Normal correlation analysis

Part 5 of regression

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]}$$

$$\parallel$$

$$f(x,y;\sigma_x^2,\sigma_y^2,\rho,\mu_x,\mu_y)$$

$$(X_i, Y_i) \qquad (x_1,y_1)\ (x_2,y_2) \ldots (x_n,y_n)$$

use MLE to estimate values of parameters $\mu_x, \mu_y, \sigma_x^2 \ldots$

$$L(\sigma_x^2,\sigma_y^2,\rho,\mu_x,\mu_y;\vec{x},\vec{y}) = \prod_{k=1}^{n} f(x_k,y_k;\sigma_x^2,\sigma_y^2,\rho,\mu_x,\mu_y)$$

maximize via $\frac{\partial}{\partial\rho}\ln L = 0$

$\frac{\partial}{\partial\mu_x}\ln L = 0 \ldots$

$$\mu_x = \bar{x} \qquad \sigma_x^2 = \frac{1}{n}\sum(x_i-\bar{x})^2 = S_{xx}$$

$$\mu_y = \bar{y} \qquad \sigma_y^2 = \frac{1}{n}\sum(y_i-\bar{y})^2 = S_{yy}$$

$$\sigma_{x,y} = \frac{1}{n}\sum(x_i-\bar{x})(y_i-\bar{y}) = S_{xy}$$

$$\rho = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

get estimators $\hat{\mu}_x = \bar{X}$ $\quad \hat{\mu}_y = \bar{y}$ $\quad \hat{\sigma}_x^2 = S_{xx}$ $\quad \hat{\sigma}_y^2 = S_{yy}$

$$\hat{\rho} = \frac{S_{xy}}{\sqrt{S_{xx} S_{xy}}}$$

<u>ex:</u> want interval estimate for σ_x^2 $\quad c\, \hat{\sigma}_x^2 / \sigma_x^2 \quad \chi^2_{n-1}$

using χ^2, t variables, can get interval estimates

can do hyp testing for values of

$\sigma_x^2, \sigma_y^2, \mu_x, \mu_y$

$$\hat{\rho} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

sample value for this is referred to by "r"

$$\sigma^2_{y|x} = \sigma_y^2 (1 - \rho^2)$$

$-1 \leq \rho \leq 1$

$\rho = 0$ (x & y indep)

$\rho = \pm 1$

$$\rho^2 = \frac{\sigma_y^2 - \sigma_{y|x}^2}{\sigma_y^2} = 1 - \underbrace{\left(\frac{\sigma_{y|x}^2}{\sigma_y^2}\right)}$$

fraction of variance "due to x"

the fraction of variance remains after x is known

Special case: if $\rho = 0$

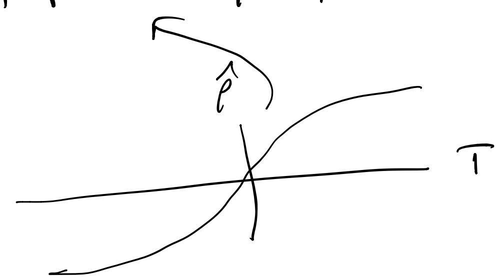
$\hat{\rho}$

$$T = \boxed{\hat{\rho}\sqrt{\frac{n-2}{1-\hat{\rho}^2}}}$$

turns out this is a t-distributed variable w/ n-2 degrees of freedom

$$\hat{\rho} = \frac{T}{\sqrt{n-2+T^2}}$$

monotonic increasing odd 1-1 function



$$P(\hat{\rho} > \lambda \mid \rho = 0)$$

$$\Updownarrow$$

$$P\left(\hat{\rho}\sqrt{\frac{n-2}{1-\hat{\rho}^2}} > \lambda\sqrt{\frac{n-2}{1-\lambda^2}}\right)$$

$$\frac{1}{T} \qquad P\left(T > \lambda\sqrt{\frac{n-2}{1-\lambda^2}}\right)$$

What if $\rho \neq 0$? $\quad \hat{\rho}$

$$P(\hat{\rho} \geq .7 \mid \rho = .5)?$$

Can write an exact expression for density func. for $\hat{\rho}$

$$f(r) = \frac{(n-2)\,\Gamma(n-1)\,(1-\rho^2)^{\frac{n-1}{2}}\,(1-r^2)^{\frac{n-4}{2}}}{\sqrt{2\pi}\;\Gamma(n-\frac{1}{2})\,(1-\rho r)^{n-3/2}} \cdot {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; \frac{2n-1}{2}, \frac{\rho r+1}{2}\right)$$

$${}_2F_1(a,b;c;x) = \sum_{k=0}^{\infty} \frac{a^{\overline{k}}\, b^{\overline{k}}}{c^{\overline{k}}} \frac{x^k}{k!}$$

$$m^{\overline{k}} = m(m+1)(m+2)\cdots(m+k-1) \quad \text{if } k \geq 1$$

$$m^{\overline{0}} = 1$$