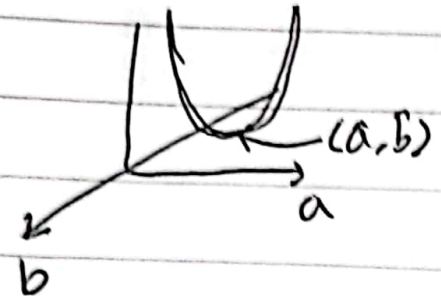
회귀분석 직선 손으로 구하기



$\hat{a} = \bar{y} - \hat{b}\bar{x}$

$\hat{b} = r\frac{S_y}{S_x}$

$RSS(a,b)$

미분에서 $y = x^2$ $y' = 2x$

$x=0$ 일때 $y=0$

미분 → $\frac{\partial RSS(a,b)}{\partial a} = 0$

$\frac{\partial RSS(a,b)}{\partial b} = 0$

} solution a, b

$$RSS(a,b) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^{n}(y_i - (a+bx_i))^2 = (y_1-a-bx_1)^2 + (y_2-a-bx_2)^2 + \cdots (y_n-a-bx_n)^2$$

① $\frac{\partial RSS(a,b)}{\partial a} =$

$2(y_1-a-bx_1)(-1) +$

$2(y_2-a-bx_2)(-1) + \cdots + 2(y_n-a-bx_n)$

$= -2\sum_{i=1}^{n}(y_i-a-bx_i) = 0$

② $\frac{\partial RSS(a,b)}{\partial b} = 2(y_1-a-bx_1)(-x_1) +$

$2(y_2-a-bx_2)(-x_2) +$

$\cdots + 2(y_n-a-bx_n)(-x_n)$

$= -2\sum_{i=1}^{n}(y_i-a-bx_i)x_i = 0$

① $\sum_{i=1}^{n}(y_i - a - bx_i) = 0 \Rightarrow \sum y_i - na - b\sum x_i = 0$

② $\sum_{i=1}^{n}(y_i - a - bx_i)(x_i) = 0 \Rightarrow \sum x_i y_i - a\sum x_i - b\sum x_i^2 = 0$

n으로 나눔

① $\frac{\sum y_i}{n} = \bar{y}$ ② $\overline{xy} - a\bar{x} - b\overline{x^2} = 0$

$\bar{y} - a - b\bar{x} = 0$

정리

① $a = \bar{y} - b\bar{x}$ ② $b\overline{x^2} = \overline{xy} - a\bar{x}$

$= \overline{xy} - \bar{x}\bar{y} + b(\bar{x})^2$

$b\overline{x^2} - b(\bar{x})^2 = \overline{xy} - \bar{x}\bar{y}$

$b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}$

$\hat{b} = r \times \frac{S_y}{S_x}$ $b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}$

$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{S_x}\right)\left(\frac{y_i - \bar{y}}{S_y}\right)$ 상관계수

$= \frac{1}{(n-1)S_x S_y}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

$\hat{b} = \frac{\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{S_x S_y} \times \frac{S_y}{S_x} = \frac{\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(S_x)^2}$

$S_x = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2}$ 표준편차

$(S_x)^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ 분산 $= \frac{1}{n-1}\sum(x_i^2 - 2x_i\bar{x} + (\bar{x})^2)$

$= \frac{1}{n-1}\left[\sum x_i^2 - 2\sum x_i\bar{x} + \sum(\bar{x})^2\right]$ $\sum(\bar{x})^2 = n(\bar{x})^2$

$$(S_x)^2 = \frac{1}{n-1}\left[\sum x_i^2 - 2\sum x_i\bar{x} + \sum \frac{(\bar{x})^2}{n}\right]$$

$$= \frac{1}{n-1}\cdot\frac{n}{n}\left[\sum x_i^2 - 2\sum x_i\bar{x} + \sum \frac{(\bar{x})^2}{n}\right]$$

$$= \frac{n}{n-1}\left(\overline{x^2} - 2(\bar{x})^2 + (\bar{x})^2\right)$$

$$= \frac{n}{n-1}\left(\overline{x^2} - (\bar{x})^2\right)$$

$$\hat{b} = \frac{\frac{1}{n-1}\sum (x_i-\bar{x})(y_i-\bar{y})}{(S_x)^2} \qquad b = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}$$

$$\hat{b} = \frac{\frac{1}{n-1}\sum (x_i-\bar{x})(y_i-\bar{y})}{\frac{n}{n-1}\left(\overline{x^2} - (\bar{x})^2\right)}$$

$$= \frac{\frac{1}{n}\sum (x_i-\bar{x})(y_i-\bar{y})}{\overline{x^2} - (\bar{x})^2} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2}$$

$$\frac{1}{n}\sum (x_iy_i - x_i\bar{y} - y_i\bar{x} + \bar{x}\bar{y})$$

$$= \frac{1}{n}\left[\sum x_iy_i - \bar{y}\sum x_i - \bar{x}\sum y_i + n\bar{x}\bar{y}\right]$$

$$= \overline{xy} - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y}$$

$$= \overline{xy} - \bar{x}\bar{y}$$