Correlation

1、 if $Y_k = aX_k + b$, $k = 1, \cdots, n$

then $\bar{\rho} = \begin{cases} 1 & a>0 \\ -1 & a<0 \end{cases}$

$k=2$, $\bar{\rho} \overset{?}{=} \dfrac{(x_1-x_2)(y_1-y_2)}{|x_1-x_2||y_1-y_2|}$

Gaussian case

$(X, Y) \sim N(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho) \rightarrow$ joint pdf exists if $|\rho| < 1$

$\Updownarrow$

$(U, V) \sim N(0, 0, 1, 1, \rho) \Leftrightarrow \mathbb{E}\, e^{i(t_1 U + t_2 V)} = \exp(-\frac{1}{2}(t_1^2 + 2\rho t_1 t_2 + t_2^2))$

$U = \dfrac{X-\mu_X}{\sigma_X}$ $V = \dfrac{Y-\mu_Y}{\sigma_Y}$

$X \sim N(0,1) \Rightarrow \mathbb{E}\, e^{itX} = e^{-\frac{t^2}{2}}$

$\Updownarrow \mathbb{E}\, e^{tX} = e^{\frac{t^2}{2}}$

<u>Thm</u> two r.v. X, Y are independent

$\Leftrightarrow \mathbb{E}\, e^{i(t_1 X + t_2 Y)} = \mathbb{E}\, e^{it_1 X} e^{it_2 Y}$ for all t_1, t_2

$\rho = 0 \Rightarrow \mathbb{E}\, e^{it_1 U + it_2 V} = e^{-\frac{t_1^2}{2}} e^{-\frac{t_2^2}{2}} = \mathbb{E}\, e^{it_1 U}\, \mathbb{E}\, e^{it_2 V}$

$\varphi(t_1, t_2) = e^{i(t_1 U + t_2 V)}$

$\mathbb{E}\, UV = \dfrac{\partial^2 \varphi}{\partial t_1 \partial t_2}\Big|_{t_1=0, t_2=0}$

$\varphi(t) = \mathbb{E}\, e^{itX}$

$\varphi'(t) = i\,\mathbb{E}\, X e^{itX}$

$\varphi'(0) = i\,\mathbb{E} X$

$\mathbb{E}(V|U) = ? \quad \rho U$

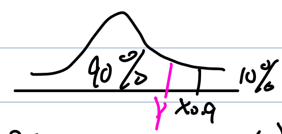
try $Z = V - aU \perp U$

$\mathbb{E} ZU = \mathbb{E} UV - a \mathbb{E} U^2 = \rho - a = 0 \Rightarrow \rho = a$

So $V - \rho U \perp V \Rightarrow \mathbb{E}(V - \rho U | U) = \mathbb{E} V - \mathbb{E} U = 0$

$\mathbb{E}(V|U) = \rho \mathbb{E}(U|U) = \rho U$

$\mathbb{E}(Y|X)$: $\dfrac{Y - \mu_Y}{\sigma_Y} = \rho \dfrac{X - \mu_X}{\sigma_X}$



know rank of X, Y, and ρ (rank 90% $x = x_{0.9}$: $P(X < x_{0.9}) = 90\%$)

$\dfrac{X - \mu_X}{\sigma_X} = z_{0.1} \Leftrightarrow x = x_{0.9} \Rightarrow \dfrac{Y - \mu_Y}{\sigma_Y} = \rho z_{0.1} < z_{0.1}$

$0 < \rho < 1$

ranking of Y is less than ranking of X
$\Leftrightarrow$ regression to the mean

Numbers

X - 90% $\Leftrightarrow z_{0.1} = 1.3$

$\rho = 0.6$

Y-value $= 0.6 \cdot 1.3 = 0.8$

$\Rightarrow$ Y-rank $= P(N(0,1) < 0.8) = 0.79$