Лекция 8

Проверка статистических гипотез

Статистическая гипотеза. Критерии, используемые для проверки однородности гидрологических рядов. Критерий Стьюдента. Критерий равенства двух дисперсий. Рангово – суммарный критерий.

(Ахметов С.К.)

Определения

Статистическая гипотеза – это предположение о каком-то свойстве генеральной совокупности

Например, что m_x генеральной совокупности равно какому–то числу

Нулевая гипотеза – это когда предполагается, что среднее значение выборки $x_1, x_2, x_3 \dots x_n$ (если n большое) мало отличаться от m_x генеральной совокупности

Альтернативная гипотеза H_1 – это когда предполагается, что $m_x \neq x_{cp}$, $m_x > x_{cp}$, $m_x < x_{cp}$

Критерий (тест) статистической гипотезы - это правило, позволяющее принять или отвергнуть гипотезу

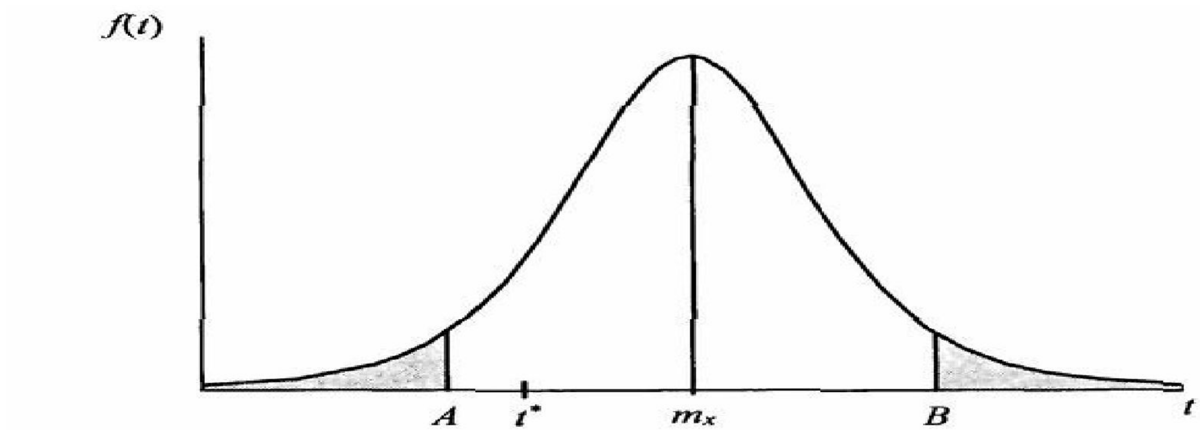
Статистики – это определенные функции $g\ (x_1, x_2, x_3 \dots x_n)$, используемые для выполнения теста

Пример статистики

- Рассмотрим выборку с параметрами $x_{ср}$, S_x *(СКО)*
- Допустим, что $m_x = 12$.
- Нулевая гипотеза (H_0). Разница $(x_{ср} - m_x)$ достаточно мала
- Эту разницу можно рассматривать в качестве анализируемой статистики
- Но на практике используют другую статистику $t = (x_{ср} - m_x)/(S_x/\sqrt{n})$, так как соответствие с **теоремой *2*** прошлой лекции заранее известно, что это выражение подчиняется распределению Стьюдента
- На использовании этой ***статистики*** базируется критерий Стьюдента, который можно использовать для проверки нулевой гипотезы
- С этой целью для конкретной реализации рассчитывают эмпирическое значение статистики Стьюдента t^*. Например, $n=36$, $x_{ср} =11$, $S_x = 5$, $t^* = 1.2$.
- Величина t^* является СВ и для выборок различной длины значение t^* будет различным
- Область возможных значений *(ОВЗ)* этой статистики - вся числовая ось
- ***ОВЗ*** делиться на две области:
  - область принятия гипотезы
  - критическая область
- Если t^* попадает в область принятия гипотезы, то H_0 не опровергается, если в критическую область, то H_0 опровергается.

Доверительная область

- ***Доверительная область (доверительный интервал)*** **-** область принятия гипотезы
- ***Доверительная вероятность ($p_д$)*** – вероятность получения значения ***t**** по произвольной выборке.
- Геометрически эта вероятность показана (заштрихована) на рисунке



Область принятия гипотезы **[A,B]**
и критическая область: **(-∞, A) + (B, + ∞)**

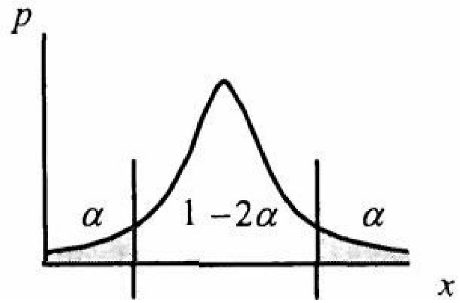
Уровень значимости

- ***Вероятность попадания t* в критическую область*** равна $\alpha = 1 - p_{\partial}$. Здесь **α – *уровень значимости***
- Если критическая область состоит из двух частей, то вместо α пишут $2\alpha = 1 - p_д$, где 2α указывает на то, что уровень значимости двухсторонний
- В гидрологической практике наиболее часто используют уровни значимости 5 и 10%
- К примеру, если принять уровень значимости $2\alpha = 10\%$ то, используя таблицу распределения Стьюдента при $\nu = n - 1 = 35$, получим ***теоретическое*** значение ***t*** – статистики равное 1,69. Это означает, что доверительная область представлена отрезком **[-1,69, + 1,69]**
- Если $t^* = -1{,}20$ ***(по фактическим данным***), значить ***t**** попадает в доверительную область, первоначальная нулевая гипотеза не опровергается, то есть можно считать, что разница между m_x и $x_{ср}$ является статистически не значимой
- Хотя в данном случае речь идет о конкретном критерии (Стьюдента), основные принципы проверки нулевой гипотезы сохраняются и при использовании других критериев.

Уровень значимости (α) и доверительная вероятность для различных альтернативных гипотез

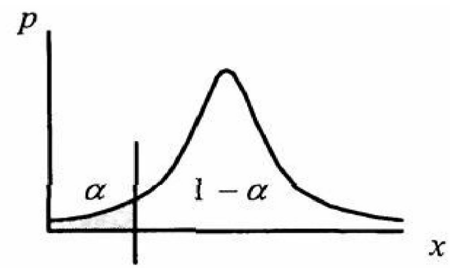
$H_0 : m_x = \bar{x}$,

$H_1 : m_x \neq \bar{x}$.



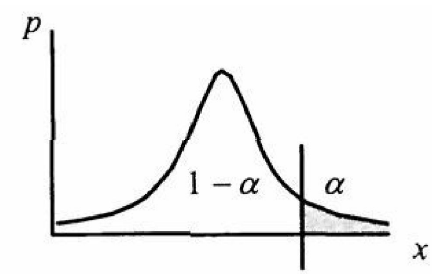
$H_0 : m_x = \bar{x}$,

$H_1 : m_x < \bar{x}$.



$H_0 : m_x = \bar{x}$,

$H_1 : m_x > \bar{x}$.



Критерии, используемые для проверки однородности гидрологических рядов

- Используются критерии двух типов: ***параметрические*** и ***непараметрические***

- ***В параметрических критериях*** используются выборочные оценки параметров распределения (критерии Стьюдента и Фишера). При этом считается, что выборка относится к генеральной совокупности с известным законом распределения (обычно с нормальным законом)

- ***Непараметрические критерии*** основываются на использовании непараметрических статистик. Статистика $g(x_1, x_2, x_3 \ldots x_n)$, является непараметрической, если ее распределение не зависит от распределения X

Критерии Стьюдента для проверки значимости различных средних значений двух выборок

- Допустим, что $(x_1, x_2, x_2 \dots x_n)$ и $(y_1, y_2, y \dots y_n)$ – выборки длиной m и n с ***неизвестными*** параметрами m_x, σ_x и m_y, σ_y, но при этом известно, что $\sigma_x = \sigma_y$, хотя СКО значение неизвестно. Поэтому можно записать, что $\sigma = \sigma_x = \sigma_y$
- Если выборки относятся к одной генеральной совокупности, то разность должна быть близка к нулю. На основе этой разности строиться статистика

$$t = (x_{cp} - y_{cp})/(\sigma_{xcp - ycp})$$

где $\sigma_{xcp - ycp}$ – СКО разности $(x_{cp} - y_{cp})$. Как следует из теоремы 2, из предыдущей лекции эта статистика подчиняется распределению Стьюдента при $v = (m + n - 2)$.

В математической статистике доказано, что

$$\sigma_{\bar{x}-\bar{y}} = S\sqrt{(m+n)/mn}$$

где S – эмпирическая оценка $\sigma_{xcp - ycp}$. Значение S определяется в зависимости от выборочных значений S_x и S_y.

$$S = \sqrt{\frac{(m-1)S_x^2 + (n-1)S_y^2}{m+n-2}}$$

Критерии Стьюдента для проверки значимости различных средних значений двух выборок (2)

В окончательном виде выражение для статистики ***t*** имеет вид

$$t = \frac{\bar{x} - \bar{y}}{S}\sqrt{\frac{mn}{m+n}}$$

- В практике гидрологических расчетов эта статистика используется для проверки однородности рядов по среднему значению
- Исходный ряд делится на две части (две выборки) Если дата возможного нарушения стока неизвестна, то ряд делится пополам.
- Уровень значимости обычно принимается ***2α = 5%*** или ***2α = 10%***
- Критерий Стьюдента рекомендуется в большинстве нормативных документов в качестве одного из официальных тестов на однородность.
- При использовании критерия Стьюдента нужно учитывать три момента:

- предполагается, что анализируемые выборки относятся к нормальным совокупностям
- предполагается, что анализируемые выборки имеют одинаковую (хотя и неизвестную) дисперсию, поэтому перед использованием критерия Стьюдента следует проверить ряд на однородность по дисперсии
- В классической статистике длина выборок предполагается значительной большей, чем та, которую мы имеем на практике

Критерий равенства двух дисперсий

- Если $(x_1, x_2, x_2 ....x_n)$ и $(y_1, y_2, y ....y_n)$ – выборки из нормальной совокупности с параметрами m_x, σ_x, и m_y, σ_y и если $\sigma_x = \sigma_y = \sigma$, то в соответствие с теоремами 1 и 3 из предыдущей лекции, отношение их выборочных дисперсий S_x^2/S_y^2 подчиняется распределению Фишера с числом степеней свободы $v_1 = m\text{-}1$ и $v_2 = n\text{-}1$
- Следовательно, при нулевой гипотезе H_0 (т.е. при предположении о том, что $S_x^2 = S_y^2$) и при уровне значимости 2α доверительная значимость для отношения S_x^2/S_y^2 определяется выражением

$$F_\alpha (v_1, v_2) \le (S_x^2/S_y^2) < F_{1-\alpha} (v_1, v_2) \quad \text{или} \quad 1/F_{1-\alpha} \le (S_x^2/S_y^2) < F_{1-\alpha}$$

- Распределение Фишера является несимметричным и для того, чтобы сократить объем таблиц, их составляют только для $F > 1$, а при сравнении S_x^2 и S_y^2 в числитель всегда подставляют большую дисперсию. В этом случае доверительная область при уровне значимости 2α определяется выражением

$$1 \le (S_x^2/S_y^2) < F_{1-\alpha}$$

- Этот критерий используется для проверки однородности гидрологических рядов по дисперсии.

Расчет по критерию Фишера

- Исходный ряд делится на две части
- Оцениваются дисперсии для каждой из частей ряда и вычисляются эмпирическое значение статистики Фишера $F^* \leq S_x^{\,2}/S_y^{\,2}$, где $S_x^{\,2} > S_y^{\,2}$.
- Полученное значение F^* сравниваются с табличным значением $F_{1-\alpha}$.
- Если при принятом уровне значимости оказывается, что $F^* < F_{1-\alpha}$, то расхождение дисперсий считается незначимым и гипотеза об однородности ряда по дисперсии не опровергается
- Критерий Фишера (также как и критерий Стьюдента) относится к категории стандартных критериев и рекомендуется в большинстве нормативных документов в качестве официального теста на однородность.

Рангово – суммарные
критерии Вилкоксона, Вилкоксона - Манна - Уитни

- Критерии используются для проверки нулевой гипотезы о том, что две независимые выборки принадлежат к совокупностям, которые имеют идентичные функции распределения

- Критерии Вилкоксона и Манна – Уитни относятся к категории непараметрических критериев и не подразумевают непосредственного расчета выборочных параметров функции распределения

- Достоинством критериев является то, что они не требуют обязательной принадлежности выборок к нормальной совокупности

Критерий Вилкоксона

- ***Статистика Вилкоксона.*** Пусть даны две выборки из совокупностей X и Y длиной m и n (m и n). Объединим эти выборки в один ряд и расположим все значения в возрастающем порядке так, что $y_1 < y_2$, $y_2 < x_1$, $x_1 < x_2$ и т.д.

Определение рангов (по Уилкоксону)

Ранжированный объединенный ряд	y_1	y_2	x_1	x_2	x_3	y_3	x_4	y_4	x_5	y_5	y_6
Ранг	1	2	3	4	5	6	7	8	9	10	11

Каждому значению нового ряда присвоим порядковый номер (ранг). Выпишем значения подученных рангов отдельно для каждой из выборок (таблица ниже) и вычислим ранговые суммы ω_1 и ω_2

Вычисление ранговых сумм (по Уилкоксону)

Выборка 1 ...	x_1	x_2	x_3	x_4	x_5		Сумма рангов ω_1
Ранг	3	4	5	7	9		28
Выборка 2	y_1	y_2	y_3	y_4	y_5	y_6	Сумма рангов ω_2
Ранг	1	2	6	8	10	11	38

Критерий Вилкоксона

Если расчет выполнен правильно, то должно выполняться равенство

$$\omega_1 + \omega_2 = N(N+1)/2$$

где $N = m + n$

В качестве анализируемой статистики ω^* рассматривается выборка

$$\omega^* = r_1 + r_2 + r_3 + \quad + r_m$$

где r_i - ранг x_i

Для статистики Вилкоксона разработаны таблицы, позволяющие определить доверительный интервал для ω (в зависимости от m, n и уровня значимости 2α). Найденное значение ω сравнивается с математическим ожиданием статистики **M(W)**, которое вычисляется по формуле $M(W) = (m+n+1)/2$

$w_В(\alpha, n_1, n_2) > w$ и $w_Н(\alpha, n_1, n_2) < w$

Верхнее критическое значение связано с нижним соотношением

$$w_В(\alpha, n_1, n_2) = 2\,M(W) - w_Н(\alpha, n_1, n_2)$$

Однако в настоящее время более широкое распространение получила статистика, предложенная Манном и Уитни, так называемая U – статистика.

Статистика Вилкоксона - Манна - Уитни (U – статистика)

Для определения U вычислим:

$$U_1 = m \bullet n + m(m+1)/2 - \omega_1$$

$$U_2 = m \bullet n + m(m+1)/2 - \omega_2$$

В качестве анализируемой статистики U^* можно использовать любое из полученных значений (U_1 и U_2). Обычно в качестве U^* принимается меньшее значение. При правильном расчете должно выполняться равенство

$$U_1 + U_2 = m \bullet n$$

Распределение U – статистики Манна – Уитни является симметричным с *МО и дисперсией*

$$m_U = (m \bullet n)/2$$

$$D_U = \sigma^2_U = m \bullet n(m+n+1)/12$$

Статистика Вилкоксона - Манна - Уитни (U – статистика)

При $m \geq 8$ и $n \geq 8$ функция распределения нормированной величины статистики U может быть аппроксимирована стандартным нормальным распределением. При этом доверительный интервал для статистики U при уровне значимости 2α имеет вид:

$$m_U - t_{1-\alpha}\,\sigma_U \leq U < m_U + t_{1-\alpha}\,\sigma_U$$

где $t_{1-\alpha}$ - квантиль стандартного нормального распределения;

m_U и σ_U – параметры, определяемые по формулам

$$m_U = (m \bullet n)/2$$

$$D_U = \sigma^2{}_U = m \bullet n(m+n+1)/12$$

Критерий Уилкоксона – Манна – Уитни можно использовать для проверки однородности гидрологических рядов. В этом случае исходный ряд разбивается на две выборки, длиной m и n.

СПАСИБО ЗА ВНИМАНИЕ!