Лекция 6

РАСПРЕДЕЛЕНИЕ ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК

1. Распределение выборочного среднего

а) известна генеральная дисперсия

Можно показать, что выборочное среднее $\overline{x}$ из n независимых испытаний из нормально распределенной генеральной совокупности с генеральным средним μ и генеральной дисперсией σ^2 — $N\left(\mu,\sigma^2\right)$, само распределено нормально с параметрами μ и σ^2/n — $N\left(\mu,\sigma^2/n\right)$, то есть

$$M\left\{\overline{x}\right\}=\mu; \qquad D\left\{\overline{x}\right\}=\sigma_{\overline{x}}^2=\sigma^2/n.$$

Если генеральная дисперсия σ^2 известна, то нормированная случайна величина

$$z=\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$$

также распределена по нормальному закону, причем $M\{z\}=0$ и $D\{z\}=1$, то есть z подчиняется нормированному нормальному распределению — $N(0,1)$.

Квантиль распределения величины z_p уровня p определяется из соотношения

$$P\left(z<z_p\right)=\Phi\left(z_p\right)=p.$$

С другой стороны,

$$P\left(\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}<z_p\right)=\Phi\left(\frac{\overline{x}_p-\mu}{\sigma/\sqrt{n}}=z_p\right)=p,$$

где $\overline{x}_p$ — квантиль уровня p выборочного среднего $\overline{x}$. Следовательно,

$$\overline{x}_p=\mu+z_p\frac{\sigma}{\sqrt{n}}.$$

Вероятность попадания z в интервал $z_{p_1}<z<z_{p_2}$

$$P\left(z_{p_1}<z<z_{p_2}\right)=P\left(z_{p_1}<\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}<z_{p_2}\right)=\Phi\left(z_{p_2}\right)-\Phi\left(z_{p_1}\right)=p_2-p_1,$$

Неравенство

$$z_{p_1} < \frac{\bar{x} - \mu}{\sigma^2 / \sqrt{n}} < z_{p_2}$$

эквивалентно неравенству

$$\bar{x} - z_{p_2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - z_{p_1} \frac{\sigma}{\sqrt{n}},$$

Следовательно, можно записать

$$P\left(z_{p_1} < z < z_{p_2}\right) = P\left(\bar{x} - z_{p_2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - z_{p_1} \frac{\sigma}{\sqrt{n}}\right) = \Phi\left(z_{p_2}\right) - \Phi\left(z_{p_1}\right) = p_2 - p_1,$$

таким образом, находим доверительный интервал для генерального среднего μ. Вероятность попадания μ в интервал

$$\bar{x} - z_{p_2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} - z_{p_1} \frac{\sigma}{\sqrt{n}}$$

равна $p = p_2 - p_1$. Границы интервала $\left(\bar{x} - z_{p_2} \sigma/\sqrt{n}\,; \bar{x} - z_{p_1} \sigma/\sqrt{n}\right)$ называются *доверительными границами*, а вероятность $p = p_2 - p_1$, которую принято записывать в виде $p = 1 - \alpha$, где α — малая величина, называют *доверительной вероятностью* или *статистической надежностью.*

При построении доверительных интервалов для генерального среднего μ выбирают

$$p_1 = \alpha/2; \quad p_2 = 1 - p_1 = 1 - \alpha/2, \qquad \left(p_2 - p_1 = 1 - \alpha\right),$$

что соответствует симметричным границам доверительных интервалов относительно выборочного среднего $\bar{x}$, причем на практике полагают $\alpha = 0,1$ или $\alpha = 0,05$ реже $\alpha = 0,01$. Например, при $\alpha = 0,05$ $p_1 = 0,025$, $p_2 = 0,975$, тогда $z_{p_1} = -1,96$ и $z_{p_2} = 1,96$,

$$P\left(\bar{x} - 1,96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95,$$

при $\alpha = 0,003$ $p_1 = 0,0015$, $p_2 = 0,9985$, тогда $z_{p_1} = -3$ и $z_{p_2} = 3$, тогда

$$P\left(\bar{x} - 3\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 3\frac{\sigma}{\sqrt{n}}\right) = 0,997.$$

б) неизвестна генеральная дисперсия

При решении практических задач неизвестны как генеральное среднее μ, так и генеральная дисперсия σ^2. Поэтому для определения доверительных интервалов для μ используют выборочную дисперсию s^2. Вводят статистику, называемую t-статистикой,

$$t = \frac{\overline{x} - \mu}{s/\sqrt{N}},$$

функция распределения, которого имеет вид

$$F(t) = \frac{\Gamma\left(\left(k+1\right)/2\right)}{\sqrt{\pi k}\,\Gamma\left(k/2\right)} \int\limits_{-\infty}^{t} \left(1 + \frac{t^2}{k}\right)^{-(k+1)/2} d\,t,$$

где k — *число степеней свободы*, определяемое как разность между объемом выборки n и числом параметров, оцениваемых по выборке (в нашем случае $k = N - 1$, так как был рассчитан один параметр — выборочная дисперсия s^2); $\Gamma(x)$ — гамма-функция

$$\Gamma(x) = \int\limits_{0}^{\infty} \mathrm{e}^{-t} t^{(x-1)} dt,$$

причем $\Gamma(x+1) = x\Gamma(x)$, при целом $x = m$ $\Gamma(m) = m!$, а для полуцелого аргумента

$$\Gamma\left(m + \frac{1}{2}\right) = \sqrt{\pi}\frac{1\cdot 3\cdot 5\ \cdot\ ...\ (2m\text{ -}1)}{2^n}.$$

Соответственно плотность вероятности t-статистики равна

$$w(t) = \frac{\Gamma\left(\left(k+1\right)/2\right)}{\sqrt{\pi k}\,\Gamma\left(k/2\right)} \left(1 + \frac{t^2}{k}\right)^{-(k+1)/2},$$

Это *распределение Стьдента* или t *-распределение*. При большом числе степеней свободы k $(k \geq 30)$ t-распределение сходится к нормированному нормальному распределению $N(0,1)$.

Проведя рассуждения, аналогичные приведенным выше, находим доверительный интервал для генерального среднего μ при неизвестной генеральной дисперсии

$$\overline{x} - \frac{s}{\sqrt{N}} t_{\alpha,\,k} \leq \mu \leq \overline{x} + \frac{s}{\sqrt{N}} t_{\alpha,\,k},$$

где $t_{\alpha,\,k}$ — квантили распределения Стьдента уровня $p = 1 - \alpha/2$ для числа степеней свободы $k = N - 1$. Их значения можно найти в статистических таблицах.

2. Распределение выборочной дисперсии

Образованная с помощью выборочной дисперсии статистика, называемая χ^2-статистикой,

$$\chi^2 = \frac{(N-1)s^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2$$

подчиняется *распределению Пирсона* или χ^2*-распределению* с числом степеней свободы $k = N-1$:

$$w\left(\chi^2\right) = \frac{1}{2^{k/2}\Gamma\left(k/2\right)}\left(\chi^2\right)^{k/2-1}\exp\left(-\chi^2/2\right).$$

Генеральное среднее χ^2-распределения равно $\mu\left(\chi^2\right) = N$, дисперсия $D\left(\chi^2\right) = 2N$. Плотность вероятности $w\left(\chi^2\right)$ асимметрична, но при $k \to \infty$, то есть при $N \to \infty$, стремится к плотности нормального распределения с математическим ожиданием $k = N-1$ и дисперсией $D = 2k = 2N-2$, но медленнее, чем распределение Стьюдента.

Поскольку χ^2-статистика вероятность $P\left(\chi^2 < \chi_p^2\right)$ равна

$$P\left(\chi^2 < \chi_p^2\right) = F\left(\chi_p^2\right) = \frac{1}{2^{k/2}\Gamma\left(k/2\right)}\int_0^{\chi_p^2}\left(\chi^2\right)^{k/2-1}\exp\left(-\chi^2/2\right)d\chi^2,$$

где χ_p^2 — квантиль χ^2-распределения, причем $\chi_{p_1}^2 < \frac{s^2}{\sigma^2}\left(N-1\right) < \chi_{p_2}^2$ при большом числе степеней свободы (большом объеме выборки) $k > 30$ справедлива аппроксимация

$$\chi_p^2 = k\left(1 - \frac{2}{9k} + z_{1-p}\sqrt{\frac{2}{9k}}\right),$$

где z_{1-p} — квантиль уровня $\left(1-p\right)$ случайной величины, распределенной по нормальному закону $N\left(0,1\right)$. Квантили распределения выборочной дисперсии s_p^2 находим из квантилей χ_p^2 по формуле

$$s_p^2 = \sigma^2\frac{\chi_p^2}{N-1}.$$

Так как неравенство

$$\chi^2_{p_1} < \frac{s^2}{\sigma^2}\left(N-1\right) < \chi^2_{p_2}$$

эквивалентно неравенству

$$\frac{s^2}{\chi^2_{p_2}}\left(N-1\right) < \sigma^2 < \frac{s^2}{\chi^2_{p_1}}\left(N-1\right),$$

находим доверительный интервал для генеральной дисперсии

$$P\left(\frac{s^2}{\chi^2_{p_2}}\left(N-1\right) < \sigma^2 < \frac{s^2}{\chi^2_{p_1}}\left(N-1\right)\right) = p_2 - p_1.$$

Доверительную вероятность записывают в виде

$$p = p_2 - p_1 = 1-\alpha,$$

причем обычно принимают $p_1 = \alpha/2$ и $p_2 = 1-\alpha/2$, а $p = 1-\alpha$ равным 0,9 или 0,95.

Функция распределения χ^2 и квантили χ^2_p табулированы и играет важную роль при проверке статистических гипотез.

Проверка статистических гипотез

Заключения о результатах диагностики, контроля или измерений строятся на основании аппарата проверки *статистических гипотез*.

Под **статистической гипотезой** подразумевают любое утверждение, подлежащее проверке о виде или свойствах распределения значений измеренных величин.

Гипотезы выдвигаются на основе теоретических соображений или предыдущего опыта. Например, можно предположить, что случайные величины, полученные в результате измерений, имеют нормальное распределение.

Если для исследуемого явления или процесса сформулирована та или иная гипотеза (ее обычно называют основной и обозначают символом H_0), необходимо определить правило, согласно которому гипотеза H_0 должна быть проверена на состоятельность, то есть принята или отвергнута. В последнем случае принимается альтернативная гипотеза H_1.

С точки зрения статистики безразлично, какую гипотезу считать основной, а какую альтернативной. С целью единообразия в статистике принято считать основной гипотезу H_0, согласно которой результат контроля согласуется с заданным распределением (генеральной или иной заданной выборкой).

Метод проверки статистической гипотезы путем анализа выборки из генеральной совокупности называется **статистическим критерием** (или просто критерием). С его помощью можно установить факт совместимости свойств данной выборки, выраженных через выборочную характеристику (статистику), с гипотезой H_0, принять эту гипотезу или отвергнуть, если значение характеристики при достоверности H_0 маловероятно.

Различают критерии

- *параметрические* — в их основе лежит предположение об известном законе распределения изучаемой характеристики, например, нормальном распределении;
- *непараметрические* — эти критерии не используют информацию о законе распределения случайных величин.

В общем случае на основе экспериментальных данных строят некоторую статистику и определяют тем или иным образом ее функцию распределения, если она неизвестна. С помощью функции распределения (или плотности вероятности) этой статистики определяется интервал значений, в котором с большой вероятностью должны находиться ее значения, если гипотеза H_0 состоятельна. При состоятельности H_0 выпадение экспериментально найденного значения статистики из этого интервала маловероятно. Эту малую вероятность выпадение статистики из указанного интервала называют *уровнем значимости* и обычно обозначают через α, а множество выпадающих значений составляют *критическую область* в отличие от *области допустимых значений*, при которых гипотеза не отвергается. Таким образом, если рассчитанное по экспериментальным данным значение статистики попадает в область допустимых значений, гипотеза H_0 с вероятностью $p = 1 - \alpha$ принимается, если не попадает — отвергается. Во всех случаях речь идет о *статистически значимом* различии величин.

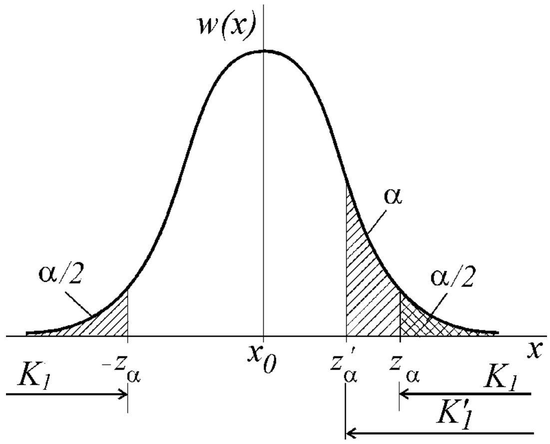

Рис. Л6.1. Критические области при использовании одностороннего и двустороннего критерия на примере нормального распределения

Критерии, определяющие критическую область, могут быть одно- и двусторонними, что поясняет рис. Л6.1.

Пусть основная гипотеза H_0 заключается в том, что наблюдаемое значение измеряемой величины, подчиняющееся нормальному закону с дисперсией σ^2, отличается от своего среднего значения x_0 с уровнем значимости α. Поскольку для нормированной случайной величины $(x - x_0)/\sigma$ справедливо

$$P\left(z_{p_1} < \frac{x - x_0}{\sigma} < z_{p_2}\right) = \Phi\left(z_{p_2}\right) - \Phi\left(z_{p_1}\right) = p_2 - p_1,$$

то

$$P\left(\frac{x - x_0}{\sigma} < z_{p_1}\right) + P\left(\frac{x - x_0}{\sigma} > z_{p_2}\right) = 1 - P\left(z_{p_1} < \frac{x - x_0}{\sigma} < z_{p_2}\right) = 1 - p_2 + p_1.$$

Полагая $p_2 = 1 - \alpha/2$ и $p_1 = \alpha/2$, будем иметь

$$P\left(x < x_0 + \sigma z_{\alpha/2}\right) + P\left(x > x_0 + \sigma z_{1-\alpha/2}\right) = \alpha/2 + \alpha/2 = \alpha,$$

Следовательно, критическая область K_1 состоит из двух подобластей, симметричных относительно x_0, — подобласть слева соответствует вероятности $\alpha/2$ того, что наблюдаемое значение x меньше $x_0 + \sigma z_{\alpha/2}$ $\left(x < x_0 + \sigma z_{\alpha/2}\right)$, подобласть справа — вероятности $\alpha/2$ того, что оно больше $x_0 + \sigma z_{1-\alpha/2}$ $\left(x > x_0 + \sigma z_{1-\alpha/2}\right)$. Через $z_{\alpha/2}$ и $z_{1-\alpha/2}$ обозначены квантили нормального распределения, причем $z_{1-\alpha/2} = -z_{\alpha/2}$.

При проверке того, что наблюдаемое значение *не больше* x_0, должен использоваться односторонний критерий. Так как

$$P\left(\frac{x - x_0}{\sigma} < z_p\right) = \Phi\left(z_p\right) = p,$$

или

$$P\left(x < x_0 + \sigma z_p\right) = p,$$

то

$$P\left(x > x_0 + \sigma z_p\right) = 1 - p.$$

Выбрав $p = 1 - \alpha$, получим

$$P\left(x > x_0 + \sigma z_{1-\alpha}\right) = \alpha$$

Критическая область K', см. рис. Л6.1, теперь является неразрывной. Аналогично, при проверке того, что результат не меньше x_0 будем иметь

$$P\left(x < x_0 + \sigma z_{\alpha}\right) = \alpha.$$

Статистический критерий основывается на свойствах выборки, то есть зависящем от случая результате, поэтому доказательство справедливости или отклонения гипотезы имеют вероятностный характер. Другими словами, принятие гипотезы не означает, что она, вообще говоря, истина, а ее отклонение не означает, что она ложна. В любом случае при использовании статистического критерия возможны ошибочные решения.

Ошибку, связанную с отклонением верной нулевой гипотезы из-за попадания статистики в критическую область, называют *ошибкой первого рода*. Вероятность ее, как следует из изложенного, равна α.

Принятие неверной нулевой гипотезы носит название *ошибки второго рода*, соответствующую вероятность обозначают через β. Вероятность $\pi = 1 - \beta$ отклонения основной гипотезы при верной альтернативной носит название функции *мощности критерия*. Чем больше значение π, тем более мощным считается критерий. Обычно задают значение α и стремятся получить возможно большее значение π.

В технической диагностике и дефектоскопии ошибка первого рода, которую также называют *ложной тревогой,* табл. Л6.1, приводит к напрасному прекращению эксплуатации объекта, его браковке, в связи с чем эту ошибку называют также *перебраковкой*. Ошибка второго рода, называемая также *пропуском сигнала*, *недобраковкой*, влечет за собой разрешение дальнейшей эксплуатации неисправного объекта или использования негодного изделия.

Таблица Л6.1. **Характеристика решений при выборе гипотез**

Истинное событие	**Принятое решение**	**Характеристика решения**
H_0 – объект исправен	H_0 – объект исправен	Правильное
H_0 – объект исправен	H_1 – объект не исправен	Ошибка 1-го рода — ложная тревога
H_1 – объект не исправен	H_0 – объект исправен	Ошибка 2-го рода — пропуск сигнала опасности
H_1 – объект не исправен	H_1 – объект не исправен	Правильное

Тяжесть, а, следовательно, и стоимость ошибок второго рода, как правило, неизмеримо более высоки по сравнению с тяжестью ошибок первого рода. Если, например, стоимость напрасно забракованного сварного шва принять равной некоторой условной единице, то согласно имеющимся оценкам стоимость ошибки второго рода, то есть пропуска дефектного сварного шва, в топливно-энергетическом комплексе обходится в пятьдесят тысяч раз дороже. Такова средняя стоимость последствий, которые могут наступить в результате разрыва дефектного шва при эксплуатации. А если такой разрыв приведет к человеческим жертвам, то с точки зрения принятой в нашей стране до последнего времени морали стоимость подобных ошибок вообще не поддается оценке. По-видимому, соответствующие расчеты должны проводиться на основе установленных страховых и штрафных сумм и санкций.

При прогнозе динамики аварийности наряду со стандартными методами математической статистики следует применять конкретный эвристический анализ ситуации. Целесообразно организовать и регулярно проводить более полный анализ ретроспективной информации. Результаты оценки степени риска суммируются по каждому типу причины аварии — коррозия, воздействие сторонней организации, качество сварки, качество проекта, материал и пр. По каждому возможному типу аварии устанавливается вероятность ее возникновения. Например, риск по каждому участку трубопровода определяется путем умножения показателя риска на частоту отдельных аварий и на стоимость последствий. Результаты такого анализа позволяют проводить ранжирование риска для всех элементов системы.