Лекция 6

Тема: «Статистические гипотезы»

План лекции:

1. Критерии проверки статистических гипотез

2. Параметрические критерии: Критерий Стьюдента, Критерий Фишера

3. Непараметрические критерии: Хи-квадрат, критерий Колмогорова-Смирнова, Критерий знаков, Критерий Мана-Уитни, критерий Уилка-Шапиро и др.

4. Применение статистических критериев в анализе почвенных данных

Основные понятия:

- Нулевая гипотеза
- Альтернативная гипотеза
- Ошибки первого и второго рода
- Уровень значимости

Этапы проверки статистических гипотез

- Формулировка основной гипотезы H_0 и конкурирующей гипотезы H_1. Гипотезы должны быть чётко формализованы в математических терминах.
- Задание вероятности α, называемой уровнем значимости и отвечающей ошибкам первого рода, на котором в дальнейшем и будет сделан вывод о правдивости гипотезы.
- Расчёт статистики φ критерия такой, что:
  - её величина зависит от исходной выборки ;
  - по её значению можно делать выводы об истинности гипотезы H_0;
  - сама статистика φ должна подчиняться какому-то известному закону распределения, т.к. сама φ является случайной в силу случайности .
- Построение критической области. Из области значений φ выделяется подмножество таких значений, по которым можно судить о существенных расхождениях с предположением. Его размер выбирается таким образом, чтобы выполнялось равенство . Это множество и называется **критической областью**.
- Вывод об истинности гипотезы. Наблюдаемые значения выборки подставляются в статистику φ и по попаданию (или непопаданию) в критическую область выносится решение об отвержении (или принятии) выдвинутой гипотезы H_0.

Статистическая гипотеза - некоторое предположение о свойствах генеральной совокупности, которой принадлежит выборка.

- Нулевая гипотеза (H_0) - предположение о том, что между генеральными параметрами сравниваемых групп разница равна нулю, или различия между выборочными показателями носят случайный характер

Если выборка из совокупности 1 имеет параметры μ_1 и σ_1, а выборка из совокупности 2 соответственно $\mu_2 \sigma_2$, то:

$$\mu_1=\mu_2,\ \sigma_1=\sigma_2$$

и

$$\mu_1-\mu_2=0,\ \sigma_1-\sigma_2=0$$

Нулевая гипотеза может иметь в виду μ=α, где α- какое-то число.

$$H_0: \mu = a.$$

Альтернативная (противоположная) гипотеза – противопоставляется нулевой гипотезе и исходит из того, что:

$$\mu_1 - \mu_2 \neq 0$$

и

$$\sigma_1 - \sigma_2 \neq 0$$

Критерии проверки гипотез:

Число степеней свободы (k) – числа, показывающие количество свободно варьирующих элементов или членов статистической совокупности, способных принимать любые произвольные значения.

Уровень значимости (α) – значение вероятности, при котором различия, наблюдаемые между выборочными показателями, можно считать несущественными, случайными.

- Критерии значимости
  - Параметрические
    - Критерий Стьюдента
    - (t)
    - Критерий Фишера
    - (F)
  - Непараметрические
    - Критерий
    - Хи-квадрат
    - (χ^2)
    -
    - Критерий
    - Колмогорова-Смирнова (d)
    - Критерий знаков
    - (z)

- Критерий Мана-Уитни (U)

- Критерий
- Уилка-Шапиро (W)

- Т-критерий Уилксона (Т)

- Параметрические критерии
  - строятся на основе параметров выборочной совокупности

- Непараметрические критерии
  - функции от вариант данной совокупности с их частотами

- Область значений случайной величины
  - Область допустимых значений

  - Область маловероятных значений

- Критическое значение – соответствует границе между областью допустимых и областью маловероятных значений.

Устанавливается в зависимости от принятого уровня значимости (α). Критерии проверки гипотез

Выделяют три вида критических областей:

- *Двусторонняя критическая область* определяется двумя интервалами, где находят из условий .
- *Левосторонняя критическая область* определяется интервалом , где x_α находят из условия $P(\varphi < x_\alpha) = \alpha$.
- *Правосторонняя критическая* область определяется интервалом , где x_α находят из условия $P(\varphi > x_\alpha) = \alpha$.

Ошибка первого рода

- Уровень значимости характеризует ту вероятность, которой решено пренебрегать в данном исследовании.

- Отклонение нулевой гипотезы при попадании значения случайной величины в критическую область нельзя рассматривать как доказательство того, что гипотеза неверна, так как значения, выходящие за пределы области принятия гипотезы H_0 могут иметь место и в случае правильности нуль-гипотезы, и вероятность такого события известна - она равна α.

- Отклоняя правильную нулевую гипотезу, мы допускаем так называемую ошибку первого рода, принятый же уровень значимости α характеризует риск допустить такую ошибку.

Ошибка второго рода

Принятие нулевой гипотезы, когда она неверна, носит название ошибки второго рода. Вероятность такой ошибки обозначается (β).

С вероятностью 1 - β принятия нулевой гипотезы, когда она верна, связывается в математической статистике понятие мощность критерия.

Уменьшая вероятность ошибки первого рода (α), мы неизбежно увеличиваем вероятность ошибки второго рода (β).

Выбор уровня значимости α (устанавливается обычно α, а не β) определяется условиями проведения эксперимента, ответственностью выводов и учетом того, ошибка какого рода наиболее нежелательна.

В большинстве случаев принимают α = 0,05 (5%), что соответствует доверительной вероятности P = 0,95.

Параметрические критерии

- Распределение Стьюдента (или t-распределение) - это распределение отклонений нормально распределенной случайной величины от генерального среднего, нормированных выборочной оценкой среднего квадратического отклонения.
- Это распределение зависит от числа степеней свободы γ, с которым найдена оценка среднего квадратического отклонения.

Классическим примером распределения Стьюдента является распределение стандартизованных отклонений

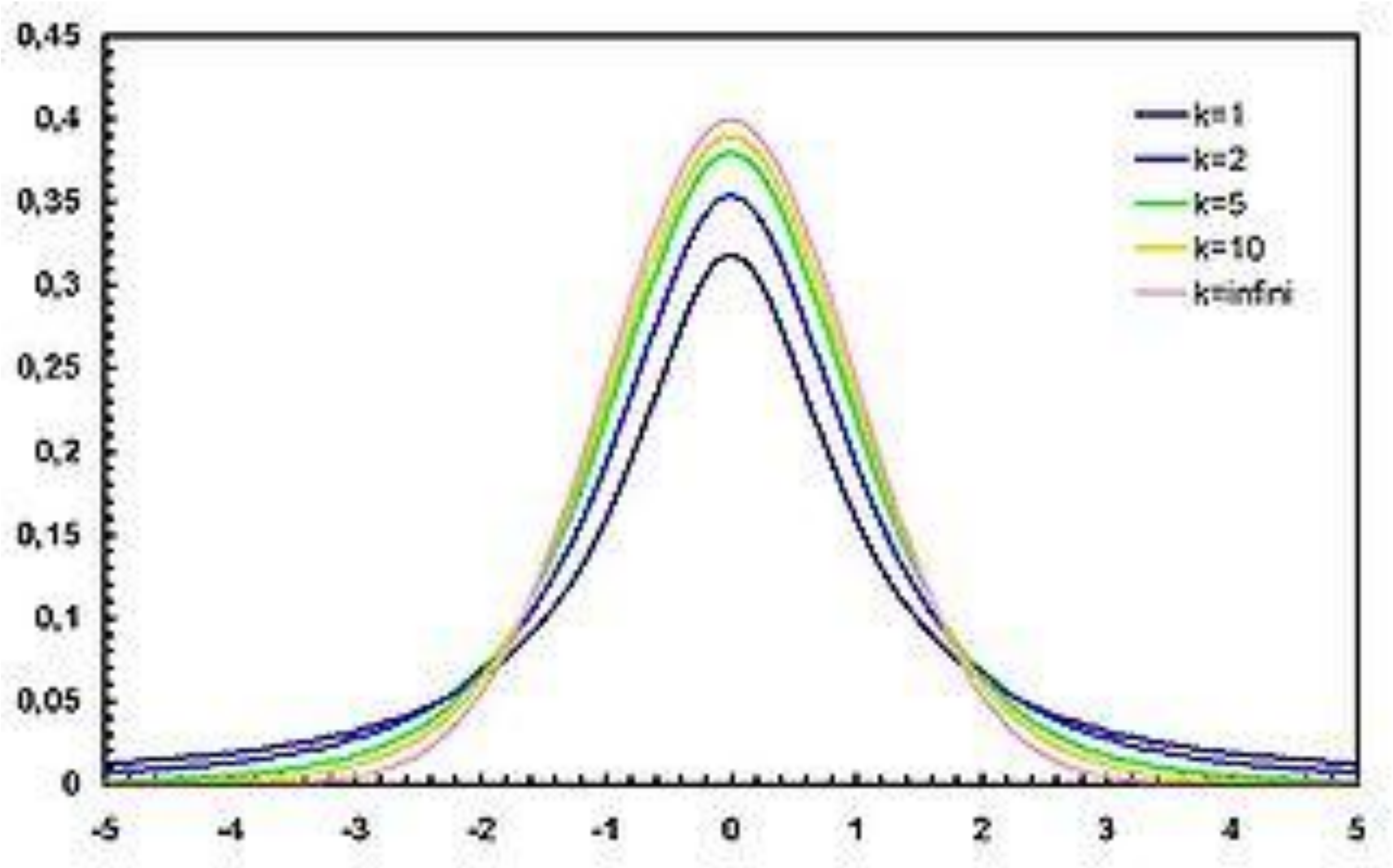
$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}},$$

где: x - нормально распределенное выборочное среднее;
μ- генеральное среднее; S_x - ошибка среднего, вычисленная по выборке объема n,
t - значение случайной величины, распределенной по Стьюденту с ν= n - 1 числом степеней свободы.

Кривая распределения Стьюдента похожа по внешнему виду на кривую нормального распределения: она одновершинна, симметрична, ее ветви асимптотически приближаются к оси абсцисс.

При $\nu \to \infty$ распределение Стьюдента стремится к нормальному распределению с параметрами $\mu = 0$ и $\sigma = 1$.

Кривые нормального распределения
(Z -сплошная линия) и распределения t-Стьюдента при v=3 (пунктирная линия)



- Наибольшее отличие распределения Стьюдента от нормального наблюдается при $\nu=1$, когда при значениях переменной величины t, близких к среднему, плотность вероятности распределения Стьюдента меньше, а при значениях, сильно отличающихся от среднего, больше, чем при нормальном распределении.

Закон распределения имеет непрерывную функцию и описывается формулой:

$$f(t) = C\left(1 + \frac{t^2}{n-1}\right)^{-\frac{n-1}{2}}$$

для $-\infty < t < +\infty$

где: C – константа, зависящая от числа степеней свободы n-1 t – распределение – частный случай нормального распределения;

t – распределение – симметрично;

t – распределение отражает специфику распределения малой выборки по нормальному закону.

Таблица 3. Критические точки t-критерия Стьюдента при различных уровнях значимости α

Число степеней свободы k	α, %			Число степеней свободы k	α, %		
	5	1	0,1		5	1	0,1
1	12,71	63,66	64,60	18	2,10	2,88	3,92
2	4,30	9,92	31,60	19	2,09	2,86	3,88
3	3,18	5,84	12,92	20	2,09	2,85	3,85
4	2,78	4,60	8,61	21	2,08	2,83	3,82
5	2,57	4,03	6,87	22	2,07	2,82	3,79
6	2,45	3,71	5,96	23	2,07	2,81	3,77
7	2,37	3,50	5,41	24	2,06	2,80	3,75
8	2,31	3,36	5,04	25	2,06	2,79	3,73
9	2,26	3,25	4,78	26	2,06	2,78	3,71
10	2,23	3,17	4,59	27	2,05	2,77	3,69
11	2,20	3,11	4,44	28	2,05	2,76	3,67
12	2,18	3,05	4,32	29	2,05	2,76	3,66
13	2,16	3,01	4,22	30	2,04	2,75	3,65
14	2,14	2,98	4,14	40	2,02	2,70	3,55
15	2,13	2,95	4,07	60	2,00	2,66	3,46
16	2,12	2,92	4,02	120	1,98	2,62	3,37
17	2,11	2,90	3,97	∞	1,96	2,58	3,29
P	0,05	0,01	0,001	–	0,05	0,01	0,001

Сравнение средних арифметических корреляционно не связанных между собой выборок, взятых из нормально распределяющихся совокупностей с их параметрами $\mu_1\sigma_1^2$ $\mu_2\sigma_2^2$ исходят из предположения , что разница между ними возникла случайно ($d=X_1-X_2$). В качестве критерия проверки гипотезы служит переменная величина:

t =t	**[(x_1-x_2)]**
	S_{X1-X2}
Где:	S_{X1-X2} =S_d– ошибка разности между выборочными средними s

Нулевая гипотеза опровергается ($\cancel{H_0}$), если $t_ф \geq t_{st}$ для принятого уровня значимости и числа степеней свободы $k=n_1+n_2-2$.

Распределение F Фишера.

Распределение представляющее собой случайную величину, распределение которой было изучено Фишером, названо его именем и обозначено буквой F.

$$F = \frac{s_1^2}{s_2^2},$$

- Если имеются две оценки S_1^2 и S_2^2 одной и той же дисперсии σ^2 нормально распределенной случайной величины, то, принимая, что $S_1^2 > S_2^2$, можно найти отношение этих оценок. При этом всегда берется отношение большей дисперсии к меньшей:

С увеличением v_1 и v_2 обе оценки стремятся к одному и тому же параметру σ^2, F при этом стремится к единице.

Чем меньше v_1 и v_2, тем больше шансов получить в случайном порядке достаточно отличные от единицы значения F.

Распределение F зависит от числа степеней свободы ν_1 и ν_2, с которыми найдены оценки дисперсий в числителе (ν_1) и в знаменателе (ν_2).

Таблица 4. Значения F-критерия Фишера при уровнях значимости α = 5% (верхняя строка) и α = 1% (нижняя строка)

κ_2	κ_1 — степени свободы для большей дисперсии							
	1	2	3	4	5	6	7	8
1	161 4052	200 4999	216 5403	225 5625	230 5764	234 5859	237 5928	239 5982
2	18,51 98,50	19,00 99,00	19,16 99,17	19,25 99,25	19,30 99,30	19,33 99,33	19,35 99,36	19,37 99,37
3	10,13 34,12	9,55 30,82	9,28 29,16	9,12 28,71	9,01 28,42	8,94 27,91	8,89 27,67	8,85 27,49
4	7,71 21,20	6,94 18,00	6,59 16,89	6,39 15,98	6,26 15,52	6,16 15,21	6,04 14,98	6,00 14,80
5	6,61 16,26	5,79 13,27	5,41 12,06	5,19 11,39	5,05 10,97	4,95 10,67	4,88 10,46	4,82 10,29
6	5,99 13,75	5,14 10,92	4,76 9,78	4,53 9,15	4,39 8,75	4,28 8,47	4,21 8,26	4,15 8,10
7	5,59 12,25	4,74 9,55	4,35 8,47	4,12 7,85	3,97 7,46	3,87 7,19	3,79 6,99	3,73 6,84
8	5,32 11,26	4,46 8,65	4,07 7,59	3,84 7,01	3,69 6,63	3,58 6,37	3,50 6,18	3,44 6,03
9	5,12 10,56	4,26 8,02	3,86 6,99	3,63 6,42	3,48 6,06	3,37 5,80	3,29 5,61	3,23 5,47
10	4,96 10,04	4,10 7,56	3,71 6,55	3,48 5,99	3,33 5,64	3,22 5,39	3,14 5,20	3,07 5,06
11	4,84 9,65	3,98 7,21	3,59 6,22	3,36 5,67	3,20 5,32	3,09 5,07	3,01 4,89	2,95 4,74
12	4,75 9,33	3,80 6,93	3,49 5,95	3,26 5,41	3,11 5,06	3,00 4,82	2,91 4,64	2,85 4,50
13	4,67 9,07	3,80 6,70	3,41 5,74	3,18 5,21	3,02 4,86	2,92 4,62	2,83 4,44	2,77 4,30

- Если выборки взяты из разных совокупностей с неравными параметрами σ_1^2 и σ_2^2, то $F_{\text{ф}} \geq F_{st}$ и нулевая гипотеза должна быть опровергнута ($\cancel{H_0}$).

Непараметрические критерии

Распределение Хи-квадрат ($\chi^2(n)$)

- Допустим, что случайная величина Z распределена нормально с параметрами . Если взять n случайных значений z и найти сумму их квадратов, то полученная сумма будет представлять собой значение некоторой случайной величины, обозначаемой χ^2 (хи-квадрат):

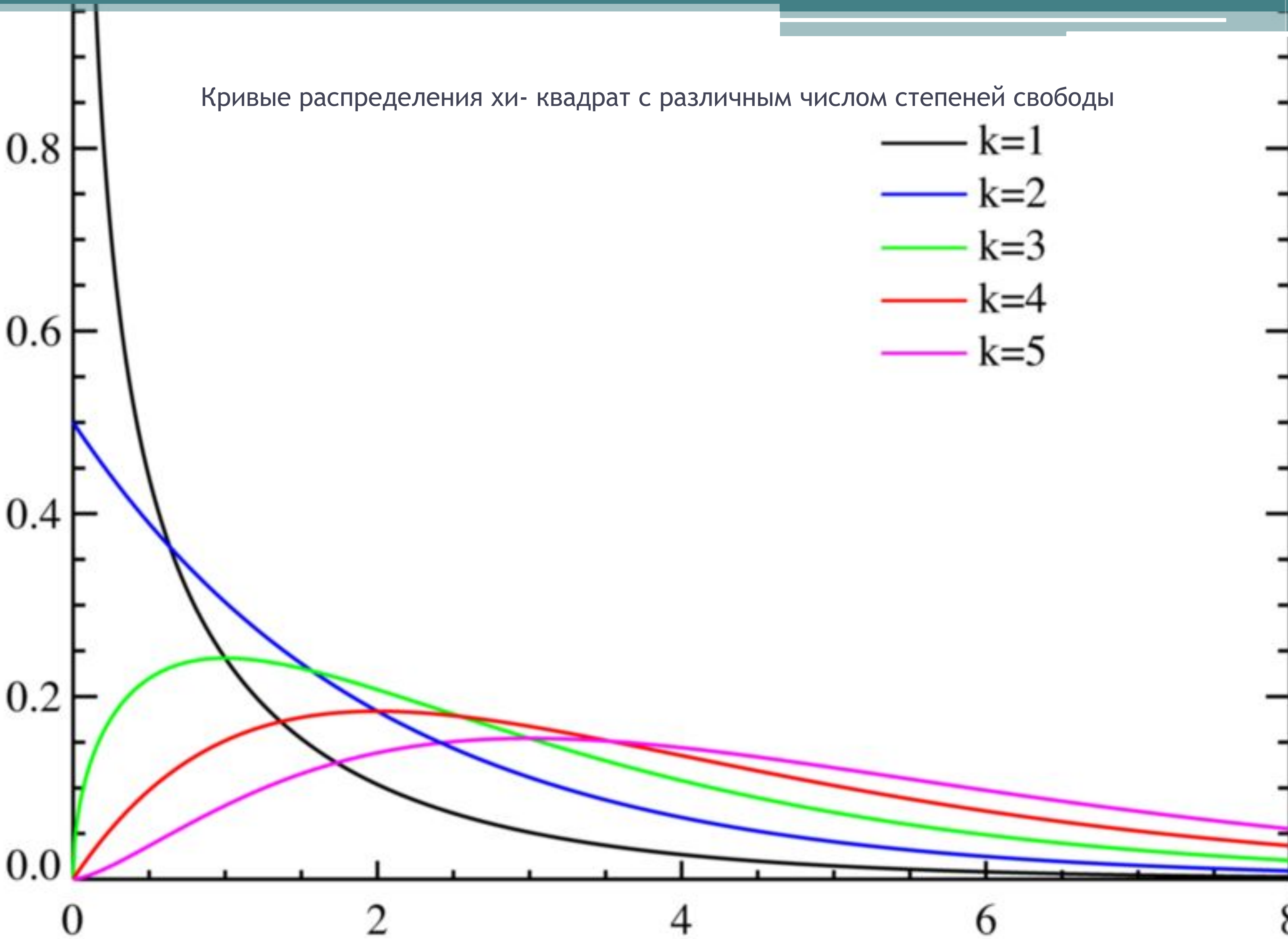
Основные свойства критерия:

Случайная величина χ^2, будучи суммой квадратов, всегда положительна и должна зависеть от числа слагаемых.

Величина χ^2 может принимать значения от 0 до ∞.

- Вид кривой распределения существенно зависит от числа слагаемых, точнее, от числа независимых слагаемых, т.е. от числа степеней свободы ν. При очень малых ν распределение сильно асимметрично, но асимметрия быстро уменьшается по мере увеличения числа степеней свободы. Для распределения χ^2 среднее число равно числу степеней свободы, а дисперсия - удвоенному числу степеней свободы:

Кривые распределения хи- квадрат с различным числом степеней свободы



- Так как закон распределения известен, то не составляет большого труда вычислить критические значения $\chi_\alpha{}^2$, случайно превысить которые при заданном ν можно с вероятностью α.

Для выборок равного объема, $n_1=n_2$ и $N= n_1+n_2$

$$\chi^2_{\phi} = 4 \cdot \sum \left(\frac{f_1^2}{f_1 + f_2} \right) - N$$

Для выборок разного объема, $n_1 \neq n_2$

$$\chi^2_{\phi} = \frac{N^2}{n_1 n_2}\left(\sum \frac{f_1^2}{f_1 + f_2} - \frac{n_1^2}{N}\right)$$

При сравнении эмпирического и теоретического распределения формула используют формулу

$$\chi^2{}_{\phi} = \sum \frac{\left(f_i - f_i^{*}\right)^2}{f_i^{*}}$$

Примите H_0, если $\chi^2_ф < \chi^2_{теор}$.

Отклоните H_0, если $\chi^2_ф \geq \chi^2_{теор}$.

U-критерий Манна-Уитни (англ. *Mann-Whitney U test*) — непараметрический

- статистический критерий, используемый для оценки различий между двумя выборками по уровню какого-либо признака, измеренного количественно. Позволяет выявлять различия в значении параметра между малыми выборками. Другие названия: критерий Манна-Уитни-Уилкоксона (англ. *Mann-Whitney-Wilcoxon*, *MWW*), критерий суммы рангов Уилкоксона (англ. *Wilcoxon rank-sum test*) или критерий Уилкоксона-Манна-Уитни (англ. *Wilcoxon-Mann-Whitney test*).

- Простой непараметрический критерий. Метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами (ранжированным рядом значений параметра в первой выборке и таким же во второй выборке).

- Чем меньше значение критерия, тем вероятнее, что различия между значениями параметра в выборках достоверны.

Для применения U-критерия Манна-Уитни нужно произвести следующие операции:

- 1. Составить единый ранжированный ряд из обеих сопоставляемых выборок, расставив их элементы по степени нарастания признака и приписав меньшему значению меньший ранг. Общее количество рангов получится равным: $N = n_1 + n_2$, где n_1 — количество единиц в первой выборке, а n_2 — количество единиц во второй выборке.

2. Разделить единый ранжированный ряд на два, состоящие соответственно из единиц первой и второй выборок. Подсчитать отдельно сумму рангов, пришедшихся на долю элементов первой выборки, и отдельно — на долю элементов второй выборки. Определить **большую** из двух ранговых сумм (T_x), соответствующую выборке с n_x единиц.

3. Определить значение U-критерия Манна-Уитни по формуле:

$$U_1 = R_1 - \frac{N_1(N_1+1)}{2} \quad \text{и} \quad U_2 = R_2 - \frac{N_2(N_2+1)}{2}$$

4. По таблице для избранного уровня статистической значимости определить критическое значение критерия для данных n_1 и n_2. Если полученное значение U **меньше** табличного или равно ему, то признается наличие существенного различия между уровнем признака в рассматриваемых выборках (принимается альтернативная гипотеза). Если же полученное значение U больше табличного, принимается нулевая гипотеза. Достоверность различий тем выше, чем меньше значение U.

5. При справедливости нулевой гипотезы критерий имеет матожидание и дисперсию и при достаточно большом объёме выборочных данных (n_1>19, n_2>19) распределён практически нормально.

Примите H_0, если $U_ф > U_{теор}$.

Отклоните H_0, если $U_ф \leq U_{теор}$.

Ограничения применимости критерия

1. В каждой из выборок должно быть не менее 3 значений признака. Допускается, чтобы в одной выборке было два значения, но во второй тогда не менее пяти.
2. В выборочных данных не должно быть совпадающих значений (все числа - разные) или таких совпадений должно быть очень мало.

Критерий Колмогорова -Смирнова

- В статистике критерий согласия Колмогорова (также известный, как критерий согласия Колмогорова-Смирнова) используется для того, чтобы определить, подчиняются ли два эмпирических распределения одному закону, либо определить, подчиняется ли полученное распределение предполагаемой модели.

Критерий Колмогорова-Смирнова о проверке гипотезы об однородности двух эмпирических законов распределения является одним из основных и наиболее широко используемых непараметрических методов, так как достаточно чувствителен к различиям в исследуемых выборках.

Критерий Колмогорова-Смирнова о проверке гипотезы об однородности двух эмпирических законов распределения является одним из основных и наиболее широко используемых непараметрических методов, так как достаточно чувствителен к различиям в исследуемых выборках.

Максимальная по модулю разность между соответствующими накопленными относительными частотами является фактическим значением критерия Колмогорова-Смирнова.

$$d_{\text{ф}} = \max|F_1 - F_2|$$

Теоретическое значение критерия Колмогорова Смирнова вычисляется по формуле:

$$d_{теор.} = \lambda \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$$

Примите H_0, если $d_ф < d_{теор}$.

Отклоните H_0, если $d_ф \geq d_{теор}$.

Спасибо за внимание!