

Критерии достоверности оценок.

Статистическая проверка гипотез.

Для сравнительной оценки генеральных параметров выборок используют нулевую гипотезу $H_0.(\mu_1 - \mu_2 = 0; \bar{x}_1 - \bar{x}_2 = 0)$

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, Хи²-критерий).

Критерий Стьюдента.

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

Пусть \bar{X}_1 и \bar{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 H_0 -гипотезы служит отношение:

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

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

H_0 -гипотезу отвергают, если фактически установленная величина t_{ϕ} -критерия превзойдет или окажется равной стандартному t_{st} - этой величины для принятого уровня значимости α и числа степеней свободы $k = n_1 + n_2 - 2$,

т. е. при условии: $t_{\phi} \geq t_{st}$, Если $t_{\phi} < t_{st}$, то H_0 -гипотеза сохраняется.

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

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

Задача:

На двух группах крыс поставлен опыт по сравнению влияния разных рационов на рост. Крысы первой группы получали рацион с высоким содержанием белка, крысы второй – с низким. Привесы за 56 дней опыта для каждой крысы составляли в (г):

Высокобелк рацион	134	146	104	119	124	161	107
Низкобелк рацион	70	118	101	85	107	132	94

Применяя t-критерий Стьюдента определить достоверность влияния высокобелкового рациона на рост крыс.

Для решения задачи составляют таблицу:

№	X_{1i}	X_{2i}	$X_{1i} - \bar{X}_1$	$X_{2i} - \bar{X}_2$	$(X_{1i} - \bar{X}_1)^2$	$(X_{2i} - \bar{X}_2)^2$
1	134	70	6	-31	38	961
2	146	118	18	17	329	289
3	104	101	-24	0	569	0
4	119	85	-9	-16	78	256
5	124	107	-4	6	15	36
6	161	132	33	31	1098	961
7	107	94	-21	-7	435	49
$\Sigma =$	895	707			2563	2552
$\bar{X} =$	128	101				
	n=	7				

Схема вычисления критерия Стьюдента:

1. Находят средние значения в первой и второй выборке (\bar{X}_1 и \bar{X}_2).
2. Находят разность между каждым значением случайной величины и средним значением в первой и второй выборке.
3. Возводят в квадрат полученные разности.
4. Суммируют значения полученных разностей в первой и второй выборке.
5. Подставляют полученные суммы в формулу критерия Стьюдента и рассчитывают фактическое значение критерия Стьюдента по формуле:

$$t_0 = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sum(x_i - \bar{x}_1)^2 + \sum(x_i - \bar{x}_2)^2}{n(n-1)}}} = \frac{|127,86 - 101|}{\sqrt{\frac{2562,86 + 2552}{42}}} = 2,43$$

Значение критерия Стьюдента для $P=0,95$ и $R=n_1+n_2-2$ числа степеней свободы: $R=7+7-2=12$, $t_{st}=2,18$

7. Делают вывод:

$t_{ф} \geq t_{st}$, Но – отвергается, высокобелковый рацион на рост крыс влияет.

F-критерий Фишера. Проверка гипотез для дисперсий.

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

$$F = \frac{\sigma_1^2}{\sigma_2^2} \quad \text{при } \sigma_1^2 \geq \sigma_2^2.$$

Так как принято брать отношение большей дисперсии к меньшей, то критерий $F \geq 1$.

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 $F_{\phi} \geq F_{st}$ и нулевая гипотеза должна быть отвергнута.

Задача:

Пусть при лечении некоторого заболевания применяются две методики: А и В. Отобраны две однородные группы больных, первая численностью $n_1=20$, а вторая- $n_2=16$ человек. Известно, что соответствующие генеральные совокупности X и Y имеют нормальное распределение. Оказалось, что для обеих групп средние значения практически равны, а выборочные дисперсии: $\sigma_1^2 = 21,5$ и $\sigma_2^2 = 32,8$. Требуется сопоставить обе методики лечения при уровне значимости $\alpha=0,1$.

Решение:

Дисперсия для метода А: $\sigma_1^2=21,5$

Дисперсия для метода В: $\sigma_2^2=32,8$

Вычисляют дисперсионное отношение $F = \frac{32,8}{21,5} = 1,526$

В таблице для 1% уровня значимости (нижняя цифра) и чисел степеней свободы $k_1=16-1=15$ (см. верхнюю строку таблицы) и $k_2=20-1=19$ (см. первую графу той же таблицы) находят $F_{st}=3,15$

Вывод: Так как $F_{\phi} < F_{st}$. нулевая гипотеза остаётся в силе. Обе методики эквивалентны друг другу.

Критерий Ван-дер-Вардена.

Этот критерий относится к группе ранговых критериев, его применяют для проверки нулевой гипотезы при сравнении друг с другом независимых выборок. Техника расчётов X -критерия Ван-дер-Вардена сводится к следующему. Сравнимые выборки ранжируют в один общий ряд по возрастающим значениям признака. Затем каждому члену ряда присваивают порядковый номер (R), отмечающий его место в общем ранжированном строю. Далее по порядковым номерам одной из выборок, обычно меньшей по объёму, находят отношение $R/(N+1)$, где $N+1=n_1+n_2+1$.

С помощью таблицы “Значения функции $\psi[R/(N+1)]$ “ находят значения функции $\psi[R/(N+1)]$ для каждого значения $R/(N+1)$. Суммируя результаты (обязательно с учётом знаков), получают величину $X_{\phi} = \sum \psi[R/(N+1)]$, которую сравнивают с критической точкой этого критерия X_{st} для принятого уровня значимости и общего числа членов сравниваемых выборок, т.е. $N = n_1 + n_2$. Критические точки X -критерия для 5%-ного и 1%-ного уровней значимости и общего числа членов двух выборок $N = n_1 + n_2$ (с учётом разности $n_1 - n_2$) содержатся в таблице “Критические значения X -критерия Ван-дер-Вардена».

Нулевая гипотеза сводится к предположению, что сравниваемые выборки извлечены из генеральных совокупностей с одинаковыми функциями распределения. Если окажется, что

$X_{\phi} \geq X_{st}$, нулевая гипотеза должна быть отвергнута на принятом уровне значимости.

Задача:

Применяя критерий Ван-дер-Вардена определить достоверность влияния токсических свойств винилпропионата на среднее время гибели мышей на двух группах животных.

Опыт	22	35	39	41	43	45	46	48	48	69
Контр	13	14	17	22	26	27	30	32	40	55

Решение: Данные заносят в таблицу.

контр	опыт	контроль (по возрастанию)	Опыт (по возрастанию)	R	R/(N+1)	\square
13	22	13		1		
14	35	14		2		
17	39	17		3		
22	41	22		4		
26	43		22	5	0,24	- 0,71
27	45	26		6		
30	46	27		7		
32	48	30		8		
40	48	32		9		
55	69		35	10	0,48	- 0,05
			39	11	0,52	0,05
		40		12		
			41	13	0,62	0,31
			43	14	0,67	0,44
			45	15	0,71	0,55
			46	16	0,76	0,71
			48	17	0,81	0,88
			48	18	0,86	1,08

		55		19		
n₁=10	n₂=10		69	20	0,95	1,64
					$\Sigma=$	4,9

План решения задачи.

1. Выстраивают данные задачи в один общий возрастающий вариационный ряд.
2. Ранжируют (нумеруют **R**) полученные данные.
3. Выбирают одну выборку (обычно меньшую по объёму). Если выборки равночисленные, то выбирают любую.
4. Рассчитывают значение $\frac{R}{N+1}$ для каждого значения выбранной выборки.
5. По таблице №6 «Значение функции $\phi(R/N+1)$ » находят значение ϕ функции для каждого $\frac{R}{N+1}$.
6. Суммируют значения ϕ функции с учётом знаков (если получается отрицательное значение, то берут модуль этой суммы) и получают фактическое значение **X_ф=4.9**
7. Находят стандартное значение **X_{ст}=3,86** (по таблице №7) для **P=0,95**
8. Делают вывод: **X_ф>X_{ст}**, H₀-гипотеза отвергается.

Ответ: Влияние токсических свойств винилпропионата на среднее время гибели мышей достоверно.

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

H₀: 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.

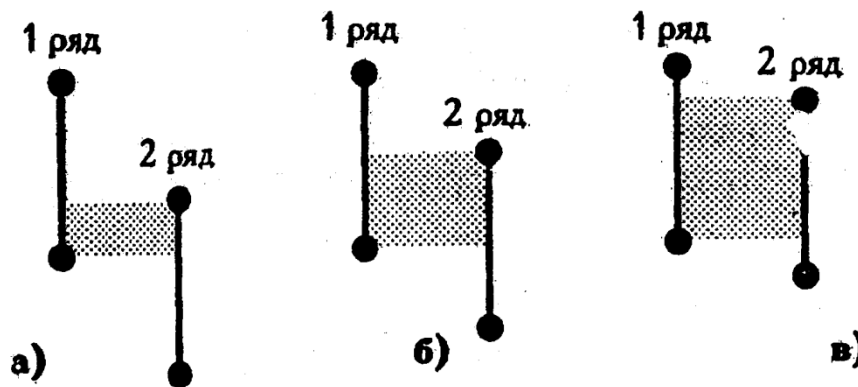


Рис1. Возможные варианты соотношений рядов значений в двух выборках; штриховкой обозначены зоны наложения.

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 обеих 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 $U_{\text{эмп}} > U_{\text{кр} 0.05}$, H_0 принимается.

9. Если $U_{\text{эмп}} \leq U_{\text{кр} 0.05}$, H_0 отвергается. Чем меньше значения U , тем достоверность различий выше.

Задача:

Имеются две группы лабораторных мышей, опытная ($n_1=9$) и контрольная ($n_2=11$). Проверить с помощью критерия U Манна-Уитни, оказывает ли действие на массу тела лабораторных мышей новый лекарственный препарат.

Масса тела мышей	Опыт	75	70	64	68	72	79	76	83	80		
	контроль	71	70	66	60	62	69	73	69	60	80	78

Для решения задачи создаём таблицу:

Масса тела мышей		Масса тела мышей (по возрастанию)		R	O	R_o	K	R_k
O	K	O	K					
75	71		60	1,5			60	1,5
70	70		60	1,5			60	1,5
64	66		62	3			62	3
68	60	64		4	64	4		
72	62		66	5			66	5
79	69	68		6,5	68	6,5		
76	73		68	6,5			68	6,5
83	69		69	8			69	8
80	60	70		9,5	70	9,5		
	80		70	9,5			70	9,5
	78		71	11			71	11
		72		12	72	12		
			73	13			73	13
		75		14	75	14		
		76		15	76	15		
			78	16			78	16
		79		17	79	17		
			80	18,5			80	18,5
		80		18,5	80	18,5		
		83		20	83	20		
						$\Sigma=116,5$		$\Sigma=93,5$

Общая сумма рангов: $\sum R = 116,5 + 93,5 = 210$.

Расчетная сумма: $\sum R_1 = \frac{N \cdot (N + 1)}{2} = \frac{20 \cdot (20 + 1)}{2} = 210$

Равенство реальной и расчетной сумм соблюдено.

Опытная группа имеет большую сумму рангов $\sum R_o = 116.5$.

Теперь необходимо сформулировать гипотезу H_0 :

H_0 : Опытная группа не превосходит контрольную группу по массе тела.

Определяем эмпирическую величину U :

$$U_{\text{эп}} = (11 \cdot 9) + \frac{9 \cdot (9 + 1)}{2} - 116.5 = 27.5$$

По табл. 8 определяем критические значения для соответствующих n_1 и n_2 , причем меньшее n принимаем за n_1 ($n_1=9$) и отыскиваем его в верхней строке табл.8, большее n принимаем за n_2 ($n_2=11$) и отыскиваем его в левом столбце табл.8.

$$U_{\text{крит}} = \begin{cases} 27 (\delta \leq 0.05) \\ 18 (\delta \leq 0.01) \end{cases}$$

Критерий U является одним из двух исключений из общего правила принятия решения о достоверности различий, а именно мы можем констатировать достоверные различия, если $U_{\text{эп}} \leq U_{\text{крит}}$.

$U_{\text{эп}} > U_{\text{крит}}$, при уровне значимости 0,05, следовательно между группами различий нет и новый лекарственный препарат не оказывает действия на массу мышей.

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.

$$f_1 \quad f_2 \quad f_3 \quad \dots \quad f_n$$

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

$$f'_1 \quad f'_2 \quad f'_3 \quad \dots \quad f'_n$$

3. Находят разность между эмпирическими и теоретическими частотами.

$$f_1 - f'_1 \quad f_2 - f'_2 \quad f_3 - f'_3 \quad \dots \quad f_n - f'_n$$

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

$$(f_1 - f'_1)^2 \quad (f_2 - f'_2)^2 \quad (f_3 - f'_3)^2 \quad \dots \quad (f_n - f'_n)^2$$

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

$$\frac{(f_1 - f'_1)^2}{f'_1} \quad \frac{(f_2 - f'_2)^2}{f'_2} \quad \frac{(f_3 - f'_3)^2}{f'_3} \quad \dots$$

6. Находят сумму отношений квадратов разности между эмпирическими и

теоретическими частотами к теоретическим частотам.

$$\frac{(f_1 - f'_1)^2}{f'_1} + \frac{(f_2 - f'_2)^2}{f'_2} + \frac{(f_3 - f'_3)^2}{f'_3} + \dots + \frac{(f_n - f'_n)^2}{f'_n}$$

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 X_{ϕ}^2 сравнить с её критическим значением X_{st}^2 , определяемом по таблице “ X^2 -распределение. Критические точки для разных значений вероятности P и чисел степеней свободы-R”.

Если $X_{\phi}^2 \geq X_{st}^2$, то нулевая гипотеза должна быть отвергнута на принятом уровне значимости с числом степеней свободы R.

Задача:

В таблице приведены эмпирические (f) и вычисленные по нормальному закону(f') частоты распределения длины тела у 267 мужчин. Применяя критерий **Хи-квадрат** выяснить, следует ли это распределение нормальному закону.

Решение:

Частоты		f-f'	(f-f') ²	$\frac{(f-f')^2}{f'}$
f-опытные	f'-теоретическ.			
12	11.6	0.4	0.16	0.01
31	34.3	3.3	10.89	0.32
71	67.8	3.2	10.24	0.15
82	77.6	4.4	19.36	0.25
46	51.2	5.2	27.04	0.53
19	19.5	0.5	0.25	0.01
6	5	1.0	1.00	0.20
$\Sigma=267$	$\Sigma=267$		$\Sigma=$	1.47

Расчёт значения **Хи-квадрат** критерия дал результат **1.47**.

В данном случае число степеней свободы **R =7-3=4**. Исходя из **5%**-ного уровня значимости из таблицы “ X^2 -распределение. Критические точки для разных значений вероятности P и чисел степеней свободы-k) находим $X_{st}^2=9.49$

Эта величина значительно превышает $X_{\phi}^2=1.47$, что не позволяет отвергнуть

H_0 -гипотезу.

Следовательно, существуют достаточные основания для утверждения, что данное распределение следует нормальному закону.

Многофункциональные статистические критерии

Многофункциональные статистические критерии – это критерии, которые могут использоваться по отношению к самым разнообразным данным, выборкам и задачам.

1. Данные могут быть представлены в любой шкале, начиная от номинативной (шкалы наименований).
2. Выборки могут быть как независимыми, так и «связанными», то есть мы можем с помощью многофункциональных критериев сравнивать и разные выборки испытуемых, и показатели одной и той же выборки, измеренные в разных условиях. Нижние границы выборок – 5 наблюдений, но возможно применение критериев и по отношению к выборкам $n=2$, с некоторыми оговорками.
3. В критерии φ^* Фишера верхней границы не существует – выборки могут быть насколько удобными большими.

К числу многофункциональных критериев относится критерий φ^* Фишера (угловое преобразование Фишера).

Многофункциональные критерии построены на сопоставлении долей выраженных в долях единицы или в процентах. Суть критериев состоит в определении того, какая доля наблюдений (реакций, выборов, испытуемых) в данной выборке характеризуется интересующим исследователя эффектом и какая доля этим эффектом не характеризуется.

Критерий φ^* применяется в тех случаях, когда обследованы две выборки испытуемых.

Критерий φ^* - угловое преобразование Фишера

Критерий Фишера предназначен для сопоставления двух выборок по частоте встречаемости интересующего исследователя эффекта.

Описания критерия

Критерий оценивает достоверность различий между процентными долями двух выборок, в которых зарегистрирован интересующий нас эффект.

Суть углового преобразования Фишера состоит в переводе процентных долей в величины центрального угла, который измеряется в радианах. Большой процентной доле будет соответствовать больший угол φ , а меньшей доле – меньший угол, но соотношения здесь не линейные:

$$\varphi = 2 \times \arcsin(\sqrt{P})$$

Где P – процентная доля, выраженная в долях единицы

При увеличении расхождения между углами φ_1 , φ_2 и увеличения

численности выборок значение критерия возрастает. Чем больше величина φ^* , тем более вероятно, что различия достоверны.

Гипотезы:

H_0 : Доля лиц, у которых проявляется исследуемый эффект, в выборке 1 не больше, чем в выборке 2.

Ограничения критериев φ^*

1. Ни одна из сопоставляемых долей не должна быть равной нулю. Формально нет препятствий для применения метода φ в случаях, когда доля наблюдений в одной из выборок равна 0. Однако в этих случаях результат может оказаться неоправданно завышенным.
2. Верхний предел в критерии φ отсутствует – выборки могут быть сколь угодно большими.
3. Нижний предел – 2 наблюдения в одной из выборок. Однако должны соблюдаться следующие соотношения в численности двух выборок:
 - а) если в одной выборке всего 2 наблюдения, то во второй должно быть не менее 30:
 $n_1 = 2 \rightarrow n_2 \geq 30$;
 - б) если в одной из выборок всего 3 наблюдения, то во второй должно быть не менее 7:
 $n_1 = 3 \rightarrow n_2 \geq 7$;
 - в) если в одной из выборок всего 4 наблюдения, то во второй должно быть не менее 5:
 $n_1 = 4 \rightarrow n_2 \geq 5$;
 - г) при $n_1, n_2 \geq 5$ возможны любые сопоставления;

В принципе возможно и сопоставление выборок, не отвечающих этому условию, например, соотношением $n_1 = 2, n_2 = 15$, но в этих случаях не удастся выявить достоверных различий.

Других ограничений у критерия φ^* нет.

Пример 1 - сопоставления выборок по качественно определяемому признаку

В данном варианте использования критерия мы сравниваем процент испытуемых в одной выборке, характеризующиеся каким-либо качеством, с процентом испытуемых в другой выборке, характеризующихся тем же качеством.

Допустим, нас интересует, различаются ли две группы студентов по успешности решения новой экспериментальной задачи. В первой группе из 20 человек с нею справилось 12 человек, а во второй выборке из 25 человек – 10. В первом случае процентная доля решивших задачу составит $12/20 \cdot 100\% = 60\%$, а во второй $10/25 \cdot 100\% = 40\%$. Достоверно ли различаются эти процентные доли при данных n_1 и n_2 ?

Казалось бы, и «на глаз» можно определить, что 60% значительно выше 40%. Однако на самом деле эти различия при данных n_1, n_2 недостоверны.

Проверим это. Поскольку нас интересует факт решения задачи, будем считать «эффектом» успех в решении экспериментальной задачи, а отсутствие эффекта - неудачу в ее решении.

Сформулируем нулевую гипотезу.

H_0 : Доля лиц, справившихся с задачей, в первой группе не больше, чем во второй группе.

Теперь построим так называемую четырехклеточную, или четырехпольную таблицу, которая фактически представляет собой таблицу эмпирических частот по двум значениям признака: «есть эффект» - «нет эффекта».

Четырехклеточная таблица для расчета критерия при сопоставлении двух групп испытуемых по процентной доле решивших задачу.

группы	«есть эффект»: задача решена			«нет эффекта»: задача не решена			Суммы
	количество испытуемых	% доля		количество испытуемых	% доля		
1 группа	12	(60%)	А	8	(60%)	Б	20
2 группа	10	(40%)	В	15	(40%)	Г	25
суммы	22			23			45

В четырехклеточной таблице, как правило, сверху отмечаются столбцы «есть эффект» и «нет эффекта», а слева – строки «1 группа» и «2 группа».

Участвуют в сопоставлениях, собственно, только поля (ячейки) А и В, то есть процентные доли по столбцу «Есть эффект».

По табл. 10 определяем величины φ , соответствующие процентным долям в каждой из групп.

$$\varphi_{1(60\%)} = 1,772$$

$$\varphi_{2(40\%)} = 1,396$$

Теперь подсчитаем эмпирическое значение φ^* по формуле:

$$\varphi^* = (\varphi_1 - \varphi_2) * \sqrt{\frac{n_1 * n_2}{n_1 + n_2}}$$

Где φ_1 – угол, соответствующий большей % доле;

φ_2 – угол, соответствующий меньшей % доле;

n_1 – количество наблюдений в выборке 1;

n_2 – количество наблюдений в выборке 2;

В данном случае:

$$\varphi^* = (1,772 - 1,396) * \sqrt{\frac{20 * 25}{20 + 25}} = 0,403 * \sqrt{11,11} = 1,34$$

По Табл.9 определяем, какому уровню значимости соответствует $\varphi^*_{эмп} = 1,34$:
 $P = 0,09$

Алгоритм расчета критерия φ^*

1. Определить те значения признака, которые будут критерием для разделения испытуемых на тех, у кого «есть эффект» и тех, у кого «нет эффекта».

2. Начертить четырехклеточную таблицу из двух столбцов и двух

строк. Первый столбец – «есть эффект»; второй столбец – «нет эффекта»; первая строка сверху – 1 группа (выборка1); вторая строка – 2 группа (выборка2).

3. Подсчитать количество испытуемых в первой группе, у которых «есть эффект», и занести это число в левую верхнюю ячейку таблицы.

4. Подсчитать количество испытуемых в первой выборке, у которых «нет эффекта», и занести это число в правую верхнюю ячейку таблицы.

5. Подсчитать сумму по двум верхним ячейкам. Она должна совпадать с количеством испытуемых в первой группе.

6. Подсчитать количество испытуемых во второй группе, у которых «есть эффект», и занести это число в левую нижнюю ячейку таблицы.

7. Подсчитать количество испытуемых во второй выборке, у которых «нет эффекта», и занести это число в правую нижнюю ячейку таблицы.

Подсчитать сумму по двум нижним ячейкам. Она должна совпадать с количеством испытуемых во второй группе (выборке).

8. Определить процентные доли испытуемых, у которых «есть эффект», путем отнесения их количества к общему количеству испытуемых в данной группе (выборке). Записать полученных процентные доли соответственно в левой верхней и левой нижней ячейках таблицы в скобках, чтобы не перепутать их с абсолютными значениями.

9. Проверить, не равняется ли одна из сопоставляемых процентных долей нулю. Если это так, отказаться от критерия φ^* и использовать критерий χ^2 .

10. Определить по табл.10 величины углов φ для каждой из сопоставленных процентных долей.

11. Подсчитать эмпирическое значение φ^* по формуле:

12. где: φ_1 – угол, соответствующий большей процентной доле;
 φ_2 - угол, соответствующий меньшей процентной доле;
 n_1 – количество наблюдений в выборке 1;
 n_2 – количество наблюдений в выборке 2.

$$\varphi^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

13. Сопоставить полученные значения φ^* с критическими значениями:

$\varphi^* \leq 1,64$ ($\rho \leq 0,05$) и $\varphi^* \leq 2,31$ ($\rho \leq 0,01$).

Если $\varphi^*_{\text{эмп}} \geq \varphi^*_{\text{кр}}$, H_0 отвергается.

При необходимости определить точный уровень значимости полученного $\varphi^*_{\text{эмп}}$ по табл. 9.

Задачи для самостоятельного решения.

1. Для определения рН использовались 2 типа электродов.

Тип электрода	Показания рН			
		5,78	5,74	5,84
2	5,82	5,87	5,96	5,89

Применяя **t**-критерий Стьюдента определить, следует ли отбросить нулевую гипотезу?

2. Изучено общее содержание азота в плазме крови крыс-альбиносов в возрасте 37 и 180 дней. Результаты выражены в граммах на 100см³ плазмы.

Возраст	1	2	3	4	5	6	7	8	9
37	0,98	0,83	0,99	0,86	0,9	0,81	0,94	0,92	0,87
180	1,2	1,18	1,33	1,21	1,2	1,07	1,13	1,12	1,3

Применяя **t**-критерий Стьюдента определить достоверность влияния возраста на содержание азота в плазме крови крыс-альбиносов.

3. У 12 работающих на ультразвуковых установках изучалось содержание сахара в крови натощак до работы и через 3 часа после работы. Определить достоверность влияния ультразвуковых установок на снижение сахара в крови, используя **t**-критерий Стьюдента.

Натощак	98	82	99	72	79	82	64	70	88	66	88	81
После 3-х час. раб.	54	67	96	59	79	76	66	66	48	61	61	50

4. Изучалось влияние на величину веса щитовидной железы белых крыс раздражения животных во время кормления слабым электрическим током. Получены следующие данные о весе (в мг) щитовидной железы указанных животных и животных контрольной группы, не подвергавшихся раздражению:

Опытная	16	21	16	16	35	24	23	23	16
Контрольная	19	10	12	13	9	8	15	13	12

Используя **t**-критерий Стьюдента, определить, являются ли различия в весе щитовидной железы животных сравниваемых групп статистически значимыми.

5. На двух группах крыс поставлен опыт по сравнению влияния разных рационов на рост. Крысы первой группы получали рацион с высоким содержанием белка, крысы второй - с низким. Привесы за 56 дней опыта для каждой крысы составили в (г):

Высокобелк. рацион	134	146	104	119	124	161	107
Низкобелк. рацион	70	118	101	85	107	132	94

Пользуясь t -критерием Стьюдента определить достоверность влияния высокобелкового рациона на рост крыс.

6. На двух группах лабораторных мышей опытной и контрольной изучалось влияние на организм нового препарата. После месячных испытаний масса тела животных (г) варьировала следующим образом:

Опыт	80	76	75	64	70	72	68	79	83
Контроль	70	78	60	80	62	68	73	60	71

Используя t -критерий Стьюдента определить достоверность влияния на организм нового препарата.

7. Пользуясь t -критерием Стьюдента, определить достоверно ли изменение содержания Na в сыворотке крови кроликов с атеросклерозом на 10-й день после перевязки коронарной артерии и 9 дневного введения нероболила.

До опыта	407	420	420	326	379	474	474	499	387	449
После опыта	382	331	360	357	350	439	450	405	382	373

8. Определялось содержание сиаловой кислоты больных инфарктом миокарда, поступивших на лечение в сроки до 3-х (X) дней и позднее 6-ти дней (Y) от начала заболевания:

X	240	235	270	280	185	287	148
Y	314	270	220	226	230	305	278

Определить достоверность влияния сроков заболевания на содержание сиаловой кислоты в крови используя t -критерий Стьюдента.

9. Температура тела разнополых тушканчиков оказалась следующей:

У самцов	37,5	37,9	37,4	37,8	36,8	37,8	37,5
У самок	37,8	38,1	37,0	37,5	37,7	37,8	37,6

Используя t -критерий Стьюдента, определить, отличаются ли самцы и самки по температуре тела.

10. Следующие данные основаны на результатах сравнительного исследования средней концентрации свинца в крови (в мг/100г) группы рабочих аккумуляторного завода, подвергавшихся профессиональному воздействию-(X) и группы рабочих текстильной фабрики не подвергавшихся профессиональному воздействию-(Y).

X	0.082	0.080	0.079	0.069	0.085	0.090	0.086
Y	0.040	0.035	0.036	0.039	0.041	0.046	0.043

Применяя t -критерий Стьюдента, определить, есть ли различие в среднем содержании свинца в сыворотке крови у рабочих двух предприятий.

11. Определение содержания основного фармакологически активного вещества в жидком лекарственном препарате двумя методами – дало следующие результаты:

1	98,2	98,30	98,30	98,40	98,40	98,40	98,50	98,50	98,60
2	98,3	98,40	98,40	98,50	98,50	98,60	98,60	98,70	98,70

С помощью критерия Фишера сравнить оба метода при доверительной вероятности $P=0.95$.

12. Измерения пульса 10 больных, проведенные после некоторой процедуры, и 12 больных контрольной группы дали следующие результаты: для 1-ой группы $\bar{X}=70$ уд/мин, для 2-ой $\bar{Y}=68$ уд/мин; дисперсии соответственно - $\sigma_x^2=9$ и $\sigma_y^2=4$. При уровне значимости $\alpha = 0,05$ с помощью критерия Фишера проверить, значимо ли различие пульса у больных этих групп.

13. Для определения содержания хлора в химическом соединении были применены методы А и В. Результаты даны в %. Применить F-критерий Фишера для сравнения методов А и В.

А	27,5	27,0	27,3	27,6	27,8			
В	27,9	26,5	27,2	26,3	27,0	27,4	27,3	26,8

14. При определении влияния фактора А на потребление кислорода кроликами по одной методике была получена величина дисперсии $\sigma_1^2=11,6$. Вторая методика дала значение $\sigma_2^2=4,3$. Численность первой и второй групп кроликов соответственно равно: $n_1=8$, $n_2=14$. Требуется установить, существенно ли различие этих дисперсий, т. е эквивалентны ли обе методики по точности?

15. Применить критерии Ван-дер-Вардена и Манна-Уитни для оценки значимости различия между % фагоцитировавших лейкоцитов у морских свинок, сенсibilизированных лошадиной сывороткой (X) и в контроле (Y). Сравнить результаты.

X	2	6	8	8	10	14
Y	22	32	36	54		

16. Изучалось влияние на поглотительные способности ретикулоэндотелиальной системы витамина B_{12} . Получены данные:

Опыт.	28	29	33	34	35	36	39	48	50	53	54	57
Контр.	40	48	50	50	51	53	55	59	60	60	62	84

Применить критерии Ван-дер-Вардена и Манна-Уитни для определения достоверности влияния витамина B_{12} на поглотительную способность ретикулоэндотелиальной системы. Сравнить результаты.

17. Изучалось влияние кобальта на массу тела кроликов. Опыт проводился на двух группах животных: опытной и контрольной. Опытные кролики ежедневно получали добавку к рациону в виде водного раствора по 0.06г хлористого кобальта на 1 кг живой массы тела. Проанализировать с

помощью X- критерия Ван-дер- Вардена и Манна-Уитни результаты о влиянии кобальта на величину массы тела кроликов.

Контроль	420	470	490	504	530	560	580	580
Опыт	561	580	621	630	640	680	692	700

18. Применяя критерии Ван-дер-Вардена и Манна-Уитни, определить достоверность влияния токсических свойств винилпропионата на среднее время гибели мышей. Сравнить результаты.

Опытная	22	35	39	41	43	45	46	48	48	69
Контрольная	13	14	17	22	26	27	30	32	40	55

19. На двух группах лабораторных мышей-опытной ($n_1=9$) и контрольной ($n_2=11$) изучали влияние на массу нового препарата. После месячных испытаний масса тела животных, выраженная в граммах, варьировала следующим образом:

Опыт	80	76	75	64	70	68	72	79	83		
Контроль	70	78	60	80	62	68	73	60	71	66	69

Применяя критерии Ван-дер-Вардена и Манна-Уитни оценить эффективность воздействия нового препарата на организм мышей.

20. Получены данные о весе разнополых тушканчиков (X-самцы, Y-самки).

X	186	190	165	182	182	180	173	157	179	164	146	173	144	156	156
Y	162	163	190	188	147	145	157	162	186	175	147	145	145	155	174

Применяя критерии Ван-дер-Вардена и Манна-Уитни определить, отличаются ли самцы от самок по весу. Сравнить результаты.

21. Дан вариационный ряд распределения початков кукурузы по длине (в мм) и теоретически вычисленный ряд в соответствии с нормальным распределением. Применяя критерий хи-квадрат, определить, подчиняется ли длина початков кукурузы нормальному закону.

Длина	100	110	120	130	140	150	160	170
Фактические частоты	17	39	44	60	42	34	29	18
Теоретические частоты	15	29	45	55	54	43	27	14

22. Изучалась поражаемость клеток при облучении ткани животного организма альфа – частицами. Проведено 517 испытаний. Теоретические значения частот вычислены по закону Пуассона. Результаты распределились следующим образом:

Число пораженных клеток	0	1	2	3	4	5	6	7
Частота поражений	112	168	130	68	32	5	1	1
Теоретические частоты	115	173	130	65	24	7	2	1

Применяя критерий **ХИ**-квадрат, определить, подчиняется ли поражаемость клеток при облучении ткани животного организма альфа-частицами закону Пуассона.

23. Урожай фасоли, полученный на делянках крупных f_1 и мелких f_2 семян, распределяется следующим образом:

Масса	125	175	225	275	325	375	425	475	525
f1	1	5	17	45	70	51	10	1	0
f2	1	3	7	22	88	69	7	2	1

Применяя критерий **хи**-квадрат, определить, какой характер различий между частотами этих рядов - случайный или систематический.

24. В таблице приведены эмпирические и вычисленные по нормальному закону частоты распределения длины тела у 267 мужчин.

Эмпирические	12	31	71	82	46	19	6
Теоретические	11,6	34,3	67,8	77,6	51,2	51,2	19,5

Пользуясь критерием **ХИ**-квадрат, выяснить, случайны или достоверны различия между частотами.

25. Даны частоты значений сердечного индекса X (л/мин.м²) по интервалам и ожидаемые частоты, посчитанные в предположении, что сердечный индекс является нормальной величиной. Используя критерий **хи**-квадрат, определить, является ли сердечный индекс нормальной случайной величиной.

Интервал	Наблюдаемая Частота	Ожидаемая Частота
0-0,5	1	7,83
0,5-1,0	9	7,38
1,0-1,5	23	11,2
1,5-2,0	17	14,67
2,0-2,5	13	16,80
2,5-3,0	12	16,46
3,0-3,5	10	14
3,5-4,0	9	10,42
4,0-4,5	9	6,61
4,5-5,0	3	3,81
5,0-5,5	3	3,81

26. В препарате через равные промежутки времени регистрируется число бактерий, попавших в поле зрения микроскопа. Получены следующие данные:

I	0	1	2	3	4	5	6	7
Наблюдаемые	112	168	130	68	32	5	1	1
Ожидаемые	115	173	130	65	24	7	2	2

Используя критерий χ^2 -критерий проверить при уровне значимости 0,05, что число бактерий, попадающих в поле зрения микроскопа в любой момент регистрации, распределено по закону Пуассона.

- 27.** Загрудинные боли наблюдались у 98 больных из 100 обследованных, а через месяц после лечения у 91 больного. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 28.** Загрудинные боли наблюдались у 98 больных из 100 обследованных, а через год после лечения у 65 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 29.** Загрудинные боли наблюдались у 88 больных из 98 обследованных, а через месяц после лечения у 81 больного. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 30.** Одышка наблюдалась у 84 больных из 100 обследованных, а через месяц после лечения у 71 больного. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 31.** Одышка наблюдалась у 84 больных из 98 обследованных, а через год после лечения у 75 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 32.** Перебои в работе сердца наблюдались у 37 больных из 88 обследованных, а через месяц после лечения у 31 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 33.** Перебои в работе сердца наблюдались у 37 больных из 98 обследованных, а через месяц после лечения у 33 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 34.** Перебои в работе сердца наблюдались у 67 больных из 98 обследованных, а через год после лечения у 53 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 35.** Перебои в работе сердца наблюдались у 67 больных из 98 обследованных, а через два года после лечения у 57 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.

-
- 36.** Сердцебиение наблюдалось у 67 больных из 88 обследованных, а через месяц после лечения у 53 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 37.** Сердцебиение наблюдалось у 80 больных из 89 обследованных, а через год после лечения у 53 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 38.** Сердцебиение наблюдалось у 77 больных из 85 обследованных, а через два года после лечения у 53 больных. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 39.** Повышенное давление наблюдалось у 45 больных гипертонией в одной группе из 99 человек, а в другой группе у 23 больных из 47. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 40.** Повышенное давление наблюдалось у всех больных гипертонией в одной группе из 99 человек, а через месяц после лечения у 23 больных из 47. Используя угловой критерий Фишера определить эффективность применяемого препарата.
- 41.** Аллергическая реакция наблюдалась у 53 больных в группе из 100 человек больных астмой, а в другой группе у 49 больных из 70 человек. Используя угловой критерий Фишера определить, есть ли разница между этими группами.
- 42.** Аллергическая реакция наблюдалась у 73 больных в группе из 100 человек больных астмой, а в другой группе у 45 больных из 70 человек. Используя угловой критерий Фишера определить, есть ли разница между этими группами.
- 43.** Аллергическая реакция наблюдалась у 65 больных в группе из 100 человек больных астмой, а в другой группе у 44 больных из 70 человек. Используя угловой критерий Фишера определить, есть ли разница между этими группами.