ЛАБОРАТОРНАЯ РАБОТА 3.
КРИТЕРИИ ЗНАЧИМОСТИ

Теоретические сведения

Основные задачи математической статистики:

- ***оценивание параметров***, т.е. получить по выборке оценки (приближенные значения) параметров, наилучшие в том или ином смысле,
- ***проверка статистических гипотез***, т.е. по выборке принять или отвергнуть некоторое предположение о генеральной совокупности, из которой извлечена выборка.

Статистическая гипотеза – любое предположение о виде (***непараметрическая гипотеза***) или параметрах (***параметрическая гипотеза***) неизвестного распределения.

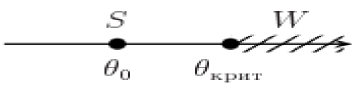
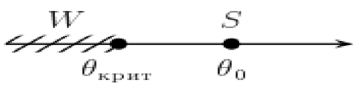
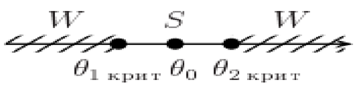
Общий принцип проверки гипотезы

1. Формулируем H_0 – основную (***нулевую***) гипотезу и альтернативную *(конкурирующую)* $\overline{H}$ гипотезу – логическое отрицание H_0.
2. Задаем уровень значимости α – вероятность отвергнуть нулевую гипотезу, при условии, что она верна.
3. Выбираем статистику – правило, по которому принимается решение принять или отклонить нулевую гипотезу – ***критерием проверки статистической гипотезы***.

Статистические критерии, с помощью которых проверяются гипотезы о значениях параметров распределения или о соотношениях между ними в предположении, что тип распределения известен, называются ***критериями значимости,*** или ***параметрическими критериями***.

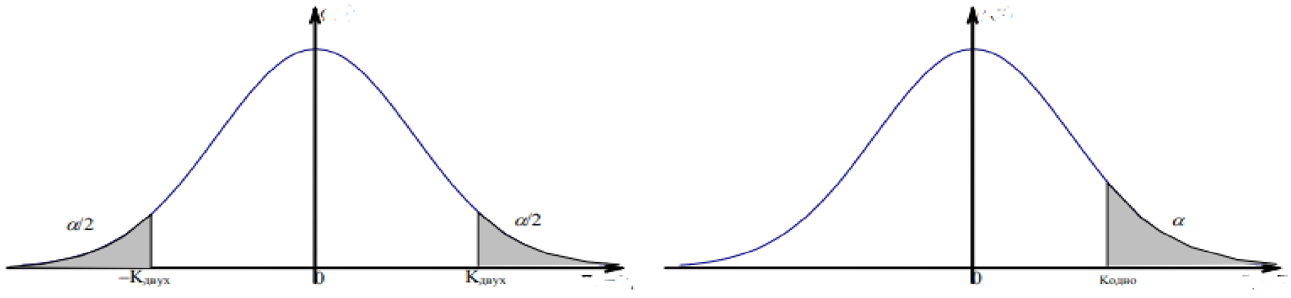
4. Определяем критическую область (табличное значение статистики).

Вид критической области W и области S принятия гипотезы зависит от вида альтернативной гипотезы.

$H_0:\ \theta=\theta_0$ $\overline{H}:\ \theta>\theta_0$ $W=\{\widehat{\theta}_n>\theta_{\text{крит}}\}$	$H_0:\ \theta=\theta_0$ $\overline{H}:\ \theta<\theta_0$ $W=\{\widehat{\theta}_n<\theta_{\text{крит}}\}$	$H_0:\ \theta=\theta_0$ $\overline{H}:\ \theta\neq\theta_0$ $W=\left\{\widehat{\theta}_n<\theta_{1\,\text{крит}}\right.$ $\left.\text{или }\widehat{\theta}_n>\theta_{2\,\text{крит}}\right\}$
		
$P\left(\widehat{\theta}_n>\theta_{\text{крит}}\mid H_0\right)=\alpha$ правосторонняя критическая область	$P\left(\widehat{\theta}_n<\theta_{\text{крит}}\mid H_0\right)=\alpha$ левосторонняя критическая область	$P\left(\widehat{\theta}_n<\theta_{1\,\text{крит}}\mid H_0\right)=$ $=P\left(\widehat{\theta}_n>\theta_{2\,\text{крит}}\mid H_0\right)=\frac{\alpha}{2}$ двусторонняя критическая область

Таким образом, в зависимости от вида альтернативной гипотезы $\overline{H}$ выбирают *правостороннюю*, *левостороннюю* или *двустороннюю* критическую область.

Другими словами, уровень значимости при двустороннем критерии для односторонней альтернативы УДВАИВАЕМ.



Уровни значимости при двустороннем (а) и одностороннем (б) критериях

5. Вычисляем по выборке расчетное значение статистики.
6. Сравниваем расчетное и табличное (критическое) значения статистики, делаем вывод о принятии нулевой или альтернативной гипотез.

Критерии значимости для проверки гипотез о значениях параметров *в случае выборок из нормального распределения.*

Нормальное распределение имеет **два параметра**: *математическое ожидание* a и *дисперсию* σ^2, которые оцениваются с помощью выборочного среднего $\bar{x}$ и выборочной дисперсии (исправленной выборочной дисперсии s^2) соответственно.

Выборочное среднее *является оценкой для* ***среднего значения измеряемой величины*** *и может служить оценкой того или иного показателя качества.*

Дисперсия *характеризует* ***разброс экспериментальных значений****, а следовательно, служит* ***мерой точности****. Например, если произведено несколько измерений одной и той же величины, то дисперсия может характеризовать точность прибора, метода измерения и т. д.*

Требуется по выборке объема n проверить гипотезу H_0 при заданном уровне значимости α.

1. Проверка гипотезы о равенстве математического ожидания нормального распределения заданному значению.

Нулевая гипотеза $H_0:\ a = a_0.$

Альтернативная гипотеза $\bar{H}:\ a \neq a_0.$

1 случай. Если дисперсия σ^2 ***известна****, то гипотеза* H_0 *принимается* (т. е. согласуется с результатами наблюдений) *при условии, что*

$$u_{\text{расч}} = \frac{|\bar{x} - a_0|}{\sqrt{\sigma^2 / n}} < u_{\text{табл}} = u_{\alpha}, \quad \text{где} \quad \Phi(u_{\alpha}) = \frac{1-\alpha}{2}. \qquad (1)$$

2 случай. Если дисперсия σ^2 ***неизвестна****, то гипотеза* H_0 *принимается при*

$$t_{\text{расч}} = \frac{|\bar{x} - a_0|}{\sqrt{s^2 / n}} < t_{\text{табл}} = t_{\alpha;\, n-1}, \qquad (2)$$

где квантиль $t_{\alpha;\, n-1}$ определяется по таблице распределения Стьюдента.

2. Проверка гипотезы о равенстве заданному значению дисперсии нормального распределения.

Нулевая гипотеза $H_0: \sigma^2 = \sigma_0^2.$

Альтернативная гипотеза $\bar{H}: \sigma^2 \neq \sigma_0^2.$

Гипотеза H_0 при заданном уровне значимости α *принимается, если*

$$\chi^2_{1-\alpha/2;\,n-1} < \chi^2_{\text{расч}} = \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\alpha/2;\,n-1}, \qquad (3)$$

где квантили $\chi^2_{1-\alpha/2;\,n-1}$ и $\chi^2_{\alpha/2;\,n-1}$ определяются по таблице распределения χ^2.

В случае альтернативы $\bar{H}_1 : \sigma^2 > \sigma_0^2$ *гипотеза* $H_0 : \sigma^2 = \sigma_0^2$ *при заданном уровне значимости* α *принимается, если*

$$\chi^2_{\text{расч}} = \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\alpha;\,n-1};$$

в случае альтернативы $\bar{H}_2 : \sigma^2 < \sigma_0^2$ гипотеза H_0 принимается, если

$$\chi^2_{\text{расч}} = \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{1-\alpha;\,n-1}.$$

3. Сравнение двух дисперсий нормально распределенных признаков.

Нулевая гипотеза $H_0 : \sigma_1^2 = \sigma_2^2$.

Альтернативная гипотеза $\bar{H} : \sigma_1^2 \neq \sigma_2^2$.

Пусть для первой дисперсии по выборке объема n_1 найдена несмещенная оценка s_1^2, для второй – по выборке объема n_2 оценка s_2^2.

Гипотеза H_0 *при заданном уровне значимости* α *принимается, если*

$$F_{\text{расч}} = \frac{s^2_{\max}}{s^2_{\min}} < F_{\text{табл}} = F_{\alpha/2;\,f_1;\,f_2}. \qquad (4)$$

Здесь $F_{\text{расч}}$ равно отношению *большей* несмещенной оценки дисперсии к *меньшей*, квантиль $F_{\alpha/2;\,f_1;\,f_2}$ определяется по таблице распределения Фишера, причем f_1 и f_2 – числа степеней свободы *соответственно* числителя и знаменателя, т. е. *большей* и *меньшей* оценок дисперсий.

Если гипотеза о равенстве дисперсий принимается, то эти дисперсии считаются ***однородными***. (Термин «однородные» в статистике означает «являющиеся оценкой одного и того же параметра».)

Замечание. Критерий Фишера (4) может применяться также для проверки гипотезы о равенстве нескольких дисперсий нормально распределенных признаков. В этом случае проверяют гипотезу о равенстве наибольшей и наименьшей из сравниваемых дисперсий. Если они признаются однородными, то можно принять гипотезу о равенстве всех сравниваемых дисперсий.

4. Сравнение двух средних в случае независимых нормально распределенных признаков.

Нулевая гипотеза $H_0 :\ a_1 = a_2$.

Альтернативная гипотеза $\bar{H} : a_1 \neq a_2$.

Требуется по выборкам объемов n_1 и n_2 проверить гипотезу H_0 при заданном уровне значимости α.

1 случай. **Если дисперсии σ_1^2 и σ_2^2 *известны***, то *гипотеза H_0 принимается при условии, что*

$$u_{\text{расч}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < u_{\text{табл}} = u_\alpha, \tag{5}$$

где квантиль u_α удовлетворяет соотношению $\Phi(u_\alpha) = \dfrac{1-\alpha}{2}$.

2 случай. **Если дисперсии σ_1^2 и σ_2^2 неизвестны и на основании проверки соответствующей гипотезы по критерию Фишера признаны *однородными***, то *гипотеза H_0 принимается при*

$$t_{\text{расч}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} < t_{\text{табл}} = t_{\alpha;\, f}, \tag{6}$$

где общая средневзвешенная дисперсия s^2 вычисляется по формуле

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

и имеет число степеней свободы $f = n_1 + n_2 - 2$; значение $t_{\alpha;f}$ определяется по таблице квантилей распределения Стьюдента.

3 случай. **Если дисперсии σ_1^2 и σ_2^2 неизвестны** и на основании проверки по критерию Фишера признаны ***неоднородными***, то проверка также проводится по критерию Стьюдента, однако этот критерий является приближенным. В этом случае *гипотеза H_0 принимается, если*

$$t_{\text{расч}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} < t_{\text{табл}} = t_{\alpha;f}, \qquad f \approx \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1-1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2-1}}. \tag{7}$$

Замечание. При сравнении двух средних в случае неизвестных дисперсий возникает необходимость проверки **двух различных гипотез** по одним и тем же данным. Сперва проверяют гипотезу о равенстве дисперсий, а затем гипотезу о равенстве средних.

5. Сравнение нескольких средних в случае независимых нормально распределенных признаков.

Для сравнения нескольких средних в случае независимых нормально распределенных признаков используется специальная статистическая процедура, которая называется ***дисперсионным анализом***.

Однако можно сделать вывод и **на основании критерия Стьюдента, проверив гипотезу о равенстве наибольшего и наименьшего средних**.

Опишем процедуру ***однофакторного дисперсионного анализа.***

Пусть имеется N независимых выборок объемов $n_1, n_2, \ldots, n_N$ соответственно и задан уровень значимости α. Обозначим через $\bar{x}_i, s_i^2$ несмещенные оценки математического ожидания и дисперсии, полученные по *i*-й выборке, $f_i = n_i - 1$.

Нулевая гипотеза $H_0 : a_1 = a_2 = \ldots = a_N$.

Альтернативная гипотеза $\bar{H}$: не все эти математические ожидания равны между собой.

Условием применимости метода дисперсионного анализа является, помимо нормальности выборок, однородность дисперсий. Следовательно, как и в случае двух выборок, процедуре сравнения средних должно предшествовать сравнение дисперсий.

Идея однофакторного дисперсионного анализа заключается в разбиении общей дисперсии, которая получается при объединении всех наблюдений в одну выборку, на два независимых слагаемых – *факторную (межгрупповую)* дисперсию $s^2_{\text{факт}}$, порождаемую различием между группами (выборками), и *остаточную (внутригрупповую)* дисперсию $s^2_{\text{ост}}$, обусловленную случайными помехами и неучтенными факторами: $s^2_{\text{общ}} = s^2_{\text{факт}} + s^2_{\text{ост}}$. Дисперсионный анализ был первоначально предложен Р. Фишером и определен им как метод «отделения дисперсии, приписываемой одной группе причин, от дисперсии, приписываемой другим группам».

Межгрупповая дисперсия рассчитывается по формуле

$$s^2_{\text{факт}} = \frac{1}{N-1}\sum_{i=1}^{N}(\bar{x}_i - \bar{\bar{x}})^2 n_i,$$

где $\bar{\bar{x}}$ – выборочное среднее, рассчитанное по объединенной выборке; число степеней межгрупповой дисперсии равно $f_{\text{факт}} = N - 1$. Остаточная дисперсия представляет собой взвешенное среднее оценок дисперсий и рассчитывается по формуле

$$s^2_{\text{ост}} = \frac{f_1 s_1^2 + f_2 s_2^2 + \ldots + f_N s_N^2}{f_1 + f_2 + \ldots + f_N}, \qquad f_{\text{ост}} = f_1 + f_2 + \ldots + f_N.$$

Гипотеза H_0 *при заданном уровне значимости* α *принимается (не противоречит экспериментальным данным), если*

$$F_{\text{расч}} = \frac{s^2_{\text{факт}}}{s^2_{\text{ост}}} < F_{\text{табл}} = F_{\alpha;\, f_{\text{факт}};\, f_{\text{ост}}},$$

где $F_{\alpha;\, f_{\text{факт}};\, f_{\text{ост}}}$ определяется по таблице квантилей распределения Фишера.

6. Сравнение двух средних в случае зависимых нормально распределенных признаков.

Имеем две выборки одинакового объема n:

$$x_{11}, \quad x_{12}, \quad \ldots, \quad x_{1n};$$

$$x_{21}, \quad x_{22}, \quad \ldots, \quad x_{2n}.$$

Поскольку значения в каждой паре x_{1i}, x_{2i} связаны (например, измерены на одном и том же объекте), то получим новую выборку с элементами $\Delta x_i = x_{1i} - x_{2i}$.

Задача сводится к проверке гипотезы о равенстве нулю среднего значения новой выборки, т. е. $H_0 : a_{\Delta x} = 0$. Эта проверка проводится по критерию (2).

Такая задача возникает, например, когда проводятся измерения одних и тех же величин на одних и тех же объектах двумя разными методами и требуется определить, одинаковы ли результаты использования двух методов измерения. Либо если проводятся измерения какой-то характеристики для одних и тех же объектов до и после некоторого воздействия и требуется определить, влияет ли это воздействие на значение характеристики.