

AYUDANTÍA 2.

A. Feedback Quiz #1.

1. $R^2 \rightarrow$ Muchos consideraron que $R^2 = \frac{SEC}{SSR} = 0$ pero la definición es

$$R^2 = \frac{SEC}{STC} = 1 - \frac{SSR}{STC}$$

é. j

2. "Si el supuesto de media condicional no se cumple, los **residuos** \neq error"

Por construcción, cuando calculamos condiciones de primer orden, los residuos son ortogonales a las v. explicativas.
¡Suman cero y son ortogonales!

Clase N°2.

B. Repaso interpretación transformaciones funcionales

- Modelo simple
- Modelo cuadrático
- Modelo con variable cualitativa
- Modelo con interacciones
- Modelo $y - \log(x)$
- Modelo $\log(y) - x$
- Modelo $\log(y) - \log(x)$

AYUDANTIA N°2 - ECONOMETRÍA

Profesor: Juan Urquiza

Ayudante: Valentina Andrade (vandrade@uc.cl)

TEMA I

Considere el siguiente modelo de regresión lineal múltiple:

$$\log(wage) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{exper}^2 + u,$$

donde $\log(wage)$ representa el logaritmo natural del salario, educ los años de educación, y exper los años de experiencia laboral de la persona. Suponga que se cumplen todos los supuestos revisados en clase.

El modelo se estima mediante Mínimos Cuadrados Ordinarios (MCO) en Stata, obteniendo lo siguiente:

¡Revisemos el output!

Source	SS	df	MS	Number of obs	=	526
Model	44.5393702	3	14.8464567	Prob > F	=	0.0000
Residual	103.790392	522	.198832168	R²	=	
Total	148.329762	525	.28253288	Adj R-squared	=	0.2963

Y l wage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
β_1 educ	β_1 .007468	.007468	12.10	0.000	.0756948 .1050368
β_2 exper	β_2 .0410089	.0051965	7.89	0.000	.0308002 .0512175
β_3 expersq	β_3 -.0007136	.0001158	-6.16	0.000	-.000941 -.0004861
_cons	β_0 .1279975	.1059323	1.21	0.227	-.0801085 .3361034

99%
confianza
rechazo H_0
 \therefore v. relev.

donde $\text{expersq} = (\text{exper} \times \text{exper})$ y $\text{l wage} = \log(\text{wage})$.

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 - \dots$$

Además, sabemos que la matriz de varianzas y covarianzas correspondiente a la estimación anterior viene dada por:

	educ	exper	expersq	_cons
educ	.00005577			
exper	-2.793e-06	.000027		
expersq	1.422e-07	-5.759e-07	1.340e-08	
_cons	-.00072042	-.00015181	1.669e-06	.01122165

donde "1.422e(-07)" equivale a "1.422*10^(-7)".

- a. ¿Qué signo espera que tenga el parámetro poblacional β_1 en el modelo presentado? Usando la información de Stata, ¿se corresponde este signo con el de la estimación de MCO? Demuestre, y luego proporcione la interpretación de $\hat{\beta}_1$.

$\beta_1 \rightarrow$ efecto marginal que tienen los años de educación sobre el log(salario).

► $\hat{\beta}_1 > 0$

$$H_0: \beta_1 = 0$$

$$t = \frac{\hat{\beta}_1 - 0}{\text{se}(\hat{\beta}_1)}$$

$$t \cdot \text{se}(\hat{\beta}_1) = \hat{\beta}_1 = 0.09$$

$$\log(\text{wage}) = \beta_0 + 0.09 \cdot \text{educ} + \dots$$

$$\Delta Y = b \cdot 100\%$$

Interpretación: Por cada año de educación, el salario aumenta en un 9%, controlando por el resto de las variables del modelo.

- b. [Escriba la expresión del efecto parcial de la experiencia laboral sobre $\log(\text{wage})$] y luego escriba la hipótesis nula de que la experiencia tiene un efecto parcial constante sobre $\log(\text{wage})$. Usando la información de Stata, ¿se puede rechazar la hipótesis nula? Justifique su respuesta.

Debemos mirar todos los coeficientes que acompañan a la variable experiencia laboral [sea lineal, cuadrática, etc]

El efecto parcial de la experiencia sobre $\log(\text{wage})$ es

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{exper}^2 + u$$

$$\frac{\partial \log(\text{wage})}{\partial \text{exper}} = \beta_2 + 2\beta_3 \text{exper}$$

- Sobre la hipótesis, ¿qué significa que sea constante?

$$H_0: \beta_3 = 0 \quad H_1: \beta_3 \neq 0$$

$p < 0.00 \rightarrow$ Puedo rechazar con 99% de confianza H_0 .

∴ Es probable, con un 99% de confianza, que la experiencia no tenga un efecto constante.

- c. De acuerdo con las estimaciones de MCO, ¿cuál es la variación esperada en el salario al pasar de **5 a 6 años** de experiencia laboral? Use la matriz estimada de varianzas y covarianzas que se proporciona en el enunciado para evaluar su significancia estadística al 5%. No olvide especificar la distribución del estadístico de contraste, incluyendo los grados de libertad.

Experiencia tiene un doble efecto

$$\Delta = \hat{\beta}_2 + 2 \cdot \beta_3 \cdot \text{exper}$$

exper: 5

$$\Delta = (\hat{\beta}_2 + 2 \cdot \beta_3 \cdot 5) \cdot 100$$

pero los salarios están en log
 \Rightarrow Cuánto aumentan los salarios cuando aumentan en 1 año la experiencia

$$(0.041 + 2 \cdot (-0.0007) \cdot 5) \cdot 100$$

3.387%

• Nota: ¡partiendo de 5 años de experiencia!

Significancia

① $H_0 : \beta_2 + 2\beta_3 \cdot 5 = 0$ $H_1 : \beta_2 + 2\beta_3 \cdot 5 \neq 0$

②

$$\hat{t} = \frac{\beta_2 + 2\beta_3 \cdot 5}{\sqrt{V(\beta_2 + 2\beta_3 \cdot 5)}} = \frac{\beta_2 + 2\beta_3 \cdot 5}{\sqrt{V(\beta_2) + 2 \cdot 2 \cdot V(\beta_3) - 2 \cdot \text{Cov}(\beta_2, \beta_3)}}$$

$$= \frac{0.0339}{0.0041} \approx 8.26$$

③

$$t_{n-k-1}^{1-\frac{\alpha}{2}} = t_{326-3-1}^{1-\frac{0.05}{2}} = t_{322}^{0.975} \approx t_{1000}^{0.975} = 1.962$$

$$n = 326$$

$$K = 3$$

$$\alpha = 0.05$$

④ $\hat{t} = 8.26$

$$\hat{t} = 1.962$$

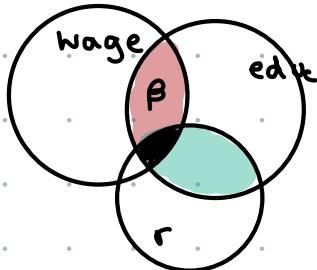
Con un 95% de confianza podemos rechazar H_0 , esto nos dice que existe evidencia de un efecto significativo y no lineal de la experiencia sobre salarios.

x_3

- d. Se está pensando en incluir la renta familiar como variable explicativa adicional. Se sabe que $\text{corr}(\text{renta familiar}, \text{educ}) > 0$. ¿Qué puede decir acerca de la varianza del estimador MCO de β_1 si se decide incluir la renta familiar? Discuta.

$$\text{Var}(\hat{\beta}_1 | x)$$

$$\text{Var}(\hat{\beta}_1 | x) = \frac{\hat{\sigma}^2}{\text{SCT}(1 - R^2)}$$



También podríamos decir que si la renta es una variable relevante, al omitirla su efecto pasa al término de error, entonces SSR crecería y por tanto también crecería la varianza de los errores $\hat{\sigma}^2$



Si la incorporo $\downarrow \text{SSR} \quad \downarrow \hat{\sigma}^2$

y en ese caso tengo dos efectos **

TEMA II

Usted obtiene una base de datos de 1,000 trabajadores en México. Cada uno de estos trabajadores trabaja en una de las siguientes tres industrias: **manufacturas**, **agricultura** o **servicios**. La base cuenta con las siguientes variables:

- **logsalario** = logaritmo natural del salario horario del individuo.
- **educ_anios** = años de educación del individuo.
- **Manufacturas** = v. binaria que toma valor 1 si el individuo trabaja en la industria manufacturera y 0 si trabaja en otro sector.
- **Agricultura** = v. binaria que toma valor 1 si trabaja en agricultura y 0 si trabaja en otro sector.
- **Servicios** = v. binaria que toma valor 1 si trabaja en sector servicios y 0 si trabaja en otro sector.
- **Mujer** = v. binaria que toma valor 1 si el individuo es mujer y 0 si es hombre.

Usted está interesado en caracterizar las variables que afectan el salario de las personas en México. Para ello, estima el siguiente modelo (M1) por Mínimos Cuadrados Ordinarios (MCO):

$$\log(\text{salario}_i) = \beta_0 + \beta_1 \text{educ}_i + \beta_2 \text{Manufacturas}_i + \beta_3 \text{Servicios}_i + \beta_4 \text{Mujer}_i + u_i,$$

obteniendo la siguiente salida de Stata:

Source	SS	df	MS	Number of obs	=	1,000
Model	187.330316	4	46.8325791	F(4, 995)	=	72.51
Residual	642.661487	995	.645890942	Prob > F	=	0.0000
Total	829.991803	999	.830822626	R-squared	=	0.2257
				Adj R-squared	=	0.2226
				Root MSE	=	.80367

logsalario	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ_anios	β_1 .0765472	.0059852	12.79	0.000	.0648022 .0882922
Manufacturas	β_2 .2583274	.0884845	2.92	0.004	.0846898 .4319649
Servicios	β_3 .4546783	.0738849	6.15	0.000	.3096902 .5996663
mujer	β_4 -.3319119	.0558438	-5.94	0.000	-.4414971 -.2223267
_cons	β_0 3.464946	.0690181	50.20	0.000	3.329508 3.600384

$$\beta_k \cdot 100 \%$$

$$\log Y = X$$

- a. Interprete los coeficientes que acompañan a las variables Manufacturas y Servicios, y luego discuta su significancia estadística individual.

Manufactura β_2

Entre los trabajadores de la manufactura y los de la agricultura existe una diferencia promedio en sus salarios de un 25, 83% ($\beta_2 \cdot 100\%$), *ceteris paribus*.

- $p = 0.004 \rightarrow$ Con un 99% de confianza, rechazamos H_0 , esto quiere decir que existe evidencia de la diferencia salarial promedio de los trabajadores de la manufactura respecto de los de la agricultura.

Servicios

Entre los trabajadores de el sector serv. y los de la agricultura existe una diferencia promedio en sus salarios de un 45, 5% ($\beta_2 \cdot 100\%$), *ceteris paribus*.

- b. En base a la tabla anterior, y controlando por los años de educación, ¿cuál es la diferencia salarial promedio entre una mujer que trabaja en Servicios y un hombre que trabaja en Manufacturas? Interprete su resultado.

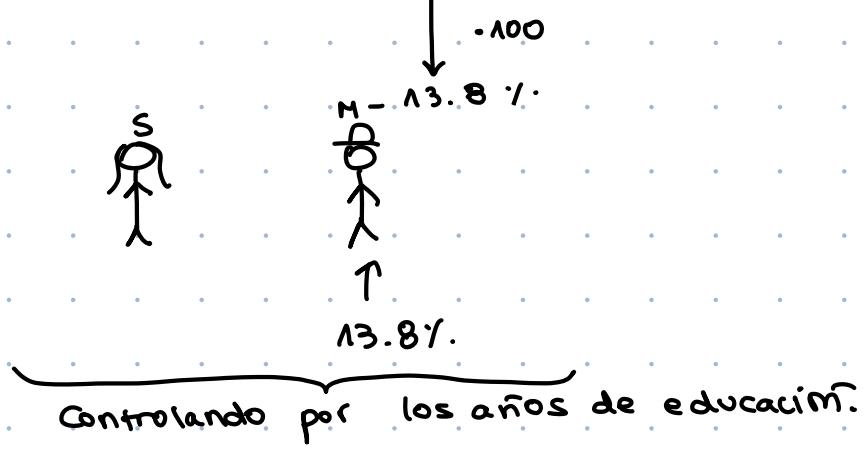
$$\hat{Y}_{MS} - \hat{Y}_{HM}$$

$$\log(\text{salario}_i) = \beta_0 + \beta_1 \text{educ}_i + \beta_2 \text{Manufacturas}_i + \beta_3 \text{Servicios}_i + \beta_4 \text{Mujer}_i + u_i$$

$$\log(wage_m) = \beta_0 + \beta_1 \text{educ} + \beta_3 \text{servicios} + \beta_4 \text{mujer}$$

$$\log(wage_H) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{Manufactura}$$

$$\begin{aligned} \log(wage_m) - \log(wage_H) &= \underbrace{\beta_3 \cdot 1}_{\text{servicios}} + \underbrace{\beta_4 \cdot 1}_{\text{mujer}} - \underbrace{\beta_2 \cdot 1}_{\text{manufactura}} \\ &= 0.454 \cdot 1 + (-0.33) \cdot 1 - 0.258 \cdot 1 \\ &= -0.138 \end{aligned}$$



ref: educ x agricultura

β_S educ x servicios
 β_M educ x manufatura

interacción : educ x sector

Una reconocida economista le plantea que las distintas industrias valoran diferente a la educación y que, por lo tanto, la remuneran de forma diferente. Usted decide hacerle caso y plantea el siguiente modelo (M2):

$$\log(\text{salario}_i) = \beta_0 + \beta_1 \text{educ}_i + \beta_2 \text{Manufacturas}_i + \beta_3 \text{Servicios}_i + \beta_4 \text{Mujer}_i \\ + \beta_5 \text{educ}X\text{manufact}_i + \beta_6 \text{educ}X\text{servicios}_i + u_i$$

obteniendo la siguiente salida de Stata:

Source	SS	df	MS	Number of obs	=	1,000
Model	196.755025	6	32.7925041	F(6, 993)	=	51.42
Residual	633.236779	993	.637700683	Prob > F	=	0.0000
Total	829.991803	999	.830822626	R-squared	=	0.2371
				Adj R-squared	=	0.2324
				Root MSE	=	.79856

logsalario	Coef.	Std. Err.	t	P> t	
educ_anios	$\beta_1 .0477684$.0142916	3.34	0.001	
Manufacturas	$\beta_2 -.3196944$.1744159	-1.83	0.067	
Servicios	$\beta_3 .3201212$.1242954	2.58	0.010	
mujer	$\beta_4 -.3169233$.0557134	-5.69	0.000	
→ manufactXeduc	$\beta_5 .0795826$.0212048	1.42	0.121	
serviciosXeduc	$\beta_6 .0254001$.0160006	1.59	0.113	
_cons	$\beta_0 3.624159$.0997043	36.35	0.000	

a. $\frac{\partial \log(\text{salario}_i)}{\partial \text{educ}} = \beta_0 + \beta_1 \text{educ}_i + \beta_2 \text{Manufacturas}_i + \beta_3 \text{Servicios}_i + \beta_4 \text{Mujer}_i$
 $+ \beta_5 \text{educ}X\text{manufact}_i + \beta_6 \text{educ}X\text{servicios}_i + u_i$

$$= \beta_1 + \beta_5 \cdot \text{manuf} + \beta_6 \cdot \text{servicios}.$$

- c. ¿Cuál es el efecto de un año adicional de educación en el salario para trabajadores en el sector manufacturero? Interprete, y luego encuentre la diferencia del efecto de un año adicional de educación en salario entre personas que trabajan en el sector manufacturero y personas que trabajan en el sector agrícola. ¿Podemos decir que esta diferencia es estadísticamente significativa? Considere un nivel de significancia del 5%

β_5

- $\frac{\partial \log(\text{salarios})}{\partial \text{educ}} = \beta_1 + \beta_5 \cdot \text{manufactura} + \beta_6 \cdot \text{servicios}$



$\text{servicios} = 0$

$$\left(\frac{\partial \log(\text{salarios})}{\partial \text{educ}} \Big|_{\text{manufacturero}} \right) = \beta_1 + \beta_5 \cdot \underbrace{\text{manufacturero}}_1$$

$$= \beta_1 + \beta_5$$

$$\log(\text{salarios}) = 0.047 + 0.079$$

$$= 0.1274$$

$$\begin{array}{r} \cdot 100 \\ \downarrow \\ 12,74\% \end{array}$$

Controlando por el género, un año adicional de educación para una persona del sector manufacturero, su salario aumenta en un 12,74%.

→ La diferencia promedio de un año de educación para alguien del sector manufactura respecto de agricultura es de $0.079 \cdot 100 \rightarrow 7,9\%$.

→ Significancia

$0.121 > 0.05 \rightarrow$ con un 95% de confianza no puedo rechazar H_0 ,

lo que significa que no es estadísticamente significativo a un 95%.

- d. En base a la información disponible, evalúe formalmente la hipótesis de que el efecto de la educación en el salario no depende del sector o de la industria en que el trabajador se desempeña. Considere un nivel de significancia del 5%, y no olvide especificar las hipótesis nula y alternativa.

TEMA III (TAREA 2)

Imagine que desea estimar la relación empírica entre las privatizaciones de los años 80 en Chile (es decir, la venta de empresas estatales a privados) y el rendimiento de las empresas en los años 90. Para ello recolecta datos de 100 empresas que eran estatales en los años 70, de las cuales la mitad fueron privatizadas en los años 80 y estima la siguiente regresión:

$$(1) \quad Y_i = \alpha + \beta \times P_i + u_i,$$

donde Y_i es el rendimiento promedio de la empresa i en los años 90 (medido como retorno sobre los activos), P_i es una variable binaria que toma el valor de 1 para aquellas empresas que fueron privatizadas, y u_i es un error con media igual a cero.

- ¿Cuál es el retorno promedio de las empresas privatizadas en términos de los parámetros del modelo (1), y qué condición debiera cumplirse para que dicho retorno sea superior al de las empresas no privatizadas?
- Usted aprende que en 1982 hubo una crisis económica y cree que a las empresas privatizadas después de la crisis les fue mejor que a las privatizadas antes de la crisis. Considere entonces el siguiente modelo:

$$(2) \quad Y_i = \phi + \gamma_1 \times P_i^{PRE} + \gamma_2 \times P_i^{POST} + u_i,$$

donde P_i^{PRE} y P_i^{POST} son variables binarias para privatizaciones antes y después de la crisis, respectivamente. Se pide entonces que exprese su hipótesis en términos de los parámetros de este modelo, y que luego explique cómo se relaciona ϕ en el modelo (2) con α en el modelo (1).

c. Una colega le propone considerar una especificación alternativa:

$$(3) \quad Y_i = \tau + \delta_1 \times P_i + \delta_2 \times (P_i \times T_i) + u_i,$$

donde P_i está definida como antes y T_i es una variable binaria que toma el valor de 1 para aquellas empresas que fueron privatizadas después de la crisis. En base a los parámetros del modelo (3), explique qué condición debiera cumplirse si a las empresas privatizadas después de la crisis les fue mejor que a las privatizadas antes de la crisis. Además, explique cómo se relaciona δ_1 en el modelo (3) con γ_1 en el modelo (2).