Geoffrey Guindine

Homework 1

9/8/24

(1)

a- A good target variable is the student's GPA.

b- It is continuous

c- A student's average standardized test scores like SAT or ACT

d- Yes, I expect the data to be in a positive linear relationship.

(2)

a- $\bar{x} = (0+1+2+3+4)/5 = \boxed{2}$

$\bar{y} = (0+2+3+8+17)/5 = \boxed{6}$

b- $S_x^2 = \sum_{i=1}^{5} (x_i - \bar{x})^2 = (-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2 / 4$

$\boxed{= 2.5}$

$S_y^2 = \sum_{i=1}^{5} (y_i - \bar{y})^2 = (-6)^2 + (-4)^2 + (-3)^2 + (2)^2 + (11)^2 / 4$

$\boxed{= 46.5}$

$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = (-2)(-6) + (-1)(-4) + 0 + (1)(2) + (2)(11)$

$/4$

$\boxed{= 10}$

c- $B_1 = \frac{S_{xy}}{S_x^2} = \frac{10}{2.5} = \boxed{4}$

$B_0 = 6 - 4(2) = \boxed{-2}$

$\boxed{Y = -2 + 4x + \varepsilon}$

d- $Y = -2 + 4(2.5) = \boxed{8}$

(3)

$$y = B_0 + B_1 x$$

a- $z(t) = z_0 e^{-\alpha t}$

$$\ln(z(t)) = \ln(z_0 e^{-\alpha t})$$

$$= \ln(z(t)) = \ln(z_0) + \ln(e^{-\alpha t})$$

$$\boxed{= \ln(z(t)) = \ln(z_0) - \alpha t}$$

b- $B_0 = \ln(z_0)$, $\quad$ $\boxed{z_0 = e^{B_0}}$

$B_1 = -\alpha$, $\quad$ $\boxed{\alpha = -B_1}$

c-
```
t = np.linspace(1, 100, 100)
z = np.linspace(1, 200, 4)
t_mean = np.mean(t)
z_mean = np.mean(z)
covar_xy = (np.sum(t - t_mean) * (np.sum(z - z_mean))
# calculates the covariance
var_x = np.sum(t - t_mean) ** 2
b_0 = np.log(covar_xy / var_x)   # log the B0
b_1 = -(z_mean - b_0 * t_mean)   # invert the B1
samples = []
for i in range(len(t)):
    samples.append(np.log(z_mean - t_mean * i))
```

(4) $Y = \beta_1 x$

a- $RSS(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta x_i)^2$

b- $\frac{d}{d\beta} RSS = \frac{d}{d\beta} \sum_{i=1}^{n} (y_i - \beta x_i)^2$

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

$$\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} \beta x_i x_i = 0$$

$$\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} \beta x_i^2$$

$$\boxed{\beta = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}}$$