HW 3

Madeline Blasingame
205412882
10/17/2021

Exercise 1

a. we want to test $H_0: \beta_1 = 0$ against $H_a: \beta_1 \neq 0$ using an F statistic

We know: $\hat{\beta}_1 \sim N\left(0, \frac{\sigma}{\sqrt{\sum(x_i-\bar{x})^2}}\right)$ under H_0

Therefore:

$$\frac{\hat{\beta}_1 - 0}{\sigma/\sqrt{\sum(x_i-\bar{x})^2}} \sim N(0,1) \quad \text{and} \quad \frac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sigma^2} \sim \chi^2_1$$

we also know:

independent

$$\frac{(n-2)\,Se^2}{\sigma^2} \sim \chi^2_{n-2}$$

Therefore:

$$\frac{\dfrac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sigma^2} \Big/ 1}{\dfrac{(n-2)\,Se^2}{\sigma^2} \Big/ n-2} = \frac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{Se^2} \sim F_{1,n-2}$$

we want to show that the F statistic can also be expressed as $F = \frac{R^2}{1-R^2}(n-2)$ where $R^2 = \frac{SSR}{SST}$

because $\hat{Y}_i = \bar{Y} + \hat{\beta}_1(x_i - \bar{x})$

$$R^2 = \frac{SSR}{SST} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i-\bar{Y})^2} = \frac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sum(Y_i-\bar{Y})^2}$$

$$\frac{R^2}{1-R^2}(n-2) = \frac{\dfrac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sum(Y_i-\bar{Y})^2}}{1 - \dfrac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sum(Y_i-\bar{Y})^2}}(n-2) = \frac{\dfrac{\hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sum(Y_i-\bar{Y})^2}}{\dfrac{\sum(Y_i-\bar{Y})^2 - \hat{\beta}_1^2 \sum(x_i-\bar{x})^2}{\sum(Y_i-\bar{Y})^2}}(n-2)$$

$$= \frac{\left(\hat{\beta}_1^2 \sum (x_i - \bar{x})^2\right)\left(\sum (Y_i - \bar{Y})^2\right) \quad (n-2)}{\left(\sum (Y_i - \bar{Y})^2\right)\left(\sum (Y_i - \bar{Y})^2 - \hat{\beta}_1^2 \sum (x_i - \bar{x})^2\right)} = \frac{(n-2)\, \hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sum (Y_i - \bar{Y})^2 - \hat{\beta}_1^2 \sum (x_i - \bar{x})^2}$$

Note:

$$\sum (Y_i - \bar{Y})^2 = \sum e_i^2 + \sum (\hat{Y}_i - \bar{Y})^2$$

$$\Rightarrow \quad \sum e_i^2 = \sum (Y_i - \bar{Y})^2 - \underbrace{\sum (\hat{Y}_i - \bar{Y})^2}$$

Note: $\hat{Y}_i = \bar{Y} + \hat{\beta}_1 (x_i - \bar{x})^2$

$$\Rightarrow \sum (\hat{Y}_i - \bar{Y})^2 = \sum (\bar{Y} + \hat{\beta}_1 (x_i - \bar{x}) - \bar{Y})^2 = \sum (\hat{\beta}_1 (x_i - \bar{x}))^2 =$$
$$\sum \hat{\beta}_1^2 (x_i - \bar{x})^2$$

Therefore: $\sum e_i^2 = \sum (Y_i - \bar{Y})^2 - \hat{\beta}_1^2 \sum (x_i - \bar{x})^2$

$$= \frac{(n-2)\, \hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sum e_i^2}$$

Note: $S_e^2 = \dfrac{\sum e_i^2}{(n-2)}$

$$= \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sum e_i^2}$$

We already showed that this term follows an F distribution, therefore:

$$\frac{R^2}{1 - R^2} (n-2) = \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sum e_i^2} \sim F_{1, n-2}$$

b. $H_0: \beta_1 = 0$
$H_a: \beta_1 \neq 0$

$$F = \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2}$$

$$E[F] = E\left[\frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2}\right] = \sum (x_i - \bar{x})^2 \, E\left[\frac{\hat{\beta}_1^2}{S_e^2}\right]$$

we know: $\hat{\beta}_1 \perp\!\!\!\perp e_i \Rightarrow \hat{\beta}_1^2 \perp\!\!\!\perp S_e^2$

Therefore:

$$= \Sigma (x_i - \bar{x})^2 \ E[\hat{\beta}_1^2] \ E[(S_e^2)^{-1}]$$

Note: $E[\hat{\beta}_1^2] = var(\hat{\beta}_1) + E(\hat{\beta}_1)^2$

$$= \frac{\sigma^2}{\Sigma (x_i - \bar{x})^2} + \beta_1^2 = \frac{\sigma^2}{\Sigma (x_i - \bar{x})^2}$$ due to the assumption under H_0

Note:

$S_e^2 \sim \Gamma\left(\frac{n-2}{2}, \frac{2\sigma^2}{n-2}\right)$ and if $Q \sim \Gamma(\alpha, \beta) \Rightarrow$

$$E(Q^k) = \frac{\Gamma(\alpha + k)\, \beta^k}{\Gamma(\alpha)}$$

Therefore:

$$E[(S_e^2)^{-1}] = \frac{\Gamma\left(\frac{n-2}{2} - 1\right)\left(\frac{2\sigma^2}{n-2}\right)^{-1}}{\Gamma\left(\frac{n-2}{2}\right)} = \frac{\Gamma\left(\frac{n-2}{2} - 1\right)}{\Gamma\left(\frac{n-2}{2}\right)} \cdot \frac{n-2}{2\sigma^2}$$

Note: $\dfrac{\Gamma(\alpha - 1)}{\Gamma(\alpha)} = \dfrac{1}{\alpha - 1}$

Therefore:

$$= \frac{1}{\frac{n-2}{2} - 1} \cdot \frac{n-2}{2\sigma^2} = \frac{1}{\frac{n-2-2}{2}} \cdot \frac{n-2}{2\sigma^2} = \frac{2}{n-4} \cdot \frac{n-2}{2\sigma^2}$$

$$= \frac{n-2}{(n-4)\sigma^2}$$

$$E[\hat{\beta}_1^2] = \Sigma (x_i - \bar{x})^2 \cdot \frac{\sigma^2}{\Sigma (x_i - \bar{x})^2} \cdot \frac{n-2}{(n-4)\sigma^2} = \frac{n-2}{(n-4)}$$

Exercise 2

$y_i = \beta_1 x_i + \varepsilon_i$

$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$ Derive the likelihood ratio test statistic

$$\lambda = \frac{\text{maximum likelihood when } H_0 \text{ is true } (\beta_1 = 0)}{\text{maximum likelihood when you estimate } \beta_1} < K$$

$$= \frac{(2\pi \hat{\sigma}_0^2)^{-n/2} e^{\frac{-\Sigma y_i^2}{2\hat{\sigma}_0^2}}}{(2\pi \hat{\sigma}^2)^{-n/2} e^{\frac{-\Sigma(y_i - \hat{\beta}_1 x_i)^2}{2\hat{\sigma}^2}}} < K$$

MLE for $\hat{\sigma}_0$:

$$L = (2\pi \hat{\sigma}_0^2)^{-n/2} e^{-\frac{\Sigma y_i^2}{2\hat{\sigma}_0^2}}$$

$$l = \ln(L) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\hat{\sigma}_0^2) - \frac{\Sigma y_i^2}{2\hat{\sigma}_0^2}$$

$$\frac{\partial l}{\partial \hat{\sigma}_0^2} = \frac{-n}{2\hat{\sigma}_0^2} + \frac{\Sigma y_i^2}{2(\hat{\sigma}_0^2)^2} = 0$$

$$\frac{\Sigma y_i^2}{2(\hat{\sigma}_0^2)^2} = \frac{n}{2(\hat{\sigma}_0^2)} \rightarrow \frac{\Sigma y_i^2}{n} = \hat{\sigma}_0^2$$

MLE for $\hat{\sigma}^2$: we know that when we are estimating β_1 that $\hat{\sigma}^2 = \frac{1}{n}\Sigma(y_i - \hat{\beta}_1 x_i)^2$ due to HW1

$$\Rightarrow \; = \frac{(2\pi \hat{\sigma}_0^2)^{-n/2} e^{-n/2}}{(2\pi \hat{\sigma}^2)^{-n/2} e^{-n/2}} < K \Rightarrow \left(\frac{(2\pi \hat{\sigma}_0^2)^{-n/2}}{(2\pi \hat{\sigma}^2)^{-n/2}}\right)^{-2/n} > K^{-2/n}$$

$$\Rightarrow \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} > K^{-2/n}$$

$$\frac{\frac{\sum y_i^2}{n}}{\frac{\sum (y_i - \hat{\beta}_1 x_i)^2}{n}} > K^{-2/n} \quad \Rightarrow \quad \frac{\sum y_i^2}{n} \cdot \frac{n}{\sum (y_i - \hat{\beta}_1 x_i)^2} > K^{-2/n}$$

$$\Rightarrow \quad \frac{\sum y_i^2}{\sum (y_i - \hat{\beta}_1 x_i)^2} > K^{-2/n}$$

Note :

$$e_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_1 x_i$$

$$\sum y_i^2 = \sum \hat{y}_i + \sum e_i^2$$

$$\Rightarrow \quad \frac{\sum \hat{y}_i + \sum e_i^2}{\sum e_i^2} > K^{-2/n} \quad \Rightarrow \quad \frac{\sum \hat{y}_i}{\sum e_i^2} + 1 > K^{-2/n}$$

$$\Rightarrow \quad \frac{\sum \hat{y}_i}{\sum e_i^2} > \boxed{K^{-2/n} - 1} \rightarrow \text{Define } K' = K^{-2/n} - 1$$

Reject H_0 if :

$$\frac{\sum \hat{y}_i}{\sum e_i^2} > K'$$

Exercise 3

a. $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$var(\varepsilon_i) = \sigma^2$

$\varepsilon_i \perp\!\!\!\perp \varepsilon_j$ if $i \neq j$

$E[\varepsilon_i] = 0$

we need: $\hat{Y}_0 = \sum a_i y_i$

$E(\hat{Y}_0) = E[Y_0]$

We know:

$$E[\hat{Y}_0] = \beta_0 + \beta_1 X_0 = \begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix} \begin{bmatrix} 1 \\ x_0 \end{bmatrix}$$

$E[Y_i] = \beta_0 + \beta_1 x_i$

$$E[\hat{Y}_0] = \sum a_i E[Y_i]$$

$$= \beta_0 \sum a_i + \beta_1 \sum a_i x_i = \begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix} \begin{bmatrix} \sum a_i \\ \sum a_i x_i \end{bmatrix}$$

So $E(Y_0) = E(\hat{Y}_0) \Rightarrow \begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix} \begin{bmatrix} 1 \\ X_0 \end{bmatrix} = \begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix} \begin{bmatrix} \sum a_i \\ \sum a_i x_i \end{bmatrix}$

$$\Rightarrow \underbrace{\begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix}}_{V_1} \underbrace{\begin{bmatrix} 1 - \sum a_i \\ x_0 - \sum a_i x_i \end{bmatrix}}_{V_2} = 0$$

From linear algebra, it follows that either $V_1 = [0 \; 0]$ or $V_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow$ however $\beta_0 = 0$ and $\beta_1 = 0$ is not necessarily true, so $V_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Therefore:

$$\begin{bmatrix} 1 \\ X_0 \end{bmatrix} = \begin{bmatrix} \sum a_i \\ \sum a_i x_i \end{bmatrix}$$

b. We know that $E(a^2) = var(a) + [E(a)]^2$

using this rule we can show:

$$E(Y_0 - \hat{Y}_0)^2 = var(Y_0 - \hat{Y}_0) + [E(Y_0 - \hat{Y}_0)]^2$$

We also know: $E(Y_0 - \hat{Y}_0) = 0$ because $E(Y_0) = E(\hat{Y}_0) = \beta_0 + \beta_1 x_0$ when $\hat{Y}_0$ is unbiased.

Therefore:

$$E(Y_0 - \hat{Y}_0)^2 = var(Y_0 - \hat{Y}_0)$$

c. We need to minimize $var(Y_0 - \hat{Y}_0)$ with the constraints $\Sigma a_i = 1$ and $\Sigma a_i x_i = x_0$

$$\min Q = var(Y_0 - \hat{Y}_0) - 2\lambda_1[\Sigma a_i - 1] - 2\lambda_2[\Sigma a_i x_i - x_0]$$
$$= var(Y_0) + var(\hat{Y}_0) + \underbrace{2\,cov(Y_0, \hat{Y}_0)} - \lambda_1[\Sigma a_i - 1] - \lambda_2[\Sigma a_i x_i - x_0]$$

$$cov(Y_0, \hat{Y}_0) = cov(Y_0, \Sigma a_i Y_i)$$
$$= \Sigma a_i \, cov(Y_0, Y_i) = 0$$
because $Y_i \perp\!\!\!\perp Y_j$ s.t. $i \neq j$

$$= var(Y_0) + var(\hat{Y}_0) - 2\lambda_1[\Sigma a_i - 1] - 2\lambda_2[\Sigma a_i x_i - x_0]$$
$$= \sigma^2 + \sigma^2 \Sigma a_i^2 - 2\lambda_1[\Sigma a_i - 1] - 2\lambda_2[\Sigma a_i x_i - x_0]$$

$$\frac{\partial Q}{\partial a_i} = 2\sigma^2 a_i - 2\lambda_1 - 2\lambda_2 x_i = 0$$
$$a_i = \frac{2\lambda_1 + 2\lambda_2 x_i}{2\sigma^2} = \frac{\lambda_1 + \lambda_2 x_i}{\sigma^2}$$

constraints: $\Sigma a_i = 1$ $\quad \Sigma a_i x_i = x_0$

$$\Sigma a_i = \Sigma \frac{\lambda_1 + \lambda_2 x_i}{\sigma^2} = \frac{1}{\sigma^2} \Sigma \lambda_1 + \lambda_2 x_i = \frac{n\lambda_1}{\sigma^2} + \frac{\lambda_2 \Sigma x_i}{\sigma^2}$$

$$\Sigma a_i x_i = \Sigma \frac{\lambda_1 + \lambda_2 x_i}{\sigma^2} x_i = \Sigma \frac{\lambda_1 x_i + \lambda_2 x_i^2}{\sigma^2} = \frac{1}{\sigma^2} \Sigma \lambda_1 x_i + \lambda_2 x_i^2$$
$$= \frac{\lambda_1 \Sigma x_i}{\sigma^2} + \frac{\lambda_2 \Sigma x_i^2}{\sigma^2}$$

Therefore:

$$\begin{bmatrix} \frac{n\lambda_1}{\sigma^2} + \frac{\lambda_2 \Sigma x_i}{\sigma^2} \\ \frac{\lambda_1 \Sigma x_i}{\sigma^2} + \frac{\lambda_2 \Sigma x_i^2}{\sigma^2} \end{bmatrix} = \begin{bmatrix} 1 \\ x_0 \end{bmatrix} \rightarrow \frac{1}{\sigma^2} \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 \\ x_0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \sigma^2 \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ x_0 \end{bmatrix}$$

this matrix will be invertable iff $n\Sigma x_i^2 - \Sigma x_i \Sigma x_i \neq 0$

$$\begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}^{-1} = \frac{1}{\underbrace{n\Sigma x_i^2 - (\Sigma x_i)^2}_{= n\Sigma(x_i-\bar{x})^2}} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \frac{\sigma^2}{n\Sigma(x_i-\bar{x})^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix} \begin{bmatrix} 1 \\ x_0 \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \frac{\sigma^2}{n\Sigma(x_i-\bar{x})^2} \begin{bmatrix} \Sigma x_i^2 - x_0 \Sigma x_i \\ -\Sigma x_i + n x_0 \end{bmatrix} = \frac{\sigma^2}{\Sigma(x_i-\bar{x})^2} = \begin{bmatrix} \bar{x}^2 - x_0\bar{x} \\ -\bar{x} + x_0 \end{bmatrix}$$

$$\lambda_1 = \frac{\sigma^2(\bar{x}^2 - x_0\bar{x})}{\Sigma(x_i-\bar{x})^2}$$

$$\lambda_2 = \frac{\sigma^2(x_0 - \bar{x})}{\Sigma(x_i-\bar{x})^2}$$

$$\hat{Y}_0 = \Sigma a_i y_i = \frac{1}{\sigma^2}\Sigma(\lambda_1 + \lambda_2 x_i) y_i = \frac{\lambda_1}{\sigma^2}\Sigma y_i + \frac{\lambda_2}{\sigma^2}\Sigma x_i y_i$$

$$= \frac{(\bar{x}^2 - x_0\bar{x})}{\Sigma(x_i-\bar{x})^2}\Sigma y_i + \frac{(x_0-\bar{x})}{\Sigma(x_i-\bar{x})^2}\Sigma x_i y_i = \frac{\bar{x}^2\Sigma y_i - x_0\bar{x}\Sigma y_i + x_0\Sigma x_i y_i + \bar{x}\Sigma x_i y_i}{\Sigma(x_i-\bar{x})^2}$$

$$= \frac{\bar{x}^2\Sigma y_i + \bar{x}\Sigma x_i y_i + \Sigma x_0 y_i(x_i - \bar{x})}{\Sigma(x_i-\bar{x})^2} = \frac{\bar{x}^2\Sigma y_i + \bar{x}\Sigma x_i y_i}{\Sigma(x_i-\bar{x})^2} + \frac{x_0\Sigma(x_i-\bar{x})y_i}{\Sigma(x_i-\bar{x})^2}$$

Note: $\hat{\beta}_1 = \dfrac{\Sigma(x_i-\bar{x})Y_i}{\Sigma(x_i-\bar{x})^2}$ $\qquad \Sigma x_i^2 = \Sigma(x_i-\bar{x})^2 + n\bar{x}^2$

$$= \frac{(\frac{1}{n}\Sigma(x_i-\bar{x})^2\,\Sigma y_i)}{\Sigma(x_i-\bar{x})^2} + \frac{(\bar{x}\,\Sigma\bar{x}y_i) - (\bar{x}\,\Sigma x_i y_i)}{\Sigma(x_i-\bar{x})^2} + \hat{\beta}_1 x_0$$

$$= \frac{\Sigma y_i}{n} - \frac{\bar{x}\,\Sigma(x_i-\bar{x})y_i}{\Sigma(x_i-\bar{x})^2} + \hat{\beta}_1 x_0 = \bar{y} - \bar{x}\hat{\beta}_1 + \hat{\beta}_1 x_0 =$$

$$\bar{y} + \hat{\beta}_1(x_0 - \bar{x}) = \hat{Y}_0$$

Exercise 4

$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ $\qquad H_0: \beta_1 = 0$

$\varepsilon_i \sim N(0, \sigma^2)$ $\qquad H_a: \beta_1 \neq 0$

what happens to the F-statistic when we work with $y - c$?

$$F = \frac{\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{S_e^2} = \frac{(n-2)\hat{\beta}_1^2 \sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y} - \hat{\beta}_1 (x_i - \bar{x}))^2}$$

what is the estimate of β_1 when we use $y_i - c$ instead of y_i?

know: $\hat{\beta}_1 = \dfrac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2}$

so if we use $y_i - c$: $\hat{\beta}_1^{new} = \dfrac{\sum (x_i - \bar{x})(y_i - c)}{\sum (x_i - \bar{x})^2} = \dfrac{\sum (x_i - \bar{x}) y_i - \overbrace{c \sum (x_i - \bar{x})}^{=0}}{\sum (x_i - \bar{x})^2}$

$$= \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} = \hat{\beta}_1$$

For simplicity we define: $\tilde{y}_i = y_i - c$

Therefore:

$$F_{new} = \frac{(n-2)\hat{\beta}_1^{new} \sum (x_i - \bar{x})^2}{\sum (\tilde{y}_i - \frac{1}{n}\sum \tilde{y}_i - \hat{\beta}_1^{new}(x_i - \bar{x}))^2}$$

above, we showed that $\hat{\beta}_1^{new} = \hat{\beta}_1$

Therefore:

$$F_{new} = \frac{(n-2)\hat{\beta}_1 \sum (x_i - \bar{x})^2}{\sum (\tilde{y}_i - \frac{1}{n}\sum \tilde{y}_i - \hat{\beta}_1 (x_i - \bar{x}))^2}$$

$$= \frac{(n-2)\hat{\beta}_1 \sum (x_i - \bar{x})^2}{\sum ((Y_i - c) - (\bar{Y} - c) - \hat{\beta}_1 (X_i - \bar{X}))^2}$$

$$\frac{1}{n}\sum \tilde{y}_i = \frac{1}{n}\sum y_i - c$$

$$= \frac{1}{n}\left[\sum y_i - nc\right]$$

$$= \frac{1}{n}\sum y_i - c = \bar{Y} - c$$

$$= \frac{(n-2)\hat{\beta}_1 \Sigma (x_i - \bar{x})^2}{\Sigma((y_i - c) - (\bar{y} - c) - \hat{\beta}_1 (x_i - \bar{x}))^2} = \frac{(n-2)\hat{\beta}_1 \Sigma (x_i - \bar{x})^2}{\Sigma((y_i - c - \bar{y} + c) - \hat{\beta}_1 (x_i - \bar{x}))^2}$$

$$= \frac{(n-2)\hat{\beta}_1 \Sigma (x_i - \bar{x})^2}{\Sigma(y_i - \bar{y} - \hat{\beta}_1 (x_i - \bar{x}))^2} = \frac{\hat{\beta}_1 \Sigma (x_i - \bar{x})^2}{S_e^2}$$

$$F_{new} = F$$

Therefore, if a constant c is subtracted from each y_i the F statistic for testing H_0 does not change.

Now let $Y_i = \beta_1 x_i + \varepsilon_i$ $H_0: \beta_1 = 0$

$\varepsilon_i \sim N(0, \sigma)$ $H_a: \beta_1 \neq 0$

What happens to the F statistic when we work with $y - c$?

$$F = \frac{\hat{\beta}_1^2 \Sigma x_i^2}{S_e^2} = \frac{(n-2)\hat{\beta}_1^2 \Sigma x_i^2}{\Sigma (y_i - \hat{\beta}_1 x_i)^2}$$

What is the estimate of β_1 when we use $y_i - c$ instead of y_i?

To simplify, we define: $\tilde{y}_1 = y_i - c$

$$\hat{\beta}_1 = \frac{\Sigma x_i y_i}{x_i^2} \quad \text{therefore} \quad \hat{\beta}_1^{new} = \frac{\Sigma (x_i)(y_i - c)}{x_i^2}$$

Because $\hat{\beta}_1 \neq \hat{\beta}_1^{new}$ it follows that F (which uses β_1) $\neq F_{new}$ (which uses β_1^{new}).

Therefore, if the model does not include an intercept β_1 then the F statistic testing H_0 will change if you subtract a constant c from each y_i.

Exercise 5

a. $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ $\quad\quad$ $H_0: \beta_1 = 0$

$\varepsilon_i \sim N(0, \sigma)$ $\quad\quad$ $H_a: \beta_1 \neq 0$

Note: $\hat{\beta}_1 \sim \left(\beta_1, \dfrac{\sigma^2}{\Sigma(x_i - \bar{x})^2}\right)$ under assumptions of $H_0: \beta_1 = 0$

$$\Rightarrow \hat{\beta}_1 \sim \left(0, \frac{\sigma^2}{\Sigma(x_i-\bar{x})^2}\right) \Rightarrow \frac{\hat{\beta}_1 - 0}{\sqrt{\sigma^2 / \Sigma(x_i-\bar{x})^2}} \sim N(0,1)$$

We know: $\dfrac{(n-2)\, S_e^2}{\sigma^2} \sim \chi^2_{n-2}$

Note: If $U \sim N(0,1)$, $V \sim \chi^2_m$, and $U \perp\!\!\!\perp V \Rightarrow \dfrac{U}{\sqrt{V/m}} \sim t_m$

Let $U = \dfrac{\hat{\beta}_1 - 0}{\sqrt{\sigma^2/\Sigma(x_i-\bar{x})^2}}$ and $V = \dfrac{(n-2)\, S_e^2}{\sigma^2}$

then
$$t_{stat} = \frac{\dfrac{\hat{\beta}_1}{\sqrt{\sigma^2/\Sigma(x_i-\bar{x})^2}}}{\sqrt{\dfrac{(n-2)S_e^2}{\sigma^2}/(n-2)}} = \frac{\hat{\beta}_1 \sqrt{\Sigma(x_i-\bar{x})^2}}{\sqrt{S_e^2}}$$

Note: $\hat{\beta}_1 = \dfrac{\Sigma(x_i - \bar{x})Y_i}{\Sigma(x_i - \bar{x})^2}$ $\quad\quad$ $S_e^2 = \dfrac{\Sigma e_i^2}{(n-2)} = \dfrac{\Sigma((Y_i - \bar{Y}) - \hat{\beta}_1(x_i - \bar{x}))^2}{(n-2)}$

$n = 30$

$\Rightarrow \hat{\beta}_1 = 0.2159301$ $\quad\quad$ $\Rightarrow S_e^2 = 411.116$

I used loops in R to calculate $\hat{\beta}_1$ and S_e^2 to solve for t_{stat}, the final value I got for t_{stat} is:

$t_{stat} = 18.22474$ $\quad\quad$ $df = n-2 = 30-2 = 28$

To calculate the p-value:

$p = 2 * pt(q = 18.22474,\ df = 28,\ lower.tail = F) = 4.5754 \times 10^{-17}$

Because our p-value is smaller than $\alpha = 0.05$ we reject $H_0: \beta_1 = 0$

b. Find: $E\left[\dfrac{\dfrac{\hat{\beta}_1 - 0}{\sigma/\sqrt{\Sigma(x_i-\bar{x})^2}}}{\sqrt{\dfrac{(n-2)S_e^2}{\sigma^2}/(n-2)}}\right] \overset{\text{simplify the ratio}}{=} E\left[\dfrac{\hat{\beta}_1\sqrt{\Sigma(x_i-\bar{x})^2}}{\sqrt{S_e^2}}\right] = \sqrt{\Sigma(x_i-\bar{x})^2}\; E\left[\dfrac{\hat{\beta}_1}{S_e^2}\right]$

we know that $\hat{\beta}_1 \perp\!\!\!\perp S_e^2$ because $\hat{\beta}_1 \perp\!\!\!\perp e_i$

$\Rightarrow \sqrt{\Sigma(x_i-\bar{x})^2}\; E[\hat{\beta}_1] \cdot E\left[\dfrac{1}{S_e^2}\right]$

$\hat{\beta}_1$ is an unbiased estimator of β_1 therefore,
$E[\hat{\beta}_1] = \beta_1$

$E\left[\dfrac{1}{S_e^2}\right] = E\left[(S_e^2)^{-1/2}\right]$

$S_e^2 \sim \Gamma\left(\dfrac{n-2}{2}, \dfrac{2\sigma^2}{n-2}\right)$

and if $Q \sim \Gamma(\alpha, \beta) \Rightarrow E(Q^k) = \dfrac{\Gamma(\alpha+k)\,\beta^k}{\Gamma(\alpha)}$

$E\left[(S_e^2)^{-1/2}\right] = \dfrac{\Gamma\left(\frac{n-2}{2} - 1/2\right)\left(\frac{2\sigma^2}{n-2}\right)^{-1/2}}{\Gamma\left(\frac{n-2}{2}\right)}$

$\sqrt{\Sigma(x_i-\bar{x})^2}\; E[\hat{\beta}_1] \cdot E\left[\dfrac{1}{S_e^2}\right] = \sqrt{\Sigma(x_i-\bar{x})^2} \cdot \beta_1 \cdot \dfrac{\Gamma\left(\frac{n-2}{2} - 1/2\right)\left(\frac{2\sigma^2}{n-2}\right)^{-1/2}}{\Gamma\left(\frac{n-2}{2}\right)}$

$= E[t_{stat}]$

c. $\beta_1 = 0.05$ $\quad \sigma^2 = 600$ $\quad \alpha = 0.05$ (significance level)

power: probability that we reject H_0 when $\beta_1 \neq 0$

Definition of a non-central t-distribution:

If $u \sim N(\gamma, 1)$, $V \sim \chi^2_m$ and $u \perp\!\!\!\perp V$ then:

$$\frac{u}{\sqrt{V/m}} \sim t_{m,\gamma} \quad \text{where } m = df \text{ and } \gamma = ncp$$

Let $u = \dfrac{\hat{\beta}_1}{\sqrt{\sigma^2/\Sigma(x_i-\bar{x})^2}}$ $\qquad V = \dfrac{(n-2)\,Se^2}{\sigma^2} \sim t_{n-2}$

$$t_{stat} = \frac{\hat{\beta}_1\sqrt{\Sigma(x_i-\bar{x})^2}}{\sqrt{Se^2}} = \frac{u}{\sqrt{V/(n-2)}}$$

$$\gamma = E[u]$$

$$E\left[\frac{\hat{\beta}_1}{\sqrt{\sigma^2/\Sigma(x_i-\bar{x})^2}}\right] = \frac{\beta_1}{\sqrt{\sigma^2/\Sigma(x_i-\bar{x})^2}} = \gamma$$

$$\gamma = \frac{0.05}{\sqrt{600/1711.317}} = 0.0844 \Rightarrow t_{stat} \sim t_{n-2,\,\gamma = 3.4932}$$

$$\text{power} = P(\underbrace{t_{stat}}_{\text{follows non-central t-distribution}} > t_\alpha) + P(t_{stat} < \underbrace{-t_\alpha}_{\text{comes from } t_{n-2}})$$

R code:

$qt(0.025, 28, lower.tail = F) = t_\alpha = 2.048407$

$\Rightarrow -t_\alpha = -2.048407$

power = pt(q = 2.048407, 28, ncp = 3.4932, lower.tail = F)

+ pt(q = -2.048407, 28, ncp = 3.4932) = 0.9208