

PS02 – Basic Econometrics

Lecturer: Luis Chávez

Los siguientes ejercicios permiten medir la capacidad analítica y procedimental. Se sugiere resolverlos en forma ascendente.

Problema 1: linealidad de modelos

A partir de los siguientes modelos, identificar cuáles son lineales, intrínsecamente lineales y no lineales.

a) $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + u_i$

b) $y_i = \ln(\beta_1) + \beta_2 x_{i2} + \beta_3 \ln(x_{i3}) + u_i$

c) $\ln(y_i) = \beta_1 x_{i1} + u_i$.

d) $y_i = x_{i1}^{\beta_1} x_{i2}^{\beta_2} u_i$

e) $\ln(y_i) = \beta_1 + \beta_2 \frac{1}{x_{i2}} + u_i$

Problema 2: regresión básica

De un grupo de 7 estudiantes, se dispone las notas de matemáticas (x) y de econometría (y):

$$x = (14, 16, 06, 11, 08, 11, 09)$$

$$y = (12, 18, 08, 15, 11, 14, 07)$$

- Realizar un scatterplot y establecer la forma funcional.
- Estimar matricialmente el modelo de regresión simple. Interpretar.
- Hallar los residuos y demostrar que $\sum \hat{u}_i = 0$.
- Hallar la nota de econometría para un estudiante que tiene 17 en matemáticas.

Problema 3: varianzas

Demostrar que las varianzas de los LSE en el modelo de regresión simple son:

$$\text{var}(\hat{\beta}_1) = \frac{\sum x_{i2}^2}{n \sum \underline{x}_{i2}^2} \sigma^2$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum \underline{x}_{i2}^2}$$

donde $\underline{x}_{i2} = x_{i2} - \bar{x}_2$ y σ^2 es la varianza de los errores. Demostrar también que su estimador es:

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n - 2}$$

Problema 4: Coeficiente de determinación

Demostrar que el coeficiente de determinación, R^2 , se puede escribir como:

$$R^2 = \hat{\beta}_2^2 \frac{\sum \underline{x}_{i2}^2}{\sum \underline{y}_i^2}$$

$$R^2 = 1 - \frac{SCR}{SCT}$$

donde $SCT = SCE + SCR$.

Problema 5: varianzas

Sea una muestra de 140 datos y un modelo de regresión simple:

- a) Si la varianza estimada de los errores es $\hat{\sigma}^2 = 2$, hallar SCR.
- b) Si $var(\hat{\beta}_2) = 0.012$, cuál es su ee y $\sum \underline{x}_{i2}$?
- c) Si $\bar{x}_2 = 70$ y $\bar{y} = 24$, hallar $\sum x_{i2}^2$.
- d) Hallar $ee(\hat{\beta}_1)$.

Problema 6: regresión simple

Para una muestra de 120 universidades, se estimó la incidencia del tamaño (en m^2) del campus (tc) sobre el número anual de egresados (ne). Los resultados básicos fueron.

$$\hat{ne}_i = 15.21 + 0.87\hat{tc}, \quad R^2 = 0.372 \quad ee = 7.4$$

- a) ¿Cuál es la predicción del modelo para una universidad de $3000m^2$?
- b) Si la media muestral del tamaño de campus es $4500m^2$, ¿cuál es la media muestral del número de egresados?
- c) Probar empíricamente usando datos de universidades peruanas.
- d) ¿Cuál es la desviación típica muestral de ne ? *Hint*: use ee y la varianza muestral de nt .
- e) ¿Cuáles serán los nuevos resultados si tc ahora está expresado en hectáreas ($he = 10000m^2$) y ne en miles?

Problema 7: regresión múltiple

Dada la información,

$$X'X = \begin{bmatrix} 62 & 0 & 0 \\ 0 & 610 & 64 \\ 0 & 92 & 985 \end{bmatrix} \quad X'y = \begin{bmatrix} 244 \\ 56 \\ 162 \end{bmatrix} \quad \sum (y - \bar{y})^2 = 1164$$

- a) ¿Cuál es el tamaño de la muestra?

- b) ¿Cuáles son los coeficientes de la regresión?
- c) ¿Cuál es el valor del ee de $\hat{\beta}_1$?
- d) Calcular la predicción de y , dado $(x_1, x_2) = (-4, 2)$.

Problema 8: ecuación de Mincer

Usted ha propuesto el modelo:

$$\ln(y_i) = \alpha + \beta_1 esc_i + \beta_2 exp_i + \beta_3 exp_i^2 + \beta_4 sex + \beta_5 edad + \beta_6 ec_i + \beta_7 ht_i + u_i$$

donde y son los ingresos, esc son los años de escolaridad, exp son los años de experiencia laboral, sex es una dummy (0 = varón, 1 = mujer), ec es el estado civil (0 = no casado, 1 = casado) y ht son la horas de trabajo. La muestra comprende un total de 2050 individuos recopilados de ENAHO. A continuación se presentan los resultados de la regresión:

Call:

$lm(formula = ly \sim esc + exp + exp2 + sex + edad + ec + ht,$
 $data = datos, weights = facpob07)$

Weighted Residuals:

Min	1Q	Median	3Q	Max
-267.214	-14.55	2.168	18.044	104.779

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	5.071392	0.142076	35.695000	< 2e-16	***
esc	0.079839	0.006623	12.055000	< 2e-16	***
exp	0.054699	0.007080	7.726000	1.73e-14	***
exp2	-0.001159	0.000188	-6.156000	8.95e-10	***
sex	-0.572269	0.050977	-11.226000	< 2e-16	***
edad	-0.002732	0.002040	-1.339000	0.181	
ec	0.026061	0.059422	0.439000	0.661	
ht	0.004563	0.000262	17.442000	< 2e-16	***

—
 Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 31.33 on 2042 degrees of freedom
 (9375 observations deleted due to missingness)

Multiple R-squared: 0.3061, Adjusted R-squared: 0.3037

F-statistic: 128.7 on 7 and 2042 DF, p-value: < 2.2e-16

- a) ¿Cuál sería la interpretación econométrica de cada coeficiente?
- b) ¿Cuál sería la interpretación económica de cada coeficiente?
- c) ¿Qué experiencia debe tener una persona para maximizar sus ingresos?
- d) ¿Qué supuestos de Gauss-Markov parecen no cumplirse en el modelo estimado?