1er Trabajo Práctico de Econometría

Mauro Ciani Juan Camilo Gutman

El presente trabajo se propuso blabla

Tabla de contenidos

0.1 Librerías

```
library(tidyverse)    #Para manejar bases de datos
library(ggplot2)      #Para graficar
library(modelsummary) #Mejores tablas de regresión
library(tinytable)    #Motor de creación de tablas
```

```
df1 <- readRDS("bases/eph_1abc.RDS")
# df2 <- readRDS("Bases/eph_1de.RDS")
```

Nos quedamos únicamente con los jefes de hogar, hombres, con edades entre 25 y 65 años, ocupados y asalariados. Habría que mencionar algo sobre la base resultante (y cuantos datos estamos descartando) y un mínimo análisis exploratorio.

0.2 Probando referencias

(Pradier et al., 2023) (Arel-Bundock, 2022)

Cras a molestie nulla. (Hemingway, 1952)

Siguiendo a (Koenker, 2005), encontramos que:

```
#PUNTO 1B
library(tidyverse)       # Para manejar bases de datos
library(ggplot2)         # Para graficar
library(modelsummary)    # Mejores tablas de regresión
library(tinytable)       # Motor de creación de tablas
library(sandwich)        # Robust Covariance Matrix Estimators
library(quantreg)        # Regresión por cuantiles
```

```
Cargando paquete requerido: SparseM
```

```
Adjuntando el paquete: 'SparseM'
```

```
The following object is masked from 'package:base':

    backsolve
```

```
eph1 <- readRDS("Bases/eph_1abc.RDS")

reg_b <- lm(logSal ~ educn + edad + I(edad^2) + est_civ + region, data = eph1)
b_coef_estim <- coef(reg_b)
b_edad_max_salario <- (-1*b_coef_estim["edad"])/(2*b_coef_estim["I(edad^2)"])
print (b_edad_max_salario)
```

```
   edad
49.2958
```

```
#Buscamos la varianza según el método delta
##Definimos a la función g de los parametros de interés como (-1*b_coef_estim ["edad"])/(2*b_
print("$$ \beta $$")
```

```
[1] "$$ \beta $$"
```

```
## _b[edad2] * 2 * edad + _b[edad] = 0
## edad = -(_b[edad]) / 2 _b[edad2]

b_dg_coef_edad <-(-1)/(2*2*b_coef_estim["I(edad^2)"])
print(b_dg_coef_edad)
```

```
I(edad^2)
 853.8809
```

```r
b_dg_coef_edad2 <-(b_coef_estim ["edad"])/(2*((b_coef_estim["I(edad^2)"])^2))
print(b_dg_coef_edad2)
```

```
  edad
168371
```

```r
b_gradiente <- c(b_dg_coef_edad, b_dg_coef_edad2)
print(b_gradiente)
```

```
  I(edad^2)        edad
   853.8809 168370.9547
```

```r
b_vcov_mat <- vcov(reg_b)
#view(b_vcov_mat)

b_vcov_mat_edad <- b_vcov_mat [3:4,3:4]
#view(b_vcov_mat_edad)

# Aplicar el método delta para obtener Varianza de g(beta1, beta2) aproximada por el método
b_var_delta <- t(b_gradiente) %*% b_vcov_mat_edad %*% b_gradiente
print(b_var_delta)
```

```
         [,1]
[1,] 67.58301
```

```r
#Armamos el intervalo de confianza utilizando la distribución normal estándar

b_alpha1 <- 0.1  # Nivel de significancia: 90% de confianza
b_alpha2 <- 0.05 # Nivel de significancia: 95% de confianza
b_alpha3 <- 0.01  # Nivel de significancia: 99% de confianza

# Estimación de g(theta) {NO}
#g_hat <- g(theta_hat) {NO}

b_alpha = b_alpha3
b_error_std <- qnorm(1 - b_alpha/ 2) * sqrt(b_var_delta)
b_int_conf <- c(b_edad_max_salario - b_error_std, b_edad_max_salario + b_error_std)
```

```
b_estad_z_crit <- qnorm(1 - b_alpha/ 2)
print(b_int_conf)
```

```
[1] 28.12019 70.47141
```

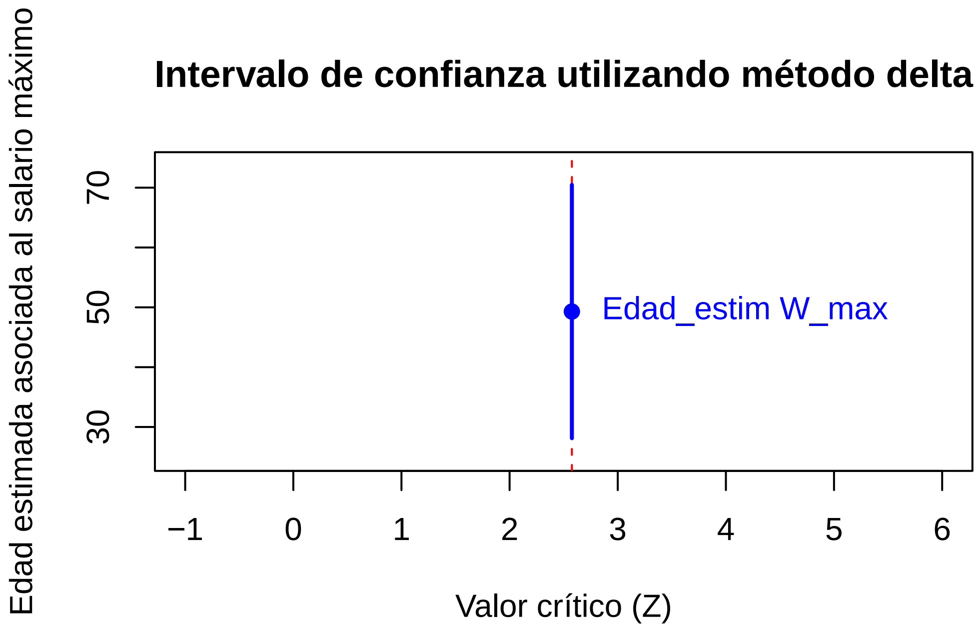
```
#Insertar gráfico

# Define los límites del gráfico
plot(c(-1, 6), c(b_edad_max_salario - 3 * sqrt(b_var_delta), b_edad_max_salario + 3 * sqrt(b_
     type = "n", xlab = "Valor crítico (Z)", ylab = "Edad estimada asociada al salario máximo
     main = "Intervalo de confianza utilizando método delta")

# Línea vertical en el valor crítico de la distribución normal estándar
abline(v = b_estad_z_crit, col = "red", lty = 2)

# Intervalo de confianza
segments(b_estad_z_crit, b_int_conf[1], b_estad_z_crit, b_int_conf[2], col = "blue", lwd = 2)

# Punto en g_hat
points(b_estad_z_crit, b_edad_max_salario, col = "blue", pch = 19)
text(b_estad_z_crit + 0.1, b_edad_max_salario, "Edad_estim W_max", pos = 4, col = "blue")
```



```
# "Para el int confianza creado a partir del metodo delta, aceptaríamos la Hip nula de que ed
# NO ME GUSTA COMO ESTA REDACTADO ESTO, es poco riguroso, hay que decirlo como que "esto sug

# Definimos la función de restricción no lineal
g <- function(coef) {
  beta1 <- coef["edad"]
  beta2 <- coef["I(edad {2)"]
  return((-1*beta1)/(2*(beta2)))
}

# Obtener las estimaciones de los coeficientes
b_betas_hat <- coef(reg_b)
b_beta1_hat <- b_betas_hat["edad"]
b_beta2_hat <- b_betas_hat["I(edad^2)"]

# Evaluar la función de restricción en las estimaciones
#restriction_value <- g(b_beta1_hat, b_beta2_hat)
b_hipotesis_nl <- g(b_betas_hat)

print(b_beta1_hat)
```

```
      edad
0.02886574
```

```
print(b_beta2_hat)
```

```
    I(edad^2)
-0.0002927809
```

```
##Armado de test de hipotesis simil 'testnl' en STATA
# transformation <- function(coef) { exp(coef["edad"]) / (1 +coef["I(edad^2)"]) }

# A MANOPLA ####
# Extraer los coeficientes
beta_hat <- coef(reg_b)

# Extraer la matriz de varianza-covarianza de los coeficientes
vcov_beta_hat <- vcov(reg_b)

# Definir la matriz de hipótesis para "2 * beta[I(edad^2)] * 50 + beta[edad] = 0"
```

```r
# Asumiendo que 'edad' es el tercer coeficiente y 'I(edad^2)' es el cuarto coeficiente
A <- matrix(c(0, 0, 1, 100, rep(0, length(beta_hat) - 4)), ncol = length(beta_hat))

# Calcular el estadístico de prueba Wald
W <- t(A %*% beta_hat) %*% solve(A %*% vcov_beta_hat %*% t(A)) %*% (A %*% beta_hat)

# El estadístico de prueba Wald sigue una distribución chi-cuadrado
# Obtener el valor p
p_value <- pchisq(W, df = 1, lower.tail = FALSE)

# Imprimir el estadístico de prueba y el valor p
print(W)
```

```
           [,1]
[1,] 0.08434709
```

```r
print(p_value)
```

```
          [,1]
[1,] 0.7714906
```

```r
# Crear un dataframe con el estadístico de prueba y el valor p
results_df <- data.frame(Estadistico_De_Prueba = W, pValor = p_value)

# Convertir el dataframe a una tabla tinytable
results_table <- tt(results_df)

# Mostrar la tabla
print(results_table)
```

```
+-----------------------+--------+
| Estadistico_De_Prueba | pValor |
+=======================+========+
| 0.0843                | 0.771  |
+-----------------------+--------+
```

(Pradier et al., 2023)

En la tabla pueden leerse los resultados de la primer regresión.

Tabla 2 — Resultados de la regresión, Ecuación de Mincer (estimación del efecto de la educación en el logaritmo de los salarios, controlando por...)

	(1)			
	Est.	p	2.5 %	97.5 %
(Intercept)	11.769***	<0.001	11.632	11.906
Edad	0.003***	0.007	0.001	0.005
Nivel Educativo				
Educn	0.063***	<0.001	0.058	0.068
Est civUnido	−0.115***	<0.001	−0.165	−0.066
Est civSeparado/Divorciado	−0.091**	0.020	−0.168	−0.014
Est civViudo	−0.013	0.892	−0.200	0.174
Est civSoltero	−0.257***	<0.001	−0.314	−0.200
regionNoroeste	−0.351***	<0.001	−0.417	−0.286
Estado Civil				
regionNoreste	−0.346***	<0.001	−0.429	−0.262
regionCuyo	−0.110***	0.006	−0.187	−0.032
regionPampeana	−0.056*	0.075	−0.117	0.006
regionPatagonia	0.299***	<0.001	0.227	0.371
Región				
Num.Obs.	3475			
R2	0.238			
R2 Adj.	0.235			
AIC	6123.3			
BIC	6203.3			
Bondad de Ajuste				

Continued on next page

Tabla 2 — Resultados de la regresión, Ecuación de Mincer (estimación del efecto de la educación en el logaritmo de los salarios, controlando por...) (Continued)

	(1)
Log.Lik.	−3048.641
F	98.208
RMSE	0.58

Notas: Acá comentarios, explicaciones, etc.
*Significant at the 10 percent level.

```
gof2 <- get_gof(reg2)
gof2 <- as.data.frame(t(gof2), optional = TRUE)
gof2$estad <- rownames(gof2)
gof2 <- gof2 %>% select(estad, everything())
names(gof2) <- NULL

gof1 <- get_gof(reg2)
gof1 <- as.data.frame(t(gof1), optional = TRUE)
gof1$estad <- rownames(gof1)
gof1 <- gof1 %>% select(estad, everything())
names(gof1) <- NULL
```

```
# gof1 <- gof1[4:nrow(gof1)-1,]
#
# gof1$orden <- c(3, 1, 2, 4, 5)
#
# gof1<- gof1 %>%
#   arrange(orden)
#
# gof1$orden <- NULL
#
```

```
# gof2 %>% tt()
```

```
gm1 <- gof_map
gm1[1,2] <- "n"
gm1[16:19,4] <- "TRUE" #Chau a los AIC y BIC
gm1[16:19,]
```

```
   raw clean fmt omit
16 AIC   AIC   1 TRUE
17 aic   AIC   1 TRUE
18 BIC   BIC   1 TRUE
19 bic   BIC   1 TRUE
```

```
gof1 <- get_gof(reg1, gof_map = gm1)
gof1 %>% tt()
```

aic	bic	r.squared	adj.r.squared	rmse	nobs	F	logLik
6174	6285	0.229	0.225	0.585	3475	64.1	-3069

```
# Anova(reg2) %>% tt()
```

testlineal.tex

texto de prueba

Bibliografía

Arel-Bundock, V. (2022). modelsummary: Data and Model Summaries in R. *Journal of Statistical Software*, *103*(1), 295-316. https://doi.org/10.1017/CBO9780511754098.011

Hemingway, E. (1952). *The Old Man and the Sea*. Charles Scribner's Sons.

Koenker, R. (2005). Quantile Regression in R: A Vignette. En *Quantile Regression* (pp. 295-316). Cambridge University Press.

Pradier, C., Weksler, G., Tiscornia, P., Shokida, N., Rosati, G., & Kozlowski, D. (2023). *ropensci/eph V1.0.0* (Versión 1.0.0) [Software]. Zenodo. https://doi.org/10.5281/zenodo.8352221