

Econometría I

Taller

Pablo Astudillo-Estevéz

Elías José Mantilla

2023-10-13

Sección I

Usted cuenta con datos de los resultados de los exámenes parciales administrados para la clase de econometría I.

Tabla 1: Resultados para 10 estudiantes de Econometría I

id	midterm	final
1	80.17720	48.65460
2	67.50057	48.96702
3	51.46045	52.50818
4	41.65793	53.34862
5	52.96675	44.94110
6	67.41659	53.72615
7	75.42283	58.93460
8	53.69872	65.76566
9	45.47207	47.00501
10	42.05456	45.02272

1. Calcular, a mano, el coeficiente de correlación de Pearson. Sírvase de la definición del coeficiente de correlación:

$$r_{xy} = \text{cov}(x, y) / \sigma_x \sigma_y = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

2. Calcular, a mano, los coeficientes del modelo de regresión lineal simple:

$$\text{final}_i = \beta_0 + \beta_1 \text{midterm}_i + \epsilon_i$$

Hint: Utilice el hecho estilizado de que $\hat{\beta} = r_{xy} \frac{\sigma_y}{\sigma_x}$.

3. Este es un claro ejemplo del fenómeno conocido como regresión a la media o regresión a la mediocridad. ¿Explique en que consiste la regresión a la media y como está implicada en este ejemplo?

Sección II

Para cada uno de los modelos presentados a continuación, realice según sea requerido:

1. Identifique, enliste e interprete cada uno de los parámetros estimados.

Determinantes del GPA

Se recopiló una muestra de 141 estudiantes universitarios con el objetivo de estimar los parámetros de un modelo para predecir el GPA universitario en función del GPA de la escuela secundaria (hsGPA) y el puntaje del examen ACT (ACT) como predictores. La variable de interés, GPA, tiene como promedio 3.1 y sus valores están acotados en el rango [2.2, 4]. El modelo estimado es presentado a continuación:

```
lm(formula = colGPA ~ hsGPA + ACT, data = gpa1)
      coef.est coef.se
(Intercept)  1.2863   0.3408
hsGPA         0.4535   0.0958
ACT           0.0094   0.0108
---
n = 141, k = 3
residual sd = 0.3403, R-Squared = 0.18
```

Salario por hora

Los parámetros del siguiente modelo fueron estimados con una base de datos de 526 individuos. El modelo busca predecir el logaritmo del salario por hora en función de la educación, la experiencia y la tenencia de la empresa. El modelo estimado es presentado a continuación:

```
lm(formula = log(wage) ~ educ + exper + tenure, data = wage1)
      coef.est coef.se
(Intercept)  0.2844   0.1042
educ         0.0920   0.0073
exper        0.0041   0.0017
tenure       0.0221   0.0031
---
n = 526, k = 4
residual sd = 0.4409, R-Squared = 0.32
```

Sección III

Considere el siguiente modelo de regresión lineal simple: $y_i = \beta x_i + \epsilon_i$. Las muestras (x_i, y_i) pueden asumirse como idéntica e independientemente distribuidas (iid). Adicionalmente, se asume que $\epsilon_i \sim N(0, \sigma^2)$; $\mathbb{E}[\epsilon|x] = 0$; $\mathbb{V}[\epsilon|x] = \sigma^2$.

1. Muestre que $\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$.
2. Muestre que se trata de un estimador insesgado. $\mathbb{E}[\hat{\beta}|x] = \beta$.
3. Encuentre la varianza del estimador $\hat{\beta}$.

Demuestre que:

1. $\sum_i^n \hat{u}_i = 0$
2. $\sum_i^n x_i \hat{u}_i = 0$

Anexo

Para referencia del ejercicio II, se debería llegar a valores iguales o similares a:

Intercepto	Pendiente	Corr. Pearson
46.61617	0.0912243	0.1935205