Галанов Ю.И.

Лабораторный практикум по мат. статистике

Проверка гипотезы о независимости

Задание

1. Опишите этапы проведения корреляционного анализа на данном примере.
2. Проверьте влияние объема выборки на чувствительность данного критерия к величине параметров линейной связи.
3. Прокомментируйте этапы составления программы.
4. Исследуйте влияние объема выборки и значение параметров модели на надежность статистических выводов.

$N := 1000 \qquad k := 0..N-1$

$w(x,y) := \sin(2\cdot\pi\cdot x)\cdot\sqrt{-2\ln(y)} \qquad k := 0..N-1$

Модель линейной регрессии $\qquad Y(x,mx,my,sx,sy,\rho) := my + \rho\cdot\frac{sy}{sx}\cdot(x-mx)$

$mx := 0 \qquad my := 1 \qquad sx := 10 \qquad sy := 50 \qquad \rho := 0$

$X_{k,0} := mx + sx\cdot w(rnd(1),rnd(1)) \qquad X_{k,1} := \left(X_{k,0}\right)^2 + sy\cdot w(rnd(1),rnd(1))$

$X_{k,1} := \left(X_{k,0}\right) + sy\cdot w(rnd(1),rnd(1))$

$X := csort(X,0)$

$Mx := mean\left(X^{\langle 0\rangle}\right) \qquad My := mean\left(X^{\langle 1\rangle}\right) \qquad Sx := stdev\left(X^{\langle 0\rangle}\right) \qquad Sy := stdev\left(X^{\langle 1\rangle}\right)$

$corr\left(X^{\langle 0\rangle},X^{\langle 1\rangle}\right) = 0.134026$

$$r := \frac{\frac{1}{N}\cdot\sum_k\left[\left[\left(X^{\langle 0\rangle}\right)_k - Mx\right]\cdot\left[\left(X^{\langle 1\rangle}\right)_k - My\right]\right]}{Sx\cdot Sy}$$

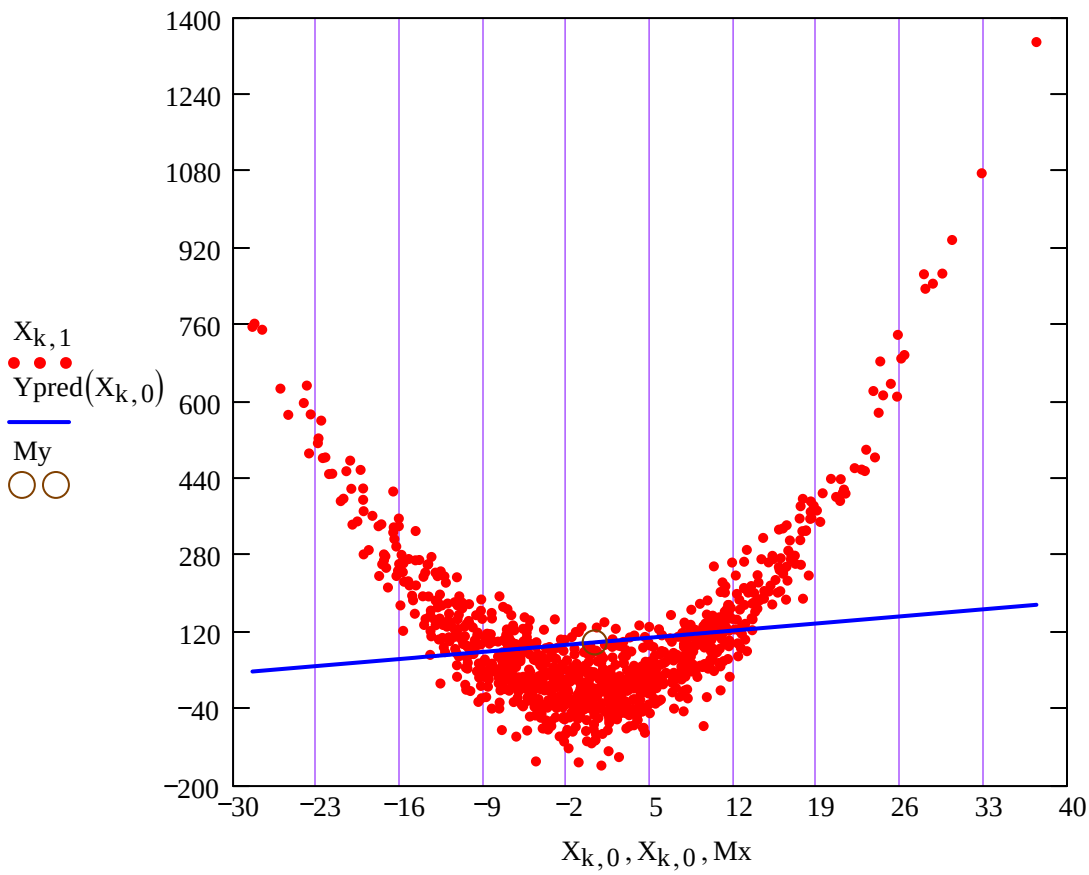
$r = 0.134026$

$$T := \frac{|r|}{\sqrt{1-r^2}} \cdot \sqrt{N-2} \qquad T = 4.272575 \qquad \alpha := 0.001 \qquad C := qt\left(1-\frac{\alpha}{2}, N-2\right)$$

$$C = 3.300302$$

Предсказанное значение

$$Ypred(x) := Y(x, Mx, My, Sx, Sy, r)$$



$mx = 0$ $\qquad Mx = 0.343491$

$my = 1$ $\qquad My = 99.199438 \qquad C = 3.300302 \quad \Gamma = 4.272575$

$sx = 10$ $\qquad Sx = 9.961014$

$sy = 50$ $\qquad Sy = 157.12902$

$\rho = 0$ $\qquad r = 0.134026$