

МЕДИЦИНСКАЯ СТАТИСТИКА: ОТ ИССЛЕДОВАТЕЛЬСКОГО ВОПРОСА К СТАТИСТИЧЕСКОМУ ВЫВОДУ

Доцент кафедры хирургии и эндоскопии
ИПК и ПКЗ УО БГМУ
Дыбов Олег Геннадьевич



ЗАЧЕМ НАМ МЕДИЦИНСКАЯ СТАТИСТИКА?

Данные пациента $\neq$ научный вывод

Статистика помогает ответить на вопросы:

Насколько наши данные типичны?

Есть ли различия между группами?

Есть ли связь между показателями?

Насколько велик эффект?

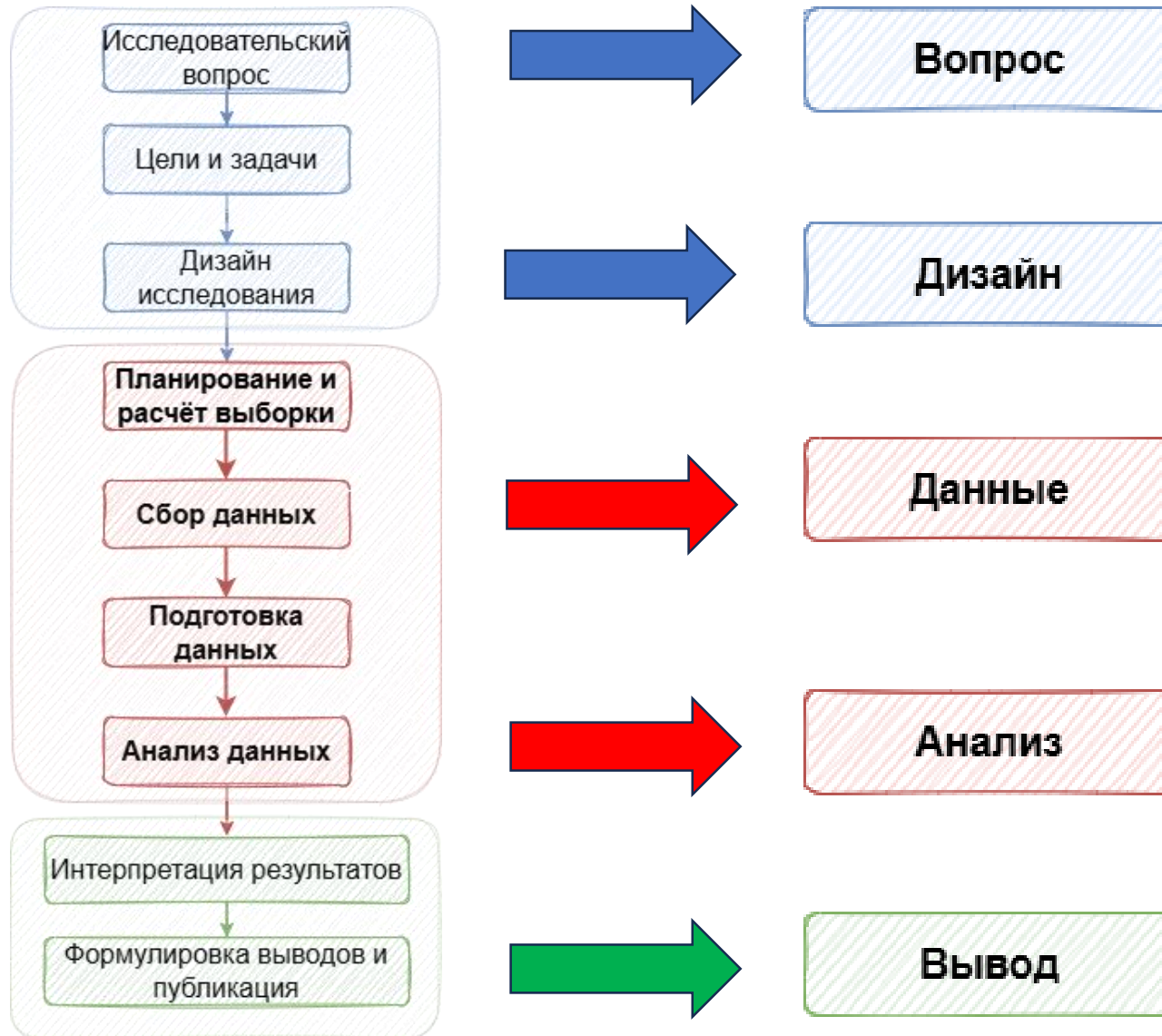
Насколько точна наша оценка?

Можно ли распространить результат на генеральную совокупность?



NB: Статистический анализ — это не поиск $p < 0,05$

Этапы научного исследования



NB: Статистический анализ начинается до сбора данных

ПЛАНИРОВАНИЕ ИССЛЕДОВАНИЯ

До начала сбора данных необходимо определить:

- исследовательский вопрос
- исход/конечную точку
- дизайн
- критерии включения и исключения
- размер выборки
- методы статистического анализа

NB: Нельзя сначала собрать данные, а потом решать, что именно исследовать

Источники информации при планировании



Научная электронная библиотека «КиберЛенинка»



КиберЛенинка — это научная электронная библиотека, построенная на парадигме [открытой науки](#) (Open Science), основными задачами которой является популяризация науки и научной деятельности, общественный контроль качества научных публикаций, развитие междисциплинарных исследований, современного института научной рецензии, повышение цитируемости российской науки и построение [инфраструктуры знаний](#).



ИССЛЕДОВАТЕЛЬСКИЙ ВОПРОС

Вопрос

Пример:

Уменьшает ли новый метод лечения частоту послеоперационных осложнений?

или

Отличается ли длительность госпитализации между двумя методами лечения?

или

Связан ли уровень CRP с длительностью госпитализации?

NB: Формулировка вопроса определяет статистический анализ

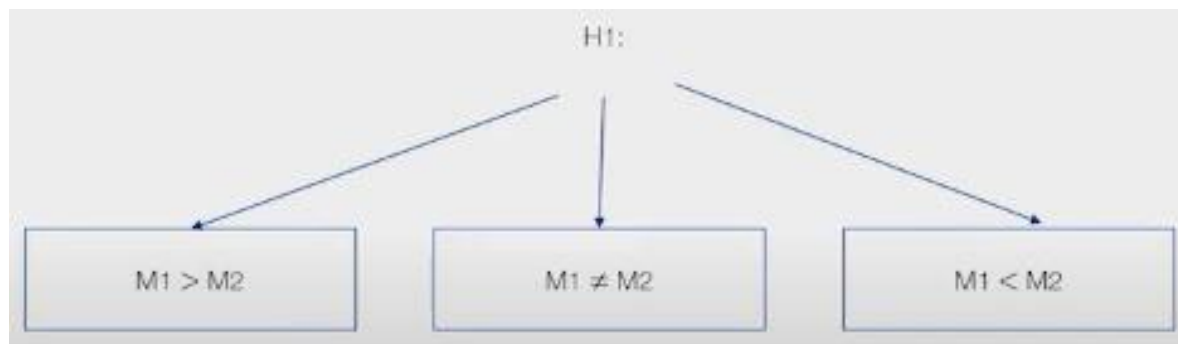
Гипотезы

Вопрос

▪ *Нулевая гипотеза H_0* – предполагается **отсутствие связи/различий/эффекта** между исследуемыми группами (~ условно процессы происходят случайно).

$$H_0: M_1 = M_2$$

▪ *Альтернативная гипотеза H_1* – предполагается **наличие связи/различий/эффекта** между группами (~ условно процессы происходят НЕ случайно).

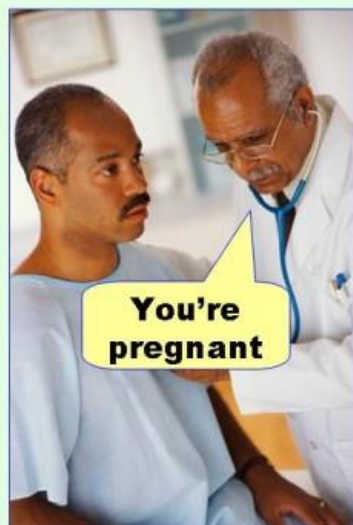


Возможные ошибки при проверке гипотез

Вопрос

		Верная гипотеза	
		H_0	H_1
Результат применения критерия	H_0	H_0 верно принята	H_0 неверно принята (Ошибка второго рода)
	H_1	H_0 неверно отвергнута (Ошибка первого рода)	H_0 верно отвергнута

Type I error
(false positive)



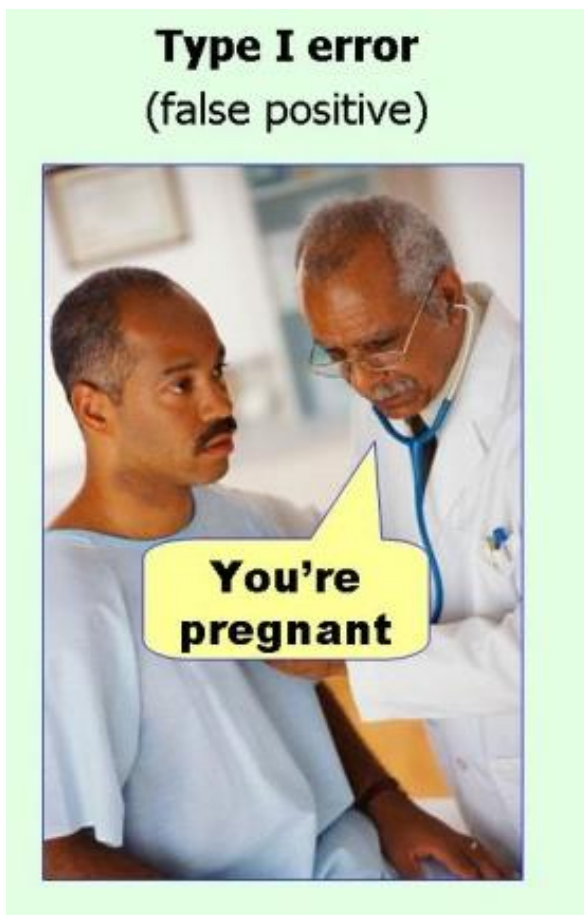
Type II error
(false negative)



Уровень значимости

Вопрос

p (Sig.) (Significance – наблюдаемый *уровень значимости*) – вероятность получить наблюдаемые или более экстремальные данные при условии, что нулевая гипотеза (H_0) верна.

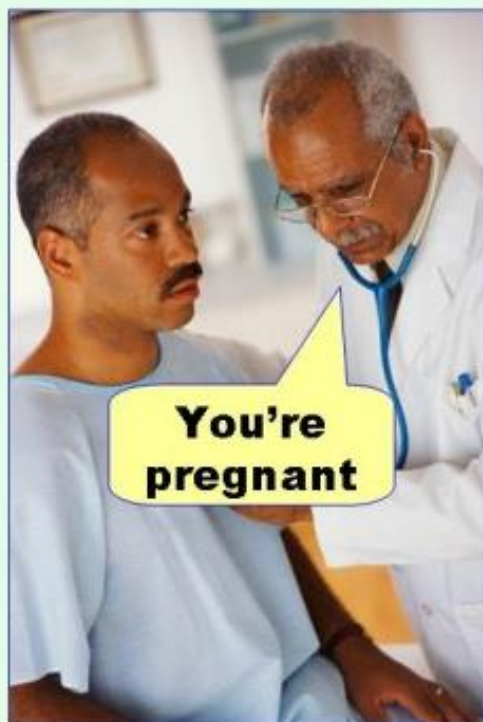


Уровень значимости

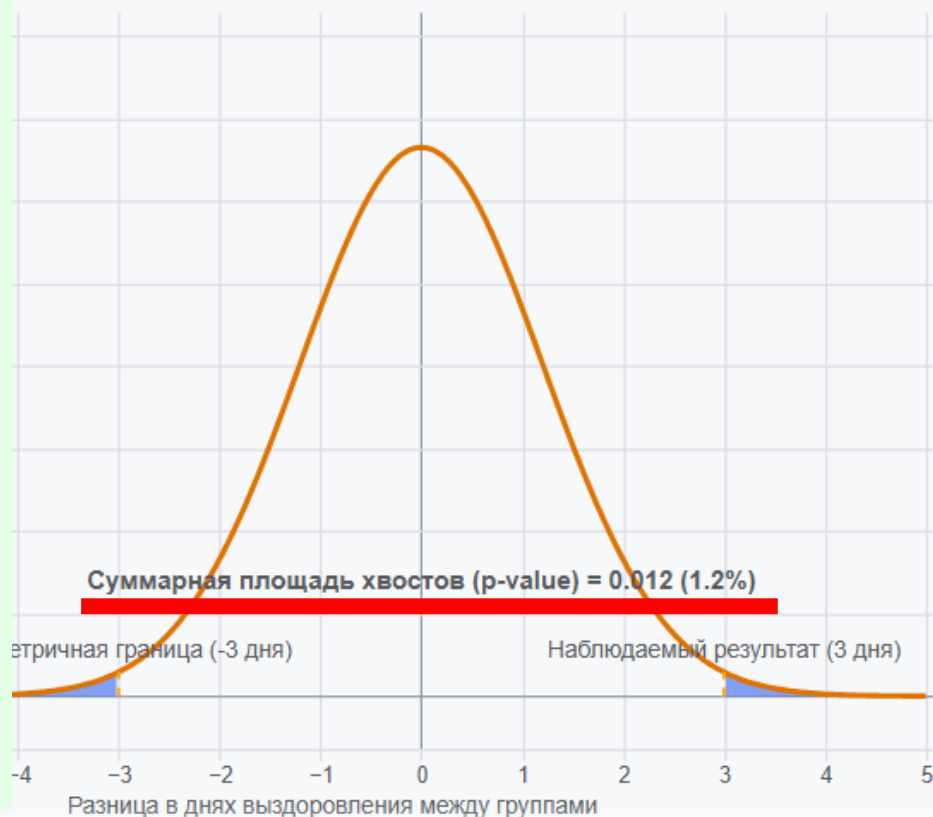
Вопрос

Пример

Type I error
(false positive)



аемые случайные колебания при Н0



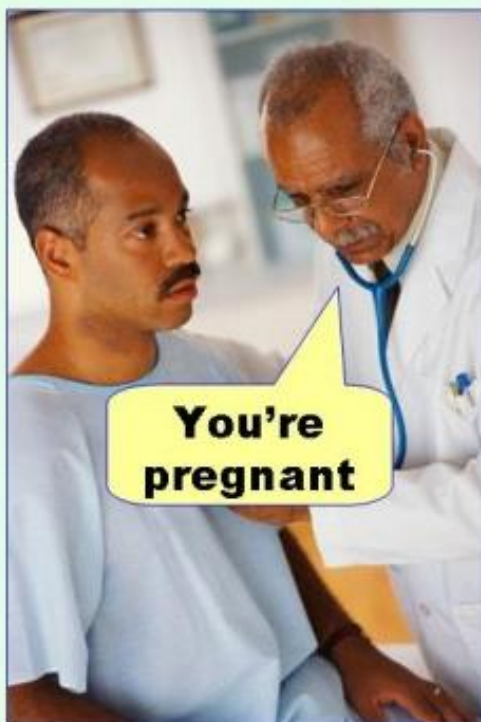
NB: одного p-value недостаточно для вывода о клинической эффективности и безопасности препарата

Уровень значимости

Вопрос

Антипример

Type I error
(false positive)



громный шум при H_0

шанс получить такую разницу в 5 единиц просто из-за случайного стечения обстоятельств составляет **целых 24%**



NB: $p > 0,05 \neq$ эффекта нет

Уровень значимости

Вопрос

Сводная таблица порог-риск

Сфера применения	Стандартный порог α	Допустимый шанс случайной ошибки	Цена ошибки
Физика (ЦЕРН)	0.0000003	1 из 3 500 000	Ложное научное открытие века
Производство	0.0000034	1 из ~300 000	Массовый отзыв техники, поломки
Медицина	0.05	1 из 20	Вред здоровью, неэффективное лечение
Бизнес / IT	0.05 – 0.10	1 из 20 или 1 из 10	Незначительная потеря прибыли
Социология	0.10	1 из 10	Неточный вывод о поведении общества

Уровень значимости

Вопрос

1. Буква α (Альфа) — это ваш внутренний «закон»

Это **порог значимости**, который вы устанавливаете сами **до начала эксперимента**. Это максимальный риск ошибки, который вы готовы себе простить.

- **Кто её задает?** Вы сами (или стандарты вашей индустрии, например, 0.05 в медицине).
- **Смысл:** «Я согласен ошибиться максимум в 5% случаев и признать пустышку работающим лекарством».

2. Буква p (p-value) — это сухой «факт» из данных

Это **наблюдаемый уровень значимости**, который выдает статистическая программа **после того, как вы собрали данные**. Это реальный размер синих углов на графике.

- **Кто её задает?** Компьютер на основе ваших формул и собранных цифр.
- **Смысл:** «В данном конкретном тесте шанс случайности равен 1.2%».

Уровень значимости

Вопрос

- Правильная формулировка №1 (если эффект есть):

«Мы установили уровень значимости $\alpha = 0.05$. В результате теста мы получили $p = 0.012$. Так как $p < \alpha$, результаты статистически значимы».

- Правильная формулировка №2 (краткая, в таблицах):

Рядом с результатами просто пишут: $p < 0.05$ или конкретное число $p = 0.03$. Букву α в таблицы обычно не вставляют, так как все и так понимают, какой порог использовался.

Уровень значимости

Вопрос

Разница между ними исключительно **смысловая** и заключается в моменте времени:

- **Уровень значимости (α)** — это пороговое значение, которое исследователь **задает до начала эксперимента**. Это ваша декларация: *«Я готов принять риск ошибки, но не более чем в 5% случаев»* (при $\alpha = 0,05$).
- **Ошибка первого рода (α)** — это само **событие или вероятность** совершить это событие (ложноположительный результат), когда в реальности верна нулевая гипотеза, но вы её отвергли.

Вопрос

Модель PICO

- **P (Population):** Кто ваши пациенты? (например, *«пациенты с сахарным диабетом 2 типа»*).
- **I/E (Intervention/Exposure):** Какое воздействие или фактор риска вы изучаете? (например, *«прием нового препарата X»* или *«курение»*).
- **C (Comparison):** С чем вы сравниваете? (например, *«прием плацебо»* или *«некурящие»*).
- **O (Outcome):** Какой клинический исход вы ждете? (например, *«снижение уровня гликированного гемоглобина»* или *«развитие инфаркта»*).

ПЛАНИРОВАНИЕ ИССЛЕДОВАНИЯ

До начала сбора данных необходимо определить:

- ~~исследовательский вопрос~~
- исход/конечную точку
- дизайн
- критерии включения и исключения
- размер выборки
- методы статистического анализа

NB: Нельзя сначала собрать данные, а потом решать, что именно исследовать

Тип конечной точки	Статус в исследовании	Что измеряет / Примеры	Главная ценность	Подводный камень (На что смотреть врачу)
Первичная (Primary)	Главный критерий успеха. Всегда заявляется строго одна (редко две) до начала теста.	- Общая выживаемость (в онкологии) - Частота инсультов за 3 года - Выздоровление (Да/Нет)	На основе её p-value регуляторы (FDA, Минздрав) одобряют или запрещают лекарство.	Если по этой точке получен провал ($p > 0.05$), исследование официально считается неудачным , даже если всё остальное супер
Вторичная (Secondary)	Дополнительные эффекты. Помогают раскрыть детали работы препарата, расширить показания.	- Снижение выраженности симптомов - Частота госпитализаций - Качество жизни по опросникам	Показывает сопутствующую пользу, помогает маркетинговым выделить препарат среди конкурентов.	Нельзя использовать как доказательство эффективности, если первичная точка провалена
Суррогатная (Surrogate)	Косвенный маркер («заменитель»). Лабораторный или инструментальный показатель.	- Уровень сахара/холестерина в крови - Артериальное давление в мм рт.ст. - Плотность костной ткани	Позволяет завершить исследование в разы быстрее и дешевле (не нужно ждать исхода годами).	«Улучшение анализов не равно излечению». Лекарство может идеально снижать холестерин, но при этом не снижать смертность.
Комбинированная (Composite)	Объединенная точка. Считается факт наступления <i>любого</i> из списка критических событий.	MACE (Смерть от инфаркта + нефатальный инфаркт + нефатальный инсульт)	Повышает статистическую мощность (вероятность поймать эффект) на небольших группах пациентов.	«Размытие результата». Общий позитивный сдвиг может произойти за счет легкого компонента (меньше госпитализаций) при этом смертность может вырасти.

Сводная таблица клинических конечных точек (Clinical Endpoints)

ПЛАНИРОВАНИЕ ИССЛЕДОВАНИЯ

До начала сбора данных необходимо определить:

- ~~исследовательский вопрос~~
- ~~исход/конечную точку~~
- дизайн
- критерии включения и исключения
- размер выборки
- методы статистического анализа

Классификация

Дизайн



Шаг 1. Сформулируйте гипотезу (Что и у кого вы ищете?)

Четко определите три элемента вашей научной работы (модель **PICO**):

- **P (Population):** Кто ваши пациенты? (например, *«пациенты с сахарным диабетом 2 типа»*).
- **I/E (Intervention/Exposure):** Какое воздействие или фактор риска вы изучаете? (например, *«прием нового препарата X»* или *«курение»*).
- **C (Comparison):** С чем вы сравниваете? (например, *«прием плацебо»* или *«некурящие»*).
- **O (Outcome):** Какой клинический исход вы ждете? (например, *«снижение уровня гликированного гемоглобина»* или *«развитие инфаркта»*).

Шаг 2. Определите характер исследования (Вмешиваетесь ли вы?)

Задайте себе вопрос: «Буду ли я сам назначать лечение/воздействие пациентам?»

- **✗ Нет, я просто наблюдаю.** Пациенты лечатся как обычно, или вредный фактор уже присутствует в их жизни. Вы выбираете **Обсервационное (наблюдательное) исследование.**
- **Да, я сам определяю, кто что получает.** Вы даете препарат, меняете диету, проводите операцию по строгому протоколу. Вы выбираете **Экспериментальное исследование (Интервенционное).**

Шаг 3. Выберите конкретный дизайн (Если исследование обсервационное)

Если вы выбрали путь наблюдения, определитесь со временем и группами:

- **Если исход (болезнь) уже наступил**, а времени и бюджета мало: наберите больных, подберите к ним здоровых («контроль») и изучите их прошлое.
👉 *Ваш дизайн: Исследование «случай-контроль».*
- **Если исхода еще нет**, но есть время и деньги на долгое наблюдение: наберите людей с фактором риска и без него, и отслеживайте их в будущем.
👉 *Ваш дизайн: Проспективное когортное исследование.*
- **Если нужно сделать «моментальный снимок» прямо сейчас**: вы берете группу людей здесь и сейчас и одновременно проверяете и наличие фактора риска, и наличие болезни (например, распространенность ожирения и гипертонии у жителей города в этом месяце).
👉 *Ваш дизайн: Одномоментное (поперечное) исследование (Cross-sectional).*

Шаг 4. Выберите конкретный дизайн (Если исследование экспериментальное)

Если вы сами управляете процессом лечения:

- **Разделяйте пациентов случайно (рандомизация).** Это исключит предвзятость, группы будут одинаковыми по возрасту, полу и тяжести болезней.
 - **Используйте ослепление (blinding), если возможно.** Одностороннее слепое — пациент не знает, что принимает (настоящую таблетку или «пустышку»-плацебо). Двойное слепое — ни пациент, ни лечащий врач не знают этого до конца исследования.
- 👉 **Ваш дизайн: Рандомизированное контролируемое испытание (РКИ).** Это вершина доказательности для оценки эффективности лечения.

Ретроспективное vs Проспективное

Дизайн

Представьте, что ученые хотят изучить связь между **энергетическими напитками** и **аритмией**:

-  **Проспективный подход:** Ученые берут 1000 здоровых студентов сегодня. Делят их на тех, кто пьет энергетики, и тех, кто нет. На протяжении следующих 5 лет они регулярно обследуют их сердца и смотрят, у кого в будущем разовьется аритмия.
-  **Ретроспективный подход:** Ученые приходят в кардиологическую клинику сегодня. Берут 100 пациентов с уже диагностированной аритмией («случаи») и 100 здоровых людей («контроль»). Затем они поднимают их старые медицинские карты или просят вспомнить, сколько энергетиков они пили последние несколько лет.

NB: Ретроспективное / проспективное — это характеристика временной организации исследования, а не отдельный класс, противопоставленный когортному/случай-контроль.

СТАНДАРТЫ

Дизайн



БЕЛОРУССКИЙ
ГОСУДАРСТВЕННЫЙ
МЕДИЦИНСКИЙ
УНИВЕРСИТЕТ

МЕДИЦИНСКИЙ ЖУРНАЛ

Расширенный поиск

ГЛАВНАЯ

О ЖУРНАЛЕ

Главная > № 3 (2026) > Дыбов

Формы публикаций результатов исследований

О. Г. Дыбов, Г. В. Бут-Гусаим, М. Ю. Ревтович, Н. В. Пехтерева, А. Ч. Шулейко

<https://doi.org/10.51922/1818-426X.2026.3.135>

ПОЛНЫЙ ТЕКСТ:

PDF (RUS)



СГЕНЕРИРОВАТЬ
QR КОД

Таблица 1. Формы публикаций и их стандарты
в соответствии с EQUATOR Network

Форма научной публикации	Стандарты
Участие в научных мероприятиях	–
Письмо в редакцию	–
Публикация клинического случая	CARE
Публикация (регистрация) плана и дизайна исследования	SPIRIT, PRISMA-P, STARD, TRIPOD
Публикация систематического обзора (мета-анализ)	PRISMA
Публикация ретроспективного, проспективного наблюдательного или сравнительного исследования	STROBE
Публикация результатов рандомизированного когортного исследования	CONSORT

ПЛАНИРОВАНИЕ ИССЛЕДОВАНИЯ

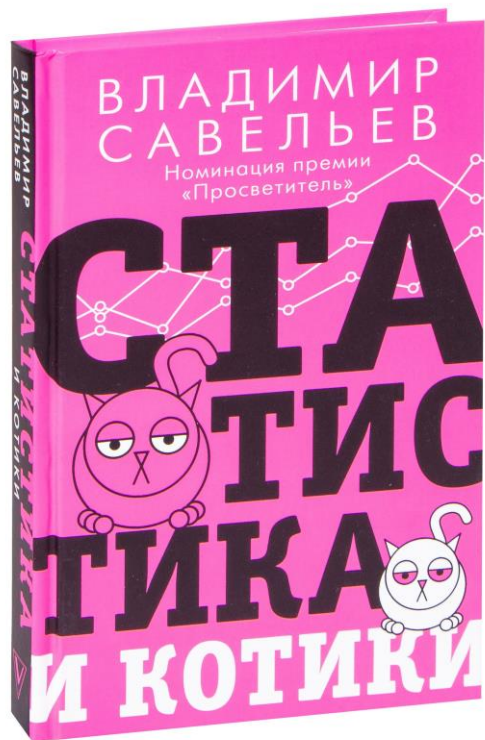
До начала сбора данных необходимо определить:

- ~~исследовательский вопрос~~
- ~~исход/конечную точку~~
- ~~дизайн~~
- **критерии включения и исключения**
- размер выборки
- методы статистического анализа

ПЛАНИРОВАНИЕ ИССЛЕДОВАНИЯ

До начала сбора данных необходимо определить:

- ~~исследовательский вопрос~~
- ~~исход/конечную точку~~
- ~~дизайн~~
- ~~критерии включения и исключения~~
- размер выборки
- методы статистического анализа



ЗАЧЕМ РАССЧИТЫВАТЬ РАЗМЕР ВЫБОРКУ?

Дизайн

Слишком маленькая выборка

- → реальный эффект можно не обнаружить
- → широкие доверительные интервалы
- → высокая вероятность ошибки II рода

Слишком большая выборка

- → исследование требует больше времени и ресурсов
- → статистически значимыми могут стать клинически несущественные различия

Type II error
(false negative)



ОТ ЧЕГО ЗАВИСИТ РАЗМЕР ВЫБОРКИ?

Дизайн

Основные параметры:

- первичная конечная точка
- ожидаемый размер эффекта (какое различие мы хотим обнаружить?)
- вариабельность данных (стандартное отклонение)
- уровень значимости α
- мощность исследования ($1-\beta$, обычно 80%)
- соотношение размеров групп
- ожидаемые потери / выбывание

ОТ ЧЕГО ЗАВИСИТ РАЗМЕР ВЫБОРКИ?

Дизайн

J. Cohen (1988) предложил рассматривать 80% мощности как разумный компромисс между риском ошибки I рода (α) и ошибки II рода (β).

При: $\alpha = 0,05 \rightarrow \beta = 0,20 \rightarrow \text{Power} = 1 - \beta = 0,80$

В исходной концепции Cohen предполагалось, что ошибка I рода может считаться примерно в 4 раза более серьёзной, чем ошибка II рода.

Но 80% — не универсальный стандарт. Требуемая мощность зависит от последствий ошибок и дизайна исследования.



NB: 80% — распространённая отправная точка, а не закон природы

ОТ ЧЕГО ЗАВИСИТ РАЗМЕР ВЫБОРКИ?

Дизайн

Желаемая мощность ($1 - \beta$)	Ошибка II рода (β)	Необходимый размер выборки (на группу)
70%	30%	~50 человек
80% (Стандарт)	20%	~64 человека
90%	10%	~85 человек
95%	5%	~105 человек

КАК ОПРЕДЕЛИТЬ РАЗМЕР ЭФФЕКТА?

Дизайн

Способ 1. Анализ научной литературы

Способ 2. Минимальное клинически значимое различие

Способ 3. Пилотное исследование (Если тема абсолютно новая)

Способ 4. Эмпирические ориентиры Коэна (Если данных нет совсем)

Если у вас нет ни литературы, ни клинических ориентиров, ни пилота, используют стандарты Джейкоба Коэна. Он предложил оценивать эффект в безразмерных единицах (индекс (d) Коэна для средних или (w) для долей) и разделил их на три категории:

Малый эффект ($d = 0.2$): Различия настолько тонкие, что их трудно заметить «на глаз» (например, разница в росте между 15-летними и 16-летними подростками). *Потребуется огромная выборка.*

Средний эффект ($d = 0.5$): Заметное различие, видимое опытному врачу (например, разница в росте между 14-летними и 18-летними). *Потребуется средняя выборка.*

Большой эффект ($d = 0.8$): Очевидное и сильное различие (например, разница в росте между девочками и парнями в старших классах). *Достаточно маленькой выборки.*

ГДЕ РАССЧИТЫВАТЬ ВЫБОРКУ?

Дизайн

Инструмент	Уровень	Для чего
ClinCalc	 Начальный	Быстрый расчёт онлайн
G*Power	 Начальный/средний	Обучение power analysis
MedCalc	  Средний	Медицинские исследования
R (<code>pwr</code> , <code>pwrss</code>)	 Средний/продвинутый	Воспроизводимые расчёты
PASS	 Продвинутый	Сложные клинические исследования

ПЛАНИРОВАНИЕ ИССЛЕДОВАНИЯ

До начала сбора данных необходимо определить:

- ~~исследовательский вопрос~~
- ~~исход/конечную точку~~
- ~~дизайн~~
- ~~критерии включения и исключения~~
- ~~размер выборки~~
- **методы статистического анализа**

Алгоритм анализа данных

1. Сбор
2. Обработка
3. Распределение
4. Описательная статистика
5. Сравнительный анализ
6. Интерпретация
7. Вывод

СБОР ДАННЫХ

Данные

Тестовая форма

* Обязательно

Ваш пол? *

☐ Мужской

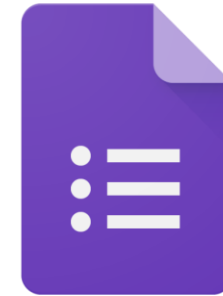
☒ Женский

☐ Другой _____

ОТПРАВИТЬ

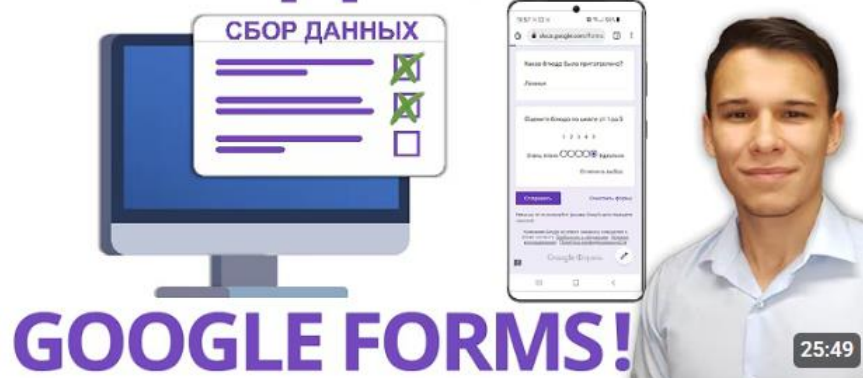
Никогда не используйте формы Google для передачи паролей.

Компания Google не имеет никакого отношения к этому контенту. Сообщение о нарушении - Условия использования - Дополнительные условия



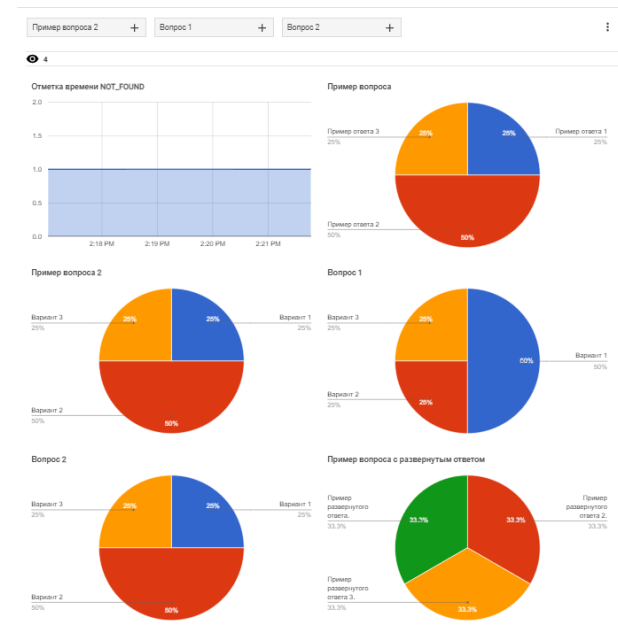
Google Forms

СБОР ДАННЫХ?



GOOGLE FORMS!

Билял Хасенов

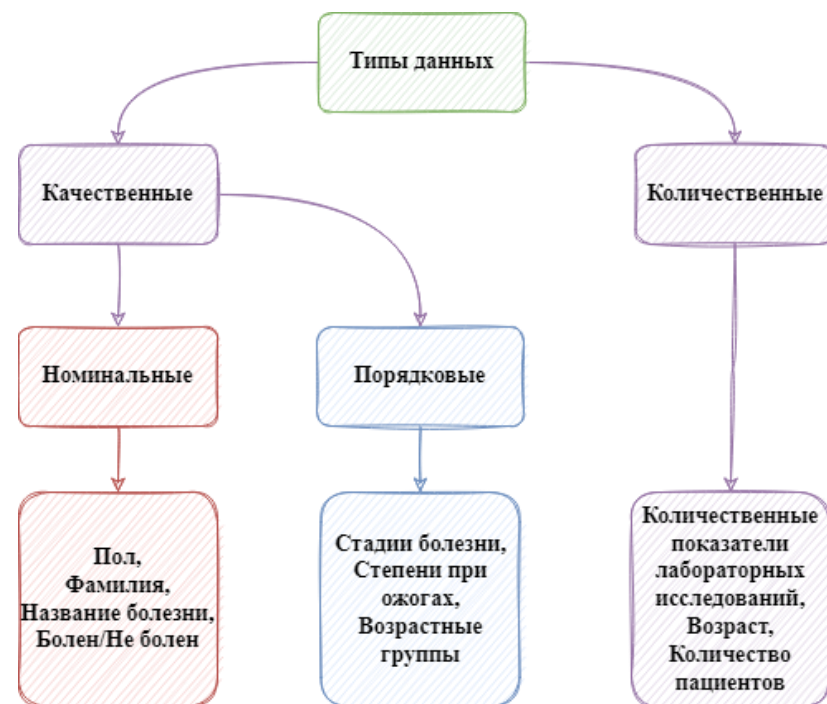


ТИПЫ ДАННЫХ

Данные



О.Ю.Реброва. Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA М., Медиафера, 2006. 312 с. 3-е издание



NB: тип переменной — один из главных факторов выбора статистического метода!!!

ЗАВИСИМЫЕ VS НЕЗАВИСИМЫЕ НАБЛЮДЕНИЯ

Данные

Независимые наблюдения (несвязанные выборки) — это данные, полученные из абсолютно разных, не связанных друг с другом групп. Результат в одной группе никак не влияет на результат в другой.:

Особенности:

- Объекты или люди в группах разные.
- Количество участников (объем) в группах может отличаться.
- Измерения проводятся случайно или раздельно.

Примеры:

- Сравнение уровня теста по математике между школьниками из города А и города Б.
- Измерение давления у мужчин и женщин.

ЗАВИСИМЫЕ VS НЕЗАВИСИМЫЕ НАБЛЮДЕНИЯ

Данные

Зависимые наблюдения (связанные выборки) — это данные, полученные от одних и тех же объектов/людей в разное время, либо от специально подобранных пар (например, близнецов). Каждому результату из первой группы соответствует один конкретный результат из второй.:

Особенности:

- Данные всегда парные, количество элементов в группах одинаково.
- Условия или состав участников связаны.

Примеры:

- Измерение веса пациентов **до** диеты и **после** диеты (одни и те же люди).
- Сравнение результатов теста у одних и тех же студентов до и после лекции.
- Сравнение показателей у братьев-близнецов.

Алгоритм анализа данных

1. ~~Сбор~~

2. Обработка

3. Распределение

4. Описательная статистика

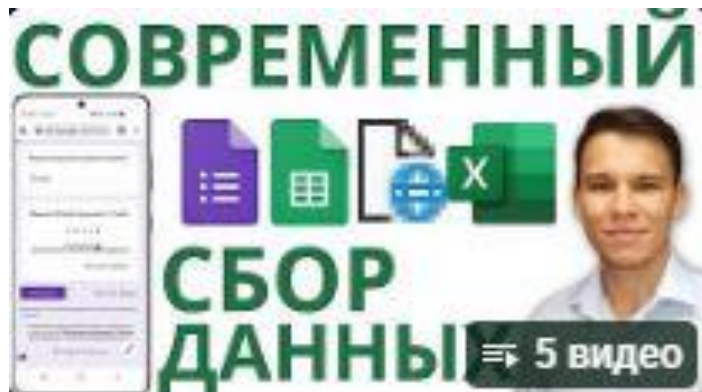
5. Сравнительный анализ

6. Интерпретация

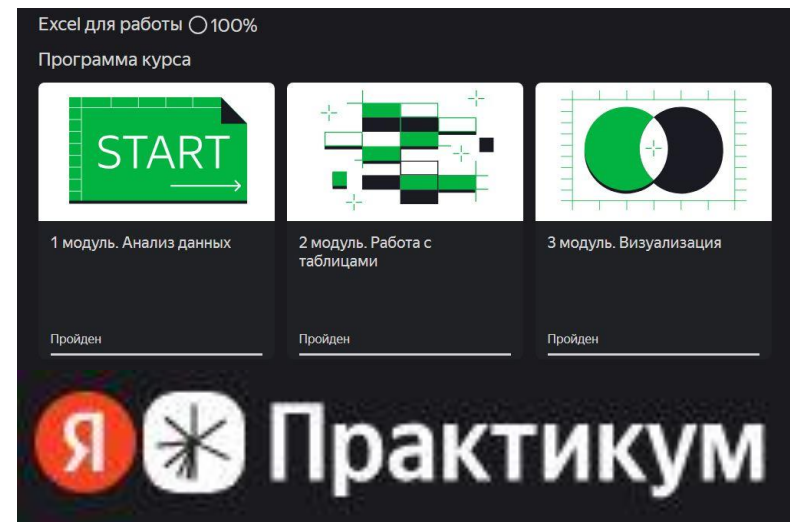
7. Вывод

ПОДГОТОВКА ДАННЫХ

Данные



Билял Хасенов



ПОДГОТОВКА ДАННЫХ

Данные

Заполнение таблицы:

- Удалить дубликаты объектов исследования в таблице.
- Один столбец = один тип данных.
- При возможности перевести текстовые данные в числовые.
- Разбить ячейки с несколькими данными на несколько ячеек.

Перед статистическим анализом:

- Проверить пропущенные значения.
- Проверить ошибки ввода.
- Проверить единицы измерения.
- Проверить кодировку категорий.
- Определить правила обработки пропущенных данных.

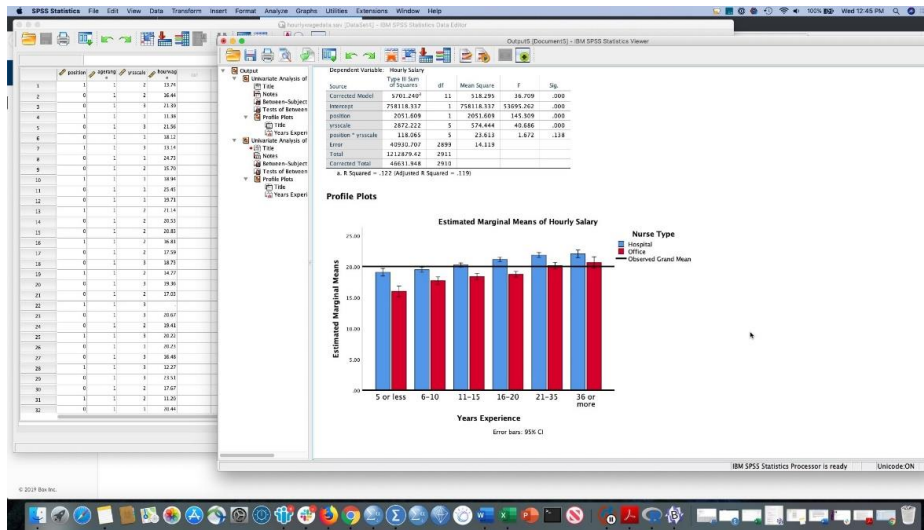
NB: Garbage in → garbage out (GIGO)!

NB: Rubbish in, rubbish out (RIRO)!

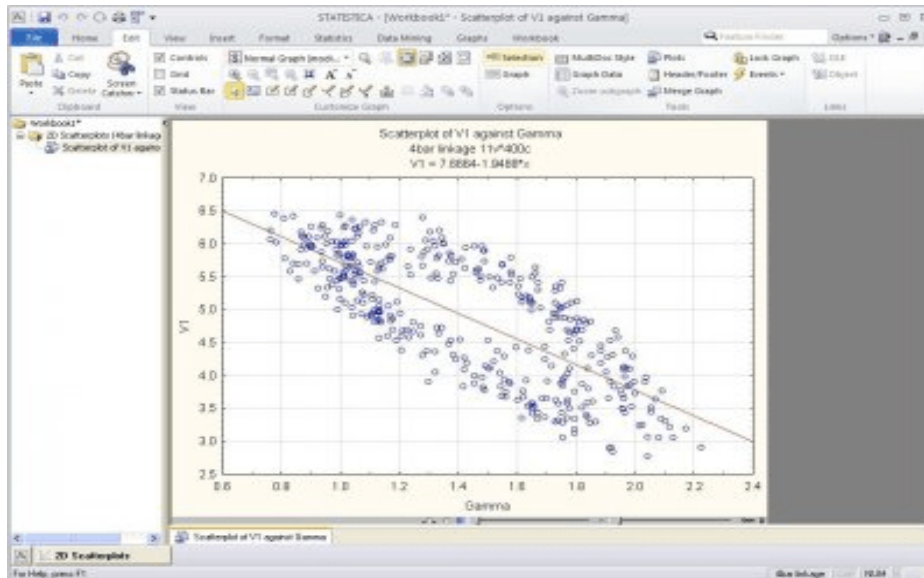
Remove to improve the **data tables** edition

ПЛАТФОРМЫ ДЛЯ АНАЛИЗА ДАННЫХ (КОММЕРЧЕСЕСКИЕ)

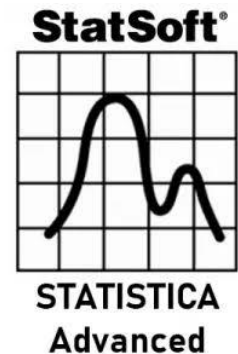
Анализ



IBM SPSS Statistics

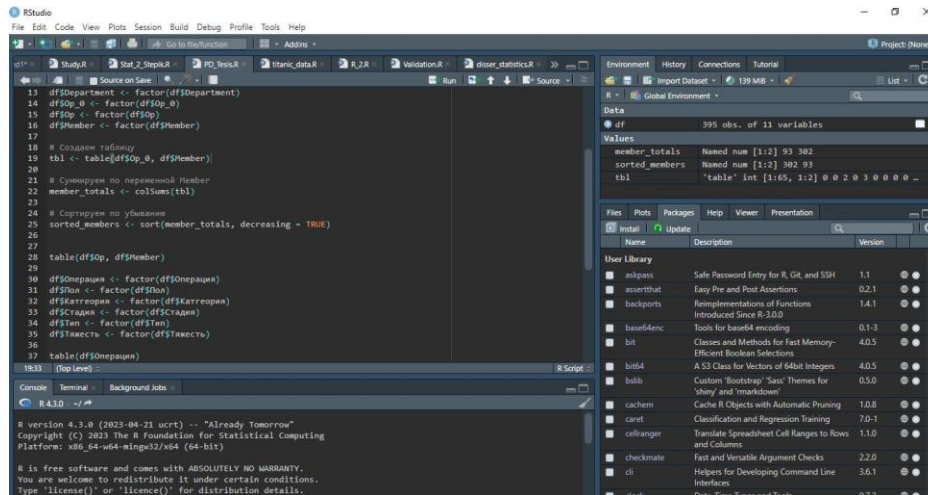


StatSoft STATISTICA



ПЛАТФОРМЫ ДЛЯ АНАЛИЗА ДАННЫХ (БЕСПЛАТНЫЕ)

Анализ



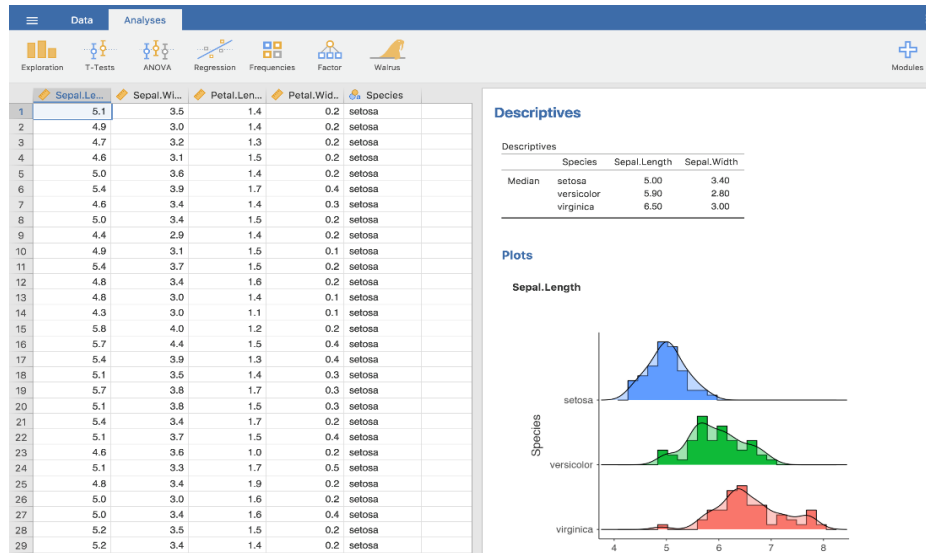
```
13 df$Department <- factor(df$Department)
14 df$Op_0 <- factor(df$Op_0)
15 df$Op <- factor(df$Op)
16 df$Member <- factor(df$Member)
17
18 # Сортируем таблицу
19 tbl <- table(df$Op_0, df$Member)
20
21 # Суммируем по первому столбцу
22 member_totals <- columns(tbl)
23
24 # Сортируем по убыванию
25 sorted_members <- sort(member_totals, decreasing = TRUE)
26
27
28 table(df$Op, df$Member)
29
30 df$Операция <- factor(df$Операция)
31 df$Стан <- factor(df$Стан)
32 df$Категория <- factor(df$Категория)
33 df$Стадия <- factor(df$Стадия)
34 df$Стан <- factor(df$Стан)
35 df$Текста <- factor(df$Текста)
36
37 table(df$Операция)
```

R 4.3.0

version 4.3.0 (2023-04-21 ucrt) -- "Already Tomorrow"
Copyright (C) 2023 The R Foundation for Statistical Computing
platform: x86_64-w64-mingw32/x64 (64-bit)

is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R-studio



jamovi



Алгоритм анализа данных

1. ~~Сбор~~

2. ~~Обработка~~

3. Распределение

4. Описательная статистика

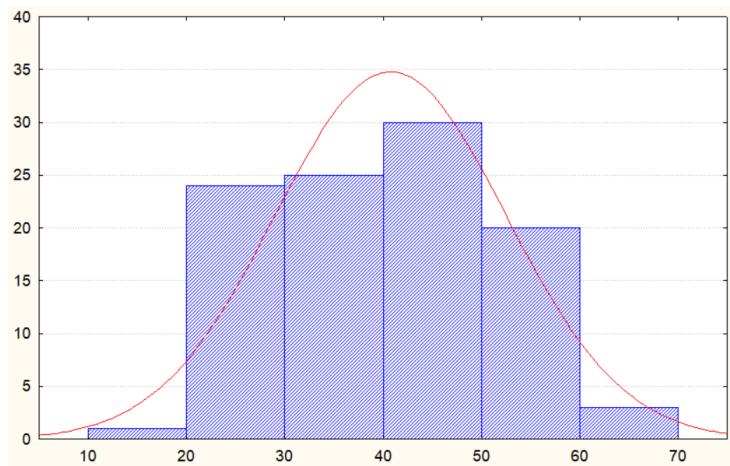
5. Сравнительный анализ

6. Интерпретация

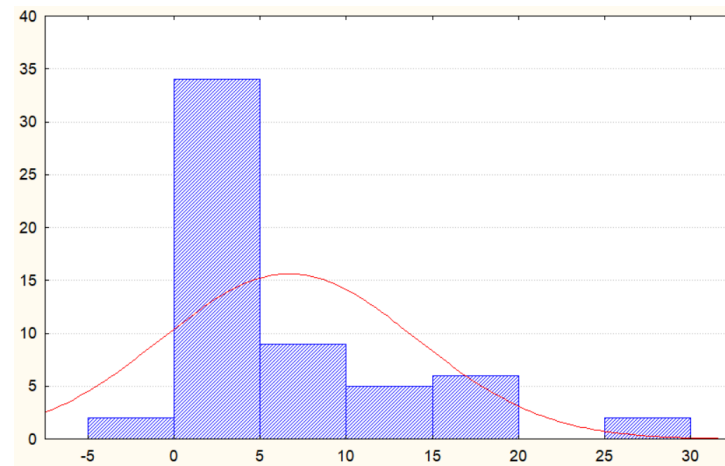
7. Вывод

Распределение

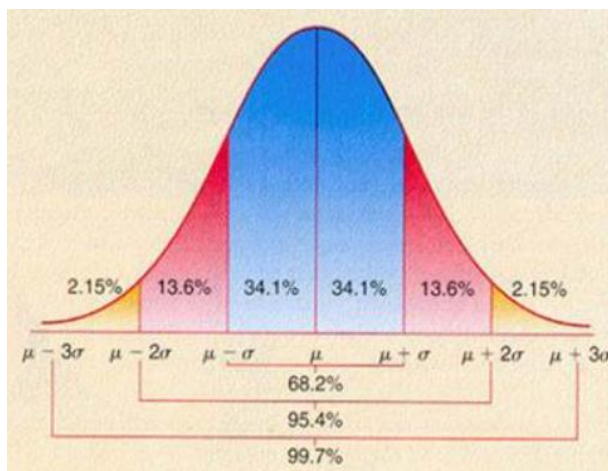
Анализ



Нормальное (Симметричное)

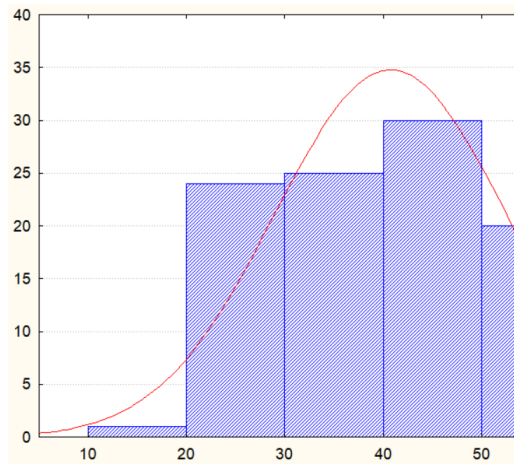


Ненормальное
(Асимметричное)

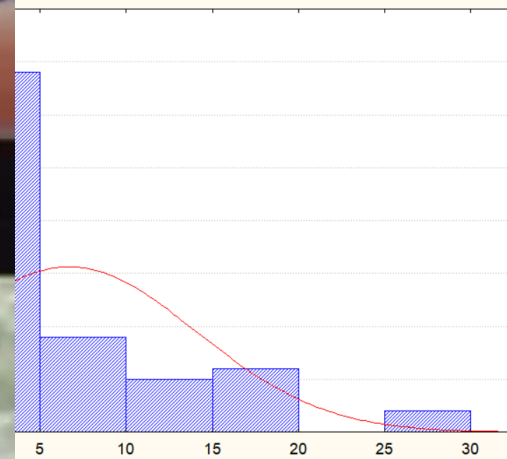


Распределение

Анализ



Нормальное (Симмет



Ненормальное
симметричное)

68.2%

95.4%

99.7%

Распределение (тесты)

Анализ

- ✓ **Колмогорова-Смирнова (Kolmogorov-Smirnov)** Рекомендуется использовать для больших выборок ($\sim$ более 50–100 элементов). На малых выборках этот тест слишком «слаб» и часто ошибочно показывает норму там, где её нет.
- ✓ **Шapiro-Уилка (Shapiro-Wilk's)** Идеален для малых и средних выборок ($\sim$ обычно до 50 измерений, максимум до нескольких сотен). Он обладает самой высокой статистической мощностью при поиске отклонений.

NB: при $p < 0.05$ распределение ненормально

Алгоритм анализа данных

1. ~~Сбор~~

2. ~~Обработка~~

3. ~~Распределение~~

4. **Описательная статистика**

5. Сравнительный анализ

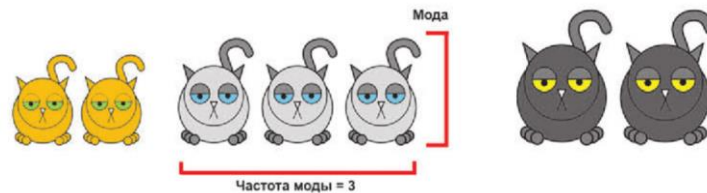
6. Интерпретация

7. Вывод

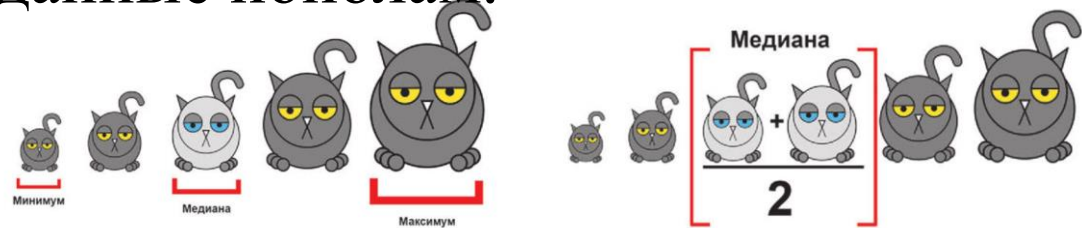
Меры центральной тенденции

Анализ

- **Мода** (Mode - Mo) – наиболее часто встречаемое значение.



- **Медиана** (Median - Me) – значение, разделяющее упорядоченные данные пополам.



- **Среднее арифметическое** (Mean - μ) – сумма всех значений, делённое на число наблюдений



Меры изменчивости

Анализ

- **Отклонение** — разность между конкретным значением и средним.



- **Дисперсия (D)** — среднее значение от квадратов отклонений.



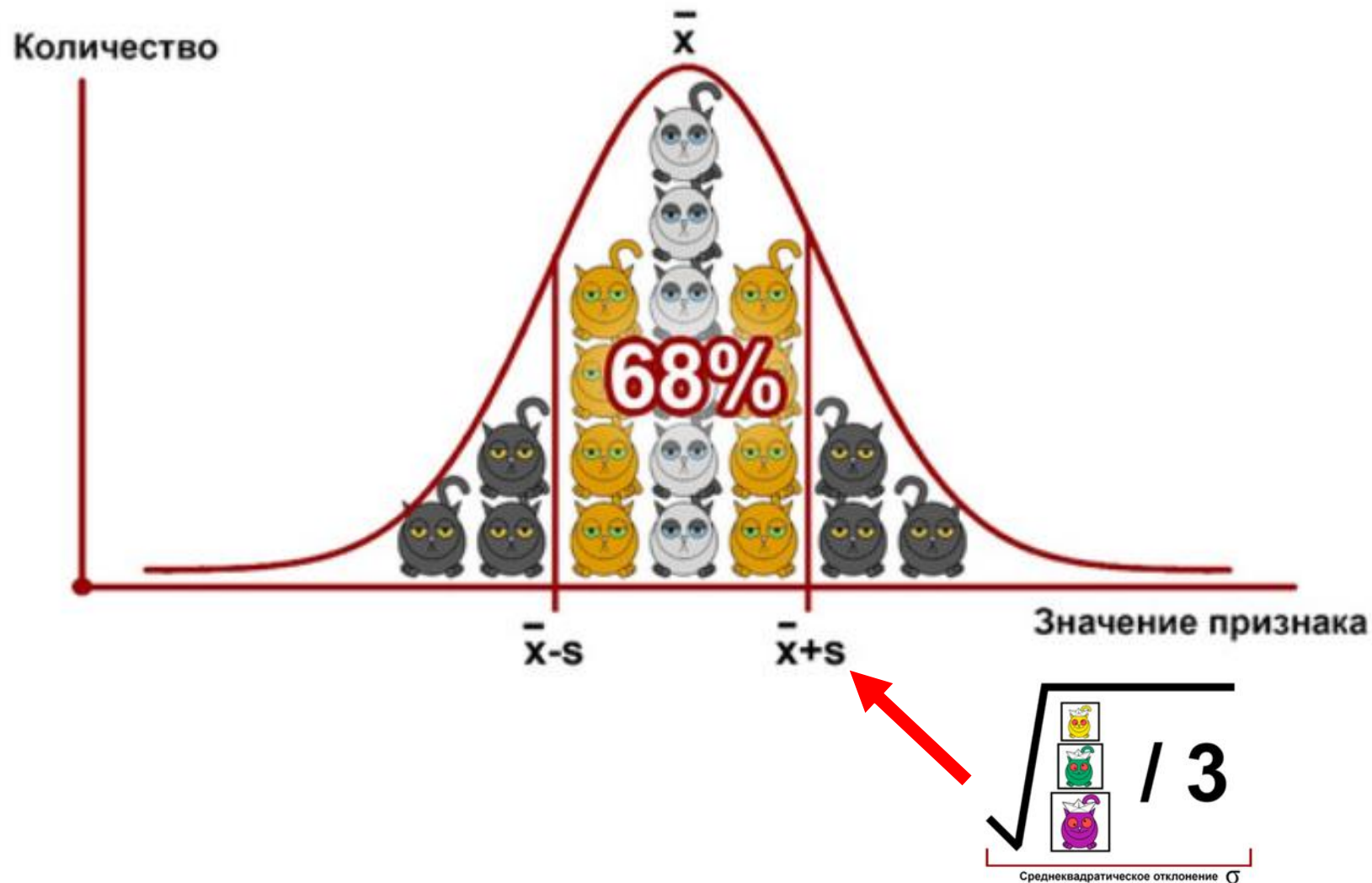
- **Среднеквадратическое отклонение (σ)** — корень из дисперсии



NB: дисперсия и среднеквадратическое отклонение так же неустойчивы к выбросам, как и среднее арифметическое

Нормальное распределение

Анализ



Генеральная совокупность vs выборка



Формы дисперсии

Анализ



Среднеквадратичное отклонение (σ) - корень из дисперсии генеральной совокупности.

Стандартное отклонение (SD) - корень из дисперсии по выборке называется.

Однако не будет большой ошибкой, если вы будете пользоваться терминами *стандартное отклонение генеральной совокупности* и *стандартное отклонение выборки*.

Меры изменчивости

Анализ

- **Стандартная ошибка среднего (SEM)** — показывает на сколько среднее значение выборки отличается от среднего значения генеральной совокупности.

Формула стандартной ошибки среднего выглядит так:

$$SEM = \frac{\sigma}{\sqrt{n}}$$

Где:

- σ (или s) — стандартное отклонение (разброс данных в выборке).
- n — размер выборки (количество наблюдений).

SEM vs SD

Анализ

Простой пример - представьте, что вы измеряете рост 100 человек в большом городе:

- **Стандартное отклонение (SD)** покажет, насколько сильно рост конкретных людей в вашей группе отличается от среднего роста по этой группе (кто-то ниже, кто-то выше).

Вариант записи: средний рост в выборке составил 176,5 см ($sd = 6,2$, $n = 100$).

- **Стандартная ошибка среднего (SEM)** покажет, с какой погрешностью средний рост вашей группы отражает средний рост всех жителей этого города. Если вы возьмете другую группу из 100 человек, их среднее значение будет немного другим — стандартная ошибка как раз оценивает этот разброс средних.

Вариант записи: средний рост в выборке составил $176,5 \pm 0,7$ см ($\mu \pm SEM$, $n = 100$).»

Доверительным интервал среднего (CI, англ. – confidence interval for mean) - это диапазон значений, в котором с заданной вероятностью находится истинное (реальное) среднее значение всей генеральной совокупности.

Из чего состоит доверительный интервал?

Формула расчета строится вокруг выборочного среднего и показателей, которые мы обсуждали ранее:

$$\text{Доверительный интервал} = \bar{X} \pm (\text{Критическое значение} \times \text{SEM})$$

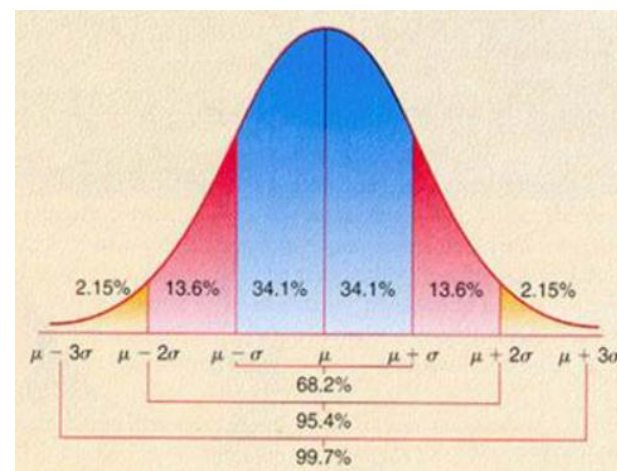
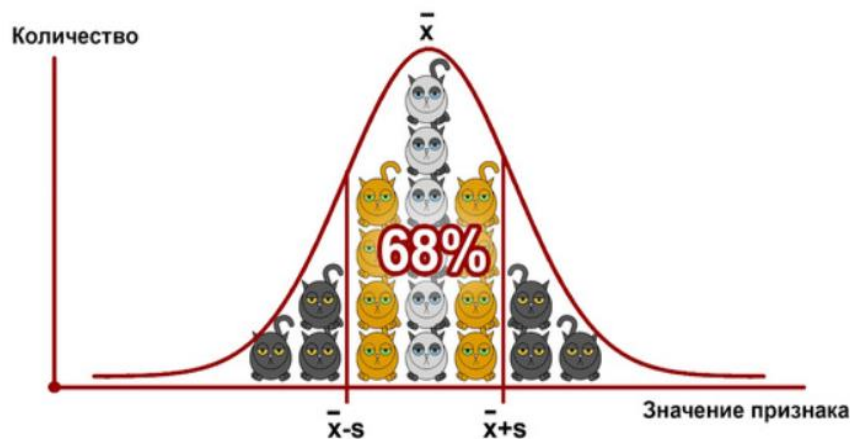
1. **Выборочное среднее ($\bar{X}$)** — центр нашего интервала.
2. **Стандартная ошибка среднего (SEM)** — показывает базовую погрешность выборки.
3. **Уровень доверия (вероятность)** — выбирается исследователем. Чаще всего берут **95%** (иногда 99% или 90%).
4. **Критическое значение (z или t)** — коэффициент, который берется из математических таблиц. Он зависит от выбранного уровня доверия и размера выборки. Для 95% уровня доверия на больших выборках этот коэффициент примерно равен **1.96**.

ДИ vs SEM

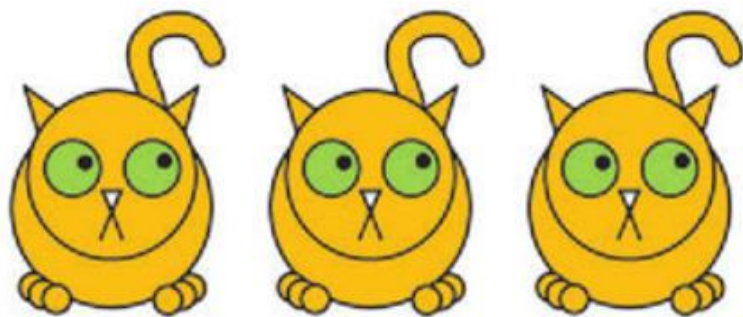
Анализ

ДИ и SEM — показывают точность оценки среднего значения, но SEM — это шаг на пути к расчету, а ДИ — это уже конечный результат, готовый для интерпретации.

Главное отличие: SEM указывает на абстрактную стандартную ошибку в пределах одного стандартного отклонения (охватывает около 68% вероятности), а ДИ (обычно 95%) очерчивает конкретные границы, внутри которых истинное



SEM	ДИ
$\mu \pm \text{SEM}$	μ ; 95% ДИ: ДИ _{min} -ДИ _{max}
рост - 175.0 ± 0.7 см	рост - 175 см; 95% ДИ: 173.6–176.4



Выброс

Меры изменчивости

Анализ



Форма записи: Me [Q_н; Q_в]

Пример: индекс массы тела - 24,2 [19,7; 25,8]

Меры изменчивости



Форма записи: $Me [Q_n; Q_v]$

Пример: индекс массы тела - $24,2 [19,7; 25,8]$

Me vs μ

Анализ

Представьте, что в небольшом отделе работают **5 человек**: четыре обычных сотрудника и один директор. Их ежемесячные зарплаты составляют:

20 000 руб., 25 000 руб., 30 000 руб., 35 000 руб. и 500 000 руб.
(директор).

- **Как считает Среднее значение:**

Мы складываем абсолютно все зарплаты и делим на 5 сотрудников:

$$(20 + 25 + 30 + 35 + 500)/5 = 610/5 = \mathbf{122\ 000\ руб.}$$

Вывод среднего: Средняя зарплата в отделе — 122 тысячи. Но реальность такова, что 4 из 5 сотрудников не получают даже трети этой суммы. Директор со своей огромной зарплатой (выбросом) исказил общую картину.

Me vs μ

Анализ

Как считает Медиана:

Мы выстраиваем зарплаты по порядку (от меньшей к большей) и берем число, которое стоит **ровно посередине**:

20 000, 25 000, **[30 000]**, 35 000, 500 000.

Вывод медианы: Медианная зарплата — **30 000 руб.** Это гораздо честнее описывает реальный доход обычного сотрудника. Ровно половина отдела получает меньше 30 тысяч, а половина — больше. Если директор поднимет себе зарплату до миллиона, медиана все равно останется равной 30 000.

Алгоритм анализа данных

1. ~~Сбор~~

2. ~~Обработка~~

3. ~~Распределение~~

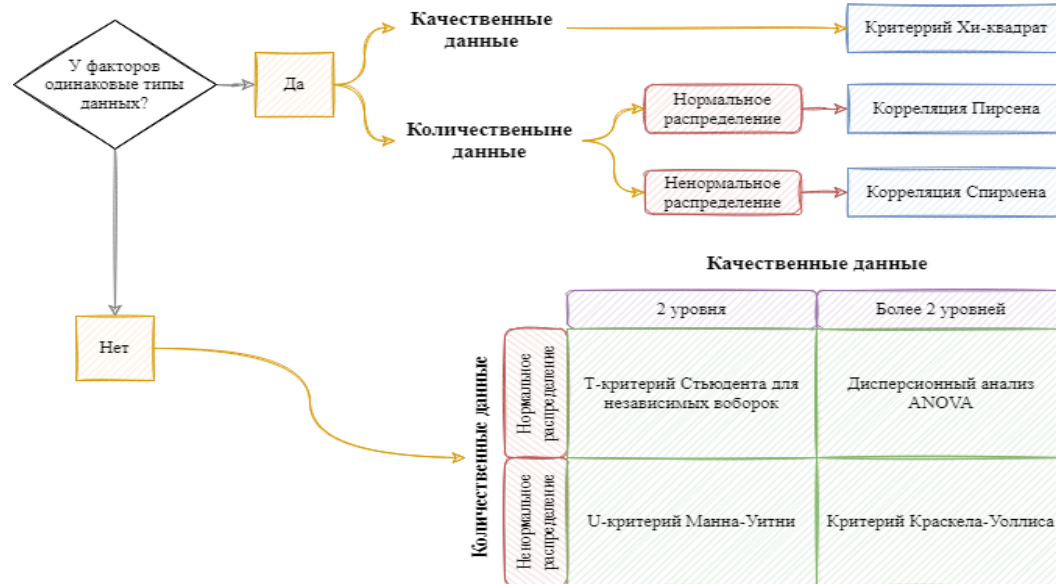
4. ~~Описательная статистика~~

5. Сравнительный анализ

6. Интерпретация

7. Вывод

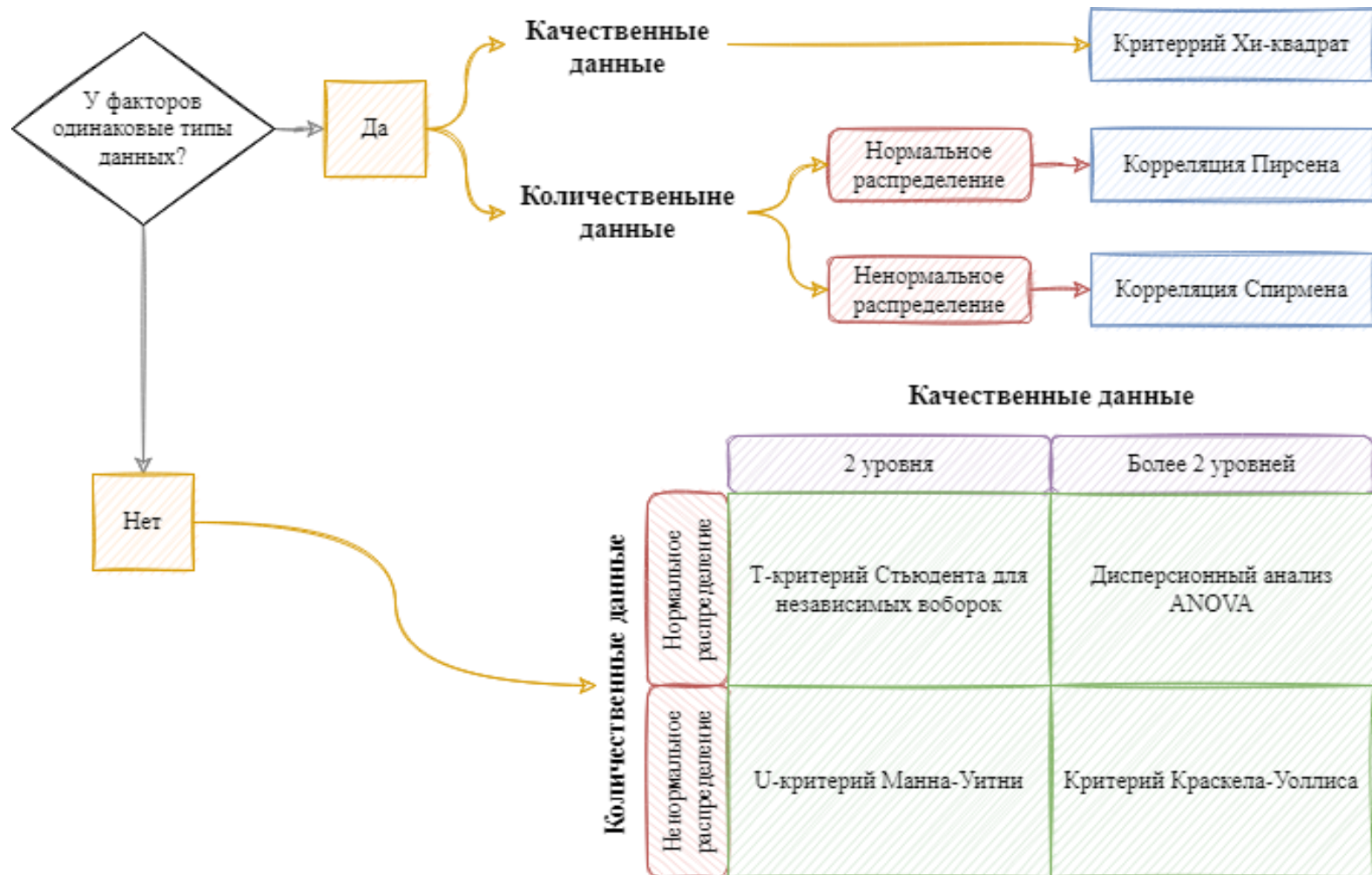
Алгоритм выбора метода анализа данных для независимых выборок



Алгоритм выбора метода анализа данных для зависимых выборок



Алгоритм выбора метода анализа данных для независимых выборок



КОЛИЧЕСТВЕННЫЙ ИСХОД 2 НЕЗАВИСИМЫЕ ГРУППЫ

Вопрос: отличается ли длительность госпитализации между двумя методами лечения?

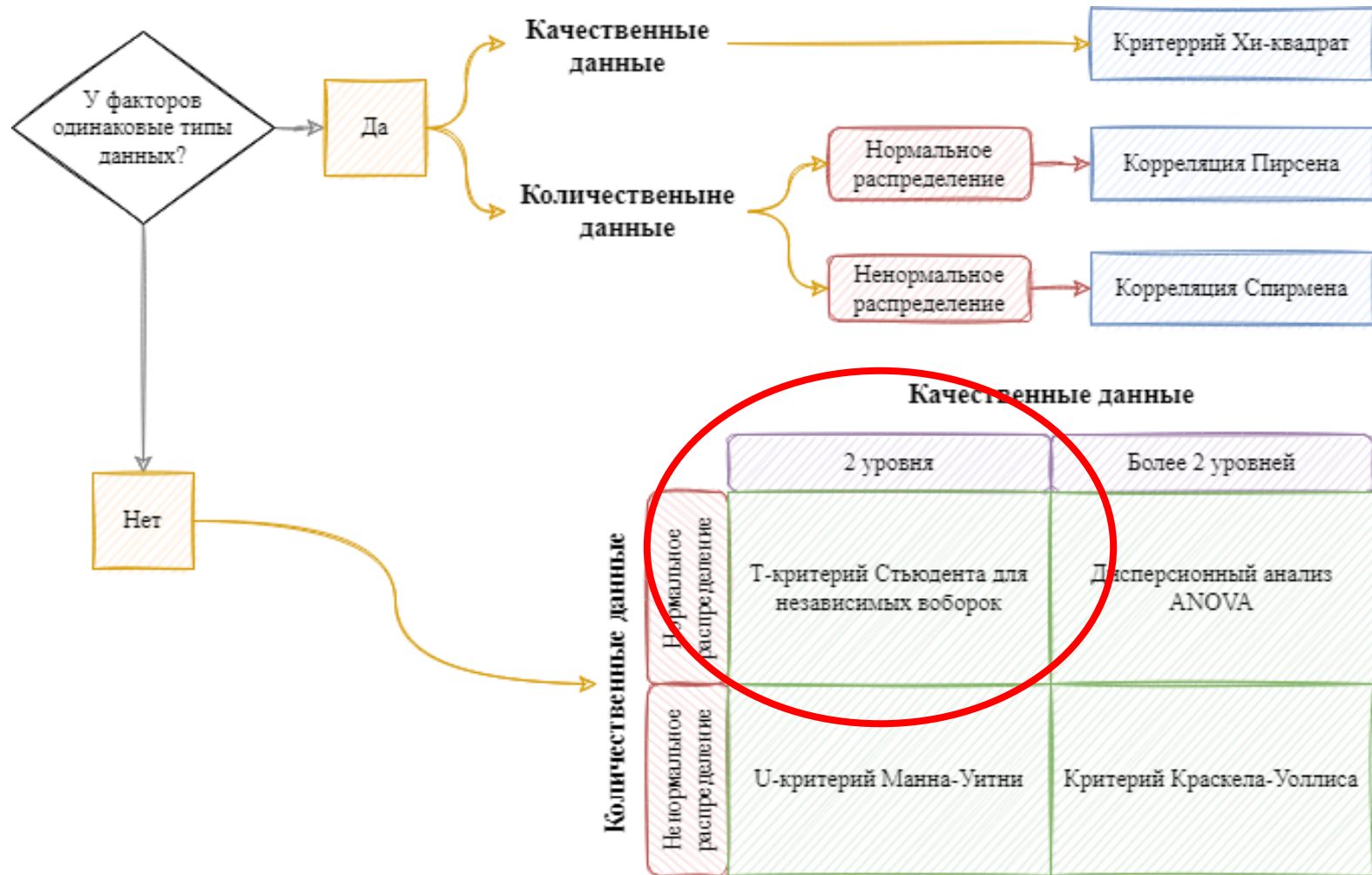
Исход
Длительность госпитализации

Группы
Метод А vs Метод В
Разные пациенты

Задача
Сравнить показатели
(длительность госпитализации)
между группами

Качественные данные		2 уровня
Количественные данные	Нормальное распределение	Т-критерий Стьюдента для независимых выборок
	Не нормальное распределение	U-критерий Манна-Уитни

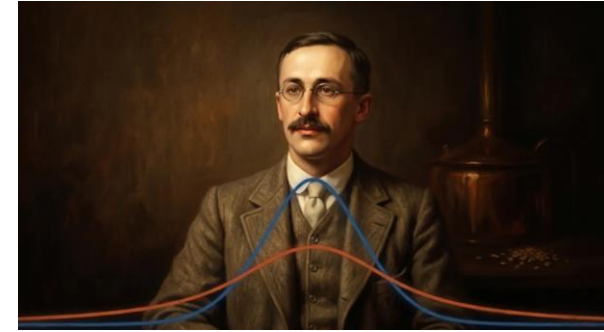
Алгоритм выбора метода анализа данных для независимых выборок



t-критерий Стьюдента (t-тест)



t-критерий Стьюдента проверяет равенство средних значений количественных данных в двух подгруппах.



Требования критерия:

- ✓ Качественные данные должны содержать **только 2 уровня (только две группы!!!)**.
- ✓ Данные **внутри** сравниваемых подгрупп должны быть **нормально распределены**.
- ✓ Каждая подгруппа должна иметь **желательно 30 значений**.



Запись результата: $\text{Mean} \pm \text{SD}$; $t(\text{df})$; 95% ДИ; p-value

t-критерий Стьюдента (t-тест)



В зависимости от характера собираемых данных применяют три основных сценария:

- ✓ **Одновыборочный (One-sample):** Сравнивает среднее значение одной группы с уже известным нормативным показателем или фиксированным числом.

Пример: Соответствует ли средний вес онкологического пациента, здоровому человеку, заявленным 70 кг?

- ✓ **Двухвыборочный для независимых выборок (Independent two-sample):** Сравнивает показатели двух совершенно разных, не связанных между собой групп людей или объектов.

Пример: Сравнение среднего веса онкологических пациентов и среднего крадиологических.

- ✓ **Парный / зависимый (Paired):** Сравнивает результаты одной и той же группы объектов, измеренные дважды (в разное время или в разных условиях).

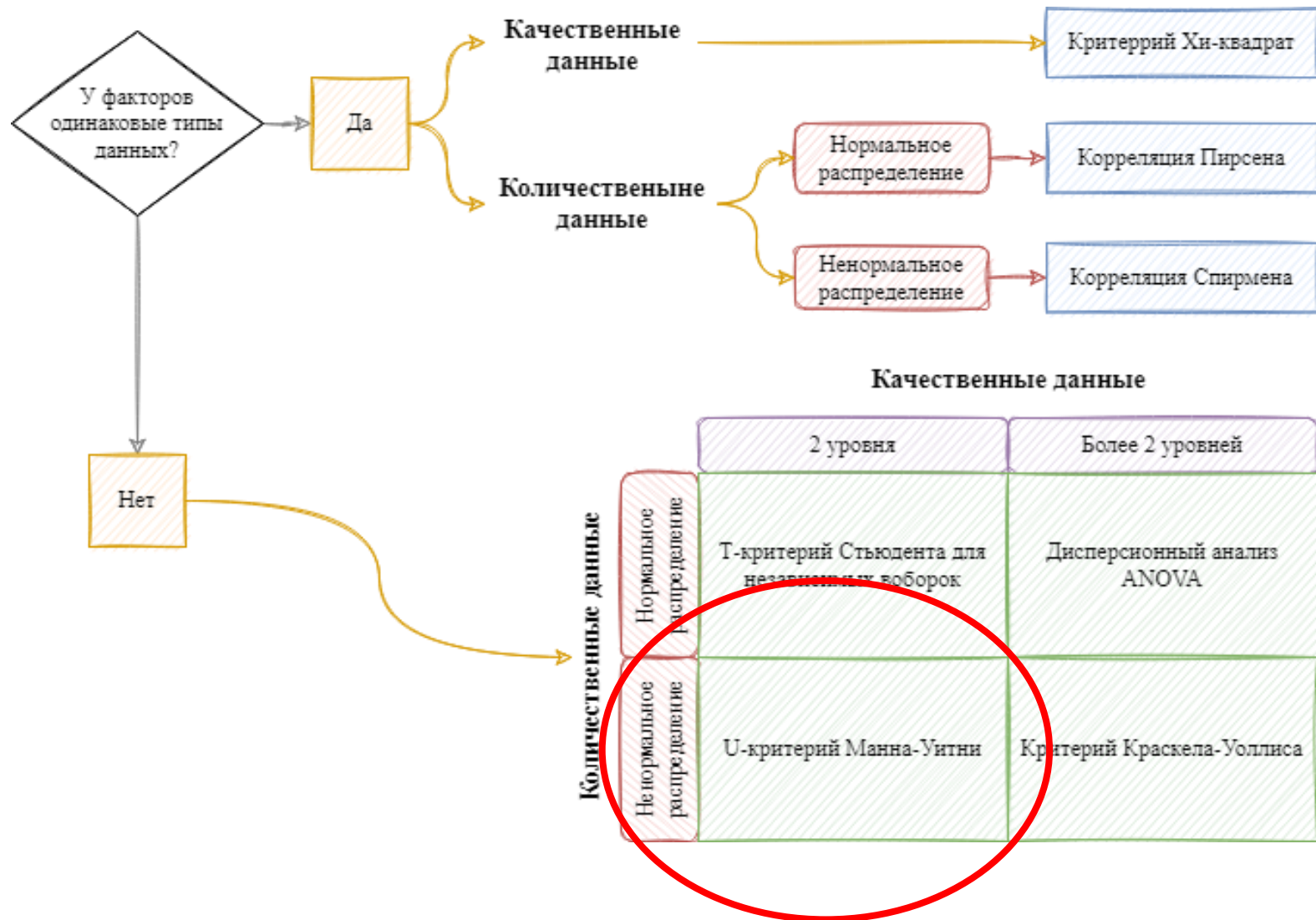
Пример: Сравнивает средний вес пациента до операции и после.



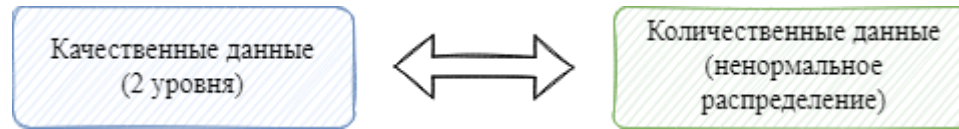
Форма записи: $\text{Mean} \pm \text{SD}$; $t(\text{df})$; 95% ДИ; p-value

Пример : «Анализ показал, что у пациентов, принимавших новый препарат, уровень гемоглобина был значимо выше ($M = 118.0$ г/л; $SD = 4.2$), чем в группе стандартной терапии ($M = 112.0$ г/л; $SD = 5.1$), $t(58) = 4.96$; $p < 0.001$; 95% ДИ [3.58; 8.42]».

Алгоритм выбора метода анализа данных для независимых выборок



U-критерий Манна-Уитни (U-тест)



U-критерий Манна-Уитни (U-тест Mann-Whitney) проверяет равенство медианных значений количественных данных в двух подгруппах.

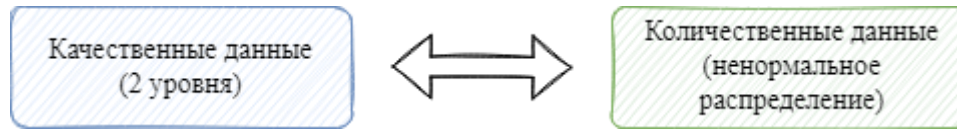
Требования критерия:

- Качественные данные должны содержать **только 2 уровня**.
- Данные **внутри** сравниваемых подгрупп могут быть **ненормально** распределены.



Запись результата: M, IQR, U, Z, p-value

U-критерий Манна-Уитни (U-тест)



U-критерий Манна-Уитни (U-тест Mann-Whitney) проверяет равенство медианных значений количественных данных в двух подгруппах.

В чем принципиальное отличие от t-теста?

t-тест Стьюдента сравнивает *средние арифметические значения* и требует, чтобы данные подчинялись закону нормального распределения (колоколообразной кривой).

U-тест Манна — Уитни игнорирует реальные значения и **сравнивает ранги (порядковые места) данных**, выстроенных по возрастанию. Ему абсолютно всё равно, какое у ваших данных распределение.



Запись результата: M, IQR, U, Z, p-value

U-критерий Манна-Уитни (U-тест)

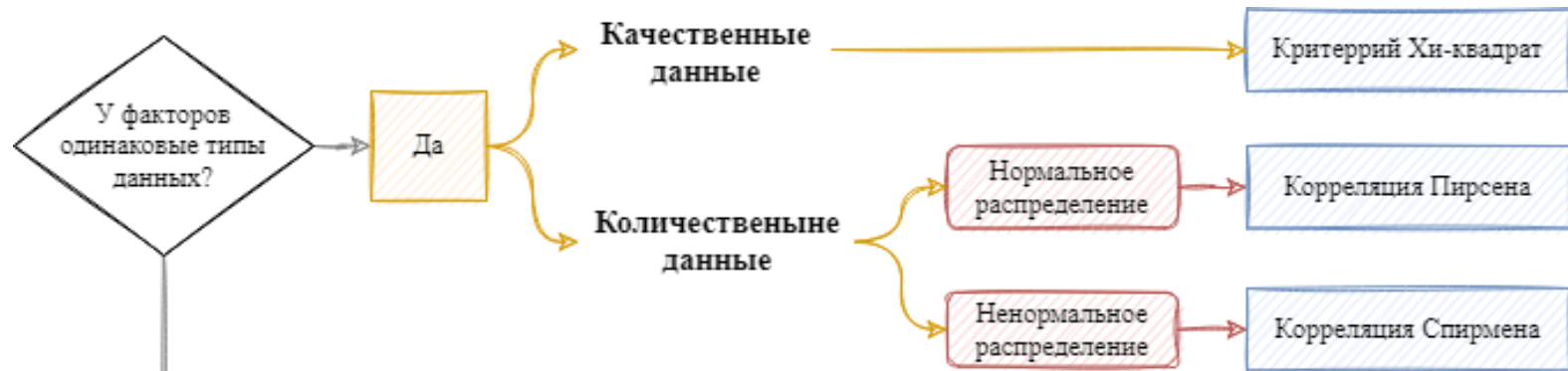
Когда именно нужен тест Манна — Уитни? Его выбирают при наличии следующих условий и ограничений:

- **Данные распределены ненормально.** Есть сильные перекосы, «длинные хвосты» или несколько пиков на графике.
- **Маленькие выборки.** Если у вас в группах всего по 5–10 человек, проверить их на нормальность распределения невозможно. Тест Манна — Уитни отлично работает на малых объемах (минимально — от 3 человек в группе).
- **Есть сильные выбросы (аномалии).** Представьте, что один пациент из-за осложнений пролежал в больнице 120 дней вместо обычных 10. t-тест «сломается», так как это число резко исказит среднее арифметическое. Тест Манна — Уитни просто присвоит этому числу самый высокий ранг, и экстремальное значение никак не повлияет на адекватность результата.
- **Данные качественные (порядковые).** Если результаты измеряются не в жестких физических величинах, а в баллах (например, оценка боли по шкале от 1 до 5, степени тяжести болезни: легкая/средняя/тяжелая).

Запись результата: M, IQR, U, Z, p-value

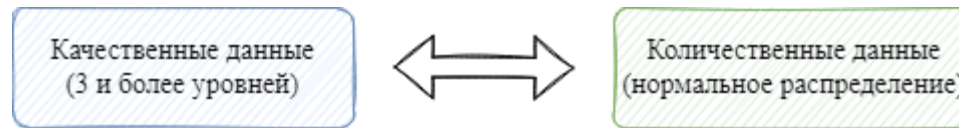
Пример : «Анализ показал, что длительность госпитализации пациентов, получавших новую терапию (M = 8 дней; IQR: 6–10), была статистически значимо ниже, чем в группе стандартного лечения (M = 12 дней; IQR: 9–15); U = 132.5; Z = -3.14; p = 0.002».

Алгоритм выбора метода анализа данных для независимых выборок



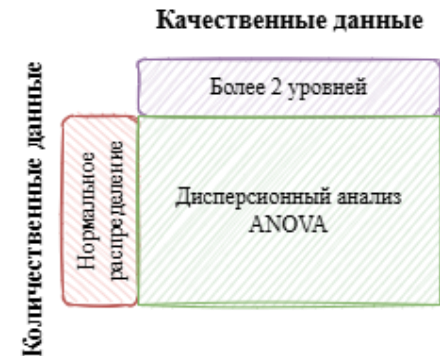
Качественные данные		Качественные данные	
		2 уровня	Более 2 уровней
Количественные данные	Нормальное распределение	Т-критерий Стьюдента для независимых выборок	Дисперсионный анализ ANOVA
	Не нормальное распределение	U-критерий Манна-Уитни	Критерий Краскела-Уоллиса

ANOVA-дисперсионный анализ



ANOVA-дисперсионный анализ (от англ. ANalysis Of VAriance — дисперсионный анализ) — это метод параметрической статистики, который используется для сравнения средних значений в трех и более независимых группах

Главная задача теста — выяснить, влияют ли качественные переменные (факторы) на количественный результат, или же все обнаруженные различия между группами являются чисто случайными.



Запись результата: $F(df, df(residuals))$, p-value, η^2

ANOVA-дисперсионный анализ

Качественные данные
(3 и более уровней)



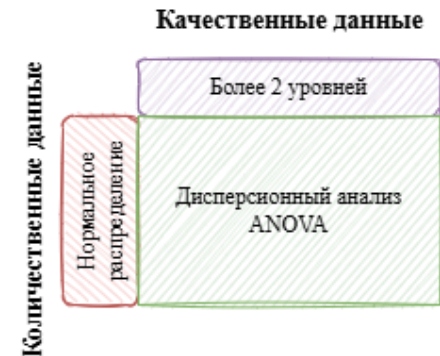
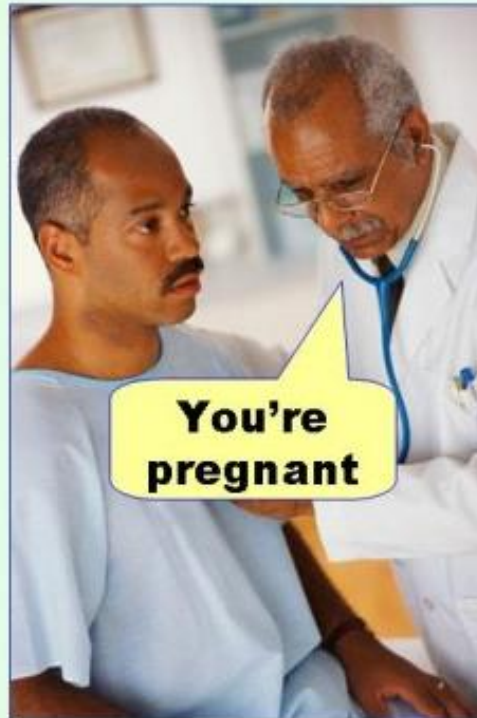
Количественные данные
(нормальное распределение)

ANOVA-дисперсионный анализ — дисперсионный анализ (Variance Analysis) — параметрической статистики, позволяющий проводить сравнения средних значений в нескольких группах.

Требования анализа:

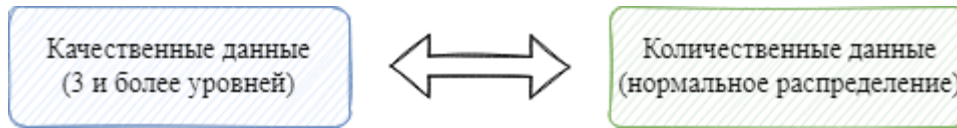
- Качественные данные должны иметь 3 и более уровней.
- Данные внутри сравниваемых групп должны быть нормально распределены.
- Каждая подгруппа должна быть достаточно большой.
- Гомогенность (равенство дисперсий) — дисперсия внутри групп должна быть приблизительно равна дисперсии между группами (проверяется тестом Левена).
- Независимость наблюдений — наблюдения в одной группе должны быть связаны друг с другом.

Type I error
(false positive)



Запись результата: $F(df, df(\text{residuals}))$, p-value, η^2

ANOVA-дисперсионный анализ



ANOVA-дисперсионный анализ

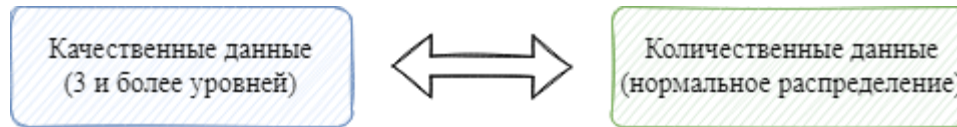
Требования анализ

- Качественные уровни.
 - Данные внутри нормально распределены.
 - Каждая подгруппа гомогенна (внутри групп должен тестом Левена).
 - Независимость наблюдений: участники в группах не должны быть связаны друг с другом.
-
- Дисперсия D



Запись результата: F(df, df(residuals)), p-value, η^2

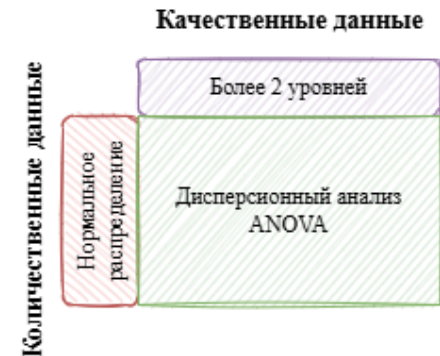
ANOVA-дисперсионный анализ



ANOVA-дисперсионный анализ (от англ. ANalysis Of VAriance — дисперсионный анализ) — это метод параметрической статистики, который используется для сравнения средних значений в трех и более независимых группах

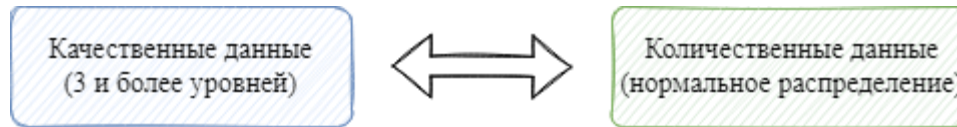
Основные виды ANOVA:

- **Однофакторный (One-Way ANOVA):** изучается влияние одного фактора на зависимую переменную (например: зависит ли уровень дохода от уровня образования — 4 группы образования). [1]
- **Двухфакторный (Two-Way ANOVA):** оценивается влияние двух факторов одновременно, а также их взаимодействие (например: зависит ли доход от уровня образования и пола сотрудника).



Запись результата: $F(df, df(residuals)), p\text{-value}, \eta^2$

ANOVA-дисперсионный анализ

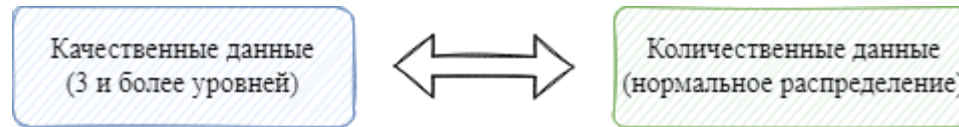


Шпаргалка: что выбирать?

Ситуация	Что делать / Какую поправку применять
Дисперсии равны, нужно сравнить все группы между собой	Классический ANOVA → Поправка Тьюки (Tukey)
Дисперсии равны, нужно сравнить только с контролем	Классический ANOVA → Критерий Даннетта (Dunnett)
Дисперсии НЕ равны	ANOVA Уэлча → Поправка Геймса-Хоуэлла (Games-Howell)
Групп мало, нужен простой и строгий контроль	Классический ANOVA → Поправка Холма или Бонферрони

Запись результата: $F(df, df(residuals)), p\text{-value}, \eta^2$

ANOVA-дисперсионный анализ

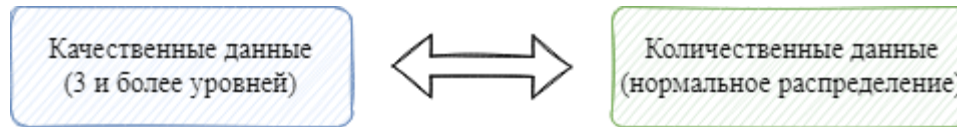


Для каждого фактора или взаимодействия вам нужно выписать четыре обязательных элемента:

1. **F** — символ критерия Фишера (пишется курсивом).
2. **Степени свободы (df)** — два числа, которые пишутся в скобках сразу после буквы *F*. Первое число — межгрупповые степени свободы (в jamovi столбец `df` в строке фактора), второе — внутригрупповые / остаточные (в jamovi строка `Residuals`).
3. **Значение F-статистики** — само число из столбца `F` (округляется до двух знаков после запятой).
4. **Значение p** — уровень значимости (курсивом, без ноля перед точкой в англоязычных статьях: $p = .006$, либо привычное $p = 0,006$ для русскоязычных).
5. **Размер эффекта (желательно)** — обычно используют частичную эта-квадрат (η_p^2) или омега-квадрат (ω^2).

Запись результата: $F(df, df(residuals)), p\text{-value}, \eta^2$

ANOVA-дисперсионный анализ



Шаблон для однофакторного ANOVA:

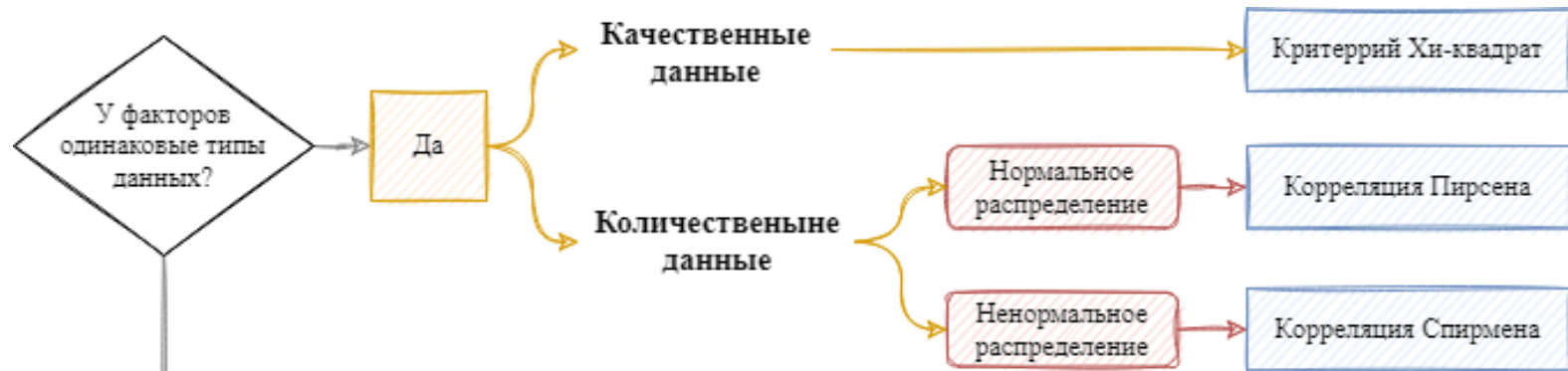
«Результаты однофакторного дисперсионного анализа показали статистически значимое влияние типа методики на уровень остаточных знаний: $F(2, 42) = 5,84$; $p = 0,006$; $\eta_p^2 = 0,22$ ».

Шаблон для многофакторного ANOVA (с взаимодействием):

«Двухфакторный дисперсионный анализ выявил значимый эффект взаимодействия факторов пола и уровня образования: $F(2, 84) = 5,51$; $p = 0,006$; $\eta_p^2 = 0,12$. При этом главные эффекты пола ($F(1, 84) = 1,20$; $p = 0,276$) и образования ($F(2, 84) = 2,15$; $p = 0,123$) оказались статистически незначимы».

Запись результата: $F(df, df(residuals))$, p-value, η^2

Алгоритм выбора метода анализа данных для независимых выборок



Качественные данные			
		2 уровня	Более 2 уровней
Количественные данные	Нормальное распределение	Т-критерий Стьюдента для независимых выборок	Дисперсионный анализ ANOVA
	Не нормальное распределение	U-критерий Манна-Уитни	Критерий Краскела-Уоллиса

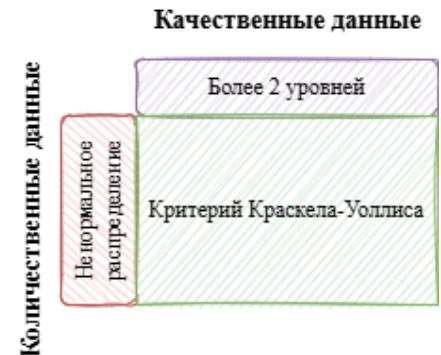
Критерий Краскела-Уоллиса



Критерий Краскела—Уоллиса (Kruskal-Wallis test/ Однофакторный ранговый дисперсионный анализ Краскела—Уоллиса) — это непараметрический статистический тест, который используется для сравнения трех и более независимых групп.

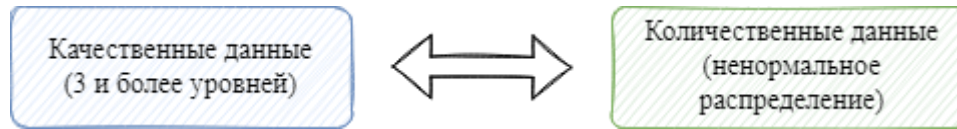
Главная задача критерия Краскела — Уоллиса — определить, есть ли статистически значимые различия между тремя или более независимыми группами по уровню исследуемого признака.

Если говорить простыми словами, тест отвечает на один конкретный вопрос: «Правда ли, что распределение значений (или медианы) хотя бы в одной из групп сильно сдвинуто относительно остальных, или же все группы примерно одинаковы?».



Запись результата: (χ^2 или H), df, p-value

Критерий Краскела-Уоллиса



Критерий Краскела—Уоллиса (Kruskal-Wallis test/ Однофакторный ранговый дисперсионный анализ Краскела—Уоллиса) — это непараметрический статистический тест, который используется для сравнения трех и более независимых групп.

Требования анализа:

- Качественные данные должны содержать **3 и более уровней**.
- **Нарушена нормальность распределения:** данные в группах распределены несимметрично, имеют тяжелые «хвосты» или выбросы (тест Шапиро — Уилка показал $p < 0.05$).
- **Маленький объем выборок:** в группах слишком мало наблюдений (например, менее 10–15 человек), из-за чего невозможно надежно проверить нормальность распределения..
- **Порядковые (ранговые) данные:** зависимая переменная измерена не в непрерывной шкале, а в баллах, оценках или рангах (например, пятибалльная шкала удовлетворенности, уровни боли от 1 до 10).

Количественные данные

Качественные данные	
Не нормальное распределение	Более 2 уровней
	Критерий Краскела-Уоллиса

Запись результата: (χ^2 или H), df, p-value

Критерий Краскела-Уоллиса



Как записывать результат в научных публикациях?

Для текста статьи вам понадобятся три значения из таблицы jamovi: величина критерия (χ^2 или H), число степеней свободы (df) и точное p -value.

«Для оценки различий между тремя группами использовался критерий Краскела — Уоллиса. Анализ выявил статистически значимые различия в уровне исследуемого признака: $\chi^2(2) = 10,45$; $p = 0,005$. Апостериорные сравнения с использованием критерия DSCF показали, что значимые различия наблюдаются только между Группой 1 и Группой 3 ($p = 0,004$), в то время как различия между остальными парами групп оказались незначимы ($p > 0,05$)».

Запись результата: (χ^2 или H), df , p -value

Поправка Бонферрони

Поправка Бонферрони — это самый простой и строгий метод защиты от **ошибки I рода** (ложноположительного результата) при множественных статистических сравнениях.

Способ 1. Коррекция порога значимости (Альфа)

Если вы хотите проводить тесты вручную и не трогать полученные программой p -значения, вы рассчитываете **новый, более строгий порог значимости ($\alpha_{\text{скорр}}$)**.

$$\alpha_{\text{скорр}} = \frac{\alpha_{\text{исходный}}}{k}$$

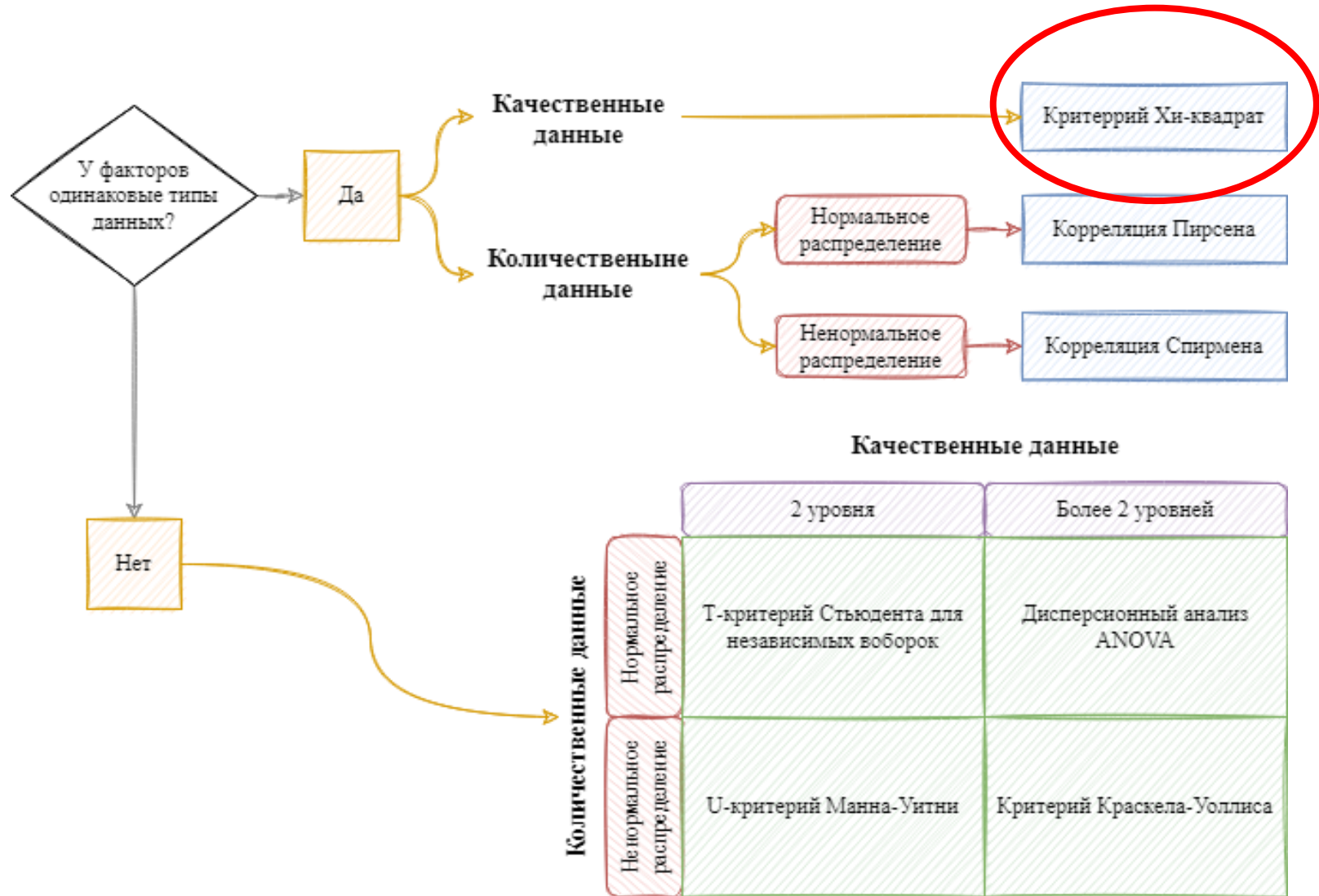
$$k = \frac{n \times (n - 1)}{2}$$

Способ 2. Коррекция самого p -value (Как делает Jamovi, SPSS, R)

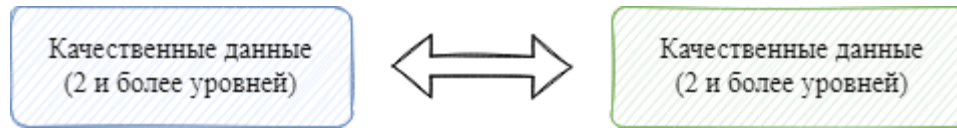
Программы для автоматизации не меняют глобальный порог 0.05, им удобнее изменить само полученное значение p . Формула «штрафа» выглядит так:

$$p_{\text{bonf}} = p_{\text{сырое}} \times k$$

Алгоритм выбора метода анализа данных для независимых выборок



Критерий Хи-квадрат (χ^2)

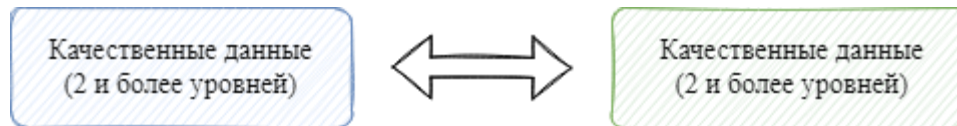


Критерий хи-квадрат (χ^2) — это один из самых популярных непараметрических методов в статистике, который используется для анализа категориальных (качественных) данных. Он позволяет проверить, насколько частоты признаков, полученные в ходе эксперимента (наблюдаемые), соответствуют теоретическим частотам (ожидаемым).

Критерий Хи-квадрат

Запись результата: χ^2 , df, p-value

Критерий Хи-квадрат (χ^2)



Два главных типа применения критерия

1. Критерий согласия (Goodness of Fit Test)

- **Зачем:** Проверяет, подчиняется ли выборка какому-то конкретному теоретическому закону распределения (например, равномерному, нормальному или менделевскому расщеплению признаков).
- *Пример:* Соответствует ли кубик стандартам казино, или какая-то грань выпадает подозрительно часто?

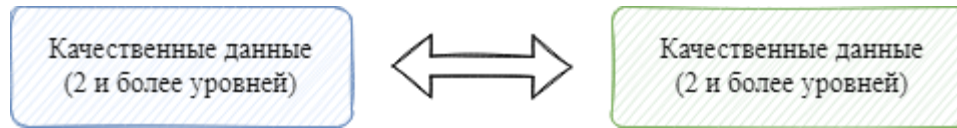
Критерий Хи-квадрат

2. Критерий независимости (Test of Independence)

- **Зачем:** Проверяет, связаны ли между собой два качественных признака.
- *Пример:* Зависит ли выбор любимого жанра кино от пола человека?

Запись результата: χ^2 , df, p-value

Критерий Хи-квадрат (χ^2)



Требования анализа:

- Тип данных должны быть представлены качественными (номинальные, порядковые) целочисленными частотами или бинарными (пол).
- Необходимо, но не строго, общее количество наблюдений было более 20.
- Неприемлемо использование данного критерия при анализе непрерывных абсолютных данных, процентов и долей.

Критерий Хи-квадрат

Запись результата: χ^2 , df, p-value

df	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.001$
1	2.706	3.841	5.024	6.635	10.828
2	4.605	5.991	7.378	9.210	13.816
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.833	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322
8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588
12	18.549	21.026	23.337	26.217	32.910
15	22.307	24.996	27.488	30.578	37.697
20	28.412	31.410	34.170	37.566	45.315

(χ^2)

качественные данные
(2 и более уровней)

для каждой ячейки: $E = \frac{\text{итого по строке} \times \text{итого по столбцу}}{\text{Всего пациентов}}$

«Лекарство / Заболели»: $(60 \times 40)/120 = 20$

«Лекарство / НЕ заболели»: $(60 \times 80)/120 = 40$

«Плацебо / Заболели»: $(60 \times 40)/120 = 20$

• Для ячейки «Плацебо / НЕ заболели»: $(60 \times 80)/120 = 40$

2. Расчет ожидаемых частот (E_i).

3. Вычисление значения χ^2 по формуле выше.

4. Определение степеней свободы (df):

- Для критерия согласия: $df = n - 1$ (где n — количество категорий).
- Для таблицы сопряженности признаков ($r \times c$): $df = (r - 1) \times (c - 1)$, где r — число строк, c — число столбцов.

5. Сравнение с критическим значением: Полученный χ^2 сравнивают со стандартным табличным значением для выбранного уровня значимости (обычно $\alpha = 0.05$) и найденного df . Если расчетный χ^2 больше критического, нулевая гипотеза отвергается.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

- O_i (Observed) — наблюдаемая частота (то, что получили на практике).
- E_i (Expected) — ожидаемая частота (то, что должно быть в теории, если нулевая гипотеза верна).

Запись результата: χ^2 , df , p -value

Критерий Хи-квадрат (χ^2)

«Применение нового препарата привело к статистически значимому снижению заболеваемости гриппом по сравнению с группой плацебо: **20,0%** (12/60) против **46,7%** (28/60) соответственно ($\chi^2 = 9,60$; $df = 1$; $p = 0,002$).»

Запись результата: χ^2 , df, p-value

Критерий Хи-квадрат (χ^2)

1. Поправка Йейтса (Continuity Correction)

- **Когда применяется:** Используется **только для таблиц 2×2** , когда общая выборка достаточно большая, но ожидаемая частота хотя бы в одной ячейке падает **ниже 5** (но остается **больше 1–2**).
- **Суть:** Это искусственное математическое «смягчение» формулы хи-квадрат (вычитание 0,5 из разности между наблюдаемым и ожидаемым значениями). Она делает критерий более консервативным (строгим).
- *Минус:* Сегодня многие статистики считают её устаревшей, так как она иногда *слишком* завышает p -value, из-за чего можно пропустить реальный медицинский эффект.

2. Точный критерий Фишера (Fisher's Exact Test)

- **Когда применяется:** Это «золотой стандарт» для медицины. Используется, когда выборка **совсем маленькая** (например, общая сумма пациентов меньше 20–30) или когда ожидаемые частоты в ячейках **меньше 5 (и особенно меньше 1)**.
- **Суть:** В отличие от хи-квадрат, который использует приближенные (асимптотические) формулы, метод Фишера рассчитывает **точную** математическую вероятность получить такое распределение данных.
- *Плюс:* Он абсолютно точен при любых, даже самых микроскопических выборках.

Запись результата: χ^2 , df, p -value

Критерий Хи-квадрат (χ^2)

В критерии хи-квадрат есть два основных способа провести post-hoc анализ:

Способ 1. Анализ стандартизированных остатков (Рекомендуемый)

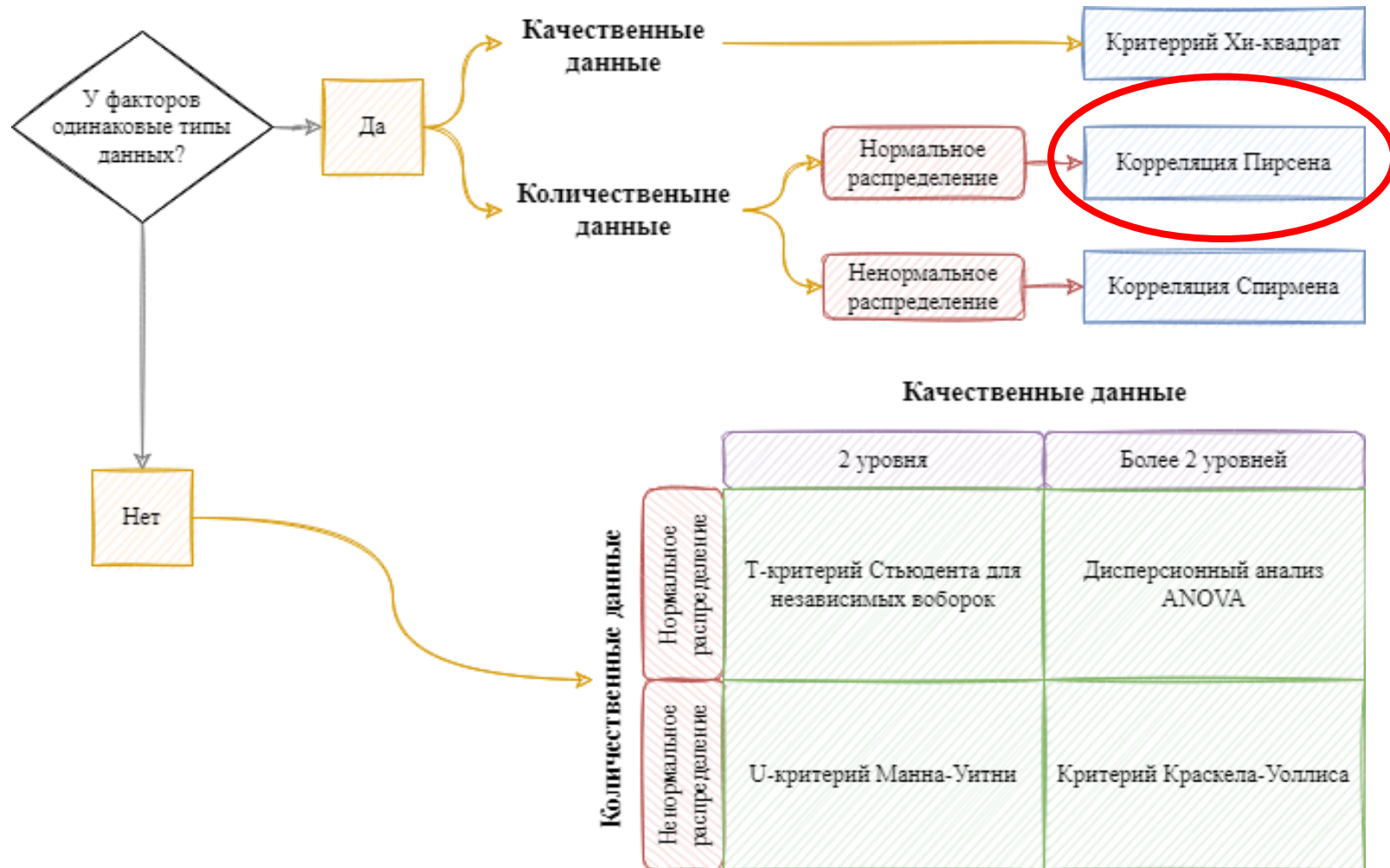
В настройках jamovi во вкладке **Cells** (Ячейки) нужно поставить галочку **Adjusted residuals** (Скорректированные остатки).

«Поскольку общее значение критерия хи-квадрат было значимым ($\chi^2 = 14,20$; $df = 2$; $p < 0,001$), для выявления межгрупповых различий был проведен post-hoc анализ с использованием поправки Бонферрони. Было установлено, что частота исхода в Группе А значимо отличалась от Группы Б ($p = 0,008$), в то время как различия между Группой Б и Группой В не достигли статистической значимости ($p = 0,420$).»

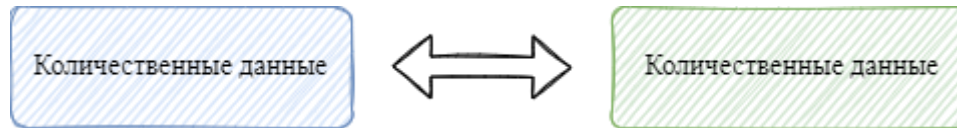
- **Важно:** при множественных сравнениях растет риск ложноположительного результата. Поэтому применяется **поправка Бонферрони**.
- **Суть поправки:** Вы делите стандартный уровень значимости (0,05) на количество сделанных сравнений. Если вы делаете 3 сравнения, то критическим уровнем p для каждого из них станет уже не 0,05, а $0,05 / 3 = 0,0167$.

Запись результата: χ^2 , df , p -value

Алгоритм выбора метода анализа данных для независимых выборок



Корреляция Пирсона

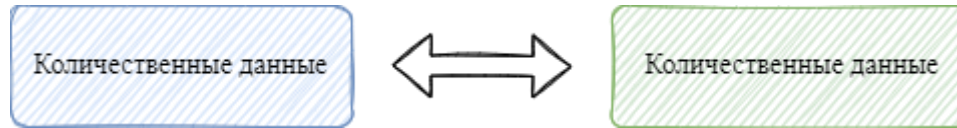


Коэффициент корреляции Пирсона (линейная корреляция) — это параметрический метод статистического анализа, который измеряет силу и направление линейной взаимосвязи между двумя непрерывными числовыми переменными.

Разработанный в 1890-х годах математиком Карлом Пирсоном, этот коэффициент обозначается буквой r и принимает значения в диапазоне от -1 до $+1$.

Запись результата: r , p -value

Корреляция Пирсона



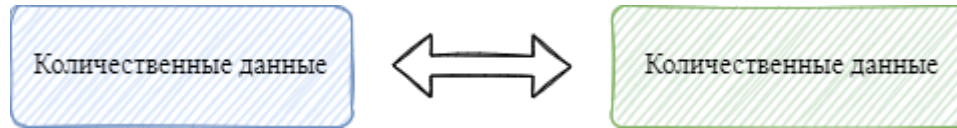
Важные условия применения

Чтобы результаты корреляции Пирсона были корректными, данные должны соответствовать следующим критериям:

1. **Непрерывный тип данных:** Обе переменные должны быть измерены в интервальной шкале или шкале отношений (например, рост, вес, температура, доход).
2. **Нормальное распределение:** Обе выборки должны быть распределены приблизительно нормально. Если распределение сильно отличается от нормального, лучше использовать ранговый [коэффициент Спирмена](#). 
3. **Отсутствие грубых выбросов:** Коэффициент Пирсона крайне чувствителен к аномалиям (выбросам), которые могут искусственно завысить или занижить показатель.

Запись результата: r , p -value

Корреляция Пирсона



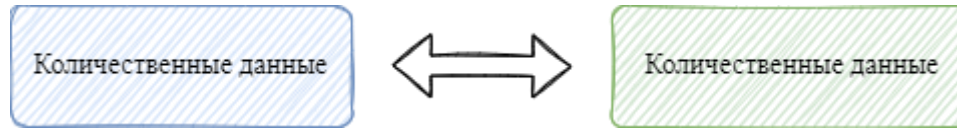
Главные отличия: Непрерывные vs Дискретные данные

Числовые данные в статистике делятся на два основных типа:

Характеристика	 Непрерывные данные	 Дискретные данные
Что делают?	Их измеряют с помощью приборов (линейки, весов, часов).	Их считают штуками или целыми единицами.
Значения	Любые числа, включая дробные и десятичные (1.5, 3.1415, 99.9).	Только изолированные значения (обычно целые числа: 1, 2, 3).
Пример	Рост человека (например, 175.45 см). Между 175 и 176 см есть бесконечно много вариантов.	Количество детей в семье (может быть 2 или 3, но не 2.5).

Запись результата: r , p -value

Корреляция Пирсона



Как интерпретировать значение r

Направление связи определяется знаком коэффициента, а сила — его абсолютной величиной:

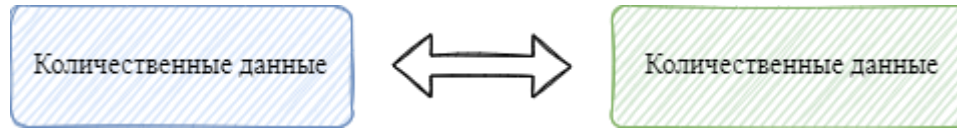
- $r = 1$: Идеальная (функциональная) **положительная** линейная связь. При увеличении одной переменной вторая строго пропорционально растет.
- $r = -1$: Идеальная **отрицательная** (обратная) линейная связь. При увеличении одной переменной вторая строго пропорционально уменьшается.
- $r = 0$: Линейная связь **отсутствует**. (При этом между переменными может существовать сильная нелинейная зависимость, например, квадратичная).

Шкала силы связи (по Чеддоку):

- До ± 0.3 — **слабая** связь
- От ± 0.3 до ± 0.7 — **умеренная** (средняя) связь
- Более ± 0.7 — **сильная** (высокая) связь
- Более ± 0.9 — **очень сильная** связь

Запись результата: r , p-value

Корреляция Пирсона



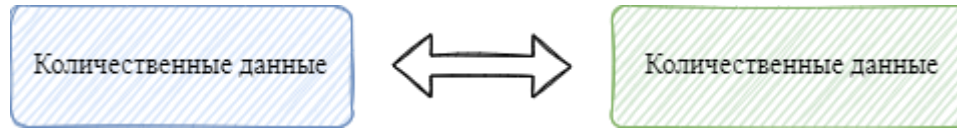
Главное правило: Корреляция $\neq$ Причинно-следственная связь

Высокий коэффициент корреляции между X и Y указывает лишь на то, что они изменяются согласованно. Это **не доказывает**, что X является причиной изменений в Y . Взаимосвязь может быть вызвана третьим (скрытым) фактором или оказаться чистым совпадением (ложная корреляция).

Пример ложной корреляции: продажи мороженого и количество лесных пожаров сильно коррелируют, но причиной обоих явлений является общая третья переменная — летняя жара.

Запись результата: r , p -value

Корреляция Пирсона



Вариант 1. Стандартный (Рекомендуемый для большинства журналов)

«Выявлена прямая (положительная) связь умеренной силы между уровнем систолического артериального давления и индексом массы тела у пациентов с гипертонической болезнью ($r = 0,482$; $p < 0,001$).»

Вариант 2. С указанием объема выборки (n)

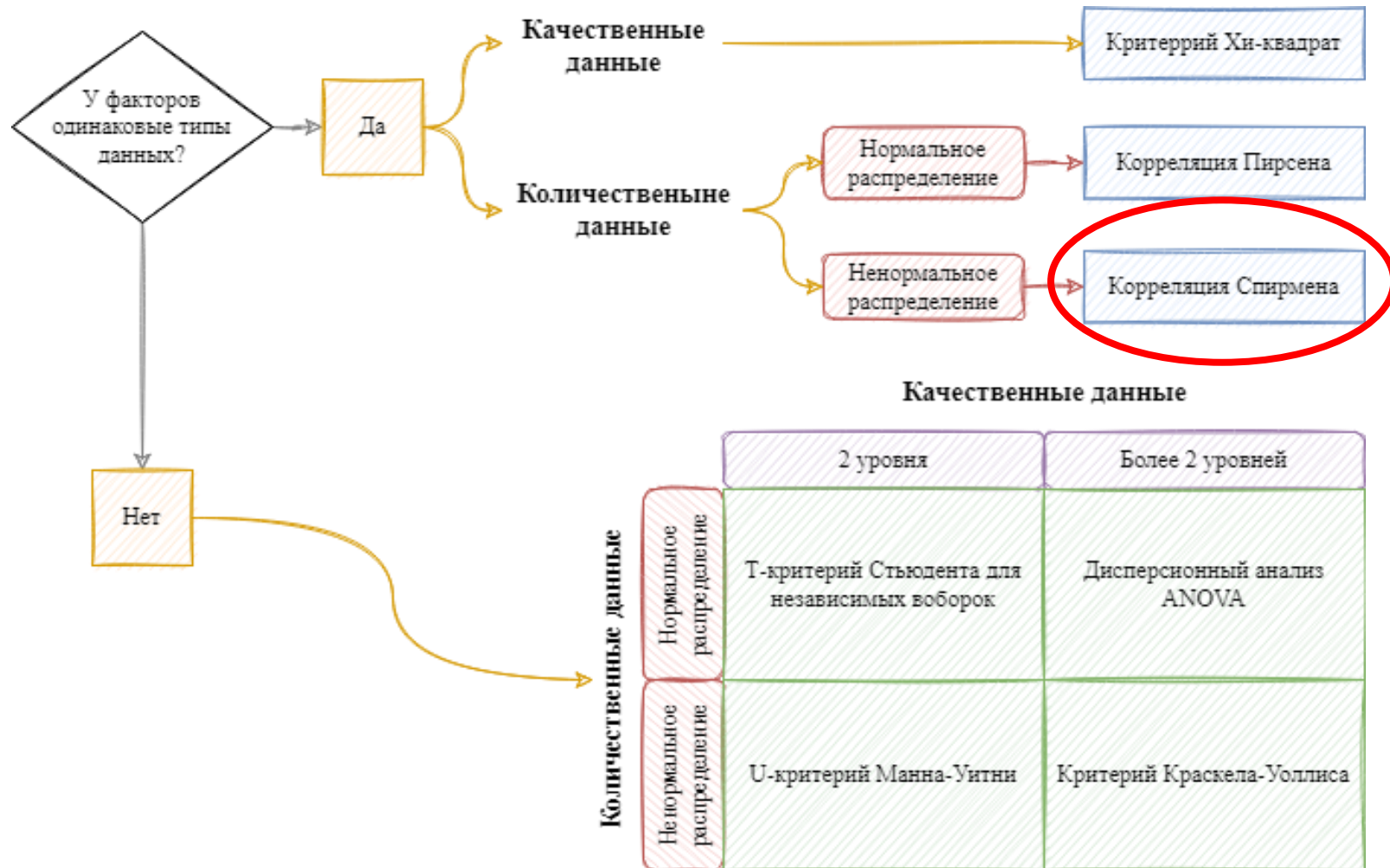
«Анализ данных показал наличие обратной (отрицательной) сильной взаимосвязи между фракцией выброса левого желудочка и уровнем мозгового натрийуретического пептида (NT-proBNP) ($r = -0,74$; $n = 120$; $p = 0,003$).»

Вариант 3. Когда связь отсутствует (Статистически не значима)

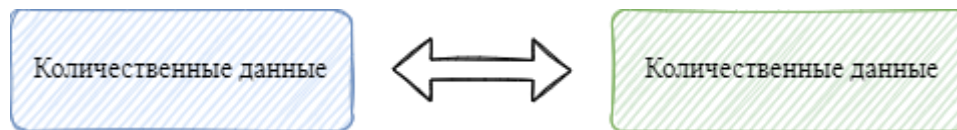
«Не было обнаружено статистически значимой линейной связи между возрастом пациентов и уровнем глюкозы в плазме крови натощак ($r = 0,08$; $p = 0,412$).»

Запись результата: r , p -value

Алгоритм выбора метода анализа данных для независимых выборок



Корреляция Спирмена



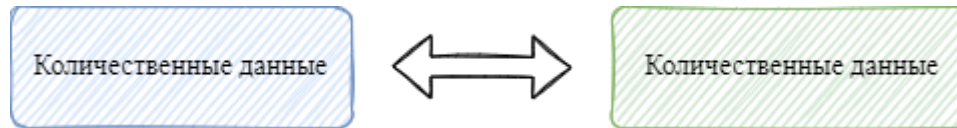
Коэффициент ранговой корреляции Спирмена — это непараметрический метод, который измеряет **силу и направление монотонной связи** между двумя переменными.

В отличие от Пирсона, он оценивает не строго линейную зависимость, а то, насколько согласованно возрастают или убывают показатели. Для этого реальные значения данных переводятся в **ранги** (порядковые номера).

Его обозначают греческой буквой **ρ (ро)** или латинским сочетанием **r_s** . Принимает значения также **от -1 до +1**.

Запись результата: χ^2 , df, p-value

Корреляция Спирмена

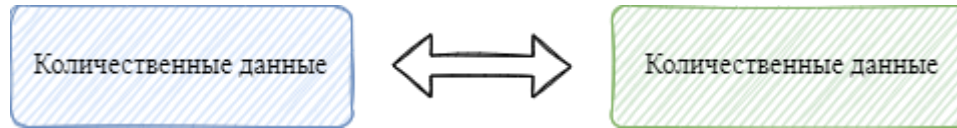


Когда в медицине выбирают именно Спирмена (вместо Пирсона)?

- **Порядковые (ранговые) шкалы:** Когда данные измерены в баллах, стадиях, степенях тяжести (например, стадии рака 1-4, баллы по шкале Апгар, уровни тревожности).
- **Ненормальное распределение данных:** Большинство медицинских показателей (уровень гормонов, антител, ферментов в крови) распределены ненормально, с сильной асимметрией. Пирсон здесь совет, а Спирмен работает корректно.
- **Наличие тяжелых выбросов:** Если у одного-двух пациентов в выборке показатели «улетели в космос» (например, экстремально высокий уровень СРБ), Спирмен сгладит этот шум, так как переведет цифры в порядковые ранги.

Запись результата: χ^2 , df, p-value

Корреляция Спирмена



Как интерпретировать значение r

Направление связи определяется знаком коэффициента, а сила — его абсолютной величиной:

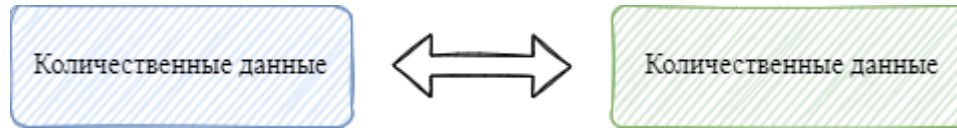
- $r = 1$: Идеальная (функциональная) **положительная** линейная связь. При увеличении одной переменной вторая строго пропорционально растет.
- $r = -1$: Идеальная **отрицательная** (обратная) линейная связь. При увеличении одной переменной вторая строго пропорционально уменьшается.
- $r = 0$: Линейная связь **отсутствует**. (При этом между переменными может существовать сильная нелинейная зависимость, например, квадратичная).

Шкала силы связи (по Чеддоку):

- До ± 0.3 — **слабая** связь
- От ± 0.3 до ± 0.7 — **умеренная** (средняя) связь
- Более ± 0.7 — **сильная** (высокая) связь
- Более ± 0.9 — **очень сильная** связь

Запись результата: r , p-value

Корреляция Спирмена



Вариант 1. Положительная (прямая) монотонная связь

«В результате анализа была обнаружена прямая связь средней силы между стадией хронической болезни почек и уровнем креатинина в сыворотке крови ($\rho = 0,541$; $p < 0,001$).»

Вариант 2. Отрицательная (обратная) монотонная связь

«Выявлена выраженная обратная зависимость между баллом по шкале оценки когнитивных функций MMSE и возрастом пациентов с деменцией ($r_s = -0,68$; $n = 84$; $p = 0,014$).»

Вариант 3. Отсутствие связи (Статистически не значимо)

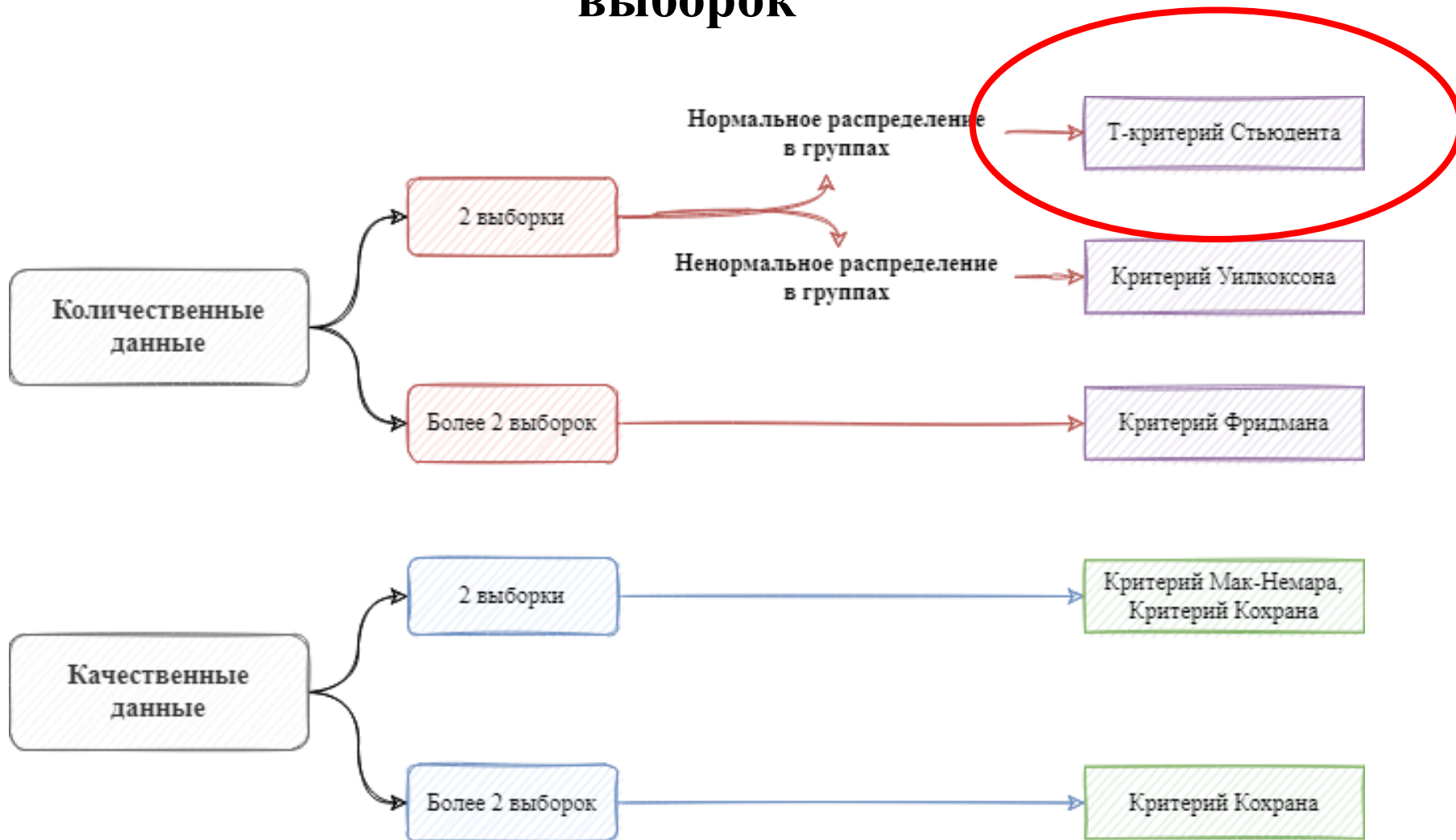
«Не установлено статистически значимой взаимосвязи между выраженностью болевого синдрома по визуально-аналоговой шкале (ВАШ) и длительностью заболевания ($\rho = 0,11$; $p = 0,320$).»

Запись результата: r , p-value

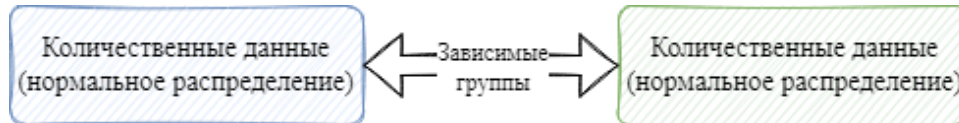
Алгоритм выбора метода анализа данных для зависимых выборок



Алгоритм выбора метода анализа данных для зависимых выборок



t-критерий Стьюдента (t-тест) для зависимых групп



t-критерий Стьюдента для зависимых (связанных) выборок — это параметрический статистический метод, который используется для сравнения средних значений двух групп, когда между элементами этих групп есть парная связь.

Требования критерия:

- ✓ **Зависимость выборок:** данные в двух группах должны быть парными (для каждого измерения в первой выборке есть строго соответствующее измерение во второй).
- ✓ **Тип данных:** признак должен быть измерен в количественной шкале (интервальной или шкале отношений).
- ✓ **Отсутствие грубых выбросов:** сильные экстремальные значения могут исказить среднее и привести к ложным выводам.
- ✓ **Нормальное распределение:** разности парных значений должны быть распределены нормально.

Запись результата: Mean ± SD; t(df); 95% ДИ; p-value

t-критерий Стьюдента (t-тест) для зависимых групп

Алгоритм проверки гипотез

1. Формулируются гипотезы:

- H_0 (нулевая): Различий между средними значениями нет ($M_1 = M_2$, то есть среднее разностей $M_d = 0$).
- H_1 (альтернативная): Различия между средними значениями статистически значимы ($M_1 \neq M_2$).

2. Рассчитывается эмпирическое значение $t_{\text{эмп}}$ по формуле выше.

3. Определяется критическое значение $t_{\text{крит}}$ по [таблице распределения Стьюдента](#) для заданного уровня значимости ($\alpha = 0.05$ или $\alpha = 0.01$) и числа степеней свободы df .

4. Принимается решение:

- Если $|t_{\text{эмп}}| \geq t_{\text{крит}}$ (или $p\text{-value} < 0.05$), то **нулевая гипотеза отвергается**. Изменения признака статистически значимы.
- Если $|t_{\text{эмп}}| < t_{\text{крит}}$ (или $p\text{-value} \geq 0.05$), то **нулевая гипотеза принимается**. Изменения случайны, эффект отсутствует.

t-критерий Стьюдента (t-тест) для зависимых групп

⚙ Алгоритм проверки гипотез

1. Формулируются гипотезы:

- H_0 (нулевая): Различий между средними значениями нет ($M_1 = M_2$), среднее разностей $M_d = 0$).
- H_1 (альтернативная): Различия между средними значениями значимы ($M_1 \neq M_2$).

2. Рассчитывается эмпирическое значение $t_{\text{эмп}}$ по формуле выше.

3. Определяется критическое значение $t_{\text{крит}}$ по [таблице распределения Стьюдента](#) для заданного уровня значимости ($\alpha = 0.05$ или $\alpha = 0.01$) и числа степеней свободы df .

4. Принимается решение:

- Если $|t_{\text{эмп}}| \geq t_{\text{крит}}$ (или $p\text{-value} < 0.05$), то **нулевая гипотеза отвергается**. Изменения признака статистически значимы.
- Если $|t_{\text{эмп}}| < t_{\text{крит}}$ (или $p\text{-value} \geq 0.05$), то **нулевая гипотеза принимается**. Изменения случайны, эффект отсутствует.

📊 Формула расчета

Суть метода заключается в том, что мы переходим от анализа двух выборок к анализу **одной выборки разностей**.

Для каждого испытуемого вычисляется разность: $d_i = x_{1i} - x_{2i}$.

Формула эмпирического значения t-критерия выглядит так:

$$t = \frac{|M_d|}{\frac{\sigma_d}{\sqrt{n}}}$$

Где:

- M_d — среднее арифметическое разностей;
- σ_d — стандартное отклонение разностей;
- n — количество пар наблюдений (объем выборки);
- $\frac{\sigma_d}{\sqrt{n}}$ — стандартная ошибка разности средних.

Форма записи: $\text{Mean} \pm \text{SD}; t(df); 95\% \text{ ДИ}; p\text{-value}$

Критические точки распределения Стьюдента.						
Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,51	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
50	1,67	2,00	2,39	2,66	3,23	3,46
60	1,66	1,98	2,36	2,62	3,17	3,37
120	1,64	1,96	2,33	2,58	3,09	3,29
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
Уровень значимости α (односторонняя критическая область)						

t-критерий Стьюдента (t-тест)

для зависимых групп

Формула расчета

Суть метода заключается в том, что мы переходим от анализа двух выборок к анализу **одной выборки разностей**.

Для каждого испытуемого вычисляется разность: $d_i = x_{1i} - x_{2i}$.

Формула эмпирического значения t-критерия выглядит так:

$$t = \frac{|M_d|}{\frac{\sigma_d}{\sqrt{n}}}$$

и значениями нет (M_1 :

Где:

средними значениями

- M_d — среднее арифметическое разностей;
- σ_d — стандартное отклонение разностей;
- n — количество пар наблюдений (объем выборки);
- $\frac{\sigma_d}{\sqrt{n}}$ — стандартная ошибка разности средних.

ие $t_{\text{эмп}}$ по формуле

3. Определяется критическое значение $t_{\text{крит}}$ по [таблице распределения Стьюдента](#) для заданного уровня значимости ($\alpha = 0.05$ или $\alpha = 0.01$) и числа степеней свободы df .

4. Принимается решение:

- Если $|t_{\text{эмп}}| \geq t_{\text{крит}}$ (или $p\text{-value} < 0.05$), то **нулевая гипотеза отвергается**. Изменения признака статистически значимы.
- Если $|t_{\text{эмп}}| < t_{\text{крит}}$ (или $p\text{-value} \geq 0.05$), то **нулевая гипотеза принимается**. Изменения случайны, эффект отсутствует.

Форма записи: $\text{Mean} \pm \text{SD}; t(df); 95\% \text{ ДИ}; p\text{-value}$

t-критерий Стьюдента (t-тест) для зависимых групп

1. **Дескриптивную статистику** до и после вмешательства: среднее значение (M) и показатель разброса — стандартное отклонение (SD) или стандартную ошибку (m или SE).
2. **Значение самого t-критерия** (желательно, но в современной западной периодике иногда опускается в пользу p-value и доверительных интервалов).
3. **Число степеней свободы (df)** или точный объем выборки (n).
4. **Точное значение уровня значимости (p)** (например, $p = 0.024$, а не просто $p < 0.05$).
5. **95% доверительный интервал (95% ДИ или CI)** для разности средних (требование большинства высокорейтинговых журналов).

Запись результата: Mean $\pm$ SD; t(df); 95% ДИ; p-value

t-критерий Стьюдента (t-тест) для зависимых групп

1. **Дескриптивную статистику** (показатель разброса – *t* или *SE*).

2. **Значение самого t-критерия** (периодике иногда опускают)

3. **Число степеней свободы**

4. **Точное значение уровня значимости** ($p < 0.05$).

5. **95% доверительный интервал (95% ДИ или CI)** для разности средних (требование большинства высокорейтинговых журналов).

Пример 1: Классический вариант (для отечественных журналов ВАК / РИНЦ)

«В результате проведенного курса терапии у пациентов с артериальной гипертензией ($n = 32$) отмечено статистически значимое снижение уровня систолического артериального давления (САД). Среднее значение САД до лечения составляло 154.2 ± 8.4 мм рт. ст., а после лечения снизилось до 132.5 ± 6.1 мм рт. ст. Для зависимых выборок **t-критерий Стьюдента составил 11.45** при числе степеней свободы **$df = 31$** (различия высокосignимы, **$p < 0.001$**). Полученные данные свидетельствуют о высокой гипотензивной эффективности исследуемого препарата».

Запись результата: Mean $\pm$ SD; t(df); 95% ДИ; p-value

t-критерий Стьюдента (t-тест) для зависимых групп

Пример 1: Классический вариант (для отечественных журналов ВАК / РИНЦ)

Пример 2: Современный международный стандарт (для журналов Scopus / WoS по стилю APA)

"After 8 weeks of lifestyle intervention, participants exhibited a statistically significant decrease in body mass index (BMI). The mean BMI decreased from $31.4 \pm 2.8 \text{ kg/m}^2$ at baseline to $29.1 \pm 2.5 \text{ kg/m}^2$ post-intervention. The paired-samples t-test revealed a significant effect of the intervention, $t(24) = 4.12$, $p = 0.0004$, 95% CI [1.14, 3.46]. These findings demonstrate that the structured lifestyle program effectively reduces BMI in obese individuals."

(Перевод: После 8 недель изменения образа жизни у участников было отмечено статистически значимое снижение индекса массы тела (ИМТ). Средний ИМТ снизился с $31.4 \pm 2.8 \text{ кг/м}^2$ в начале исследования до $29.1 \pm 2.5 \text{ кг/м}^2$ после его завершения. Парный t-критерий Стьюдента выявил значимый эффект вмешательства: $t(24) = 4.12$, $p = 0.0004$, 95% ДИ [1.14, 3.46]. Здесь цифра 24 в скобках — это число степеней свободы $df = n - 1$).

иентов с
этистически
иального давления
ляло $154.2 \pm 8.4 \text{ мм}$
 мм рт. ст. Для
гавил 11.45 при
значимы, $p < 0.001$).
гипотензивной

Запись результата: Mean $\pm$ SD; $t(df)$; 95% ДИ; p-value

t-критерий Стьюдента (t-тест) для зависимых групп

Пример 1: Классический вариант (для отечественных журналов ВАК / РИНЦ)

Пример 2: Современный международный стандарт (для журналов Scopus / WoS по стилю APA)

"After
signif
from
The p
interv
demo
in obe

(Пер
отме
(ИМТ
до 29
Стью
0.00
степе

Оформление в таблицах

Если показателей много, их выносят в таблицу, а в тексте делают краткую ссылку. В таблицах t-критерий указывают в отдельной колонке или выносят информацию о нем в примечание под таблицей.

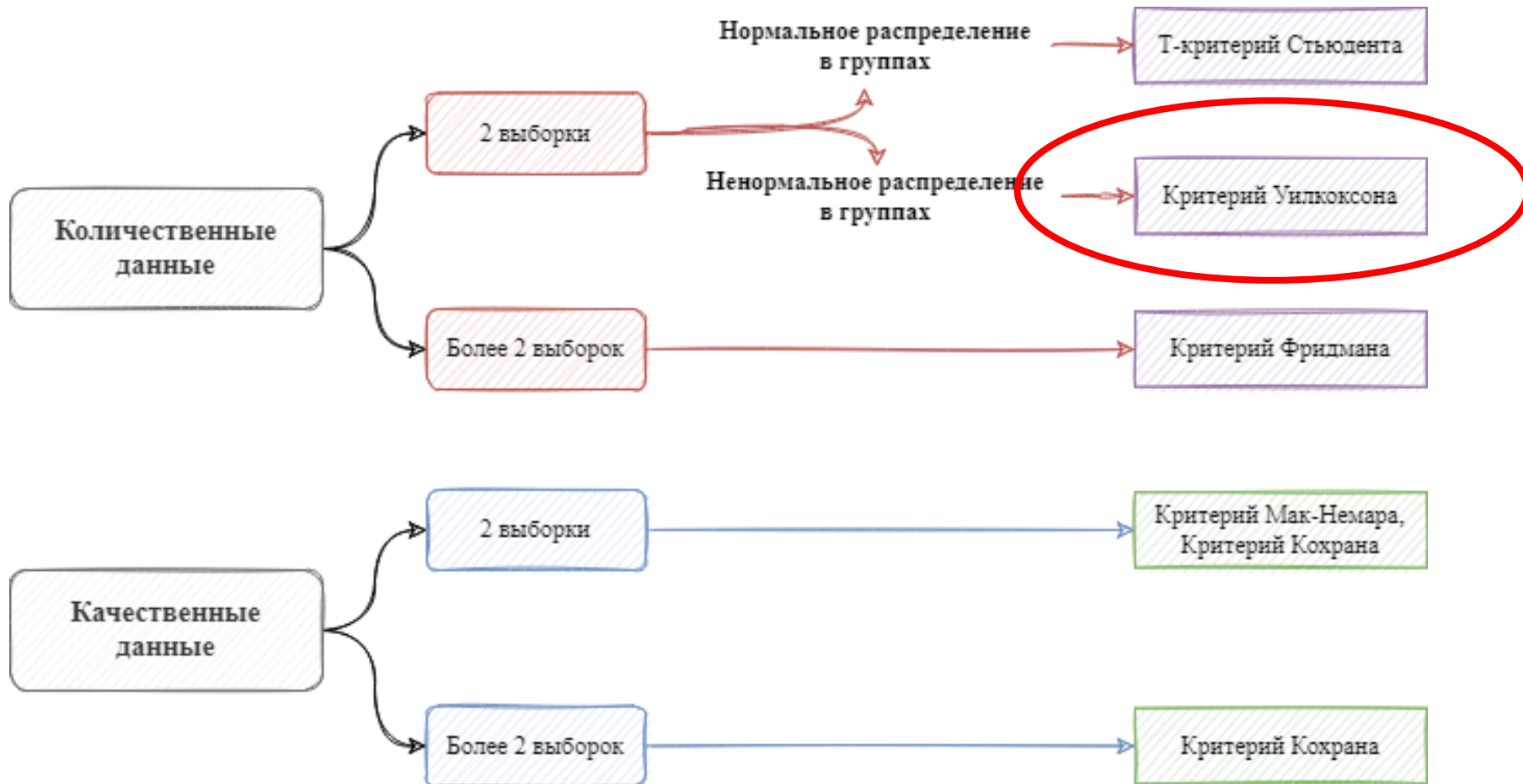
Показатель	До лечения ($M \pm SD$)	После лечения ($M \pm SD$)	Изменение, %	Значение t	Значимость, p
Холестерин, ммоль/л	6.2 ± 0.9	5.1 ± 0.7	-17.7%	5.24	0.012
Триглицериды, ммоль/л	2.4 ± 0.5	1.9 ± 0.4	-20.8%	3.89	0.038

авления
8.4 мм
для
при
< 0.001).
ной

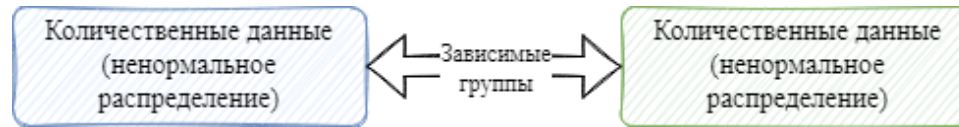
иентов с

Запись результата: Mean ± SD; t(df); 95% ДИ; p-value

Алгоритм выбора метода анализа данных для зависимых выборок



Критерий Вилкоксона



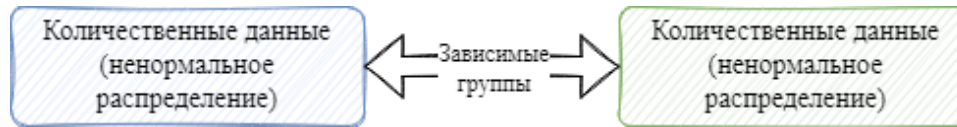
Критерий Вилкоксона (Уилкоксона) — это непараметрический статистический тест, который используется для сравнения двух выборок по уровню количественного или порядкового признака. Его главное преимущество — он **не требует нормального распределения** данных (в отличие от параметрического t-критерия Стьюдента).

Требования критерия:

- ✓ **Тип и шкала измерения данных:** данные должны представлять собой числа или баллы, которые можно логически расставить по порядку.
- ✓ **Зависимость и парность выборок.**
- ✓ **Ограничения на размер выборки.**
- ✓ **Распределение:** распределение данных хотя бы одной из групп не является ненормальным.
- ✓ **Нулевые сдвиги:** если у кого-то из испытуемых показатель «до» и «после» совпал, такие пары обычно исключаются из анализа, что уменьшает реальный размер выборки (N).

Запись результата: M, IQR, Z (T), p-value

Критерий Вилкоксона



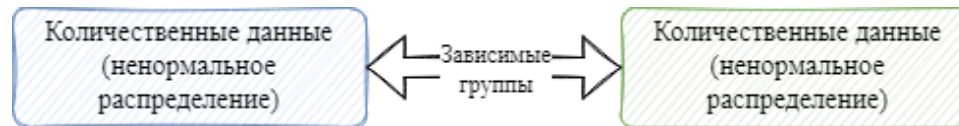
Т-критерий Вилкоксона (для зависимых выборок)

Используется для сравнения одной и той же группы объектов **в двух разных условиях или в режиме «до — после»** (например, замер уровня тревожности до психологического тренинга и после него).

- **Суть метода:** Тест оценивает не только направление изменений (увеличился показатель или уменьшился), но и их выраженность (интенсивность сдвига). Для этого рассчитываются разности парных значений, их модули ранжируются по возрастанию, а затем суммируются ранги для положительных и отрицательных сдвигов.
- **Статистика Т:** Равна сумме рангов того направления, которое произошло реже (так называемый «нетипичный сдвиг»).
- **Правило принятия решения:** Если полученное эмпирическое значение $T_{\text{эмп}}$ **меньше или равно** критическому табличному значению $T_{\text{кр}}$, то сдвиг признака считается статистически достоверным (нулевая гипотеза об отсутствии различий отвергается).

Запись результата: M, IQR, Z (T), p-value

Критерий Вилкоксона



Правила оформления в статьях

1. **Обоснование выбора:** В разделе «Материалы и методы» обязательно укажите, почему выбран именно этот критерий (например: «Ввиду распределения признаков, отличного от нормального, для сравнения зависимых выборок использовался T-критерий Вилкоксона»).
2. **Формат представления данных:** Поскольку распределение ненормальное, среднее арифметическое (M) и стандартное отклонение (SD) указывать некорректно. Данные приводят в виде **медианы и интерквартильного размаха**: **$Me [Q_1; Q_3]$** или **$Me (25\%; 75\%)$** .
3. **Статистические показатели:** В тексте или таблице необходимо привести:
 1. Объём выборки (n);
 2. Значение эмпирической статистики (T или Z -критерий при аппроксимации);
 3. Точный уровень значимости (p). Точное значение (например, $p = 0,024$) всегда предпочтительнее размытого $p < 0,05$.

Запись результата: M , IQR , Z (T), p -value

Критерий Вилкоксона

«На фоне проводимой терапии у пациентов с артериальной гипертензией отмечено статистически значимое снижение уровня высокочувствительного С-реактивного белка (вч-СРБ). Исходно концентрация вч-СРБ составляла **3,4 [2,1; 4,8] мг/л**, а через 14 дней лечения снизилась до **1,9 [1,2; 2,6] мг/л** ($T = 42$, $n = 28$, $p = 0,012$, Т-критерий Вилкоксона). Полученные данные свидетельствуют об уменьшении выраженности системного воспалительного ответа».

Запись результата: M, IQR, Z (T), p-value

Критерий Вилкоксона

«На фоне проводимой терапии у пациентов с артериальной гипертензией отмечено статистически значимое снижение уровня высокочувствительного С-реактивного белка (вч-СРБ). Исходно концентрация вч-СРБ составляла **3,4 [2,1; 4,8] мг/л**, а через 14 дней лечения снизилась до **1,5 [0,8; 2,2] мг/л** (критерий Вилкоксона). При уменьшении выраженности

Если в вашей выборке более 30 человек, современные статистические программы (SPSS, Statistica, R) рассчитывают Z-статистику (нормальную аппроксимацию). В журналах это оформляется так:

«Оценка интенсивности болевого синдрома по ВАШ (визуально-аналоговой шкале) у прооперированных больных в первые и третьи сутки показала выраженную положительную динамику. Медиана боли в 1-е сутки составила **6,0 (5,0; 7,0) баллов**, а к 3-м суткам значимо уменьшилась до **3,0 (2,0; 4,0) балла** ($Z = -4,12, p < 0,001$, T-критерий Вилкоксона, $n = 54$).»

Запись результата: M, IQR, Z (T), p-value

Критерий Вилкоксона

«На фоне проводимой терапии у пациентов с артериальной гипертензией отмечено статистически значимое снижение уровня высокочувствительного С-реактивного белка (вч-СРБ). Исходно концентрация вч-СРБ составляла **3,4 [2,1; 4,8] мг/л**, а через 14 дней лечения снизилась до **1,9 [1,1; 2,6] мг/л**»

критерий
уменьшения

Показатель	До лечения, <i>Me</i> [<i>Q</i> ₁ ; <i>Q</i> ₃]	После лечения, <i>Me</i> [<i>Q</i> ₁ ; <i>Q</i> ₃]	Статистика критерия (<i>T</i> или <i>Z</i>)	Значимость, <i>p</i>
Систолическое АД, мм рт. ст.	155,0 [145,0; 165,0]	132,0 [125,0; 138,0]	<i>Z</i> = − 3,89	< 0,001
Диастолическое АД, мм рт. ст.	95,0 [90,0; 100,0]	82,0 [80,0; 85,0]	<i>Z</i> = − 3,12	0,002
ЧСС, уд. в мин.	78,0 [72,0; 84,0]	74,0 [68,0; 78,0]	<i>T</i> = 114,0	0,064

раммы
и). В

и

Запись результата: *M*, *IQR*, *Z* (*T*), *p*-value

Алгоритм выбора метода анализа данных для зависимых выборок



Критерий Фридмана

Критерий Фридмана — это непараметрический статистический тест, который используется для сравнения трех или более зависимых (связанных) выборок по уровню количественного или порядкового признака

Простыми словами, если критерий Вилкоксона сравнивает показатели одной группы в двух точках (например, «До» и «После» лечения), то критерий Фридмана является его прямым расширением и применяется, когда точек измерения три и более (например, «До лечения», «Через 2 недели» и «Через 6 месяцев»).

Требования критерия:

- ✓ **Зависимость выборок:** сравнение происходит внутри одной и той же группы объектов (пациентов, животных, образцов) в разные моменты времени или в разных условиях.
- ✓ **Количество условий:** строго 3 или более групп/точек замера. Если условий всего 2, тест математически сводится к критерию Вилкоксона.
- ✓ **Шкала измерения:** порядковая (баллы шкал, опросников, стадии) или количественная (интервалы, отношения). Нормальное распределение данных не требуется.
- ✓ **Размер выборки (n):** минимально — от 4–5 объектов, однако для адекватной мощности теста и использования стандартного распределения хи-квадрат (χ^2) рекомендуется выборка от 10–15 объектов.
- ✓ **Отсутствие пропусков:** для каждого объекта должны быть заполнены данные во всех точках измерения. Пациенты с утерянными данными на одном из этапов полностью исключаются из анализа.

Запись результата: M, IQR, χ^2 , df, p-value, n

Критерий Фридмана

Пример 2: Оформление в виде таблицы

Если показателей или точек контроля много, результаты сводят в таблицу. Важно помнить: если критерий Фридмана показал общую значимость ($p < 0,05$), далее проводят **попарные post-hoc сравнения** (например, критерием Вилкоксона с поправкой Бонферрони), чтобы указать, между какими именно точками есть различия.

Лабораторный показатель	Исходно	3-и сутки	7-е сутки	Критерий Фридмана, χ^2 (df=2)	Значимость, p
Уровень ЛДГ, Ед/л	420 [380; 490]	510 [440; 580]*	390 [340; 430]#	18,45	< 0,001
Уровень КФК, Ед/л	165 [130; 190]	180 [145; 210]	150 [120; 175]	4,12	0,127

* — различие значимо по сравнению с исходным уровнем ($p < 0,017$); # — различие значимо по сравнению с 3-ми сутками ($p < 0,017$) с учетом поправки Бонферрони.

Общий объем выборки: $n = 18$.

Запись результата: M, IQR, χ^2 , df, p-value, n

Алгоритм выбора метода анализа данных для зависимых выборок



Критерий Мак-Нимара

Критерий Мак-Нимара (McNemar's test) — это непараметрический статистический тест, который используется для сравнения двух связанных (зависимых) выборок, когда признак измеряется в номинальной (качественной) шкале и имеет всего два исхода (дихотомический).

Простыми словами: если критерий Вилкоксона нужен для сравнения количественных или балльных показателей «до и после» (например, давление или боль), то критерий Мак-Нимара используется для ответов формата «да/нет», «есть симптом / нет симптома», «здоров/болен» у одних и тех же пациентов в двух разных условиях или точках времени.

Основные требования к данным

1. **Зависимость выборок:** Оценка проводится на одной и той же группе людей/объектов (исследования типа «до — после», «до лечения — после лечения») либо на строго подобранных парах «случай-контроль».
2. **Дихотомическая шкала:** Признак должен иметь строго 2 варианта ответа (0 и 1, Да и Нет, Положительно и Отрицательно).
3. **Четырехпольная таблица сопряженности (2 × 2):** Данные сводятся в матрицу, где анализируются не просто исходы, а *траектория изменений* каждого пациента.

Запись результата: χ^2 , df, p-value

Критерий Формула расчета статистики хи-квадрат (χ^2) Мак-Немара:

Клинический пример: оценка эффекта препарата

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

Подсчет критерия Мак-Немара

Для расчета мы берем только изменившиеся результаты (ячейки **b** и **c**):

- $b = 5$ (ухудшение)
- $c = 65$ (улучшение)

Подставляем значения в формулу:

$$\chi^2 = \frac{(b - c)^2}{b + c} = \frac{(5 - 65)^2}{5 + 65} = \frac{(-60)^2}{70} = \frac{3600}{70} \approx 51.43$$

...

Итого ПОСЛЕ

75

25

N = 100

Запись результата: M, IQR, χ^2 , df, p-value, n

Критерий Мак-Нимара

«До начала
субфебриль
62,5% (20/3
сохранялся
помощью кр
значимое сн
14,06, df = 1,
эффективно

Динамика переносимости физической нагрузки у пациентов после реабилитации (n = 45)

	После реабилитации: Одышка ЕСТЬ	После реабилитации: Одышки НЕТ	Всего исходно
-			
Исходно: Одышка ЕСТЬ	12 (пациент не улучшился)	22 (пациент улучшился)	34
Исходно: Одышки НЕТ	2 (пациенту стало хуже)	9 (пациент остался в норме)	11
Всего на выходе	14	31	45

Примечание к таблице: Анализ изменений с помощью критерия Мак-Нимара выявил статистически значимое преобладание пациентов с положительной динамикой над ухудшившимися ($\chi^2 = 15,04$, $df = 1$, $p < 0,001$).

Запись результата: χ^2 , df, p-value

Алгоритм выбора метода анализа данных для зависимых выборок



Критерий Кохрейна

Критерий Q Кохрейна (Cochran's Q test) — это непараметрический статистический тест, который используется для сравнения трех или более зависимых (связанных) выборок, когда признак измеряется в номинальной (дихотомической) шкале с двумя исходами (например, «да/нет», «успех/неуспех», «есть симптом / нет симптома»).

Основные требования к данным

1. **Зависимость выборок:** Исследование проводится на одной и той же группе объектов (пациентов) в разные периоды времени или в разных условиях.
2. **Тип данных:** Строго бинарный/дихотомический код. Обычно в матрицу вносятся цифры: **1** (наличие признака / успех) и **0** (отсутствие / неуспех).
3. **Количество условий:** Три или более повторных измерений (контрольных точек).
4. **Размер выборки:** Процедура точна при условии, что произведение количества объектов (N) на количество точек измерения (k) достаточно велико (желательно $N \times k \geq 24$).
5. **Отсутствие пропусков:** Каждый пациент должен пройти все этапы контроля. Если данные по какому-то визиту утеряны, строка этого пациента полностью удаляется из анализа.

Запись результата: Q , df , p -value, n

Критерий Мак-Нимара

«В ходе исследования оценивалась динамика частоты выявления изжоги у (исходно, непараметрические регистрировались снизилась 10,5% (4/40) высокозначительно подтверждено»

Пример оформления таблицы в статье

Динамика частоты выявления кашля у пациентов на этапах терапии (n = 40)

Симптом	Исходно, абс. (%)	14-й день, абс. (%)	28-й день, абс. (%)	Критерий Q Кохрейна	Значимость, p
Наличие кашля	34 (85,0%)	18 (45,0%)*	6 (15,0%)* #	Q = 33,82 (df = 2)	< 0,001

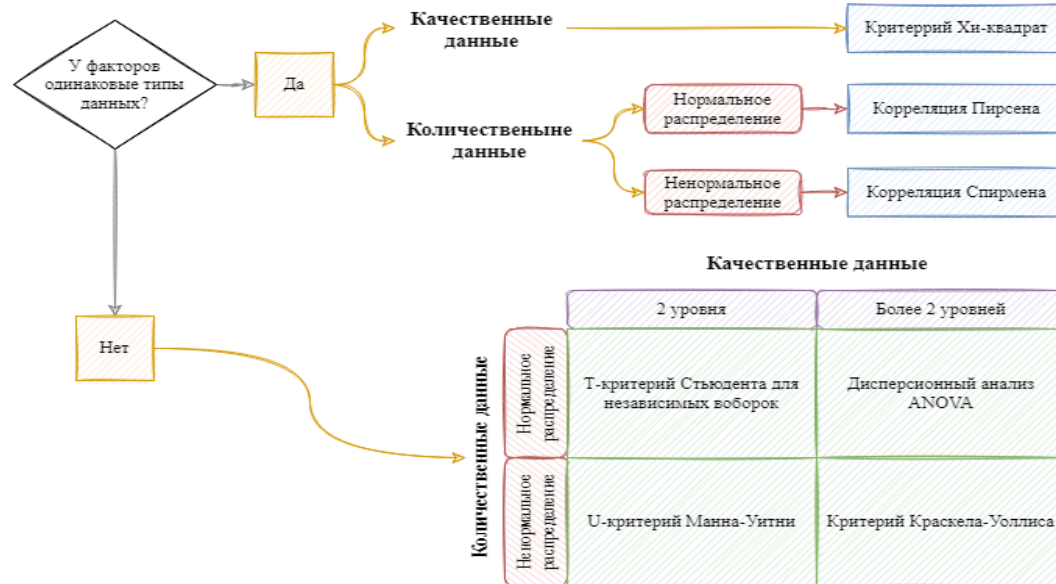
Примечание к таблице:

- различия статистически значимы по сравнению с исходным уровнем ($p < 0,017$);
- # — различия статистически значимы по сравнению с 14-м днем терапии ($p < 0,017$).

Попарные *post-hoc* сравнения выполнены с помощью критерия Мак-Нимара с поправкой Бонферрони для множественных сравнений (критический уровень значимости $\alpha = 0,05/3 = 0,017$).

Запись результата: Q, df, p-value, n

Алгоритм выбора метода анализа данных для независимых выборок



Алгоритм выбора метода анализа данных для зависимых выборок

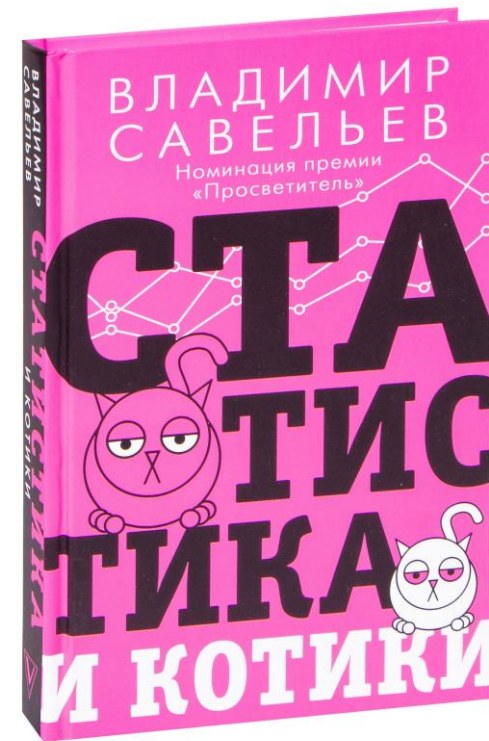
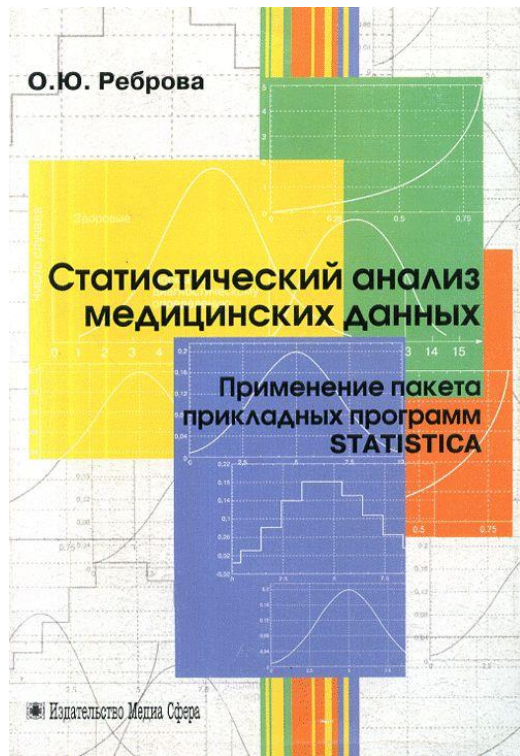


Алгоритм анализа данных

1. Сбор
2. Обработка
3. Распределение
4. Описательная статистика
5. Сравнительный анализ
6. Интерпретация
7. Вывод

NB: Не запоминайте критерии – научитесь проходить алгоритм

Что почитать/посмотреть



stepik Каталог ▾ Преподавание

Поиск...

Русский ▾

Основы статистики

Курс знакомит слушателей с основными понятиями и методами математической статистики. В течение трех недель мы рассмотрим наиболее широко используемые статистические методы и принципы, стоящие за ними. Полученных знаний будет достаточно для решения широкого круга задач, возникающих в рамках исследовательской работы.

Начать просмотр

Начальный уровень 3-4 часа в неделю Сертификат Stepik

★★★★★ 4.9 302 458 учащихся 5 682 отзыва

Что почитать/посмотреть



Спасибо за внимание!

МЕДИЦИНСКАЯ СТАТИСТИКА: ОТ ИССЛЕДОВАТЕЛЬСКОГО ВОПРОСА К СТАТИСТИЧЕСКОМУ ВЫВОДУ

Доцент кафедры хирургии и эндоскопии
ИПК и ПКЗ УО БГМУ
Дыбов Олег Геннадьевич



2026

