PSTAT 126 Exam I 4/28/2022

(a) $\hat{\beta}_1 = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2} = \frac{654}{1400} = 0.467$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 19 - (0.467)(24) = 7.792$

$\Rightarrow \hat{y} = 7.792 + 0.467x$

(b) $S_{tot} = S_{yy} = \Sigma(y_i-\bar{y})^2 = 454$

$RSS = 148.489 \Rightarrow SS_{reg} = SS_{Tot} - RSS = 454 - 148.489$
$= 305.511$

Source of variation	SS	df	MS
Regression	305.511	1	305.511
Error	148.489	12	12.374
Total	454	13	

~~(c) $H_0: \beta_1 = 0$ vs $H_1: \beta_1 > 0$ $\alpha = 0.05$~~

(c) $R^2 = \frac{SS_{reg}}{S_{yy}} = \frac{305.511}{454} = 0.673 = 67.3\%$

$\Longrightarrow$ 67.3% of the y-variability was explained by the fitted linear regression.

(d) $r = +\sqrt{0.673} = 0.82$

(e) $H_0: \beta_1 = 0$ vs $H_1: \beta_1 > 0$ $\alpha = 0.05$

Reject H_0 if $t > t_{n-2, 0.05} = t_{12, 0.05} = 1.782$

$$t = \frac{\hat{\beta}_1}{\hat{\sigma}/\sqrt{\sum(x-\bar{x})^2}} = \frac{0.467}{\sqrt{12.374}/\sqrt{1400}} = 4.967 > 1.782$$

$\Longrightarrow$ Reject H_0

(f) $x_0 = 30 \Longrightarrow \hat{y}_0 = 7.792 + 0.467(30) = 21.802$

$100(1-\alpha)\%$ CI for $E(Y \mid x = 30) = \beta_0 + \beta_1(30)$:

$$\hat{y}_0 \mp t_{n-2,\alpha/2}\,\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}}$$

$$\Longrightarrow 21.802 \mp 2.719\sqrt{12.374}\sqrt{\frac{1}{14} + \frac{(30-24)^2}{1400}}$$

$$\Longrightarrow 21.802 \mp 2.389 \Longrightarrow (19.413,\ 24.191)$$

(g) $Y_0 = \beta_0 + \beta_1 x_0 + e_0$ $\qquad \hat{Y}_0 = 21.802$

$100(1-\alpha)\%$ PI for $Y_0 = \beta_0 + \beta_1 x_0 + e_0$:

$$\hat{y}_0 \mp t_{n-2,\alpha/2}\,\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}}$$

$$\Longrightarrow 21.802 \mp 2.719\sqrt{12.374}\sqrt{1 + \frac{1}{14} + \frac{(30-26)^2}{1400}}$$

$$\Longrightarrow 21.802 \mp 8.029 \Longrightarrow (13.773,\ 29.841)$$

(h) $\bar{Y}_0^{(5)} = \frac{1}{5}\left(Y_0^{(1)} + Y_0^{(2)} + Y_0^{(3)} + Y_0^{(4)} + Y_0^{(5)}\right)$ $\qquad \hat{\bar{Y}}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0 = 21.802$

$100(1-\alpha)\%$ PI for $\bar{Y}_0^{(5)}$:

$$\hat{y}_0 \mp t_{n-2,\alpha/2}\,\hat{\sigma}\sqrt{\frac{1}{5} + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}}$$

$$\Rightarrow 21.802 \mp 2.179 \sqrt{12.394} \sqrt{\frac{1}{5} + \frac{1}{14} + \frac{(30-24)^2}{1400}}$$

$$\Rightarrow 21.802 \mp 4.178 \Rightarrow (17.624, 25.980)$$

(j) $100(1-\alpha)$ PI for $\sum_{i=1}^{5} y_0^{(i)} = m\bar{y}_0^{(m)} = 5\bar{y}_0^{(5)}$

$$\hat{y}_0 - \bar{y}_0^{(m)} \sim N\left(0, \left(\frac{1}{m} + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}\right)\sigma^2\right)$$

$$\Rightarrow m\left(\hat{y}_0 - \bar{y}_0^{(m)}\right) \sim N\left(0, m^2\left(\frac{1}{m} + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}\right)\sigma^2\right)$$

$$\Rightarrow \frac{m\hat{y}_0 - m\bar{y}_0^{(m)} - 0}{m\hat{\sigma}\sqrt{\frac{1}{m} + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}}} \sim t_{n-2}$$

$\Rightarrow$ $100(1-\alpha)\%$ PI for $\sum_{i=1}^{5} y_0^{(i)} = m\bar{y}_0^{(m)}$ is

$$m\hat{y}_0 \mp t_{n-2,\alpha/2}\, \hat{\sigma}\, m \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x-\bar{x})^2}}$$

$$\Rightarrow 5(21.802) \mp (2.179)\sqrt{12.394}\,(5)\sqrt{\frac{1}{5} + \frac{1}{14} + \frac{(30-24)^2}{1400}}$$

$$\Rightarrow 109.01 \mp 20.891$$

$$\Rightarrow (88.119, 129.901)$$

(j) $Y_0^{(20)} = \beta_0 + \beta_1(20) + e_0^{(20)}$

$Y_0^{(30)} = \beta_0 + \beta_1(30) + e_0^{(30)}$

$e_0^{(20)}, e_0^{(30)} \overset{iid}{\sim} N(0, \sigma^2)$

$$\Rightarrow \left\{\begin{array}{l} Y_0^{(20)} \sim N(\beta_0 + 20\beta_1, \sigma^2) \\ Y_0^{(30)} \sim N(\beta_0 + 30\beta_1, \sigma^2) \end{array}\right\} \text{indep}$$

$$\Rightarrow Y_0^{(20)} + Y_0^{(30)} \sim N(2\beta_0 + 50\beta_1, 2\sigma^2)$$

$\hat{Y}_0^{(20)} = \hat{\beta}_0 + \hat{\beta}_1(20)$

$\hat{Y}_0^{(30)} = \hat{\beta}_0 + \hat{\beta}_1(30)$

$\hat{Y}_0^{(20)} + \hat{Y}_0^{(30)} = \hat{\beta}_0 + 20\hat{\beta}_1 + \hat{\beta}_0 + 30\hat{\beta}_1 = 2\hat{\beta}_0 + 50\hat{\beta}_1 \sim$ Normal

$$E\left(\hat{Y}_0^{(20)} + \hat{Y}_0^{(30)}\right) = E(2\hat{\beta}_0 + 50\hat{\beta}_1) = 2\beta_0 + 50\beta_1 = E\left(Y_0^{(20)} + Y_0^{(30)}\right)$$

$$Var\left(\hat{Y}_0^{(20)} + \hat{Y}_0^{(30)}\right) = Var(2\hat{\beta}_0 + 50\hat{\beta}_1) = Var\left[2(\bar{Y} - \hat{\beta}_1\bar{x}) + 50\hat{\beta}_1\right]$$

$$= Var\left(2\bar{Y} + (50 - 2\bar{x})\hat{\beta}_1\right) = 4\,Var(\bar{Y}) + (50 - \bar{x})^2 Var(\hat{\beta}_1)$$

$$= \frac{4}{n}\sigma^2 + (50 - 2\bar{x})^2 \frac{\sigma^2}{\sum(x - \bar{x})^2}$$

$$= \sigma^2\left(\frac{4}{n} + \frac{(50 - 2\bar{x})^2}{\sum(x - \bar{x})^2}\right)$$

Consider $\left(\hat{Y}_0^{(20)} + \hat{Y}_0^{(30)}\right) - \left(Y_0^{(20)} + Y_0^{(30)}\right) \sim$ Normal

$$E\left(\left(\hat{Y}_0^{(20)} + \hat{Y}_0^{(30)}\right) - \left(Y_0^{(20)} + Y_0^{(30)}\right)\right) = (2\beta_0 + 50\beta_1) - (2\beta_0 + 50\beta_1) = 0$$

$$Var\left[\left(\hat{Y}_0^{(20)}+\hat{Y}_0^{(30)}\right)-\left(Y_0^{(20)}+Y_0^{(30)}\right)\right]$$

$$=Var\left(\hat{Y}_0^{(20)}+\hat{Y}_0^{(30)}\right)+Var\left(Y_0^{(20)}+Y_0^{(30)}\right)$$

$$=\sigma^2\left(\frac{4}{n}+\frac{(50-2\bar{x})^2}{\sum(x-\bar{x})^2}\right)+2\sigma^2$$

$$=\sigma^2\left(2+\frac{4}{n}+\frac{(50-2\bar{x})^2}{\sum(x-\bar{x})^2}\right)$$

$$\Rightarrow \left(\hat{Y}_0^{(20)}+\hat{Y}_0^{(30)}\right)-\left(Y_0^{(20)}+Y_0^{(30)}\right)\sim N\left(0,\sigma^2\left(2+\frac{4}{n}+\frac{(50-2\bar{x})^2}{\sum(x-\bar{x})^2}\right)\right)$$

$$\Rightarrow \frac{\left(\hat{Y}_0^{(20)}+\hat{Y}_0^{(30)}\right)-\left(Y_0^{(20)}+Y_0^{(30)}\right)}{\sigma\sqrt{2+\frac{4}{n}+\frac{(50-2\bar{x})^2}{\sum(x-\bar{x})^2}}}\sim N(0,1)$$

$$\Rightarrow \frac{\left(\hat{Y}_0^{(20)}+\hat{Y}_0^{(30)}\right)-\left(Y_0^{(20)}+Y_0^{(30)}\right)}{\hat{\sigma}\sqrt{2+\frac{4}{n}+\frac{(50-2\bar{x})^2}{\sum(x-\bar{x})^2}}}\sim t_{n-2}$$

$\Rightarrow 100(1-\alpha)\%$ CI for $Y_0^{(20)}+Y_0^{(30)}$:

$$\left(\hat{y}_0^{(20)}+\hat{y}_0^{(30)}\right)\mp t_{n-2,\alpha/2}\,\hat{\sigma}\sqrt{2+\frac{4}{n}+\frac{(50-2\bar{x})^2}{\sum(x-\bar{x})^2}}$$

$$\hat{y}_0^{(20)}=7.792+0.467(20)=17.132$$
$$\hat{y}_0^{(30)}=7.792+0.467(30)=21.802$$

$$\overset{PI}{\Rightarrow}(17.132+21.802)\mp 2.179\sqrt{12.374}\sqrt{2+\frac{4}{14}+\frac{(50-48)^2}{1400}}$$

$$\Rightarrow 38.934\mp 11.596\Rightarrow(27.338,\ 50.530)$$