Untitled Page.

Saturday, May 22, 2021 9:08 AM

$y \rightarrow$ ice creams sold

$x \rightarrow$ daily temp

$y \rightleftarrows x$

Regression

$y = \underset{\downarrow \text{ intercept}}{\alpha} + \underset{\downarrow \text{ slope}}{\beta} x \rightarrow$ expl. variable

Assumptions

- → (x, y) need to be bivariate normal 1
- → errors are normally distributed $e_i \sim N(0, \sigma^2)$ 2
- → ~~No multicollinearity~~ 3
- → No heteroscedasticity
- → x & y to be linear

$Y_i = \alpha + \beta x_i + e_i$

$y = 2 + 3x$

$x = 2, \quad \hat{y} = 2 + 3(2) = 8$

$y = 8.5$

$y - \hat{y} = 8.5 - 8 = 0.5$

$e_i = 0.5$ constant variance

$l_2 = 3$

$\vdots$

$l_n = 2$

Pearson coeff

$$r = \frac{Cov(x, y)}{S_x \cdot S_y}$$

$$= \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum_i (x_i-\bar{x})^2} \cdot \sqrt{\sum_i (y-\bar{y})^2}}$$

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

$S_{xx} \to \sum(x-\bar{x})^2 = \sum x^2 - n\bar{x}^2$

$S_{yy} \to \sum(y-\bar{y})^2 = \sum y^2 - n\bar{y}^2$

$S_{xy} \to \sum(x-\bar{x})(y-\bar{y}) = \sum xy - n\bar{x}\bar{y}$

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

$\hat{\beta} = \frac{S_{xy}}{S_{xx}}$

Least squares estimation

$y_i = \alpha + \beta x_i + e_i$

$e_i = y_i - \alpha - \beta x_i$

Square both sides,

$e_i^2 = (y_i - \alpha - \beta x_i)^2$

$\sum e_i^2 = \sum(y_i - \alpha - \beta x_i)^2$ — eqn

Diff w.r.t α

$$\frac{d e_i^2}{d\alpha} = \sum 2(y_i - \alpha - \beta x_i)(-1)$$

set to zero —

$\sum \cancel{2}(y_i - \alpha - \beta x_i)\cancel{(-1)} = 0$

$\sum y_i - n\alpha - \beta \sum x_i = 0$ — (1)

Diff w.r.t β —

$$\frac{d\sum e_i^2}{d\beta} = \sum 2(y_i - \alpha - \beta x_i)(-x_i)$$

setting to zero —

$-\sum x_i y_i + \alpha \sum x_i + \beta \sum x_i^2 = 0$ — (2)

$$-\sum x_i y_i + \alpha \sum x_i + \beta \sum x_i^2 = 0 \text{ --- } (2)$$

(1) —

$$\sum y_i - n\alpha - \beta \sum x_i = 0$$

$$\frac{\sum y_i - \beta \sum x_i}{n} = \alpha$$

$$\alpha = \frac{\sum y_i}{n} - \beta \cdot \frac{\sum x_i}{n}$$

$$\alpha = \bar{y} - \beta \bar{x} \text{ --- } (3)$$

(2) —

$$-\sum x_i y_i + \alpha \sum x_i + \beta \sum x_i^2 = 0$$

$$-\sum x_i y_i + (\bar{y} - \beta \bar{x}) \cdot \sum x_i + \beta \sum x_i^2 = 0$$

$$-\sum x_i y_i + (\bar{y} - \beta \bar{x})(n\bar{x}) + \beta \sum x_i^2 = 0$$

$$-\sum x_i y_i + n\bar{x}\bar{y} - \beta n \bar{x}^2 + \beta \sum x_i^2 = 0$$

$$\beta(\sum x_i^2 - n\bar{x}^2) = \sum x_i y_i - n\bar{x}\bar{y}$$

$$\beta = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{S_{xy}}{S_{xx}}$$

$$\beta = \frac{S_{xy}}{S_{xx}}$$

$$S_{xy} = \sum xy - n\bar{x}\bar{y}$$

$$S_{xx} = \sum x^2 - n\bar{x}^2$$

$$\sum xy = 65.607$$

$$S_{xx} = \sum x^2 - n\bar{x}^2$$

$\div$ by $(n-1)$ —

$$\frac{S_{xx}}{n-1} = \frac{\sum x^2 - n\bar{x}^2}{n-1} \leftarrow \text{Sample variance}$$

$$\frac{S_{xx}}{n-1} = s_x^2 \qquad \therefore S_{xx} = s_x^2(n-1)$$