

Taller Práctico Regresión Lineal Simple (2) *

Estadística II Universidad Nacional de Colombia, Sede Medellín

Este documento corresponde al segundo 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 el análisis de varianza, intervalos de confianza y predicción-. **Monitor:** *Santiago Carmona Hincapié*.

Keywords: regresión, análisis varianza

Información general

Con el propósito de profundizar en los conceptos del modelo de regresión lineal simple vistos en clase, se propone afrontar un problema práctico, haciendo uso del conjunto de datos del taller anterior (**primer taller**).

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
83.5	94.2	88.3	58.3	90.7	89.2	76.0	86.9	77.3	76.3	95.0	90	80
83.0	94.2	91.3	98.1	70.4	64.6	95.7	89.3	60.7	80.8	86.4	85	80
82.6	93.5	83.4	94.3	78.0	82.4	91.7	91.3	62.8	74.5	79.2	90	70
80.0	82.2	73.4	94.0	79.2	90.5	90.3	84.9	69.1	80.1	86.4	70	60
79.2	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 dos covariables y de respuesta a los siguientes planteamientos:*

1. Realice un breve análisis descriptivo, ¿Qué covariables podría escoger para un análisis de regresión lineal simple? *Seleccione dos.*
2. Realice una partición de la información en un conjunto de entrenamiento y prueba. *(Se sugiere emplear el 75% – 80% de los datos como conjunto de entrenamiento).*

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

3. Escriba la ecuación del modelo de regresión lineal, considerando los supuestos asociados. Ajuste el modelo lineal para cada una de las covariables seleccionadas. ¿Son comparables ambos modelos? ¿Tienen las mismas unidades?
4. Realice la prueba de significancia para *la pendiente* en alguno de los modelos, luego una prueba de significancia para la regresión mediante el análisis de varianza. ¿Existe alguna relación entre ambas pruebas? Brinde una interpretación para los parámetros (β_0, β_1) de uno de los modelos considerados.
5. Calcule la métrica R^2 para ambos modelos. Realice una interpretación. ¿Qué se puede concluir de una comparación del R^2 ?
6. Seleccione uno de los modelos planteados. Realice una predicción puntual para \hat{Y}_i empleando el conjunto de prueba. Brinde los correspondientes intervalos de confianza y predicción. ¿Qué diferencias existen entre ambos?