90623-95123

Correlation & Regression (Advanced Level)

Algorithm is one of the main decider

PDF
PMF

Random variable $X \rightarrow$ mean μ, σ^2

$Y = \alpha + \beta X$ $\quad -\infty < \alpha < \infty$, $\beta > 0$

Select α, β $\quad Y$ has 0 mean, 1 variance

X is R.V. $\quad E(X) = \mu$, $V(X) = \sigma^2$

$E(Y) = \alpha + \beta E(X) = (\alpha + \beta\mu)$ $\quad E(Y) = \mu$

$V(Y) = E[Y - E(Y)]^2 = E[\beta(X - E(X))]^2$

$= \beta^2 \text{var}(X) = \beta^2\sigma^2$

We chose α, β

$E(Y) = 0$ $\quad V(Y) = 1$

$\beta^2\sigma^2 = 1$

$E(Y)=0$ $\quad V(Y)=\cdot$

$\alpha+\beta\mu=0$ $\quad \beta^2\sigma^2=1$

$\tilde{\beta}=\frac{1}{\sigma^2}$ $\quad \boxed{\beta=\frac{1}{\sigma}}$ $\quad \alpha=-\beta\mu=-\frac{\mu}{\sigma}$

NEXT ANALYSIS (with a change)..

- X, Y are correlated ??..

$Cov(X,Y)=E(X-E(X))(Y-E(Y))$

$=E[(X-E(X))\{\beta(X-E(X))\}]$

$=\beta E[X-E(X)]^2=\beta\cdot var(X)$

$=\beta\cdot\sigma^2$

Now, $V(Y)=\beta^2\sigma^2$

$\rho(X,Y)=\frac{\beta\sigma^2}{\sigma\cdot\beta\sigma}=1$

As, $Y=\alpha+\beta X$ → This is a perfect linear Relationship..

$\beta \quad \frac{dY}{dX}=\beta \rightarrow r$

??..

$\rho(X,Y+Z)=\frac{\sigma_Y}{\sigma_{Y+Z}}r(X,Y)+\frac{\sigma_Z}{\sigma_{Y+Z}}r(X,Z)$

$r(X,Y+Z)=\frac{Cov(X,Y+Z)}{\ }=\frac{Cov(X,Y)+Cov(X,Z)}{\ }$

$$r(X,Y+Z) = \frac{Cov(X,Y+Z)}{\sigma_X \sigma_{Y+Z}} = \frac{Cov(X,Y)+Cov(X,Z)}{\sigma_X \cdot \sigma_{Y+Z}}$$

$$= \frac{\sigma_Y}{\sigma_{Y+Z}} \cdot \frac{Cov(X,Y)}{\sigma_X \sigma_Y} + \frac{\sigma_Z}{\sigma_{Y+Z}} \frac{Cov(X,Z)}{\sigma_X \sigma_Z}$$

$$= \frac{\sigma_Y}{\sigma_{Y+Z}} r(X,Y) + \frac{\sigma_Z}{\sigma_{Y+Z}} r(X,Z)$$

$Y = a+bX+cX^2$ (Quadratic function)

find the correlation b/w X,Y

$Corr(X,Y)$

(I) X,Y uncorrelated

(II) X,Y perfectly correlated ..

$X \sim N(0,1)$

$E(X) = E(X^3) = 0$ $\quad E(X^2) = \mu_2 = \sigma_X^2 = 1$

$E(X^4) = \mu_4 = 3\sigma^4 = 3$

$Y = (a+bX+cX^2)$

$E(Y) = a+bE(X)+cE(X^2)$

$a+0+c \cdot 1$

$E(XY) = E[ax+bx^2+cx^3]$

$$E(XY) = E[ax + bx^2 + cx^3]$$
$$= aE(x) + bE(x^2) + cE(x^3)$$
$$= 0 + b \cdot 1 + c \cdot 0$$
$$= b$$

$$E(y^2) = E(a + bx + c x^2)^2$$
$$= E[a^2 + b^2 x^2 + c^2 x^4 \cancel{+ 2abx} + 2ac x^2 + 2bc x^3]$$
$$= a^2 + b^2 + 3c^2 + 2ac$$

$$v(y) = E(y^2) - [E(y)]^2 = a^2 + b^2 + 3c^2 + 2ac - (a+c)^2$$
$$= b^2 + 2c^2.$$

$$Cov(x, y) = E(xy) - E(x)E(y) = b$$

$$\rho(x,y) = \frac{b}{\sqrt{1 \cdot (b^2 + 2c^2)}}$$

Case 1

$$cov(x,y) = 0 \qquad b = 0$$

Case 11

$$\sqrt{b^2 + 2c^2} = b$$
$$\cancel{b^2} + 2c^2 = \cancel{b^2} \qquad c = 0$$

Regression

$Y = \cancel{\alpha + \beta x + \mu}$

meaning fluctuation of y..

$Y = \alpha + \beta X + \epsilon$

$\sum y = \alpha + \beta \sum x + \sum \varepsilon = 0$

Why Sum of fluctuation = 0 But a sense of relief ??

+ + +

− − −

$y \neq \alpha + \beta x$

5PM → 5.10 → 5.20 → 5.30 → 5.40 → 6.40

7

Time

5 km

uber

Comedy

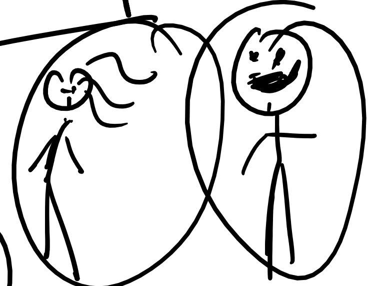
freelancer

Youth friendly

Coffee Shop → close

Date

5 8 7



y

Y X

$\sum \varepsilon = 0$

dependent ← $Y = f(X)$ ← independent

leader ... $= f(Y)$

deendor $\quad x = f(y)$

Y on X and X on Y

$\tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma_x \sigma_y}{\sigma_x^2+\sigma_y^2}\right)$

Slope $\quad y - \bar{y} = b_{yx}(x-\bar{x})$

$$m_1 = b_{yx} = r\cdot\frac{\sigma_y}{\sigma_x}$$

$$m_2 = \frac{\sigma_y}{r\cdot\sigma_x}$$

$$\tan\theta = \frac{m_1 \sim m_2}{1+m_1 m_2} = \frac{r\cdot\frac{\sigma_y}{\sigma_x} - \frac{\sigma_y}{r\cdot\sigma_x}}{1+\frac{\sigma_y^2}{\sigma_x^2}}$$

$$= \frac{\sigma_x \sigma_y}{\sigma_x^2+\sigma_y^2}\left(\frac{1-r^2}{r}\right)$$

2 Regression Variables (Common Mean)

$$Y = aX + b \qquad X = \alpha Y + \beta$$

Show that, $\frac{b}{\beta} = \frac{1-a}{1-\alpha}$

Ans: $\bar{x} = \bar{y} = \mu$

$\bar{y} = a\bar{x} + b$

$\mu = a\mu + b$

$(1-a)\mu = b \quad --(1)$

$(1-\alpha)\mu = \beta$

$(1-\alpha)\mu = \beta$

Simly

$$\mu = \frac{b}{1-a} = \frac{\beta}{1-\alpha} \Rightarrow \frac{b}{\beta} = \frac{1-a}{1-\alpha}$$

Prove that

If $a_1x + b_1y + c_1 = 0$ $\quad a_1b_2 \le a_2b_1$

$a_2x + b_2y + c_2 = 0$

y on x $\quad y = -\frac{a_1}{b_1}x - \frac{c_1}{b_1}$ $\quad b_{yx} = -\frac{a_1}{b_1}$

$b_{xy} = -\frac{b_2}{a_2}$

$r^2 = b_{xy}b_{yx} = +\left(\frac{a_1b_2}{b_1a_2}\right)$

$0 \le r^2 \le 1$ $\quad \frac{a_1b_2}{b_1a_2} \le 1$ $\quad \boxed{a_1b_2 \le b_1a_2}$

$b_1a_2 \ge a_1b_2$

$f(x,y) = y\exp[-y(1+x)]$

$= ye^{-y(1+x)} \longrightarrow$ Joint PDF

find Regression equation.

Ans: marginal PDF of X, Y

$$f_X(x) = \int^{\infty} f(x,y)\,dy = \int_0^{\infty} ye^{-(1+x)y}\,dy$$

but

$f_X(x) = \int_0^\infty f(x,y)\,dy = \int_0 y e$

$\Gamma(2) = 1!$

$\Gamma(3) = (3-1)! = 2!$

$= \frac{\Gamma(2)}{(1+x)^2} = \frac{1}{(1+x)^2}$

Sat 7PM

90623
95123

$f_Y(y) = \int_0^\infty f(x,y)\,dx$

$= y e^{-y} \int_0^\infty e^{-yx} dx$

$= y e^{-y} \left| \frac{e^{-yx}}{-y} \right|_{0=x}^{\infty}$

$= e^{-y}$

for conditional distribution

$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{y e^{-y(1+x)}}{\frac{1}{(1+x)^2}}$

$= y(1+x)^2 e^{-y(1+x)}$

$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{y e^{-y(1+x)}}{e^{-y}}$

$= y e^{-yx}$

$E(Y|X=x) = \int_0^\infty y f(y|x)\,dy$

$= (1+x)^2 \int_0^\infty y^2 e^{-(1+x)y} dy$

$= (1+x)^2 \frac{\Gamma(3)}{(1+x)^3} = \frac{2}{1+x}$

Regr of y on x: $y = E(Y|X=x) = \frac{2}{1+x}$

$E(X|Y=y) = \int_0^\infty x f(x|y)\,dx = y \int_0^\infty x e^{-yx} dx$

$E(X|Y=y) = \int_0 x f(x|y)\,dx = y \int_0 x e \, dx$

$= y \frac{\Gamma(2)}{y^2} = \frac{1}{y}$

$\boxed{X \text{ on } Y} \rightarrow x = E(X|Y=y) = \frac{1}{y}$

$\boxed{xy = 1}$

96623
95723

$f(x,y) = \frac{4}{5}(x+3y)\,e^{-x-2y}$

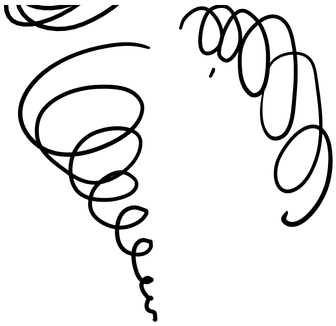
Y on X $\rightarrow$ $y = \frac{x+3}{2x+3}$

marg. PDF

chatGPT

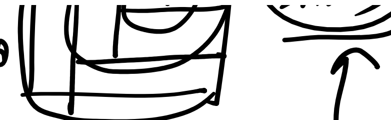
BARD

$\int (f(x,y))\,dy$



$f_X = \int_0^\infty f(x,y)\,dy$

$= \frac{y}{7} e^{-x} \left[x \left| \frac{e^{-xy}}{-2} \right|_0^\infty + \frac{1}{2} \int_0^\infty y\, e^{-xy}\, dy \right]$



dynamic spring

Beta

~~$B(m,n)$~~

$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1}\, dx$

$\Gamma(0)$ ~~$\Gamma(1)=$~~ $0! = 1$ $\quad -5! = ??$

$1! = 1$ $\quad \Gamma(-3)\ \times$

$\Gamma(n+1) = n\Gamma(n)$

$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$

$\int_0^\infty e^{-x^2} \cdot x^9\, dx$

$x^2 = t$

$2x\,dx = dt$

$dx = \frac{1}{2} x^{-1}\, dt$

$\int_0^\infty e^{-x^2} \cdot x^9\, dx = \int_0^\infty e^{-t}\, t^{9/2}\, \frac{1}{2}\, t^{-1/2}\, dt$

$= \frac{1}{2} \int_0^\infty e^{-t}\, t^4\, dt$

$\Gamma \int_0^\infty e^{-t}\, t^{5-1}\, dt$

$$= \frac{1}{2} \int_0^{\infty} e^{-t} t^{5-1}\, dt$$

$$= \frac{1}{2} \Gamma(5) = \frac{1}{2}(4!) = \frac{24}{2} = 12$$