Прикладная статистика. Занятие 2. Параметрическая проверка гипотез.

21 февраля 2012 г.

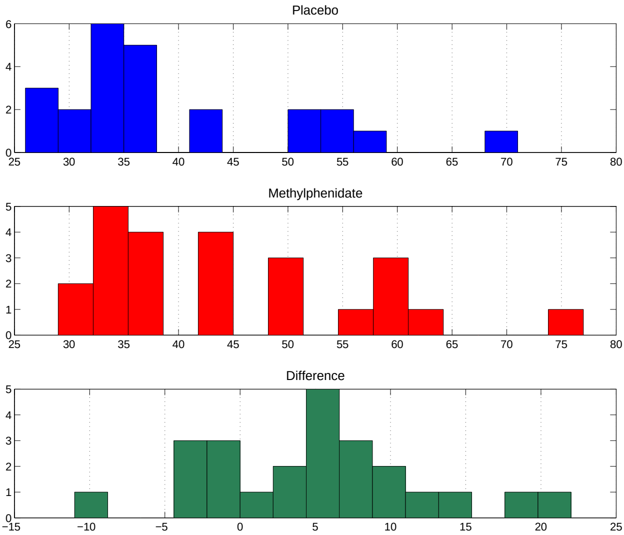
Метилфенидат

Pearson et al, 2003, Treatment effects of methylphenidate on behavioral adjustment in children with mental retardation and ADHD: исследовалось влияние метилфенидата на способность к подавлению импульсивных поведенческих реакций умственно отсталых детей с синдромом дефицита внимания и гиперактивности. Каждый испытуемый в течение недели принимал либо препарат, либо плацебо, а в конце недели проходил тест. На втором этапе плацебо и препарат менялись, после недельного курса каждый испытуемый проходил второй тест.

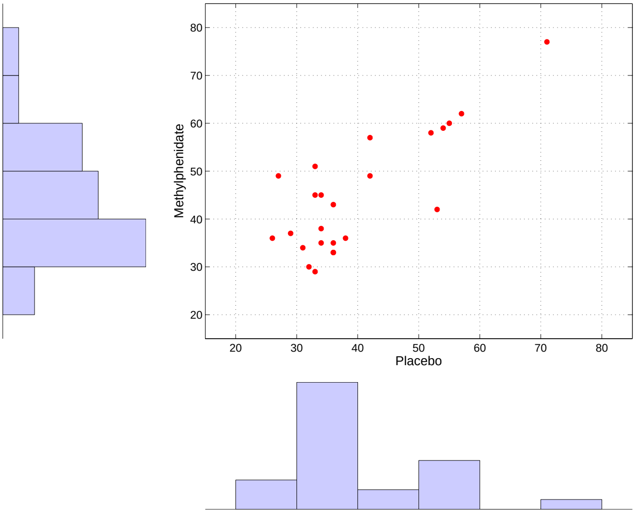
Для 24 испытуемых известны результаты в норме и после недельного курса препарата.

Эффективен ли препарат? Каков его эффект?

Метилфенидат



Метилфенидат



Парный критерий Стьюдента

H_0: терапия неэффективна, способность к подавлению импульсивных поведенческих реакций не меняется.

H_1: в результате терапии способность к подавлению импульсивных поведенческих реакций увеличивается.
Применяем парный критерий Стьюдента: $p = 0.0019$; верхний 95% доверительный предел для увеличения — $CU = -2.3212$.

H_1: в результате терапии способность к подавлению импульсивных поведенческих реакций меняется.
Применяем парный критерий Стьюдента: $p = 0.0038$; 95% доверительный интервал для изменения — $CI = [-8.1414, -1.7752]$.

Игнорируем связь между выборками и применим обычный двухвыборочный критерий Стьюдента:

- для односторонней альтернативы $p = 0.0766$, $CU = 0.7734$;
- для двусторонней альтернативы $p = 0.1532$, $CI = [-11.8313, 1.9146]$.

Доверительный интервал

$H_0\colon \mu = 0, \quad H_1\colon \mu \neq 0.$
Если нулевая гипотеза справедлива:

$$T(X) = \frac{\sqrt{x}\left(\bar{X} - \mu\right)}{s} \sim t_{n-1} \Rightarrow$$

$$P\left(-t_{1-\frac{\alpha}{2},n-1} \leq \frac{\sqrt{n}\left(\bar{X} - \mu\right)}{s} \leq t_{1-\frac{\alpha}{2},n-1}\right) = 1 - \alpha,$$

$$P\left(\bar{X} - t_{1-\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{1-\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right) = 1 - \alpha.$$

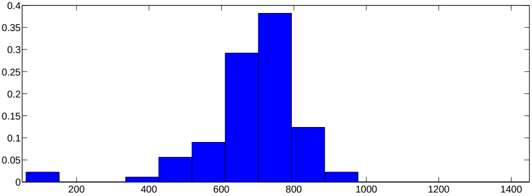
Таким образом, $\bar{X} \pm t_{1-\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}$ — $100\,(1-\alpha)\%$ доверительный интервал для среднего.
Односторонние доверительные пределы аналогично выводятся из критерия, проверяющего гипотезу против односторонней альтернативы.

Продолжительность жизни крыс

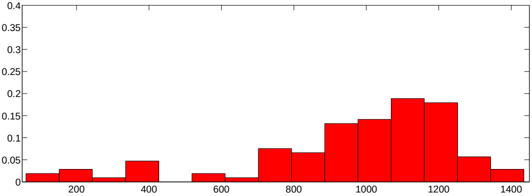
Walford and Weindruch, 1988, The Retardation of Aging and Disease by Dietary Restriction: в исследовании принимало участие 194 крысы. 105 из них держали на строгой диете, оставшиеся 89 — на диете ad libitum. Имеющиеся данные: продолжительность жизни крыс в каждой из групп.

Вопрос: влияет ли диета на продолжительность жизни?

Продолжительность жизни крыс

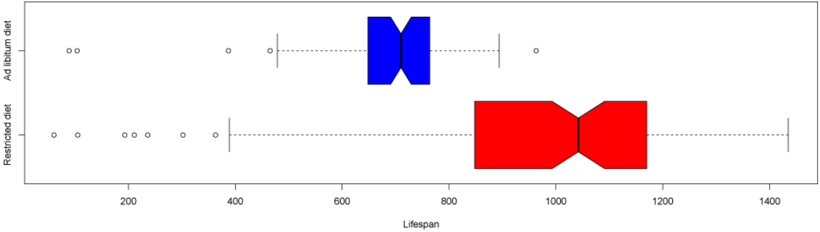


Продолжительность жизни крыс на диете ad libitum $(n = 89)$



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

Продолжительность жизни крыс



Ящик с усами

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.

Двухвыборочный критерий Стьюдента

H_0: продолжительность жизни крыс не меняется при ограничении диеты.

H_1: крысы на строгой диете живут дольше.
Применяем критерий Стьюдента для двух выборок с неизвестной неравной дисперсией: $p = 2 \times 10^{-15}$; нижний 95% доверительный предел для увеличения продолжительности жизни — $CL = 227$.

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 — $CI = [217, 344]$.

Проверка нормальности

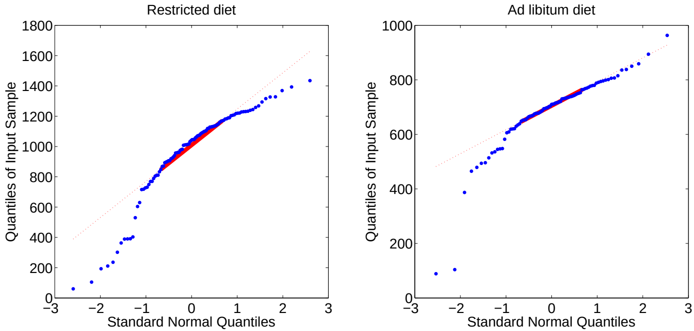
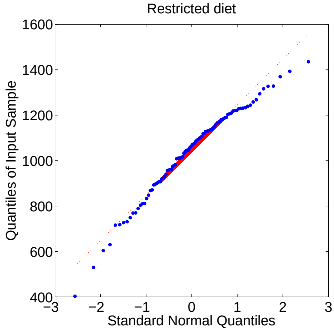
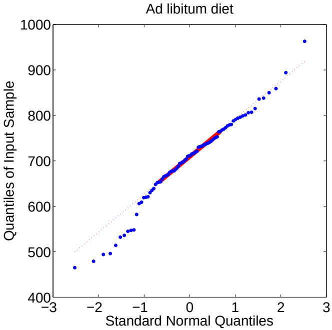


График Q-Q (квантиль-квантиль)

Критерий Шапиро-Уилка отклоняет гипотезу нормальности: $p_1 = 1.7 \times 10^{-6}, p_2 = 1.5 \times 10^{-7}$.

Исключение выбросов





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

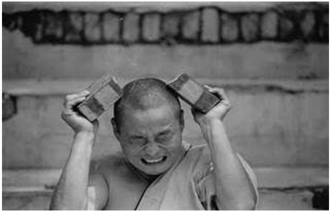
Критерий Шапиро-Уилка: $p_1 = 0.0443, p_2 = 0.0960.$
Критерий Стьюдента:

- для односторонней альтернативы $p = 4 \times 10^{-32}, \quad CL = 298;$
- для двусторонней альтернативы $p = 9 \times 10^{-32}, \quad CI = [290, 382]$.

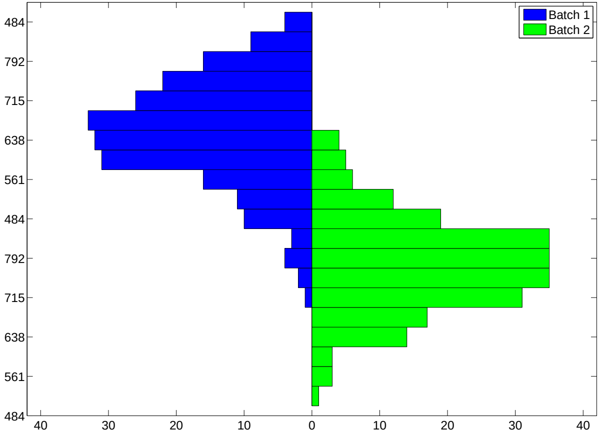
Прочность керамики

NIST industry ceramics consortium for strength optimization of ceramic strength, 1996: собраны данные о прочности материала 440 керамических изделий из двух партий по 220 в каждой.

Цель — проверить, одинакова ли дисперсия прочности в разных партиях.



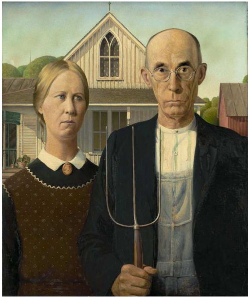
Прочность керамики



Гипотеза нормальности не отклоняется критерием Шапиро-Уилка ($p_1 = 0.2062, p_2 = 0.7028$).
Критерий Фишера: $p = 0.1721, \quad CI = [0.9225, 1.5690]$.

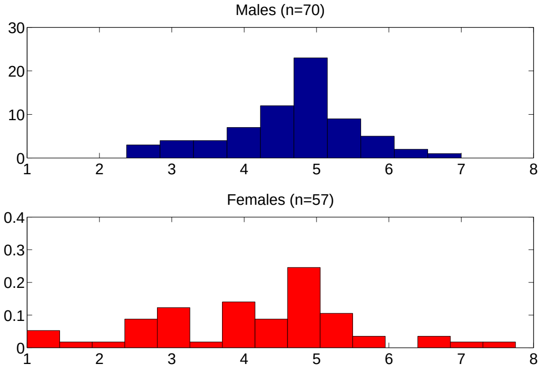
Тревожность пожилых людей

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 70 мужчин и 57 женщин.

Тревожность пожилых людей



Гипотеза нормальности не отклоняется критерием Шапиро-Уилка ($p_1 = 0.077, p_2 = 0.5774$).
Критерий Фишера: $p = 0.00039151, \quad CI = [0.2431, 0.6651]$.

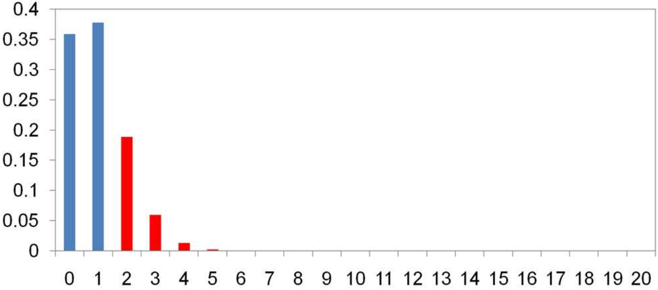
Доля дефектных изделий

Кобзарь, 2006, Прикладная математическая статистика, задача 227: нормируемый уровень дефектных изделий в партии $p_0 = 0.05$. Из партии извлечена выборка $n = 20$ изделий, в которой обнаружены при проверке $m = 2$ дефектных.

Необходимо проверить гипотезу о том, что доля дефектных изделий в партии не превосходит нормируемого значения.

$H_0: p \leq p_0, \quad H_1: p > p_0.$

Точный критерий



$$p - value = \sum_{k=m}^{n} \frac{p_0^k (1 - p_0)^{n-k} n!}{k! (n-k)!} = \sum_{k=2}^{20} \frac{0.05^k 0.95^{20-k} 20!}{k! (20-k)!} \approx 0.26.$$

Нулевая гипотеза не отклоняется.

Приближения

При больших объёмах выборки вычисление достигаемого уровня значимости точного критерия затруднительно.
Используют приближение нормальным распределением:

$$m \sim Bin\left(n, p\right) \approx N\left(np, np\left(1-p\right)\right).$$

При $m \sim Bin\left(n, p_0\right)$

$$z = \frac{m - np_0}{\sqrt{np_0\left(1-p_0\right)}} \approx N\left(0, 1\right).$$

Для одностороннего варианта гипотезы используется статистика:

$$z = \frac{m - np_0 - \frac{1}{2}}{\sqrt{np_0\left(1-p_0\right)}}.$$

Гипотеза $H_0\colon p \leq p_0$ отклоняется в пользу альтернативы $H_1\colon p > p_0$, если $z > z_\alpha$.

В задаче $z \approx 0.51, p - value \approx 0.31$.

Доверительные интервалы

$\frac{m-np}{\sqrt{np(1-p)}} \approx N(0,1)$.

$$P\left(-z_{1-\frac{\alpha}{2}} \leq \frac{m-np}{\sqrt{np(1-p)}} \leq z_{1-\frac{\alpha}{2}}\right) \approx 1-\alpha,$$

$$P\left(\frac{m}{n} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \leq p \leq \frac{m}{n} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right) \approx 1-\alpha;$$

подставим вместо p оценку $\hat{p} = \frac{m}{n}$, получим

$$\hat{p} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

— $100(1-\alpha)$% доверительный интервал Вальда (приближённый).

В задаче 95% доверительный интервал Вальда $CI_{Wald} = [-0.0315, 0.2315]$.

Доверительные интервалы

Более точный доверительный интервал Уилсона:

$$\frac{\hat{p} + \frac{1}{2n} z_{1-\frac{\alpha}{2}}^2 \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{1-\frac{\alpha}{2}}^2}{4n^2}}}{1 + \frac{1}{n} z_{1-\frac{\alpha}{2}}^2}.$$

В задаче 95% доверительный интервал Уилсона $CI_{Wilson} = [0.0279, 0.3010]$.

Доли дефектных изделий

Кобзарь, 2006, Прикладная математическая статистика, задача 226: в двух партиях объёмами $n_1 = 100$ шт. и $n_2 = 200$ шт. обнаружено соответственно $m_1 = 3$ и $m_2 = 5$ дефектных приборов.

Необходимо проверить гипотезу о равенстве долей дефектных приборов в партиях.

$H_0\colon p_1 = p_2, \quad H_1\colon p <\neq> p_0.$

Приближённый двухвыборочный критерий

$$z = \frac{\frac{m_1}{n_1} + \frac{1}{2n_1} - \frac{m_2}{n_2} - \frac{1}{2n_2}}{\sqrt{\frac{m_1+m_2}{n_1+n_2}\frac{n_1+n_2-m_1-m_2}{n_1+n_2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}.$$

Если $|z| > z_{\frac{1+\alpha}{2}}$, $z > z_\alpha$ или $z < z_{1-\alpha}$, нулевая гипотеза отклоняется в пользу альтернатив

$$H_1 : p_1 \neq p_2, \quad H_1 : p_1 > p_2, \quad H_1 : p_1 < p_2.$$

В задаче $z = 0.376, p - value = 0.71$ (двусторонняя альтернатива).

Прикладная статистика
Семинар 2. Параметрическая проверка гипотез.

Рябенко Евгений
riabenko.e@gmail.com