

1. En este ejercicio, se ilustra que bajo ciertos supuestos el estimador de mínimos cuadrados ordinarios de \mathbf{b} es *incondicionalmente insesgado*

i) Generar (usando algún paquete econométrico) una muestra i.i.d.:

$$\left\{ \begin{pmatrix} Y_i \\ X_{1i} \\ X_{2i} \end{pmatrix} \right\}_{i=1}^{100}$$

$$\text{donde } \begin{pmatrix} Y_i \\ X_{1i} \\ X_{2i} \end{pmatrix} \sim N_3 \left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0.8 & 0.4 & -0.2 \\ 0.4 & 1.0 & -0.8 \\ -0.2 & -0.8 & 2.0 \end{pmatrix} \right)$$

```
library(mvtnorm)
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", "X1", "X2")
```

ii) Con los datos obtenidos en i), estimar por mínimos cuadrados ordinarios el modelo de regresión lineal:

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + U_i$$

$$\text{con } i = 1, 2, \dots, 100. \text{ Obteniendo así el estimador } \hat{b}_{MCO} = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \end{pmatrix}$$

```
linear_regression <- lm(Y~X1 + X2, data = Z)
b_hat_mco <- linear_regression$coefficients
names(b_hat_mco) <- c("bhat0", "bhat1", "bhat2")
b_hat_mco
```

```
##      bhat0      bhat1      bhat2
## 0.96230529 0.43835976 0.01469546
```

iii) Con los parámetros dados en i), encontrar el valor verdadero de \mathbf{b} .

Podemos utilizar la siguiente fórmula, compuesta por los elementos dados en i):

$$\mathbf{b} = \begin{pmatrix} \mu_Y - \Sigma_{YX} \Sigma_X^{-1} \mu_X \\ \Sigma_X^{-1} (\Sigma_{YX})' \end{pmatrix}$$

```

#Calculamos b_0
mu_y <- 1
Sigma_YX <- matrix(c(0.4, -0.2),
                    nrow = 1)
Sigma_X <- matrix(c(1, -0.8, -0.8, 2),
                  byrow = T,
                  nrow = 2)
mu_x <- matrix(c(0, 2),
               ncol = 1)
b0 <- mu_y - Sigma_YX%%solve(Sigma_X)%%mu_x

#Calculamos b_1 y b_2
B <- solve(Sigma_X)%%t(Sigma_YX)

# Unimos y encontramos b
b <- matrix(c(b0,
              solve(Sigma_X)%%t(Sigma_YX)))
b

```

```

##           [,1]
## [1,] 0.82352941
## [2,] 0.47058824
## [3,] 0.08823529

```

```

bhats_DF <- data.frame(t(b_hat_mco))
for (i in 1:999) {

  Z <- data.frame(
    rmvnorm(
      n,
      mean = mu,
      sigma = Sigma)
    )
  names(Z) <- c("Y", "X1", "X2")

  linear_regression <- lm(Y~X1 + X2, data = Z)
  bhats_DF[i+1,] <- linear_regression$coefficients
}

head(bhats_DF)

```

iv) Repetir 999 veces los pasos i) y ii) para obtener 1000 estimadores $\{\hat{b}_{MCO_i}\}_{i=1}^{1000}$ de b

```

##      bhat0      bhat1      bhat2
## 1 0.9623053 0.4383598 0.01469546
## 2 0.8504011 0.4914479 0.08472991
## 3 0.9018382 0.3874991 0.05215762
## 4 0.7407816 0.5204219 0.09817845
## 5 1.0431266 0.4038970 -0.02806452
## 6 0.9470779 0.3662170 0.07251694

```

```
tail(bhats_DF)
```

```
##          bhat0      bhat1      bhat2
## 995  0.7918503  0.4056112  0.05071403
## 996  0.6257880  0.5108283  0.19001092
## 997  0.8315591  0.5018796  0.09048536
## 998  0.8751455  0.5678748  0.06819700
## 999  0.9194414  0.4888945  0.08809414
## 1000 0.7886932  0.5119912  0.10691558
```

v) ¿Es $\frac{1}{1000} \sum_{i=1}^{1000} \hat{b}_{MCO_i}$ muy cercano al valor verdadero b ?

Los valores reales son:

```
b <- data.frame(t(b))
names(b) <- c("b0", "b1", "b2")
print.data.frame(b, row.names = F)
```

```
##          b0          b1          b2
## 0.8235294 0.4705882 0.08823529
```

Mientras que la media de \hat{b}_{MCO_i} es:

```
colMeans(bhats_DF)
```

```
##          bhat0      bhat1      bhat2
## 0.82260820 0.47371262 0.08845258
```

Los valores son cercanos dado que los estimadores son insesgados.

vi) ¿Qué supuestos mínimos garantizan que el estimador \hat{b}_{MCO} sea insesgado?

Los supuestos 1 a 3 que vimos en clase:

- S1: Los vectores aleatorios \mathbf{Z}_i son iid
- S2: $Y = Xb + U$ y $E[U|X] = 0$, y
- S3: $X'X$ es invertible

2. En este ejercicio, se ilustra que bajo ciertos supuestos el estimador de mínimos cuadrados ordinarios de σ_u^2 es *incondicionalmente insesgado*

i) Generar (usando algún paquete econométrico) una muestra i.i.d.

$$\left\{ \begin{pmatrix} Y_i \\ X_{1i} \\ X_{2i} \end{pmatrix} \right\}_{i=1}^{100}$$

donde $\begin{pmatrix} Y_i \\ X_{1i} \\ X_{2i} \end{pmatrix} \sim N_3 \left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0.8 & 0.4 & -0.2 \\ 0.4 & 1.0 & -0.8 \\ -0.2 & -0.8 & 2.0 \end{pmatrix} \right)$

```
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", "X1", "X2")
```

ii) Con los datos obtenidos en i), estimar por mínimos cuadrados ordinarios el modelo de regresión lineal:

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + U_i$$

con $i = 1, 2, \dots, 100$. Obteniendo así el estimador $\hat{b}_{MCO} = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \end{pmatrix}$ y $\hat{\sigma}_u^2 = \frac{1}{n-k} \hat{U}' \hat{U}$

```
# estimador b_mco
linear_regression <- lm(Y~X1 + X2, data = Z)
b_hat_mco <- linear_regression$coefficients
names(b_hat_mco) <- c("bhat0", "bhat1", "bhat2")
b_hat_mco
```

```
##      bhat0      bhat1      bhat2
## 0.95226992 0.40564668 0.02699275
```

```
# estimador sigma_u
Y <- Z$Y
X <- cbind(rep(1,n),
           Z[,-1])
U <- Y - as.matrix(X)%*%as.matrix(b_hat_mco)
k <- 3 #dimensión de los vectores aleatorios Z_i
sigma_u_mco <- 1/(n-k)*(t(U)%*%U)
sigma_u_mco
```

```
##           [,1]
## [1,] 0.5886415
```

iii) Con los parámetros dados en i), encontrar el valor verdadero de σ_u^2

En clase se vio que si $\{U_j\}_{j=1}^n$ son independientes, entonces

$$\sigma_u^2 = \text{var}(U_j) = \sigma_Y^2 - \Sigma_{YX} \Sigma_X^{-1} (\Sigma_{YX})'$$

Dado que los 100 vectores aleatorios \mathbf{Z} que estimamos son i.i.d., podemos aplicar la fórmula

```
sigma_y <- 0.8
Sigma_YX <- matrix(c(0.4, -0.2),
                    nrow = 1)
Sigma_X <- matrix(c(1, -0.8, -0.8, 2),
                  byrow = T,
                  nrow = 2)

sigma_u <- sigma_y - Sigma_YX%%solve(Sigma_X)%%t(Sigma_YX)

sigma_u
```

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

```
for (i in 1:999) {

  Z <- data.frame(
    rmvnorm(
      n,
      mean = mu,
      sigma = Sigma)
    )
  names(Z) <- c("Y", "X1", "X2")
  Y <- Z$Y
  X <- cbind(rep(1,n),
             Z[,-1])

  linear_regression <- lm(Y~X1 + X2, data = Z)
  bhat <- linear_regression$coefficients

  U <- Y - as.matrix(X)%%as.matrix(bhat)
  k <- 3 #dimensión de los vectores aleatorios Z_i
  sigma_u_mco <- rbind(sigma_u_mco,
                      1/(1000-k)*(t(U)%%U))
}

head(sigma_u_mco)
```

iv) Repetir 999 veces los pasos i) y ii) para obtener 1000 estimadores $\{\hat{\sigma}_{uMCO_i}^2\}_{i=1}^{1000}$ de σ_u^2

```
##           [,1]
## [1,] 0.58864154
## [2,] 0.05273801
## [3,] 0.05302320
## [4,] 0.05674314
## [5,] 0.06821358
## [6,] 0.06123840
```

```
tail(sigma_u_mco)
```

```
##           [,1]
## [995,] 0.06741127
## [996,] 0.05782102
## [997,] 0.07726208
## [998,] 0.05315073
## [999,] 0.05998869
## [1000,] 0.06890939
```

v) ¿Es $\frac{1}{1000} \sum_{i=1}^{1000} \hat{\sigma}_{uMCO_i}^2$ muy cercano al valor verdadero σ_u^2 ?

El valor real de σ_u^2 es:

```
sigma_u
```

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

Mientras que el promedio de los estimadores del experimento es:

```
mean(sigma_u_mco)
```

```
## [1] 0.06199038
```

Son valores cercanos dado que los estimadores son insesgados.

vi) ¿Qué supuestos mínimos garantizan que el estimador $\hat{\sigma}_{uMCO}^2$ sea insesgado?

Los supuestos 1 a 3 que vimos en clase y que mencionamos en el ejercicio anterior, más el supuesto 4:

- S4: $\text{var}(U|X) = \sigma_u^2 \mathbb{I}_n$ (homocedasticidad condicional)

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 (como se indica en la página XV del prefacio del libro).

`library(wooldridge)`

3. Problema 1, capítulo 2, pág. 53

Let `kids` denote the number of children ever born to a woman, and let `educ` denote years of education for the woman. A simple model relating fertility to years of education is

$$kids = \beta_0 + \beta_1 educ + u$$

where u is the unobserved error.

- (i) What kinds of factors are contained in u ? Are these likely to be correlated with level of education?

Existen variables que podrían o no estar relacionadas con la variable independiente asociada a la fertilidad `kids` o con la variable dependiente asociada a los años de educación `educ`. Al ser el modelo muy simple, es decir, un modelo que intenta explicar una variable compleja, como la fertilidad, en función de una única variable de índole más social, es probable que existan más factores no contenidos en el modelo, pero sí en u , que estén correlacionados con el nivel de educación.

- (ii) Will a simple regression analysis uncover the *ceteris paribus* effect of education on fertility?

Explain. Es posible que este modelo capture el efecto *ceteris paribus* siempre y cuando se restrinja a que $\mathbb{E}[u|educ] = 0$, lo que implicaría que $\mathbb{E}[kids|educ] = \beta_0 + \beta_1 educ$. De otro modo, no sería posible capturar tal efecto.

4. Problema 4, capítulo 2, pág. 54

The data set **BWGHT** contains data on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces (**bwght**), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy (**cigs**). The following simple regression was estimated using data on $n=1,388$ births:

$$\widehat{bwght} = 119.77 - 0.514 * cigs$$

(i) What is the predicted birth weight when $cigs = 0$? What about when $cigs = 20$ (one pack per day)? Comment on the difference.

- Si $cigs = 0$, entonces la predicción para el peso al nacer es $\widehat{bwght}\big|_{cigs=0} = 119.77$
- Si $cigs = 20$, entonces la predicción para el peso al nacer es $\widehat{bwght}\big|_{cigs=20} = 119.77 - 0.514(20) = 119.77 - 10.28 = 109.49$

(ii) Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mother's smoking habits? Explain.

Puede existir causalidad entre ambas variables, pero solamente captura una correlación entre éstas, considerando los supuestos vistos en clase.

(iii) To predict a birth weight of 125 ounces, what would $cigs$ have to be? Comment.

Resolvamos:

$$125 = 119.77 - 0.514 * cigs \quad 0.514 * cigs = 119.77 - 125 \quad cigs = -\frac{5.23}{0.514} = -10.1751$$

Lo anterior implica que se necesitan fumar “menos” diez cigarrillos; algo que no tiene sentido.

(iv) The proportion of women in the sample who do not smoke while pregnant is about .85. Does this help reconcile your finding from part (iii)?

Afecta, debido a que el modelo de regresión utiliza datos donde no hay varianza en el 85% de los datos, es decir, donde $cigs=0$. Para analizar el efecto entre las variables del modelo, sería más apropiado estimar una regresión en ese 15% de mujeres embarazadas que sí fuman, sin utilizar un intercepto. No obstante, el valor calculado en (iii) seguramente sería menor (mayor en términos absolutos), lo que es indicativo de que se requieren más variables para explicar el peso del recién nacido, y así obtener resultados más intuitivos que una cantidad de cigarros negativa.

5. Problema 6, capítulo 2, pág. 54

Using data from 1988 for houses sold in Andover, Massachusetts, from Kiel and McClain (1995), the following equation relates housing price (**price**) to the distance from a recently built garbage incinerator (**dist**):

$$\widehat{\log(\text{price})} = 9.40 + 0.312 * \log(\text{dist})$$

$$n = 135, R^2 = 0.162$$

(i) Interpret the coefficient on $\log(\text{dist})$. Is the sign of this estimate what you expect it to be?

El coeficiente indica que el si la distancia de la casa a un incinerador construido recientemente se incrementa en un 1%, el valor de la casa incrementa un 32% en promedio. El signo del estimador coincide con lo que esperaríamos ver en la realidad, considerando que la demanda para casas ubicadas cerca de ambientes contaminantes es menor.

(ii) Do you think simple regression provides an unbiased estimator of the *ceteris paribus* elasticity of **price** with respect to **dist**? (Think about the city's decision on where to put the incinerator.)

Seguramente ninguna ciudad colocaría un incinerador cerca de una zona residencial. Mucho menos lo colocaría cerca de una zona residencial de alto valor, dado que sus pobladores tendrían quizás un mayor poder de negociación. Entonces, es posible que existan en u variables correlacionadas con dist que harían que no se cumpla que $\mathbb{E}[u|\text{dist}] = 0$; es decir, no sería insesgado el estimador.

(iii) What other factors about a house affect its price? Might these be correlated with distance from the incinerator?

Factores que pudieran estar correlacionados con la distancia y que a su vez afecte el precio, es el nivel de ingresos promedio en la zona residencial donde se ubica la casa. Pensaríamos que entre mayor sea este nivel de ingresos, mayor es el precio de la casa (solamente los ricos viven en mansiones). Además, una zona con nivel de ingresos alto puede tener influencia en las decisiones de la ciudad, como por ejemplo en decidir dónde ubicar los incineradores.

6. Problema 7, capítulo 2, pág. 54

Consider the savings function

$$sav = \beta_0 + \beta_1 inc + u, \quad u = \sqrt{inc} * e,$$

where e is a random variable with $\mathbb{E}[e] = 0$ and $Var(e) = \sigma_e^2$. Assume that e is independent of inc .

- (i) Show that $\mathbb{E}[u|inc] = 0$, so that the key zero conditional mean assumption (Assumption SLR.4) is satisfied. [Hint: If e is independent of inc , then $\mathbb{E}[e|inc] = \mathbb{E}[e]$.]

$$\begin{aligned} \mathbb{E}[u|inc] &= \mathbb{E}[\sqrt{inc} * e|inc] \\ &= \sqrt{inc} * \mathbb{E}[e|inc] \\ &= \sqrt{inc} * \mathbb{E}[e] \\ &= \sqrt{inc} * 0 \\ &= 0 \quad \blacksquare \end{aligned}$$

- (ii) Show that $Var[u|inc] = \sigma_e^2 inc$, so that the homoskedasticity Assumption SLR.5 is violated. In particular, the variance of sav increases with inc . [Hint: $Var[e|inc] = Var[e]$ if e and inc are independent.]

$$\begin{aligned} Var[u|inc] &= Var[\sqrt{inc} * e|inc] \\ &= (\sqrt{inc})^2 * Var[e|inc] \\ &= inc * Var[e] \\ &= inc * \sigma_e^2 \\ &= \sigma_e^2 inc \quad \blacksquare \end{aligned}$$

- (iii) Provide a discussion that supports the assumption that the variance of savings increases with family income.

La idea de que los ahorros se vuelven más volátiles a medida que el ingreso familiar aumenta, puede explicarse si consideramos que, en el caso de los hogares con ingresos bajos, existen restricciones que limitan el uso de sus ingresos para fines distintos al consumo (casi todo su ingreso lo gastan en ello), lo que implicaría un ahorro bajo y constante. En cambio, a medida que los ingresos se incrementan, existe un sobrante que puede destinarse al ahorro, o a otros fines como la inversión y el consumo suntuoso; esto es, existe una discrecionalidad mayor respecto al uso del ingreso sobrante, que puede ahorrarse o no, lo que finalmente resulta en una variación mayor del ahorro.

7. Problema C6, capítulo 2, pág. 57

We used the data in MEAP93 for Example 2.12. Now we want to explore the relationship between the math pass rate (`math10`) and spending per student (`expend`).

- (i) Do you think each additional dollar spent has the same effect on the pass rate, or does a diminishing effect seem more appropriate? Explain.

Tendría sentido esperar retornos decrecientes en las tasas de aprobación por cada dólar gastado. A medida que el gasto por alumno incrementa, puede ser que ese gasto extra se destine a fines que contribuyan menos al rendimiento educativo del alumno, a diferencia del gasto realizado al comienzo, cuando se destina a elementos básicos como pupitres, libros, servicios, etc.

- (ii) In the population model

$$\text{math10} = \beta_0 + \beta_1 \log(\text{expend}) + u,$$

argue that $\beta_1/10$ is the percentage point change in `math10` given a 10% increase in `expend`.

Es posible demostrar lo anterior con derivadas:

$$\begin{aligned} \frac{\partial \text{math10}}{\partial \text{expend}} &= \frac{\beta_1}{\text{expend}} \\ \Rightarrow \partial \text{math10} &= \frac{\beta_1}{\text{expend}} \partial \text{expend} = \beta_1 \frac{\partial \text{expend}}{\text{expend}} \end{aligned}$$

$$\therefore \Delta \text{math10} = \beta_1 \frac{\Delta \text{expend}}{\text{expend}} = \beta_1 \Delta \% \text{expend}$$

Entonces, con un incremento porcentual del 10%:

$$\Delta \text{math10} = \beta_1 * 10\% = \beta_1 * \frac{1}{10} = \frac{\beta_1}{10} \quad \blacksquare$$

- (iii) Use the data in MEAP93 to estimate the model from part (ii). Report the estimated equation in the usual way, including the sample size and *R*-squared.

```
model_ii <- lm(math10 ~ log(expend), data = meap93)
coefs <- model_ii$coefficients
results <- summary(model_ii)
results

##
## Call:
## lm(formula = math10 ~ log(expend), data = meap93)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -22.343  -7.100  -0.914   6.148  39.093
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -69.341     26.530  -2.614 0.009290 **
## log(expend)   11.164       3.169   3.523 0.000475 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.35 on 406 degrees of freedom
## Multiple R-squared:  0.02966,    Adjusted R-squared:  0.02727
## F-statistic: 12.41 on 1 and 406 DF,  p-value: 0.0004752
```

Entonces, el modelo estimado es

$$\widehat{\text{math10}} = -69.34 + 11.16 * \log(\text{expend})$$

El tamaño de la muestra es 408 y además, $R^2 = 0.0296633$

- (iv) How big is the estimated spending effect? Namely, if spending increases by 10%, what is the estimated percentage point increase in `math10`?

El efecto sería

$$\Delta \widehat{\text{math10}} \Big|_{\beta_1 = \hat{\beta}_1} = \frac{\hat{\beta}_1}{10} = \frac{11.16}{10} = 1.1164402$$

- (v) One might worry that regression analysis can produce fitted values for `math10` that are greater than 100. Why is this not much of a worry in this data set?

Analicemos lo que debería ocurrir con `expend` si el estimado para `math10` fuera mayor a 100:

$$\begin{aligned} \widehat{\text{math10}} > 100 &\implies -69.34 + 11.16 * \log(\text{expend}) > 100 \\ &\implies \log(\text{expend}) > \frac{100 + 69.34}{11.16} \\ &\implies \text{expend} > e^{\frac{100+69.34}{11.16}} \\ &\implies \text{expend} > 3.8668687 \times 10^6 \end{aligned}$$

Es decir, la variable de gasto por estudiante (`expend`) tendría que ser mayor a 3,866,869... y en esta base de datos, la más 7,419. Por lo tanto, no puede ocurrir que el estimado sea mayor a 100.

8. Problema C9, capítulo 2, pág. 58

Use the data in COUNTYMURDERS to answer this questions. Use only the data for 1996.

- (i) How many counties had zero murders in 1996? How many counties had at least one execution? What is the largest number of executions?

Número de condados con cero muertes en 1996: 1051

```
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

countymurders %>%
  filter(year == 1996) %>%
  filter(murders == 0) %>%
  distinct(countyid) %>%
  count()

##           n
## 1 1051
```

Número de condados con al menos una ejecución en 1996: 31

```
countymurders %>%
  filter(year == 1996) %>%
  filter(execs>0) %>%
  distinct(countyid) %>%
  count()

##           n
## 1 31
```

Número más grande de ejecuciones (en 1996): 3

```
countymurders %>%
  filter(year == 1996) %>%
  filter(execs == max(execs))
```

```
##   arrests countyid density popul perc1019 perc2029 percblack percmale
## 1      34      45019 302.7346 277721 12.68575 17.78944 36.26301 48.23906
##   rpcincmaint rpcpersinc rpcunemins year murders murdrate arrestrate statefips
## 1      253.924   14554.06    37.004 1996      32 1.152236    1.22425      45
##   countyfips execs   lpopul   execrate
## 1          19      3 12.53437 0.1080221
```

(ii) Estimate the equation

$$\text{murders} = \beta_0 + \beta_1 \text{execs} + u$$

by OLS and report the results in the usual way, including sample size and R -squared.

```
murders_96 <- countymurders %>%
  filter(year == 1996)
model_murders <- lm(murders ~ execs, data = murders_96)
summary(model_murders)

##
## Call:
## lm(formula = murders ~ execs, data = murders_96)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -149.12   -5.46    -4.46    -2.46  1338.99
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.4572     0.8348   6.537 7.79e-11 ***
## execs         58.5555     5.8333  10.038 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 38.89 on 2195 degrees of freedom
## Multiple R-squared:  0.04389,    Adjusted R-squared:  0.04346
## F-statistic: 100.8 on 1 and 2195 DF,  p-value: < 2.2e-16
```

(iii) Interpret the slope coefficient reported in part (ii). Does the estimated equation suggest a deterrent effect of capital punishment?

Aparentemente, en el año 1996, en promedio, cuando se realizaba una ejecución, ocurrieron aproximadamente 58-59 asesinatos. Esto parecería sugerir que la pena capital está altamente correlacionada con los asesinatos.

(iv) What is the smallest number of murders that can be predicted by the equation? What is the residual for a county with zero executions and zero murders?

Redondeando, el número más pequeño de asesinatos es 5. El residuo para un condado con cero ejecuciones y cero muertes, es precisamente -5.

(v) Explain why a simple regression analysis is not well suited for determining whether capital punishment has a deterrent effect on murders.

Sencillamente porque es difícil explicar un fenómeno tan complejo como los asesinatos con tan solo una variable explicativa, dado que es posible que se omita el efecto de muchas otras variables. Sugerir que la pena capital tiene un pernicioso sobre el número de asesinatos pasa por alto el hecho de que precisamente en condados donde ocurren muchos asesinatos, es que posiblemente se aplica la pena capital con mayor frecuencia; es decir, la causalidad podría ser al revés.