Критерии достоверности оценок.

Статистическая проверка гипотез.

Для сравнительной оценки генеральных параметров выборок используют нулевую гипотезу H_0.(μ_1-μ_2=0; $\overline{x_1} - \overline{x_2}$ =0)

Для проверки принятой гипотезы, используют функции распределения, которые называются критериями достоверности (статистические критерии).

Статистические критерии:

- параметрические (t – критерий Стьюдента, F – критерий Фишера);
- непараметрические (X – критерий Ван-дер-Вардена, Манна-Уитни, Хи^2-критерий).

Критерий Стьюдента.

Критерий Стьюдента применяется для сравнения двух независимых выборок, взятых из нормально распределяющихся совокупностей.

Пусть $\overline{X}_1$ и $\overline{X}_2$ - средние значения выборок, взятых из генеральных совокупностей со средними μ_1 **и** μ_2. Нулевая гипотеза сводится к предположению, что $\mu_1 = \mu_2$.

Критерием для проверки $\mathbf{H_0}$-гипотезы служит отношение:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_{\bar{x}_1 - \bar{x}_2}} \qquad \text{где} \qquad S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sum (Xi - \overline{X}_1)^2 + \sum (X_i - \overline{X}_2)^2}{n(n-1)}}$$

$$t_\phi = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sum (Xi - \overline{X}_1)^2 + \sum (X_i - \overline{X}_2)^2}{n(n-1)}}}$$

$\mathbf{H_0}$-гипотезу отвергают, если фактически установленная величина $\mathbf{t_\phi}$-критерия превзойдёт или окажется равной стандартному $\mathbf{t_{st}}$- этой величины для принятого уровня значимости α и числа степеней свободы $\mathbf{k=n_1+n_2-2}$, т. е. при условии: $\mathbf{t_\phi \geq t_{st}}$, Если $\mathbf{t_\phi < t_{st}}$, то $\mathbf{H_0}$-гипотеза сохраняется.

В случае не равночисленных выборок, т. е. при $\mathbf{n_1 \neq n_2}$

$$t_\phi = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sum (Xi - \overline{X}_1)^2 + \sum (X_i - \overline{X}_2)^2}{n_1 + n_2 - 2}\left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}}$$

F-критерий Фишера. Проверка гипотез для дисперсий.

Для проверки $\mathbf{H_0}$ –гипотезы о равенстве генеральных дисперсий ($\mathbf{S_1= S_2}$) нормально распределяющихся совокупностей **t**-критерий оказывается недостаточно точным, особенно при оценке разности дисперсий малочисленных выборок. Д. Снедокер предложил использовать отношение выборочных дисперсий, обозначив этот показатель в честь Фишера буквой **F** т.е. $F = \frac{\sigma_1^2}{\sigma_2^2}$ при $\sigma_1^2 \geq \sigma_2^2$.

Так как принято брать отношение большей дисперсии к меньшей, то критерий **F≥1**.

Если $\sigma_1^2 = \sigma_2^2$,то **F=1**. Чем значительнее неравенство между выборочными дисперсиями, тем больше будет и величина **F**, и, наоборот, чем меньше окажется разница между дисперсиями, тем меньше будет величина **F**.

Функция **F-** распределения табулирована для **5%**-ного и **1%**-ного уровней значимости и чисел степеней свободы $k_1=n_1-1$ для большей дисперсии и $k_2=n_2-1$ для меньшей. Критические точки для **F**-критерия содержатся в таблице "Значения **F**-критерия Фишера и при уровнях значимости $\alpha=5\%$ (верхняя строка) и $\alpha=1\%$ (нижняя строка)". В этой таблице степени свободы для большей дисперсии k_1 расположены в верхней строке (по горизонтали), а степени свободы для меньшей дисперсии k_2- в первой графе (по вертикали).

Если сравниваемые выборки извлечены из одной и той же генеральной совокупности или из разных совокупностей с дисперсиями σ_1^2 и σ_2^2 равными друг другу: $\sigma_1^2 = \sigma_2^2$, то величина **F**-критерия не превысит критические точки F_{st} , указанные в таблице. Если же сравниваемые выборки взяты из разных совокупностей с их параметрами σ_1^2 и σ_2^2 не равными друг другу, то $F_ф \geq F_{st}$ и нулевая гипотеза должна быть отвергнута.

Критерий U Манна-Уитни

Назначение критерия

Критерий предназначен для оценки различий между *двумя* независимыми выборками по *уровню* какого-либо признака, количественно измеренного. Он позволяет выявлять различия между *малыми* выборками, когда $n_1,n_2 \geq 3$ или $n_1=2$, $n_2 \geq 5$.

Описание критерия

Этот метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами. Чем меньше область перекрещивающихся значений, тем более вероятно, что различия достоверны. Иногда эти различия называют различиями в *расположении* двух выборок.

Эмпирическое значение критерия U отражает то, насколько велика зона совпадения между рядами. Поэтому *чем меньше* $U_{эпр}$, *тем более* вероятно, что различия *достоверны.*

Гипотезы

H_0: Уровень признака в группе 2 не ниже уровня признака в группе 1.

Графическое представление критерия U

На Рис.1 представлены три из множества возможных вариантов соотношения двух рядов значений.

В варианте (а) второй ряд ниже первого, и ряды почти не перекрещиваются. Область наложения слишком мала, чтобы скрадывать различия между рядами.

Есть шанс, что различия между ними достоверны. Точно определить это можно с помощью критерия U.

В варианте (б) второй ряд тоже ниже первого, но и область перекрещивающихся значений у двух рядов достаточно обширна. Она может ещё не достигать критической величины, когда различия придётся признать несущественными. Но так ли это, можно определить только путём точного подсчета критерия U.

В варианте (в) второй ряд ниже первого, но область наложения настолько обширна, что различия между рядами скрадываются.

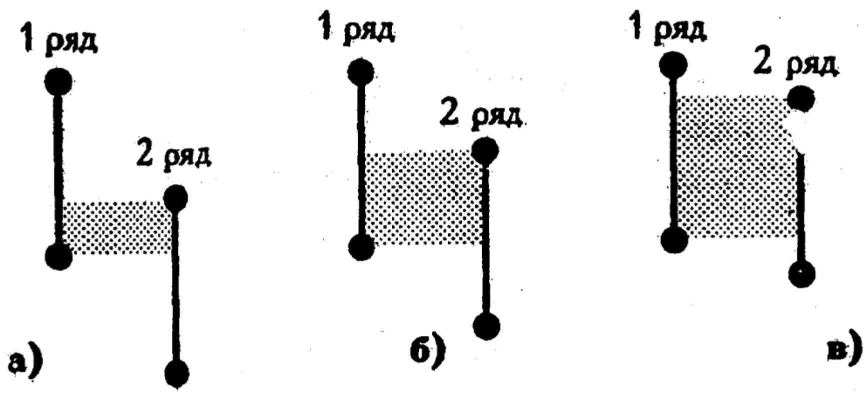

Рис1. Возможные варианты соотношений рядов значений в двух выборках; штриховкой обозначены зоны наложения.

Ограничения критерия U

1. В каждой выборке должно быть не менее 3 наблюдений: $n_1,n_2 \geq 3$; допускается, чтобы в одной выборке было 2 наблюдения, но тогда во второй их должно быть не менее 5.
2. В каждой выборке должно быть не более 60 наблюдений; $n_1,n_2 \leq 60$. Однако уже при $n_1,n_2>20$ ранжирование становится достаточно трудоёмким.

АЛГОРИТМ

подсчета критерия U Манна-Уитни

1. Расположить все данные обоих выборок по степени нарастания признака, не считаясь с тем, к какой выборке они относятся, как если бы мы работали с одной большой выборкой.
2. Проранжировать значения, приписывая меньшему значению меньший ранг. Всего рангов получится столько, сколько у нас (n_1+n_2).
3. Подсчитать сумму рангов отдельно по первой и второй выборке отдельно. Проверить, совпадает ли общая сумма рангов с расчетной.
4. Расчетная сумма: $\sum R_1 = \frac{N \cdot (N+1)}{2}$
5. Определить большую из двух ранговых сумм.
6. Определить значение U по формуле:

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x$$

n_1-количество испытуемых в выборке 1

n_2- количество испытуемых в выборке 2

T_x-большая из двух ранговых сумм

n_x-количество испытуемых в группе с большей суммой рангов.

7. Определить критические значения U по табл. 8.
8. Если $U_{эмп}>U_{кр}$ 0.05, H_0 принимается.
9. Если $U_{эмп}\leq U_{кр\ 0.05}$, H_0 отвергается. Чем меньше значения U, тем достоверность различий выше.

Критерий Хи-квадрат

Критерий Хи-квадрат применяется для проверки правильности выбранного закона распределения.

Схема вычисления критерия **Хи-квадрат:**

1. Определяют опытные (эмпирические) частоты встречи данной случайной величины.

f_1 f_2 f_3 … f_n

2. Выбирают закон распределения случайной величины в качестве предполагаемого и в соответствии с выбранным законом рассчитывают теоретические частоты.

f_1' $F\ _2'$ f_3' … f_n'

3. Находят разность между эмпирическими и теоретическими частотами.

$f_1- f_1'$ $f_2- f_2'$ $f_3- f_3'$ … $f_n- f_n'$

4. Возводят в квадрат разность между эмпирическими и теоретическими частотами.

$(f_1- f_1')^2$ $(f_2- f_2')^2$ $(f_3- f_3')^2$ … $(f_n- f_n')^2$

5. Находят отношение квадрата разности между эмпирическими и теоретическими частотами к теоретическим частотам.

$$\frac{(f_1 - f_1')^2}{f_1'} \quad \frac{(f_2 - f_2')^2}{f_2'} \quad \frac{(f_3 - f_3')^2}{f_3'} \quad \cdots$$

6. Находят сумму отношений квадратов разности между эмпирическими и теоретическими частотами к теоретическим частотам.

$$\frac{(f_1 - f_1')^2}{f_1'} + \frac{(f_2 - f_2')^2}{f_2'} + \frac{(f_3 - f_3')^2}{f_3'} + \cdots \quad \frac{(f_n - f_n')^2}{f_n'}$$

7. Получают формулу критерия **Хи-квадрат:** $X^2 = \sum_{i=1}^{k} \frac{(f - f')^2}{f'}$

Число степеней свободы **R** при оценке эмпирических распределений, следующих нормальному закону, **R=N-3** . Если же оценке подлежит распределение, следующее закону Пуассона, число степеней свободы уменьшается на единицу, т. е. **R=N-2.** В других случаях число степеней свободы устанавливается особо.

Нулевая гипотеза сводится к предположению, что различия, наблюдаемые между эмпирическими и вычисленными частотами, носят исключительно случайный характер. Для проверки нулевой гипотезы нужно фактически полученную

величину $\mathbf{X_{ф}^2}$ сравнить с её критическим значением $\mathbf{X_{st}^2}$, определяемом по таблице “$\mathbf{X^2}$- распределение. Критические точки для разных значений вероятности Р и чисел степеней свободы-R)”.
Если $\mathbf{X_{ф}^2 \geq X_{st}^2}$ **,** то нулевая гипотеза должна быть отвергнута на принятом уровне значимости с числом степеней свободы R**.**