

# Методы выявления межгрупповых различий

*Курс лекций  
«Методы прикладной статистики  
для социологов»  
(лекция 4)*

Преподаватель: **Цихончик Надежда  
Васильевна**, старший преподаватель  
кафедры философии и социологии ВШ  
СГНиМК САФУ

# План лекции

1. Независимые и зависимые выборки
2. Общий принцип проверки статистических гипотез
3. Статистические критерии



# 1. Зависимые и независимые выборки

✓ Пять типовых ситуаций

# Просто пример

Выборочное исследование показало, что девочки пропустили в среднем 3,9 учебных дня в году, а мальчики 3,6. В исследовании участвовало 16 девочек и 22 мальчика. Стандартные отклонения составили 0,6 и 0,8 соответственно.

Исследователь считает, что среди учеников средней школы девочки чаще чем мальчики прогуливают занятия. Подтверждают ли результаты эксперимента мнение исследователя?

Другими словами, имеются ли основания полагать, что различие в средних двух генеральных совокупностей статистически значимо?

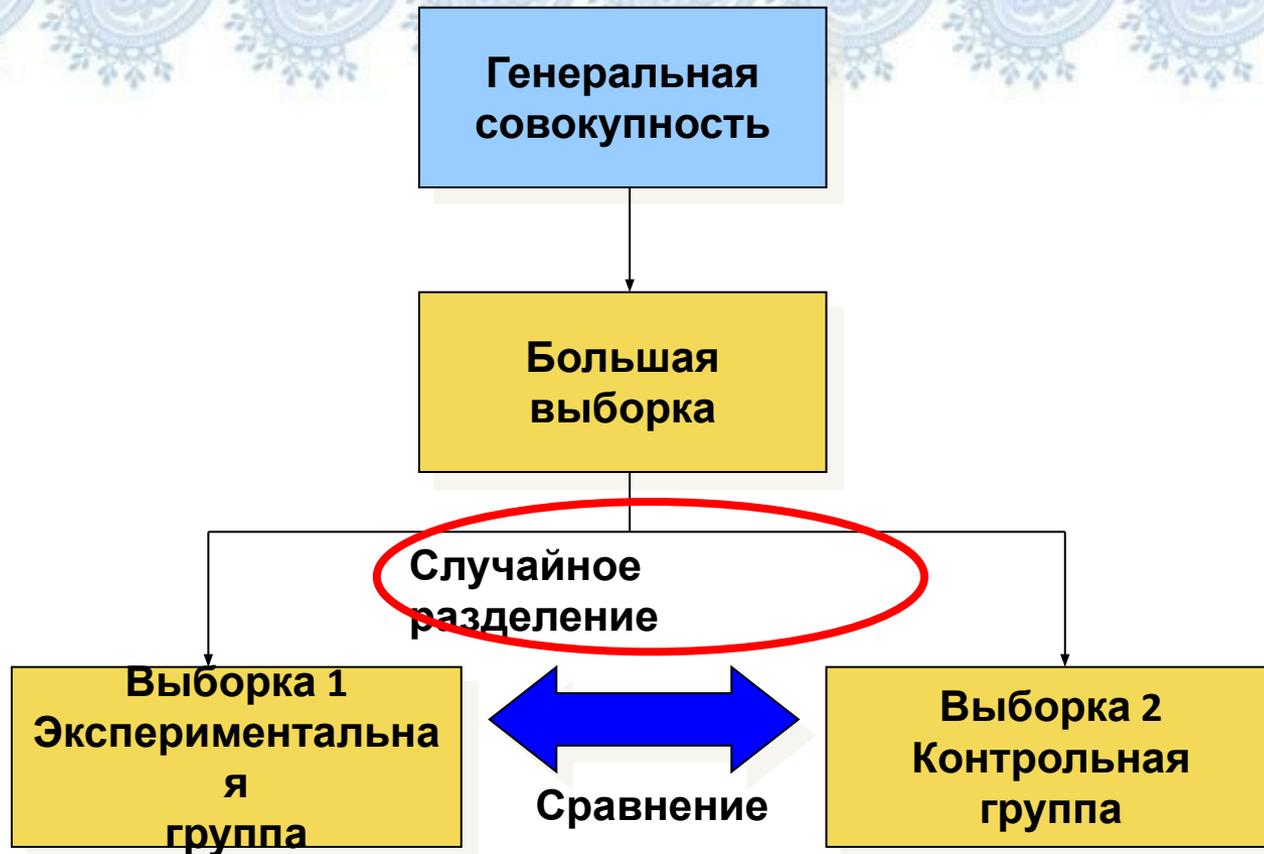
# 1. Независимые выборки

1.1. Две генеральные совокупности, две независимые выборки



# 1. Независимые выборки

1.2. Одна генеральная совокупность, две независимые выборки



# 2. Зависимые выборки

2.1. Две генеральные совокупности, две зависимые выборки



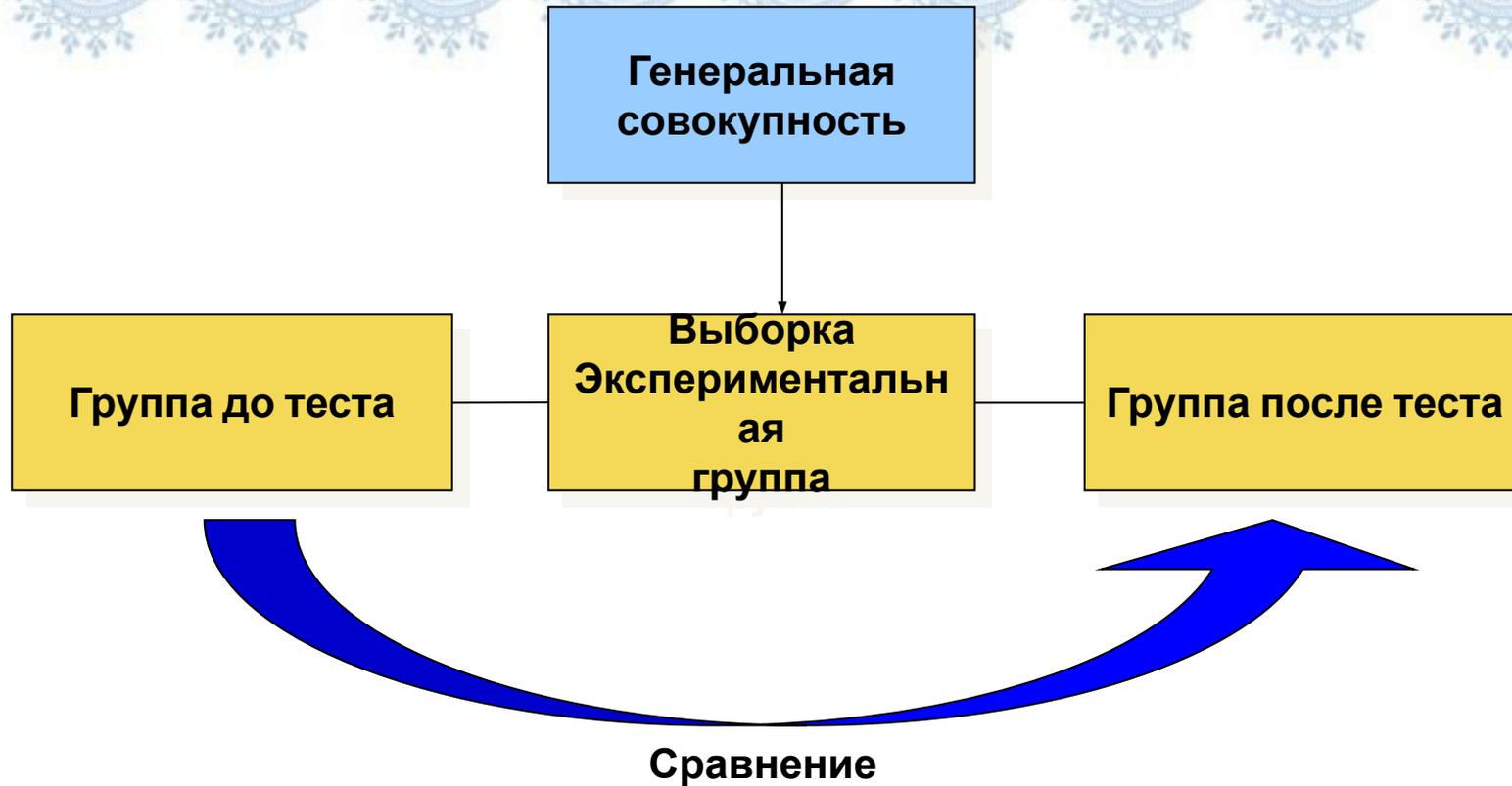
# 2. Зависимые выборки

2.2. Одна генеральная совокупность, две зависимые выборки



## 2. Зависимые выборки

2.3. Одна генеральная совокупность, две зависимые выборки до и после теста



# Сравнительный анализ

- Методы сравнения двух выборок
- Однофакторный ANOVA и непараметрические аналоги
- Многофакторный ANOVA
- Многомерный ANOVA
- Дискриминантный анализ
- ANOVA с повторными измерениями

# Классификация методов сравнения (с. 113)

Количество выборок (градаций $X$ )		Две выборки		Больше двух выборок	
Зависимость выборок		Независимые	Зависимые	Независимые	Зависимые
Признак $Y$	метрический	Параметрические методы сравнения			
		$t$ -Стьюдента, для независимых выборок	$t$ -Стьюдента, для зависимых выборок	ANOVA	ANOVA, с повторными измерениями
	ранговый	Непараметрические методы сравнения			
		$U$ -Манна-Уитни, критерий серий	$T$ -Вилкоксона, критерий знаков	$H$ -Краскала-Уоллеса	$\chi^2$ -Фридмана

Если  $Y$  – метрическая переменная (распределение приблизительно нормальное), то применяются методы сравнения средних.

Если  $Y$  – порядковая переменная (выбросы, асимметрия распределения...), или  $N < 20-25$ , то применяются ранговые методы (критерии) сравнения, предполагающие предварительное ранжирование  $Y$ .



## **2. Общий принцип проверки статистических гипотез**

- ✓ Основная и альтернативная гипотеза
- ✓ Уровень значимости
- ✓ Статистика
- ✓ Критическая область
- ✓ Этапы проверки гипотезы

# Статистическая гипотеза

- **Статистической гипотезой (statistical hypothesis)** мы называем любое предположение о свойствах и характеристиках исследуемых генеральных совокупностей, которое может быть проверено на основе анализа выборок.

# Статистическая гипотеза

- предположение о свойствах случайных величин или событий, которое мы хотим проверить по имеющимся данным
- некоторое предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки
- утверждение в отношении неизвестного параметра, сформулированное на языке математической статистики

# Как проверяют статистические гипотезы

1. Формулируем основную и альтернативную гипотезы

2. Задаем уровень значимости

3. Выбираем статистику - критерий проверки гипотезы

4. Определяем критическую область

5. Вычисляем значение статистики по выборке

6. Сравниваем значение статистики с критической областью, делаем вывод

- Проверка статистических гипотез состоит из шести этапов.
- Нам необходимо изучить каждый сначала теоретически, а затем применять в различных задачах.

# Статистические гипотезы

направленные

ненаправленные

нулевая

об  
отсутствии  
различий  
(опровергнуть)

альтернативная

о значимости  
различий  
(доказать)

нулевая

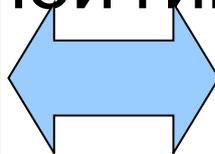
альтернативная

# 1. Основная и альтернативная гипотезы

## ГИПОТЕЗЫ

- Проверяемая гипотеза в статистике называется **основной** (или нулевой) гипотезой. **Основная гипотеза  $H_0$**  подлежит проверке, по результатам которой ее можно принять либо отклонить. «Принять» означает «не получить убедительных аргументов для отклонения гипотезы».
- **Альтернативная гипотеза  $H_1$**  принимается только тогда, когда есть убедительное статистическое доказательство для отклонения основной гипотезы.

Принять  
основную гипотезу  
 $H_0$



Принять  
альтернативную  
гипотезу  $H_1$

# Ситуация А. Новая методика преподавания

- Исследователь хочет проверить, повлияет ли новая методика преподавания на уровень успеваемости студентов. Повысится или понизится успеваемость у студентов, прослушавших курс по новой методике?
- Исследователю известно, что средняя успеваемость без нововведений составляет 4,23 балла. Гипотезы в этом случае будут сформулированы следующим образом:
  - $H_0: \mu = 4,23$
  - $H_1: \mu \neq 4,23$

# Ситуация Б. Аккумуляторы для ноутбуков

- Производители аккумуляторных батарей для ноутбуков утверждают, что разработали принципиально новый тип батареи, которая существенно дольше может работать без подзарядки. Из предыдущих исследований известно, что среднее время работы существующих аккумуляторов составляет 2,5 часа, после чего их требуется заряжать.
- Гипотезы будут сформулированы так:
  - $H_0: \mu \leq 2,5$
  - $H_1: \mu > 2,5$

# Ситуация В. Расходы на канцелярию

- Менеджер бюро переводов хочет снизить расходы компании на канцелярские принадлежности. В среднем эти расходы составляют 5 300 рублей в неделю. После принятия определенных мер по экономии бумаги и скрепок менеджер хотел бы проверить, снизились ли расходы или остались на прежнем уровне.
- Гипотезы будут записаны так:
  - $H_0: \mu \geq 5300$
  - $H_1: \mu < 5300$

# Ошибки первого и второго рода

- **Ошибка первого рода (type I error)** происходит, если мы отвергаем верную нулевую гипотезу.
- **Ошибка второго рода (type II error)** происходит, если мы принимаем нулевую гипотезу.

	Основная гипотеза верна	Основная гипотеза неверна
Мы приняли основную гипотезу	Верное решение	Ошибка II рода
Мы отклонили основную гипотезу	Ошибка I рода	Верное решение

## 2. Уровень значимости гипотезы

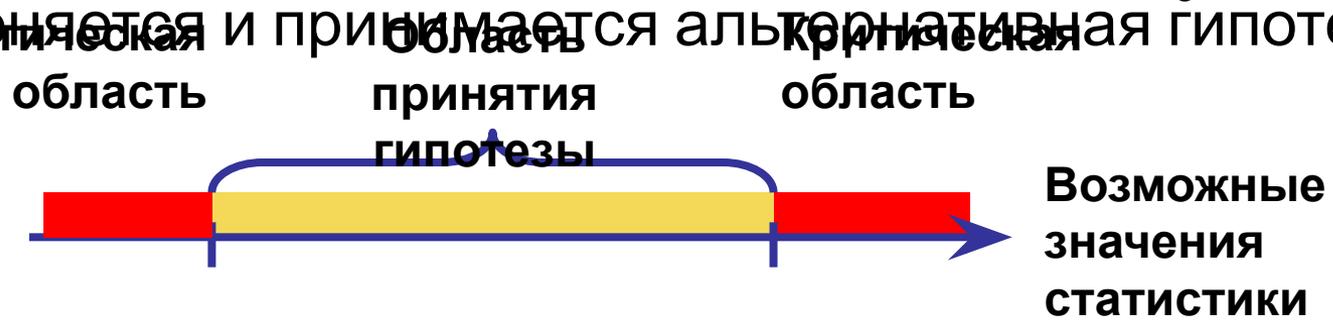
- **Уровнем значимости (level of significance)** гипотезы называют вероятность совершить ошибку первого рода, то есть отклонить верную нулевую гипотезу.
- Обозначение:  $\alpha$ .
- Значение  $\alpha$  обычно выбирается небольшим: 10%, 5% или 1%.

### 3. Статистика - критерий проверки гипотезы

- **Статистика (критерий, statistical test)** есть специальная функция от элементов выборки, по значениям которой принимают решение о принятии или отклонении основной гипотезы.
- Статистика зависит от выборки, поэтому является случайной функцией.

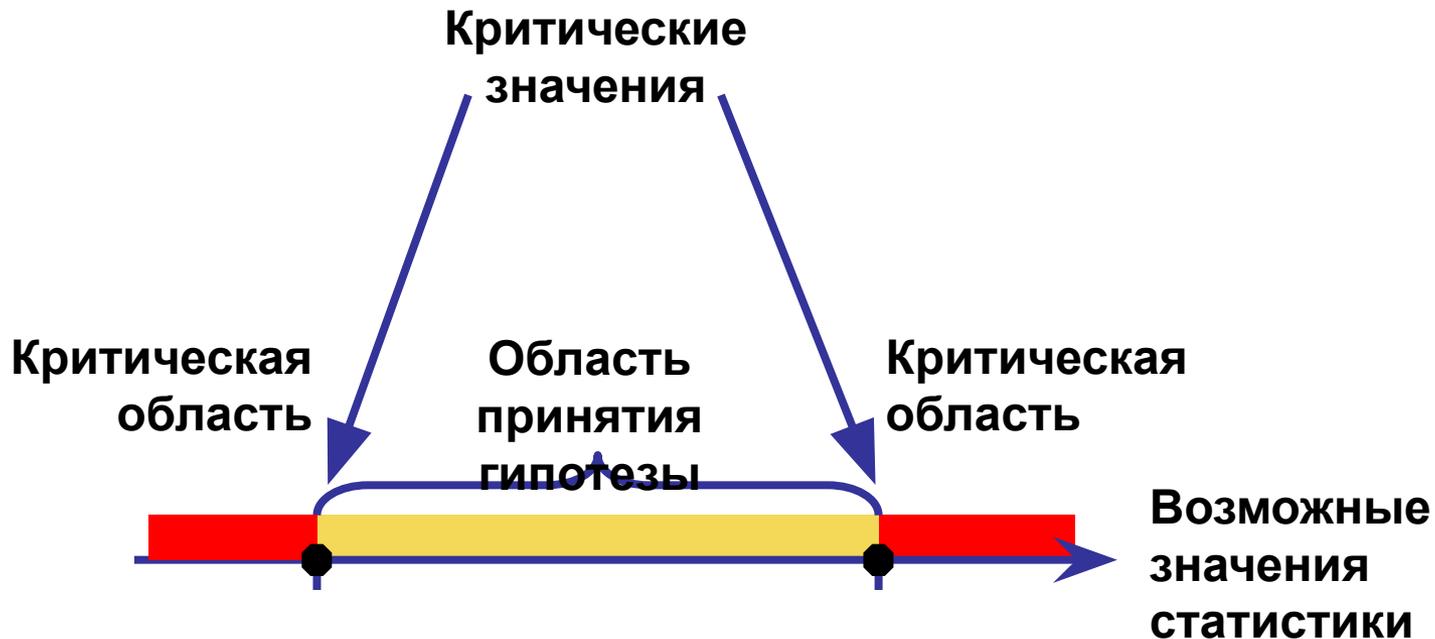
# 4. Критическая область

- Множество значений статистики включает две области:
- **Область принятия гипотезы**, то есть множество тех значений статистики, при которых гипотеза  $H_0$  принимается,
- **Критическую область**, то есть множество тех значений статистики, при которых гипотеза  $H_0$  отклоняется и принимается альтернативная гипотеза.



# Критические значения

- **Критические значения (critical value(s))** отделяют критическую область от области принятия гипотезы.



# Вид критической области

- Критическая область строится, исходя из имеющихся знаний о законе распределения статистики, и зависит от:
- объема выборки,
- уровня значимости, задаваемого исследователем,
- вида альтернативной гипотезы.

Левосторонняя критическая область



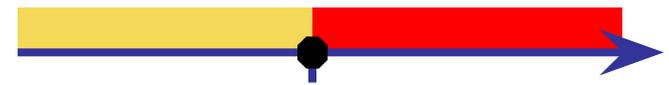
$$H_1 : \mu < \mu_0$$

Двусторонняя критическая область



$$H_1 : \mu \neq \mu_0$$

Правосторонняя критическая область



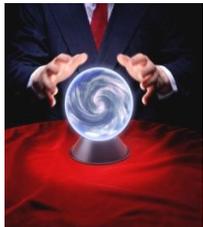
$$H_1 : \mu > \mu_0$$

## 5-6. Вычисление статистики и ВЫВОД

- После построения критической области вычисляют значение статистики по выборке и сравнивают его с критической областью.
- **Если значение статистики попало в область принятия гипотезы, то гипотеза  $H_0$  принимается**
  - **Если значение статистики попало в критическую область,**
  - **то гипотеза  $H_0$  отклоняется и принимается альтернативная гипотеза  $H_1$**

# Последовательность действий

- Шаг 1.** Сформулировать основную и альтернативную гипотезы.
- Шаг 2.** Задать уровень значимости  $\alpha$ .
- Шаг 3.** По таблице найти критические значения и построить критическую область.
- Шаг 4.** По выборке сосчитать значение статистики.
- Шаг 5.** Сравнить полученное значение с критической областью. Если значение попало в критическую область – отклонить основную гипотезу, не попало – принять.
- Шаг 6.** Написать ответ.





### **3. Статистические критерии**

- ✓ Понятие статистического критерия.
- ✓ Виды критериев и статистические задачи

# Статистические критерии

- строгое математическое правило, по которому принимается или отвергается та или иная статистическая гипотеза с известным уровнем значимости
- это МЕТОД расчета определенного числа и само это число
- 1) эмпирическое значение критерия – это определенное число, рассчитанное по данному методу (получаемо в исследовании)
- Виды критериев: значимости, согласия, однородности, сдвига, симметричности и т. Д.

# Статистические критерии

- Критерий (греч. kriterion - средство для решения) - признак, на основании которого производится оценка, определение, классификация чего-н., мерило
- 2) Критическое значение – мера, предел, лимит
- По соотношению эмпирического (1) и критического (2) значений критерия судить о том, подтверждается или опровергается нулевая гипотеза

$$\chi^2_{\text{эмп}} > \chi^2_{\text{кр}} \text{ M}$$

# Статистические критерии

1. Параметрические критерии – критерии, включающие в формулу расчета параметры распределения, т.е. средние и дисперсии
2. Непараметрические критерии – критерии, не включающие в формулу расчета параметров распределения и основанные на оперировании частотами или рангами

## Возможности и ограничения параметрических и непараметрических критериев

ПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ	НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ
1. Оценивают различия в средних, полученных в двух выборках (t-критерий Стьюдента)	Оценивают лишь средние тенденции (чаще ли в выборке А встречаются более высокие, а в выборке Б – более низкие значения признака (критерии Розенбаума (Q), Манна Уитни (U), угловое преобразование Фишера ( $\varphi$ ) и др.)
2. Позволяет прямо оценить различия в дисперсиях (критерий Фишера)	Позволяют оценить лишь различия в диапазонах вариативности признака (критерий $\varphi^*$ )
3. Позволяют выявить тенденции изменения признака лишь при условии - нормального распределения признака - значения признака измерены по интервальной шкале	-при любом распределении признака; -значения признака могут быть представлены в любой шкале
4. Математические расчеты довольно сложны	Математические расчеты более просты и занимают мало времени

# Четыре преимущества непараметрических методов

1. Они могут использоваться для проверки гипотез о параметрах генеральной совокупности, когда переменная **не распределена нормально.**
2. Они могут использоваться для **номинальных и порядковых** данных.
3. Они могут использоваться для проверки гипотез, которые **не связаны с параметрами** генеральной совокупности.
4. В большинстве случаев для непараметрических методов **подсчеты проще,** чем для параметрических. Они более **понятны.**

# Три недостатка непараметрических методов

1. **Они менее точны**, чем соответствующие параметрические методы. Следовательно, требуются более значительные отклонения, чтобы отвергнуть нулевую гипотезу.
2. **Они менее информативны**, чем параметрические критерии. Например, критерий знаков позволяет исследователю определить, превосходит значение данных медиану или нет, но не отвечает – на сколько именно.
3. **Они менее эффективны**, чем соответствующие параметрические критерии. Например, непараметрический критерий знаков дает лишь 60% эффективности от того, что можно получить, используя его параметрическое соответствие –  $z$ -критерий. Требуется больший объем выборки, чтобы компенсировать утрату информации: нужна выборка из 100 человек для критерия знаков, в то время, как для аналогичных результатов при использовании  $z$ -критерия достаточно было бы выборки из 60 человек.

# Сравнение эффективности

Приложения	Параметрический тест	Непараметрический тест	Эффективность
Парные выборки	t-тест или z-тест	Критерий знаков	0,63
		Знако-ранговый критерий	0,95
Две независимые выборки	t-тест или z-тест	Критерий Вилкоксона	0,95
Несколько независимых выборок	Дисперсионный анализ (F-тест)	Критерий Краскела-Уоллиса	0,95
Корреляция	Линейная корреляция	Ранговая корреляция	0,91
Проверка случайности	Нет параметрических	Тест Руна	Нет базы сравнения

Эффективность непараметрических тестов оценивалась в сравнении с параметрическими для нормально распределенной генеральной совокупности.

# Классификация задач и методов их решения

задачи	условия	методы
1. Выявление различий в уровне исследуемого признака	А) 2 выборки испытуемых <hr/> Б) 3 и более выборки испытуемых	Q – критерий Розенбаума; U – критерий Манна-Уитни; φ – критерий (угловое преобразование Фишера); <hr/> S – критерий тенденций Джонкира; H – критерий Крускала-Уоллиса.
2. Оценка сдвига значений исследуемого признака	А) 2 замера на одной и той же выборке испытуемых <hr/> Б) 3 и более замеров на одной и той же выборке испытуемых	T – критерий Вилкоксона; G – критерий знаков; φ – критерий (угловое преобразование Фишера). <hr/> χ <sup>2</sup> – критерий Фридмана; L – критерий тенденций Пейджа.

задачи	условия	методы
3. Выявление различий в распределении признака	<p>А) при сопоставлении эмпирического распределения с теоретическим</p> <hr/> <p>Б) при сопоставлении двух эмпирических распределений</p>	<p><math>\chi^2</math> - критерий Пирсона;  <math>\lambda</math> – критерий Колмогорова-Смирнова.</p> <hr/> <p><math>\chi^2</math> - критерий Пирсона;  <math>\lambda</math> – критерий Колмогорова-Смирнова;  <math>\varphi</math> – критерий (угловое преобразование Фишера).</p>
4. Выявление степени согласованности и изменений	<p>А) двух признаков</p> <hr/> <p>Б) двух иерархий или профилей</p>	<p><math>r</math> – коэффициент ранговой корреляции Спирмена.</p> <hr/> <p><math>r</math> – коэффициент ранговой корреляции Спирмена.</p>
5. Анализ изменений признака под влиянием контролируемых условий	<p>А) под влиянием одного фактора</p> <hr/> <p>Б) под влиянием двух факторов одновременно</p>	<p><math>S</math> – критерий тенденций Джонкира;  <math>L</math> – критерий тенденций Пейджа;  Однофакторный дисперсионный анализ Фишера</p> <hr/> <p>Двухфакторный дисперсионный</p>

## Для выявления различий в уровне исследуемого признака

- по полу;
- по возрасту;
- по национальности;
- и т.д.

Критерий Розенбаума (Q),

Критерий Манна-Уитни (U),

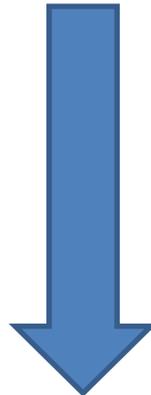
Критерий тенденций Джонкира (S),

Критерий Крускала-Уоллиса (H),

Угловое преобразование Фишера ( $\phi$ ).

## Выбор критерия достоверности различий между независимыми выборками по уровню признака

- 2 выборки: объем выборки больше 11, но не больше 60!- Критерий Розенбаума (Q), критерий Манна-Уитни (U)



- Если различия не выявляются, то используете – **угловое преобразование Фишера**

# U – критерий Манна-Уитни

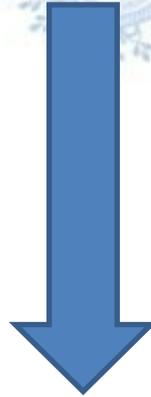
- Критерий предназначен для оценки различий между *двумя выборками* по *уровню* какого-либо признака, количественно измеренного. Он позволяет выявлять различия между *малыми* выборками, когда  $n_1$  и  $n_2$  больше или равны 3 (либо  $n_1 = 2$ , а  $n_2$  тогда больше или равно 5)

-

# Что проверяет критерий Манна-Уитни

- Критерий Манна-Уитни проверяет гипотезу об однородности для двух независимых выборок: совпадают ли законы распределения генеральных совокупностей, из которых взяты эти выборки.
- Гипотезы формулируются следующим образом:
- $H_0$ : выборки взяты из одной генеральной совокупности
- $H_1$ : выборки взяты из разных генеральных совокупностей

- 3 и более выборок: критерий тенденций Джонкира (S), критерий Крускала-Уоллиса (H)



- Если различия не выявляются, то используете – **угловое преобразование Фишера** *многофункциональный критерий*

# Уровни статистической значимости

- Статистическая значимость результата представляет собой оцененную меру уверенности в его правильности
- величину называют **статистически значимой**, если мала вероятность чисто случайного возникновения её или ещё более крайних величин
- Популярными уровнями значимости ( $p$ -уровень) являются 10%, 5%, 1%, и 0,1%

# Уровни статистической значимости

- Уровень значимости – это вероятность отклонения от нулевой гипотезы, в то время как она верна.
- **Ошибка, состоящая в том, что мы отклонили нулевую гипотезу, в то время как она верна, называется ошибкой 1 рода**
- низшим уровнем статистической значимости принято считать 5%-ный уровень, достаточный – 1%-ный и высшим 0,1 процентный., поэтому в таблицах критических значений обычно приводятся соответствующие значения
- До тех пор, пока уровень значимости не достигнет 0,05, мы еще не имеем право отклонить нулевую гипотезу

# Уровни статистической значимости



# Мощность критерия

- это способность критерия выявлять различия, если они есть
- это вероятность правильно отвергнуть нулевую гипотезу, то есть отвергнуть ее, когда она неверна
- **Ошибка, состоящая в том, что мы приняли нулевую гипотезу, в то время как она неверна, называется ошибкой 2 рода**