Статистические гипотезы

Любое суждение о генеральной совокупности называется **статистической гипотезой.**

Типы гипотез:

1) ***о законе распределения генеральной совокупности***

2) ***о значениях её параметров.***

3) ***….***

Проверка статистической гипотезы состоит в выяснении совместимости выдвинутого ***предположения*** с результатами ***наблюдений***.

Проверка гипотезы базируется на полученной **выборке**.

Всегда возможно расхождение между теоретическим предположением и результатами измерений из-за того, что **элементы выборки – случайные величины**.

Поэтому, при ***малых*** расхождениях теоретических и экспериментальных величин отвергать гипотезу ***не следует***.

Необходимо определить, какие расхождения можно полагать ***пренебрежимо малыми***, а какие – ***существенными*** для отбрасывания выдвинутой гипотезы.

Так как элементы выборки (результаты экспериментов) являются **случайными величинами**, то определенной величине расхождения соответствует некоторая **вероятность**.

Следовательно, выводы о принятии и отвержении гипотезы утверждаются с некоторой **вероятностью**.

Выводы о результате проверки **статистической гипотезы** основаны на **принципе практической невозможности**.

«***Случайное событие с малой вероятностью в однократном испытании произойти не может***» (Чебышев, 1845г.).

Уровень значимости α.

Величина **α** – вероятность **практически невозможного** события в **однократном** испытании.

Если **вероятность** различия теоретических и экспериментальных результатов окажется меньше величины **α**, то это различие объясняется **флуктуациями** элементов выборки и объявляется «**незначимым**».

Статистическая гипотеза ***принимается*** с вероятностью

$$\mathbf{\eta = 1 - \alpha}.$$

В противоположном случае, когда **вероятность** различия теоретических и экспериментальных результатов больше величины **α**, то различие эмпирического и теоретического распределений объявляется «**значимым**», т.е. необъяснимым флуктуациями элементов выборки.

Тогда выдвинутая гипотеза ***отвергается*** на принятом **уровне значимости α**.

Начальная гипотеза, которая проверяется называется **нулевой H_0**

Принятие или отбрасывание нулевой гипотезы осуществляется с помощью некоторого **критерия**.

Критерием K называется некоторая **статистика** (т.е. случайная величина, построенная из элементов выборки), чей закон распределения вероятности известен из теории вероятностей.

На множестве всевозможных значений критерия **{K}** выделяется подмножество **$\{K_0\}$** , называемое **критической областью**.

Критическая область строится так, чтобы вероятность попадания случайного значения критерия **K** в область **$\{K_0\}$** (при условии справедливости нулевой гипотезы) равнялась выбранному уровню значимости **α**.

$$\mathbf{P[\,K \in \{K_0\}: H_0\,] = \alpha} \qquad (1)$$

Алгоритм проверки статистической гипотезы

Формулируется **нулевая гипотеза H_0**.

Выбирается **уровень значимости α**.

Определяется **критическая область $\{K_0\}$** ,

Вычисляется значение критерия **K*** на базе полученной выборки.

Проверяется попадание вычисленного значения **K*** в критическую область **$\{K_0\}$.**

При попадании (**$K^* \in \{K_0\}$**) нулевая гипотеза **H_0** ***отвергается***.

В противоположном случае – нулевая гипотеза **H_0** не отвергается (т.е. ***принимается***).

Статистические ошибки 1-го рода и 2-го рода

Выдвинутая статистическая гипотеза называется **нулевой**.

Ошибка 1-го рода: отбрасывание истинной гипотезы. Равна выбранному уровню значимости **α**.

Кроме нулевой гипотезы всегда существует **альтернативная гипотеза**.

Ошибка 2-го рода: принятие нулевой гипотезы, когда она ложна (т.е. когда верна альтернативная гипотеза).

При выборе критической области (при фиксированном уровне значимости **α**) необходимо ***максимально уменьшать*** **ошибку 2-го рода**:

$$\mathbf{P[\ K \notin \{K_0\}: H_1\ ] = \beta} \qquad (2)$$

Вероятность $\mathbf{P[\ K \in \{K_0\}: H_1\ ] = 1 - \beta}$ (3)

называется **мощностью критерия**.

Практически всегда при уменьшении ошибки 1-го рода **α** начинает возрастать ошибка 2-го рода. Поэтому требуется искать компромисс между величинами ошибок 1-го рода и 2-го рода.

В частности, если $\mathbf{P[\ K \in \{K_0\}:\ H_1\ ] \leq \alpha}$, то отвергать нулевую гипотезу $\mathbf{H_0}$ в пользу альтернативной $\mathbf{H_1}$ было бы принципиально неверно, так как вероятность события $\mathbf{K \in \{K_0\}}$ при альтернативной гипотезе $\mathbf{H_1}$ еще меньше, чем при нулевой гипотезе $\mathbf{H_0}$.

Наилучшим для проверки статистической гипотезы было бы такое **критическое событие**, которое имело бы ***малую*** вероятность при нулевой гипотезе $\mathbf{H_0}$ и ***большую*** вероятность при альтернативной гипотезе $\mathbf{H_1}$.

$$\mathbf{P[\ K \in \{K_0\}:\ H_0\ ] = \alpha} \qquad (1)$$

$$\mathbf{P[\ K \in \{K_0\}:\ H_1\ ] = 1 - \beta} \qquad (3)$$

Критерий согласия Пирсона

Критерий «хи-квадрат»

Область изменения значений *генеральной совокупности* разбивается на R конечных интервалов s_k ($k = 1, 2, \ldots, R$).
Для каждого интервала подсчитываются:

во-первых, **вероятность** попадания значения случайной величины в данный интервал :

$$p_k = P[x \in s_k] \qquad k = 1, 2, \ldots, R \qquad (4)$$

во-вторых, **частота** попадания в данный интервал элементов полученной выборки:

$$v_k = n_k / n \qquad k = 1, 2, \ldots, R \qquad (5)$$

Критерий Пирсона задается формулой

$$K = n\sum_{k=1}^{R}\frac{(\nu_k - p_k)^2}{p_k} \qquad (6)$$

Частоты (5) – случайные величины.
Следовательно, величина (6) – ***случайная***.

Теорема Пирсона.

Случайная величина (6) при $n \longrightarrow \infty$, имеет распределение «**хи-квадрат**» с числом степеней свободы: $R - 1$

Гипотетические распределения в практических задачах часто содержат **параметры**.

Неизвестные генеральные значения параметров приходится заменять их **оценками**, полученными из выборки.

Тогда вероятности, вычисленные по формулам (4) получат ***случайный*** разброс.

Каков тогда **закон распределения** критерия (6) ?

Теорема Фишера.

Случайная величина (5) при $n \to \infty$ имеет распределение «**хи-квадрат**» с числом степеней свободы

$$R - 1 - q\,, \tag{7}$$

где q – количество **параметров** генерального распределения, которые заменены выборочными точечными **оценками**.

Задача о погибших кавалеристах

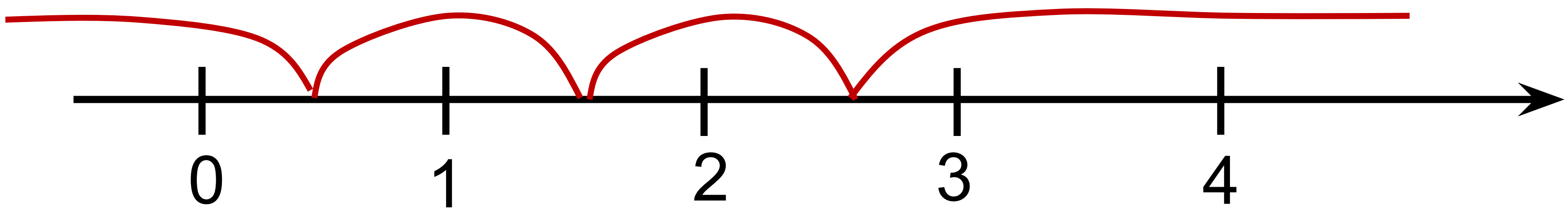
20 лет собирались сведения о количестве кавалеристов прусской армии, погибших в результате гибели под ними коня.

Данные извлекались из ежегодных донесений **10**-и армейских корпусов, что в целом составило **200** донесений.

***k* - количество погибших в год**	**0**	**1**	**2**	**3**	**4**	**>4**
n_k – соответствующее число донесений	**109**	**65**	**22**	**3**	**1**	**0**

Разбиение генеральной совокупности на **интервалы** и расчет **частот**

$R = 4$



s_k	$-\infty$; 0,5	0,5; 1,5	1,5; 2,5	2,5; ∞
k	0	1	2	3; 4
v_k	0,545	0,325	0,11	0,02

Нулевая гипотеза H_0:

Распределение погибших подчиняется закону **Пуассона**

$$P(k) = \frac{a^k}{k!} \exp(-a) \tag{8}$$

Параметр a по смыслу является **математическим ожиданием** пуассоновской случайной величины и его значение неизвестно.

Заменим неизвестный параметр a его приближенным значением – **средним статистическим**

$$a \approx \bar{k} = \frac{\sum k \cdot n_k}{\sum n_k} \approx 0{,}61 \tag{9}$$

$$P(k) = \frac{\bar{k}^{\,k}}{k!} \exp(-\bar{k}) \tag{10}$$

Рассчитаем вероятности по предыдущей формуле для тех же интервалов

s_k	$-\infty; 0,5$	$0,5; 1,5$	$1,5; 2,5$	$2,5; \infty$
v_k	0,545	0,325	0,11	0,02
k	0	1	2	3; 4
p_k	0,543	0,331	0,101	0,024

Вычислим значение критерия Пирсона по данным таблицы

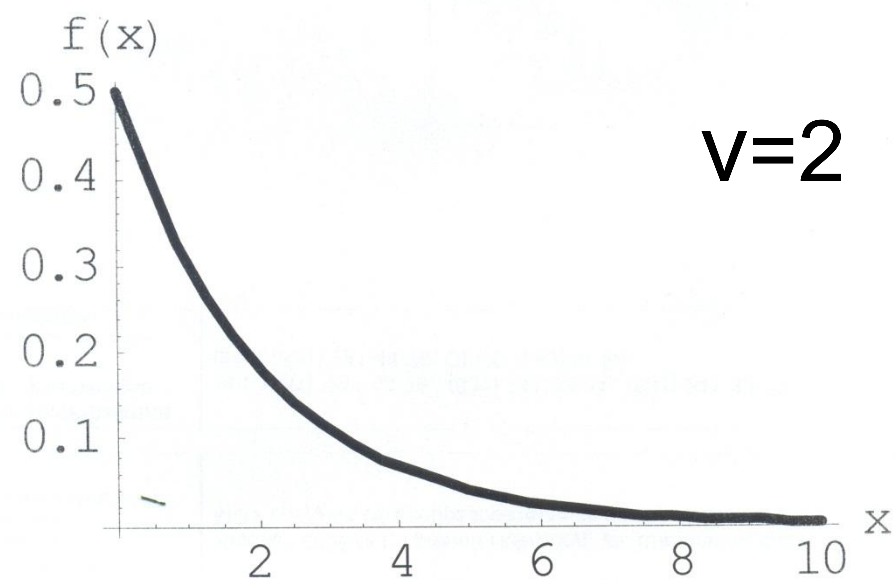
$$K^* = n\sum_{k=1}^{R}\frac{(\nu_k - p_k)^2}{p_k} \approx 0{,}32 \qquad (11)$$

Критерий Пирсона является случайной величиной, распределенной по закону «**хи-квадрат**».

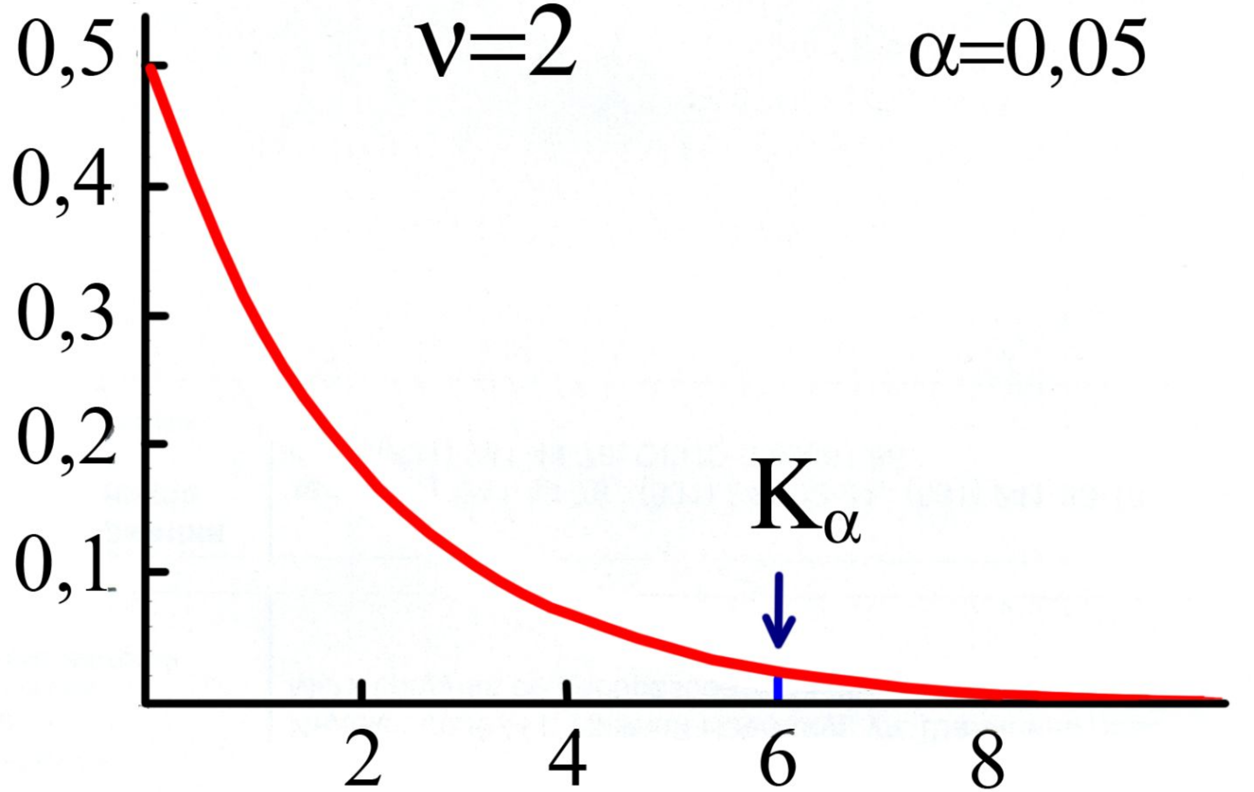
В данной задаче число степеней свободы:

$$R - 1 - 1 = 4 - 1 - 1 = \mathbf{2}$$

Плотность случайной величины **«хи-квадрат»** с числом степеней свободы **2**



Критическую область выбираем в области больших значений критерия (K_α ; $+\infty$)



По заданному уровню значимости **α = 0,05**

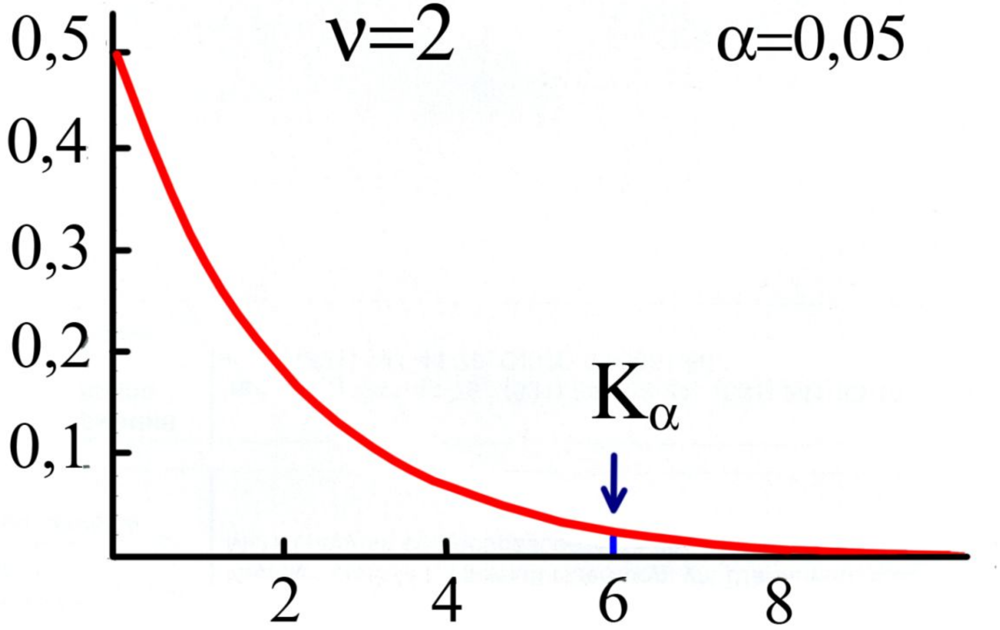
находится предел значимости: $\mathbf{K_\alpha \approx 6}$

Сравнение значения критерия $\mathbf{K^* = 0{,}32}$

с пределом значимости $\mathbf{K_\alpha = 6}$

позволяет **принять нулевую гипотезу** $\mathbf{H_0}$.

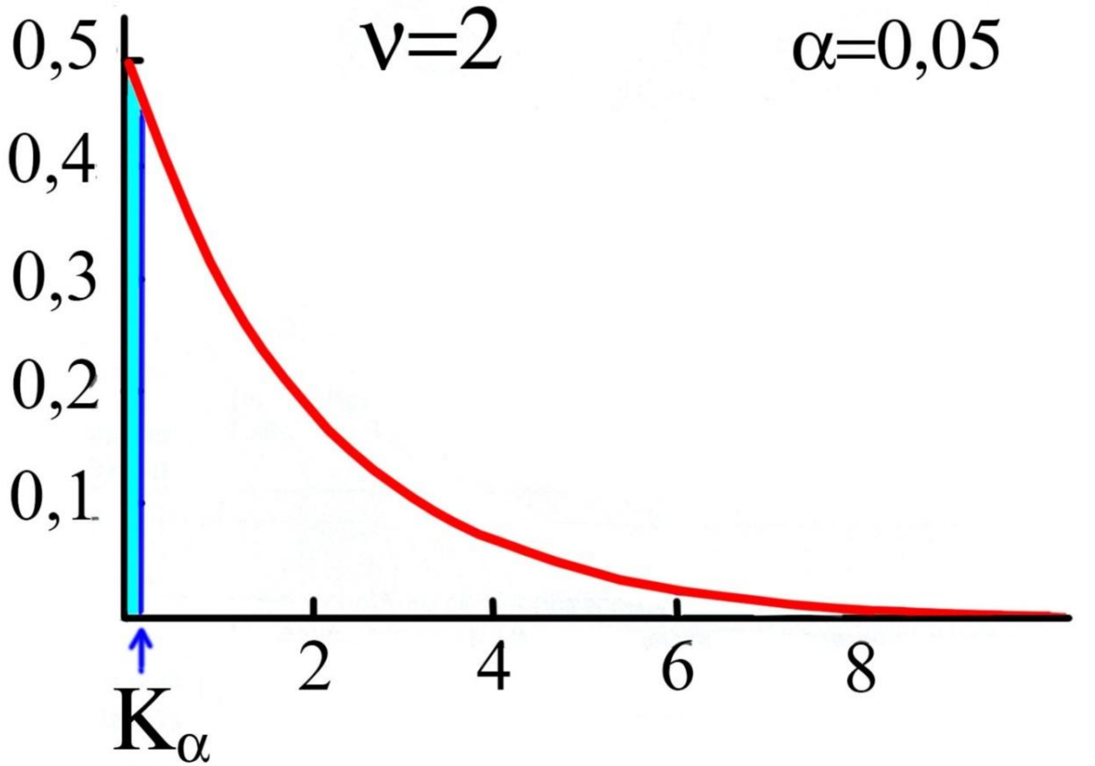
Очевидно, что в данной задаче критическую область следует взять в области больших значений критерия.



$$K = n\sum_{k=1}^{R}\frac{(\nu_k - p_k)^2}{p_k}$$

При большом различии частот ν_k и вероятностей p_k величина критерия Пирсона K будет высока.

Если взять критическую область в области малых значений критерия, то данная гипотеза будет *отвергнута* при почти точном совпадении частот и вероятностей.



$$K = n\sum_{k=1}^{R}\frac{(\nu_k - p_k)^2}{p_k}$$

Результат парадоксальный и абсолютно неверный.