

Ejercicios wooldridge..

Juliana Galindo-Juliana Chisnes

2024-03-03

Table of Contents

Ejercicio C1	1
1.1	1
1.2	2
1.3	2
Ejercicio C2	3
2.1	3
2.2	4
2.3	4
2.4	5
2.5	5
2.6	5

Solución de ejercicios C1 y C2 del capítulo 2 del libro “Introductory econometrics: a modern approach” de Jeffrey M. Wooldridge.

```
library(modelsummary)
library(wooldridge)
library(tidyverse)
library(latex2exp)
```

Ejercicio C1

1.1

(i) What is the most likely sign for b_2 ?

Según los datos arrojados por R, y teniendo en cuenta que la variable faminc se refiere a los ingresos mensuales de las familias estudiadas; se puede interpretar que el signo para β_2 es positivo ya que existe una correlación entre los ingresos de la persona y las condiciones de salud en las que se gestará y nacerá el bebé

1.2

- (ii) Do you think `cigs` and `faminc` are likely to be correlated? Explain why the correlation might be positive or negative.

```
cor(bwght$cigs, bwght$faminc)
```

```
## [1] -0.1730449
```

Los resultados de correlación entre estas dos variables indican una relación negativa entre ellas. Cuando la correlación es negativa, implica que cuando una variable aumenta, la otra tiende a disminuir. En este caso, se sugiere tentativamente que a medida que el ingreso aumenta, el consumo de cigarrillos tiende a disminuir. Es fundamental recordar que la correlación no implica causalidad, y existen otros factores que pueden influir en la relación entre el consumo de cigarrillos y los ingresos familiares.

1.3

- (iii) Now, estimate the equation with and without `faminc`, using the data in `BWGHT`. Report the results in equation form, including the sample size and R-squared. Discuss your results, focusing on whether adding `faminc` substantially changes the estimated effect of `cigs` on `bwght`.

```
reg <- lm(bwght ~ cigs, data=bwght)
```

```
summary(reg)
```

```
##
## Call:
## lm(formula = bwght ~ cigs, data = bwght)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -96.772 -11.772   0.297  13.228 151.228
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 119.77190    0.57234  209.267  < 2e-16 ***
## cigs        -0.51377    0.09049   -5.678 1.66e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 20.13 on 1386 degrees of freedom
## Multiple R-squared:  0.02273,    Adjusted R-squared:  0.02202
## F-statistic: 32.24 on 1 and 1386 DF,  p-value: 1.662e-08
```

```
reg <- lm(bwght ~ cigs + faminc , data=bwght)
```

```
summary(reg)
```

```
##
## Call:
## lm(formula = bwght ~ cigs + faminc, data = bwght)
##
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -96.061 -11.543   0.638  13.126 150.083
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 116.97413    1.04898 111.512 < 2e-16 ***
## cigs        -0.46341    0.09158  -5.060 4.75e-07 ***
## faminc       0.09276    0.02919   3.178 0.00151 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 20.06 on 1385 degrees of freedom
## Multiple R-squared:  0.0298, Adjusted R-squared:  0.0284
## F-statistic: 21.27 on 2 and 1385 DF, p-value: 7.942e-10
```

Segun los resultados obtenidos desde la ecuación $\text{bwght} = \beta_0 + \beta_1 \text{cigs} + u$, se puede observar un coeficiente de valor -0.51 para la variable cigs ; lo que significa que en promedio, por cada cigarrillo adicional que la persona consuma, el peso en el nacimiento disminuirá en 0.51 onzas.

Por otro lado, cuando se añade la variable “faminc” a la ecuación:

$\text{bwght} = \beta_0 + \beta_1 \text{cigs} + \beta_2 \text{faminc} + u$, se puede evidenciar que hay una modificación en el efecto estimado de cigs sobre el peso corporal, siendo ahora el coeficiente ligeramente menor para la variable cigs (-0.46) y para la variable faminc (0.092). Esto sugiere que el efecto del consumo de cigarrillos en el peso corporal puede estar influenciado a su vez por el nivel de ingresos de la familia; según esto, por cada 0.092 unidades adicionales de ingreso en la familia, la madre tiene mas potestad para adquirir una mayor cantidad de cigarrillos y el consumo de estos a su vez hace que el peso disminuya en 0.46.

Siguiendo esta observación, se estima que al añadir la variable faminc a la ecuación, si hay una modificación en el efecto estimado del consumo de cigarrillos sobre el peso corporal; sin embargo el impacto general sigue siendo pequeño, lo que indicará que existen otros factores que pueden ser mas influyentes.

Ejercicio C2

2.1

(i) Write out the results in equation form.

```
view(hprice1)

modelo2 <- lm(formula = price ~ sqrft + bdrms, data = hprice1)
summary(modelo2)

##
## Call:
## lm(formula = price ~ sqrft + bdrms, data = hprice1)
##
## Residuals:
```

```
##      Min      1Q   Median      3Q      Max
## -127.627 -42.876  -7.051   32.589  229.003
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -19.31500    31.04662  -0.622   0.536
## sqrft        0.12844     0.01382   9.291 1.39e-14 ***
## bdrms        15.19819     9.48352   1.603   0.113
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 63.04 on 85 degrees of freedom
## Multiple R-squared:  0.6319, Adjusted R-squared:  0.6233
## F-statistic: 72.96 on 2 and 85 DF,  p-value: < 2.2e-16
```

A partir de los resultados obtenidos, se observa que, en promedio, por cada aumento de una unidad en el área en pies cuadrados, el precio de la vivienda aumenta en 0.128 miles de dólares. Aunque ambas variables pueden tener influencia, la significancia estadística, medida por los p-valores (1.39e-14 para `sqrft` y 0.113 para `bdrms`), sugiere que el precio se verá afectado principalmente por el tamaño del área en pies cuadrados de la vivienda, más que por la cantidad de dormitorios.

2.2

- (ii) What is the estimated increase in price for a house with one more bedroom, holding square footage constant?

```
intercepto <- -19.31500
coef_sqrft <- 0.12844
coef_bdrms <- 15.19819

precio2 <- coef_bdrms + 1
summary(precio2)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      16.2    16.2    16.2    16.2    16.2    16.2
```

Teniendo en cuenta que la medida de los pies cuadrados se mantiene, el aumento del valor de la vivienda si se añade una habitación adicional, será de 16.2 miles de dolares.

2.3

- (iii) What is the estimated increase in price for a house with an additional bedroom that is 140 square feet in size? Compare this to your answer in part (ii).

```
precio1 <- coef_sqrft*140
summary(precio1)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.1284  0.1284  0.1284  0.1284  0.1284  0.1284

habitacion_adicional <- precio1*140
summary(habitacion_adicional)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    17.98  17.98   17.98   17.98  17.98   17.98
```

Teniendo en cuenta que el valor del pie cuadrado se estima en 0.128 miles de dolares, si se añade una habitación más a la vivienda, y esta es de 140 pies cuadrados, el precio total de la vivienda aumenta en promedio 17.98 miles de dolares.

2.4

- (iv) What percentage of the variation in price is explained by square footage and number of bedrooms?

Teniendo en cuenta que el R-cuadrado ajustado se emplea para identificar el nivel de variabilidad de las variables estudiadas y que este proporciona una medida de la calidad de ajuste al considerar la cantidad de variables así como también los grados de libertad. Podemos afirmar que aproximadamente el 62.33% de la variación del precio de la vivienda se ve afectado por las variables pies cuadrados y número de dormitorios; lo que se traduce en que estas dos variables tienen un impacto significativo en el precio de las viviendas.

2.5

- (v) The first house in the sample has `sqrft` 2,438 and `bdrms` 4. Find the predicted selling price for this house from the OLS regression line.

```
sqrft_casa_uno <- 2438
bdrms_casa_uno <- 4

precio_estimado_casa_uno <- intercepto + coef_sqrft*sqrft_casa_uno +
coef_bdrms*bdrms_casa_uno
summary(precio_estimado_casa_uno)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    354.6  354.6   354.6   354.6  354.6   354.6
```

Para una casa de dimensión de 2,438 pies cuadrados y cuatro habitaciones, se estima que el precio de venta se fijará en 354.6 miles de dolares.

2.6

- (vi) The actual selling price of the first house in the sample was \$300,000 (so `price` 300). Find the residual for this house. Does it suggest that the buyer underpaid or overpaid for the house?

```
precio_venta_casa_uno <- 300000
residual_casa_uno <- precio_venta_casa_uno - precio_estimado_casa_uno
summary(residual_casa_uno)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   299645  299645  299645  299645  299645  299645
```

Aunque existe una diferencia entre los residuos y el precio de venta de la primera vivienda en la muestra, esta no es significativa. Esto implica que el comprador pagó un precio justo, sin sufrir un sobreprecio ni un descuento significativo.