Мощность статистического теста. Размер эффекта. Трансформация данных.

Тестирование статистических гипотез
(повторение)

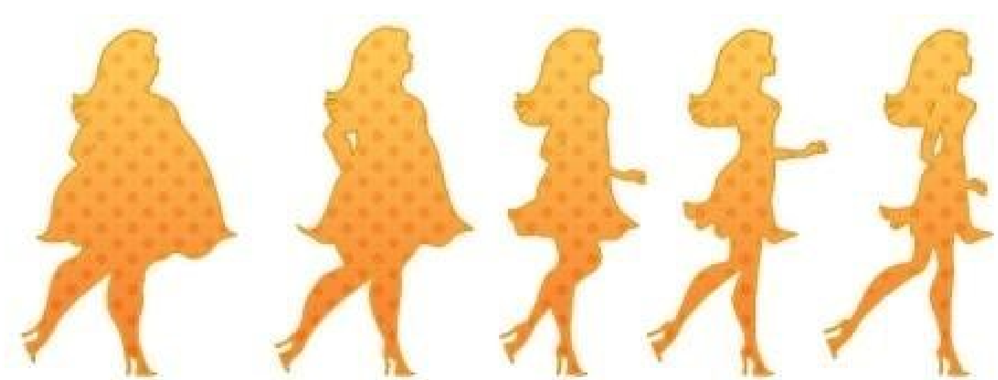
1. Формулируем H_0 и H_1
2. Подбираем статистику критерия (определяем критическую область, критические значения)
3. На основе выборки считаем значение статистики и сравниваем с критическими значениями; получаем точное значение Р.
4. Отвергаем или не отвергаем H_0

В этой процедуре есть трудности

Критика тестирования гипотез

1. Результат теста (точное *P* и вывод) **зависит от размера выборки**: чем больше N, тем меньшие различия могут быть признаны «достоверными».

Dr. Nostat изобрёл гипнотическое устройство для похудания; если его положить под подушку, за месяц оно достоверно понижает массу тела на 1 г (доктор испытывал устройство на выборке N=6000).



Размер эффекта

Критика тестирования гипотез

2. **Нулевая** **гипотеза никогда** **не может быть верна**.

Средние значения в популяциях никогда не могут совпасть совершенно; параметр не может быть в точности равен заданному числу и т.п.

Какой же смысл её тестировать?

самка

самец

Размер эффекта

Критика тестирования гипотез

3. **Уровень значимости** (ошибки 1-го рода) **ничем не оправдан, это просто договорённость**.

Причём *P* нам показывает вероятность получить нашу выборку, если верна H_0.

А на самом деле нас интересует вероятность H_0, если дана наша выборка.

Это не одно и то же! (*хотя в простых случаях этим как раз можно пренебречь*)

⇨ *Без договорённости не обойтись*

⇨ *Байесовские подходы*

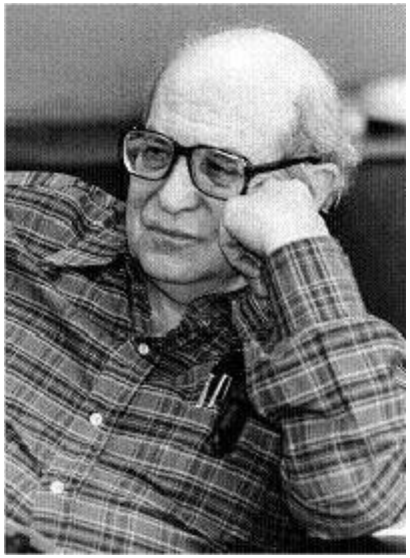
Размер эффекта (effect size)

1. На самом деле всегда есть какой-то **порогowый эффект** (различия между значениями параметров и пр.), значения меньше которого **не имеют биологического смысла**.
2. Мы можем оценить **размер эффекта после** проведённого теста (чтобы верно интерпретировать результат и его практический смысл).
3. Мы можем задать размер эффекта, который будет иметь для нас смысл, и исходя из этого **планировать исследование** (оценить размер выборки).

Это очень важный раздел, который многие игнорируют, особенно в отечественных исследованиях!

Размер эффекта

Чтобы как-то оценить научный смысл различий (отклонений от H_0), придуманы специальные **индексы**, **для разных статистических тестов – разные**.



Jacob Cohen
(1923 –1998)

Размер эффекта для сравнения средних значений:

Cohen's d

Одновыборочный тест

$$d = \frac{\left|\bar{X}_{obs} - \mu\right|}{s}$$

Двухвыборочный тест

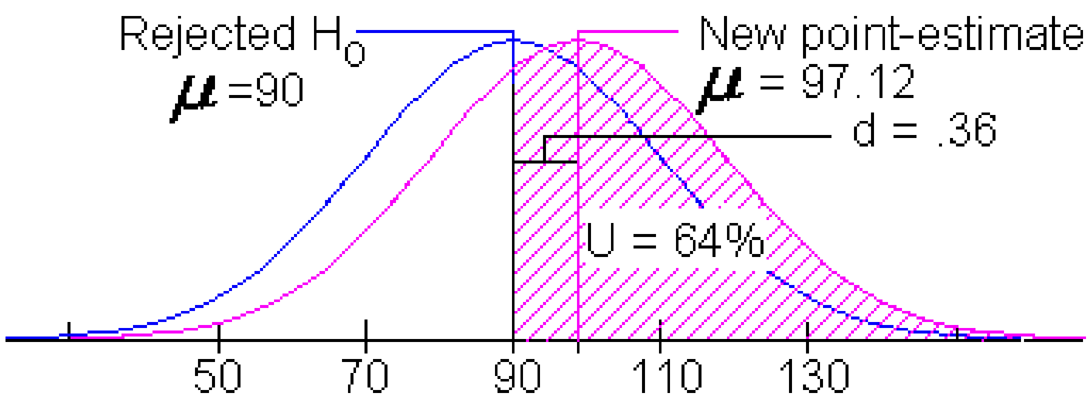
$$d = \frac{\left(\bar{X}_1 - \bar{X}_2\right)_{obs}}{s_{pooled}}$$

Размер эффекта

Просто считаем разницу между средними. Для удобства сравнения, стандартизируем её (делим на SD).

d похоже на t, но <u>не зависит от размера выборки</u>, в отличие от t!

$$d = \frac{\left(\overline{X}_1 - \overline{X}_2\right)_{obs}}{s_{pooled}}$$

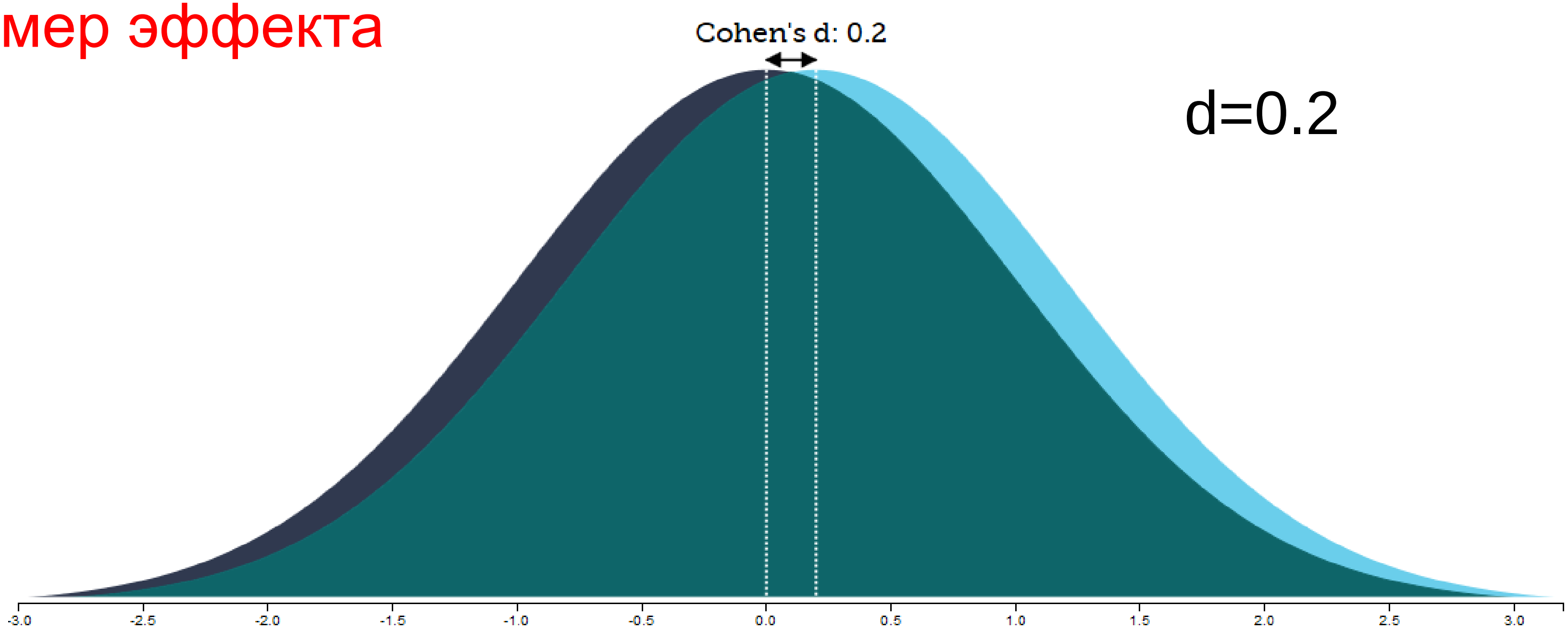


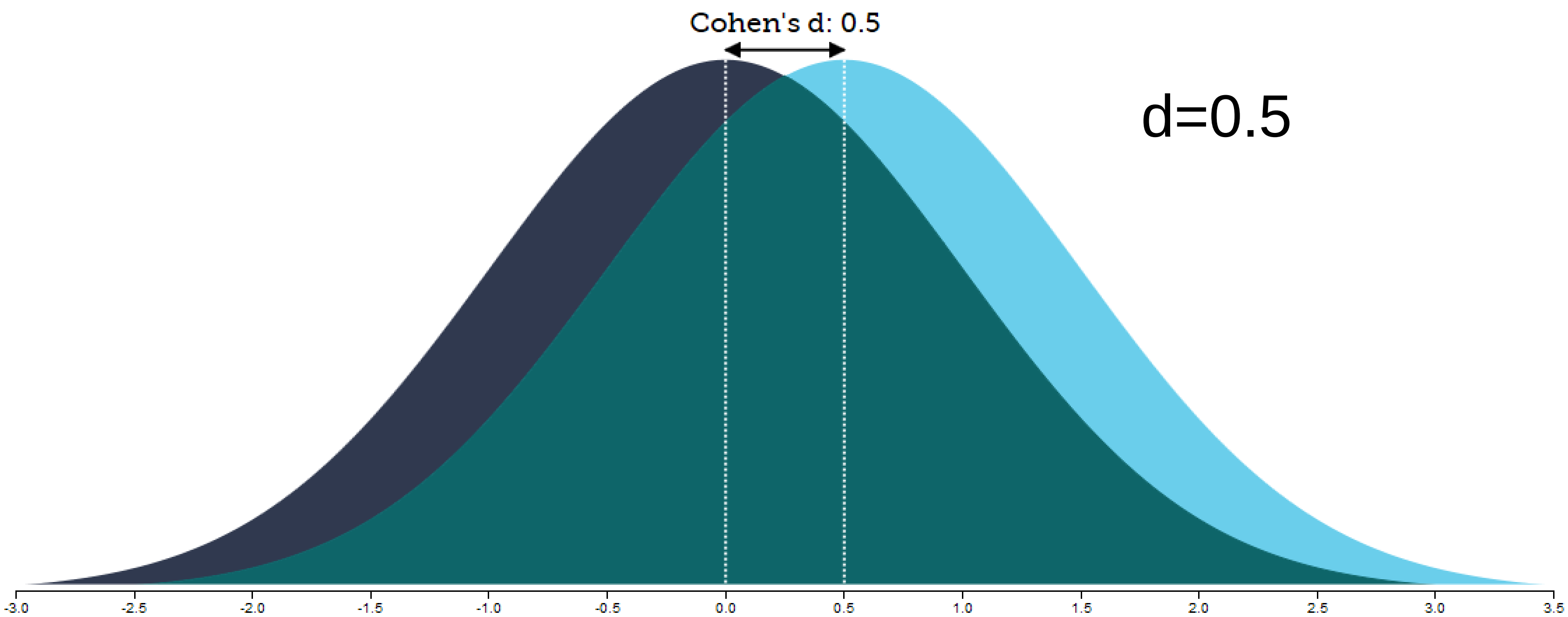
$$t = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{X}_1 - \overline{X}_2}}$$

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{s^2_{pooled}}{n_1} + \frac{s^2_{pooled}}{n_2}}$$

d =0.20 – маленький размер эффекта;
d =0.50 – средний;
d =0.80 - большой; (Cohen, 1988)

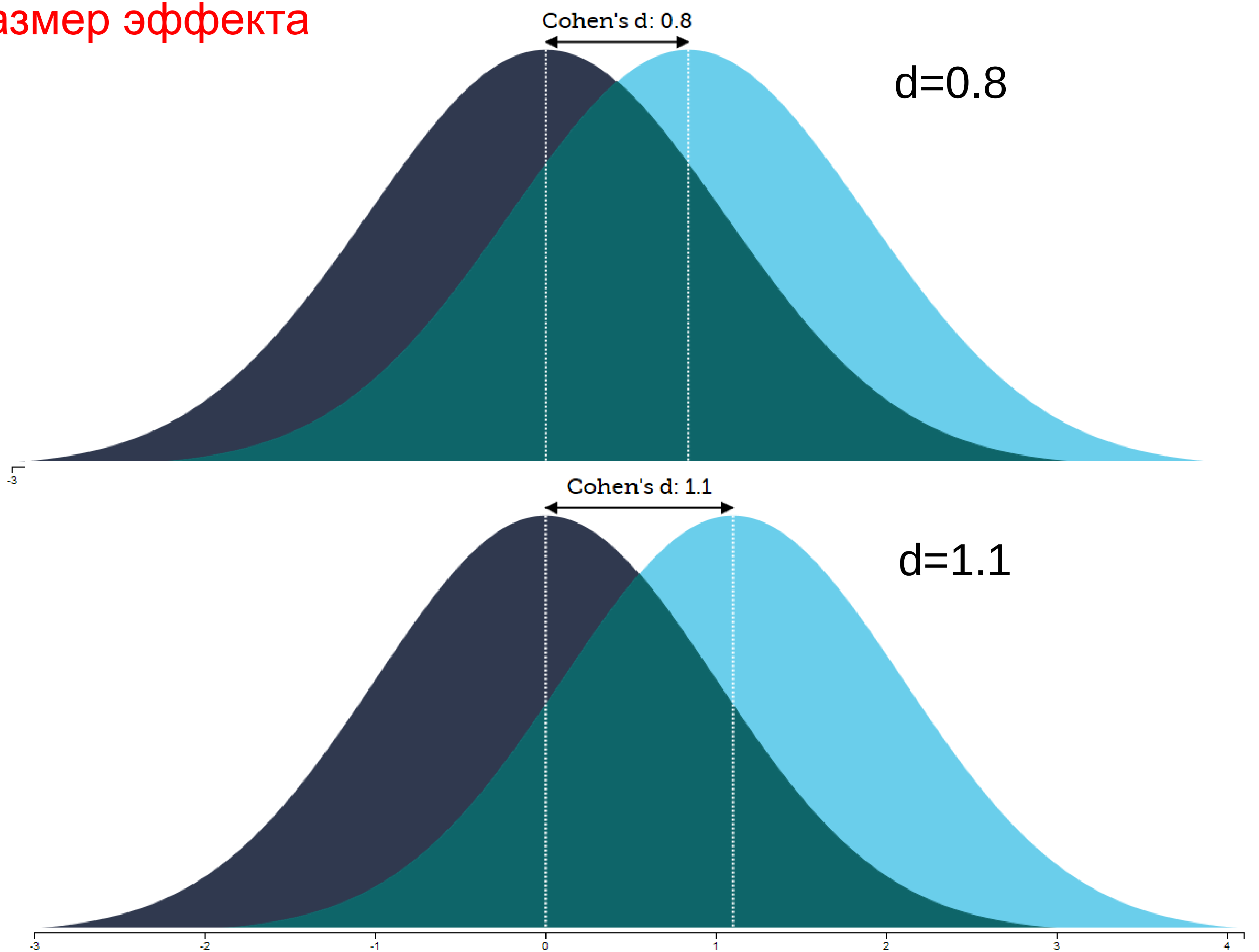
Размер эффекта





Kristoffer Magnusson, http://rpsychologist.com/d3/cohend/

Размер эффекта



Kristoffer Magnusson, http://rpsychologist.com/d3/cohend/

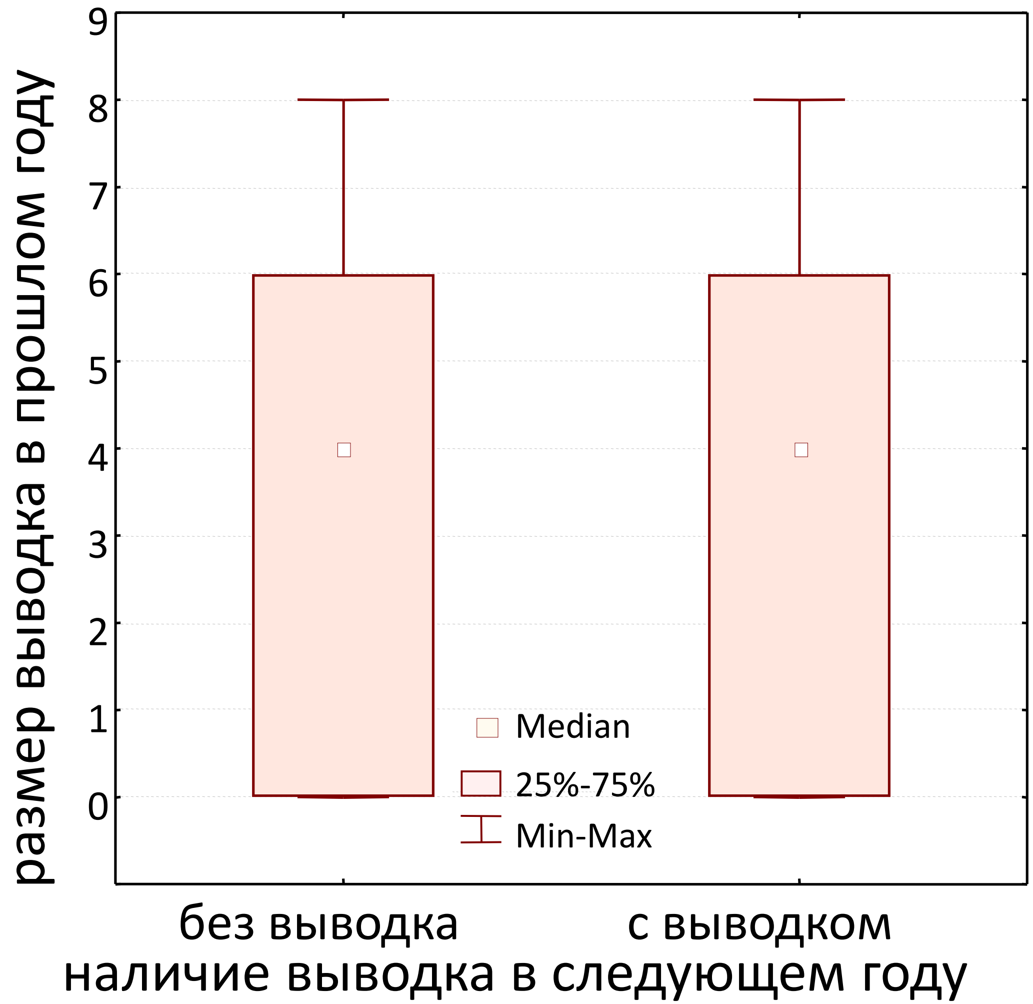
Размер эффекта (ES)

Оценка размера эффекта **ПОСЛЕ тестирования гипотезы**:

1. Очень важна в случае «недостоверного» результата:
   - ✓ если мы не отвергли H_0 (не нашли различий), а ES средний и больше, нельзя говорить и том, что различий действительно нет: это <u>повод провести дополнительное исследование, увеличив N</u>;
   - ✓ если H_0 не отвергли, а ES маленький, можно говорить, что различий, скорее всего, действительно нет. *Часто это единственный способ опубликовать недостоверные результаты!*

Размер эффекта (ES)

Здесь размер эффекта настолько мал, что каждому ясно: даже если выборку увеличить, различие средних значений будет очень мало.

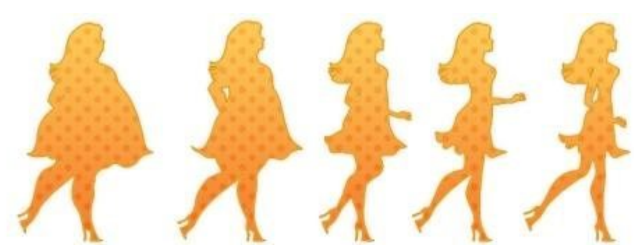


Размер эффекта (ES)

Оценка размера эффекта **ПОСЛЕ тестирования гипотезы**:

2. в случае «достоверного» результата:

- ✓ если мы отвергли H_0 (нашли различия), а ES маленький, надо хорошо подумать, имеют ли такие различия биологическое значение (может, они вообще в пределах ошибки измерения!);
- ✓ если H_0 отвергли, а ES большой, можно говорить, что различия, скорее всего, действительно есть.



В наши дни злополучный Dr. Nostat уже не опубликует свои результаты в приличном журнале.

Размер эффекта (ES)

Оценка размера эффекта **ДО тестирования гипотезы**:

1. Особенно актуальна в экспериментальных исследованиях;
2. Если мы прикинем, какие различия (с учётом разброса) для нас стоят того, чтобы вообще затевать всё это, мы можем рассчитать **примерный размер выборки**, который понадобится, чтобы их обнаружить!
3. Это полезно ещё и тем, что заставляет чётко сформулировать H_0 и обдумать, каким методом будут проанализированы данные.

Однако, для расчёта N нам кое-чего не хватает! Нужно задать мощность теста.

Про фотоловушки

Мощность статистического теста

Мощность - ВЕРОЯТНОСТЬ отвергнуть H_0 в эксперименте, когда H_0 действительно **не верна**.

	Истинное (но неизвестное нам) положение дел	
	Верна H_0	Верна H_1
Мы «приняли» H_0	ПРАВИЛЬНО! 1-α	ОШИБКА 2-го рода β
Мы отвергли H_0	ОШИБКА 1-го рода (уровень значимости) α	ПРАВИЛЬНО! (мощность критерия) 1-β

«Мощный» статистический критерий – такой, при котором эта вероятность высока (например, 80%).

Мощность

Расчёт мощности

Например, масса землероек на острове на самом деле **больше**, чем 90 г. Например, 94 г.
Мощность – вероятность того, что проведённое нами исследование установит этот факт.

H_0: $\mu \leq 90$ г;
H_1 : $\mu > 90$ г

(Односторонний тест – исключительно для удобных картинок!)

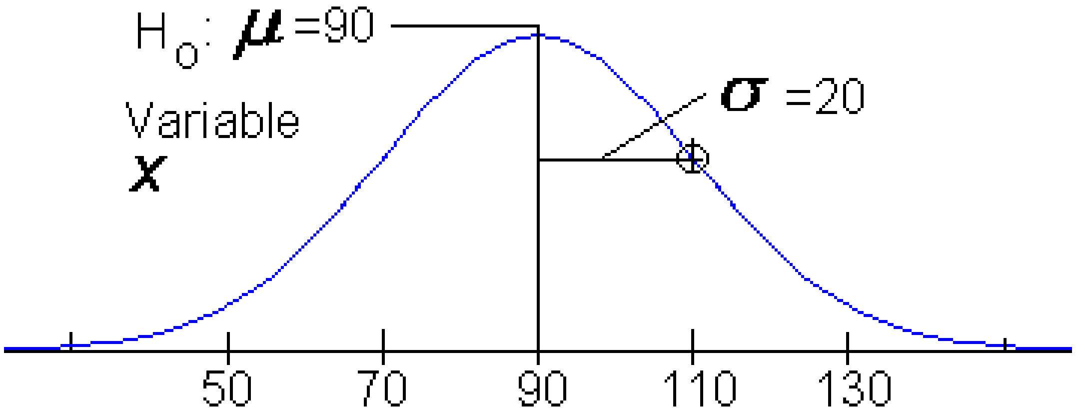
Ошибка 2-го рода + мощность = 1

$$\beta \qquad + (1-\beta) \qquad = 1$$

(это 2 возможных результата теста, ведь в нашем случае H_0 не верна)

Мощность

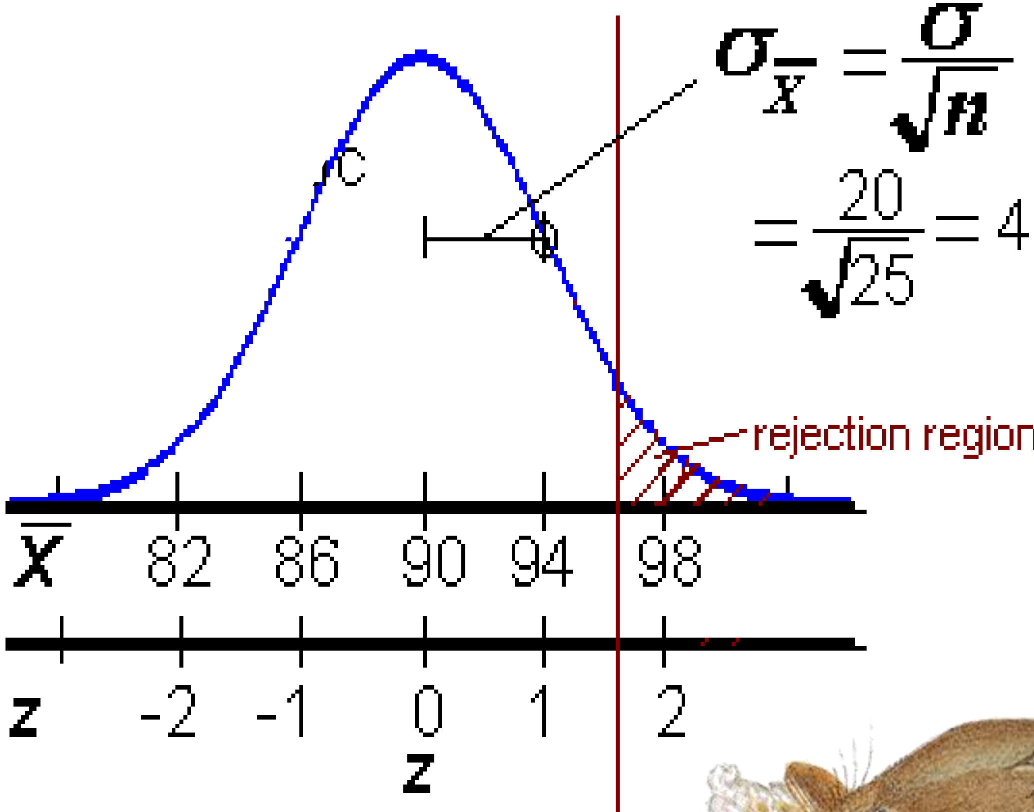
Во всей популяции землероек μ = 90 г.
Пусть «реальное значение» средней массы в популяции на острове μ = 94 г.



Это распределение масс зверьков согласно H_0

Мощность

Нарисуем распределения **выборочных средних** для μ = 90 и μ = 94 (стандартное отклонение σ = 20).

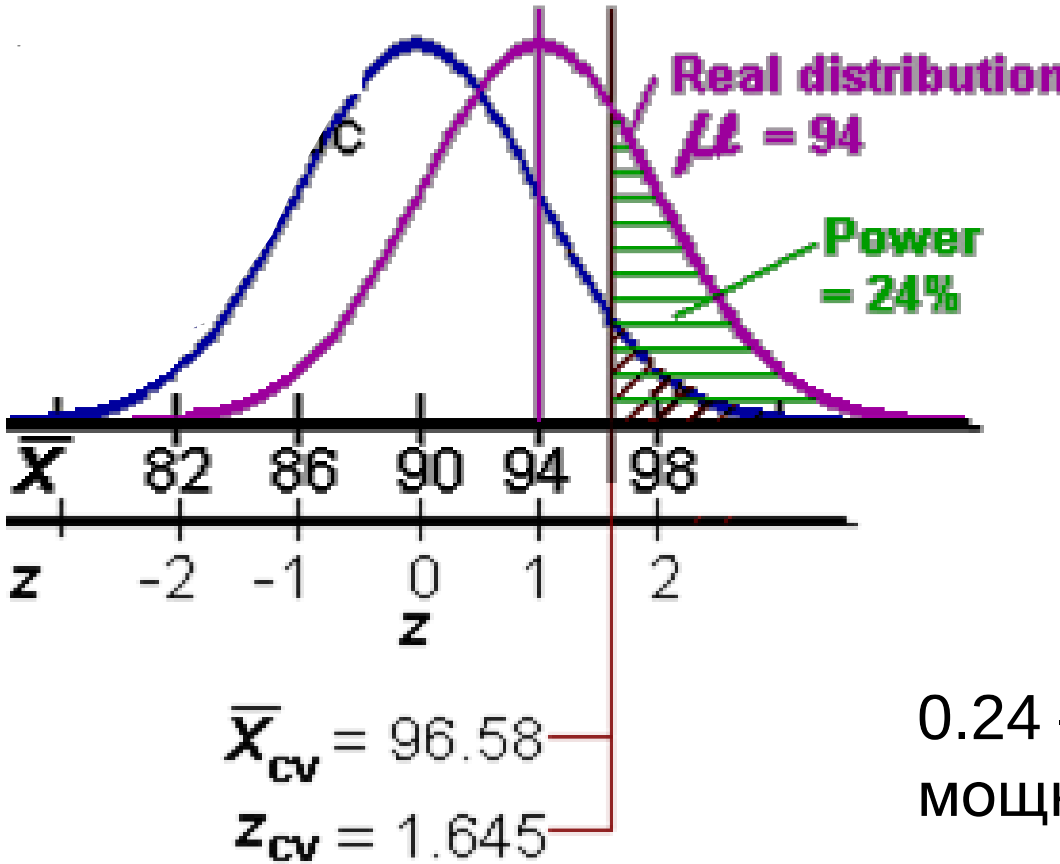


$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{25}} = 4$$

Размер выборки n = 25 зверей

Мощность

Если мы поймаем 25 землероек в заповеднике, у нас есть вероятность лишь 24%, что мы найдём различия! Т.к. лишь в 24% случаев **среднее в будущей выборке попадёт в критическую область**.



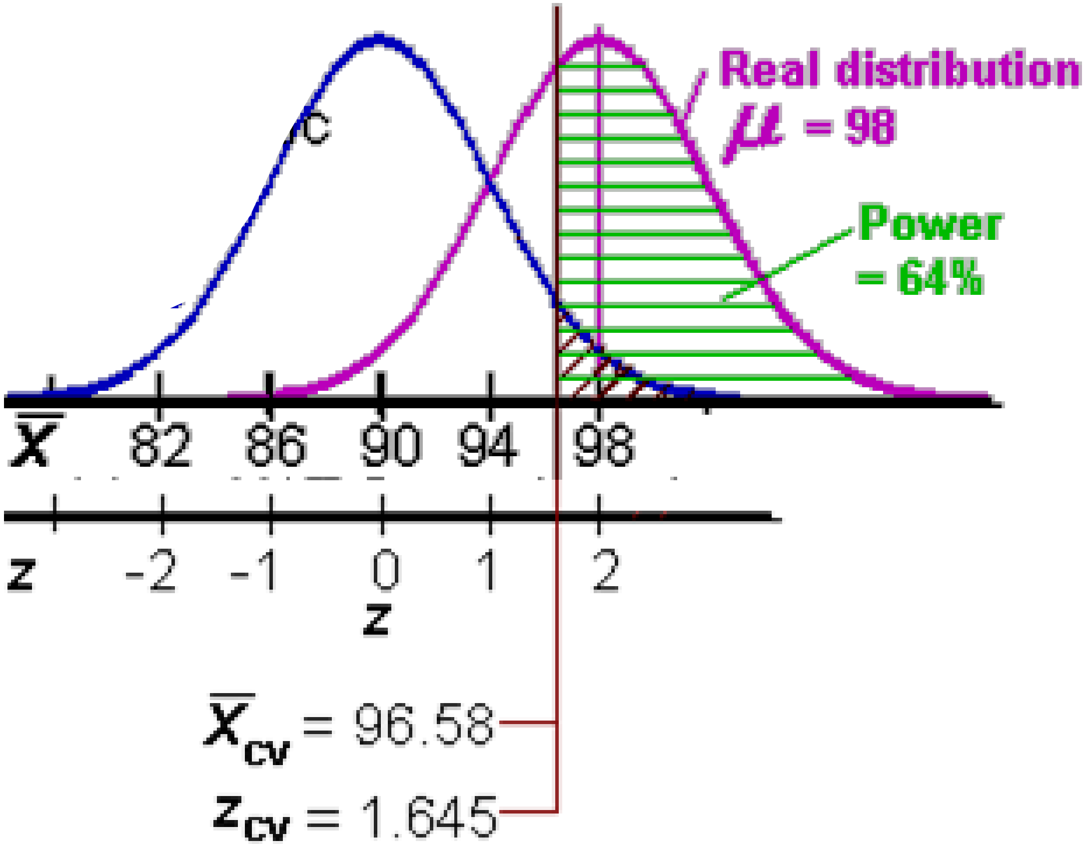
0.24 – низкая мощность

От чего зависит мощность?

1. От <u>размера выборки</u> (в случае *t*-тестов - делает «уже» распределения выборочных средних);
2. От <u>разброса</u> (поэтому надо минимизировать постороннюю дисперсию);
3. От <u>выбора статистического теста</u> (напр., для связанных выборок);
4. При сравнении групп важно выровнять группы по <u>размеру</u>;
5. При сравнении ≥3 групп разом, <u>чем больше групп</u>, тем меньше мощность.
6. От различий между популяциями (effect size);
7. От уровня значимости (α=0.05 а не α=0.01);
8. У одностороннего теста мощность выше, чем у двустороннего.

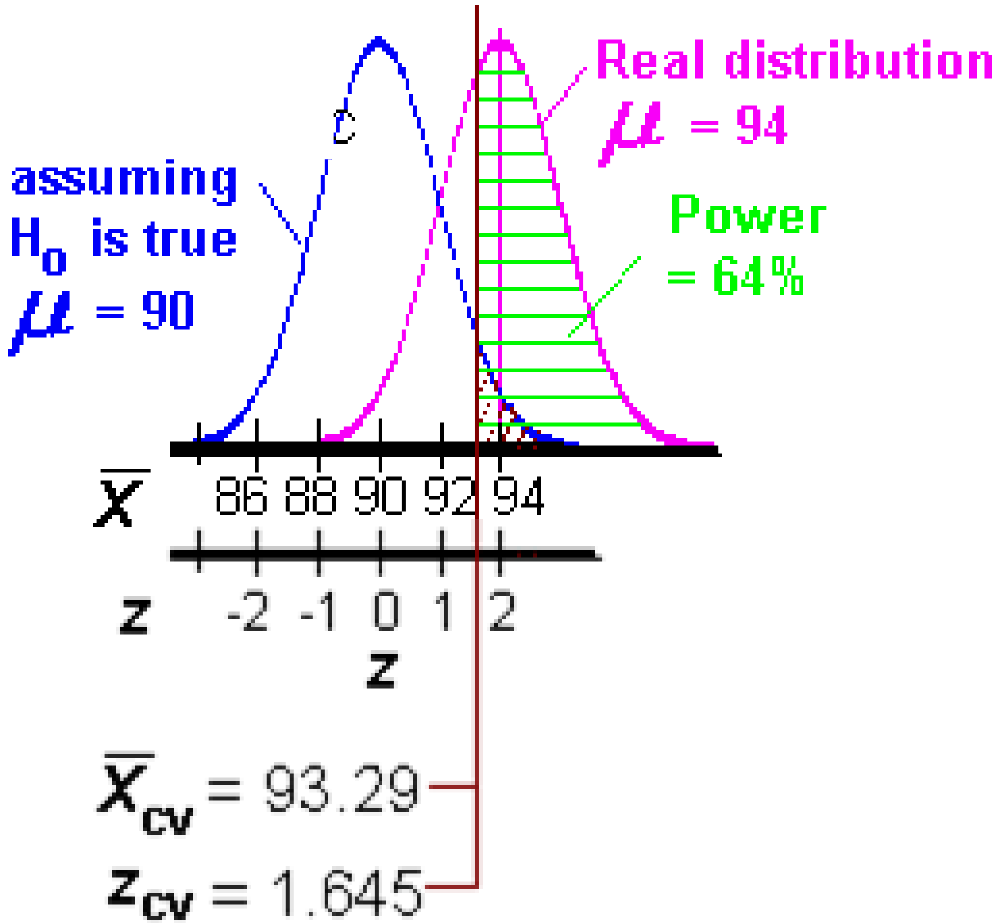
Мощность

Если в действительности средняя масса землероек в популяции на острове равна 98 г, мощность теста будет уже 64%.



Мощность

Здесь стандартное отклонение в популяции вдвое меньше исходного, и мощность теста тоже 64%.



Единственный способ на практике уменьшить стандартное отклонение – сделать выборку более гомогенной (для эксперимента подобрать зверей одного возраста, пола, и др.)

Мощность

Два раздела и цели анализа мощности:

1. В планировании исследования. Мы можем:

   - ✓ Прикинув размер эффекта (имеющий биологический смысл!), задав желаемую мощность (обычно, 0.8), и оценив примерный разброс (из литературы или предварительных данных) рассчитать необходимый **РАЗМЕР ВЫБОРКИ** для исследования!
   - ✓ Если размер выборки ограничен (скажем, нет денег на лабораторных животных и пр.), прикинув размер эффекта, разброс и задав N, можно рассчитать **МОЩНОСТЬ ТЕСТА**, и решить: стоит ли вообще браться за эту работу.

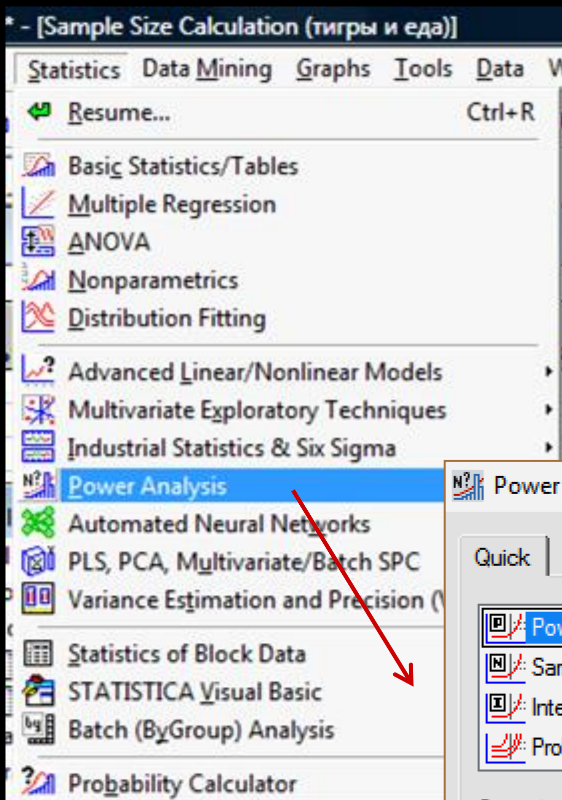
Мощность

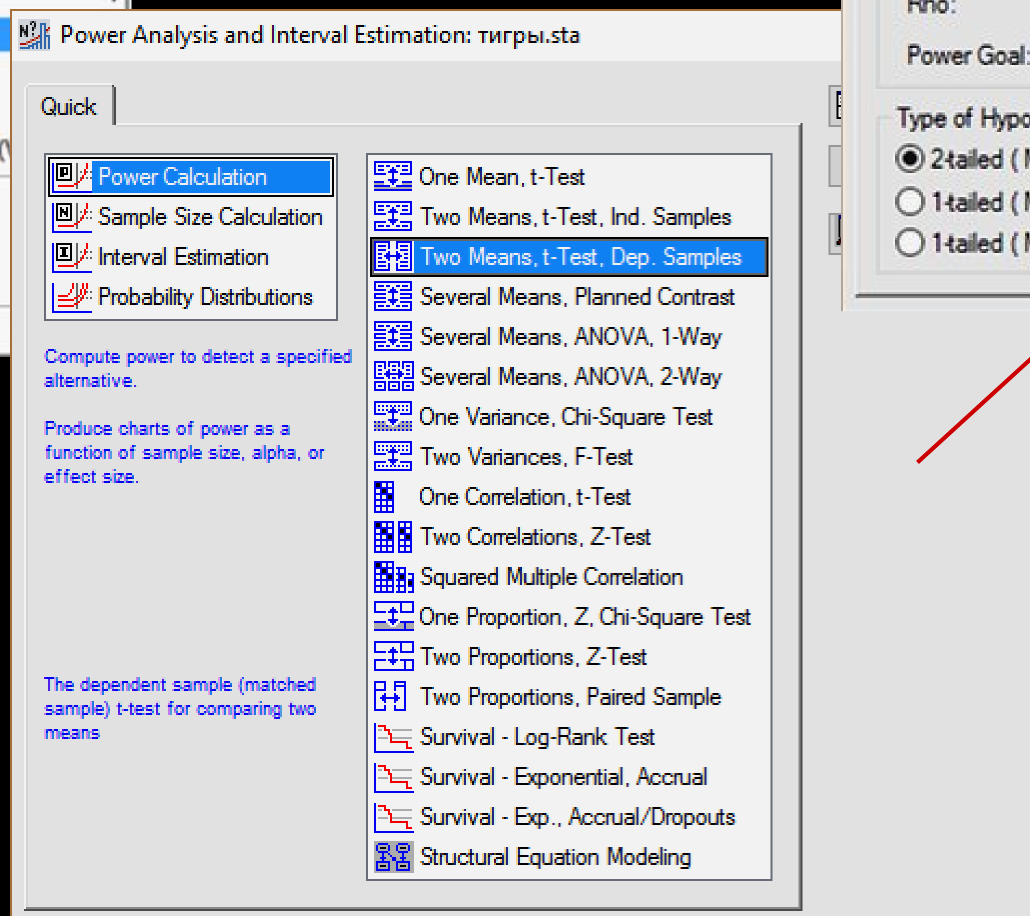
Спланируем эксперимент, чтобы оценить эффект устройства для похудения.

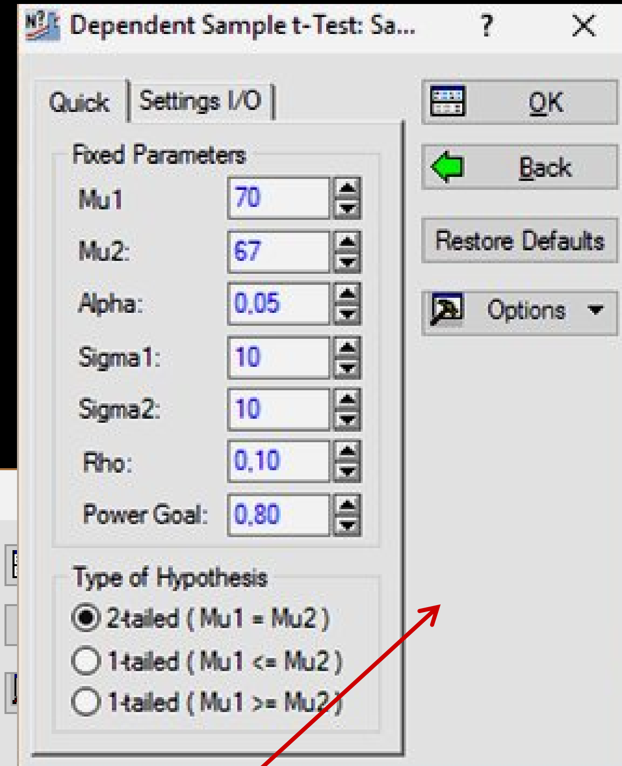
1. Прикинем абсолютный **размер эффекта**: пусть похудение за месяц на 1 кг будет считаться достойным результатом (ES=3 кг, стартовая масса 70 кг).
2. Зададим **мощность** (обычно задают 0.8).
3. Примерно оценим **изменчивость** в выборке (скажем, посчитаем SD массы тела в небольшой группе добровольцев).

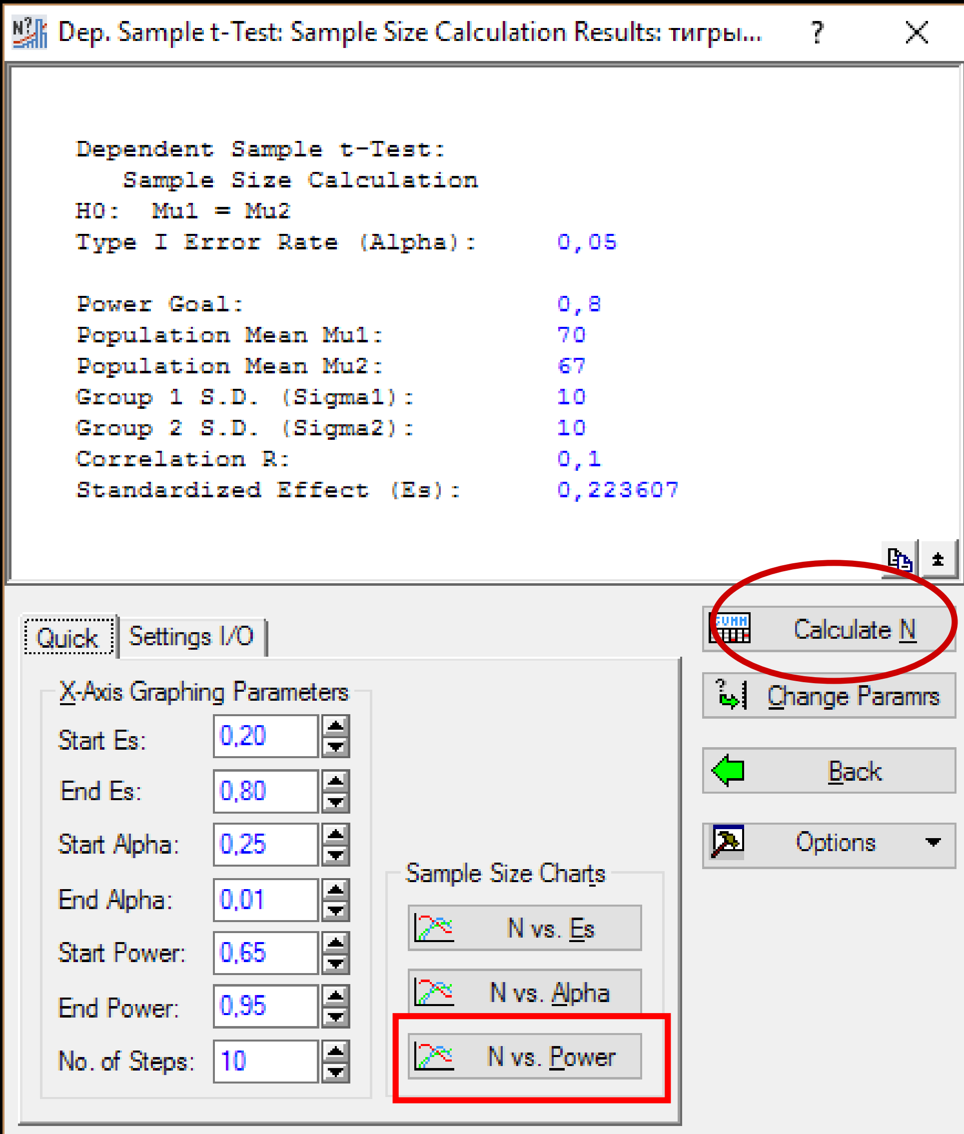
Теперь поместим все эти цифры в специальный калькулятор, и он посчитает нам N!

Расчёт РАЗМЕРА ВЫБОРКИ для заданных различий и мощности



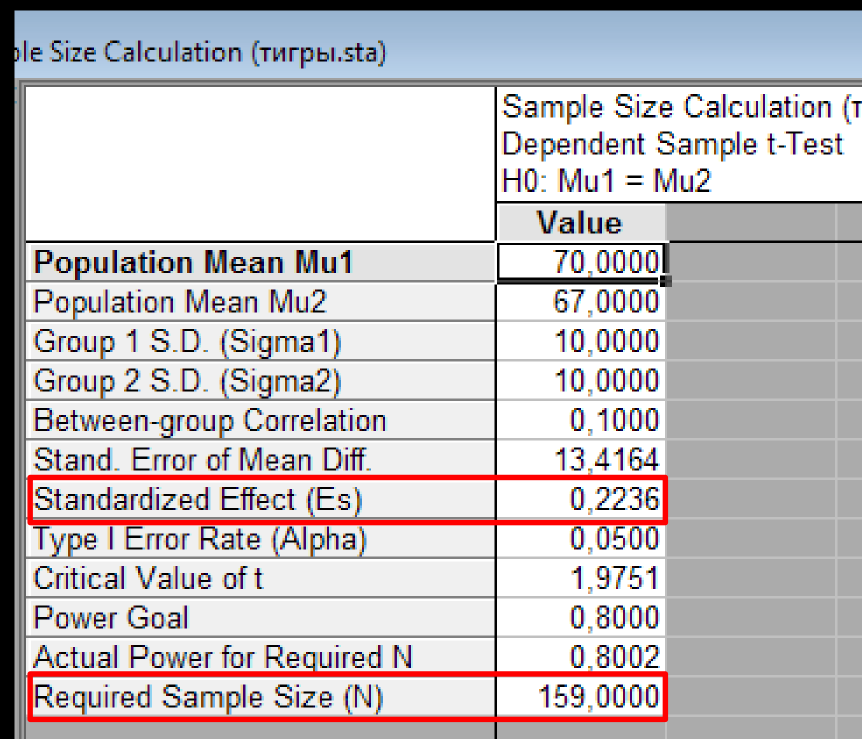






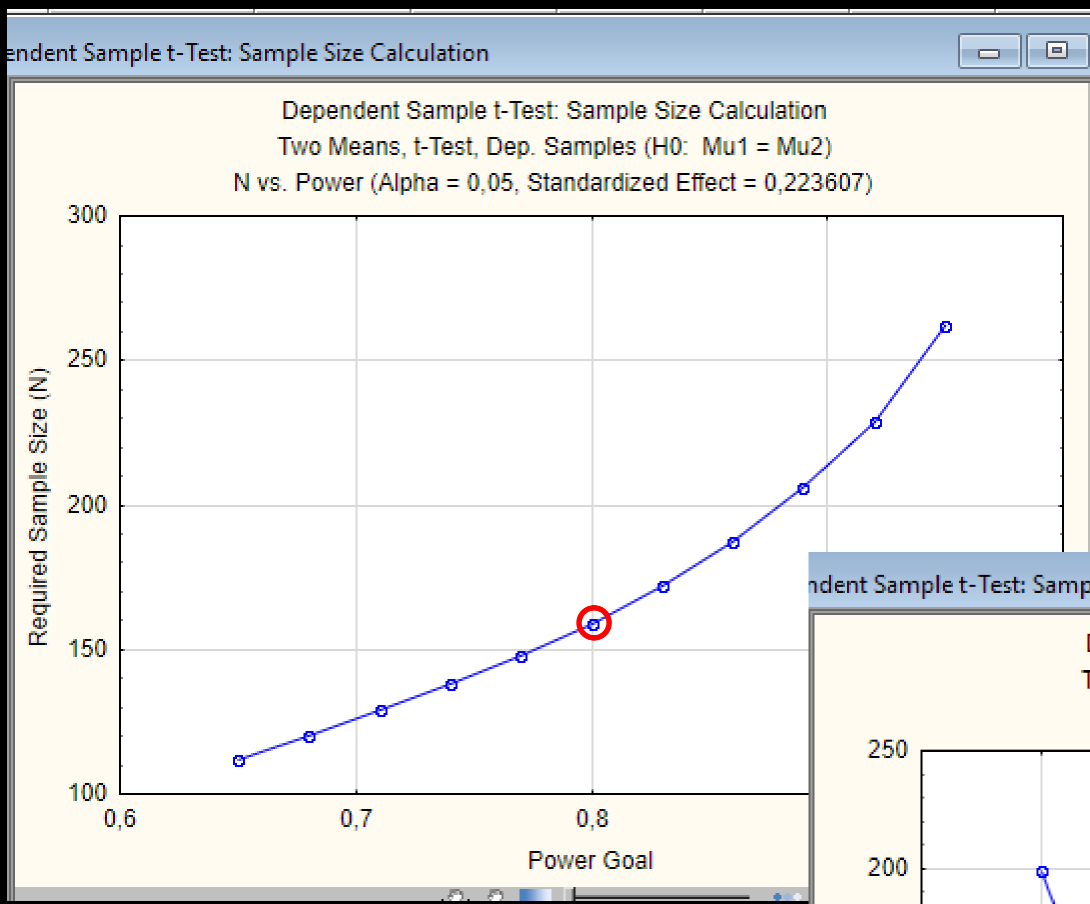
$$n_2 = \frac{nn_1}{2n_1 - n}$$

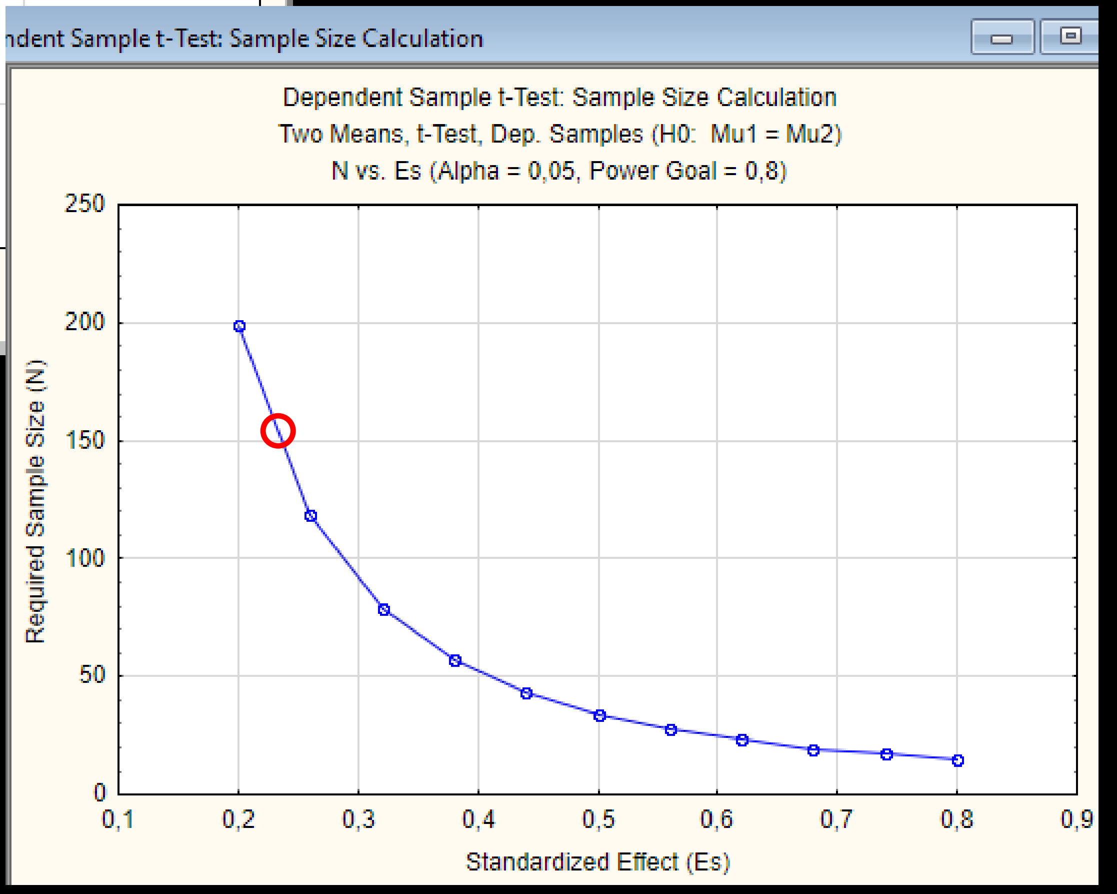
Для независимых выборок при фиксированном n_1 можно сначала рассчитать общее N, а потом n_2

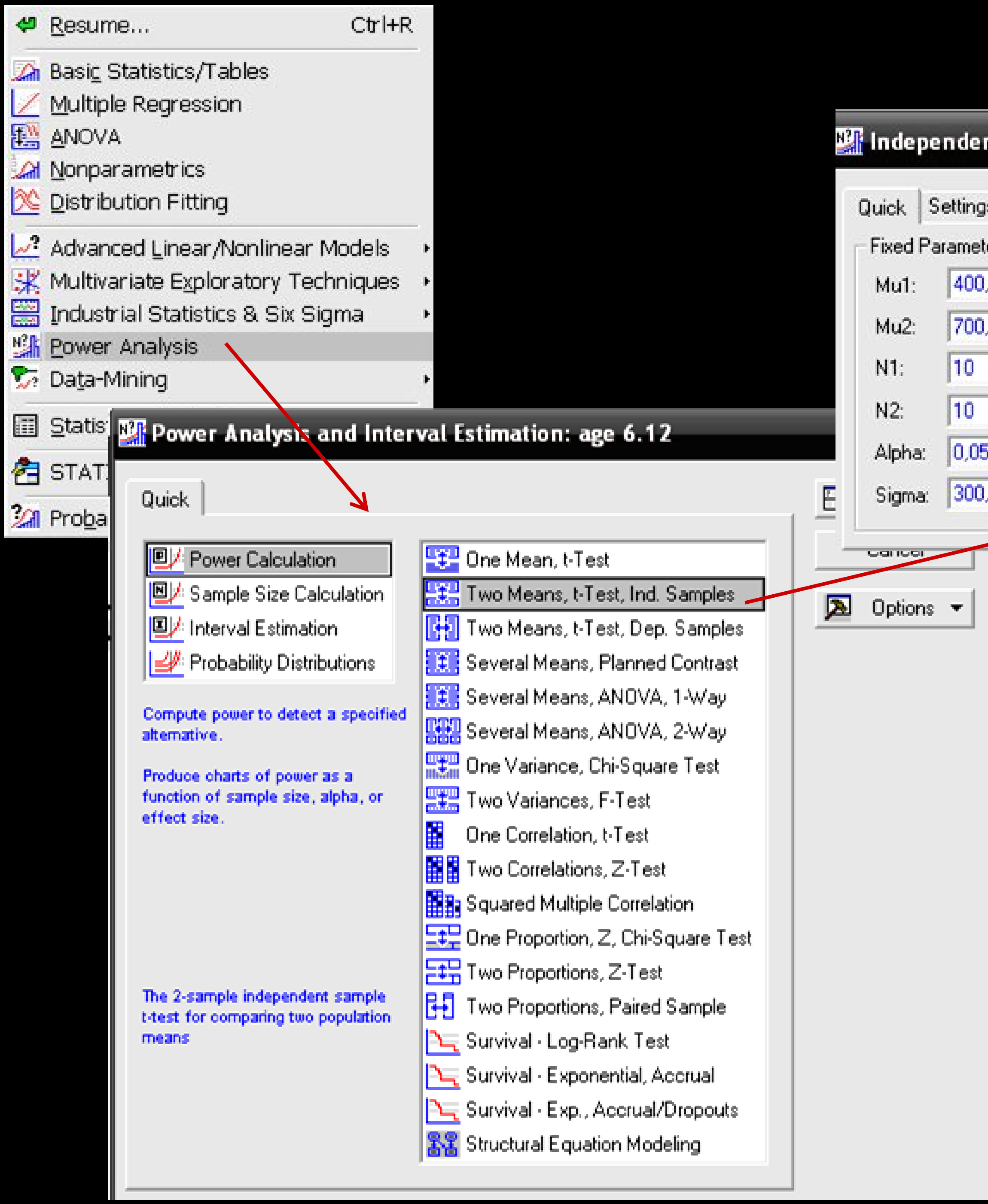

ple Size Calculation (тигры.sta)

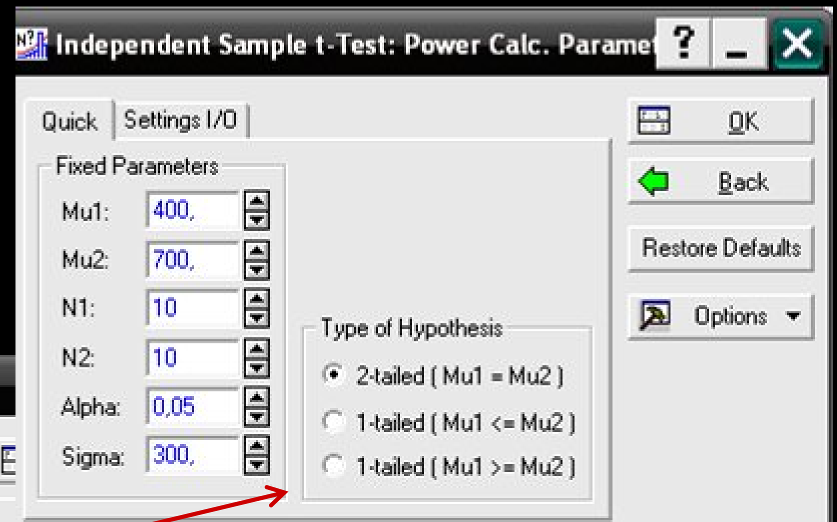
	Sample Size Calculation (ти Dependent Sample t-Test H0: Mu1 = Mu2 Value
Population Mean Mu1	70,0000
Population Mean Mu2	67,0000
Group 1 S.D. (Sigma1)	10,0000
Group 2 S.D. (Sigma2)	10,0000
Between-group Correlation	0,1000
Stand. Error of Mean Diff.	13,4164
Standardized Effect (Es)	0,2236
Type I Error Rate (Alpha)	0,0500
Critical Value of t	1,9751
Power Goal	0,8000
Actual Power for Required N	0,8002
Required Sample Size (N)	159,0000

Поскольку мы оценивали размер эффекта примерно, удобно построить картинки и посмотреть, как N зависит от мощности и размера эффекта.

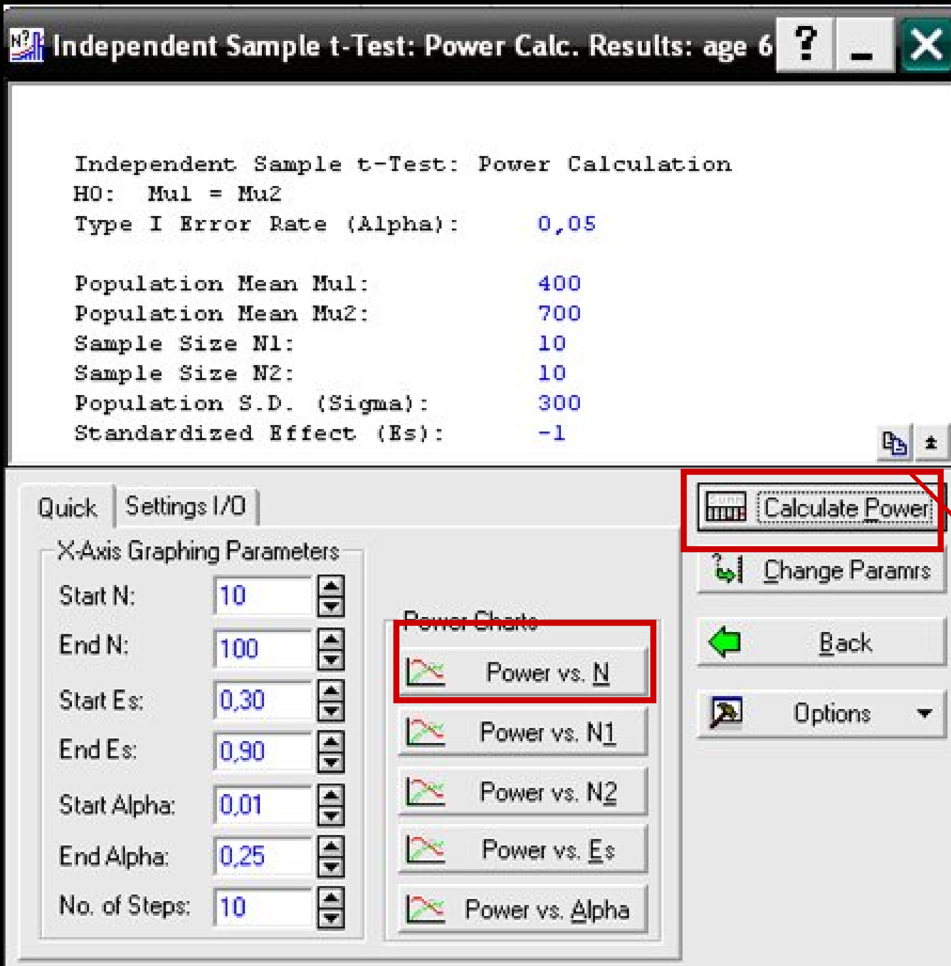


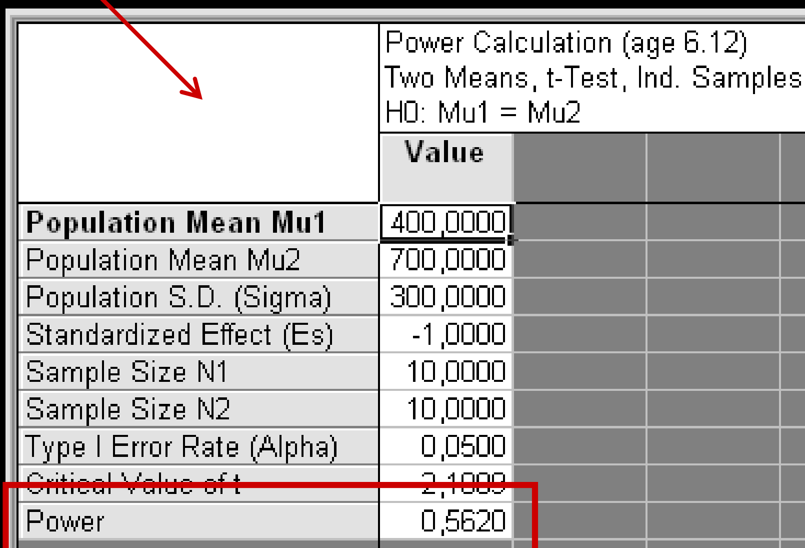






РАСЧЁТ МОЩНОСТИ для двухвыборочного t-критерия для независимых выборок.





	Power Calculation (age 6.12) Two Means, t-Test, Ind. Samples H0: Mu1 = Mu2 Value
Population Mean Mu1	400,0000
Population Mean Mu2	700,0000
Population S.D. (Sigma)	300,0000
Standardized Effect (Es)	-1,0000
Sample Size N1	10,0000
Sample Size N2	10,0000
Type I Error Rate (Alpha)	0,0500
Critical Value of t	2,1009
Power	0,5620

Мощность

Два раздела и цели анализа мощности:

2. Когда тест уже проведён, мы можем:

✓оценить фактический **РАЗМЕР ЭФФЕКТА** (это мы уже обсуждали);

✓ если результат «недостоверный», мы можем посчитать, какой будет **МОЩНОСТЬ** теста для имеющего биологический смысл размера эффекта (здесь размер выборок и разброс уже для нас известны!), и понять, был ли наш тест вообще в силах обнаружить этот эффект;

✓ часто программы считают «наблюдаемую мощность» теста; этот показатель в точности коррелирует с *P* и абсолютно неинформативен.

Для публикаций

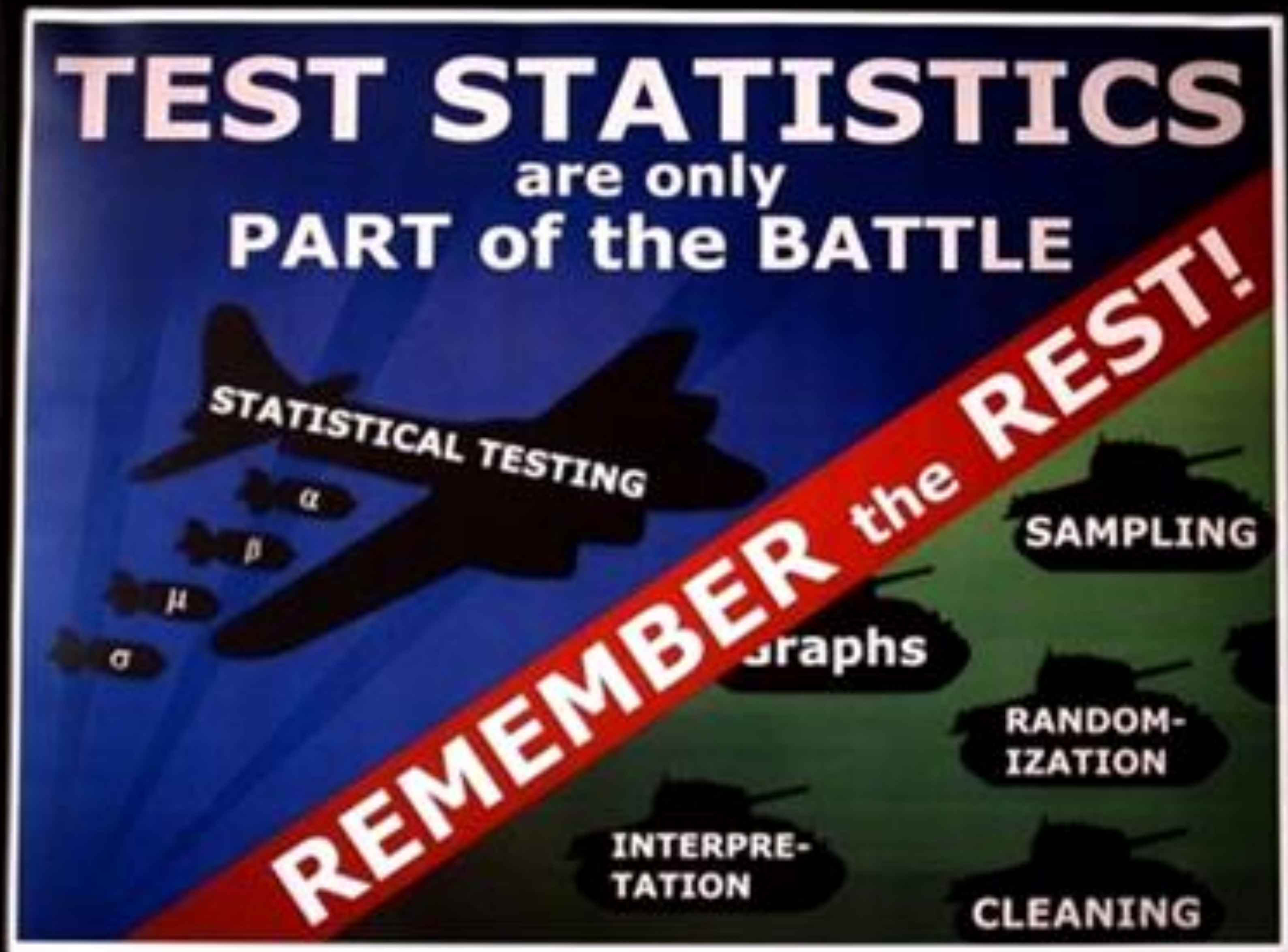
Многие серьёзные журналы ТРЕБУЮТ приводить размер эффекта и/или оценку мощности теста, особенно в случае недостоверных результатов.

В секции **методов**: написать, какой индекс размера эффекта использовали; по каким стартовым значениям оценивали мощность.
В **результатах**: приводить ES индекс вместе со статистикой критерия и *P*.

«…We calculated the effect size for all tests performed. For comparisons of the two mean values, we used Cohen's d as an effect size measurement (a d value of 0.20 indicates a small effect, 0.50 indicates a medium-sized effect, and 0.80 indicates a large effect) (Cohen 1988).»

Nakagawa, S., Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol.Rev.Camb.Philos.Soc. 82: 591–605.

McGraw, K.O., Wong, S.P. 1992. A common language effect size statistic. Psychol. Bull. 111: 361–365.



http://www.etsy.com/shop/NausicaaDistribution

Прочие трудности анализа данных и пути их решения

1. Наличие посторонних факторов, затрудняющих трактовку результатов (confounding factors). Отсутствие контрольной группы.

Бывает, что действие исследуемого фактора невозможно отделить от какого-то другого фактора.

Например:

- ✓ в контрольную и экспериментальную группы попали животные из разных источников;
- ✓ детёныши, которые больше играли, лучше выживали; делают вывод, что игра влияет на выживание;
- ✓ …

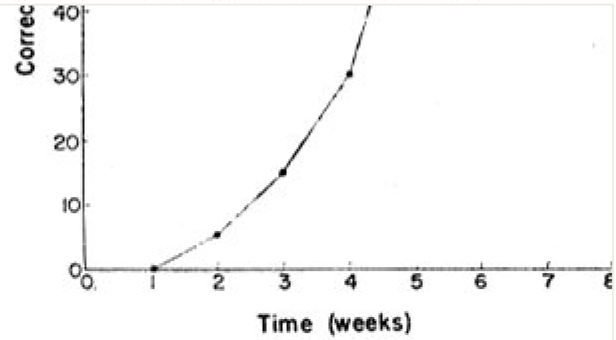
Что делать: 1. всегда думать о таких факторах до начала исследования; 2. иногда эти факторы возможно вставить как дополнительную переменную в модель.

Memory and Hibernation in Citellus lateralis

M. Colleen McNamara; M. L. Riedesel

Science, New Series, Vol. 179, No. 4068 (Jan. 5, 1973), 92-94.

of 12 hours of light followed by 12 hours of darkness for an 11-day period. All animals spent the same amount of time in the cold. For the first few days the animals had free access to food. Subsequently, food was withdrawn in varying amounts to encourage entrance into hibernation. Some animals hibernated while others did not. Bouts of hibernation were monitored by placing



. At the end of the 11 days in ld all animals were housed in the tory for 24 hours before testing tention (test A), and given free to food and water. Two days esting the animals were returned hibernaculum for another 11-day xposure and retested 24 hours test B).

he retention tests A, B, and C, nimal was given four trials in the

During the first cold-exposure period the mean hibernation period was 12 days for males and 8 days for females. Ten animals had hibernated. One male and one female died during the first 11 days of cold exposure. Animals that hibernated had better retention than did those that did not (test A in Table 1).

During the second cold-exposure period the mean hibernation period was 8 days for males and females; 11 ani-

2. «Псевдорепликация» данных (pseudoreplication).

Наблюдения в выборке не должны быть связаны друг с другом.

Например:

- ✓ измерения от одной и той же особи связаны;
- ✓ потомки одной особи (один выводок) связаны;
- ✓ если контрольная группа содержится в одном аквариуме, а экспериментальная - в другом, измерения в группах связаны;
- ✓ …

Что делать: 1. для каждой особи (выводка, аквариума…) брать среднее значение и составлять выборку из таких средних;
2. номер особи (выводка и пр.) можно вставить в анализ как дополнительный фактор особого типа (см. Занятие 4).

3. Большая посторонняя изменчивость.

Бывает, что в выборке оказываются особи разного возраста, пола, гормонального статуса, в разных фазах эстрального цикла, пойманные в разные сезоны и пр.

В результате:

- ✓распределение переменной оказывается не нормальным (и даже мультимодальным);
- ✓ действие фактора маскируется большой внутригрупповой изменчивостью, мощность теста падает.

Что делать: 1. для эксперимента подбирать как можно более одинаковых особей; 2. добавлять в модель дополнительные переменные – факторы (пол, возраст и пр.).

4. Несоответствие переменной нормальному распределению

Если размер сравниваемых групп одинаков (особенно если N>20), распределения симметричные, соблюдается равенство дисперсий – не критично (*кроме регрессионного анализа!*).

Бывает, что:

- ✓ переменная в принципе не распределена по нормальному закону (доли, пропорции, частоты, концентрации гормонов);

Что делать: трансформация данных; использование непараметрических тестов (обсудим позже).

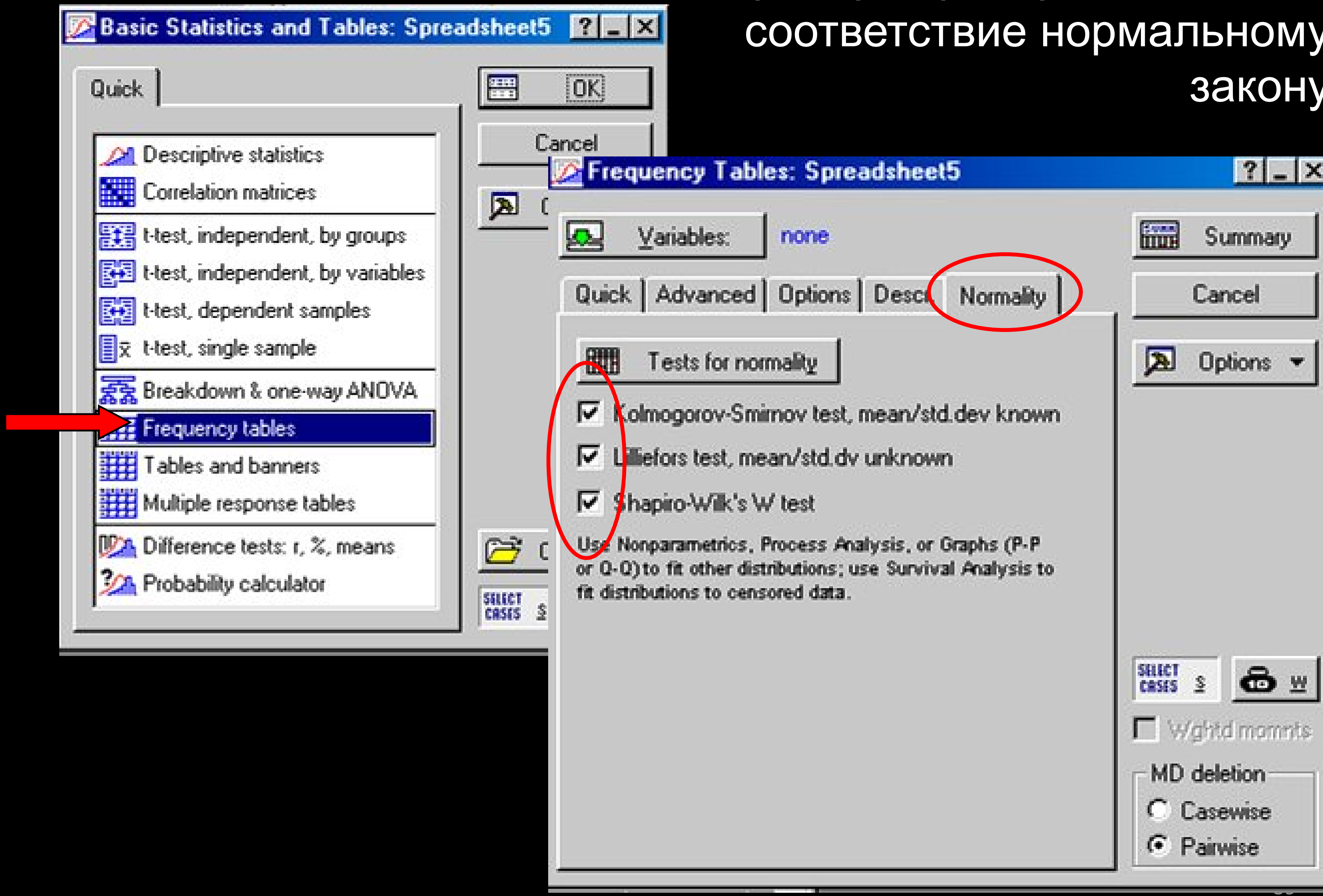
- ✓ мультимодальное распределение (уже обсуждали);
- ✓ в выборке МНОГО НУЛЕЙ.

Что делать: это проблема! Если нулей очень много, надо делить анализ на два этапа: 1. разбивать выборку на категории (1. нули; 2. не нули), анализировать; 2. отдельно анализировать взаимосвязи в ненулевой группе.

Проверка на соответствие нормальному распределению:

- Тест Колмогорова-Смирнова (Kolmogorov-Smirnov test) D-статистика. Маломощный, не рекомендуется (Zar, 2010).
- Lilliefors test – «улучшенный К-С тест»
- **Shapiro-Wilk's W test** (самый мощный, размер выборки до 5000) – наиболее предпочтительный.
- Обязательное графическое исследование (**гистограмма**)!

Проверка распределения на соответствие нормальному закону



Data: Tests of Normality (зубы)

	Tests of Normality (зубы)					
Variable	N	max D	K-S p	Lilliefors p	W	p
weight	186	0,074337	p > .20	p < ,05	0,952717	0,000007

маленькое *p* говорит о том, что данные не соответствуют нормальному распределению (распределение достоверно отличается от нормального распределения с таким же средним и SD)

5.Гомогенность дисперсии (*homogeneity = homoscedasticity*)

При сравнении выборок, у совокупностей, из которых они сформированы, дисперсии должны быть равны между собой. Если дисперсии не равны это называется гетерогенность (*heterogeneity = heteroscedasticity*).

Для регрессий/корреляций дисперсия не должна зависеть от самих переменных. Проверка – на основе анализа остатков и скаттерплота (об этом подробно - позже).

Что делать: 1. выровнять выборки по размеру; 2. убедиться в нормальном распределении; 3. трансформация данных; 4. использование непараметрических критериев.

Проверка **равенства дисперсий** при сравнении групп: вставлена в Статистике в блоки с соответствующими параметрическими тестами (*t*-тест, ANOVA)

- ✓ *F-test* – для двух групп;
- ✓ *Levene’s test* – более надёжный, подходит для двух и более групп;
- ✓ *Brown & Forsythe's test* – подходит для выборок разного размера
- ✓ *Barlett’s test* – для трёх и более групп
- ✓Обязательное графическое исследование (**гистограмма и усатые ящики**)!

В методах: «In all parametrical tests the samples were homoscedastic (Levene's test, p>0.05)».

6. Аутлаеры

Значения, дальше, чем 1,5 межквартильных размахов выше третьей и ниже первой квартилей.

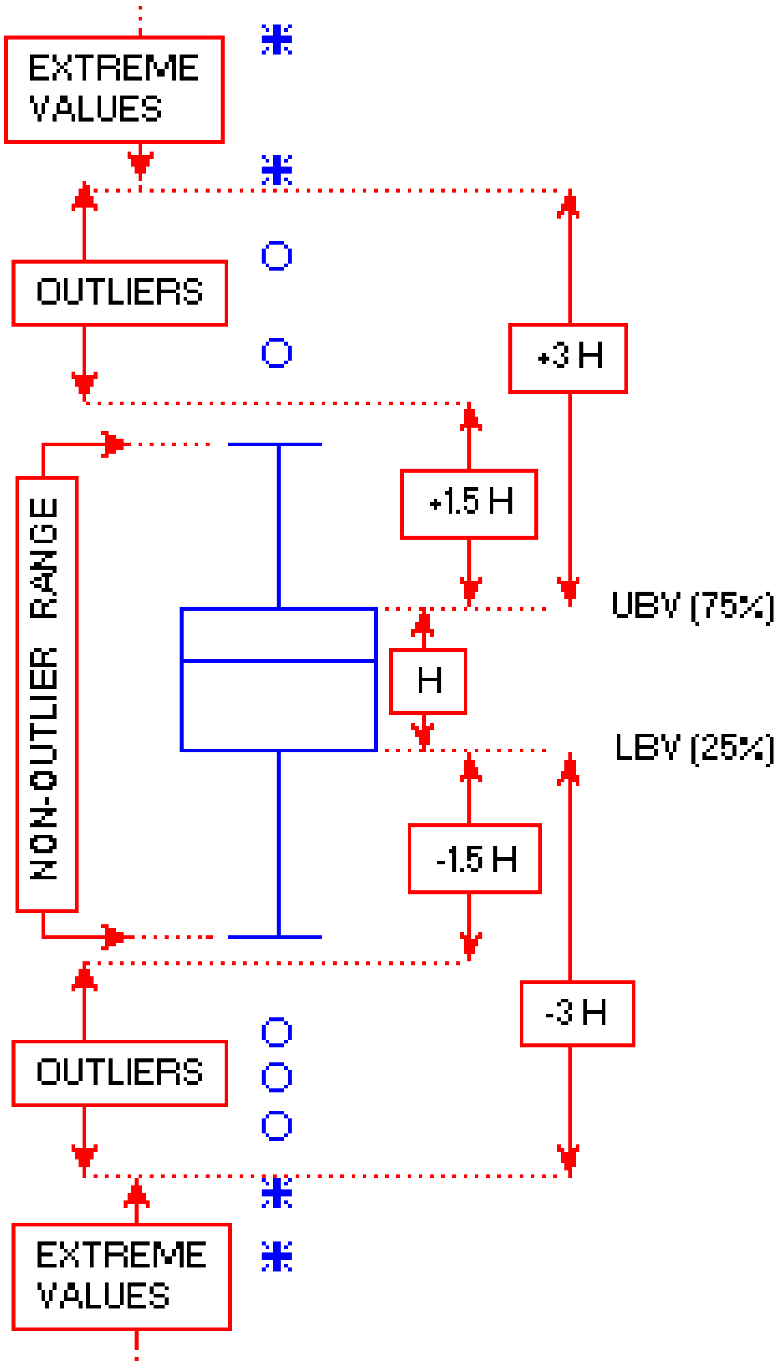
Бывают из-за:

- ✓ ошибок;
- ✓ влияния сторонних факторов (болезнь, принадлежность к другой возрастной группе…)
- ✓ природы данных (в сильно скошенном распределении они могут быть по естественным причинам).

Что делать: 1. обязательно построить картинку; 2. проверить на предмет ошибок; 3. трансформация данных.

Отбрасывать аутлаеры без оснований нельзя! Если 1-2 аутлаера меняют результат теста, но это не ошибки, лучше провести дополнительное исследование.

Аутлаеры


EXTREME VALUES
OUTLIERS
+3 H
+1.5 H
NON-OUTLIER RANGE
UBV (75%)
H
LBV (25%)
-1.5 H
-3 H
OUTLIERS
EXTREME VALUES

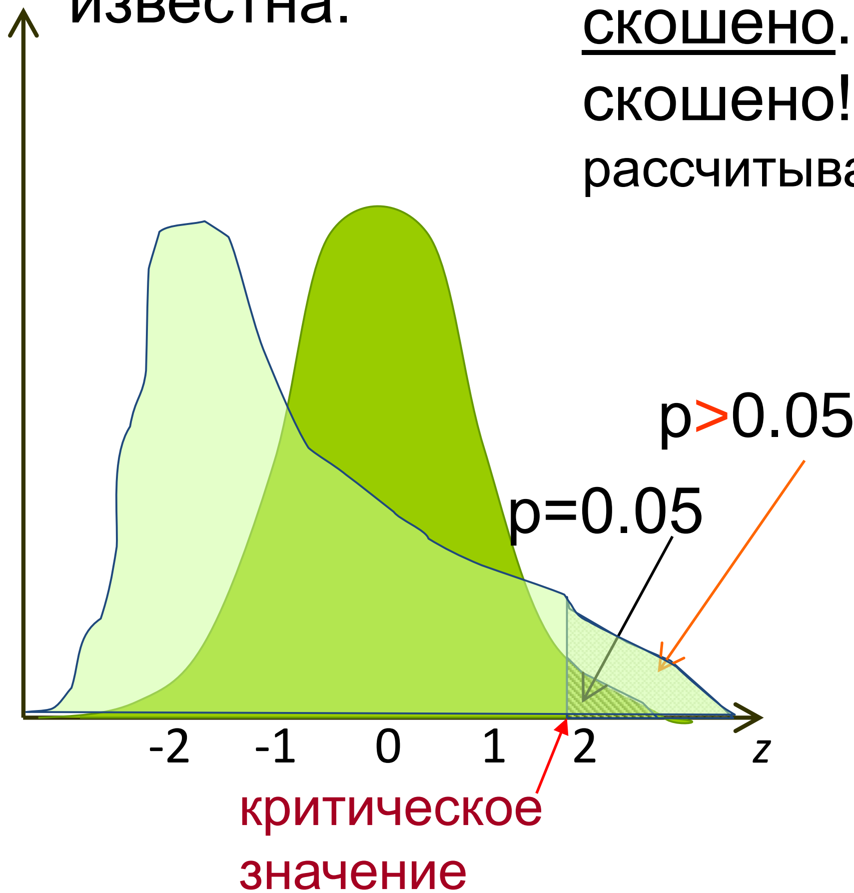
Трансформация данных

повторение

$\mathbf{H_0}$: $\mu \leq 90$ г;
$\mathbf{H_1}$: $\mu > 90$ г
Пусть $\boldsymbol{\sigma}$ известна.

*Распределение **статистики критерия** Z не будет нормальным, если в выборке не нормальное распределение.*

Пусть распределение в популяции <u>скошено</u>. Z-распределение тоже будет скошено! А критическое значение рассчитывается для нормального распределения.



Вероятность, что среднее в выборке попадёт в критическую область (рассчитанную для нормального распределения), будет <u>выше</u>, чем 0.05 – непредсказуемо **увеличится ошибка 1-го рода**

Распределения бывают

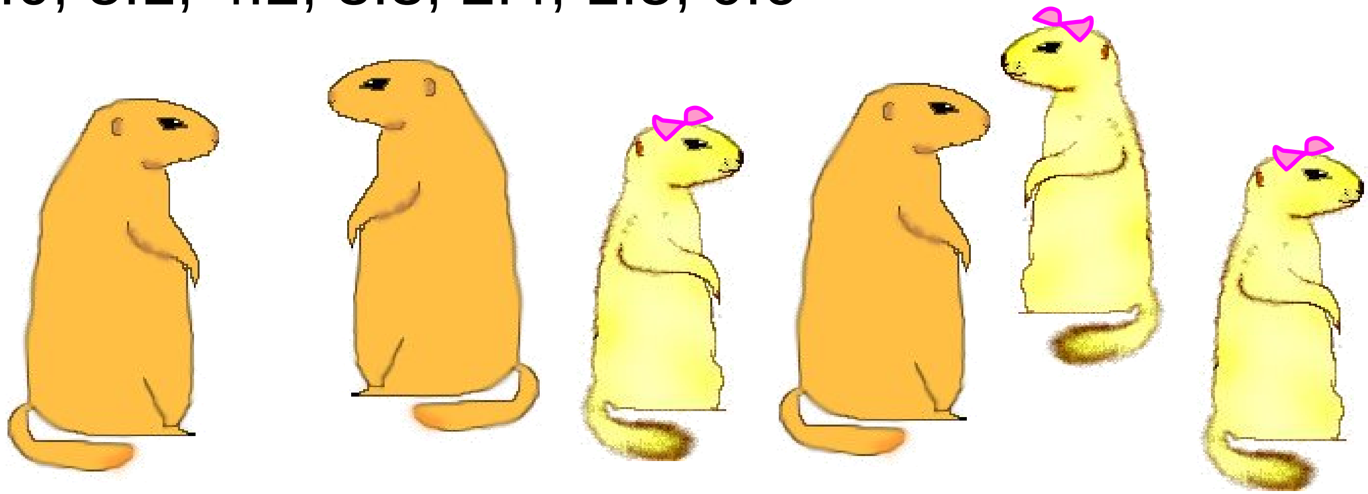
природные
(нормальное, биномиальное, Пуассона, экспоненциальное, лог-нормальное)

распределения статистик критериев
(t, F, U …)

Биномиальное распределение

Пример: рассмотрим выводки из 6 детёнышей каждый. Возможное соотношение самцов и самок в выводке:

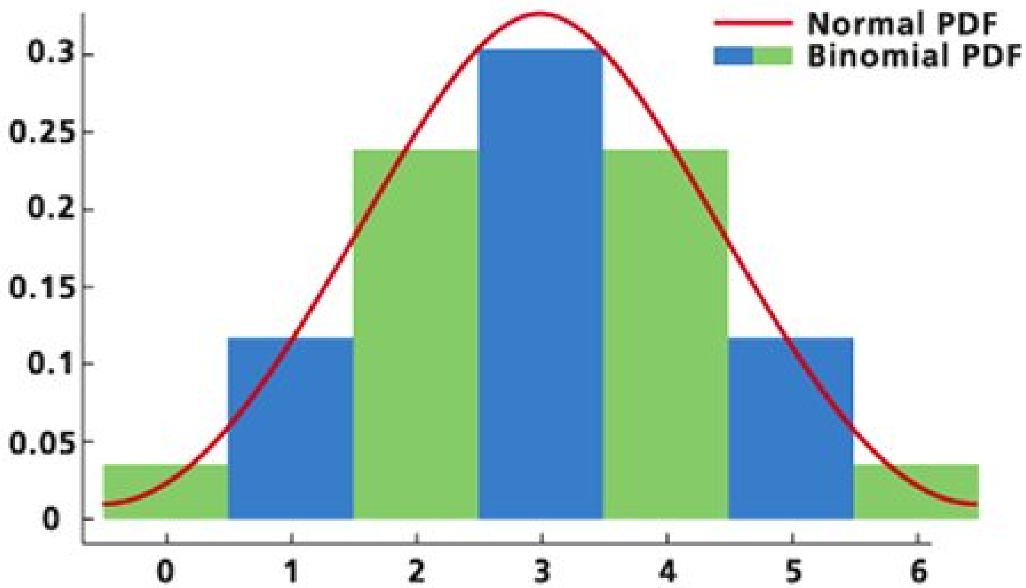
6:0; 5:1; 4:2; 3:3; 2:4; 1:5; 0:6



Биномиальное распределение

распределение количества самцов в N выводков (независимых случайных экспериментов) из $n = 6$ зверьков, таких что вероятность рождения самца постоянна и равна p, а вероятность рождения самки $q = 1 - p$.

Вероятность такого выводка



Количество самцов в выводке из 6 зверьков

Isaac Newton

Если p мало, ситуация лучше описывается распределением Пуассона

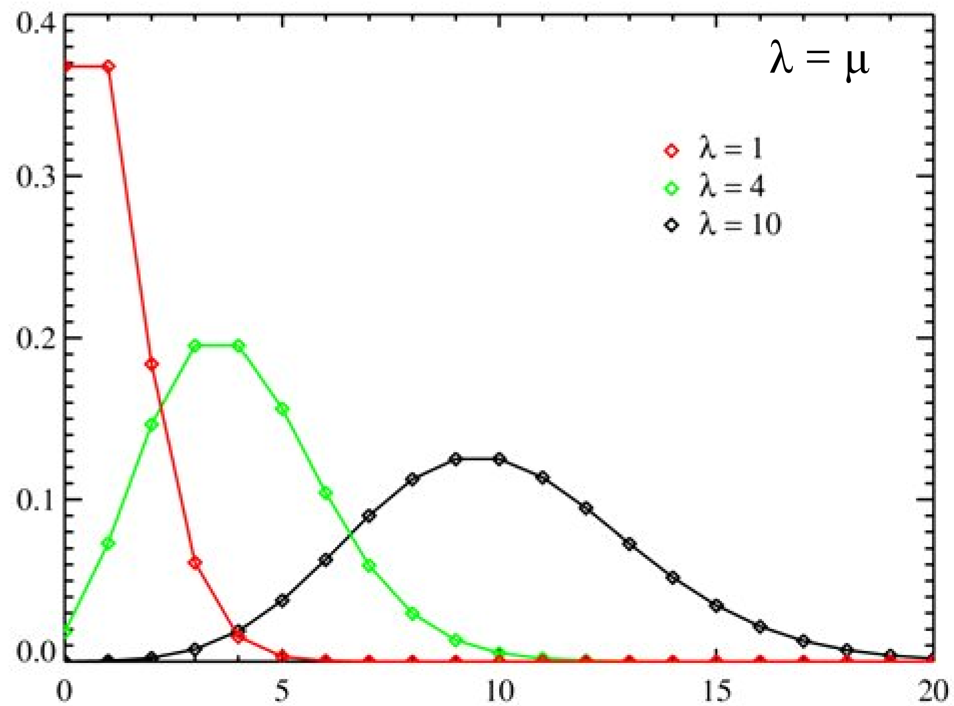
Биномиальному распределению обычно соответствуют доли, частоты, пропорции

Распределение Пуассона

Показывает вероятность того или иного количества независимых друг от друга редких и случайных событий (особей, контактов, мутаций и пр.) на заданном интервале времени (участке пространства, объёме...).

Siméon Denis Poisson

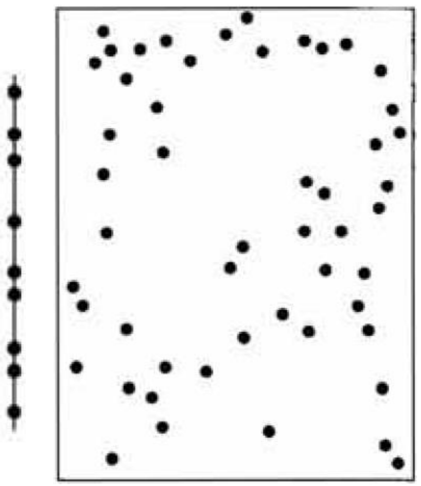
$$\boxed{\mu = \sigma^2}$$



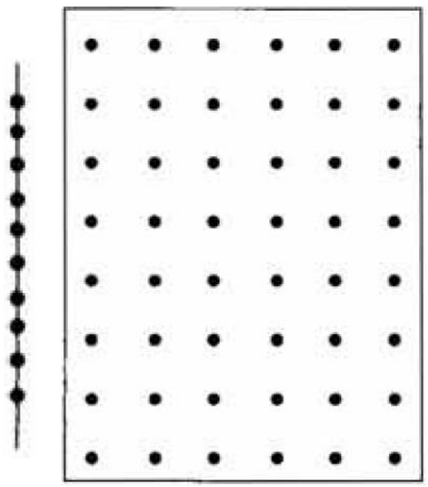
Распределению Пуассона соответствуют частоты, количества случайно распределённых объектов

Распределение Пуассона

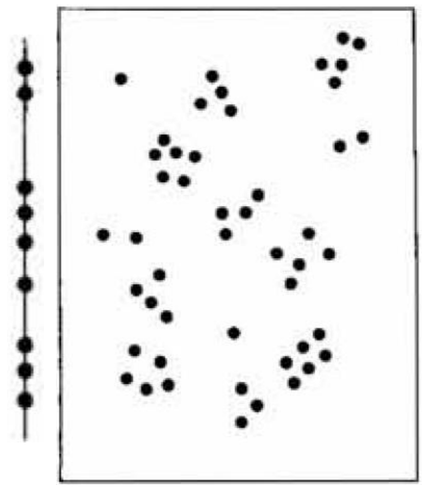
Сравнение распределения объектов во времени и пространстве со случайным распределением (testing for randomness)



$$\sigma^2 = \mu$$



$$\sigma^2 < \mu$$

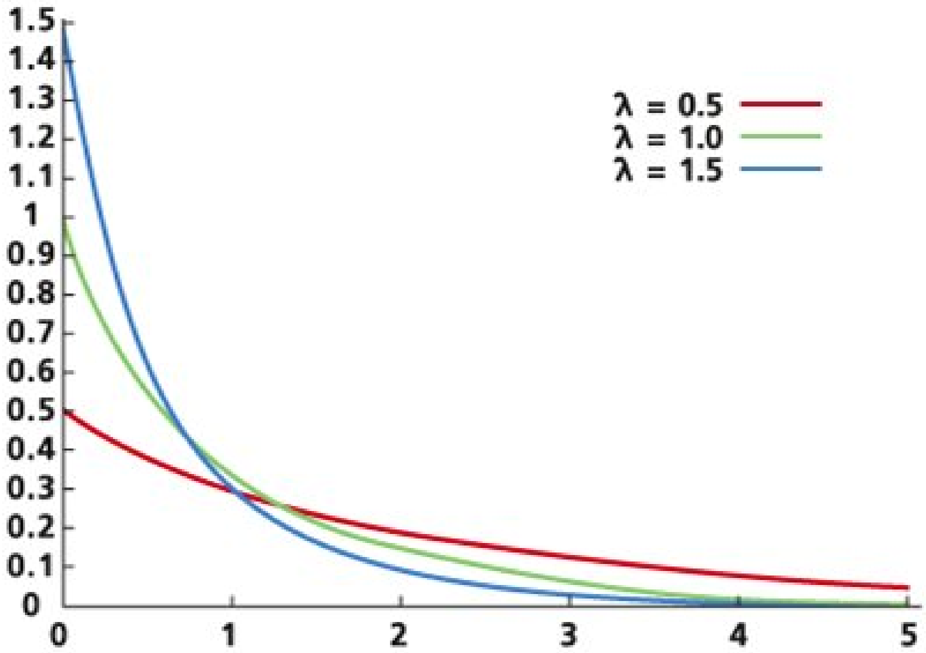


$$\sigma^2 > \mu$$

Важно: следует задавать размер элементарной единицы пространства (времени и пр.), напр., квадрата, так, чтобы μ ≈ 1

Экспоненциальное распределение

Хорошо описывает распределение промежутков времени (расстояний) между случайными событиями с заданной средней частотой событий.



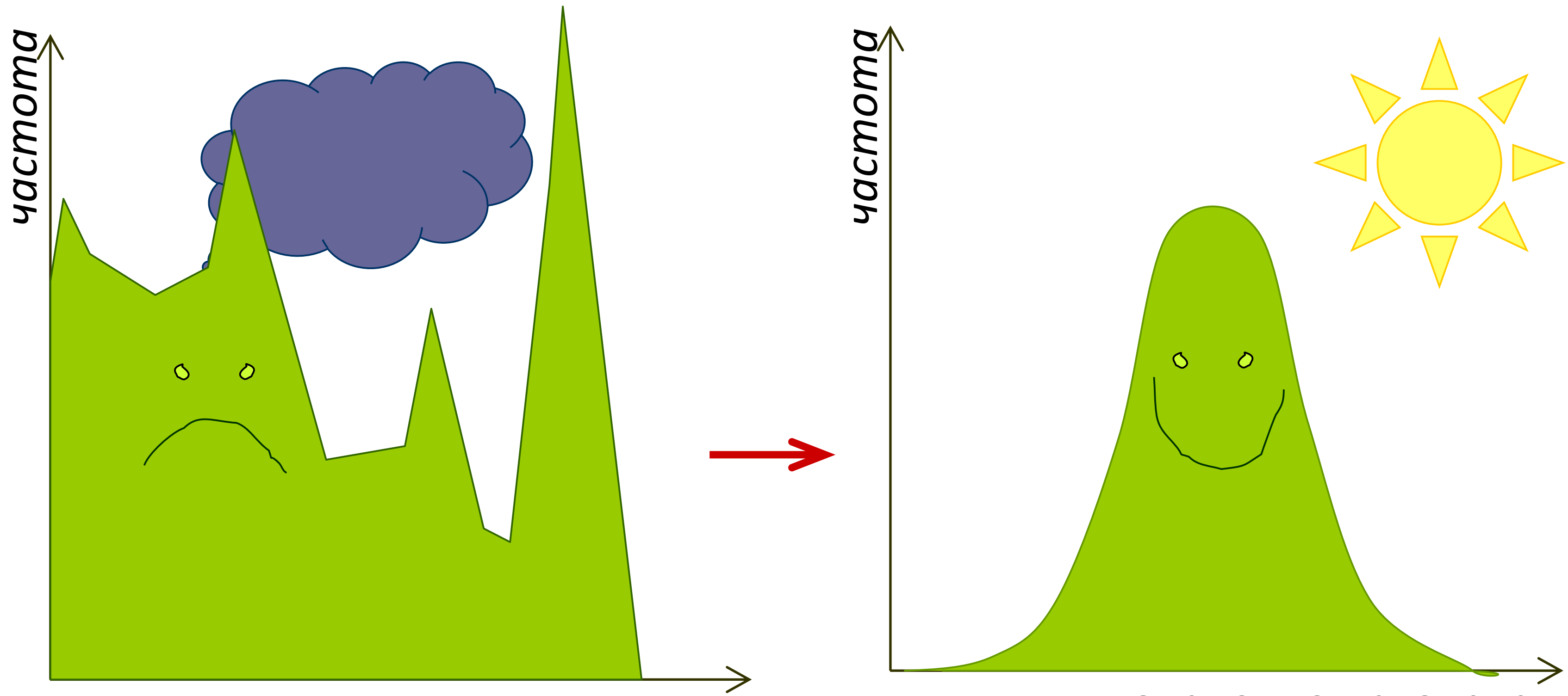
Трансформация данных

Логнормальное распределение

Скошенное вправо распределение; при логарифмировании данных получается нормальное распределение. Часто встречается в природе (концентрации, частоты…)

Трансформация данных

Если распределение отлично от нормального, дисперсии не гомогенны, можно **ТРАНСФОРМИРОВАТЬ данные**



Прекрасное свойство: часто трансформация данных приводит одновременно к нормальному распределению, гомогенности и аддитивности факторов в ANOVA

Трансформация данных

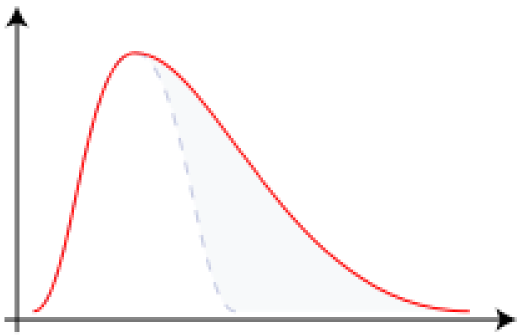
1. Логарифмическая трансформация *(logarithmic transformation):*

•Делает симметричным скошенное вправо (positively skewed) распределение.

•Хорошо работает, когда среднее значение в группах прямо пропорционально стандартному отклонению.

$$\boxed{X_i' = \lg X_i}$$

$$\boxed{X_i' = \lg\left(X_i + 1\right)}$$



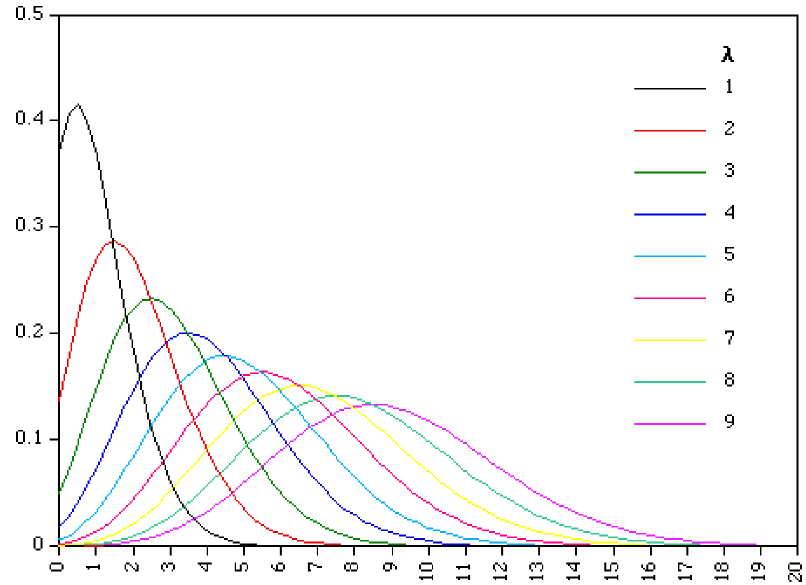
Если в результате логарифмирования получилось нормальное распределение, исходное распределение было **логнормальным.**

2. Извлечение квадратного корня *(square root transformation)*

- Используется, для скошенных вправо распределений, когда среднее значение в группе прямо пропорционально дисперсии.
- обычно такое явление свойственно выборкам из распределения Пуассона (т.е., данные представляют собой количества случайных событий, объектов…)

$$X_i' = \sqrt{X_i}$$

$$X_i' = \sqrt{X_i + 0{,}5}$$

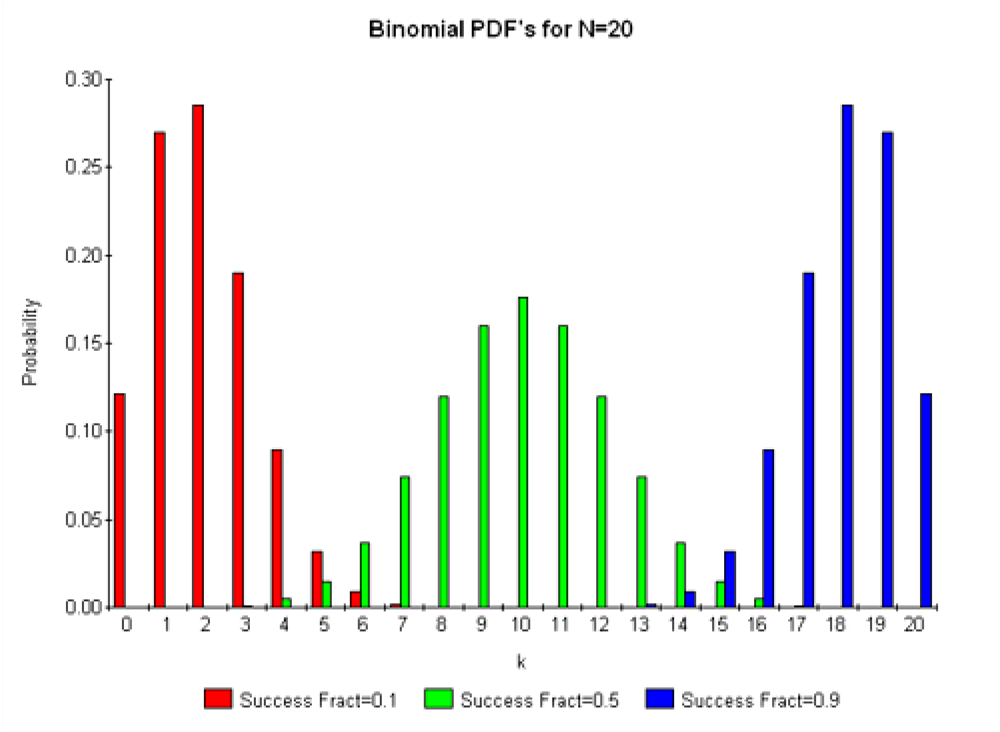


Например, количество социальных контактов в час.

Трансформация данных

3. Арксинусная трансформация (*arcsine transformation*) применяется для процентов и долей ($X_i \leq 1$), которые обычно формируют биномиальное распределение.

$$X_i' = \arcsin\sqrt{X_i}$$

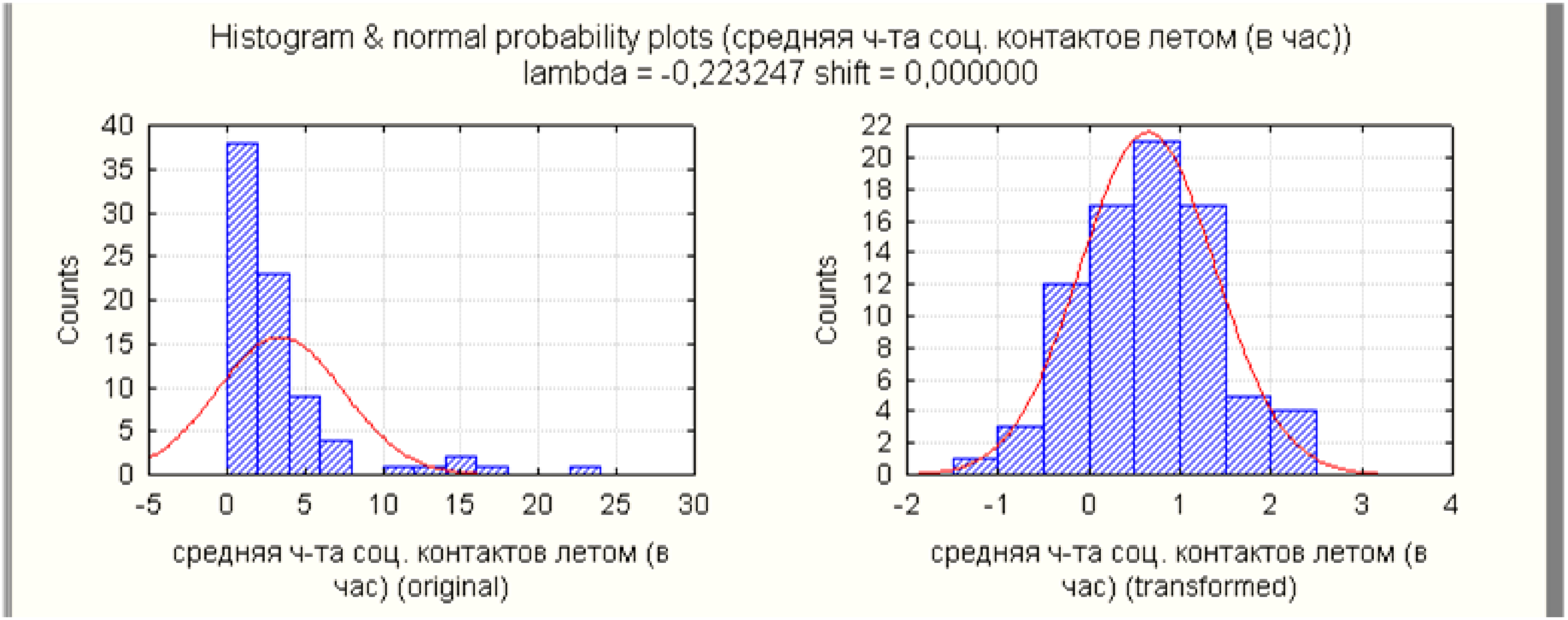


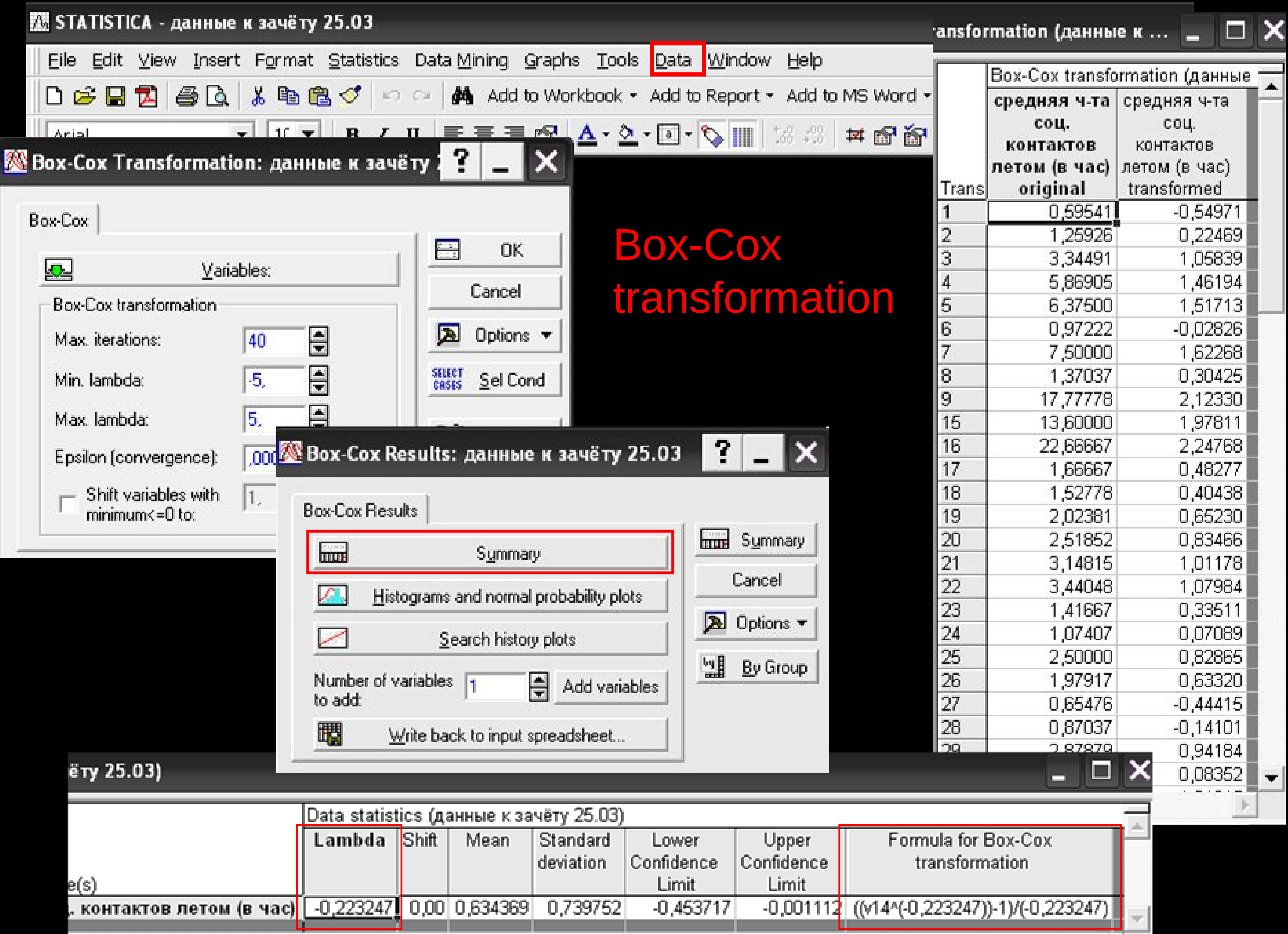
Например, мы исследуем долю самцов или долю переживших зиму детёнышей в выводках сурков.

Прочие трансформации см. Zar, 2010 (1999)

4. Box-Cox transformation

Универсальная трансформация данных, в которой программа методом проб подбирает наилучшие параметры и способ трансформации для конкретных данных (ищется особый параметр λ), лучшее при неизвестном характере распределения.

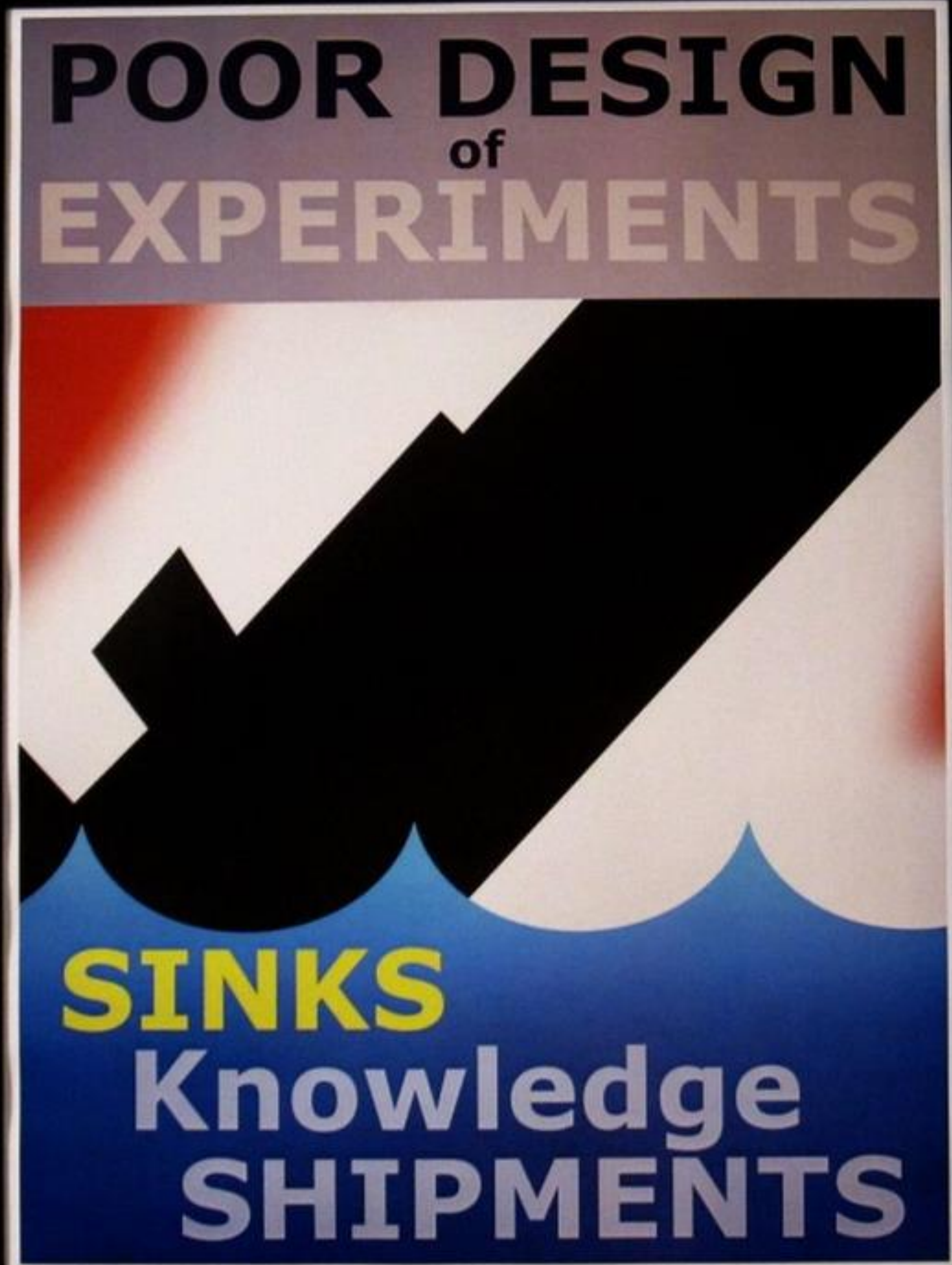



Box-Cox transformation
STATISTICA - данные к зачёту 25.03
File Edit View Insert Format Statistics Data Mining Graphs Tools Data Window Help
Add to Workbook Add to Report Add to MS Word
Box-Cox Transformation: данные к зачёту
Box-Cox
Variables:
Box-Cox transformation
Max. iterations: 40
Min. lambda: -5,
Max. lambda: 5,
Epsilon (convergence): ,000
Shift variables with minimum<=0 to: 1,
OK
Cancel
Options
Sel Cond
Box-Cox Results: данные к зачёту 25.03
Box-Cox Results
Summary
Histograms and normal probability plots
Search history plots
Number of variables to add: 1
Add variables
Write back to input spreadsheet...
Summary
Cancel
Options
By Group
ansformation (данные к ...
Box-Cox transformation (данные
Trans
средняя ч-та соц. контактов летом (в час) original
средняя ч-та соц. контактов летом (в час) transformed
1 0,59541 -0,54971
2 1,25926 0,22469
3 3,34491 1,05839
4 5,86905 1,46194
5 6,37500 1,51713
6 0,97222 -0,02826
7 7,50000 1,62268
8 1,37037 0,30425
9 17,77778 2,12330
15 13,60000 1,97811
16 22,66667 2,24768
17 1,66667 0,48277
18 1,52778 0,40438
19 2,02381 0,65230
20 2,51852 0,83466
21 3,14815 1,01178
22 3,44048 1,07984
23 1,41667 0,33511
24 1,07407 0,07089
25 2,50000 0,82865
26 1,97917 0,63320
27 0,65476 -0,44415
28 0,87037 -0,14101
0,94184
0,08352
ёту 25.03)
Data statistics (данные к зачёту 25.03)
Lambda
Shift
Mean
Standard deviation
Lower Confidence Limit
Upper Confidence Limit
Formula for Box-Cox transformation
. контактов летом (в час) -0,223247 0,00 0,634369 0,739752 -0,453717 -0,001112 ((v14^(-0,223247))-1)/(-0,223247)

5. Ранговая трансформация

Ранжирование (ranking) наблюдений: выстраиваем их по порядку от самого маленького значения к наибольшему, или наоборот.
После этого можно анализировать параметрическими тестами! Но при этом происходит большая потеря информации.

Очень важно после трансформации проверить, стало ли лучше! (Может, стало хуже.)


POOR DESIGN
of
EXPERIMENTS
SINKS
Knowledge
SHIPMENTS

Практическое занятие

1. Power analysis для т-тестов: вставить какие-нибудь произвольные значения и посчитать мощность и необходимый размер выборки, построить кривые взаимосвязей между N, Effect size, Power.
2. Файл crabs: проверить на нормальное распределение, outliers (переменная «вес»), нарисовать чудесные картинки из Описательной статистики
3. Crabs: трансформация Box-Cox; Rank - ранговая трансформация (сначала создать столбик); Стандартизация.