrt{100}} = 2.5 > Z_{1-p/2},$$

получаем $2.5 = Z_{1-p/2}$, откуда $1 - p/2 = 0.994$ (см. табл. II Приложения), $p = 0.012$. Таким образом, мы получаем наименьший уровень, при котором $H_0$ может быть отвергнута. С предыдущим выводом это согласуется, так как $0.05 > 0.012$.

### 2.2.18. Понятие мощности критерия

Пусть имеем $H_0 : \theta = \theta_0$;

$H_1 : \theta = \theta_1$.

Предположим, что плотность распределения $f(x, \theta_1)$ при конкурирующей гипотезе $H_1$ получена сдвигом вправо по оси *X* плотности $f(x, \theta_0)$ (рис. 2.15).

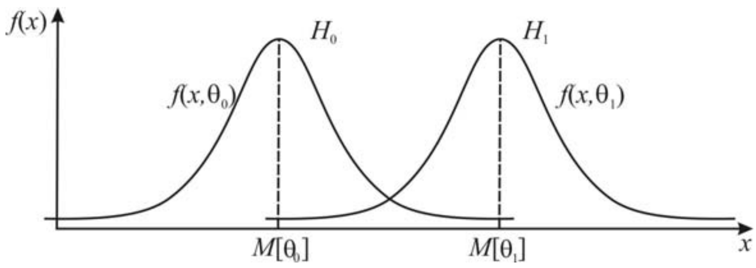


Рис. 2.15

Таким образом, математическое ожидание $M_{\theta_0}[X]$ меньше математического ожидания $M_{\theta_1}[X]$.

В качестве критического множества следует выбрать множество всех точек на числовой прямой, лежащих правее некоторой точки $x_c$, которая находится из условия, согласно которому при нулевой гипотезе (т. е. при $\theta = \theta_0$) вероятность того, что случайная величина $X$ примет значение из критического множества, равна $\alpha$. Точка $x_c$ называется *критической точкой*.

Таким образом, $x_c$ должно удовлетворять уравнению

$$\int_{x_c}^{\infty} f(x,\theta_0)\, dx = \alpha .$$

Предположим, что в действительности имеет место гипотеза $H_1$, т. е. случайная величина $X$ имеет плотность $f(x,\theta_1)$. При гипотезе $H_1$ вероятность $\beta$ того, что наблюдение попадает в область принятия $H_0$
$x < x_c$, равна площади под кривой $f(x,\theta_1)$, слева от точки $x_c$, т. е.

$\beta = \int_{-\infty}^{x_c} f(x,\theta_1)\, dx$ (рис. 2.16).

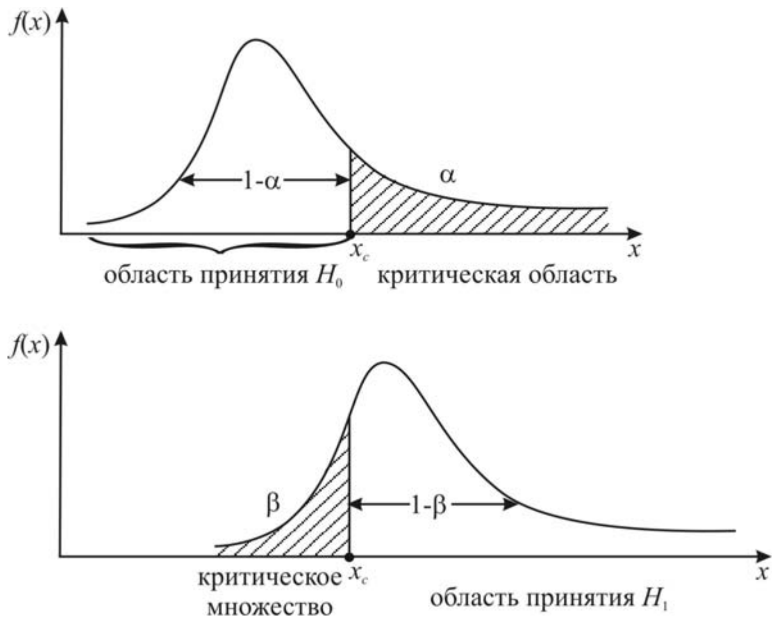


Рис. 2.16

Вероятность правильного отклонения нулевой гипотезы $H_0$, когда она неверна, называется *мощностью критерия. Указанная вероятность обозначается* $\pi$.

Таким образом, мощность критерия равна вероятности того, что наблюдение попадает в критическую область, когда справедлива гипотеза $H_1$, т. е.

$$\pi = 1 - \beta = \int_{x_c}^{\infty} f(x, \theta_1)\, dx .$$

Конечно, всегда желательно иметь возможно большую мощность, например, равную 0.9 или 0.95, при заданном уровне значимости $\alpha$.

Обычно при проверке гипотез уровень значимости $\alpha$ выбирают достаточно малым, скажем, 0.05 или 0.01.

Затем для конкретных $\theta_0$ и $\theta_1$ и данной $f(x,\theta)$ вычисляют *ошибку второго рода* $\beta$, которая часто получается значительно больше желаемого значения. В таком случае нужно быть осторожным, отвергая гипотезу $H_1$, ибо, принимая гипотезу $H_0$, мы не исключаем возможности того, что на самом деле верна гипотеза $H_1$.

На рис. 2.17 $x$ – экспериментальное значение случайной величины $X$ – принадлежит области принятия гипотезы $H_0$, поэтому мы отвергаем гипотезу $H_1$. Но $x$ могло быть значением случайной величины, распределенной по закону $f(x,\theta_1)$, и вероятность $\int_{-\infty}^{x_c} f(x,\theta_1)\,dx$ не так уж мала, тем не менее принимают гипотезу $H_0$ и верят в нее, пока не получат наблюдения, которые заставят отвергнуть ее.

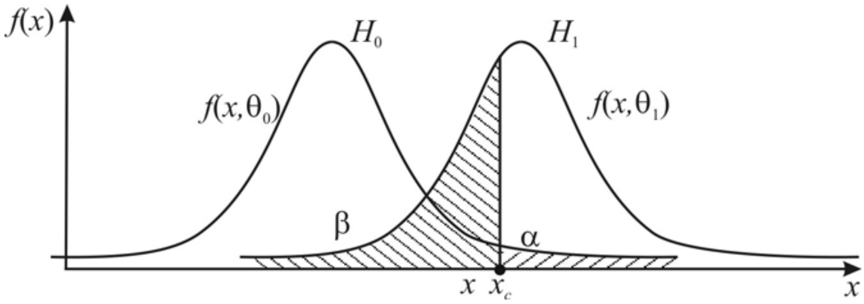


Рис. 2.17

При фиксированном объеме выборки можно сделать как угодно малой только одну из величин $\alpha$ или $\beta$. В этом случае с уменьшением одной – другая увеличивается. Лишь при увеличении объема выборки можно уменьшить одновременно $\alpha$ и $\beta$.

### 2.2.19. Задание для самостоятельной работы

1. Дать пример, когда ошибка 1-го рода более важна, чем ошибка 2-го рода.

2. Дать пример, когда ошибка 2-го рода более важна, чем ошибка 1-го рода.

3. Дано: $\overline{X} = 8^2$, $\sigma = 15$, $n = 25$.
   – Проверить гипотезу, что $\mu = 86$.
   – Определить критическую точку.

4. На станке изготавливаются детали с номинальным контролируемым размером 14 мм. Известно, что распределение контролируемого размера является нормальным $\mu = 14$ мм, $\sigma = 0.5$ мм. В течение суток было отобрано 90 деталей и сделаны замеры контролируемого параметра, средний размер которого оказался $\overline{X} = 14.3$. Можно ли считать, что станок изготавливает детали увеличенного размера с уровнем значимости 0.05.

5. Крупная торговая фирма желает открыть в новом районе города филиал. Известно, что фирма будет работать прибыльно, если ежене-

дельный средний доход жителей района превышает 400 марок. Также известно, что дисперсия дохода $\sigma^2 = 400$ марок:

– определить правило принятия решения, с помощью которого, основываясь на выборке $n = 100$ и уровне значимости $\alpha = 0.05$, можно установить, что филиал будет работать прибыльно;

– определить вероятность того, что при применении правила принятия решения, полученного при ответе на первый вопрос, будет совершена ошибка 2-го рода, если в действительности средний доход за неделю достигает 406 марок.

6. Сравниваются средние доходы фирм $X$ и $У$ в двух разных, но однотипных отраслях, имеющие нормальное распределение с $\mu_X = 1$ млн дол., $\mu_Y = 2$ млн дол. В первой отрасли по выборке из 20-ти фирм получен средний доход $\overline{X} = 1$ млн дол., а во 2-й (по выборке из 25-ти фирм) получен средний доход $\overline{Y} = 9$ млн дол. Есть ли основание отклонить гипотезу $H_0 : \mu_X = \mu_Y$ для $\alpha = 0.05$?

7. Для проверки эффективности новой технологии отобраны две группы рабочих: в 1-й группе численностью $n_1 = 50$ человек, где применялась новая технология, выборочная средняя выработка составила $\overline{X} = 85$ (изделий), во 2-й группе численностью $n_2 = 70$ человек выборочная средняя $\overline{Y} = 78$ изделий. Предварительно установлено, что дисперсии выработки в группах равны соответственно $\sigma_x^2 = 100$ и $\sigma_y^2 = 74$. На уровне значимости $\alpha = 0.05$ выяснить влияние новой технологии на среднюю производительность.