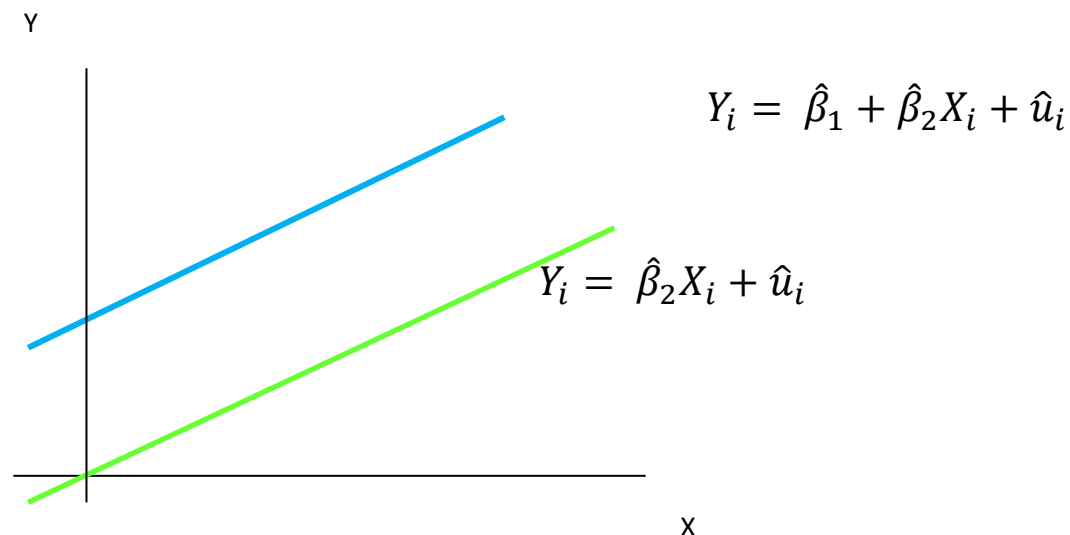


EXTENSIÓN DEL MODELO DE REGRESIÓN LINEAL DE DOS VARIABLES

REGRESIÓN A TRAVÉS DEL ORIGEN



“A MENOS QUE EXISTA UNA EXPECTATIVA A PRIORI MUY FUERTE ES ACONSEJABLE APEGARSE AL MODELO CONVENCIONAL CON PRESENCIA DE INTERSECCIÓN”

¿QUÉ OCURRE CON LOS MODELOS SIN INTERCEPTO?

- Si el modelo que se va estimar teóricamente debe llevar intersección y ésta no se incluye, se puede cometer un error de especificación y con ello violar un supuesto del MCRL, dando origen a estimadores **NO MELI**.
- La $\sum \hat{u}_i$ que siempre es cero en modelos con intercepto puede no serlo cuando está ausente $\hat{\beta}_1$. Es decir que los residuos ya no tienen media muestral igual a cero $E(u_i|X_i) \neq 0$.
- El coeficiente de determinación (R^2) que siempre es no negativo podría llegar a ser negativo, por tanto el coeficiente de determinación convencional no resulta adecuado para evaluar la bondad de ajuste del modelo. Para ello se utiliza el R^2 simple, que se interpreta de la manera ya conocida pero no se puede comparar con el R^2 tradicional.

$$R^2_{simple} = \frac{(\sum X_i Y_i)^2}{\sum X_i^2 \sum Y_i^2}$$

- Si en el modelo poblacional el intercepto es diferente de cero, y se usa un modelo muestral con $\hat{\beta}_1 = 0$, los estimadores MCO de la pendiente estarán sesgados.
- El costo de estimar $\hat{\beta}_1$ cuando realmente $\beta_1 = 0$, es que las varianzas de los estimadores MCO son mayores (no mínimas).

En resumen:

MODELO CON INTERCEPTO	MODELO SIN INTERCEPTO
$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$	$Y_i = \hat{\beta}_2 X_i + \hat{u}_i$
$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$	$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$
$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$	$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_i^2}$
$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n - 2}$	$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n - 1}$

“Si la intercepción efectivamente está ausente, el coeficiente de la pendiente puede ser estimado con mucha más precisión que cuando el término de la intersección está incluido” H. Theil (1978)

Ejemplo 1¹: Considere el modelo de asignación de precios de activos de capital (CAMP) de la teoría moderna de portafolios, expresada

$$(ER_i - r_f) = \beta_i(ER_m - r_f) \quad (2)$$

El modelo puede expresarse como:

$$R_i - r_f = \beta_i(R_m - r_f) + u_i$$

$$R_i - r_f = \alpha_i + \beta_i(R_m - r_f) + u_i \quad \text{MODELO DEL MERCADO}$$

Si se cumple el CAMP, entonces $\alpha_i = 0$

La siguiente regresión muestra los datos referentes a los rendimientos excedentes (Y_t) en %, de índice de 104 acciones del sector de bienes de consumo cíclico y los rendimientos excedentes (X_t) en % del índice de todo el mercado de valores del Reino Unido para 1980:01-1999:12

¹ Revisar Haim Levy y Marshall Sarnta, (1984). *Portfolio and Investment Selection: Theory and Practice*. Cap. 14

² ER_i : tasa esperada de rendimiento del título i . ER_m : tasa esperada del rendimientos del portafolios del mercado como se representa. r_f : tasa de rendimiento libre de riesgo. B_i : coeficiente BETA (medida de riesgo sistemático, es decir, riesgo que no se ha eliminado con la diversificación)

Dependent Variable: Y
Sample: 1980M01 1999M12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	1.155512	0.074396	15.53200	0.0000
R-squared	0.500309	Mean dependent var	0.499826	
Adjusted R-squared	0.500309	S.D. dependent var	7.849594	
S.E. of regression	5.548786	Akaike info criterion	6.269193	
Sum squared resid	7358.578	Schwarz criterion	6.283696	
Log likelihood	-751.3032	Hannan-Quinn criter.	6.275037	
Durbin-Watson stat	1.972853			

Dependent Variable: Y
Sample: 1980M01 1999M12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.447481	0.362943	-1.232925	0.2188
X	1.171128	0.075386	15.53500	0.0000
R-squared	0.503480	Mean dependent var	0.499826	
Adjusted R-squared	0.501394	S.D. dependent var	7.849594	
S.E. of regression	5.542759	Akaike info criterion	6.271160	
Sum squared resid	7311.877	Schwarz criterion	6.300165	
Log likelihood	-750.5392	Hannan-Quinn criter.	6.282847	
F-statistic	241.3363	Durbin-Watson stat	1.984746	
Prob(F-statistic)	0.000000			

¿Qué diferencias encuentra en los dos modelos?

Ejemplo 2: PIB en millones e inversión doméstica privada bruta millones (GPDIM), EEUU 1988-1997.

Dependent Variable: PIB

Sample: 1988 1997

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GPDIM	0.007259	0.001385	5.240237	0.0005
R-squared	-54.203821	Mean dependent var	6.440900	
Adjusted R-squared	-54.203821	S.D. dependent var	0.455016	
S.E. of regression	3.380733	Akaike info criterion	5.368702	
Sum squared resid	102.8642	Schwarz criterion	5.398960	
Log likelihood	-25.84351	Hannan-Quinn criter.	5.335508	
Durbin-Watson stat	0.550763			

Dependent Variable: PIB

Sample: 1988 1997

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.018338	0.245104	28.63408	0.0000
GPDIM	-0.000840	0.000318	-2.645248	0.0295
R-squared	0.466572	Mean dependent var	6.440900	
Adjusted R-squared	0.399894	S.D. dependent var	0.455016	
S.E. of regression	0.352485	Akaike info criterion	0.929238	
Sum squared resid	0.993965	Schwarz criterion	0.989755	
Log likelihood	-2.646191	Hannan-Quinn criter.	0.862851	
F-statistic	6.997338	Durbin-Watson stat	0.724787	
Prob(F-statistic)	0.029472			

¿Cuál de los dos
modelos es mejor?
¿Por qué?

ESCALAS Y UNIDADES DE MEDICIÓN

Los cambios en las escalas y en las unidades de medida se realizan con la finalidad expresar números en menores cantidades que permitan ser manejados fácilmente.

$$\text{Sea } Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i \quad (1)$$

Defina:

$$\begin{aligned} Y_i^* &= w_1 Y_i \\ X_i^* &= w_2 X_i \\ \hat{u}_i^* &= w_1 \hat{u}_i \end{aligned}$$

w_1 y w_2 Se denominan **FACTORES DE ESCALA** ($w_1 = 1$ ó $w_1 \neq w_2$)

$$Y_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* X_i^* + \hat{u}_i^* \quad (2)$$

Aplicando en método de MCO a (2) se obtiene:

$$\hat{\beta}_1^* = \bar{Y}^* - \hat{\beta}_2^* \bar{X}^*$$

$$\hat{\beta}_2^* = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}}$$

$$\text{Var}(\hat{\beta}_1^*) = \frac{\sum X_i^{*2}}{n \sum x_i^{*2}} \sigma^{*2}$$

$$Var(\hat{\beta}_2^*) = \frac{\sigma^{*2}}{\sum x_i^{*2}}$$

$$\sigma^{*2} = \frac{\sum \hat{u}_i^{*2}}{n - k}$$

Se pueden establecer las relaciones entre los dos conjuntos de parámetros estimados:

$$\hat{\beta}_2^* = \left(\frac{w_1}{w_2}\right) \hat{\beta}_2$$

$$\hat{\beta}_1^* = w_1 \hat{\beta}_1$$

$$\sigma^{*2} = w_1^2 \hat{\sigma}^2$$

$$Var(\hat{\beta}_1^*) = w_1^2 Var(\hat{\beta}_1)$$

$$Var(\hat{\beta}_2^*) = \left(\frac{w_1}{w_2}\right)^2 Var(\hat{\beta}_2)$$

$$R^2_{xy} = R^2_{x^*y^*}$$

Cambios en el origen:

$$Y_i^* = a + Y_i$$

$$X_i^* = b + X_i$$

Si ($a =$ ó $\neq b$)

$$\hat{\beta}_2^* = \hat{\beta}_2$$

$$\rho_{xy} = \rho_{x^*y^*}$$

**“LA TRANSFORMACION DE LA ESCALA
(Y, X) A LA ESCALA (Y*, X*) NO AFECTA
LAS PROPIEDADES DE LOS
ESTIMADORES MCO”**

Ejemplo 3: Inversión doméstica privada bruta (GPDIBL) en función del PIB (GDPB), ambos en miles de millones de dólares del 2000, EEUU 1990-2005.

Dependent Variable: GPDIBL
Method: Least Squares
Sample: 1990 2005
Included observations: 16

$$\rho = 0.9822$$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-926.0904	116.3577	-7.958996	0.0000
GDPB	0.253524	0.012947	19.58237	0.0000
R-squared	0.964777	Mean dependent var	1329.144	
Adjusted R-squared	0.962261	S.D. dependent var	341.9404	
S.E. of regression	66.42693	Akaike info criterion	11.34655	
Sum squared resid	61775.51	Schwarz criterion	11.44312	

Inversión doméstica privada bruta (GPDIM) en función del PIB (GDPPM) ambos en millones de dólares, EEUU 1990-2005.

Dependent Variable: GPDIM
Method: Least Squares
Sample: 1990 2005
Included observations: 16

$$\rho = 0.9822$$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-926090.4	116357.7	-7.958996	0.0000
GDPPM	0.253524	0.012947	19.58237	0.0000
R-squared	0.964777	Mean dependent var	1329144.	
Adjusted R-squared	0.962261	S.D. dependent var	341940.4	
S.E. of regression	66426.93	Akaike info criterion	25.16206	
Sum squared resid	6.18E+10	Schwarz criterion	25.25863	

REGRESIÓN SOBRE LAS VARIABLES ESTANDARIZADAS

La influencia que tiene sobre la interpretación de los efectos de las variables las unidades de medida en que éstas sean expresadas, puede evitarse si se expresan ambas variables como **variables estandarizadas**.

Una variable estandarizada se expresa como:

$$Y_i^* = \frac{Y_i - \bar{Y}}{ee_y}$$

$$X_i^* = \frac{X_i - \bar{X}}{ee_x}$$

La propiedad más resaltante de las variables estandarizadas es que el valor de su media siempre es cero y su desviación estándar es siempre 1.

$$Y_i^* \sim (0,1)$$

$$X_i^* \sim (0,1)$$

La regresión con variables estandarizadas queda:

$$Y_i^* = \beta_2^* X_i^* + u_i^*$$



¿POR QUÉ?, ¿QUÉ
PASA CON EL
COEFICIENTE DEL
INTERCEPTO?

PARA TENER EN CUENTA:

- Los coeficientes de regresión de las variables estandarizadas se conocen como coeficientes beta³.
- Después de estandarizadas las variables ya no importa las unidades de medida.
- Los coeficientes beta se interpretan en términos de desviaciones estándar, es decir, se mide el efecto en las unidades de la desviación estándar.
- Los coeficientes dentro de la misma ecuación son comparables entre sí. Se pueden utilizar como una medida de fuerza relativa de las distintas regresoras.

³ No se deben confundir estos coeficientes beta con los coeficientes beta de la teoría del portafolio.
Prof. Laura Castillo

Ejemplo 4: Regresión de inversión sobre PIB

VARIABLES ORIGINALES					VARIABLES ESTANDARIZADAS				
Matriz de correlación					Matriz de correlación				
	PIB	INV				INV	PIB		
PIB	1.000000	0.164923			INV	1.000000	0.164923		
INV	0.164923	1.000000			PIB	0.164923	1.000000		
Dependent Variable: INV Method: Least Squares					Dependent Variable: INV Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	900.2423	58.96998	15.26611	0.0000	C	-1.33E-10	0.330817	-4.02E-10	1.0000
PIB	0.001281	0.002708	0.472948	0.6489	PIB	0.164923	0.348712	0.472948	0.6489
R-squared	0.027200	Mean dependent var	916.1100		R-squared	0.027200	Mean dependent var	-1.00E-10	
Adjusted R-squared	-0.094401	S.D. dependent var	146.5929		Adjusted R-squared	-0.094401	S.D. dependent var	1.000000	
S.E. of regression	153.3561	Akaike info criterion	13.08026		S.E. of regression	1.046136	Akaike info criterion	3.104940	
Sum squared resid	188144.8	Schwarz criterion	13.14078		Sum squared resid	8.755204	Schwarz criterion	3.165457	
Log likelihood	-63.40130	F-statistic	0.223680		Log likelihood	-13.52470	F-statistic	0.223680	
Durbin-Watson stat	0.269266	Prob(F-statistic)	0.648885		Durbin-Watson stat	0.269266	Prob(F-statistic)	0.648885	

¿Cómo interpretaría los parámetros estimados de cada regresión?

Lecturas obligatorias:

- ✓ Gujarati, D. y Porter, D. (2010). *Econometría*. 5ta. Edición McGraw Hill. Capítulo 6.
- ✓ Wooldridge, J. (2010). *Introducción a la econometría un enfoque moderno*. Capítulos 6.