

# Taller de R Studio

## Aplicación

## Modelos Lineales

Edinson Tolentino

email: [edinson.tolentino@gmail.com](mailto:edinson.tolentino@gmail.com)

Twitter: @edutoleraymondi

Educate Peru

4 de septiembre de 2024



Introducción

Data y Variables

Pregunta 1

Pregunta 2

Pregunta 3



Introducción

Data y Variables

Pregunta 1

Pregunta 2

Pregunta 3

# Introducción





- ▶ Jacob Mincer publico su libro "Schooling, Experience and Earnings" (1974)



- ▶ Jacob Mincer publico su libro “Schooling, Experience and Earnings” (1974)
- ▶ Mincer modelo el logaritmo de ingresos como una funcion de años de educacion y años potenciales de experiencia en el mercado laboral
- ▶ Por lo tanto, la ecuación de Mincer sera:

$$\log w_x = X\beta + rSch + \beta_1 Exp - \beta_2 Exp^2$$





- ▶ La consultora **Marilyn Loden** , quien acuñó el termino **techo de cristal** en 1978 , es también la autora del libro Liderazgo femenino o ¿cómo triunfar en los negocios sin ser uno de los chicos?.



# Introducción





- ▶ En 1991, la Comisión de **Techos de Vidrio (glass ceiling)** del Departamento de Trabajo de EE. UU. (1991-1996) definió el **techo de vidrio** como:



- ▶ En 1991, la Comisión de **Techos de Vidrio (glass ceiling)** del Departamento de Trabajo de EE. UU. (1991-1996) definió el **techo de vidrio** como:
- ▶ ... esas barreras artificiales basadas en prejuicios organizacionales o de actitud que impiden que las personas calificadas asciendan en su organización a puestos de nivel gerencial.



- ▶ En 1991, la Comisión de **Techos de Vidrio (glass ceiling)** del Departamento de Trabajo de EE. UU. (1991-1996) definió el **techo de vidrio** como:
- ▶ ... esas barreras artificiales basadas en prejuicios organizacionales o de actitud que impiden que las personas calificadas asciendan en su organización a puestos de nivel gerencial.
- ▶ El concepto de **techo de cristal** se convirtió en el objeto de un estudio de investigación sistemático desde finales de los años ochenta.



- ▶ En 1991, la Comisión de **Techos de Vidrio (glass ceiling)** del Departamento de Trabajo de EE. UU. (1991-1996) definió el **techo de vidrio** como:
- ▶ ... esas barreras artificiales basadas en prejuicios organizacionales o de actitud que impiden que las personas calificadas asciendan en su organización a puestos de nivel gerencial.
- ▶ El concepto de **techo de cristal** se convirtió en el objeto de un estudio de investigación sistemático desde finales de los años ochenta.
- ▶ Los sociólogos estuvieron a la vanguardia de las primeras investigaciones y (algunos) economistas se interesaron a fines de la década de 1990.



- ▶ En 1991, la Comisión de **Techos de Vidrio (glass ceiling)** del Departamento de Trabajo de EE. UU. (1991-1996) definió el **techo de vidrio** como:
- ▶ ... esas barreras artificiales basadas en prejuicios organizacionales o de actitud que impiden que las personas calificadas asciendan en su organización a puestos de nivel gerencial.
- ▶ El concepto de **techo de cristal** se convirtió en el objeto de un estudio de investigación sistemático desde finales de los años ochenta.
- ▶ Los sociólogos estuvieron a la vanguardia de las primeras investigaciones y (algunos) economistas se interesaron a fines de la década de 1990.
- ▶ El enfoque original de los economistas en esta área era la promoción, no el pago.



- ▶ En 1991, la Comisión de **Techos de Vidrio (glass ceiling)** del Departamento de Trabajo de EE. UU. (1991-1996) definió el **techo de vidrio** como:
- ▶ ... esas barreras artificiales basadas en prejuicios organizacionales o de actitud que impiden que las personas calificadas asciendan en su organización a puestos de nivel gerencial.
- ▶ El concepto de **techo de cristal** se convirtió en el objeto de un estudio de investigación sistemático desde finales de los años ochenta.
- ▶ Los sociólogos estuvieron a la vanguardia de las primeras investigaciones y (algunos) economistas se interesaron a fines de la década de 1990.
- ▶ El enfoque original de los economistas en esta área era la promoción, no el pago.
- ▶ A principios de la década de 2000, las regresiones cuantílicas se utilizaban de forma rutinaria en esta investigación para centrarse en la remuneración.







- ▶ La socióloga Catherine Berheide, quien acuñó la frase **piso pegajoso** (sticky floor) en 1992



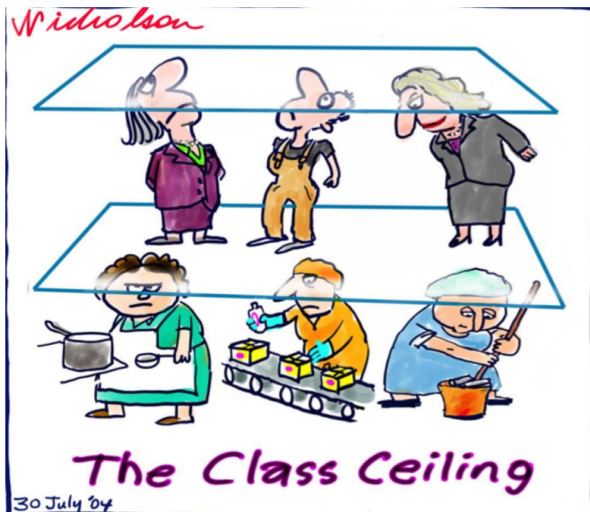
- ▶ La socióloga Catherine Berheide, quien acuñó la frase **piso pegajoso** (sticky floor) en 1992
- ▶ El **piso pegajoso** se refiere a los casos en los que las mujeres ocupan puestos de baja categoría, bajos salarios y de baja movilidad en el mercado laboral.



- ▶ La socióloga Catherine Berheide, quien acuñó la frase **piso pegajoso** (sticky floor) en 1992
- ▶ El **piso pegajoso** se refiere a los casos en los que las mujeres ocupan puestos de baja categoría, bajos salarios y de baja movilidad en el mercado laboral.
- ▶ Donde se dice:



- ▶ La socióloga Catherine Berheide, quien acuñó la frase **piso pegajoso** (sticky floor) en 1992
- ▶ El **piso pegajoso** se refiere a los casos en los que las mujeres ocupan puestos de baja categoría, bajos salarios y de baja movilidad en el mercado laboral.
- ▶ Donde se dice:
- ▶ ... la mayoría de las mujeres deberían tener la suerte de tener el techo de cristal como su problema ... muchas están atascadas en el piso pegajoso.



# Roadmap



Introducción

Data y Variables

Pregunta 1

Pregunta 2

Pregunta 3



- La información que se utilizará es proveniente de la base de datos de la Encuesta Nacional de Hogares (ENAHOG). Se procesa la base de datos del modulo 300 y 500 donde se analizará los ingresos mensuales de los trabajadores.

**Cuadro:** Descripción de variables

Variables	Descripción
ln wage	Logaritmo ingreso mensual (Soles)
reduca	Años de educación
rmujer	==1, mujer
redad	Años de edad
redadsq	Años de edad al cuadrado
rpareja	== 1, con pareja
rDpto	Según departamentos(dummies)



**Cuadro:** Estadísticas descriptivas

	Personas	Promedio	Mediana	Min.	Max.	Std
r6	24774	1,443.89	996.04	0.00	52,063.25	1,745
reduca	24774	8.70	11.00	0.00	18.00	5
rmujer	24774	0.28	0.00	0.00	1.00	0
redad	24774	46.80	47.00	18.00	69.00	12
redadsq	24774	2,330.71	2,209.00	324.00	4,761.00	1,109
rpareja	24774	0.66	1.00	0.00	1.00	0

Fuente: ENAHO - 2021.

Elaboracion: Autor





**Cuadro:** Estadísticas descriptivas

	Personas	Promedio	Mediana	Min.	Max.	Std
r6	24774	1,443.89	996.04	0.00	52,063.25	1,745
reduca	24774	8.70	11.00	0.00	18.00	5
rmujer	24774	0.28	0.00	0.00	1.00	0
redad	24774	46.80	47.00	18.00	69.00	12
redadsq	24774	2,330.71	2,209.00	324.00	4,761.00	1,109
rpareja	24774	0.66	1.00	0.00	1.00	0

Fuente: ENAHO - 2021.

Elaboracion: Autor

- ▶ Alrededor de 28 % de los trabajadores son mujeres.
- ▶ En promedio, la experiencia de los trabajadores es de 47 años y sus años de educación son de 9 años.
- ▶ El 66 % de los trabajadores tiene una condición civil : con pareja.

# Introducción





- ▶ Se realizará la estimación de la ecuación de salarios de los trabajadores peruanos utilizando la información 2020 (ENAH0)



- ▶ Se realizará la estimación de la ecuación de salarios de los trabajadores peruanos utilizando la información 2020 (ENAH0)
- ▶ Para realizar el análisis se propone tres ecuaciones:

$$\ln wage_i = \alpha_0 + \alpha_1 reduca_i + \mu_i \quad (1)$$

$$\begin{aligned} \ln wage_i = & \alpha_0 + \alpha_1 reduca_i + \pi mujer_i \\ & + \alpha_2 redad_i + \alpha_3 redad_i^2 + \alpha_4 rpareja_i + \mu_i \end{aligned} \quad (2)$$

$$\begin{aligned} \ln wage_i = & \alpha_0 + \alpha_1 reduca_i + \pi mujer_i \\ & + \alpha_2 redad_i + \alpha_3 redad_i^2 + \alpha_4 rpareja_i + \sum_{i=2}^{25} \lambda_i dpto_i + \mu_i \end{aligned} \quad (3)$$

# Roadmap



Introducción

Data y Variables

Pregunta 1

Pregunta 2

Pregunta 3

# Pregunta 1



- 1 Estime la regresión de la ecuación 2 usando OLS. Programación en STATA
  - ▶ Use el nivel de significancia de 0.05 para testear la presencia de heterocedasticidad.
  - ▶ ¿Qué es lo que usted concluye?



**Cuadro: Modelo Lineal**

	(1)	(2)
reduca	0.072*** (0.00)	0.071*** (0.00)
rmujer		-0.560*** (0.02)
redad		0.095*** (0.00)
redadsq		-0.001*** (0.00)
rpareja		-0.134*** (0.02)
Constant	6.085*** (0.02)	4.450*** (0.11)
Observations	24774	24774

Errores estandar en parentesis.

Fuente: EnAHO 2021.

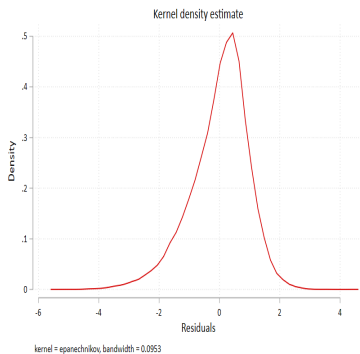
Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

# Pregunta 1



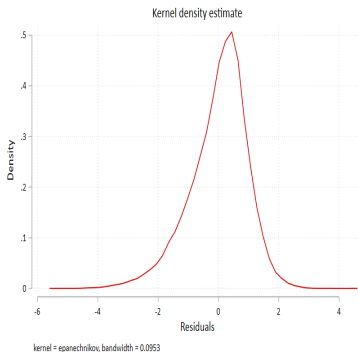
**Figura 1:**







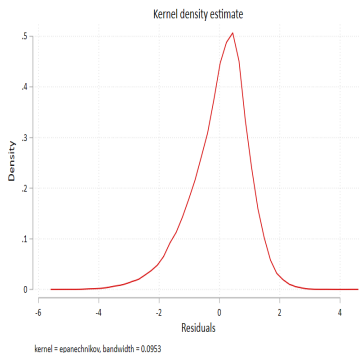
**Figura 1:**



- La distribución es muy puntiaguda con colas más alargadas en comparación con una distribución normal.



**Figura 1:**

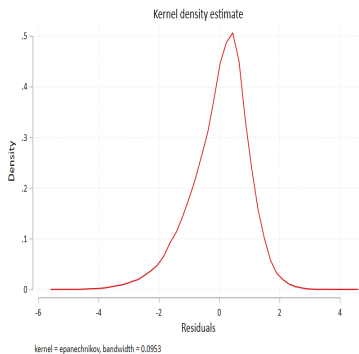


- ▶ La distribución es muy puntiaguda con colas más alargadas en comparación con una distribución normal.
- ▶ La prueba confirma que las dos proposiciones clave que gobiernan la normalidad (es decir, simetría y mesokurtosis) son ambas decididamente rechazadas por los datos en este caso.

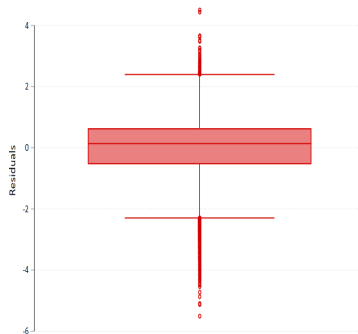
# Pregunta 1



**Figura 1:**



**Figura 2:**





- ▶ ¿Cuál es la diferencia entre los test Breusch-Pagan/Cook-Weisberg y el test de White/Koenker ?
- ▶ El Breusch-Pagan/Cook-Weisberg test asume que los errores de la ecuación original se distribuyen de manera normal.
- ▶ Por otro lado, el test de White/Koenker solo asume que los errores de la ecuación original son idéntica e independientemente distribuidos (i.i.d).
- ▶ Por lo tanto, es útil probar los residuos del modelo de regresión de la ecuación original para determinar la **normalidad** para decidir cuál de estas pruebas de heterocedasticidad usar.
- ▶ El comando relevante en STATA, dada la normalidad es violada, es:

*hettest mujer neduca exper expersq civil, iid*

## Pregunta 1



- El test de Koenker/White heteroscedasticity requiere la estimación del siguiente modelo de regresión auxiliar :

$$\varepsilon_i^2 = \delta_0 + \delta_1 \text{reduca}_i + \phi \text{rmujer}_i + \delta_2 \text{redad}_i + \delta_3 \text{redad}_i^2 + \delta_4 \text{rpareja}_i + \xi_i \quad (4)$$

- Hipotesis

$$H_o: \delta_0 = \delta_1 = \phi = \delta_2 = \delta_3 = \delta_4 = 0$$

vs  $H_a: H_o$  no es verdad

- Bajo la hipotesis Nula , el test de LM es definido como

$$n \times R^2 \sim \chi_{k-1}^2$$

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: neduca mujer exper expersq civil

chi2(5) = 407.51

Prob > chi2 = 0.0000

## Pregunta 1



- El test de Koenker/White heteroscedasticity requiere la estimación del siguiente modelo de regresión auxiliar :

$$\varepsilon_i^2 = \delta_0 + \delta_1 \text{reduca}_i + \phi \text{rmujer}_i + \delta_2 \text{redad}_i + \delta_3 \text{redad}_i^2 + \delta_4 \text{rpareja}_i + \xi_i \quad (4)$$

- Hipotesis

$$H_o: \delta_0 = \delta_1 = \phi = \delta_2 = \delta_3 = \delta_4 = 0$$

vs  $H_a: H_o$  no es verdad

- Bajo la hipotesis Nula , el test de LM es definido como

$$n \times R^2 \sim \chi^2_{k-1}$$

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

H0: Constant variance

Variables: neduca mujer exper expersq civil

chi2(5) = 407.51

Prob > chi2 = 0.0000

- La hipotesis nula de homocedasticidad tendra que ser analizada para el modelo de log(wage)

# Pregunta 1



Cuadro: Modelo Lineal

	$\varepsilon^2_{-i}$
reduca	0.016** (0.01)
rmujer	1.260*** (0.10)
redad	-0.090*** (0.02)
redadsq	0.001*** (0.00)
rpareja	1.263*** (0.09)
Constant	1.674*** (0.53)
Observations	24774

Errores estandar en paréntesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

# Pregunta 1



Cuadro: Modelo Lineal

	$\varepsilon^2_{-i}$
reduca	0.016** (0.01)
rmujer	1.260*** (0.10)
redad	-0.090*** (0.02)
redadsq	0.001*** (0.00)
rpareja	1.263*** (0.09)
Constant	1.674*** (0.53)
Observations	24774

Errores estandar en paréntesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

► Por tanto,  $n \times R^2$  es igual a 24,774



# Pregunta 1



Cuadro: Modelo Lineal

	$\varepsilon^2_{-i}$
reduca	0.016** (0.01)
rmujer	1.260*** (0.10)
edad	-0.090*** (0.02)
redadsq	0.001*** (0.00)
rpareja	1.263*** (0.09)
Constant	1.674*** (0.53)
Observations	24774

Errores estandar en parentesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- ▶ Por tanto,  $n \times R^2$  es igual a 24,774
- ▶ Dado un valor de significancia de 0.05, la distribución de la tabla  $\chi^2_{k-1}$  arroja el valor de 11.07 ( $\chi^2_5$ )



**Cuadro:** Modelo Lineal

	$\varepsilon^2_{-i}$
reduca	0.016** (0.01)
rmujer	1.260*** (0.10)
edad	-0.090*** (0.02)
redadsq	0.001*** (0.00)
rpareja	1.263*** (0.09)
Constant	1.674*** (0.53)
Observations	24774

Errores estandar en parentesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- ▶ Por tanto,  $n \times R^2$  es igual a 24,774
- ▶ Dado un valor de significancia de 0.05, la distribución de la tabla  $\chi^2_{k-1}$  arroja el valor de 11.07 ( $\chi^2_5$ )
- ▶ Evidenciando que se rehaza la  $H_0$ (homocedasticidad)



**Cuadro:** Modelo Lineal

	$\varepsilon^2_{-i}$
reduca	0.016** (0.01)
rmujer	1.260*** (0.10)
redad	-0.090*** (0.02)
redadsq	0.001*** (0.00)
rpareja	1.263*** (0.09)
Constant	1.674*** (0.53)
Observations	24774

Errores estandar en paréntesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- ▶ Por tanto,  $n \times R^2$  es igual a 24,774
- ▶ Dado un valor de significancia de 0.05, la distribución de la tabla  $\chi^2_{k-1}$  arroja el valor de 11.07 ( $\chi^2_5$ )
- ▶ Evidenciando que se rehaza la  $H_0$ (homocedasticidad)
- ▶ Concluyendo que existe presencia de **heterocedasticidad**



**Cuadro:** Modelo Lineal

	$\varepsilon^2_{-i}$
reduca	0.016** (0.01)
rmujer	1.260*** (0.10)
edad	-0.090*** (0.02)
redadsq	0.001*** (0.00)
rpareja	1.263*** (0.09)
Constant	1.674*** (0.53)
Observations	24774

Errores estandar en paréntesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.

- ▶ Por tanto,  $n \times R^2$  es igual a 24,774
- ▶ Dado un valor de significancia de 0.05, la distribución de la tabla  $\chi^2_{k-1}$  arroja el valor de 11.07 ( $\chi^2_5$ )
- ▶ Evidenciando que se rehaza la  $H_0$ (homocedasticidad)
- ▶ Concluyendo que existe presencia de **heterocedasticidad**
- ▶ Se necesita corregir la matriz de varianza-covarianza usando la opción robust antes del análisis

# Roadmap



Introducción

Data y Variables

Pregunta 1

**Pregunta 2**

Pregunta 3



- 2 Use el comando robust para corregir la presencia de heterocedasticidad y re-estime el modelo (2). Luego responda las siguientes preguntas
  1. Use la estructura del Test de Wald, realice usando un nivel de significancia de 0.05 para testear:

$$H_o : \alpha_4 = \pi = 0$$

2. Utilizando el Test de Wald en STATA, Python (test), ¿Qué es lo que usted concluye?



- Según el modelo :

$$\log r6_i = \alpha_0 + \alpha_1 \text{reduca}_i + \pi \text{rmujer}_i + \alpha_2 \text{redad}_i + \alpha_3 \text{redadsq}_i + \alpha_4 \text{rpareja}_i + \mu_i$$

Cuadro: Modelo Lineal robusto

	ln wage		ln wage		ln wage	
reduca	0.072***	(0.00)	0.072***	(0.00)	0.071***	(0.00)
rmujer			-0.476***	(0.02)	-0.560***	(0.03)
redad					0.095***	(0.01)
redadsq					-0.001***	(0.00)
rpareja					-0.134***	(0.02)
Constant	6.085***	(0.02)	6.218***	(0.01)	4.450***	(0.11)
Observations	24774		24774		24774	

Errores estandar en parentesis.

Fuente: EnAHO 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the 1 %, 5 % and 10 % levels respectively for zero.



- ▶ La hipótesis nula es expresada como:

$$H_o : \alpha_4 = \pi = 0$$

$$H_a : H_o - no - verdad$$

- ▶ La matriz de la forma del test de Wald es especificado como:

$$Wald = [(\hat{\alpha}_4 - \alpha_4) \quad (\hat{\pi} - \pi)] \hat{V}_{robust}^{-1} [(\hat{\alpha}_4 - \alpha_4) \quad (\hat{\pi} - \pi)]'$$

- ▶ El test de Wald implementa las restricciones de la  $H_o$

$$Wald = [\hat{\alpha}_4 \quad \hat{\pi}] \hat{V}_{robust}^{-1} [\hat{\alpha}_4 \quad \hat{\pi}]'$$

- ▶ El valor del test de Wald para una significancia conjunta de los dos coeficientes es **934.1019** dado una distribución de chi-cuadrado con 2 grados de libertad de **5.99** (ver tabla de distribución de una  $\chi^2$ )
  - ▶ STATA : 646.08286
  - ▶ Python : 934.1019



# Roadmap



Introducción

Data y Variables

Pregunta 1

Pregunta 2

Pregunta 3

## Pregunta 3



- 2 Usando el t-test asintótico y al nivel de significancia de 0.05 se le pide testear la proposición: el log de salario alcanza un maximo alrededor de 50 años. Reporte todos los cálculos relevantes. ¿Qué es lo que usted concluye?



Cuadro: Modelos Lineales Robustos

	(1)	(2)	(3)
reduca	0.07*** (0.00)	0.07*** (0.00)	0.07*** (0.00)
rmujer		-0.48*** (0.02)	-0.56*** (0.03)
Observaciones	24774	24774	24774
R <sup>2</sup>	0.0858	0.111	0.130
Controls			√

Fuente: ENAHO - 2021.

Elaboracion: Autor

\*\*\*, \*\*, \* denote statistical significance at the  
1 %, 5 % and 10 % levels respectively for zero.



- ▶ La derivada obtenida , esta determinada como:

$$\frac{\partial \log(wage)}{\partial redad} = \hat{\beta}_2 + 2\hat{\beta}_3 redad$$

- ▶ El punto de estado estacionario es calculado al igualar en cero la derivada obtenida y despejar el valor de la variable *exper*
- ▶ Para la presente aplicación, se tiene:

$$redad_{estacionario} = \frac{\hat{\beta}_2}{-2\hat{\beta}_3} = \frac{0.095}{2 \times (-0.001)} = 43 = \hat{\Delta}$$

- ▶ Por tanto, deseamos testear la proposición planteada a través de la hipótesis

$$H_o : \Delta = 50 \text{ vs } H_a : \Delta \neq 50$$

### Pregunta 3



- Usando el **Metodo Delta**, la varianza muestral esta dado por:

$$\begin{aligned} Var(\hat{\Delta}) &= \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} \right)^2 Var(\hat{\gamma}_2) + \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} \right)^2 Var(\hat{\gamma}_3) \\ &+ 2 \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} \right) \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} \right) Cov(\hat{\gamma}_2, \hat{\gamma}_3) \end{aligned}$$

- Recordando:

$$\hat{\Delta} = -\frac{\hat{\gamma}_2}{2\hat{\gamma}_3}$$

- Donde, para el presente caso:

### Pregunta 3



- Usando el **Metodo Delta**, la varianza muestral esta dado por:

$$\begin{aligned} Var(\hat{\Delta}) &= \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} \right)^2 Var(\hat{\gamma}_2) + \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} \right)^2 Var(\hat{\gamma}_3) \\ &+ 2 \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} \right) \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} \right) Cov(\hat{\gamma}_2, \hat{\gamma}_3) \end{aligned}$$

- Recordando:

$$\hat{\Delta} = -\frac{\hat{\gamma}_2}{2\hat{\gamma}_3}$$

- Donde, para el presente caso:

$$\frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} = -\frac{1}{2\hat{\gamma}_3}$$

### Pregunta 3



- Usando el **Metodo Delta**, la varianza muestral esta dado por:

$$\begin{aligned} Var(\hat{\Delta}) &= \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} \right)^2 Var(\hat{\gamma}_2) + \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} \right)^2 Var(\hat{\gamma}_3) \\ &+ 2 \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} \right) \left( \frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} \right) Cov(\hat{\gamma}_2, \hat{\gamma}_3) \end{aligned}$$

- Recordando:

$$\hat{\Delta} = -\frac{\hat{\gamma}_2}{2\hat{\gamma}_3}$$

- Donde, para el presente caso:

$$\frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_2} = -\frac{1}{2\hat{\gamma}_3}$$

$$\frac{\partial \hat{\Delta}}{\partial \hat{\gamma}_3} = \frac{\hat{\gamma}_2}{2\hat{\gamma}_3^2}$$

## Pregunta 3



- El valor de las derivadas encontradas son calculadas:



### Pregunta 3



- El valor de las derivadas encontradas son calculadas:

$$\frac{\widehat{\partial \Delta}}{\widehat{\partial \beta_2}} = -\frac{1}{2\widehat{\beta_3}} = 452.16$$



- El valor de las derivadas encontradas son calculadas:

$$\frac{\partial \hat{\Delta}}{\partial \hat{\beta}_2} = -\frac{1}{2\hat{\beta}_3} = 452.16$$

$$\frac{\partial \hat{\Delta}}{\partial \hat{\beta}_3} = -\frac{\hat{\beta}_2}{2\hat{\beta}_3^2} = 39031.71$$

- Los elementos obtenidos a través de la **robust** de la matriz de varianza-covarianza son:

	const	number_reduca	number_rmujer	number_redad	number_redadsq	number_rpareja
const	0.013004	-2.401348e-05	-1.583186e-04	-5.641416e-04	5.928826e-06	-1.540953e-04
number_reduca	-0.000024	2.053916e-06	-2.325714e-06	1.616597e-07	-1.187166e-10	2.660956e-07
number_rmujer	-0.000158	-2.325714e-06	6.486965e-04	-9.613457e-06	1.026090e-07	3.902510e-04
number_redad	-0.000564	1.616597e-07	-9.613457e-06	2.622671e-05	-2.823949e-07	-1.115742e-05
number_redadsq	0.000006	-1.187166e-10	1.026090e-07	-2.823949e-07	3.099567e-09	1.258704e-07
number_rpareja	-0.000154	2.660956e-07	3.902510e-04	-1.115742e-05	1.258704e-07	4.884660e-04

## Pregunta 3



- Los elementos obtenidos a través de la **robust** de la matriz de varianza-covarianza son:

	const	number_reduca	number_rmujer	number_redad	number_redadsq	number_rpareja
const	0.013004	-2.401348e-05	-1.583186e-04	-5.641416e-04	5.928826e-06	-1.540953e-04
number_reduca	-0.000024	2.053916e-06	-2.325714e-06	1.616597e-07	-1.187166e-10	2.660956e-07
number_rmujer	-0.000158	-2.325714e-06	6.486965e-04	-9.613457e-06	1.026090e-07	3.902510e-04
number_redad	-0.000564	1.616597e-07	-9.613457e-06	2.622671e-05	-2.823949e-07	-1.115742e-05
number_redadsq	0.000006	-1.187166e-10	1.026090e-07	-2.823949e-07	3.099567e-09	1.258704e-07
number_rpareja	-0.000154	2.660956e-07	3.902510e-04	-1.115742e-05	1.258704e-07	4.884660e-04

- Donde:

$$\text{Var}(\hat{\beta}_2) = 2.6226714554508038e - 05 = 0.000026226$$

$$\text{Var}(\hat{\beta}_3) = 3.099567058003456e - 09 = 0.000000003099$$

$$\text{Cov}(\hat{\beta}_2, \hat{\beta}_3) = -2.8239486279027537e - 07 = -0.0000002823$$

### Pregunta 3



- ▶ Por tanto, la varianza muestral es calculada reemplazando en el método delta:

$$\begin{aligned} \text{Var}(\hat{\Delta}) &= (452.16)^2 \times 0.000026226 + (39031.71)^2 \times 0.000000003099 + \dots \\ &= \dots + 2 \times (452.16 \times 39031.71) \times (-0.0000002823) \equiv 0.1164 \end{aligned}$$

- ▶ Nosotros deseamos testear la hipótesis:

$$H_o : \Delta = 50 \text{ vs } H_a : \Delta \neq 50$$

- ▶ El test-t es expresado en el presente caso como:

$$t = \frac{\hat{\Delta} - \Delta}{\sqrt{\text{var}(\Delta)}} = \frac{43 - 50}{\sqrt{0.1164}} = -20.51 \sim t_{\infty}$$

- ▶ **Dado el valor critico de  $\pm 1.96$ , por tanto se puede rechazar la hipótesis nula :**

$$|-20.57| > |-1.96| \Rightarrow |t_{\text{value}}| > |t_{\text{critico}}|$$

- ▶ Qué concluye?

## Pregunta 3



Figura : Perfil de salarios-edad para empleados de Perú

