



САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
МЕДИЦИНСКИЙ УНИВЕРСИТЕТ
ИМЕНИ АКАДЕМИКА И. П. ПАВЛОВА



Проверка статистических гипотез



Статистическая гипотеза -- это предположение о генеральной совокупности, высказанное на основании статистических выборочных данных.

Статистическая проверка гипотез -- это процедура обоснованного сопоставления высказанной гипотезы с имеющимися выборочными данными.

Например: исследуем влияние нового лекарственного препарата на снижение артериального давления.

$X\{x_1, x_2, \dots, x_{n_1}\}$ -- контрольная группа (выборка, объёмом n_1)

$Y\{y_1, y_2, \dots, y_{n_2}\}$ -- опытная группа (выборка объёмом n_2)



Высказываются две альтернативные гипотез

H_0 : -- различия между выборками **статистически не значимы** (т.е. носят случайный характер).

H_1 : -- различия между выборками **статистически значимы** (т.е., например, препарат эффективен)

Чтобы принять или опровергнуть эти предположения, используют **статистические критерии**.

Статистический критерий -- это случайная величина, закон распределения которой известен, т.е. каждому значению критерия поставлена в соответствие вероятность, с которой он эти значения принимает.



Для каждого критерия существует **таблица**, в которой содержатся **критические значения критерия**. Каждое критическое значение соответствует определённому **уровню значимости α** и **числу степеней свободы V** (или k)

$$V = n - a$$

где a -- число наложенных связей или ограничений.

$\alpha = 1 - P_d$ -- это вероятность ошибочно отклонить нулевую гипотезу (ошибка первого рода).

Сравнение значения критерия, вычисленного по выборке, с табличным (критическим) значением критерия, позволяет сделать вывод о правомерности выдвигаемой гипотезы для данного уровня значимости.



Например:

Хотим доказать статистическую значимость различия между выборками

$X\{x_1, x_2, \dots, x_{n1}\}$ и $Y\{y_1, y_2, \dots, y_{n2}\}$ с $P_d=0,95$
(это значит, что влияние препарата достоверно (эффективно) на 95%).

Если в результате проверки выяснилось, что вычисленному значению критерия соответствует вероятность бОльшая, чем заданный уровень значимости ($\alpha=1-0,95=0,05$), то нулевая гипотеза принимается.



Основные этапы проверки статистических гипотез.

- 1) Выдвигается гипотеза H_0 .
- 2) Подсчитывается экспериментальное значение критерия по имеющимся выборкам (для каждого критерия существует формула или алгоритм для определения значения критерия).
- 3) Выбирается величина уровня значимости α ($\alpha=1-P_{\text{д}}$).
- 4) По заданному α и числу степеней свободы ν (или k) в таблице находим критическое (табличное) значение критерия.
- 5) Сравнить экспериментальное и критическое значения критерия и сделать вывод о правомерности гипотезы H_0 .



Критерии значимости подразделяются на **параметрические** и **непараметрические**

Параметрические критерии для вычисления экспериментального значения используют статистические параметры: \bar{x} , S_n^2 , S_n , $S_{\bar{x}}$

Они могут использоваться для выборочных совокупностей, **распределённых по закону близкому к нормальному (Гаусса)**.

Непараметрические критерии не требуют вычисления выборочных параметров, указанных выше. Они менее точны, но их можно применять к выборкам, закон распределения которых неизвестен (**не обязательно нормальное распределение**). Если исследуемые выборки распределены нормально, то выводы параметрических и



1. Критерии согласованности с нормальным распределением

Асимметрия и эксцесс – основные показатели, наиболее чувствительные к отклонению от нормальности.



1.1. Коэффициент асимметрии

Кроме среднего арифметического, существуют такие статистические характеристики совокупности как **медиана** и **мода**.

Медиана разделяет ранжированный ряд на две равные части. Если ряд содержит четное значение, берется среднее арифметическое между средними значениями в ряду.

Мода — значение признака, имеющее наибольшую частоту в статистическом ряду распределения.



В симметричном распределении среднее арифметическое, медиана и мода совпадают



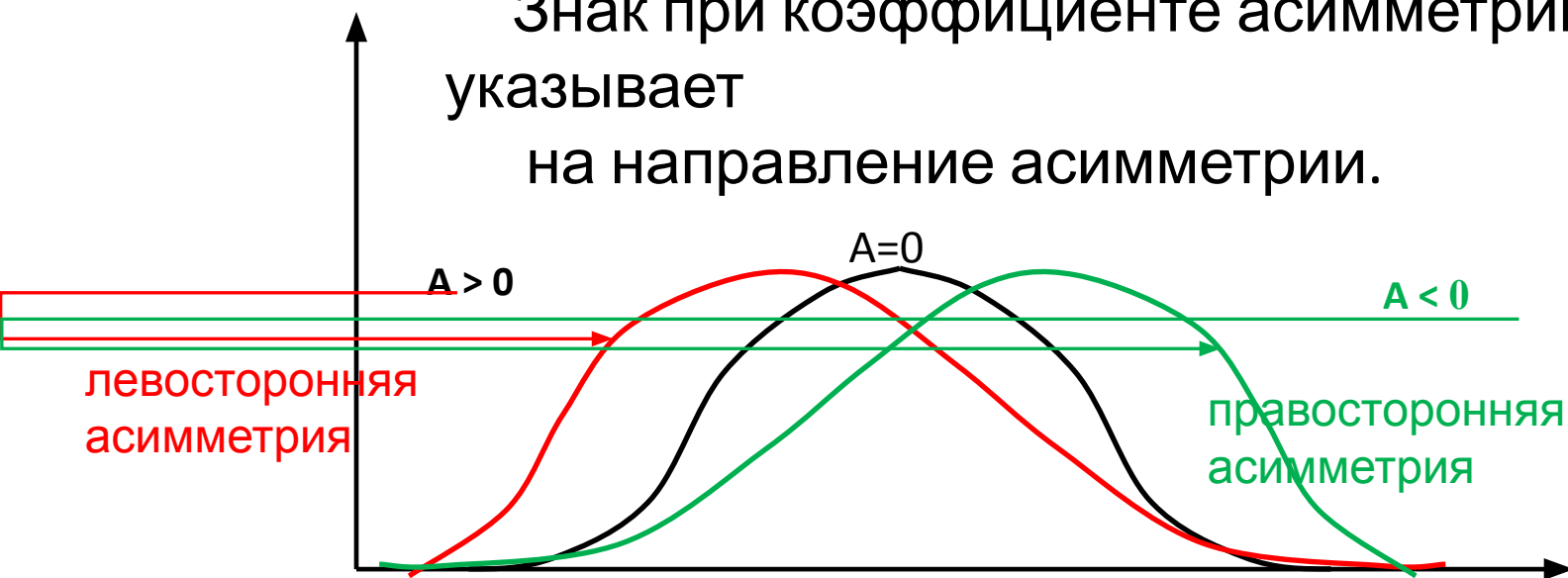
Если же наблюдается асимметрия, то **среднее арифметическое и мода смещаются относительно медианы.**

Асимметрию оценивают по формуле:

$$A = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n \cdot \sigma^3} = \frac{\sum_{i=1}^k (x_i - \bar{x})^3 \cdot p_i}{\sigma^3}$$

где K – количество интервалов

Знак при коэффициенте асимметрии указывает на направление асимметрии.





H_0 : Отличие коэффициента асимметрии от нуля статистически не значимо, то есть распределение нормально по асимметрии.

Вычисляем коэффициент асимметрии по экспериментальным данным по формуле:

$$A_{\text{эксп}} = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{nS_x^3}$$

где K – количество интервалов

Сравниваем $A_{\text{эксп}}$ с табличным (критическим) значением, которое находим в таблице критерия асимметрии для заданного уровня значимости α .



Таблица значений асимметрии

N	$\alpha=0,05$	$\alpha=0,01$
10	1,13	1,49
20	0,92	1,21
30	0,79	1,05
40	0,71	0,93
50	0,63	0,84
60	0,59	0,78
80	0,52	0,68
100	0,47	0,62

Если $|A_{\text{эксп}}| \leq A_{\text{крит}} \Rightarrow H_0$ принимаем.

Вывод: экспериментальное распределение соответствует нормальному по асимметрии.

Если $|A_{\text{эксп}}| > A_{\text{крит}} \Rightarrow H_0$ отвергаем.

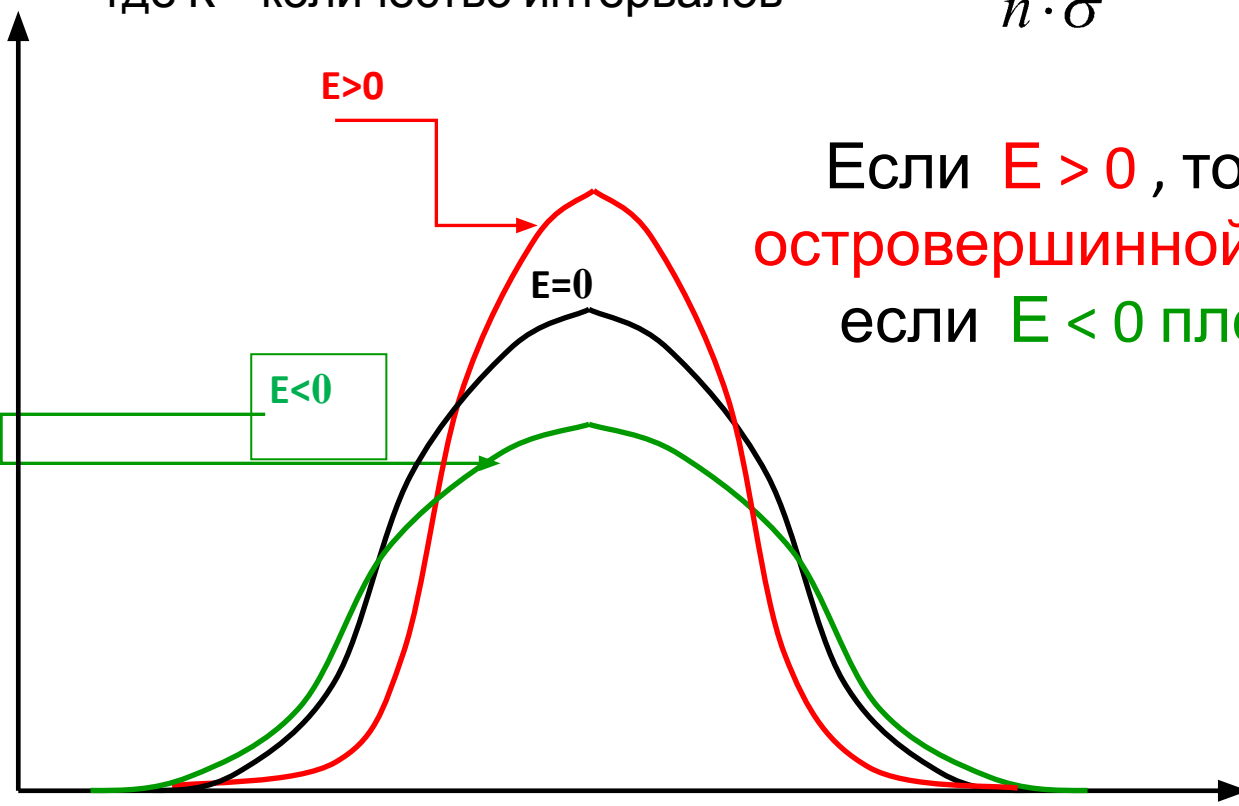
Вывод: экспериментальное распределение не соответствует нормальному по асимметрии.

1.2 Эксцесс.

Иногда этот показатель называют крутостью кривой. Эксцесс вычисляется по формуле:

$$E = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n \cdot \sigma^4} - 3 = \frac{\sum_{i=1}^k (x_i - \bar{x})^4 \cdot p_i}{\sigma^4} - 3$$

где k – количество интервалов



Если $E > 0$, то кривая называется **островершинной**,
если $E < 0$ плосковершинной.



H_0 : Отличие эксцесса от нуля носит случайный характер, то есть распределение **нормально по эксцессу**.

Вычисляем **эксцесс** по экспериментальным данным по формуле:

$$E_{\text{эксп}} = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{nS_x^4} - 3$$

где K – количество интервалов

Сравниваем $E_{\text{эксп}}$ с табличным (критическим) значением, которое находим в таблице критерия эксцесса для заданного уровня значимости α .



Таблица значений эксцесса

N	$\alpha=0,05$	$\alpha=0,01$
10	1,43	
20	1,41	1,95
30	1,31	1,78
40	1,19	1,62
50	1,11	1,50
60	1,05	1,42
80	0,94	1,25
100	0,85	1,14

Если $|E_{\text{экс}}| \leq E_{\text{крит}} \Rightarrow H_0$ принимаем.

Вывод: экспериментальное распределение соответствует нормальному по эксцессу.

Если $|E_{\text{экс}}| > E_{\text{крит}} \Rightarrow H_0$ отвергаем.

Вывод: экспериментальное распределение не соответствует нормальному по эксцессу.



Проверка гипотез о законе распределения

Проверку гипотезы о законе распределения (то есть, соответствует ли выборочная совокупность какому либо определённом распределению) проводят с помощью критерия соответствия (предложен К.Пирсоном в 1900г.).

Критерий χ^2 Пирсона

H_0 заключается в том, что различие между наблюдаемыми экспериментальными частотами m_i попадания вариант выборки в интервалы вариационного ряда от вычисленных теоретических частот $m_{i \text{ теор}} = n \cdot P_{i \text{ теор}}$ статистически не значимо (т.е. носит случайный характер). Другими словами:

H_0 : экспериментальные данные соответствуют предложенному теоретическому закону распределения.



Экспериментальное значение критерия вычисляется по формуле:

$$\chi_{\text{эксп}}^2 = \sum_{i=1}^k \frac{(m_{\text{эксп}} - m_{\text{теор}})^2}{m_{\text{теор}}} = \sum_{i=1}^k \frac{(m_{\text{эксп}} - n \cdot P_{\text{теор}})^2}{n \cdot P_{\text{теор}}}$$

где $\sum_{i=1}^k m_i = n$ -- объём выборки, k -- количество интервалов,

$P_{\text{теор}}$ -- вероятность попадания в интервал для теоретического распределения.

Затем, по таблице критерия Пирсона для **заданного уровня значимости α и числа степеней свободы $\nu = k - a$**

где a -- **число наложенных связей**, находим $\chi_{\text{крит}}^2 = \chi_{\text{табл}}^2$



если теоретическое распределение произвольное, то $a=1$,

$$v = k - 1$$

если теоретическое распределение распределено по нормальному закону Гаусса, то $a=3$ -- числу наложенных связей, необходимых для вычисления вероятности: $n, M[X]$, и σ

$$[X], \quad v = k - 3$$

$$\text{Если } \chi_{\text{эксп}}^2 \leq \chi_{\text{крит}}^2 \Rightarrow$$

H_0 принимаем.

Вывод: экспериментальное распределение соответствует теоретическому.

$$\text{Если } \chi_{\text{эксп}}^2 > \chi_{\text{крит}}^2 \Rightarrow$$

H_0 отвергаем.

Вывод: экспериментальное распределение не соответствует теоретическому.



Пример: Изучался рост 50 человек. В таблице приведены экспериментальные частоты попадания в интервал m_i и теоретические частоты, рассчитанные из вероятностей попадания в интервал для распределения Гаусса. $K=5$, $n=50$.
 $v=5-3=2$

№ интервала	1	2	3	4	5
m_i практические	5	9	22	8	6
m_i теоретические	5	10	20	10	5
	0,1	0,2	0,4	0,2	0,1



H_0 : Экспериментальное (практическое) распределение соответствует распределению Гаусса.

Из таблицы для $\nu=5-3=2$ и $\alpha=0,05$ находим

$$\begin{aligned} \chi_{\text{эксп}}^2 &= \frac{(5-5)^2}{5} + \frac{(9-10)^2}{10} + \frac{(22-20)^2}{20} + \frac{(8-10)^2}{10} + \frac{(6-5)^2}{5} = 0 + \frac{1}{10} + \frac{4}{20} + \frac{4}{10} + \frac{1}{5} \\ &= \frac{1+2+4+2}{10} = 0,9 \end{aligned}$$

Т.к. $\chi_{\text{эксп}}^2 < \chi_{\text{крит}}^2 \Rightarrow H_0$ принимаем.

Вывод: исследуемое выборочное распределение соответствует распределению Гаусса.



Значения критерия Пирсона (критерия χ^2)

Число степеней свободы, ν	$\alpha=0,05$	$\alpha=0,01$
1	3,84	6,63
2	5,99	9,21
3	7,81	11,3
4	9,49	13,3
5	11,1	15,1
6	12,6	16,8
7	14,1	18,5
8	15,5	20,1
9	16,9	21,7
10	18,3	23,2



Параметрические критерии. Критерий Фишера

Этот параметрический критерий служит для проверки нулевой гипотезы о равенстве дисперсий нормально распределенных генеральных совокупностей.

$$\text{H}_0: D[X] = D[Y] \quad (\sigma_1^2 = \sigma_2^2)$$

Экспериментальное значение критерия вычисляется по формуле:

$$F_{\text{экс}} = \frac{\sigma_1^2}{\sigma_2^2}$$

$$\sigma_1^2 \geq \sigma_2^2, F \geq 1, \nu_1 = (n_1 - 1), \nu_2 = (n_2 - 1)$$

где n_1, n_2 – объемы выборок,

ν_1, ν_2 – числа степеней свободы для этих выборок.



Сравниваем $F_{\text{эксн}}$ с табличным (критическим) значением, которое находим в таблице критерия Фишера для заданного уровня значимости α и числа степеней свободы ν_1 и ν_2 .

При пользовании таблицами следует обратить внимание, что число степеней свободы для выборки с большей по величине дисперсией выбирается как номер столбца таблицы, а для меньшей по величине дисперсии как номер строки таблицы

Если $F_{\text{эксн}} \leq F_{\text{крит}} \Rightarrow H_0$ принимаем.

Вывод: дисперсии двух генеральных совокупностей можно считать равными.

Если $F_{\text{эксн}} > F_{\text{крит}} \Rightarrow H_0$ отвергаем.

Вывод: дисперсии двух генеральных совокупностей не равны.



Пример. Изучали влияние пищевых добавок на массу тела лабораторных животных. Опыт проводился на двух группах животных: опытной и контрольной. В опытной группе животные получали пищевую добавку к рациону. За время опыта прибавки в весе составили в граммах:

X: Опыт	Y: Контроль
580	500
690	560
700	420
619	621
703	580
560	530
	450



$$H_0 : D[X] = D[Y]$$

$$\bar{X} = 642, \bar{Y} = 523$$

$$\sigma_{\text{экс}}^2 = S_X^2 = 4097, \sigma_Y^2 = S_Y^2 = 5143, F = \frac{\sigma_y^2}{\sigma_x^2} = \frac{5123}{4097} = 1,25$$

ДЛЯ $\alpha = 0,05, \nu_1 = 7 - 1 = 6, \nu_2 = 6 - 1 = 5, F_{\text{крит}} = 4,95$

Так как $F_{\text{экс}} < F_{\text{крит}} \Rightarrow H_0$ принимаем

Вывод: дисперсии двух генеральных совокупностей можно считать равными.



Таблица критерия Фишера ($\alpha=0,05$)

v_2	Число степеней свободы v_1									
	1	2	3	4	5	6	7	8	9	10
1	161	200	216	225	230	234	237	239	241	242
2	18,51	19,0	19,2	19,3	19,3	19,3	19,6	19,4	19,4	19,4
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35



Контрольные вопросы.

1. Что такое статистическая гипотеза и критерии проверки статистических гипотез?
2. Основные этапы проверки статистических гипотез.
3. Критерий Асимметрии.
4. Критерий Эксцесса.
5. Критерий Пирсона .
6. Критерий Фишера.

