PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
CIENCIAS SOCIALES
CICLO 2022-2

Fundamentos en Econometría
Práctica Dirigida 2 - Solucionario

Profesor: Juan Palomino
Jefe de práctica: Tania Paredes

---

1. **Descomposición de la suma de cuadrados y R²**

Se tiene el siguiente modelo: $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$, donde Y es la demanda de alimentos y X es el ingreso disponible. Se sabe que:

$$\sum_{i=1}^{20} X_i Y_i = 1973.67$$
$$\sum_{i=1}^{20} Y_i^2 = 1813.53$$
$$\bar{Y} = 8.765$$
$$\bar{X} = 9.56$$
$$\sum_{i=1}^{20} X_i^2 = 2165.18$$
$$n = 20$$

a. Estime los parámetros $\widehat{\beta_1}$ y $\widehat{\beta_2}$ por MCO.

SOL:

Para estimar por MCO, planteamos el modelo en desviaciones: (visto en la PD1) y en clase

$$\boxed{y_i = \hat{\beta}_2 x_i + \hat{\varepsilon}_i}$$

De este, definimos la Suma de Cuadrados Residuales (SCR)

$$\boxed{SCR = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_2 x_i)^2}$$

→ expresado en desviaciones

Después de realizar el proceso de min SCR y de derivar respecto a $\hat{\beta}_2$:

$$\boxed{\hat{\beta}_2 = \frac{\sum_{i=1}^{n}(y_i x_i)}{\sum_{i=1}^{n} x_i^2}} \qquad \boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}}$$

Por otra parte, podemos reescribir $\hat{\beta}_2$ (con la finalidad de poder utilizar los datos del ejercicio:

$$\boxed{\hat{\beta}_2 = \frac{\sum X_i Y_i - \bar{Y}\sum X_i}{\sum X_i^2 - \bar{X}\sum X_i}}$$

Necesitamos $\sum X_i$. Tenemos como dato que $\bar{X} = \frac{1}{n}\sum_i X_i = 9.56$

y dado que $n = 20$, obtenemos que $\sum x_i = 191.2$

Reemplazamos datos en $\hat{\beta}_2$

$$\boxed{\hat{\beta}_2 = \frac{\sum X_i Y_i - \bar{Y}\sum X_i}{\sum X_i^2 - \bar{X}\sum X_i} = \frac{1973.67 - 8.765(191.2)}{2165.18 - 9.56(191.2)} = \frac{297.802}{337.308} = 0.883}$$

$$\boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} = 8.765 - 0.883(9.56) = 0.324}$$

b. Realice una interpretación económica de los resultados encontrados.

SOL:

$$\frac{\partial Y}{\partial X} = \frac{\Delta Y}{\Delta X} = \beta_2 \rightarrow \text{Espero } \beta_2 > 0 \quad ; \beta_2 = 0.883$$

Un incremento en 1 unidad respecto a mi ingreso disponible generará un incremento de β_2 (0.883) en la variable Y (demanda de alimentos). Podría decirse que los bienes son normales, dado que $\uparrow X \rightarrow \uparrow Y$
Ejemplo: Y_i podría representar el consumo en arroz

c. Calcule el SCT, SCE, SCR y obtenga el valor del R-cuadrado

SOL:

Partiendo del modelo en desviaciones y elevamos este al cuadrado

$$(y_i = \hat{\beta}_2 x_i + \hat{\varepsilon}_i)^2$$

$$\boxed{y_i^2 = \hat{\beta}_2^2 x_i^2 + 2\hat{\beta}_2 x_i \hat{\varepsilon}_i + \hat{\varepsilon}_i^2}$$

Aplicamos sumatoria a la expresión anterior

$$\sum y_i^2 = \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{\varepsilon}_i^2 + 2\hat{\beta}_2 \underbrace{\sum x_i \hat{\varepsilon}_i}_{=0}$$

$=0 \rightarrow$ Por ortogonalidad
Prueba de ecuación normal

Finalmente obtenemos:

$$\boxed{\sum y_i^2 = \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{\varepsilon}_i^2}$$

SCT = SCE + SCR

Teniendo la forma funcional de la SCT, SCE y SCR, procedemos a reemplazar:

$$* \ SCT = \sum y_i^2 = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - 2\bar{Y}\sum Y_i + \sum \bar{Y}^2,$$

Y dado que $n\bar{Y} = \sum Y_i$

$$= \sum Y_i^2 - 2\bar{Y}(n\bar{Y}) + n\bar{Y}^2$$

$$= \sum Y_i^2 \underbrace{- 2n\bar{Y}^2 + n\bar{Y}^2}$$

$$= \sum Y_i^2 - n\bar{Y}^2$$

$$= 1813.53 - 20(8.765)^2$$

$$\boxed{SCT = 277.025}$$

$$* SCE = \hat{\beta}_2^2 \sum x_i^2 = \hat{\beta}_2^2 \sum (X_i - \bar{X})^2 = \hat{\beta}_2^2 \left(\sum X_i^2 - 2\bar{X}\sum X_i + \sum \bar{X}^2\right)$$

$$= (0.883)^2\left(\sum X_i^2 \underbrace{- 2n\bar{X}^2 + n\bar{X}^2}\right)$$

$$= (0.883)^2\left(\sum X_i^2 - n\bar{X}^2\right)$$

$$= (0.779)\left(2165.18 - 20(9.56)^2\right)$$

$$= (0.779)\underbrace{\left(2165.18 - 1827.87\right)}$$

$$= (0.779) \quad (337.308)$$

$$\boxed{SCE = 262.763}$$

(*)

Recordamos de la parte a) que $n\bar{X} = \sum X_i$

$* SCR = SCT - SCE$

$$\boxed{SCR = 277.025 - 262.763 = 14.262}$$

* R-cuadrado

$$\boxed{R^2 = \frac{SCE}{SCT} = \frac{262.763}{277.025} = 0.95}$$

d. Estime la varianza del término de perturbación

SOL:

Procederemos a estimar la varianza del término de perturbación

$SCR = SCT - SCE = 14.262$

$$\boxed{\underbrace{\hat{\sigma}^2 = \frac{SCR}{n-2}} = \frac{14.262}{18} = 0.792}$$

Formula de varianza del término de error

e. Halle la varianza de $\widehat{\beta_1}$ y $\widehat{\beta_2}$

SOL:

Para hallar la varianza de $\hat{\beta}_1$ y $\hat{\beta}_2$, se tienen las siguientes formas funcionales y en estas procederemos a reemplazar los valores que tenemos).

* $\boxed{Var(\hat{\beta}_1)} = \hat{\sigma}^2\left(\frac{1}{n} + \bar{X}^2/\sum x_i^2\right) = 0.792\left(\frac{1}{20} + \frac{(9.56)^2}{337.308}\right)$

$= \boxed{0.2541}$

* $\boxed{Var(\hat{\beta}_2)} = \hat{\sigma}^2/\sum x_i^2 = 0.792/337.308 = \boxed{0.0023}$

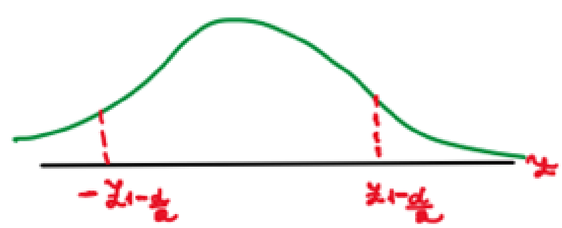
OJO: $\sum x_i^2 \rightarrow$ en desviaciones

f. Encuentre el intervalo de confianza para la estimación de β_2 al 95% y 99%

SOL:

Para el intervalo 95% encuen—

* Definido el intervalo, lo que nos interesa hallar es $\boxed{Pr\left(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}\right) = 1-\alpha}$



A partir de la función de densidad normal estándar, se definirá una variable aleatoria con distribución t-student para hallar los intervalos de confianza.

Para el intervalo de confianza al 95 %.

* El primer paso será definir el nivel de confianza de $\boxed{100(1-\alpha)\%}$

↳ $100(1-\alpha) = 95\%$, por lo que para que ello se cumpla $\alpha = 0.05$

$Pr(\hat{\beta}_k - se(\hat{\beta}_k) \cdot t_{1-\frac{\alpha}{2}}(n-k) < \beta_k < \hat{\beta}_k + se(\hat{\beta}_k) \cdot t_{1-\frac{\alpha}{2}}(n-k)$ el que ← Tenemos Z, pero dado que σ^2 es población y no conocido, se usa $\hat{\sigma}^2$ construyendo una t-student (n-k) grados

* Además se debe tener en cuenta que $\boxed{t_{1-\frac{0.05}{2}}(\overbrace{n-k}^{18}) = 2{,}1009}$ (0.025) → El valor se halla en la tabla para 18 grados de libertad.

* Resolviendo:

$$Pr\left[\hat{\beta}_2 - t_{1-\frac{0.05}{2}}(18)\sqrt{Var(\hat{\beta}_2)} < \beta_2 < \hat{\beta}_2 + t_{1-\frac{0.05}{2}}(18)\sqrt{Var(\hat{\beta}_2)}\right] = 0.95$$

$$Pr\left[0.883 - (2{,}1009)\sqrt{0.0023} < \beta_2 < \hat{\beta}_2 + (2{,}1009)\sqrt{0.0023}\right] = 0.95$$

$$\boxed{Pr\left[0.7820 < \beta_2 < 0.9838\right] = 0.95}$$

(b) Para el intervalo de confianza al 99 %

* El primer paso será definir el nivel de confianza de $\boxed{100(1-\alpha)\%}$
↳ $100(1-\alpha) = 99\%$, por lo que para que ello se cumpla $\alpha = 0.01$

* Además se debe tener en cuenta que $\boxed{t_{1-\frac{0.01}{2}}(\overbrace{n-k}^{18}) = 2{,}87844}$ (0.005) → El valor se halla en la tabla para 18 grados de libertad.

Por analogía de lo realizado para el intervalo de confianza de 95%.

$$Pr\left[0.883 - 2.87844\sqrt{0.0023} < \beta_2 < 0.883 - 2.8784\sqrt{0.0023}\right] = 0.99$$

$$\boxed{Pr\left[0.7419 < \beta_2 < 1.02\right] = 0.99}$$

2. Demanda real de dinero e intervalos de confianza

Suponga que un investigador estima por MCO el siguiente modelo econométrico:

$$m_i^d = \beta_1 + \beta_2 Y_i + e_i$$

Donde m_i^d es la demanda real de dinero el individuo i e Y_i representa el ingreso del individuo i. Luego de realizar una encuesta a 20 individuos, el investigador obtuvo los siguientes resultados:

$$m_i^d = 0.15 + 0.2Y_i + e_i \; (\widehat{\beta_1} = 0.15, \widehat{\beta_2} = 0.2)$$

$$R^2 = 0.4$$

$$\hat{\sigma} = 0.9$$

a. De acuerdo con estos resultados, cuál es el efecto de 1 unidad adicional de ingreso sobre la demanda real de dinero.
SOL:

Una unidad adicional del ingreso, aumenta en promedio 0.2 unidades de la demanda real de dinero.

b. Construya un intervalo de confianza de 95% y al 99% para β_2.
SOL:

Para construir el intervalo de confianza, necesitemos conocer: $Var(\hat{\beta}_2)$.

* Partimos de $\hat{\sigma}^2 = \frac{SCR}{n-2} \rightarrow (0.9)^2 = \frac{SCR}{18}$

$\rightarrow n = 12, \; k = 2$

Por lo que $\boxed{SCR = 14.58}$

Además $R^2 = 0.4 \rightarrow$ y $R^2 = 1 - \frac{SCR}{SCT} = 1 - \frac{14.58}{SCT}$

$0.4 = 1 - \frac{14.58}{SCT}$

$$\boxed{SCT = \frac{14.58}{0.6} = 24.3}$$

Por otra parte, recordamos: SCT = SCE + SCR

$24.3 = SCE + 14.58$

$$\boxed{SCE = 9.72}$$

Y dado que $SCE = \beta_2^2 \sum x_i^2 = \underbrace{(0.2)^2}_{0.04} \sum x_i^2 = 9.72$

$$\boxed{\sum x_i^2 = 243}$$ → Con ello podemos hallar $Var(\hat{\beta}_2)$

* $Var(\hat{\beta}_2) = S^2 / \sum x_i^2 = S^2 / \sum (X_i - \bar{X})^2 = (0.9)^2 / 243$

$$\boxed{Var(\hat{\beta}_2) = 0.0033}$$

* Dado que ya obtuvimos el valor de $Var(\hat{\beta}_2)$, definimos los intervalos para el 95% y 99%.

Para 95%

$Pr[0.2 - 2.1009(\sqrt{0.0033}) < \beta_2 < 0.2 + 2.1009(\sqrt{0.0033})] = 0.95$

$$\boxed{Pr[0.0793 < \beta_2 < 0.3206] = 0.95}$$

Para 99%:

$Pr[0.2 - 2.8784(\sqrt{0.0033}) < \beta_2 < 0.2 + 2.8784(\sqrt{0.0033})] = 0.99$

$$\boxed{Pr[0.03465 < \beta_2 < 0.3654] = 0.99}$$

3. Laboratorio

En la siguiente tabla se muestran el salario por hora en soles percibido por un grupo de 10 trabajadores según años de educación. Realice en Excel los siguientes pasos.

Salario (soles por hora)	Educación (años de estudio)
58	11
70	12
90	18
65	10

100	20
75	16
47	11
92	20
75	14
73	15

a. Calcule los valores de los parámetros estimados por MCO del modelo bivariado:

$$W_i = \beta_1 + \beta_2 A_i + e_i$$

Donde W_i y A_i denotan el salario y el número de años de educación del individuo i, respectivamente.

b. En base a los parámetros estimados, calcule los valores estimados del salario, los residuos y los residuos al cuadrado

c. Realice una gráfica de la recta de regresión estimada.

d. ¿Existe otro par de valores para β_1 y β_2 que produzcan un SCR de menor valor?

e. Realice la estimación en Stata

Sol excel

$$\boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}}$$

$$\boxed{\hat{\beta}_2 = \frac{\sum_{i=1}^{N} y_i x_i}{\sum_{i=1}^{N} x_i^2} = \frac{\sum_{i=1}^{N} (Y-\bar{Y})(X-\bar{X})}{\sum_{i=1}^{N} (X-\bar{X})^2}}$$