Econometría Básica

Capítulo 02: Modelo de Regresión Lineal Simple

José Valderrama & Freddy Rojas
jtvalderrama@gmail.com & frojasca@gmail.com
Universidad Católica Santo Toribio de Mogrovejo

Septiembre de 2021

Contenido

Contenido

Contenido

Terminología

El modelo $y = \beta_0 + \beta_1 x + u$

- y Variable dependiente, variable explicada.
- x Variable independiente, regresor, variable explicativa, variable de control o covariado.
- u Término de error poblacional.
- β_0 Intercepto, constante o coeficiente no asociado con variables.
- β_1 Pendiente o coeficiente relacionado a x.

Terminología

El modelo $y = \beta_0 + \beta_1 x + u$

- y Variable dependiente, variable explicada.
- x Variable independiente, regresor, variable explicativa, variable de control o covariado.
- u Término de error poblacional.
- β_0 Intercepto, constante o coeficiente no asociado con variables.
- β_1 Pendiente o coeficiente relacionado a x.

Terminología

El modelo $y = \beta_0 + \beta_1 x + u$

- y Variable dependiente, variable explicada.
- x Variable independiente, regresor, variable explicativa, variable de control o covariado.
- u Término de error poblacional.
- β_0 Intercepto, constante o coeficiente no asociado con variables.
- β_1 Pendiente o coeficiente relacionado a x.

Terminología

El modelo $y = \beta_0 + \beta_1 x + u$

- y Variable dependiente, variable explicada.
- x Variable independiente, regresor, variable explicativa, variable de control o covariado.
- u Término de error poblacional.
- β_0 Intercepto, constante o coeficiente no asociado con variables.
- β_1 Pendiente o coeficiente relacionado a x.

USAT Universidad Católica Santo Toribio de Mogrovejo

Terminología

El modelo $y = \beta_0 + \beta_1 x + u$

- y Variable dependiente, variable explicada.
- x Variable independiente, regresor, variable explicativa, variable de control o covariado.
- u Término de error poblacional.
- β_0 Intercepto, constante o coeficiente no asociado con variables.
- β_1 Pendiente o coeficiente relacionado a x.

Terminología

El modelo $y = \beta_0 + \beta_1 x + u$

- y Variable dependiente, variable explicada.
- x Variable independiente, regresor, variable explicativa, variable de control o covariado.
- u Término de error poblacional.
- β_0 Intercepto, constante o coeficiente no asociado con variables.
- β_1 Pendiente o coeficiente relacionado a x.

Terminología

El MRLS es

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

El subíndice i va desde $i = 1$ a $i = n$ (última observación). Y es la variable dependiente; X es la variable dependiente; β_0 es el intercepto, constante o coeficiente no asociado con variables; β_1 es la pendiente o el coeficiente relacionado a X_i y u_i es el témino error. Y y X son datos. También, podemos denotar $\bar{Y}$ y $\bar{X}$ como el promedio muestral de las variables Y y X respectivamente. En término matriciales, tenemos

$$\underset{\rightarrow}{Y} = \begin{bmatrix} 50 \\ 30 \\ 20 \\ \vdots \\ 100 \end{bmatrix} \quad \text{y} \quad \underset{\rightarrow}{X} = \begin{bmatrix} 2 \\ 4 \\ 1 \\ \vdots \\ 6 \end{bmatrix}$$

Contenido

Contenido

Estimación: Mínimos cuadrados ordinarios

Dado que se quiere ajustar un grupo de puntos (muestra) a una línea, lo lógico sería entonces minimizar los errores cometidos por tratar de aproximar los puntos por una recta.

$$\hat{Y} = \alpha + \widehat{\beta} X$$

Error:

$$\varepsilon_i = Y_i - \hat{Y}_i = Y_i - (\alpha + \widehat{\beta} X)$$

Objetivo es minimizar

$$\sum_{i=1}^{n} \varepsilon_i^2 = \text{Min} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \text{Min} \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$$

¿por qué no minimizar los errores sin ninguna potencia? ¿y con otra potencia? ¿y si se minimizan los valores absolutos?

MCO simple

- Modelo Poblacional

$$y = \beta_0 + \beta_1 x + \mu$$

- Modelo muestral

$$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta} 1 x + \epsilon$$

- Ecuaciones normales: $\sum \epsilon_i = 0$; $\sum \epsilon_i x_i = 0$
- Pendiente

$$\widehat{\beta}_1 = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{\sum(x_i - \overline{x})y_i}{\sum(x_i - \overline{x})^2}$$

- Intercepto

$$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x}$$

USAT Universidad Católica Santo Toribio de Mogrovejo

Valores predichos y residuales

Una vez tengamos las estimaciones (basados en la data, porsupuesto) podemos calcular los valores predichos ($\widehat{y}$) y los residuos ($\widehat{u}$)

$$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x$$
$$\widehat{u} = y - \widehat{y}$$

El término de error y los residuales no son lo mismo. El primero está asociado al PGD; este último asociado al estimador (hecho por mortales)

¿Cómo obtenemos las estimaciones de MCO

Consideremos la siguiente data de corte transversal (notas sobre el cuestionario 1). Descargar la data quí

i	y_i	x_i	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$	$(y_i - \overline{y})$
1	4.4	3			
2	4.1	3			
3	5.2	5			
4	2.7	3			
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
80	3.1	4			

USAT Universidad Católica Santo Toribio de Mogrovejo

¿Cómo obtenemos las estimaciones de MCO?

Una vez que tengamos las estimaciones para β_0 y β_2 podemos completar la siguiente tabla.

i	y_i	x_i	β_0	$\beta_1 x_i$	$\widehat{y}_i$
1	4.4	3			
2	4.1	3			
3	5.2	5			
4	2.7	3			
⋮	⋮	⋮	⋮	⋮	⋮
80	3.1	4			

Pero, ¿serán insesgado y eficientes?

La relación entre Y y X

Comencemos con el modelo que explica (en términos estadísticos) los determinantes de las calificaciones en nuestra clase de econometría.

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

Observe a β_1, acompaña a la variable de interés y nos dice qué tanto cambia Y ante cambios en X; es decir

$$\Delta Y = \beta_1 \cdot \Delta X$$

Por lo tanto

$$\beta_1 \rightarrow \text{ es el objetivo}$$

Pero no conocemos al verdadero valor de β_1 (solo Dios lo sabe); sin embargo, podemos encontrar un estimador $\hat{\beta}_1$ (los economistas lo saben) que nos acerque al verdadero valor de β_1

Contenido

USAT
Universidad Católica
Santo Toribio de Mogrovejo

Contenido

USAT Universidad Católica Santo Toribio de Mogrovejo

BETA VS BETA *SOMBRERO*

- β_0 y β_1 son los parámetros poblacionales (los verdadero parámetros del *Proceso Generador de Datos o PGD*)
- $\widehat{\beta}_0$ y $\widehat{\beta}_1$ son los estimadores de esos parámetros poblacionales ("Los *betas* hechos por los mortales")
- Nos gustaría conocer los verdaderos parámetros (¡realmente lo haríamos!) Pero solo nosotros podemos hablar de estimaciones
- Recuerde que nos gustaría tener un estimador insesgado y eficiente.

Contenido

Propiedades algebraicas

- $\sum_{i=1}^{n} \hat{u}_i = 0 \Leftrightarrow \frac{\sum_{i=1}^{n} \hat{u}_i}{n} = 0$
- $\sum_{i=1}^{n} x_i \hat{u}_i = 0$
- $\overline{y} = \hat{\beta}_o + \hat{\beta}_1 \overline{x}$

USAT Universidad Católica Santo Toribio de Mogrovejo

Suma Cuadrada

Cada observación se puede descomponer de una parte explicada y otra no explicada: $yi = \hat{y}_i + \hat{u}_i$. A partir de esto definimos:

- **STC** Suma total de cuadrados $\sum_{i=1}^{n}(y_i - \overline{y})^2$
- **SEC** Suma explicada al cuadrado $\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$
- **SRC** Suma de residuos al cuadrado $\sum_{i=1}^{n}(\hat{u}_i)^2$

Finalmente STC=SEC+SRC Probar!

- **R^2** Es el indicador de ajuste más popular empleado para medir que tan bien el modelo se ajusta a los datos: $R^2 = \frac{SEC}{STC} = 1 - \frac{SRC}{STC}$

Contenido

Formas funcionales

Modelo	Regresión	Variable Dep. (Y)	Variable Indep. (X)	Interpretación del regresor (β_1)
Nivel - Nivel	$Y_i = \beta_0 + \beta_1 X_i + u_i$	Y	X	$\Delta Y = \beta_1 \Delta X$
Nivel – Log	$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i$	Y	log(X)	$\Delta Y = \left(\frac{\beta_1}{100}\right) \%\Delta X$
Log - Nivel	$\log(Y_i) = \beta_0 + \beta_1 X_i + u_i$	log(Y)	X	$\%\Delta Y = (100\beta_1)\Delta X$
Log-Log	$\log(Y_i) = \beta_0 + \beta_1 \log(X_i) + u_i$	log(Y)	Log(X)	$\%\Delta Y = \beta_1 \%\Delta X$

- El modelo Nivel-Nivel representa las variables en su forma original (regresión en forma lineal). Es decir, un cambio de una unidad en X, afecta en β_1 unidades a Y.
- El modelo Nivel-Log se interpreta como un incremento del 1 % de cambio en X es asociado a un cambio en Y de $0{,}01 \cdot \beta_1$.
- El modelo Log-Nivel es el menor frecuentemente utilizado y se conoce como la semielasticidad de Y respecto a X. Se interpreta como un incremento de 1 unidad en X es asociado a un cambio en Y de $(100 \cdot \beta_1)$ %.
- El modelo Log-Log es atribuye a β_1 la elasticidad de Y, respecto a X. Se interpreta como un incremento del 1 % en X es asociado a un cambio en Y de β_1 %.

Contenido

Contenido

USAT Universidad Católica Santo Toribio de Mogrovejo

Supuestos 1-4

1. Linealidad de los parámetros

$$y = \beta_0 + \beta_1 x + \mu$$

2. Muestreo aleatorio
3. Variación muestral de la variable explicativa
4. Media condicional cero del error:

$$E(\mu/x) = E(\mu) = 0$$

Supuestos 1-4

Con los supuestos 1 a 4 se prueba que los estimadores son insesgados:

$$E(\hat{\beta}_1) = \beta_1$$

$$E(\hat{\beta}_0) = \beta_0$$

Estimadores insesgados

$$\begin{aligned}
\widehat{\beta}_1 &= \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} \\
&= \frac{\sum(x_i - \overline{x})\{\beta_1(x_i - \overline{x}) + u_i\}}{\sum(x_i - \overline{x})^2} \\
&= \beta_1 \frac{\sum(x_i - \overline{x})^2}{\sum(x_i - \overline{x})^2} + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2} \\
&= \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2} \qquad (2)
\end{aligned}$$

Tomando el esperado y asumiendo que X no es estocástico (¡por cierto, una suposición fuerte!)

$$E[\widehat{\beta}_1] = \beta_1 + E\left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}\right]$$

USAT Universidad Católica Santo Toribio de Mogrovejo

Estimadores insesgados

Finalmente

$$
\begin{aligned}
E[\widehat{\beta}_1] &= \beta_1 + E\left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}\right] \\
&= \beta_1 + \frac{\sum(x_i - \overline{x})E[u_i]}{\sum(x_i - \overline{x})^2} \\
&= \beta_1
\end{aligned}
$$

Hemos demostrado que el estimador MCO es un estimador insesgado. Pero, todavía necesitamos la variación.

Contenido

Supuestos 5: Homocedasticidad

Supuesto 5:

$$Var(\mu/x) = \sigma^2$$

Sabiendo que:

$$\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})\mu_i}{\sum(x_i - \overline{x})^2}$$

Se demuestra que:

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(x_i - \overline{x})^2}$$

σ^2 no es conocido, pero puede ser estimado a partir de los residuales ϵ.

Estimadores eficientes

Veamos detenidamente la expresión (2)

$$\widehat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2} \quad \rightarrow \quad \widehat{\beta}_1 - \beta_1 = \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$$

llevando cuadrados a ambas expresiones

$$\left[\widehat{\beta}_1 - \beta_1\right]^2 = \left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}\right]^2$$

tomando el valor esperado

$$E\left[\widehat{\beta}_1 - \beta_1\right]^2 = E\left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}\right]^2 = \frac{\sum(x_i - \overline{x})^2 E(u_i^2)}{\left[\sum(x_i - \overline{x})^2\right]^2}$$

Nota: $E\left[\widehat{\beta}_1 - \beta_1\right]^2 \equiv \widehat{\sigma}^2_{\widehat{\beta}_1}$

USAT Universidad Católica Santo Toribio de Mogrovejo

Estimadores eficientes

reescribiendo la expresión anterior y teniendo en cuenta la característica i.i.d de u

$$
\begin{aligned}
E\left[\widehat{\beta}_1 - \beta_1\right]^2 &= \frac{\sum(x_i - \overline{x})^2 E(u_i^2)}{\left[\sum(x_i - \overline{x})^2\right]^2} \\
&= \frac{\sum(x_i - \overline{x})^2 E[u_i^2]}{\left[\sum(x_i - \overline{x})^2\right]^2} \\
&= E[u_i^2] \frac{\sum(x_i - \overline{x})^2}{\left[\sum(x_i - \overline{x})^2\right]^2} \qquad (3)
\end{aligned}
$$

finalmente

$$
E\left[\widehat{\beta}_1 - \beta_1\right]^2 = \sigma^2 \frac{1}{\sum(x_i - \overline{x})^2}
$$

donde $E[u^2] = \sigma^2 \equiv s_u^2$. Estamos asumiendo homocedasticidad.

Estimadores eficientes

¡Casi ahí!. Solo lo que necesitamos saber es el error estándar del error (u). Sabemos que la desviación estándar de la estimación MCO -en el caso de un regresor- es

$$Var(\widehat{\beta}_1) = \sigma^2 \frac{1}{\sum (x_i - \overline{x})^2} \tag{4}$$

¿Qué es σ^2?. Necesitamos un estimador (insesgado) para σ^2. Entonces, tenemos la siguiente propuesta:

$$\widehat{\sigma}^2 = \frac{\sum u^2}{(n-2)}$$

Entonces:

$$Var(\widehat{\beta}_1) = \widehat{\sigma}^2 \frac{1}{\sum (x_i - \overline{x})^2}$$

Contenido

Estimador de σ^2

$$\epsilon = y - \hat{y} \tag{1}$$

$$= (\beta_o + \beta_1 x + \mu) - (\hat{\beta}_0 + \hat{\beta}_1 x) \tag{2}$$

$$\epsilon = (\beta_0 - \hat{\beta}_0) - (\hat{\beta}_1 - \beta_1) + \mu \tag{3}$$

$$\overline{\epsilon} = (\beta_0 - \hat{\beta}_0) - (\hat{\beta}_1 - \beta_1) + \overline{\mu}, \quad promedio \ (3) \tag{4}$$

$$\epsilon - \overline{\epsilon} = (\mu - \overline{\mu}) - (\hat{\beta}_1 - \beta_1)(x - \overline{x}), \quad (3-4) \tag{5}$$

$$\epsilon = (\mu - \overline{\mu}) - (\hat{\beta}_1 - \beta_1)(x - \overline{x}), \quad Ec.Normal \tag{6}$$

Estimador de σ^2

Finalmente:

$$\epsilon^2 = (\mu - \overline{\mu})^2 + (\hat{\beta}_1 - \beta_1)^2 (x - \overline{x})^2 - 2(\mu - \overline{\mu})(\hat{\beta}_1 - \beta_1)(x - \overline{x})$$

Aplicando sumatoria:

$$\begin{aligned} \sum \epsilon^2 = & \sum (\mu - \bar{\mu})^2 \\ & + (\hat{\beta}_1 - \beta_1)^2 \sum (x - \bar{x})^2 \\ & - 2(\mu - \bar{\mu})(\hat{\beta}_1 - \beta_1) \sum (x - \bar{x}) \end{aligned}$$

Esperanza matemática:

$$\begin{aligned} E\left[\sum \epsilon^2\right] = & \sum E(\mu - \bar{\mu})^2 \\ & + E(\hat{\beta}_1 - \beta_1)^2 \sum (x - \bar{x})^2 \\ & - 2E(\mu - \bar{\mu})(\hat{\beta}_1 - \beta_1) \sum (x - \bar{x}) \end{aligned}$$

USAT Universidad Católica Santo Toribio de Mogrovejo

Estimador de σ^2

Se puede probar que:

$$\sum E(\mu - \bar{\mu})^2 = (n-1)\sigma^2$$

$$E(\hat{\beta}_1 - \beta_1)^2 \sum (x - \bar{x})^2 = \sigma^2$$

$$2E(\mu - \bar{\mu})(\hat{\beta}_1 - \beta_1) \sum (x - \bar{x}) = 2\sigma^2$$

Finalmente:

$$E\left[\sum \epsilon^2\right] = (n-1)\sigma^2 + \sigma^2 - 2\sigma^2$$

$$E\left[\sum \epsilon^2/(n-2)\right] = \sigma^2$$

$\hat{\sigma}^2 = \sum \epsilon^2/(n-2)$ es un estimador insesgado de σ^2.

Contenido

Caso particular: Modelo ingenuo

Función a minimizar

$$g(\alpha) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - \alpha)^2$$

CPO:

$$\frac{\partial g(\alpha)}{\partial \alpha} = \sum_{i=1}^{n} \frac{\partial \varepsilon_i^2}{\partial \alpha} = \sum_{i=1}^{n} \frac{\partial \varepsilon_i^2}{\partial \varepsilon_i} \frac{\partial \varepsilon_i}{\partial \alpha} = 0$$

C2O:

$$\frac{\partial^2 g(\alpha)}{\partial^2 \alpha} > 0$$

Contenido

Referencias

- Stock and Watson (2011). Introduction to Econometrics. Third Edition; The Addison-Wesley Series in Economics.
- P-values. *Click aquí*
- Tablas estadísticas. *Click aquí*
- Grados de Libertad. *Click aquí*
- Tests de significancia *Click aquí*
- Cameron, C. and P. Trivedi (2011). Microeconometrics Using STATA. STATA press.

└─Referencias

Referencias

- Stock and Watson (2011). Introduction to Econometrics. Third Edition; The Addison-Wesley Series in Economics.
- P-values. *Click aquí*
- Tablas estadísticas. *Click aquí*
- Grados de Libertad. *Click aquí*
- Tests de significancia *Click aquí*
- Cameron, C. and P. Trivedi (2011). Microeconometrics Using STATA. STATA press.

- Agregar alguna nota

Econometría Básica

Capítulo 02: Modelo de Regresión Lineal Simple

José Valderrama & Freddy Rojas
jtvalderrama@gmail.com & frojasca@gmail.com
Universidad Católica Santo Toribio de Mogrovejo

Septiembre de 2021