

Regresión lineal simple

Lino AA Notarantonio (lino@tec.mx)

25 March 2021

Estimación por Mínimos Cuadrados Ordinarios (MCO)

- Se dispone de una muestra aleatoria $(ingreso_i, gasto_i)_{i=1}^N$, N tamaño muestral.

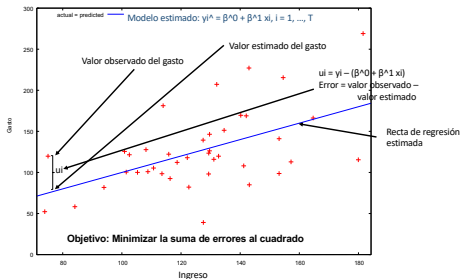


Figure 1: Gasto como función del ingreso

Estimación por Mínimos Cuadrados Ordinarios (MCO)

1. Estimar el modelo

$$gasto = \beta_0 + \beta_1 ingreso + u$$

por MCO.

2. Optimizar la función (*SRC*: suma de residuos cuadrados)

$$SRC(\beta_0, \beta_1) = \sum_{i=1}^N [y_i - (\beta_0 + \beta_1 x_i)]^2$$

Cargar los datos en RStudio

1. Bajar el archivo de datos `gastos.csv` y guardarlo en el Escritorio.
2. Abrir RStudio; en la barra de menu dar click en File
3. Seleccionar Import Dataset y después From Text (readr)
4. Se abre una ventana que le permitirá cargar el archivo en RStudio.
5. Estimar el modelo $Gasto = \beta_0 + \beta_1 Ingreso + u$.

Estimar el modelo con R

- ▶ Copiar y pegar el código a continuación en la consola.
- ▶ El nombre del archivo de datos es `gasto`.
- ▶ El código se debe ejecutar secuencialmente en la consola.

```
m <- lm(Gasto ~ Ingreso, data = gasto)
```

```
summary(m)
```

Modelo estimado

##

Call:

lm(formula = Gasto ~ Ingreso, data = gasto)

##

Residuals:

##	Min	1Q	Median	3Q	Max
##	-12.990	-3.563	-1.084	3.211	14.511

##

Coefficients:

##		Estimate	Std. Error	t value	Pr(> t)
##	(Intercept)	7.38870	4.00705	1.844	0.073002 .
##	Ingreso	0.23220	0.05528	4.200	0.000155 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

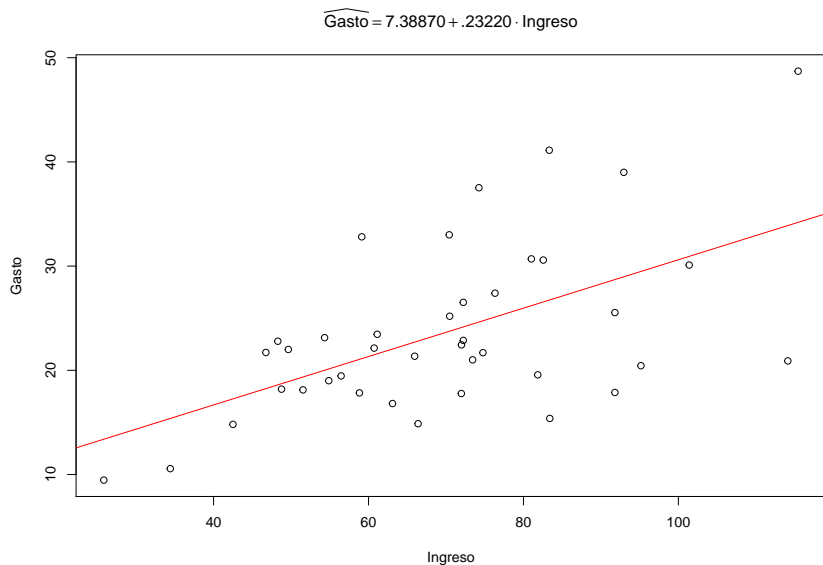
##

Residual standard error: 6.845 on 38 degrees of freedom

Multiple R-squared: 0.3171, Adjusted R-squared: 0.2991

F-statistic: 17.64 on 1 and 38 DF, p-value: 0.0001551

Modelo estimado: Plot



Tarea

Tarea: Ecuaciones normales

Minimizar (SRC: suma de los residuos cuadrados)

$$SRC(\beta_0, \beta_1) = \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2.$$

Punto crítico de la función SRC : Sistema lineal 2×2

$$\frac{\partial SRC}{\partial \beta_0} = (-2) \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)] = 0$$

$$\frac{\partial SRC}{\partial \beta_1} = (-2) \sum_{i=1}^n x_i [y_i - (\beta_0 + \beta_1 x_i)] = 0$$

(β_0, β_1) : incógnitas)

Tarea: Ecuaciones normales

- Mostrar que el sistema lineal en la diapositiva anterior se puede escribir como

$$\begin{aligned} n\beta_0 + \left(\sum_{i=1}^n x_i\right) \beta_1 &= \sum_{i=1}^n y_i \\ \left(\sum_{i=1}^n x_i\right) \beta_0 + \left(\sum_{i=1}^n x_i^2\right) \beta_1 &= \sum_{i=1}^n x_i y_i \end{aligned}$$

- Mostrar que el sistema lineal tiene solución si, y sólo si, la varianza muestral de la variable independiente x es diferente de cero.

Entrega Lunes 5 de abril. antes de las 09:00AM en el espacio de Tareas de Canvas.