

СТАТИСТИЧЕСКИЙ АНАЛИЗ НАУЧНЫХ ИССЛЕДОВАНИЙ ДЛЯ НАЧИНАЮЩИХ

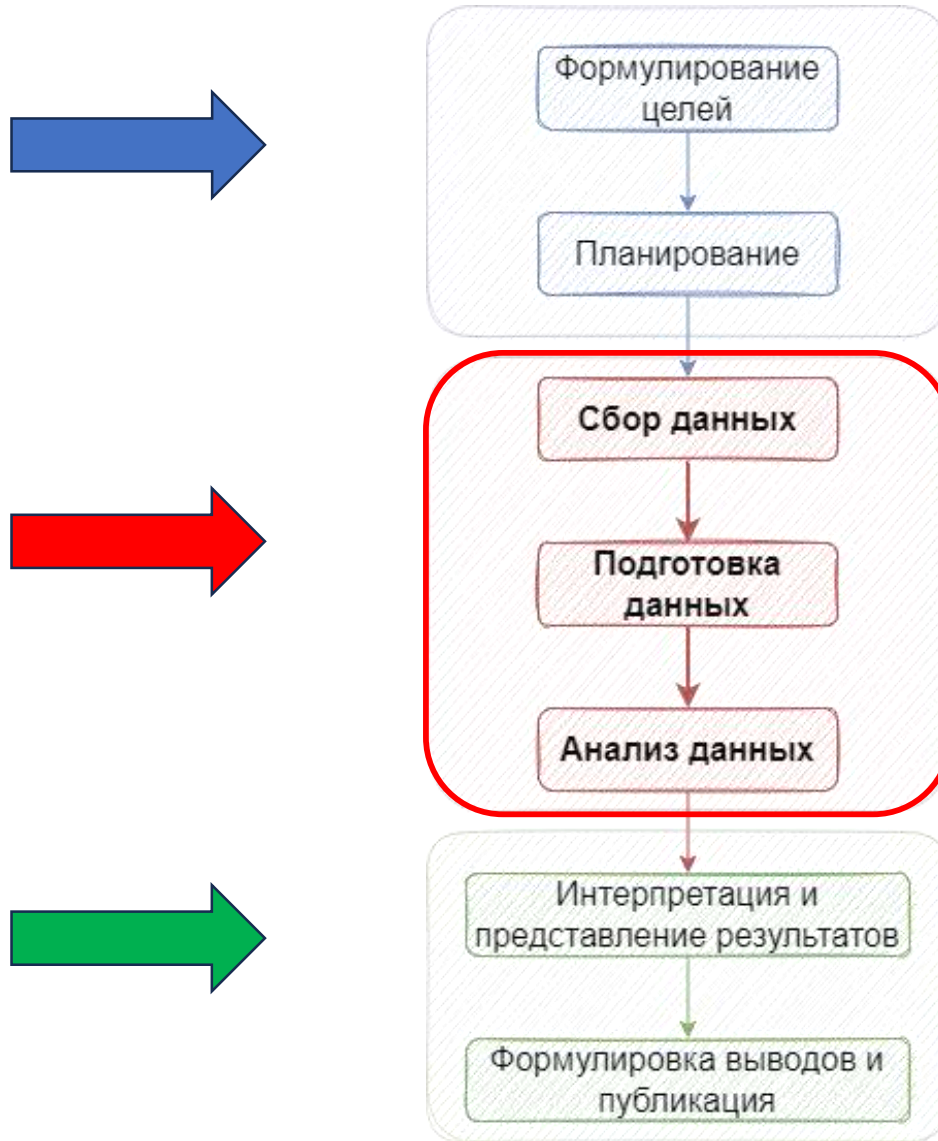
Старший преподаватель кафедры хирургии и эндоскопии
ИПКиПКЗ УО БГМУ
Дыбов Олег Геннадьевич



2025 г.

	Экспериментальное ЛС		?
	Контрольная группа	Экспериментальная группа	
Рост (см)	180	170	?

Этапы исследования

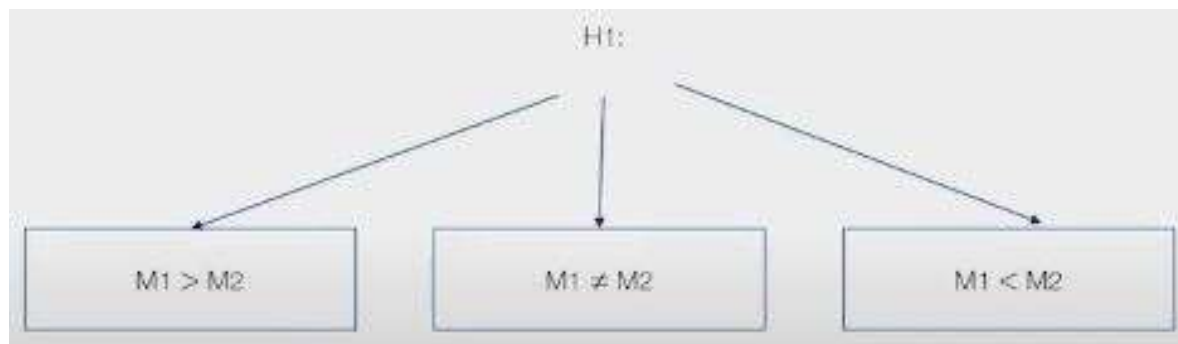


Гипотезы

- *Нулевая гипотеза H_0* , которая указывает **на отсутствие связи/различий** между исследуемыми группами (процесс происходит случайно).

$$H_0: M_1 = M_2$$

- *Альтернативная гипотеза H_1* , которая указывает **на наличие связи/различий** между группами (процессы происходят НЕ случайно).



Уровень значимости

p (Sig.) (Significance – наблюдаемый *уровень значимости*) – это вероятность допустить ошибку, утверждая, что фактор влияет на отклик.



Источники информации (планирование)



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](#).



Сбор данных

Тестовая форма

* Обязательно

Ваш пол? *

☐ Мужской

☒ Женский

☐ Другой _____

ОТПРАВИТЬ

Никогда не используйте формы Google для передачи тарелок.

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



Google Forms

СБОР ДАННЫХ?

СБОР ДАННЫХ

GOOGLE FORMS 12 видео

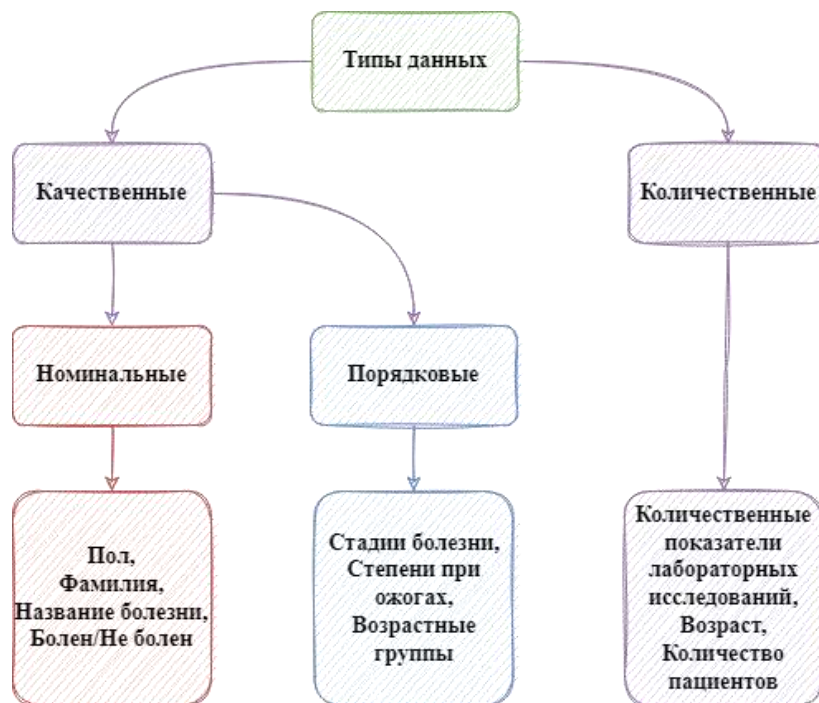
<https://www.youtube.com/@BilyalKhassenov>



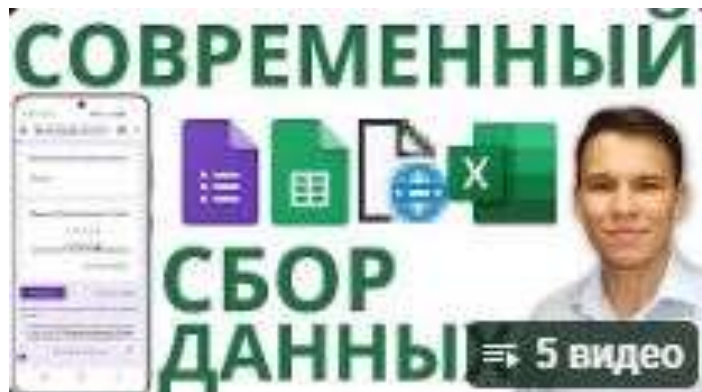
Типы данных



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



Подготовка данных

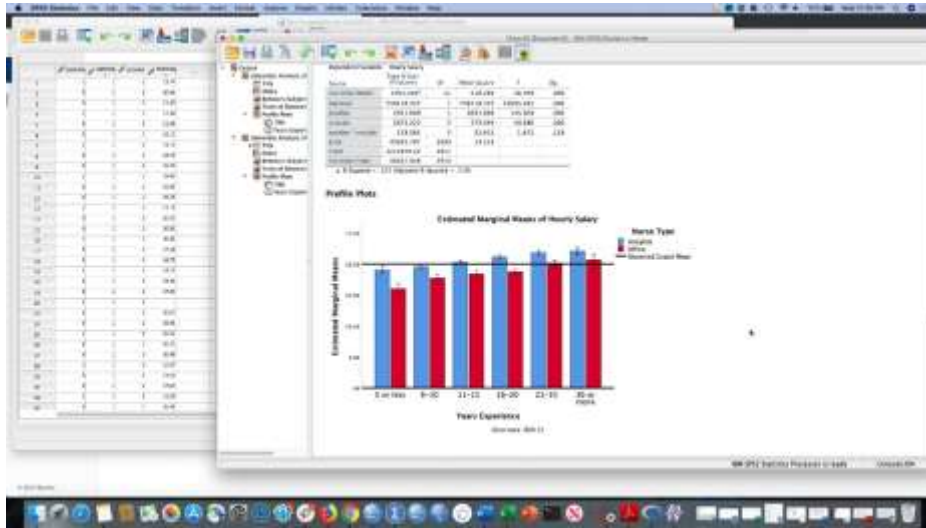


<https://www.youtube.com/@BilyalKhassenov>

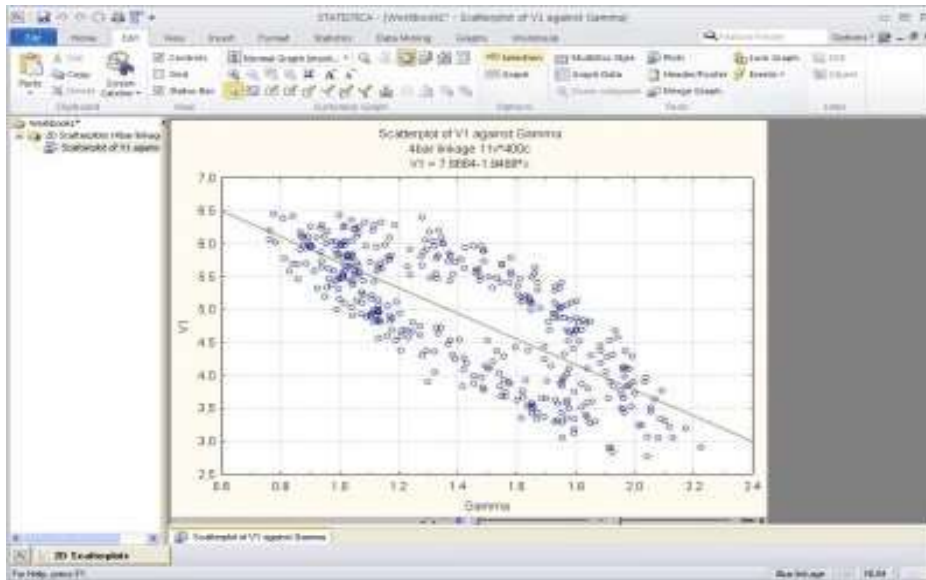


Remove to improve the **data tables** edition

Платформы для анализа данных (платные)

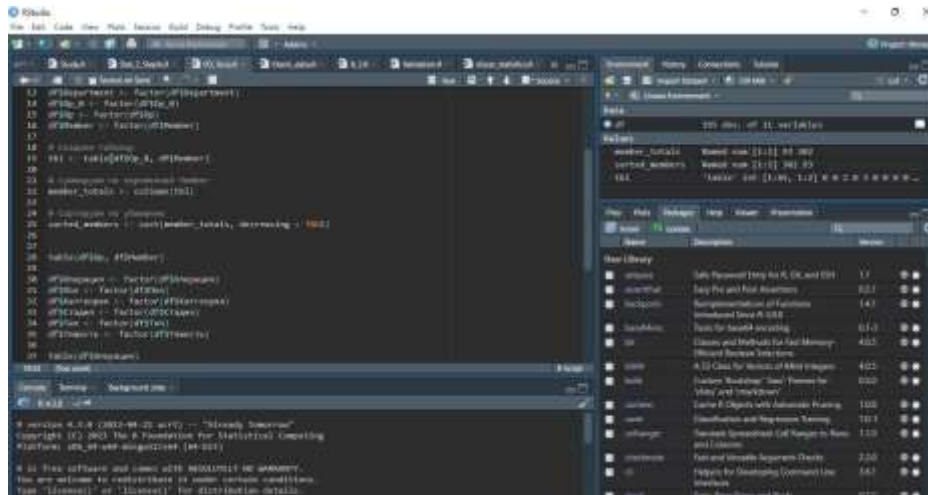


IBM SPSS Statistics

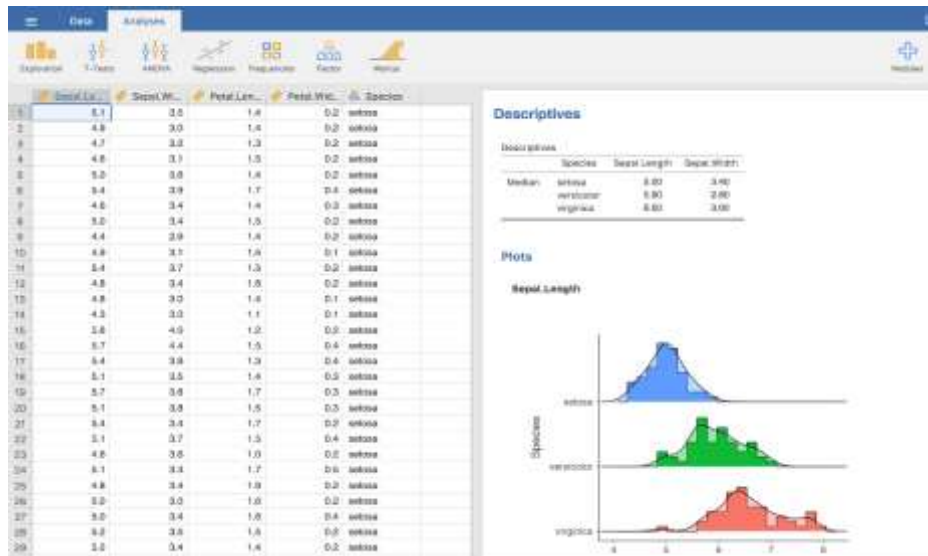


StatSoft STATISTICA

Платформы для анализа данных (бесплатные)



R-studio

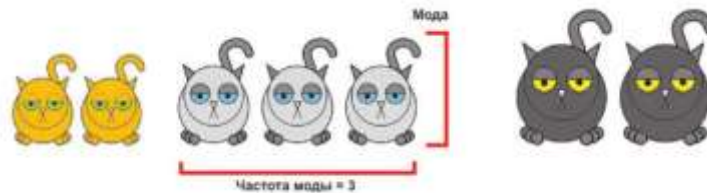


jamovi

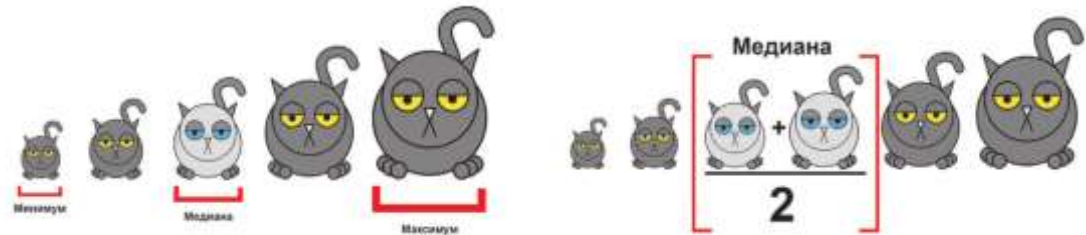


Меры центральной тенденции

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



- **Медиана** (Median - Me) – середина упорядоченного ряда значений.

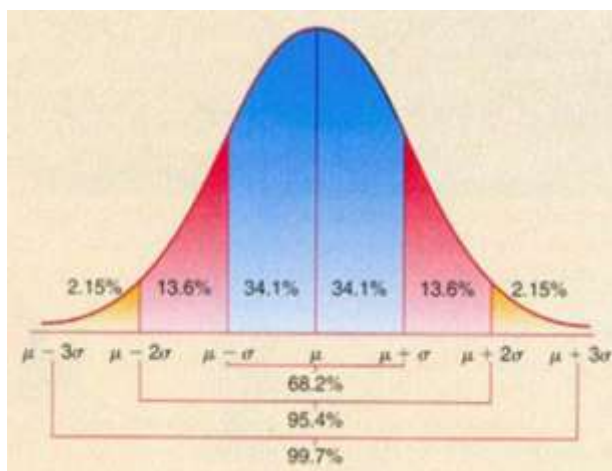


- **Среднее арифметическое** (Mean - μ) – сумма значений, делённое на их количество



Меры центральной тенденции

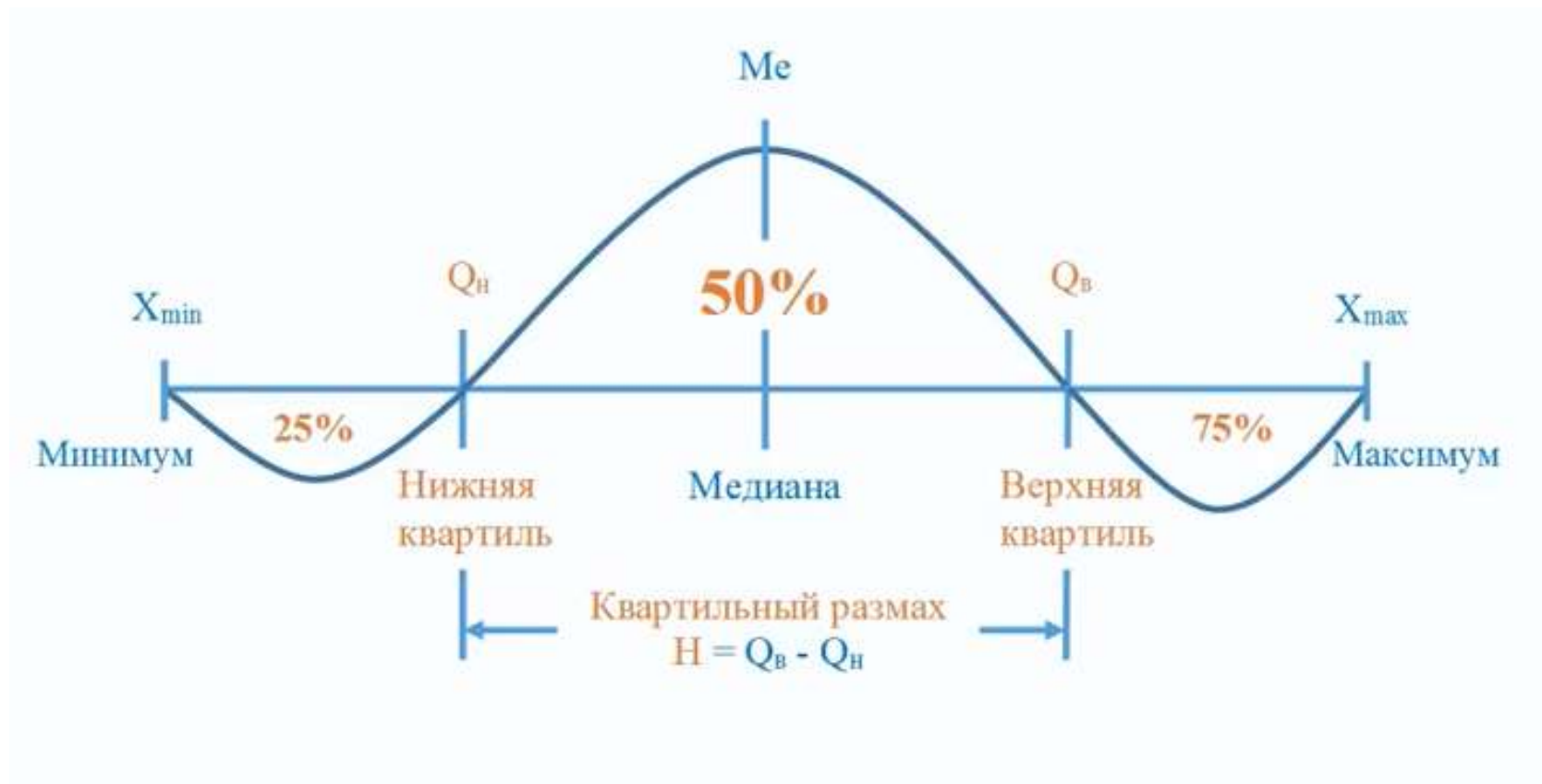
Доверительным интервал среднего (CI, англ. – confidence interval for mean). Он задаёт область вокруг среднего с определённой вероятностью (как правило 95% - $\mu \pm (1,96 \times \text{SEM})$), в котором содержится «истинное» среднее выборки.



Форма записи: μ ; 95% $\text{CI}_{\min} - \text{CI}_{\max}$, где SEM – стандартная ошибка среднего

Пример: возраст - 39,3; 95% ДИ 35,9-42,7

Меры центральной тенденции



Форма записи: Me [Q_н; Q_в]

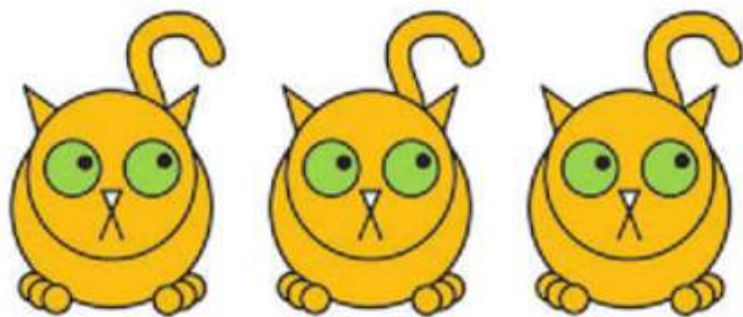
Пример: индекс массы тела - 24,2 [19,7; 25,8]

Меры центральной тенденции



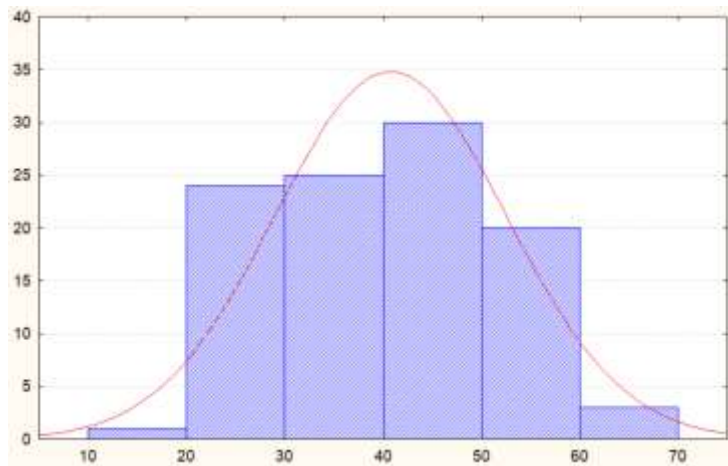
Форма записи: $Me [Q_n; Q_v]$

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

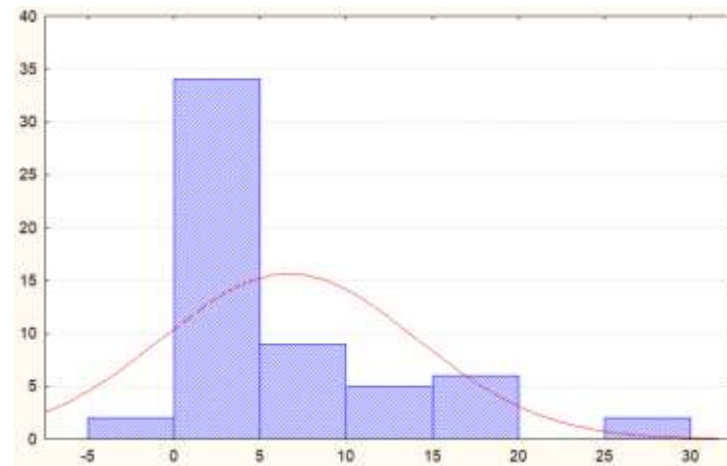


Выброс

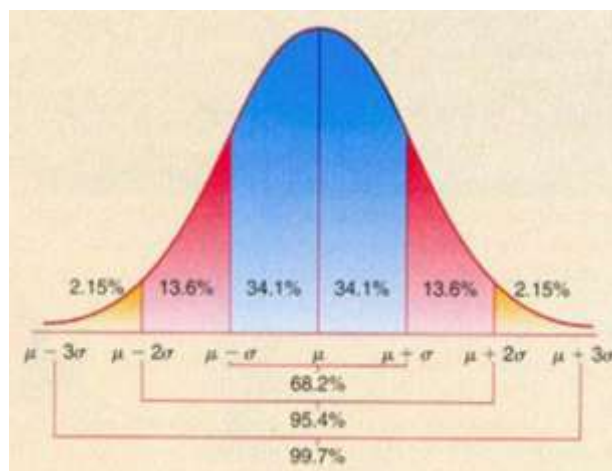
Распределение



Нормальное



Ненормальное



Распределение (тесты)

- ✓ *Колмогорова-Смирнова* (Kolmogorov-Smirnov) – при $p < 0.2$ распределение ненормально. При количестве наблюдений > 60 .
- ✓ *Шapiro-Уилка* (Shapiro-Wilk's) – при $p < 0.05$ распределение ненормально. При количестве наблюдений < 60 .

Сравнение средних

Нормальное распределение → Т-критерий Стьюдента (t-Student) проверяет равенство средних значений количественных данных в двух подгруппах.

Требования критерия:

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

Сравнение средних

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

Требования критерия:

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

		Экспериментальное ЛС		p
		Экспериментальная группа	Контрольная группа	
Рост (см)	Иssl. 1	180	170	$p_{t-Student} < 0.05$
	Иssl. 2	160	180	$p_{Mann-Whitney} > 0.05$

Интерпретация

- ✓ Исследование № 1 → в экспериментальной группе статистически значимо ($p_{t-Student} < 0.05$) средний рост «пациента» был на 10 см больше/выше, чем в контрольной
- ✓ Исследование № 2 → в экспериментальной группе средний рост «пациента» был на 20 см меньше/ниже, чем в контрольной, однако разница не является статистически значимой ($p_{Mann-Whitney} > 0.05$)

		Экспериментальное ЛС		p
		Экспериментальная группа	Контрольная группа	
Рост (см)	Иssl. 1	180	170	$p_{t-Student} < 0.05$
	Иssl. 2	160	180	$p_{Mann-Whitney} > 0.05$

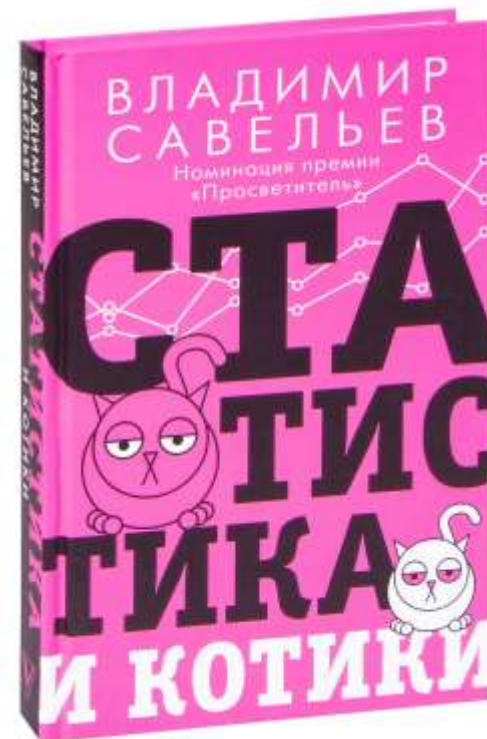
Вывод

- ✓ Исследование № 1 → экспериментальное ЛС статистически значимо влияло на рост испытуемых
- ✓ Исследование № 2 → экспериментальное ЛС статистически значимо не влияло на рост испытуемых

Алгоритм анализа данных

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

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

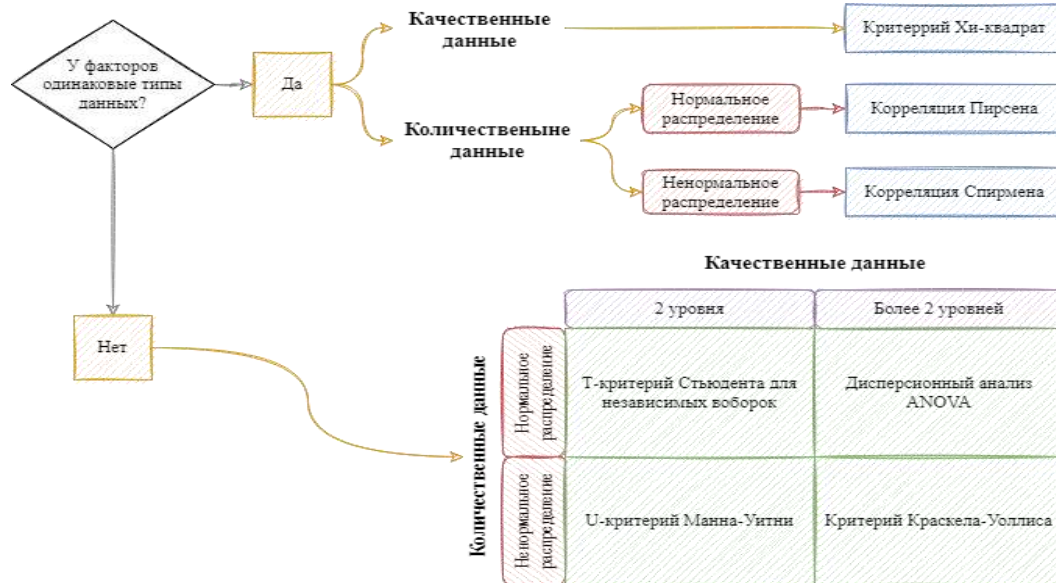
Начальный уровень 3-4 часа в неделю Сертификат Stepik

★★★★★ 4.9 302 450 учащихся 5 662 отзыва



BCÉ!

Алгоритм выбора метода анализа данных для **независимых** выборок



Алгоритм выбора метода анализа данных для **зависимых** выборок

