Jack Rodoni
STAT 608
Assignment 1
2/1/22

1.)

(a) 95% CI for β_1: (0.9514971, 1.0126660)

- Yes, 1 is a plausible value for β_1 as 1 is contained in our 95% CI.

(b) $H_0: \beta_0 = 10{,}000$, $H_a: \beta_0 \neq 10000$

$$T = \frac{\hat{\beta}_0 - 10000}{S_{\beta_0}} = \frac{6804.886 - 10000}{9929.018} = -0.3217858$$

$$P(t \geq |T|) = 2^*(1 - pt(abs(-0.3217858), 16)) = 0.7517807$$

- We would fail to reject $H_0: \beta_0 = 10000$. We would fail to reject the hypothesis that if last weeks sales were \$0, on average this weeks sales will be \$10,000.

(c) Our fitted regression eq. is:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad w/\ \hat{\beta}_0 = 6804.886,\ \hat{\beta}_1 = 0.9820815$$

- For $x = \$400{,}000$, we get a fitted value

$$E[y \mid x = 400000] = \$399{,}637.50 = \hat{y}$$

- 95% Prediction Interval: (\$359832.8, \$439,442.20)

No, \$450,000 doesn't seem like a feasible value for the gross box office results in the current week, for a production w/ \$400,000 gross box office the previous week, b/c \$450,000 is not contained w/in our 95% prediction interval.

(d) That rule seems fairly reasonable seems reasonable for 2 reasons:

① Our 95% CI on B_1 tells us that we are 95% confident that a \$1 increase in sales last week corresponds, on average, to an increase in this week sales by (\$0.95, \$1.01). Note \$1 is contained in this interval.

② Our 95PI contains last weeks gross box office sales (see part c)

2.) $Var(y_i|x_i) = var(\beta_0 + \beta_1 x_i + \varepsilon_i | x_i)$

$$= \cancelto{0}{var(\beta_0|x_i)} + \cancelto{0}{var(\beta_1 x_i|x_i)} + var(\varepsilon_i|x_i)$$
$$+ \cancelto{0}{cov(\beta_0, \beta_1 x_i|x_i)} + \cancelto{0}{cov(\beta_0, \varepsilon_i|x_i)} + \cancelto{0}{cov(\beta_1 x_i, \varepsilon_i|x_i)}$$
$$= var(\varepsilon_i|x_i)$$

$var(y_i|x_i) = var(\varepsilon_i)$ * b/c the epsilons are independent of x_i

3.) The least squares criterion tells us that when we are fitting a regression model, we choose the model that minimizes the sum of the squared residuals.

4.) (a) (i) $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

(ii) WTS $\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$ [* see chp 2 notes (slides 24-26)*]

• $RSS(\beta) = \sum e_i^2 = e'e = (y - X\beta)'(y - X\beta)$
$$= y'y - y'X\beta - \beta'X'y + \beta'X'X\beta'$$
$$= y'y - 2y'X\beta' + \beta'X'X\beta$$

• $\frac{\partial RSS(\beta)}{\partial \beta} = -2(y'X)' + (X'X + (X'X)')\beta = -2(y'X)' + (X'X + X'X)\beta := 0$

$X'X\hat{\beta} = X'y \iff \boxed{\hat{\beta} = \frac{X'y}{X'X}}$

[* recall chp 2 notes, slide 22 & 24
① $\frac{\partial}{\partial x}(a'x) = (a')' = a$
② $\frac{\partial}{\partial x} x'Ax = (A + A')x$]

• In our case: $X' = [x_1\ x_2 \cdots x_n]$, $y' = [y_1\ y_2 \cdots y_n]$

$$X'y = [x_1\ x_2 \cdots x_n]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum x_i y_i$$

$$X'X = [x_1\ x_2 \cdots x_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum x_i^2$$

$$\Rightarrow \boxed{\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}}$$

4.) (b) (i) WTS $E[\hat{\beta}|X] = \beta$

- From 4(a)(ii) we know: $\hat{\beta} = (X'X)^{-1}X'y$

$$\Rightarrow E[\hat{\beta}|X] = E[(X'X)^{-1}X'y \mid X]$$

$$= (X'X)^{-1}X' E[y|X]$$ * can pull the X's out of the E[ ] b/c they are given.

* $y = X\beta + \varepsilon$

$$= (X'X)^{-1}X' E[X\beta + \varepsilon \mid X]$$

$$= (X'X)^{-1}X' (E[X\beta|X] + E[\overset{0}{\cancel{\varepsilon}}|X])$$

$$= (X'X)^{-1}X'X\beta \qquad * A^{-1}A = I$$

$$= I\beta$$

$$\boxed{E[\hat{\beta}|X] = \beta}$$ QED

(ii) WTS $var(\hat{\beta}|X) = \sigma^2 / \sum x_i^2$

$$var(\hat{\beta}|X) = var((X'X)^{-1}X'y|X) = (X'X)^{-1}X' var(y|X) X (X'X)^{-1}$$

$$= (X'X)^{-1}X' var(X\beta + \varepsilon | X) X (X'X)^{-1}$$

$$= (X'X)^{-1}X' [\overset{0}{\cancel{var(X\beta|X)}} + \overset{\sigma^2}{\cancel{var(\varepsilon|X)}} + 2\overset{0}{\cancel{cov(X\beta, \varepsilon|X)}}] X (X'X)^{-1}$$

$$= (X'X)^{-1}X' \sigma^2 X (X'X)^{-1}$$

$$= \sigma^2 (X'X)^{-1} \cancel{X'X (X'X)^{-1}}$$ * can move σ^2 b/c it is just a scalar

$$= \frac{\sigma^2}{X'X}$$

$$\boxed{var(\hat{\beta}|X) = \frac{\sigma^2}{\sum x_i^2}}$$

(iii) show $(\hat{\beta}|X) \sim N(\beta, \frac{\sigma^2}{\sum x_i^2})$ [* see chp 2 notes slide 49]

$$(\hat{\beta}|X) = \left(\frac{\sum x_i y_i}{\sum x_i^2} \middle| X\right) = \sum \frac{x_i}{a} y_i$$ where $a = \sum x_i^2$ (constant b/c X is given)

let $c_i = x_i/a$

$$= \sum_{i=1}^{n} c(y_i|X)$$ * from chp 2 slide 49 we know $y_i|X \sim N$

thus, $\hat{\beta}|X$ is a linear combination of Normally distributed variables $\Rightarrow$

$$(\hat{\beta}|X) \sim N(\beta, \sigma^2/\sum x_i^2)$$

from (b)(i) & (ii)

5.) (a) $X' = [1 \; 1 \cdots 1]$

$$\underset{\sim}{y} = X\beta + \varepsilon \iff \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [\beta_0] + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\iff y_i = \beta_0 + \varepsilon_i \checkmark$$

(b) $\hat{\beta} = (X'X)^{-1} X' \underset{\sim}{y}$

$$\hat{\beta} = [1 \; 1 \cdots 1] \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [1 \; 1 \cdots 1] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \frac{1}{n} \sum y_i = \bar{y}$$

$$\boxed{\hat{\beta} = \bar{y}}$$

6.) Prove $\text{Cov}(aX, bY) = ab\,\text{cov}(X, Y)$:

$$\text{cov}(aX, bY) = E[aX\,bY] - E[aX]E[bY]$$
$$= abE[XY] - abE[X]E[Y]$$
$$= ab(E[XY] - E[X]E[Y])$$
$$\boxed{\text{cov}(aX, bY) = ab(\text{cov}(X, Y))}$$

7.) A CI is a confidence statement about the mean of a R.V. Thus, if we have a lot of data, it is entirely possible that we could be very confident about what the average response will be (thus have very narrow confidence bands) but there still exists large variation in the individual responses, resulting in us having 95% of our data fall outside our confidence interval.

8.) Show $\hat{\beta}_0 = \bar{y} - \frac{Sxy}{Sxx}\bar{x}$ for the simple linear regression case.

- Simple lin Reg => $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ =>

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

- $\hat{\underline{\beta}} = (X'X)^{-1} X'\underline{y}$

- $X'X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \sum x_i^2 \end{bmatrix}$

note: $\sum (x_i - \bar{x})^2 = \sum x_i^2 - 2x_i\bar{x} + \bar{x}^2 = \sum x_i^2 - 2\bar{x}\sum x_i + \sum \bar{x}^2$

$= \sum x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$

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

$\Rightarrow \sum x_i^2 = \sum (x_i - \bar{x})^2 + n\bar{x}^2$

$$X'X = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \sum x_i^2 \end{bmatrix} = n \begin{bmatrix} 1 & \bar{x} \\ \bar{x} & \frac{\sum (x_i - \bar{x})^2}{n} + \bar{x}^2 \end{bmatrix} = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & Sxx + n\bar{x}^2 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{n\,Sxx + n\bar{x}^2 - n\bar{x}^2} \begin{bmatrix} Sxx + n\bar{x}^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix} = \frac{1}{n\,Sxx} \begin{bmatrix} Sxx + n\bar{x}^2 & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix}$$

$$X'\underline{y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} = \begin{bmatrix} n\bar{y} \\ \sum y_i x_i - n\bar{y}\bar{x} + n\bar{y}\bar{x} \end{bmatrix}$$

$$= \begin{bmatrix} n\bar{y} \\ \sum (x_i - \bar{x})(y_i - \bar{y}) + n\bar{x}\bar{y} \end{bmatrix} = \begin{bmatrix} n\bar{y} \\ Sxy + n\bar{x}\bar{y} \end{bmatrix}$$

$$(X'X)^{-1} X'\underline{y} = \frac{1}{n\,Sxx} \begin{bmatrix} \frac{Sxx}{n} + \bar{x}^2 & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix} \begin{bmatrix} n\bar{y} \\ Sxy + n\bar{x}\bar{y} \end{bmatrix} = \frac{1}{Sxx} \begin{bmatrix} \left(\frac{Sxx}{n} + \bar{x}^2\right)(n\bar{y}) - \bar{x}(Sxy + n\bar{x}\bar{y}) \\ -n\bar{x}\bar{y} + Sxy + n\bar{x}\bar{y} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{y} - \bar{x}\frac{Sxy}{Sxx} \\ \frac{Sxy}{Sxx} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} \Rightarrow \hat{\beta}_0 = \bar{y} - \bar{x}\frac{Sxy}{Sxx}$$

a) $\text{var}(\underline{a}'\hat{\underline{\beta}} \mid X) = \sigma^2 \underline{a}'(X'X)^{-1}\underline{a}$ where $\underline{a}$ is a scalar vector

$\text{var}(\underline{a}'\hat{\underline{\beta}} \mid X) = \underline{a}' \text{var}(\hat{\beta} \mid X)\underline{a}$ (b/c $\underline{a}$ is a constant vector)

$\text{var}(\hat{\beta} \mid X) = \sigma^2 (X'X)^{-1}$ * proved in 4 (b) (ii)

$$\boxed{\text{var}(\underline{a}'\hat{\underline{\beta}} \mid X) = \sigma^2 \underline{a}'(X'X)^{-1}\underline{a}}$$