Проверка гипотезы о равенстве дисперсий (критерий Фишера и критерий Ансари-Брэдли)

1. Пусть выборка $X_1, \dots, X_n$ порождена СВ X с непрерывным распределением F(t-μ), а выборка $Y_1, \dots, Y_n$ - СВ Y с распределением F($\frac{t-\mu}{\Delta}$) , $\Delta > 0$. Предполагается, что DX$< \infty$, и выполняется условие $\int_{-\infty-}^{\infty} tf(t)dt = 0$. Покажите, что из справедливости гипотезы H_1: $\Delta < 1$ следует, что DX>DY.

2. Согласно опросам 29 семей, проводившимся в 1968 году в юго-западном регионе Англии, выборочное среднее арендной платы за меблированную квартиру составило 2,5£, а выборочная дисперсия 0,67 £^{2}. В Уэльсе выборочное среднее арендной платы 16 семей составило 2,06£, а выборочная дисперсия 0,42 £^{2}. Проверьте гипотезу о равенстве дисперсий арендной платы в двух регионах Великобритании. Уровень значимости считать равным 0.05. Предполагается, что все наблюдения имеют гауссовское распределение

3. Станок штампует детали, размер которых соответствует заданному нормативу, т.е. вероятность превышения и занижения нормативного размера одинакова. Технологи провели наладку станка для того, чтобы уменьшить отклонения размеров изготовленных деталей от размера, требуемого стандартом. До и после наладки случайным образом было выбрано по 11 деталей. Оказалось, что размер деталей, выбранных для наладки, составил (в мм): 52,4; 56,1; 48,6; 46,5; 46,0; 42,2; 48,8; 56,6; 59,8; 49,7; 51,6. Размер деталей, изготовленных после наладки станка (в мм): 49,3; 47,7; 52,9; 48,3; 49,1; 46,4; 47,0; 52,0; 51,5; 51,2; 49,8. Можно ли считать, опираясь на эти данные, что точность изготовления деталей увеличилась после наладки? Уровень значимости считать равным 0,05.

Домашнее задание

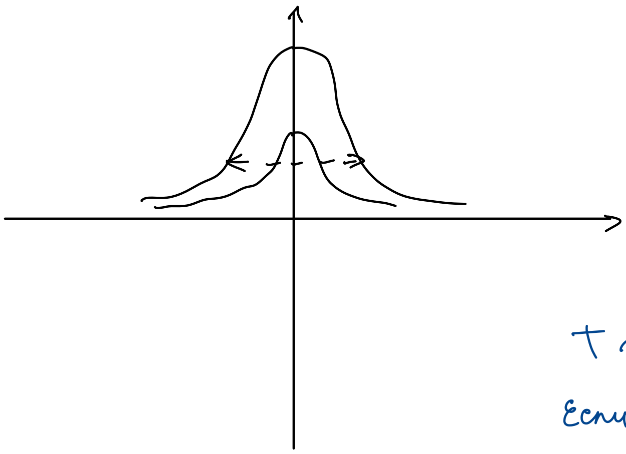
1. В метеорологии принято характеризовать температуру месяца ее средним значением (среднее значение температуры месяца равно сумме температур всех дней данного месяца, деленной на число дней в этом месяце). В таблице ниже приведены значения средней температуры января в г. Саратове и г. Алатыре. Проверьте равенство дисперсий среднеянварских температур в городах Саратове и Алатыре.

Год	1891	1892	1893	1894	1895	1896	1897
Саратов	−19,2	−14,8	−19,6	−11,1	−9,4	−16,9	−13,7
Алатырь	−21,8	−15,4	−20,8	−11,3	−11,6	−19,2	−13,0

Год	1899	1911	1912	1913	1914	1915	
Саратов	−4,9	−13,9	−9,4	−8,3	−7,9	−5,3	
Алатырь	−7,4	−15,1	−14,4	−11,1	−10,5	−7,2	

2. Решите предыдущую задачу, применяя критерий Ансари-Брэдли. Предварительно центрируйте данные выборочными медианами.

3. За последние 5 лет выборочная дисперсия доходности актива А составила 0.04, а выборочная дисперсия доходности актива Б составила 0.05.Есть ли основание утверждать (на уровне значимости 0.05), что вложения в актив А менее рискованны, чем вложения в актив Б? Предполагается, что доходности активов являются гауссовскими СВ.



$H_0: \Delta = \frac{\sigma_y}{\sigma_x}$

$H_a:$ $\Delta > 1$
$\Delta < 1$
$\Delta \neq 1$

$$T(x,y) = F_{n,m} = \frac{\frac{1}{n-1}\sum (y_i - \bar{y})^2}{\frac{1}{m-1}\sum (x_i - \bar{x})^2} = \frac{S_y^2}{S_x^2} \sim$$

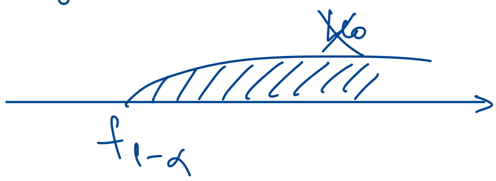
$\sim F(n-1; m-1)$
k_1, k_2

$T \sim F(n,m) \Rightarrow \frac{1}{T} \sim F(m,n)$

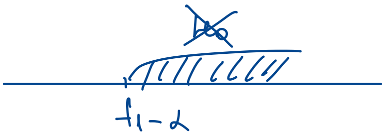
Если $H_a: \frac{\sigma_y}{\sigma_x} = \Delta < 1$

и $S_y^2 < S_x^2$, то берём $F_{m,n} = \frac{1}{F_{n,m}} = \frac{S_x^2}{S_y^2} \sim$

$\sim F(m-1, n-1)$



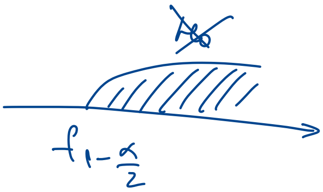
Если $H_a: \Delta > 1$ и $S_y^2 > S_x^2 \Rightarrow$ берём $F_{n,m} = \frac{S_y^2}{S_x^2} \sim F(n-1, m-1)$



Если $H_a: \Delta \neq 1$

и $S_y^2 > S_x^2 \Rightarrow T = F_{n,m} = \frac{S_y^2}{S_x^2} \sim F(n-1, m-1)$

$S_y^2 < S_x^2 \Rightarrow T = \frac{1}{F_{n,m}} = \frac{S_x^2}{S_y^2} \sim F(m-1, n-1)$



Проверка гипотезы об однородности против альт. растяж./сжатие

$X_1, \dots, X_m \sim F(t-\mu)$

$Y_1, \dots, Y_n \sim F\left(\frac{t-\mu}{\Delta}\right) \qquad \Delta > 0$

„μ" - мешающий параметр; $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ незав.

Если $\int_{-\infty}^{+\infty} t f(t)\,dt = 0$ и $\exists DX$

$$EX = \int_{-\infty}^{+\infty} t f(t-\mu)\,dt = \langle z = t-\mu \rangle = \int_{-\infty}^{+\infty} (z+\mu) f(z)\,dz = \mu$$

Тогда $DX = \int_{-\infty}^{+\infty} (t-\mu)^2 f(t-\mu)\,dt = \int_{-\infty}^{+\infty} z^2 f(z)\,dz$

$$DY = \int_{-\infty}^{+\infty} (t-\mu)^2 \frac{1}{\Delta} f\left(\frac{t-\mu}{\Delta}\right) dt = \int_{-\infty}^{+\infty} \Delta^2 z^2 f(z)\,dz =$$

$z = \frac{t-\mu}{\Delta} \quad dz = \frac{1}{\Delta} dt \qquad = \Delta^2 DX; \quad \frac{DY}{DX} = \Delta^2$

$H_0: \Delta = 1$

$H_1: \Delta < 1 \quad (DY < DX)$

$H_2: \Delta > 1 \quad (DY > DX)$

$H_3: \Delta \neq 1 : \quad (DY \neq DX)$

$$\int_{-\infty}^{+\infty} \cancel{(t-\mu)}\, f(t-\mu)\,dt =$$

$=$

ДО СЮДА

КОНСПЕКТ МАКСИМА КЛЫКОВА

2. Согласно опросам 29 семей, проводившимся в 1968 году в юго-западном регионе Англии, выборочное среднее арендной платы за меблированную квартиру составило 2,5£, а выборочная дисперсия 0,67 £^{2}. В Уэльсе выборочное среднее арендной платы 16 семей составило 2,06£, а выборочная дисперсия 0,42 £^{2}. Проверьте гипотезу о равенстве дисперсий арендной платы в двух регионах Великобритании. Уровень значимости считать равным 0.05. Предполагается, что все наблюдения имеют гауссовское распределение

$n = 16, \; m = 29$

$\hat{D}x = 0{,}67, \quad \hat{D}y = 0{,}42$

$H_0: \Delta = \frac{\sigma_y}{\sigma_x} = 1$

$H_a: \Delta \neq \frac{\sigma_y}{\sigma_x} \neq 1$

$$T = F_{n,m} = \frac{s_y^2}{s_x^2} = \frac{\frac{1}{n-1}\sum(y_i - \bar{y})^2}{\frac{1}{m-1}\sum(x_i - \bar{x})^2} \sim F(n-1, m-1)$$

Ответ: дисперсии ... короче нормальный ответ.

$\hat{D}x = \frac{1}{n}\sum(\ )^2$

$s_x^2 = \hat{D}x \frac{m}{m-1} = 0{,}67 \cdot \frac{29}{28} = 0{,}694$

$s_y^2 = \hat{D}y \frac{n}{n-1} = 0{,}42 \cdot \frac{16}{15} = 0{,}448$ } несмещенная

$t_{кр} = f_{1-\frac{\alpha}{2}}(28, 15) = f_{0{,}975}(28; 5) = 2{,}64$

$s_x^2 > s_y^2 \Rightarrow T = \frac{1}{F_{n,m}} = F_{m,n}$

$t_{набл} = \frac{s_x^2}{s_y^2} = \frac{0{,}694}{0{,}448} = 1{,}549$