

2.  $\left\{ \begin{pmatrix} Y_j \\ \mathbf{X}_j \end{pmatrix} \right\}_{j=1}^n$  vectores aleatorios i.i.d. dimensión  $k$ .

$$\mathbf{Y} = \mathbf{X}\mathbf{b} + \mathbf{U}, \mathbb{E}[\mathbf{U}|\mathbf{X}] = 0.$$

$$\mathbb{P}[\text{Rango}(\mathbf{X}) = k] = 1$$

$$\text{Var}(\mathbf{U}|\mathbf{X}) = \sigma_u^2 \mathbb{I}_n \text{ (homocedasticidad condicional).}$$

$$\mathbf{U}|\mathbf{X}=\mathbf{X} \sim N_n \quad \forall \mathbf{X},$$

donde  $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_k \end{pmatrix}$

i) Para  $i \in 1, \dots, k$  proponer una prueba de hipótesis, i.e. *una regla de decisión* para aceptar o rechazar la hipótesis,

$$\mathbf{H}_0 : b_i = b_i^*$$

contra la alternativa

$$\mathbf{H}_1 : b_i > b_i^*$$

con un nivel de significancia  $\alpha$

Podemos utilizar el pivote:

$$\frac{\hat{b}_{iMCO} - b_i}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{ii}}} \sim t_{n-k}$$

Lo anterior, al contener a  $b_i$  (no conocido aún) no es un estadístico. No obstante, bajo  $\mathbf{H}_0$  sabemos que  $b_i = b_i^*$ , de modo que el pivote anterior se convierte en un estadístico pues ya no contiene parámetros desconocidos:

$$\tau_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{ii}}} \sim t_{n-k}$$

Sabemos, por el teorema de Gauss-Markov (“GM”), que  $\hat{b}_{iMCO}$  es el mejor estimador (el más eficiente) lineal condicionalmente insesgado (*BLCUE*) de  $b_i$ ; es decir:  $\hat{b}_{iMCO} \approx b_i$ . De lo anterior, y de la hipótesis nula, tenemos las siguientes implicaciones:

$$\mathbf{H}_0 : b_i = b_i^* \xrightarrow{GM} \hat{b}_{iMCO} \approx b_i = b_i^* \implies \hat{b}_{iMCO} - b_i^* \approx 0$$

Es decir, bajo la hipótesis nula, el estadístico  $\tau_{\hat{b}_i}$  es un número pequeño; si es lo suficientemente pequeño, no podemos rechazar la hipótesis nula. Es decir, no podemos rechazar si, dado un nivel de significancia  $\alpha$ , tenemos que  $\tau_{\hat{b}_i} < t_\alpha$  (al ser una prueba de una cola).

Finalmente, bajo la hipótesis alternativa, tenemos las siguientes implicaciones:

$$\mathbf{H}_1 : b_i > b_i^* \xrightarrow{GM} \hat{b}_{iMCO} \approx b_i > b_i^* \implies \hat{b}_{iMCO} - b_i^* > 0$$

Por lo que el estadístico de prueba,  $\tau_{\hat{b}_i}$ , sería positivo (mayor a cero). Entonces, requerimos que el estadístico calculado anteriormente  $\tau_{\hat{b}_i}$  sea un número suficientemente grande, mayor al valor crítico  $t_\alpha$ , para poder rechazar la hipótesis nula.

En conclusión, la **regla de decisión** es tal que:

$$\text{Si } \tau_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{ii}^{-1}}} < t_\alpha, \text{ no rechazamos } \mathbf{H}_0 : b_i = b_i^*.$$

$$\text{Si } \tau_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{ii}^{-1}}} \geq t_\alpha, \text{ rechazamos } \mathbf{H}_0 : b_i = b_i^* \text{ en favor de } \mathbf{H}_1 : b_i > b_i^*.$$

ii) Para  $i \in 1, \dots, k$  proponer una prueba de hipótesis, i.e. ***una regla de decisión para aceptar o rechazar la hipótesis***,

$$\mathbf{H}_0 : b_i = b_i^*$$

contra la alternativa

$$\mathbf{H}_1 : b_i < b_i^*$$

con un nivel de significancia  $\alpha$

Al igual que en el inciso anterior, podemos utilizar el pivote:

$$\frac{\hat{b}_{iMCO} - b_i}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{ii}^{-1}}} \sim t_{n-k}$$

Lo anterior, al contener a  $b_i$  (no conocido aún) no es un estadístico. No obstante, bajo  $\mathbf{H}_0$  sabemos que  $b_i = b_i^*$ , de modo que el pivote anterior se convierte en un estadístico pues ya no contiene parámetros desconocidos:

$$\tau_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{ii}^{-1}}} \sim t_{n-k}$$

Sabemos, por el teorema de Gauss-Markov (“GM”), que  $\hat{b}_{iMCO}$  es el mejor estimador (el más eficiente) lineal condicionalmente insesgado (*BLCUE*) de  $b_i$ ; es decir:  $\hat{b}_{iMCO} \approx b_i$ . De lo anterior, y de la hipótesis nula, tenemos las siguientes implicaciones:

$$\mathbf{H}_0 : b_i = b_i^* \xrightarrow{GM} \hat{b}_{iMCO} \approx b_i = b_i^* \implies \hat{b}_{iMCO} - b_i^* \approx 0$$

Es decir, bajo la hipótesis nula, el estadístico  $\tau_{\hat{b}_i}$  es un número pequeño; si es lo suficientemente pequeño, no podemos rechazar la hipótesis nula. Es decir, no podemos rechazar si, dado un nivel de significancia  $\alpha$ , tenemos que  $\tau_{\hat{b}_i} > -t_\alpha$  (al ser una prueba de cola izquierda).

Finalmente, bajo la hipótesis alternativa, tenemos las siguientes implicaciones:

$$\mathbf{H}_1 : b_i < b_i^* \xrightarrow{GM} \hat{b}_{iMCO} \approx b_i < b_i^* \implies \hat{b}_{iMCO} - b_i^* < 0$$

Por lo que el estadístico de prueba,  $\tau_{\hat{b}_i}$ , sería negativo (menor a cero). Entonces, requerimos que el estadístico calculado anteriormente  $\tau_{\hat{b}_i}$  sea un número suficientemente “negativo”, menor (mayor en términos absolutos) al valor crítico  $-t_\alpha$ , para poder rechazar la hipótesis nula.

En conclusión, la **regla de decisión** es tal que:

$$\text{Si } \tau_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{ii}}} > -t_\alpha, \text{ no rechazamos } \mathbf{H}_0 : b_i = b_i^*.$$

$$\text{Si } \tau_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{ii}}} \leq -t_\alpha, \text{ rechazamos } \mathbf{H}_0 : b_i = b_i^* \text{ en favor de } \mathbf{H}_1 : b_i < b_i^*.$$

Los siguientes problemas están en Wooldridge, J.M. *Introductory Econometrics: A Modern Approach. 6th Edition.* South-Western. CENGAGE Learning. Las bases de datos se encuentran en [www.cengagebrain.com](http://www.cengagebrain.com)

### 3. Problema 2, capítulo 4, pág. 141.

Consider an equation to explain salaries of CEOs in terms of annual firm sales, return on equity (*roe*, in percentage form), and return on the firm's stock (*ros*, in percentage form):

$$\log(\text{salary}) = \beta_0 + \beta_1 \log(\text{sales}) + \beta_2 \text{roe} + \beta_3 \text{ros} + u$$

i) In terms of the model parameters, state the null hypothesis that, after controlling for *sales* and *roe*, *ros* has no effect on CEO salary. State the alternative that better stock market performance increases a CEO's salary.

Null hypothesis:

$$\mathbf{H}_0 : \beta_3 = 0$$

Alternative hypothesis

$$\mathbf{H}_1 : \beta_3 > 0$$

ii) Using the data in CEOSAL1, the following equation was obtained by OLS:

$$\begin{aligned} \widehat{\log(\text{salary})} &= 4.32 + .280 \log(\text{sales}) + .0174 \text{roe} + .00024 \text{ros} \\ &\quad (.32) \quad (.035) \quad (.0041) \quad (.00054) \\ n &= 209, R^2 = .283 \end{aligned}$$

By what percentage is *salary* predicted to increase if *ros* increases by 50 points? Does *ros* have a practically large effect on *salary*?

Following the correspondent interpretation (that of a Log-Level model), we can calculate the effect as follows:

$$\begin{aligned} \% \Delta \text{salary} &= (100\beta_3) * \Delta \text{ros} \implies \% \Delta \text{salary} = (100 * .00024) * 50 \\ &= (0.024) * 50 \\ &= 1.2 \end{aligned}$$

That is, if *ros* increases 50 points, salary increases by about 1.2%, which is a reasonably small effect.

iii) Test the null hypothesis that *ros* has no effect on *salary* against the alternative that *ros* has a positive effect. Carry out the test at the 10% significance level.

We can directly use the results from exercise 2.i) of this homework, where the decision rule would be:

$$\text{If } \tau_{\hat{b}_3} = \frac{\hat{b}_{3MCO} - b_3^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{44}}} < t_\alpha, \text{ we can't reject } \mathbf{H}_0 : b_3 = b_3^*.$$

$$\text{If } \tau_{\hat{b}_3} = \frac{\hat{b}_{3MCO} - b_3^*}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{44}}} \geq t_\alpha, \text{ we reject } \mathbf{H}_0 : b_3 = b_3^* \text{ in favor of } \mathbf{H}_1 : b_3 > b_3^*.$$

Where  $b_i^* = 0$ , and  $\alpha = 10\%$ .

Let's compute all we need.

Data:

```
rm(list = ls())
library(wooldridge)
Z <- ceosal1[, c("lsalary", "lsales", "roe", "ros")]
Z <- as.matrix(Z)
X <- cbind(1, Z[, -1])

n <- nrow(Z)
k <- ncol(Z)
```

Given that  $\alpha = 10\%$ , we can compute  $t_{\alpha/2}$ :

```
t_q <- qt(p = 0.1, df = n-k, lower.tail = F)
t_q
```

```
## [1] 1.285695
```

Now we'll compute the estimators.

```
#Matriz (X'X)^{-1}
XX_inv <- solve(t(X)%*%X)

#Estimadores de MCO
b_mco <- XX_inv%*%t(X)%*%Z[,1]
b_mco
```

```
##           [,1]
## 4.3117124632
## lsales 0.2803148895
## roe    0.0174167522
## ros    0.0002416552
```

Now calculate  $\hat{\sigma}_{uMCO}^2$ .

```
#Errores de MCO
U_mco <- Z[,1] - X%*%b_mco
#Varianza de los errores de MCO
var.u_mco <- 1/(n-k)*t(U_mco)%*%U_mco
#Desviación estándar de los errores de MCO
sd.u_mco <- sqrt(var.u_mco)
sd.u_mco
```

```
##           [,1]
## [1,] 0.4831847
```

And then we can compute the statistic:

```
#estadístico de prueba para \beta_3
t_b3 <- b_mco[4]/(sd.u_mco*sqrt(XX_inv[4,4]))
t_b3
```

```
##           [,1]
## [1,] 0.4460213
```

Hence:

$$\frac{\hat{b}_{3MCO}}{\hat{\sigma}_{uMCO}\sqrt{(X'X)^{-1}_{44}}} = \frac{2.416552 \times 10^{-4}}{0.4831847 * 0.0011213} = 0.4460213 < 1.2856949$$

That is, since  $0.4460213 < 1.2856949$ , we can't reject  $\mathbf{H}_0 : b_3 = 0$ .

Following Wooldridge, another method to calculate the statistic is:

$$\frac{\hat{\beta}_{3MCO}}{\sqrt{\text{var}(\hat{\beta}_{3MCO}|\mathbf{X})}} \sim t_{n-k}$$

That is:

```
#Matriz de varianza de estimadores de MCO
var.b_mco <- as.numeric(var.u_mco)*XX_inv
#errores estandar de los estimadores de MCO
sd.b_mco <- sqrt(diag(var.b_mco))
#estadístico de prueba
b_mco[4]/sd.b_mco[4]
```

```
##           ros
## 0.4460213
```

Hence:

$$\frac{\hat{\beta}_{3MCO}}{\sqrt{\text{var}(\hat{\beta}_{3MCO}|\mathbf{X})}} = \frac{2.416552 \times 10^{-4}}{5.4180195 \times 10^{-4}} = 0.4460213$$

The rest of the hypothesis test is the same as we already did.

**iv) Would you include *ros* in a final model explaining CEO compensation in terms of firm performance? Explain.**

Recall the discussion on statistical significance vs economic or practical significance: while the first is entirely determined by  $\tau_{\hat{\beta}_3}$ , the second is related to the size and sign of  $\hat{\beta}_3$ . We've seen that the coefficient is not statistically significant even at the 10% level; furthermore, we also know that although the sign of  $\hat{\beta}_3$  is as expected, its size isn't relevant, as was demonstrated in ii). Given these two arguments, we could incline towards removing the variable from the model.

Nonetheless, though we can't reject the null hypothesis that  $\beta_3 = 0$ , we can't remove *ros* from our model unless we're sure that  $\mathbb{E}[ros|sales,roe] = 0$ . If it is not the case, removing *ros* and sending it to *u* would make us violate the assumption that  $\mathbb{E}[u|\mathbf{X}] = 0$ .

#### 4. Problema 4, capítulo 4, pág. 141.

Are rent rates influenced by the student population in a college town? Let  $rent$  be the average monthly rent paid on rental units in a college town in the United States. Let  $pop$  denote the total city population,  $avginc$  the average city income, and  $pctstu$  the student population as a percentage of the total population. One model to test for a relationship is

$$\log(rent) = \beta_0 + \beta_1 \log(pop) + \beta_2 \log(avginc) + \beta_3 pctstu + u$$

---

i) State the null hypothesis that size of the student body relative to the population has no ceteris paribus effect on monthly rents. State the alternative that there is an effect.

Null hypothesis:

$$H_0 : \beta_3 = 0$$

Alternative hypothesis

$$H_1 : \beta_3 \neq 0$$

ii) What signs do you expect for  $\beta_1$  and  $\beta_2$ ?

Positive for both; the larger the city population, the greater the demand for rental units. Also, as the average city income increases, citizens can pay higher rents for higher quality rental units.

iii) The equation estimated using 1990 data from RENTAL for 64 college towns is

$$\widehat{\log(rent)} = \underset{(.844)}{.043} + \underset{(.039)}{.066} \log(pop) + \underset{(.081)}{.507} \log(avginc) + \underset{(.0017)}{.0056} pctstu$$
$$n = 64, R^2 = .458$$

**What is wrong with the statement: “A 10% increase in population is associated with about a 6.6% increase in rent”?**

Recall the interpretation of the coefficients in the case of a log-log model: if we have  $\log(y)$  as the dependent variable, and  $\log(x)$  as the independent variable, then the interpretation of the associated coefficient  $\beta$  is such that  $\% \Delta y = \beta * \% \Delta x$ . Considering this, in combination with the linear model results and a 10% increase in population, we have the following:

$$\begin{aligned} \% \Delta rent &= \beta_1 * \% \Delta pop \implies \% \Delta rent = .066 * 10\% \\ &= .066 * .1 \\ &= 0.0066 \\ &= 0.66\% \end{aligned}$$

That is to say, a 10% increase in population is actually associated with about a 0.66% increase in rent, not 6.6% as the statement claimed.

iv) Test the hypothesis stated in part (i) at the 1% level.

Following the same methodology as in *Tarea 15*, we can compute a statistic of the form:

$$\frac{\hat{\beta}_{3MCO}}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{44}^{-1}}} \sim t_{n-k}$$

With the decision rule being:

If  $\left| \frac{\hat{\beta}_{3MCO}}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{44}^{-1}}} \right| < t_{\alpha/2}$  we can't reject  $\mathbf{H}_0 : b_3 = 0$ .

If  $\left| \frac{\hat{\beta}_{3MCO}}{\hat{\sigma}_{uMCO} \sqrt{(X'X)_{44}^{-1}}} \right| \geq t_{\alpha/2}$  we reject  $\mathbf{H}_0 : b_3 = 0$  in favor of  $\mathbf{H}_1 : b_3 \neq 0$

We can compute  $\hat{\beta}_{3MCO}$  (for more precision),  $\hat{\sigma}_{uMCO}$  and  $\sqrt{(X'X)_{44}^{-1}}$ .

First we'll get the data and define  $n$  and  $k$ .

```
rm(list = ls())
library(wooldridge)
Z <- rental[rental$y90==1, c("lrent", "lpop", "lavginc", "pctstu")]
Z <- as.matrix(Z)
X <- cbind(1, Z[, -1])

n <- nrow(Z)
k <- ncol(Z)
```

Given that we're looking to test the hypothesis with a 1% significance level, we can compute  $t_{\alpha/2}$  where  $\alpha = 1\%$

```
t_q <- qt(p = 0.01/2, df = n-k, lower.tail = F)
t_q
```

```
## [1] 2.660283
```

Now we'll compute the estimators.

```
#Matriz (X'X)^{-1}
XX_inv <- solve(t(X)%*%X)

#Estimadores de MCO
b_mco <- XX_inv%*%t(X)%*%Z[,1]
b_mco
```

```
##           [,1]
## 0.042780259
## lpop 0.065867849
## lavginc 0.507015040
## pctstu 0.005629696
```



Now calculate  $\hat{\sigma}_{uMCO}^2$ .

```
#Errores de MCO
U_mco <- Z[,1] - X%*%b_mco
#Varianza de los errores de MCO
var.u_mco <- 1/(n-k)*t(U_mco)%*%U_mco
#Desviación estándar de los errores de MCO
sd.u_mco <- sqrt(var.u_mco)
sd.u_mco
```

```
##           [,1]
## [1,] 0.1512196
```

And then we can compute the statistic:

```
#estadístico de prueba para \beta_3
t_b3 <- b_mco[4]/(sd.u_mco*sqrt(XX_inv[4,4]))
t_b3
```

```
##           [,1]
## [1,] 3.231619
```

Hence:

$$\left| \frac{\hat{b}_{3MCO}}{\hat{\sigma}_{uMCO} \sqrt{(X'X)^{-1}_{44}}} \right| = \left| \frac{0.0056297}{0.1512196 * 0.0115201} \right| = |3.2316194| > 2.660283$$

That is, since  $|3.2316194| > 2.660283$ , we reject  $\mathbf{H}_0 : b_3 = 0$  in favor of  $\mathbf{H}_1 : b_3 \neq 0$ .

Following Wooldridge, another method to calculate the statistic is:

$$\frac{\hat{\beta}_{3MCO}}{\sqrt{\text{var}(\hat{\beta}_{3MCO}|\mathbf{X})}} \sim t_{n-k}$$

That is:

```
#Matriz de varianza de estimadores de MCO
var.b_mco <- as.numeric(var.u_mco)*XX_inv
#errores estandar de los estimadores de MCO
sd.b_mco <- sqrt(diag(var.b_mco))
```

Hence:

$$\frac{\hat{\beta}_{3MCO}}{\sqrt{\text{var}(\hat{\beta}_{3MCO}|\mathbf{X})}} = \frac{0.0056297}{0.0017421} = 3.2316194$$

The rest of the hypothesis test is the same as we already did.

5.  $\left\{ \begin{pmatrix} Y_j \\ X_{1j} \\ X_{2j} \end{pmatrix} \right\}_{j=1}^{100}$  vectores aleatorios i.i.d. dimensión 3.  $\begin{pmatrix} Y_j \\ X_{1j} \\ X_{2j} \end{pmatrix} \sim N_3$ .

$$Y_j = b_1 + b_2 X_{1j} + b_3 X_{2j} + U_j, \quad \mathbb{E}[U_j | X_{1j}, X_{2j}] = 0, \quad j = 1, 2, \dots, 100$$


---

i) Generar una muestra i.i.d.  $\left\{ \begin{pmatrix} Y_j \\ X_{1j} \\ X_{2j} \end{pmatrix} \right\}_{j=1}^{100}$  donde  $\begin{pmatrix} Y_j \\ X_{1j} \\ X_{2j} \end{pmatrix} \sim N_3 \left( \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0.8 & 0.4 & -0.2 \\ 0.4 & 1.0 & -0.8 \\ -0.2 & -0.8 & 2.0 \end{pmatrix} \right)$ .

```
library(mvtnorm)
set.seed(2)
rm(list = ls())

n <- 100
mu <- c(1, 0, 2)
Sigma <- matrix(c(0.8, 0.4, -0.2,
                  0.4, 1, -0.8,
                  -0.2, -0.8, 2),
                nrow = 3,
                byrow = T)
Z <- data.frame(
  rmvnorm(
    n,
    mean = mu,
    sigma = Sigma
  )
)
names(Z) <- c("Y", "X_1", "X_2")

Z <- as.matrix(Z)
X <- cbind(1, Z[, -1])

n <- nrow(Z)
k <- ncol(Z)
```

ii) Se quiere probar la hipótesis:

$$\mathbf{H}_0 : \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.5 \\ 1.0 \end{pmatrix}$$

contra la alternativa

$$\mathbf{H}_1 : \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \neq \begin{pmatrix} 0.0 \\ 0.5 \\ 1.0 \end{pmatrix}$$

Determinar  $\mathbf{R}$  y  $\mathbf{r}$  cuando la hipótesis se escribe de la forma  $\mathbf{H}_0 : \mathbf{R}\mathbf{b} = \mathbf{r}$

Dado que en el ejemplo tenemos  $m = 3$  restricciones y  $k = 3$ ,  $\mathbf{R}$  es una matriz de dimensión  $3 \times 3$ , y  $\mathbf{r} \in \mathbb{R}^3$ . Es decir:

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{I}_3$$

$$\mathbf{r} = \begin{pmatrix} 0.0 \\ 0.5 \\ 1.0 \end{pmatrix}$$

Entonces:

$$\mathbf{H}_0 : \mathbf{Rb} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.5 \\ 1.0 \end{pmatrix} = \mathbf{r}$$

¿Se cumple que  $\text{rango}(\mathbf{R}) < k$ ?

No, dado que  $\text{rango}(\mathbf{R}) = \text{rango}(\mathbb{I}_3) = 3 = k$

iii) Probar  $\mathbf{H}_0$  con un nivel de significancia del 5% usando los estadísticos:

$$\text{a) } \frac{(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})}{m \hat{\sigma}_{u_{MCO}}^2} \sim F_{m,n-k}$$

La regla de decisión es tal que:

$$\text{Si } F_{\hat{b}} = \frac{(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})}{m \hat{\sigma}_{u_{MCO}}^2} < F_{\alpha} \text{ no rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}.$$

$$\text{Si } F_{\hat{b}} = \frac{(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})}{m \hat{\sigma}_{u_{MCO}}^2} \geq F_{\alpha} \text{ rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r} \text{ en favor de } \mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$$

Comencemos con los cálculos para construir el estadístico  $F_{\hat{b}}$ . Definamos los elementos  $\mathbf{R}$  y  $\mathbf{r}$

```
R <- diag(3)
r <- matrix(c(0, 0.5, 1))
m <- nrow(r)
```

Estimemos por MCO:

```
#Matriz (X'X)^{-1}
XX_inv <- solve(t(X)%*%X)

#Estimadores de MCO
b_mco <- XX_inv%*%t(X)%*%Z[,1]
```

Calculamos la varianza de los errores de MCO:

```
#Errores de MCO
U_mco <- Z[,1] - X%*%b_mco
#Varianza de los errores de MCO
var.u_mco <- 1/(n-k)*t(U_mco)%*%U_mco
```

El estadístico  $F_{\hat{b}}$  es:

```
F_b.hat <- (t(R%%b_mco - r)%%solve(R%%XX_inv%%t(R))%%(R%%b_mco - r)) / (m*var.u_mco)
F_b.hat
```

```
##           [,1]
## [1,] 117.4748
```

Y con  $\alpha = 5$  calculemos  $F_\alpha$ :

```
alpha <- 0.05
F_alpha <- qf(alpha, df1 = m, df2 = n-k, lower.tail = F)
F_alpha
```

```
## [1] 2.698398
```

Es decir:

$$F_{\hat{b}} = 117.4747905 > 2.6983975 = F_\alpha$$

Entonces, rechazamos  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$

$$\text{b) } \frac{\hat{\mathbf{U}}'_{MCR} \hat{\mathbf{U}}_{MCR} - \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \sim F_{m, n-k}$$

La regla de decisión es tal que:

$$\text{Si } F_{\hat{b}} = \frac{\hat{\mathbf{U}}'_{MCR} \hat{\mathbf{U}}_{MCR} - \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} < F_\alpha \text{ no rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}.$$

$$\text{Si } F_{\hat{b}} = \frac{\hat{\mathbf{U}}'_{MCR} \hat{\mathbf{U}}_{MCR} - \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \geq F_\alpha \text{ rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r} \text{ en favor de } \mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$$

Hemos visto en la tarea 14 que el numerador de este inciso es igual al numerador del inciso anterior, además de que, por definición, el denominador es una reexpresión del denominador anterior. En esencia, el estadístico de prueba debería ser igual. No obstante, para efectos del ejercicio, calcularemos todo lo necesario como si no supieramos el hecho mencionado. Dicho esto, solamente nos hace falta calcular  $\hat{\mathbf{U}}_{MCR} \equiv \mathbf{Y} - \mathbf{X}\hat{\mathbf{b}}_{MCR}$  donde  $\hat{\mathbf{b}}_{MCR} = \hat{\mathbf{b}}_{MCO} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})$ .

Comencemos con  $\hat{\mathbf{b}}_{MCR}$

```
b_mcr <- b_mco - XX_inv%%t(R)%%solve(R%%XX_inv%%t(R))%%(R%%b_mco-r)
```

Ahora  $\hat{\mathbf{U}}_{MCR}$

```
U_mcr <- Z[,1] - X%%b_mcr
```

Calculemos entonces el estadístico  $F_{\hat{b}}$

```
F_b.hat <- (t(U_mcr)%%U_mcr - t(U_mco)%%U_mco) / (m*1/(n-k)*t(U_mco)%%U_mco)
F_b.hat
```

```
##           [,1]
## [1,] 117.4748
```

Dado que no hemos cambiado el nivel de significancia, podemos concluir:

$$F_{\hat{b}} = 117.4747905 > 2.6983975 = F_{\alpha}$$

Entonces, rechazamos  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$ . Observamos que el estadístico de prueba es exactamente igual al del inciso anterior.

$$c) \frac{(\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})' \mathbf{X}' \mathbf{X} (\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \sim F_{m,n-k}$$

La regla de decisión es tal que:

$$\text{Si } F_{\hat{b}} = \frac{(\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})' \mathbf{X}' \mathbf{X} (\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \sim F_{m,n-k} < F_{\alpha} \text{ no rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}.$$

$$\text{Si } F_{\hat{b}} = \frac{(\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})' \mathbf{X}' \mathbf{X} (\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \sim F_{m,n-k} \geq F_{\alpha} \text{ rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r} \text{ en favor}$$

de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$

Calculemos:

```
numerador <- t(b_mcr - b_mco)%%solve(XX_inv)%%(b_mcr - b_mco)
denominador <- (m*1/(n-k))*t(U_mco)%%U_mco
F_b.hat <- numerador / denominador
F_b.hat
```

```
##           [,1]
## [1,] 117.4748
```

Dado que no hemos cambiado el nivel de significancia, podemos concluir:

$$F_{\hat{b}} = 117.4747905 > 2.6983975 = F_{\alpha}$$

Entonces, rechazamos  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$ . Observamos que el estadístico de prueba es exactamente igual al del inciso anterior.

iv) Se quiere probar lo hipótesis conjunta:

$$\mathbf{H}_0 : b_1 + b_3 = 1 \text{ y } b_1 - b_2 + b_3 = 0.5$$

contra la alternativa

$$\mathbf{H}_1 : b_1 + b_3 \neq 1 \text{ o } b_1 - b_2 + b_3 \neq 0.5$$

Determinar  $\mathbf{R}$  y  $\mathbf{r}$  cuando la hipótesis se escribe de la forma  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$

Dado que en el ejemplo tenemos  $m = 2$  restricciones y  $k = 3$ ,  $\mathbf{R}$  es una matriz de dimensión  $2 \times 3$ , y  $\mathbf{r} \in \mathbb{R}^2$ . Es decir:

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$\mathbf{r} = \begin{pmatrix} 1.0 \\ 0.5 \end{pmatrix}$$

Entonces:

$$\mathbf{H}_0 : \mathbf{Rb} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 + b_3 \\ b_1 - b_2 + b_3 \end{pmatrix} = \begin{pmatrix} 1.0 \\ 0.5 \end{pmatrix} = \mathbf{r}$$

¿Se cumple que  $\text{rango}(\mathbf{R}) < k$ ?

Sí, dado que  $\text{rango}(\mathbf{R}) = 2 < k = 3$

```
R <- matrix(c(1,0,1,
              1,-1,1),
            nrow = 2,
            byrow = T)
r <- matrix(c(1, 0.5))
#rango de R
fBasics::rk(R)
```

```
## [1] 2
```

v) Probar  $\mathbf{H}_0$  con un nivel de significancia del 10% usando los estadísticos:

$$\text{a) } \frac{(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})}{m \hat{\sigma}_{u_{MCO}}^2} \sim F_{m,n-k}$$

La regla de decisión es tal que:

$$\text{Si } F_{\hat{b}} = \frac{(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})}{m \hat{\sigma}_{u_{MCO}}^2} < F_{\alpha} \text{ no rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}.$$

$$\text{Si } F_{\hat{b}} = \frac{(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})}{m \hat{\sigma}_{u_{MCO}}^2} \geq F_{\alpha} \text{ rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r} \text{ en favor de } \mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$$

Calculemos lo que nos hace falta.

```
m <- nrow(r)
```

El estadístico  $F_{\hat{b}}$  es:

```
F_b.hat <- (t(R%%b_mco - r)%%solve(R%%XX_inv%%t(R))%%(R%%b_mco - r)) / (m*var.u_mco)
F_b.hat
```

```
## [1]
## [1,] 4.062545
```

Y con  $\alpha = 10$  calculemos  $F_{\alpha}$ :

```
alpha <- 0.1
F_alpha <- qf(alpha, df1 = m, df2 = n-k, lower.tail = F)
F_alpha
```

```
## [1] 2.358119
```

Es decir:

$$F_{\hat{b}} = 4.0625448 > 2.3581192 = F_{\alpha}$$

Entonces, rechazamos  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$

$$\text{b) } \frac{\hat{\mathbf{U}}'_{MCR} \hat{\mathbf{U}}_{MCR} - \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \sim F_{m, n-k}$$

La regla de decisión es tal que:

$$\text{Si } F_{\hat{b}} = \frac{\hat{\mathbf{U}}'_{MCR} \hat{\mathbf{U}}_{MCR} - \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} < F_{\alpha} \text{ no rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}.$$

$$\text{Si } F_{\hat{b}} = \frac{\hat{\mathbf{U}}'_{MCR} \hat{\mathbf{U}}_{MCR} - \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}}{m \frac{1}{n-k} \hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}} \geq F_{\alpha} \text{ rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r} \text{ en favor de } \mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$$

Recalculemos lo que falta  $\hat{\mathbf{U}}_{MCR} \equiv \mathbf{Y} - \mathbf{X}\hat{\mathbf{b}}_{MCR}$  donde

$$\hat{\mathbf{b}}_{MCR} = \hat{\mathbf{b}}_{MCO} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})$$

Comencemos con  $\hat{\mathbf{b}}_{MCR}$

```
b_mcr <- b_mco - XX_inv*t(R)%solve(R%XX_inv*t(R))%(R%b_mco-r)
```

Ahora  $\hat{\mathbf{U}}_{MCR}$

```
U_mcr <- Z[,1] - X%b_mcr
```

Calculemos entonces el estadístico  $F_{\hat{b}}$

```
F_b.hat <- (t(U_mcr)%U_mcr - t(U_mco)%U_mco) / (m*(1/(n-k))*t(U_mco)%U_mco)
F_b.hat
```

```
## [1]
## [1,] 4.062545
```

Dado que no hemos cambiado el nivel de significancia, podemos concluir:

$$F_{\hat{b}} = 4.0625448 > 2.3581192 = F_{\alpha}$$

Entonces, rechazamos  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$ . Observamos que el estadístico de prueba es exactamente igual al del inciso anterior.

$$c) \frac{(\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})' \mathbf{X}' \mathbf{X} (\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO}) / m}{(\hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}) / (n - k)} \sim F_{m, n-k}$$

La regla de decisión es tal que:

$$\text{Si } F_{\hat{b}} = \frac{(\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})' \mathbf{X}' \mathbf{X} (\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO}) / m}{(\hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}) / (n - k)} < F_{\alpha} \text{ no rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}.$$

$$\text{Si } F_{\hat{b}} = \frac{(\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO})' \mathbf{X}' \mathbf{X} (\hat{\mathbf{b}}_{MCR} - \hat{\mathbf{b}}_{MCO}) / m}{(\hat{\mathbf{U}}'_{MCO} \hat{\mathbf{U}}_{MCO}) / (n - k)} \geq F_{\alpha} \text{ rechazamos } \mathbf{H}_0 : \mathbf{Rb} = \mathbf{r} \text{ en favor de } \mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$$

Calculemos:

```
numerador <- (t(b_mcr - b_mco)%%solve(XX_inv)%%(b_mcr - b_mco))/m
denominador <- (t(U_mco)%%U_mco)/(n-k)
F_b.hat <- numerador / denominador
F_b.hat
```

```
##           [,1]
## [1,] 4.062545
```

Dado que no hemos cambiado el nivel de significancia, podemos concluir:

$$F_{\hat{b}} = 4.0625448 > 2.3581192 = F_{\alpha}$$

Entonces, rechazamos  $\mathbf{H}_0 : \mathbf{Rb} = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{Rb} \neq \mathbf{r}$ . Observamos que el estadístico de prueba es exactamente igual al del inciso anterior.

vi) Usando los parámetros dados en i), calcular los valores verdaderos de los parámetros  $b_1$ ,  $b_2$ ,  $b_3$  y  $\sigma_u^2$ .

Comencemos recordando que

$$\mathbf{b} = \begin{pmatrix} \mu_Y - \Sigma_{YX} \Sigma_X^{-1} \mu_X \\ \Sigma_X^{-1} \Sigma'_{YX} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Y además

$$\sigma_u^2 = \sigma_Y^2 - \Sigma_{YX} \Sigma_X^{-1} \Sigma'_{YX}$$

Calculemos entonces los valores:

```
#Calculamos b_1
b_1 <- as.numeric(mu[1] - Sigma[1,2:3]%%solve(Sigma[-1,-1])%%mu[-1])

#Calculemos b_2 y b_3
b_2 <- (solve(Sigma[-1,-1])%%Sigma[1,2:3])[1]
b_3 <- (solve(Sigma[-1,-1])%%Sigma[1,2:3])[2]

#Armamos vector b
b <- t(cbind(b_1, b_2, b_3))
b
```

```
##           [,1]
## b_1 0.82352941
## b_2 0.47058824
## b_3 0.08823529
```



Es decir,  $b_1 = 0.8235294$ ,  $b_2 = 0.4705882$  y  $b_3 = 0.0882353$

Ahora calculemos la varianza de los errores

```
var_u <- Sigma[1,1] - Sigma[1,2:3]%%solve(Sigma[-1,-1])%%as.matrix(Sigma[1,2:3])
var_u
```

```
##           [,1]
## [1,] 0.6294118
```

Es decir,  $\sigma_u^2 = 0.6294118$

Podemos ver entonces que

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0.8235294 \\ 0.4705882 \\ 0.0882353 \end{pmatrix} \neq \begin{pmatrix} 0.0 \\ 0.5 \\ 1.0 \end{pmatrix}$$

Ademas, podemos ver que

$$\begin{pmatrix} b_1 + b_3 \\ b_1 - b_2 + b_3 \end{pmatrix} = \begin{pmatrix} 0.9117647 \\ 0.4411765 \end{pmatrix} \neq \begin{pmatrix} 1.0 \\ 0.5 \end{pmatrix}$$

Es decir, las pruebas de hipótesis concluyeron lo correcto en iii) y v).

6.  $\left\{ \begin{pmatrix} Y_i \\ \mathbf{X}_i \end{pmatrix} \right\}_{i=1}^{100}$  vectores aleatorios i.i.d. dimensión  $k$ .  $k < 100$

Se tiene el modelo de regresión lineal:

$$\mathbf{Y} = \mathbf{X}\mathbf{b} + \mathbf{U}, \mathbb{E}[\mathbf{U}|\mathbf{X}] = 0.$$

$$\mathbb{P}[\text{Rango}(\mathbf{X}) = k] = 1$$

$$\text{Var}(\mathbf{U}|\mathbf{X}) = \sigma_u^2 \mathbb{I}_{100} \text{ (homocedasticidad condicional).}$$

$$\mathbf{U}|\mathbf{X}=\mathbf{X} \sim N_n(\mathbf{0}, \sigma_u^2 \mathbb{I}_{100}) \quad \forall \mathbf{X},$$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & \mathbf{X}'_1 \\ 1 & \mathbf{X}'_2 \\ \vdots & \vdots \\ 1 & \mathbf{X}'_n \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1,k-1} \\ 1 & X_{21} & X_{22} & \dots & X_{2,k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{n,k-1} \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

Sabemos que  $\sigma_{u0}^2 = 0.16$

i) Para  $i \in 1, \dots, k$  proponer una prueba de la hipótesis (i.e. definir una regla de decisión para aceptarla o rechazarla)

$$\mathbf{H}_0 : b_i = b_i^0$$

contra la alternativa

$$\mathbf{H}_1 : b_i \neq b_i^0$$

con un nivel de significancia  $\alpha = 0.10$

Podemos utilizar el pivote siguiente (tomado de la tarea 4):

$$q_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i}{\sigma_u \sqrt{(\mathbf{X}'\mathbf{X})_{ii}^{-1}}} \sim N(0, 1)$$

Al contener los parámetros  $b_i$  y  $\sigma_u$ , el pivote no es un estadístico. No obstante, es posible transformarlo en un estadístico bajo  $\mathbf{H}_0 : b_i = b_i^0$  y utilizando el hecho de que conocemos  $\sigma_{u0}^2 = 0.16 \implies \sigma_{u0} = \sqrt{0.16} = 0.4$ . Es decir:

$$q_{\hat{b}_i} = \frac{\hat{b}_{iMCO} - b_i^0}{0.4 \sqrt{(\mathbf{X}'\mathbf{X})_{ii}^{-1}}} \sim N(0, 1)$$

Sabemos, por el teorema de Gauss-Markov ("GM"), que  $\hat{b}_{iMCO}$  es el mejor estimador (el más eficiente) lineal condicionalmente insesgado (BLCUE) de  $b_i$ ; es decir:  $\hat{b}_{iMCO} \approx b_i$ . De lo anterior, y de la hipótesis nula, tenemos las siguientes implicaciones:

$$\mathbf{H}_0 : b_i = b_i^0 \xrightarrow{GM} \hat{b}_{iMCO} \approx b_i = b_i^0 \implies \hat{b}_{iMCO} - b_i^0 \approx 0$$

Es decir, bajo la hipótesis nula, el estadístico  $q_{\hat{b}_i}$  es un número pequeño (positivo o negativo); si es lo suficientemente pequeño (en valor absoluto), no podemos rechazar la hipótesis nula. Es decir, no podemos

rechazar si, dado un nivel de significancia  $\alpha$ , tenemos que  $|q_{\hat{b}_i}| < q_{\alpha/2}$ , donde  $q_{\alpha/2}$  es el valor crítico de la prueba (al ser una prueba de dos colas).

Finalmente, bajo la hipótesis alternativa, tenemos las siguientes implicaciones:

$$\mathbf{H}_1 : b_i \neq b_i^0 \xrightarrow{GM} \hat{b}_{iMCO} \approx b_i \neq b_i^0 \implies \hat{b}_{iMCO} - b_i^0 \not\approx 0$$

Por lo que el estadístico de prueba,  $q_{\hat{b}_i}$ , sería aproximadamente distinto de cero (negativo o positivo). Entonces, requerimos que el estadístico calculado anteriormente  $q_{\hat{b}_i}$  sea un número suficientemente grande (en valor absoluto), mayor o igual al valor crítico  $q_{\alpha/2}$ , para poder rechazar la hipótesis nula.

Ahora bien, con  $\alpha = 10\%$ , podemos calcular el valor crítico  $q_{\alpha/2}$

```
alpha <- 0.1
q_bhat <- qnorm(p = alpha/2, lower.tail = F)
```

Es decir,  $q_{\alpha/2} = 1.6448536$

En conclusión, la **regla de decisión** es tal que:

$$\text{Si } |q_{\hat{b}_i}| = \left| \frac{\hat{b}_{iMCO} - b_i^0}{0.4\sqrt{(\mathbf{X}'\mathbf{X})_{ii}^{-1}}} \right| < 1.6448536, \text{ no rechazamos } \mathbf{H}_0 : b_i = b_i^0.$$

$$\text{Si } |q_{\hat{b}_i}| = \left| \frac{\hat{b}_{iMCO} - b_i^0}{0.4\sqrt{(\mathbf{X}'\mathbf{X})_{ii}^{-1}}} \right| \geq 1.6448536, \text{ rechazamos } \mathbf{H}_0 : b_i = b_i^0 \text{ en favor de } \mathbf{H}_1 : b_i \neq b_i^0.$$

## ii) Proponer una prueba de la hipótesis

$$\mathbf{H}_0 : \mathbf{c}'\mathbf{b}_0 = d$$

contra la alternativa

$$\mathbf{H}_1 : \mathbf{c}'\mathbf{b}_0 \neq d$$

donde  $\mathbf{c}$  es un vector en  $\mathbb{R}^k$ , con nivel de significancia  $\alpha = 0.10$

Podemos utilizar el pivote siguiente (tomado de la tarea 4):

$$q_{\hat{b}} = \frac{\mathbf{c}'(\hat{\mathbf{b}}_{MCO} - \mathbf{b}_0)}{\sigma_u \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} = \frac{\mathbf{c}'\hat{\mathbf{b}}_{MCO} - \mathbf{c}'\mathbf{b}_0}{\sigma_u \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim N(0, 1)$$

Al contener los parámetros  $\mathbf{b}_0$  y  $\sigma_u$ , el pivote no es un estadístico. No obstante, es posible transformarlo en un estadístico bajo  $\mathbf{H}_0 : \mathbf{c}'\mathbf{b}_0 = d$  y utilizando el hecho de que conocemos  $\sigma_{u0}^2 = 0.16 \implies \sigma_{u0} = \sqrt{0.16} = 0.4$ . Es decir:

$$q_{\hat{b}} = \frac{\mathbf{c}'\hat{\mathbf{b}}_{MCO} - d}{0.4\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim N(0, 1)$$

Sabemos, por el teorema de Gauss-Markov (“GM”), que  $\hat{\mathbf{b}}_{MCO}$  es el mejor estimador (el más eficiente) lineal condicionalmente insesgado (*BLCUE*) de  $\mathbf{b}_0$ ; es decir:  $\hat{\mathbf{b}}_{MCO} \approx \mathbf{b}_0$ . De lo anterior, y de la hipótesis nula, tenemos las siguientes implicaciones:

$$\mathbf{H}_0 : \mathbf{c}'\mathbf{b}_0 = d \xrightarrow{GM} \mathbf{c}'\hat{\mathbf{b}}_{MCO} \approx \mathbf{c}'\mathbf{b}_0 = d \implies \mathbf{c}'\hat{\mathbf{b}}_{MCO} - d \approx 0$$

Es decir, bajo la hipótesis nula, el estadístico  $q_{\hat{b}}$  es un número pequeño (positivo o negativo); si es lo suficientemente pequeño (en valor absoluto), no podemos rechazar la hipótesis nula. Es decir, no podemos rechazar si, dado un nivel de significancia  $\alpha$ , tenemos que  $|q_{\hat{b}}| < q_{\alpha/2}$ , donde  $q_{\alpha/2}$  es el valor crítico de la prueba al nivel de significancia  $\alpha$  (al ser una prueba de dos colas).

Finalmente, bajo la hipótesis alternativa, tenemos las siguientes implicaciones:

$$\mathbf{H}_1 : \mathbf{c}'\mathbf{b}_0 \neq d \xrightarrow{GM} \mathbf{c}'\hat{\mathbf{b}}_{MCO} \approx \mathbf{c}'\mathbf{b}_0 \neq d \implies \mathbf{c}'\hat{\mathbf{b}}_{MCO} - d \not\approx 0$$

Por lo que el estadístico de prueba,  $q_{\hat{b}}$ , sería aproximadamente distinto de cero (negativo o positivo). Entonces, requerimos que el estadístico calculado anteriormente  $q_{\hat{b}}$  sea un número suficientemente grande (en valor absoluto), mayor o igual al valor crítico  $q_{\alpha/2}$ , para poder rechazar la hipótesis nula.

Ya hemos calculado el valor crítico de la prueba con un nivel de significancia del 10%. Es decir,

$$q_{\alpha/2} = 1.6448536$$

En conclusión, la regla de decisión es tal que:

$$\text{Si } |q_{\hat{b}}| = \left| \frac{\mathbf{c}'\hat{\mathbf{b}}_{MCO} - d}{0.4\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \right| < 1.6448536, \text{ no rechazamos } \mathbf{H}_0 : \mathbf{c}'\mathbf{b}_0 = d.$$

$$\text{Si } |q_{\hat{b}}| = \left| \frac{\mathbf{c}'\hat{\mathbf{b}}_{MCO} - d}{0.4\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \right| \geq 1.6448536, \text{ rechazamos } \mathbf{H}_0 : \mathbf{c}'\mathbf{b}_0 = d \text{ en favor de } \mathbf{H}_1 : \mathbf{c}'\mathbf{b}_0 \neq d.$$

### iii) Proponer una prueba de la hipótesis

$$\mathbf{H}_0 : \mathbf{R}\mathbf{b}_0 = \mathbf{r}$$

contra la alternativa

$$\mathbf{H}_1 : \mathbf{R}\mathbf{b}_0 \neq \mathbf{r}$$

donde  $\mathbf{R}$  es una matriz  $q \times k$  y  $\mathbf{r} \in \mathbb{R}^q$ , con nivel de significancia  $\alpha = 0.10$

Podemos utilizar el pivote siguiente (tomado de la tarea 4):

$$\tau_{\hat{\mathbf{b}}} = (\hat{\mathbf{b}}_{MCO} - \mathbf{b}_0)' \mathbf{R}' \frac{1}{\sigma_u^2} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} \mathbf{R}(\hat{\mathbf{b}}_{MCO} - \mathbf{b}_0) \sim \chi_q^2$$

$$\implies \tau_{\hat{\mathbf{b}}} = (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{R}\mathbf{b}_0)' \frac{1}{\sigma_u^2} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{R}\mathbf{b}_0) \sim \chi_q^2$$

Al contener los parámetros  $\mathbf{b}_0$  y  $\sigma_u^2$ , el pivote no es un estadístico. No obstante, es posible transformarlo en un estadístico bajo  $\mathbf{H}_0 : \mathbf{R}\mathbf{b}_0 = \mathbf{r}$  y utilizando el hecho de que conocemos  $\sigma_{u0}^2 = 0.16$ . Es decir:

$$\tau_{\hat{\mathbf{b}}} = (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})' \frac{1}{0.16} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r}) \sim \chi_q^2$$

Sabemos, por el teorema de Gauss-Markov ("GM"), que  $\hat{\mathbf{b}}_{MCO}$  es el mejor estimador (el más eficiente) lineal condicionalmente insesgado (*BLCUE*) de  $\mathbf{b}_0$ ; es decir:  $\hat{\mathbf{b}}_{MCO} \approx \mathbf{b}_0$ . De lo anterior, y de la hipótesis nula, tenemos las siguientes implicaciones:

$$\mathbf{H}_0 : \mathbf{R}\mathbf{b}_0 = \mathbf{r} \xrightarrow{GM} \mathbf{R}\hat{\mathbf{b}}_{MCO} \approx \mathbf{R}\mathbf{b}_0 = \mathbf{r} \implies \mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r} \approx 0$$

Es decir, bajo la hipótesis nula, el estadístico  $\tau_{\hat{\mathbf{b}}}$  es un número pequeño y positivo; si es lo suficientemente pequeño, no podemos rechazar la hipótesis nula. Es decir, no podemos rechazar si, dado un nivel de significancia  $\alpha$ , tenemos que  $\tau_{\hat{\mathbf{b}}} < \chi_{q,\alpha}^2$ , donde  $\chi_{q,\alpha}^2$  es el valor crítico con un nivel de significancia  $\alpha$ .

Finalmente, bajo la hipótesis alternativa, tenemos las siguientes implicaciones:

$$\mathbf{H}_1 : \mathbf{R}\mathbf{b}_0 \neq \mathbf{r} \xrightarrow{GM} \mathbf{R}\hat{\mathbf{b}}_{MCO} \approx \mathbf{R}\mathbf{b}_0 \neq \mathbf{r} \implies \mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r} \not\approx 0$$

Por lo que el estadístico de prueba,  $q_{\hat{\mathbf{b}}}$ , sería aproximadamente distinto de cero y positivo. Entonces, requerimos que el estadístico calculado anteriormente  $q_{\hat{\mathbf{b}}}$  sea un número suficientemente grande, mayor o igual al valor crítico  $\chi_{q,\alpha}^2$ , para poder rechazar la hipótesis nula.

Podemos calcular el valor crítico si conocemos  $q$ . Por el momento dado que es desconocido en este ejemplo, concluimos que la **regla de decisión** es tal que:

Si  $\tau_{\hat{\mathbf{b}}} = (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})' \frac{1}{0.16} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r}) < \chi_{q,\alpha}^2$ , no rechazamos  $\mathbf{H}_0 : \mathbf{R}\mathbf{b}_0 = \mathbf{r}$ .

Si  $\tau_{\hat{\mathbf{b}}} = (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r})' \frac{1}{0.16} [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} (\mathbf{R}\hat{\mathbf{b}}_{MCO} - \mathbf{r}) \geq \chi_{q,\alpha}^2$ , rechazamos  $\mathbf{H}_0 : \mathbf{R}\mathbf{b}_0 = \mathbf{r}$  en favor de  $\mathbf{H}_1 : \mathbf{R}\mathbf{b}_0 \neq \mathbf{r}$ .

Los siguientes problemas están en Wooldridge, J.M. *Introductory Econometrics: A Modern Approach. 6th Edition.* South-Western. CENGAGE Learning. Las bases de datos se encuentran en [www.cengagebrain.com](http://www.cengagebrain.com)

## 7. Problema C6, capítulo 4, pág. 147.

Use the data in WAGE2 for this exercise.

i) Consider the standard wage equation

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u.$$

State the null hypothesis that another year of general workforce experience has the same effect on  $\log(wage)$  as another year of tenure with the current employer.

Null hypothesis:

$$\mathbf{H}_0 : \beta_2 = \beta_3$$

Against the alternative hypothesis:

$$\mathbf{H}_1 : \beta_2 \neq \beta_3$$

ii) Test the null hypothesis in part (i) against a two-sided alternative, at the 5% significance level, by constructing a 95% confidence interval. What do you conclude?

Null hypothesis can be written as follows:

$$\begin{aligned} \mathbf{H}_0 : \beta_2 = \beta_3 &\implies \beta_2 - \beta_3 = 0 \\ &\implies (0 \quad 0 \quad 1 \quad -1) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = 0 \\ &\implies \mathbf{c}'\mathbf{b} = 0 \end{aligned}$$

Next, we can use the following pivot to construct the confidence interval:

$$\frac{\mathbf{c}'\hat{\mathbf{b}}_{OLS} - \mathbf{c}'\mathbf{b}}{\hat{\sigma}_{uOLS} \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim t_{n-k}$$

We've shown in *tarea 15.3.i* that

$$\mathbb{P} \left[ \mathbf{c}'\hat{\mathbf{b}}_{OLS} - t_{\alpha/2} \hat{\sigma}_{uOLS} \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} < \mathbf{c}'\mathbf{b} < \mathbf{c}'\hat{\mathbf{b}}_{OLS} + t_{\alpha/2} \hat{\sigma}_{uOLS} \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \right] = 1 - \alpha$$

Which implies that the confidence interval (CI) is

$$CI = \left( \mathbf{c}'\hat{\mathbf{b}}_{OLS} - t_{\alpha/2} \hat{\sigma}_{uOLS} \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \quad , \quad \mathbf{c}'\hat{\mathbf{b}}_{OLS} + t_{\alpha/2} \hat{\sigma}_{uOLS} \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \right)$$

In order to calculate the CI, we need to compute all of its elements first.

Start by pulling the data

```
rm(list = ls())
library(wooldridge)

Z <- wage2[, c("lwage", "educ", "exper", "tenure")]
Z <- as.matrix(Z)

X <- cbind(1, Z[, -1])
n <- nrow(Z)
k <- ncol(Z)
```

Given that  $\alpha = 5\%$ , we can compute  $t_{\alpha/2}$  as

```
alpha <- 0.05
t_q <- qt(p = alpha/2, df = n-k, lower.tail = F)
t_q
```

```
## [1] 1.962515
```

Obtain the OLS estimators

```
XX_inv <- solve(t(X)%*%X)
b_OLS <- XX_inv%*%t(X)%*%Z[,1]
rownames(b_OLS) <- c("beta_0", "beta_1",
                    "beta_2", "beta_3")
b_OLS
```

```
##           [,1]
## beta_0 5.49669575
## beta_1 0.07486377
## beta_2 0.01532847
## beta_3 0.01337480
```

Calculate  $\hat{\sigma}_{uOLS}^2$

```
U_OLS <- Z[,1] - X%*%b_OLS
var.u_OLS <- 1/(n-k)*t(U_OLS)%*%U_OLS
sd.u_OLS <- sqrt(var.u_OLS)
sd.u_OLS
```

```
##           [,1]
## [1,] 0.3877293
```

Define  $c$

```
c <- matrix(c(0,0,1,-1))
c
```

```
##           [,1]
## [1,] 0
## [2,] 0
## [3,] 1
## [4,] -1
```

Therefore, calculate the CI as follows

```
c.XX_inv.c <- t(c)%*%XX_inv%*%c

CI <- c(
  t(c)%*%b_OLS - t_q*sd.u_OLS*sqrt(c.XX_inv.c),
  t(c)%*%b_OLS + t_q*sd.u_OLS*sqrt(c.XX_inv.c)
)
CI
```

```
## [1] -0.00735536 0.01126271
```

That is

$$\mathbb{P}\left[-0.0073554 < \mathbf{c}'\mathbf{b} < 0.0112627\right] = 95$$

And so the CI is

$$CI = \left(-0.0073554, 0.0112627\right)$$

It is clear that  $0 \in CI$ ; that is to say, since 0 is included in the confidence interval for  $\mathbf{c}'\mathbf{b} = \beta_2 - \beta_3$  with a significance level of 5%, then we can conclude that  $\mathbf{H}_0 : \beta_2 - \beta_3 = 0$  can't be rejected.



## 8. Problema C8, capítulo 4, pág. 147.

The data set 401KSUBS contains information on net financial wealth (*nettfa*), age of the survey respondent (*age*), annual family income (*inc*), family size (*fsize*), and participation in certain pension plans for people in the United States. The wealth and income variables are both recorded in thousands of dollars. For this question, use only the data for single-person households (so *fsize* = 1).

---

i) How many single-person households are there in the data set?

```
library(dplyr)
rm(list = ls())

single.person.size <- k401ksubs %>%
  filter(fsize==1) %>%
  nrow()
single.person.size
```

```
## [1] 2017
```

That is, there are 2017 single-person households in the data set.

ii) Use OLS to estimate the model

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 age + u$$

and report the results using the usual format. Be sure to use only the single-person households in the sample. Interpret the slope coefficients. Are there any surprises in the slope estimates?

```
# variables
Z <- k401ksubs %>%
  filter(fsize==1) %>%
  select(nettfa, inc, age)
Z <- as.matrix(Z)
X <- cbind(1, Z[, -1])
n <- nrow(Z)
k <- ncol(Z)

#OLS (use lm for convenience)
model <- summary(lm(nettfa ~ inc + age,
  data = as.data.frame(Z)))
b_OLS <- model$coefficients[,1]
b_std.errors <- model$coefficients[,2]
b_t.values <- model$coefficients[,3]
b_p.values <- model$coefficients[,4]
r.sq <- model$r.squared
```

Results:

$$\widehat{nettfa} = \underset{(4.0803931)}{-43.0398119} + \underset{(0.0597307)}{0.7993167 inc} + \underset{(0.0920169)}{0.8426563 age}$$

$$n = 2017, R^2 = 0.1193432$$

- If annual family income increases by 1 unit (\$1,000 USD), then, in average, net financial wealth of the individual would increase by about 0.799 units (\$799 USD).
- If the age of the respondent increases by 1 year, then, in average, net financial wealth of the individual would increase by about 0.842 units (\$842 USD).

It surprises me the size of  $\beta_1$ . It seems to imply that much of the increase in the individual's family income is actually absorbed by her net financial wealth. On the other hand, it is rather small the size of  $\beta_2$ ; a full year of life increases wealth by just \$842 dollars. Surely the average individual in the data set isn't part of the top 1%.

iii) Does the intercept from the regression in part (ii) have an interesting meaning? Explain.

It does indeed; it implies that a newborn with no family income ( $age = 0$  and  $inc = 0$ ) has, in average, a net wealth of MINUS 43 units; that is to say, a debt of \$43,039 USD. The average person in this data set sure owes much to the very existence.

iv) Find the  $p$ -value for the test  $H_0 : \beta_2 = 1$  against  $H_1 : \beta_2 < 1$ . Do you reject  $H_0$  at the 1% significance level?

From 2.ii of this *tarea*, we know that the decision rule for the hypothesis is:

If  $\tau_{\hat{\beta}_2} = \frac{\hat{\beta}_{2OLS} - 1}{\hat{\sigma}_{uOLS} \sqrt{(X'X)^{-1}_{33}}} > -t_\alpha$ , we can't reject  $H_0 : \beta_2 = 1$ .

If  $\tau_{\hat{\beta}_2} = \frac{\hat{\beta}_{2OLS} - 1}{\hat{\sigma}_{uOLS} \sqrt{(X'X)^{-1}_{33}}} \leq -t_\alpha$ , reject  $H_0 : \beta_2 = 1$  in favor of  $H_1 : \beta_2 < 1$ .

To calculate the  $p$ -value, we just need to compute the area under the curve associated with  $\tau_{\hat{\beta}_2}$ . Then, to find out whether we reject the null hypothesis at the 1% significance level, we solely need to follow the decision rule with  $\alpha = 1\%$ .

First, compute  $\tau_{\hat{\beta}_2}$

```
XX_inv <- solve(t(X)%*%X)

U_OLS <- Z[,1] - X%*%b_OLS

var.u_OLS <- 1/(n-k)*t(U_OLS)%*%U_OLS
sd.u_OLS <- sqrt(var.u_OLS)

t_beta2 <- (b_OLS[3] - 1) / (sd.u_OLS*sqrt(XX_inv[3,3]))
t_beta2

##           [,1]
## [1,] -1.709944
```

Following Wooldridge, another method to calculate the statistic is:

$$\frac{\hat{\beta}_{2MCO}}{\sqrt{\text{var}(\hat{\beta}_{2MCO}|\mathbf{X})}} \sim t_{n-k}$$

That is:

```
var.b_OLS <- as.numeric(var.u_OLS)*XX_inv
sd.b_OLS <- sqrt(diag(var.b_OLS))
t_woold <- (b_OLS[3] - 1) / sd.b_OLS[3]
names(t_woold) <- "t_beta"
t_woold
```

```
##      t_beta
## -1.709944
```

Hence

$$\tau_{\hat{\beta}_2} = \frac{\hat{\beta}_{2OLS} - 1}{\hat{\sigma}_{uOLS} \sqrt{(X'X)^{-1}_{33}}} = \frac{0.8426563 - 1}{44.6827505 * 0.0020593} = -1.7099439$$

Now, for the  $p$ -value, we can compute it as follows

```
pval <- pt(t_beta2, df = n-k)
pval
```

```
##           [,1]
## [1,] 0.04371514
```

That is,  $\mathbb{P}[t \leq \tau_{\hat{\beta}_2}] = 0.0437151$ ; which is our  $p$ -value.

Now, can we reject the null at 1% significance level? In principle, NO, since we've seen that the  $p$ -value (the minimum significance level at which we reject the null) is 4.3715139% > 1%.

If we follow the decision rule, we should conclude the same. First, calculate  $-t_\alpha$  with  $\alpha = 1\%$

```
t_1pct <- qt(0.01, df = n-k)
t_1pct
```

```
## [1] -2.328201
```

We see that  $\tau_{\hat{\beta}_2} = -1.7099439 > -2.328209 = -t_{1\%}$ . In conclusion, we can't reject at  $\alpha = 1\%$ , but at  $\alpha = 4.3715139\%$

v) If you do a simple regression of *nettfa* on *inc*, is the estimated coefficient on *inc* much different from the estimate in part (ii)? Why or why not?

```
new_model <- summary(lm(nettfa ~ inc, data = as.data.frame(Z)))
b_OLS_new <- new_model$coefficients[,1]
b_std.errors_new <- new_model$coefficients[,2]
r_sq_new <- new_model$r.squared
```

Results:

$$\widehat{nettfa} = -10.5709522 + 0.8206815 inc$$

$(2.0606775) \qquad (0.0609)$

$$n = 2017, R^2 = 0.082673$$

The estimated coefficient is in fact not much different from the estimate in part (ii). This is a consequence and a cause of our assumption  $\mathbb{E}[inc|u] = 0$ , where  $u$  contains *age* in this model. That is, correlation between *inc* and *age* should be zero or close:

```
cor(Z)
```

```
##          nettfa          inc          age
## nettfa 1.0000000 0.28752914 0.20257906
## inc    0.2875291 1.00000000 0.03905864
## age    0.2025791 0.03905864 1.00000000
```

We see that, in fact, 0.0390586 is close to zero. Also, following Wooldridge, we know that  $\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}_1$ , where  $\hat{\beta}_i$  stands for the estimators from part (ii), and  $\tilde{\beta}_i$  is the same but this part;  $\tilde{\delta}_1$  stands for the slope of the regression of *age* on *inc*. Since  $\hat{\beta}_1 \approx \tilde{\beta}_1$ , we would expect  $\tilde{\delta}_1$  to be small.

```
test_model <- lm(age ~ inc, data = as.data.frame(Z))
delta_1 <- test_model$coefficients[2]
as.numeric(delta_1)
```

```
## [1] 0.02535407
```

So we confirm  $\tilde{\delta}_1$  is small (as expected due to size of the correlation between the 2 variables).

Also

```
b_OLS_new[2]
```

```
##          inc
## 0.8206815
```

```
b_OLS[2] + b_OLS[3]*delta_1
```

```
##          inc
## 0.8206815
```