

## Прикладная статистика. Занятие 2. Параметрическая проверка гипотез.

15 февраля 2011 г.

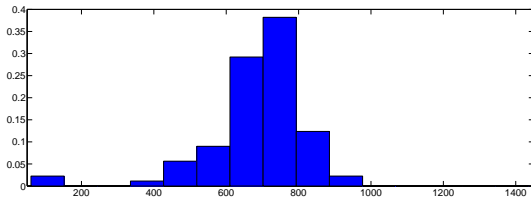
## Продолжительность жизни крыс

В исследовании принимало участие 194 крысы.  
105 из них держали на строгой диете, оставшиеся 89 — на диете *ad libitum* (ели, сколько хотели).

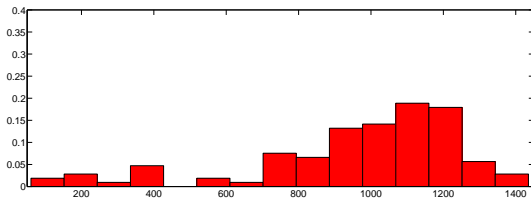
Имеющиеся данные: продолжительность жизни крыс в каждой из групп.  
Вопрос: влияет ли диета на продолжительность жизни?



## Продолжительность жизни крыс

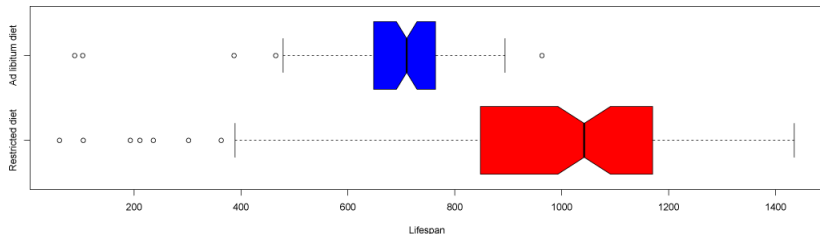


Продолжительность жизни крыс на диете ad libitum ( $n = 89$ )



Продолжительность жизни крыс на строгой диете ( $n = 105$ )

## Продолжительность жизни крыс



Boxplot

От центра к краям: медиана; 95% доверительный интервал для медианы; квартили; точки данных, ближайшие (изнутри) к концу отрезка длиной  $1.5 \times IQR$ ; точки, не попадающие в этот интервал.

## Продолжительность жизни крыс

$H_0$ : продолжительность жизни крыс не меняется при ограничении диеты.

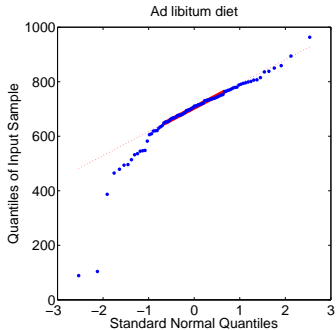
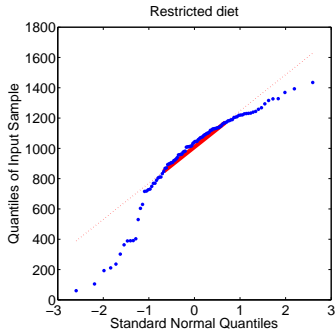
$H_1$ : крысы на строгой диете живут в среднем дольше.

Применяем критерий Стьюдента для двух выборок с неизвестной неравной дисперсией:  $p = 2 \times 10^{-15}$ ; односторонний 95% доверительный интервал для увеличения продолжительности жизни —  $[227, \infty]$ .

$H_1$ : средняя продолжительность жизни крыс меняется при ограничении диеты.

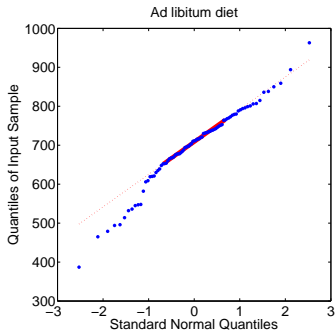
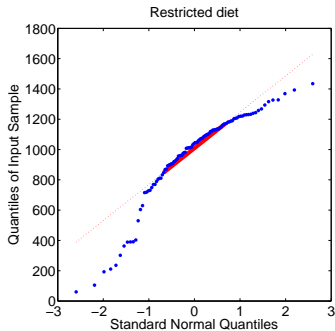
Применяем критерий Стьюдента для двух выборок с неизвестной неравной дисперсией:  $p = 4 \times 10^{-15}$ ; 95% доверительный интервал для изменения продолжительности жизни —  $[217, 344]$ .

# Продолжительность жизни крыс



Quantile-quantile plots

## Продолжительность жизни крыс

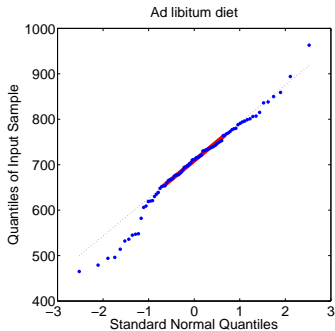
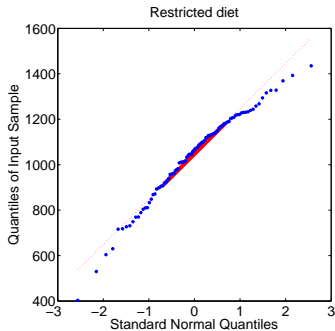


$$n_1 = 105, \quad n_2 = 87.$$

Критерий Стьюдента (односторонняя альтернатива):

$$p = 6 \times 10^{-15}; \quad ci = [216, \infty]$$

## Продолжительность жизни крыс



$$n_1 = 96, \quad n_2 = 86.$$

Критерий Стьюдента (односторонняя альтернатива):

$$p = 4 \times 10^{-31}; \quad ci = [299, \infty]$$



## Лечение анорексии

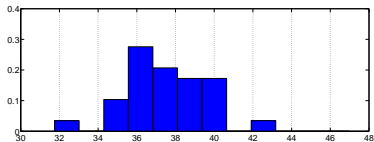
Institute of Psychiatry, London: 29 девушек до 17 лет проходили 12-ти недельный курс лечения от анорексии (семейная терапия).

Известен вес до начала и после окончания терапии.

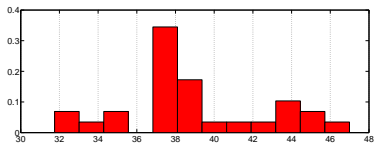
Эффективна ли терапия? Каков её эффект?



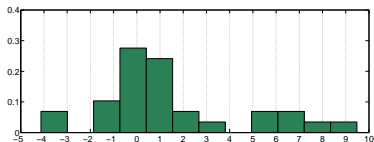
# Лечение анорексии



Вес до начала лечения

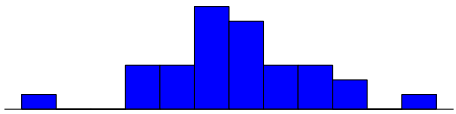
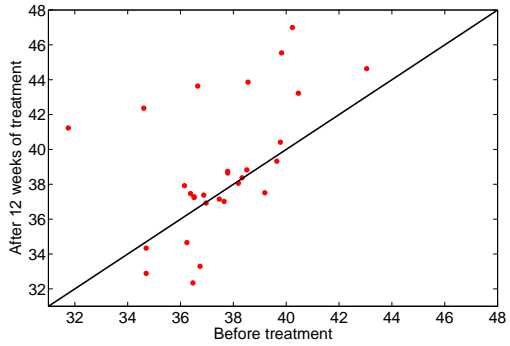
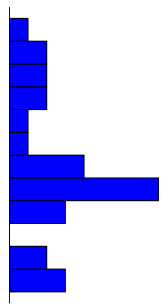


Вес после 12 недель терапии



Прирост веса

# Лечение анорексии



## Лечение анорексии

$H_0$ : терапия неэффективна, вес не меняется.

$H_1$ : в результате терапии вес в среднем увеличивается.

Применяем парный критерий Стьюдента:  $p = 0.0175$ ; односторонний 95% доверительный интервал для прироста веса —  $[0.3167, \infty)$ .

$H_1$ : в результате терапии средний вес как-то меняется.

Применяем парный критерий Стьюдента:  $p = 0.035$ ; 95% доверительный интервал для изменения веса —  $[0.1029, 2.625]$ .

Игнорируем связь между выборками и применим обычный двухвыборочный критерий Стьюдента:

- для односторонней альтернативы  $p = 0.0496$ ,  $ci = [0.0036, \infty)$ .
- для двусторонней альтернативы  $p = 0.0991$ ,  $ci = [-0.2653, 2.9932]$ .

## Тревожность пожилых людей

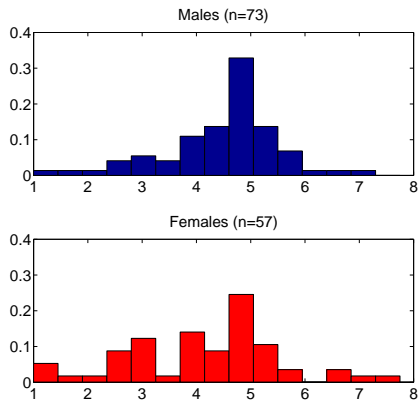
University of Massachusetts Medical School, 1999: изучалась вариация психологических и физиологических факторов в зависимости от времени года.

Один из измеренных признаков — тревожность (шкала от 1 до 9).

Отберём испытуемых старше 60 лет, возьмём среднее значение тревожности за 4 сезона и проверим, отличается ли его дисперсия в выборках из 73 мужчин и 57 женщин.



## Тревожность пожилых людей

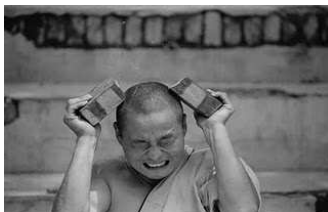


Критерий Фишера:  $p = 0.0308$ ,  $ci = [1.0516, 2.8496]$ .

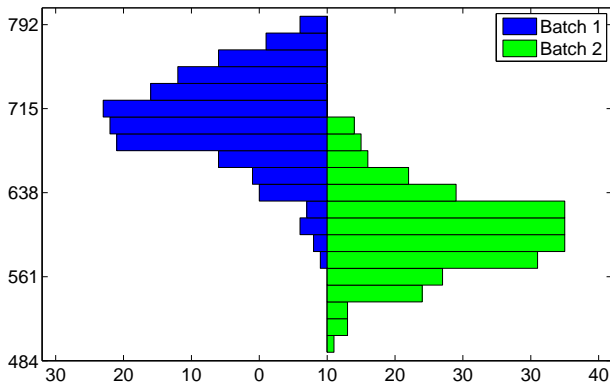
## Прочность керамики

NIST industry ceramics consortium for strength optimization of ceramic strength, 1996

Были собраны данные о прочности материала 440 керамических изделий из двух партий по 220 в каждой, цель — проверить, одинакова ли дисперсия прочности в разных партиях.



## Прочность керамики



Гипотеза нормальности не отклоняется критерием Лиллиефорса.  
Критерий Фишера:  $p = 0.1721$ ,  $ci = [0.9225, 1.5690]$ .



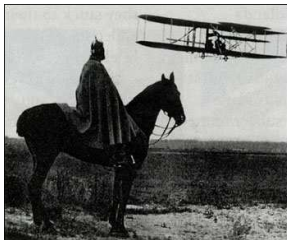
## Гибель кавалеристов

Bortkiewicz L., Das Gesetz der kleinen Zahlen, Leipzig, 1898.

В Генеральный штаб прусской армии ежегодно в течение 20 лет поступали от 10 кавалерийских корпусов данные о количестве кавалеристов, погибших в результате гибели под ними коня (всего 200 донесений).

Количество погибших	0	1	2	3	4	5	6	Всего
Количество донесений	109	65	22	3	1	0	0	200

Можно ли описать эти данные распределением Пуассона?



## Гибель кавалеристов

Распределение Пуассона:

$$\mathbb{P}(Y = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Оценим параметр  $\lambda$  методом максимального правдоподобия:

$$L(\lambda, Y) = e^{-n\lambda} \prod_{i=1}^n \frac{\lambda^{Y_i}}{Y_i!},$$

$$LL(\lambda, Y) = -n\lambda + \sum_{i=1}^n Y_i \ln \lambda - \sum_{i=1}^n \ln Y_i!,$$

$$\frac{\partial LL}{\partial \lambda} = -n + \sum_{i=1}^n \frac{Y_i}{\lambda},$$

$$\hat{\lambda} = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \text{ — оценка максимального правдоподобия.}$$

$$\bar{Y} = 0.61.$$

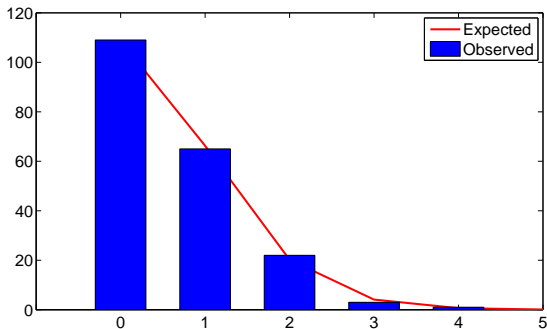
## Гибель кавалеристов

Проверим гипотезу  $H_0: Y \sim P(0.61)$  против общей альтернативы.

Критерий хи-квадрат:

$$O: [109, 65, 26],$$

$$E: [108.6702, 66.28, 25.0410].$$



$$p = 0.7734.$$

Прикладная статистика  
Семинар 2. Параметрическая проверка гипотез.

Рябенко Евгений  
riabenko.e@gmail.com