

# Taller Práctico Regresión Lineal Simple (1) \*

**Estadística II** Universidad Nacional de Colombia, Sede Medellín

Este documento corresponde al primer taller práctico del curso de **Estadística II** para la Universidad Nacional de Colombia, Sede Medellín, en el periodo 2024 - 2. Se brinda una introducción al análisis de regresión. El enfoque de este taller está sobre las componentes asociadas al modelo de regresión lineal simple -especialmente los parámetros-. **Monitor:** Santiago Carmona Hincapié.

*Keywords:* regresión, parámetros

## Información general

Con el propósito de profundizar en los conceptos del modelo de regresión lineal simple vistos en clase, se propone afrontar dos problemas prácticos. El primero consta de una simulación que ahonda en las propiedades de los parámetros del modelo de regresión. El segundo, es una aproximación al uso del modelo en situaciones de la vida real.

**La solución para cada uno de los problemas se efectúa a partir del software estadístico R.**

### *Ejercicio con datos reales*

El índice de libertad económica es una medida que *evalúa el grado de libertad económica* en diferentes países. Se presentan diversos atributos, cuya descripción puede encontrarse **en el siguiente enlace:** <https://www.kaggle.com/datasets/mlippo/freedom-economic-index/data>

Table 1: Información en análisis

OS	PR	GI	JE	TB	GS	FH	BF	LF	MF	TF	IF	FF
<b>83.5</b>	94.2	88.3	58.3	90.7	89.2	76.0	86.9	77.3	76.3	95.0	90	80
<b>83.0</b>	94.2	91.3	98.1	70.4	64.6	95.7	89.3	60.7	80.8	86.4	85	80
<b>82.6</b>	93.5	83.4	94.3	78.0	82.4	91.7	91.3	62.8	74.5	79.2	90	70
<b>80.0</b>	82.2	73.4	94.0	79.2	90.5	90.3	84.9	69.1	80.1	86.4	70	60
<b>79.2</b>	96.9	84.9	95.8	64.6	40.6	97.6	89.5	57.7	69.1	79.2	95	80

Considere a ‘Overall Score’ como la variable respuesta. *Escoja una covariable y de respuesta a los siguientes planteamientos:*

1. Realice un **breve análisis descriptivo**. ¿Un modelo de regresión lineal simple podría ser adecuado en este caso? ¿Por qué?

\*El material asociado a este taller puede encontrarse en el repositorio del curso, (<https://github.com/Itssach/Estadistica-II>)

2. Escriba la ecuación del modelo de regresión lineal, considerando los supuestos asociados. Obtenga los valores calculados de los parámetros,  $\bar{y}$ ,  $\hat{y}$ ,  $\hat{\varepsilon}_i$  y analícelos.
3. Determine si los parámetros del modelo  $\beta_0, \beta_1$  son significativos, considerando  $\alpha = 0.05$ . Realice una interpretación en relación al problema. **¿Estos parámetros tienen sentido?**
4. Calcule un intervalo de confianza -considerando  $\alpha = 0.05$ - para ambos parámetros. ¿Puede concluir a partir de este resultado si los parámetros son significativos?

### *Ejercicio de simulación*

A partir de una simulación se pretende ilustrar algunos de los principios teóricos que comprenden el análisis de regresión lineal simple. Así, se plantea un modelo  $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$ ,  $\varepsilon_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$ , tal que  $\beta_0 = 50, \beta_1 = 10, \sigma^2 = 16$ . Suponga que se emplean  $n = 20$  observaciones para ajustar el modelo. **Genere 500 muestras de 20 observaciones de tal manera que  $x = 1, 1.5, 2, \dots, 10$  para cada muestra.**

1. Para cada muestra calcule  $\hat{\beta}_0, \hat{\beta}_1$ . Construya un histograma para cada parámetro estimado y concluya.
2. Para cada muestra, calcule un intervalo de confianza al 90% para  $\beta_1$ . *¿Cuántos de estos intervalos contienen al verdadero valor del parámetro? ¿Se corresponde la teoría con la práctica?*