Taller Econometría

**Geraldine Wilches Acevedo - Isabella Bernal Roldán**

2024-02-27

library(wooldridge)

## Warning: package 'wooldridge' was built under R version 4.3.2

**C1.**

**(i). What is the most likely sign for B2?**

Esta ecuación referida al peso al nacer en onzas utiliza las variables para B1 Y B2 de “cigs” refiriéndose a los cigarrillos fumados durante un día en el embarazo y “faminc” refiriéndose a los ingresos familiares en dólares, estas son consideradas variables independientes que se relacionan entre sí. Se podría interpretar que existe una relación positiva, cuando los ingresos de una familia aumentan (B2), mejor será el acceso al cuidado prenatal, es decir, el peso al nacer en onzas será más saludable.

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

Las variables “cigs” y “faminc” están correlacionadas negativamente, ya que los cigarrillos tienen un efecto negativo sobre el peso del bebé y el ingreso familiar tiene un efecto positivo al poder tener mayores cuidados prenatales. En otras palabras, los ingresos familiares aumentan el peso del bebé mientras que, los cigarrillos disminuyen su peso.

cor(bwght$cigs, bwght$faminc)

## [1] -0.1730449

**(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.**

### Con faminc:

uno <- lm(bwght ~ cigs + faminc, data = bwght)  
summary(uno)

##   
## 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

### Sin faminc:

dos <- lm(bwght ~ cigs, data = bwght)  
summary(dos)

##   
## 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

Por un lado, en la tabla sin “faminc”, al tener un aumento en el número de cigarrillos fumados por la madre durante su gestación, representa una disminución en promedio de 0.51377 en el peso al nacer del bebé.

Por otro lado, en la tabla con “faminc”, al tener un aumento en el número de cigarrillos fumados por la madre durante su gestación, representa una disminución en promedio de 0.41341 en el peso al nacer del bebé, pero al tener un aumento en los ingresos familiares, representa el aumento en promedio de 0.09276 en el peso del bebé.

El R² siempre aumentará, aunque añada o no una variable, podríamos darnos cuenta ya que, sin “faminc” se podría decir que el modelo explica en un 2.2% la varianza del peso del bebé en onzas. El R² para “faminc”, explica el modelo en un 2.9% la varianza del peso del bebé en onzas.

**C2.**

**(i). Write out the results in equation form.**

data("hprice1")

ec1 <- lm(price ~ sqrft + bdrms, data = hprice1)  
summary(ec1)

##   
## 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

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

El incremento en precio estimado para una casa con una habitación más, manteniendo constante la superficie en pies cuadrados (sqrft) es de 15.198 dólares.

**(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).**

El aumento en el precio de la vivienda con una habitación adicional y una superficie de 140 sqrft es de 33.179 dólares. Este precio será mayor al del (ii) debido a que allí se mantenía el mismo tamaño de superficie, aunque hubiese una habitación adicional.

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

El R² podría explicar las variaciones en el tamaño de la casa, sumado con las variaciones en el número de habitaciones que explican el precio de la casa en un 63.19%.

**(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.**

**(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?**

El valor residual da como resultado -54.600 dólares, por lo que el comprador pagó un precio demasiado bajo para el tamaño de la superficie y la cantidad de habitaciones en la misma.