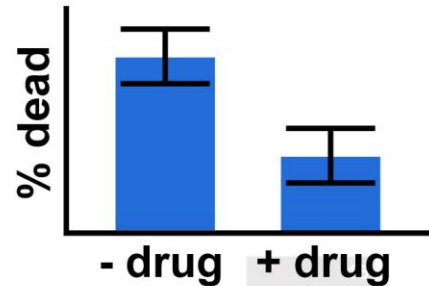


Статистика – наука о государстве (ранее)

Statistical Significance Tests

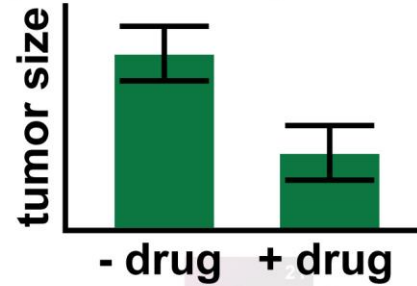
Fisher's Exact Test

x-axis: 2 categories (- or + drug)
y-axis: 2 categories (dead or alive)



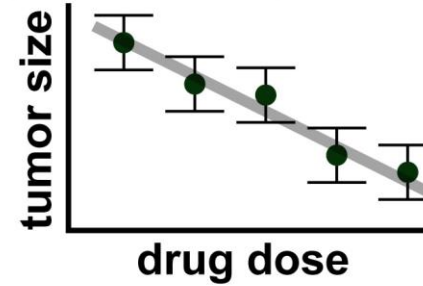
Students T-Test

x-axis: 2 categories (- or + drug)
y-axis: continuous (tumor size)



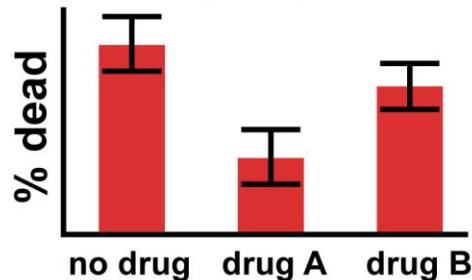
Regression Analysis

x-axis: continuous (drug dose)
y-axis: continuous (tumor size)



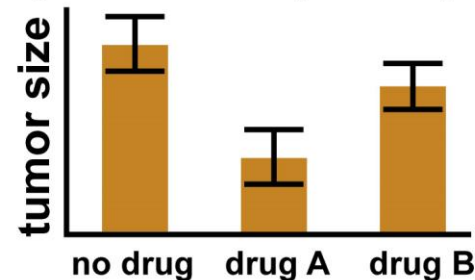
Chi-Square Test

x-axis: >2 categories (drug A, B, etc.)
y-axis: >2 categories (dead or alive)



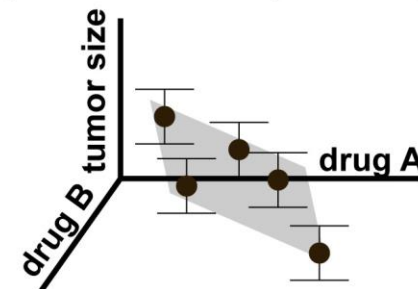
ANOVA

x-axis: >2 categories (drug A, B, etc.)
y-axis: continuous (tumor size)



Multiple Regression

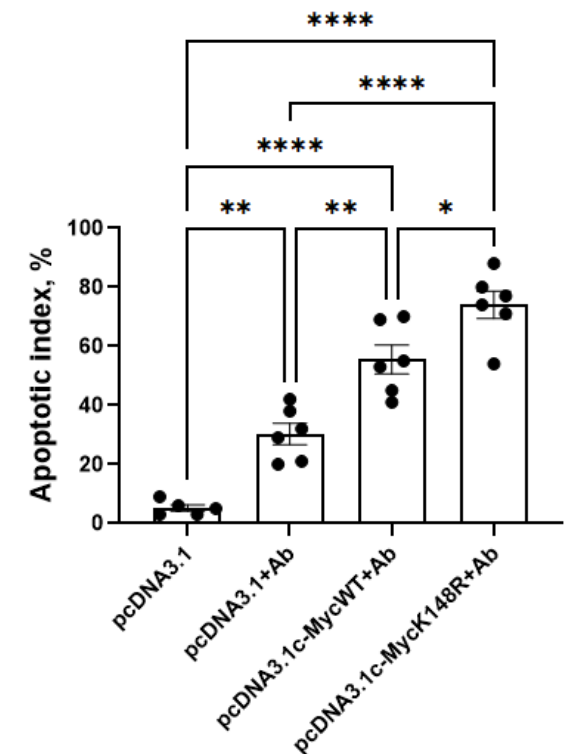
x/z-axis: >2 continuous (drug dose)
y-axis: continuous (tumor size)



Статистическое сравнение средних значений экспериментальных групп и визуализация результатов

- В экспериментальных данных нашей лаборатории преобладают группы непрерывных величин, характеризующих биологические образцы, организованные в табличные колонки
- Группы разными типами воздействия, включая отсутствие воздействия (контроль)
- Как правило, мы хотим знать, насколько значимо влияют экспериментальные манипуляции на интересующую нас величину (обычно какую-то одну), т.е. какова вероятность, что отличие средних значений групп не является случайным (нулевая гипотеза).

	pcDNA3.1	pcDNA3.1+Ab	pcDNA3.1c-MycWT+Ab	pcDNA3.1c-MycK148R+Ab
1	6	29	53	74
2	3	42	45	80
3	3	38	69	54
4	9	20	55	88
5	5	32	41	77
6		21	70	71



Приложения для анализа и визуализации данных

Microsoft Excel, LibreOffice Calc, Google Sheets ...

Statistica, SPSS/PSPP, Origin Pro, Stata ...

SigmaPlot, **GraphPad Prism**

Jamovi, JASP, scatterplot.bar

Mathematica, Matlab, Maple, GNU Octave, Maxima ...

R, Python, Julia ...

GraphPad Prism

Позволяет как проводить статистические тесты, так и строить графики

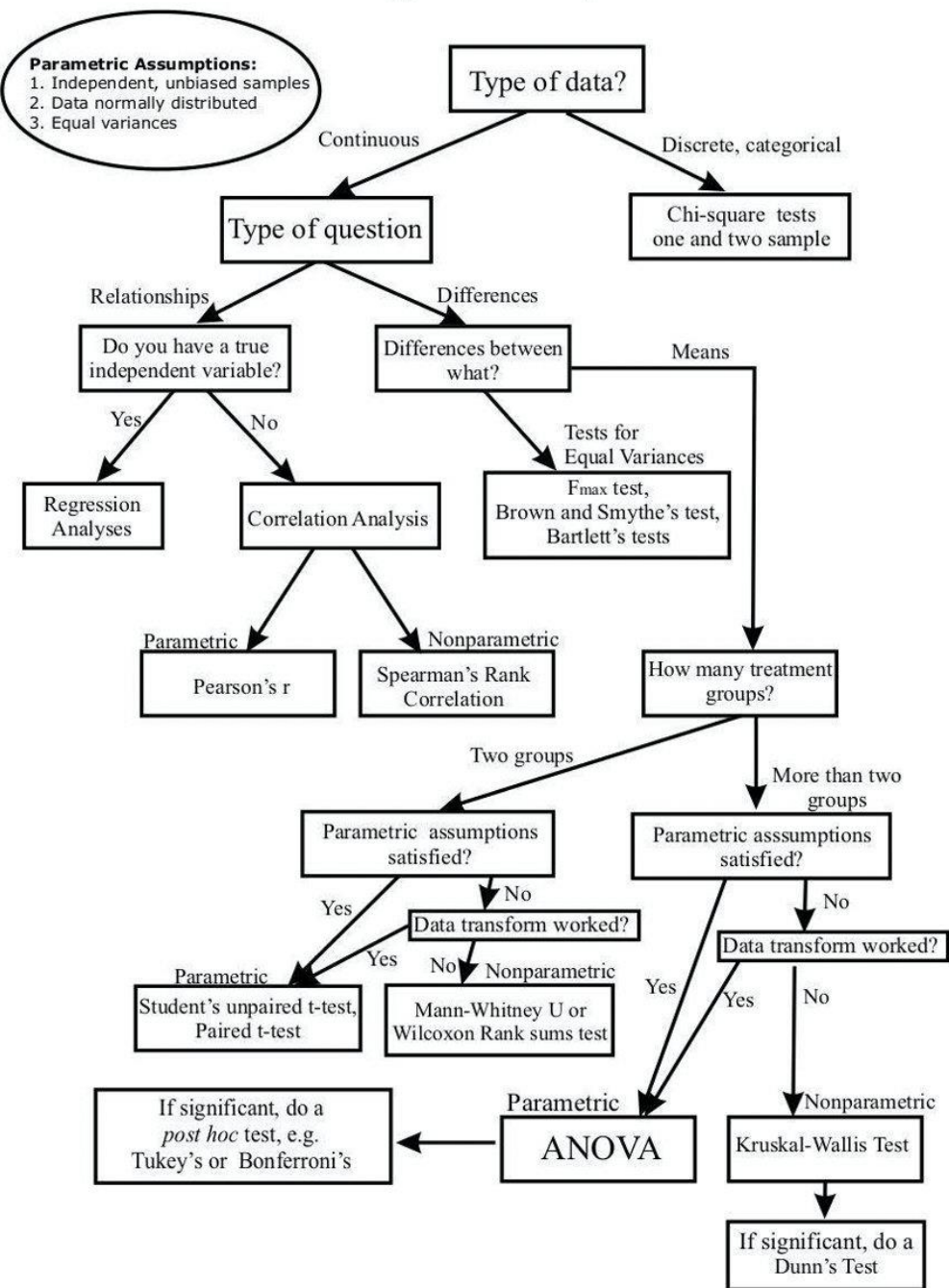
Широкие возможности настройки отображения графиков

Автоматическое добавление сравнений групп (в последних версиях)

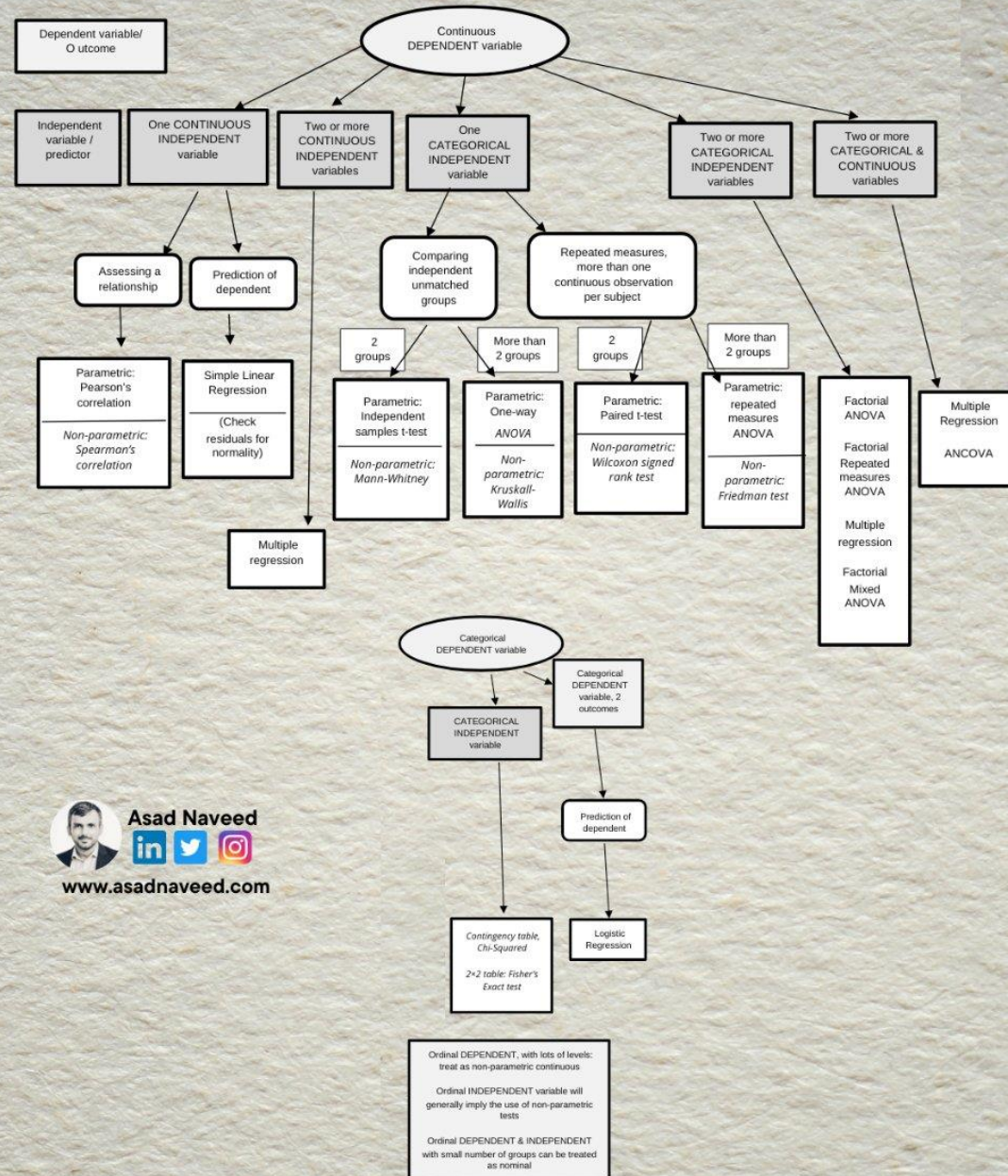
Удобство импорта данных

Главный минус – использование пиратской версии

Flow Chart for Selecting Commonly Used Statistical Tests



HOW TO CHOOSE A STATISTICAL TEST



Главные вопросы для выбора теста

- Сколько групп сравнивается – 2 или больше?
- Нормально ли распределение величин?
- Отличается ли в группах дисперсия/variance (обычно не отличается)

Что такое

- Нормальное распределение
- Дисперсия
- Стандартная ошибка среднего
- Стандартное отклонение
- Тест Стьюдента
- One-way ANOVA
- Апостериорный критерий (post hoc test)

Нормальное (Гауссово) распределение

$$f = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

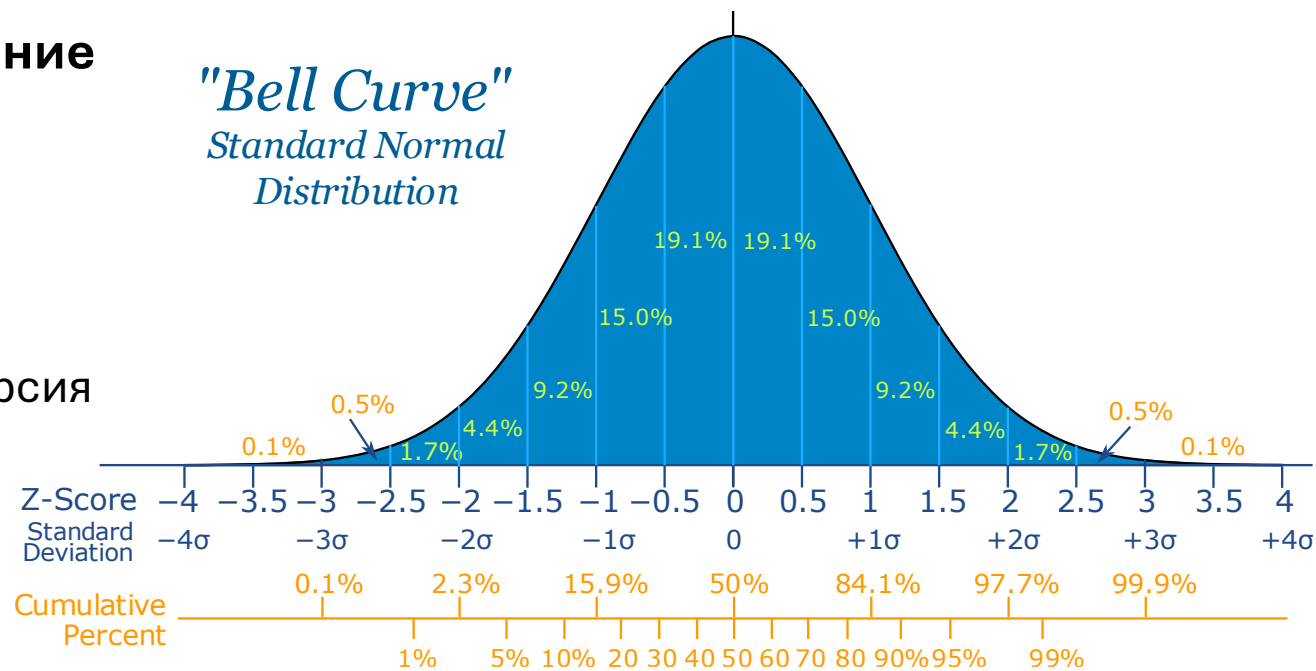
σ^2 - квадрат стандартного отклонения, или дисперсия
 μ - среднее, или мат. ожидание

СТАНДАРТНОЕ ОТКЛОНЕНИЕ

$$\sigma = \sqrt{\frac{\sum (x_i - M)^2}{n - 1}}$$

x_i - каждое наблюдаемое значение признака;
 M - среднее арифметическое;
 n - количество наблюдений (объем выборки).

"Bell Curve"
Standard Normal
Distribution



Стандартная ошибка среднего

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

- Нулевая гипотеза – исходит из того, что мат. ожидания среднего значения величины в рассматриваемых выборках одинаковы
- $P < 0,05$ – вероятность того, что нулевая гипотеза верна, не превышает 5%
- T-test Стьюдента и One-way ANOVA (Analysis of variance, дисперсионный анализ) проверяют нулевую гипотезу в предположении независимости, нормальности и равенства дисперсий выборок
- В случае двух выборок One-way ANOVA сводится к t-тесту. Следует всегда использовать two-tailed t-test
- Апостериорные критерии (post-hoc) тесты после One-way ANOVA используются для попарного сравнения конкретных, интересующие нас групп. Тест Даннетта сравнивает каждую из групп с назначенным контролем. Тест Тьюки сравнивает все выборки попарно.
- Мощность выборки – определяет вероятность отвергнуть нулевую гипотезу, если она неверна (т.е. если эффект реально есть). Используется для определения минимального объема выборки (основное в нашем случае) и других параметров (например, размера эффекта).
- Выбросы (Outliers) – экспериментальные величины, сильно отклоняющиеся от среднего значения по случайным причинам. Обычно подлежат исключению. Для определения выбросов лучше использовать специальные тесты, а не выявлять выбросы «на глаз».

Случай
отсутствия
нормального
распределения

2 группы – тест Манна-Уитни

3+ группы – тест
Краскелла-Уоллиса

Тест Данна (post hoc)

Другие типы данных

Повторные измерения
величины (динамика) –
RM ANOVA

Сравнение двух
величин – Two-way
ANOVA (ANOVA on ranks)

Демонстрация
работы в
GraphPad Prism

