EXTENSIÓN DEL MODELO DE REGRESIÓN CLÁSICO: MODELO DE REGRESIÓN MÚLTIPLE

El modelo de regresión lineal múltiple se utiliza para estudiar la relación que existe entre una variable dependiente y un conjunto de variables independientes, de forma estocástica.

El modelo de interés es:

$$\boldsymbol{E(Y_i|X_{ik}) = \hat{\beta}_1 X_{i0} + \hat{\beta}_2 X_{i1} + \hat{\beta}_3 X_{i2} + \cdots + \hat{\beta}_k X_{ik} + \hat{\varepsilon}_i}$$

Coeficientes de Regresión Parcial

Donde $\boldsymbol{X_{i0} = 1 \ \forall \ i = 1, 2, \dots, n}$

$$\underset{n*1}{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}} = \underset{n*k}{\begin{bmatrix} 1 & X_{12} & X_{13} & \dots & X_{1k} \\ 1 & X_{22} & X_{23} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n2} & X_{n3} & \dots & X_{nk} \end{bmatrix}} \underset{k*1}{\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}} + \underset{n*1}{\begin{bmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \hat{\varepsilon}_n \end{bmatrix}}$$

$Y = X\hat{\beta} + \hat{\varepsilon}$ ([1]) → **Es ahora nuestro modelo de interés**

Coeficientes de regresión parcial → $\beta_k = \frac{\partial(Y_i)}{\partial(X_{ki})}$

$E(Y_i|X_{i2}, X_{i3}, \dots, X_{ik})$ PERMITE OBTENER LA MEDIA CONDICIONAL O EL VALOR PROMEDIO ESPERADO DE Y CONDICIONADO A LOS VALORES DADOS O FIJOS DE LAS VARIABLES X_{ik}

Los coeficientes de regresión parcial $(\hat{\beta}_k)$ se interpretan como:

$\hat{\beta}_2 =$ El cambio (el valor promedio) en Y producto de un cambio unitario en X_{i2}, manteniendo constante X_{i3}.

$\hat{\beta}_3 =$ El cambio (el valor promedio) en Y producto de un cambio unitario en X_{i3}, manteniendo constante X_{i2}.

NOTA: El poder del análisis de regresión múltiple es que proporciona una interpretación *ceteris paribus* aun cuando los datos no hayan sido recolectados de manera *ceteris paribus.* Bajo esta interpretación la regresión

[1] En el caso de los modelos de regresión múltiple es preferible usar la notación matricial, pues dicha forma permite expresar el modelo en una forma más compacta y que con un poco de conocimiento del algebra matricial los resultados se simplifican considerablemente.

múltiple permite hacer en un ambiente no experimental, lo que en ciencias naturales puede hacerse con experimentos controlados en un laboratorio: MANTENER CONSTANTE OTROS FACTORES.

Supuestos del MCRL múltiple:

1. Linealidad en parámetros del modelo de regresión: $\boldsymbol{Y = X\beta + \varepsilon}$

2. Los valores de las variables explicativas son fijos en repetidas muestras, no estocásticos. REGRESORES NO ESTOCÁSTICOS, X es una matriz conocida de orden *n*k*. Además,

3. Valor medio de los residuos, igual a cero: $E(\boldsymbol{\varepsilon}_i|X_{ik}) = 0$

$$E(\varepsilon|X) = \begin{bmatrix} E(\varepsilon_1|X) \\ E(\varepsilon_2|X) \\ \vdots \\ E(\varepsilon_n|X) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

4. Homoscedasticidad [2] en las perturbaciones: $Var(\varepsilon_n|X) = \sigma^2$ *para i=1,2,…,n*

$$E(\varepsilon\varepsilon'|X) = \begin{bmatrix} E({\varepsilon_1}^2|X) & E(\varepsilon_1\varepsilon_2|X) & \dots & E(\varepsilon_1\varepsilon_n|X) \\ E(\varepsilon_2\varepsilon_1|X) & E({\varepsilon_2}^2|X) & \dots & E(\varepsilon_2\varepsilon_n|X) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_n\varepsilon_1|X) & E(\varepsilon_n\varepsilon_2|X) & \dots & E({\varepsilon_n}^2|X) \end{bmatrix}$$

$$E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

$E(\varepsilon\varepsilon') = \sigma^2\mathrm{I}.$ Varianza homoscedástica

Matriz de varianzas y covarianzas

5. No correlación serial de residuos: $Cov\left(\varepsilon_i, \varepsilon_j\right) = 0$

$$E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

[2] Vea McCulloch 1985

SUPUESTOS 4 Y 5 SE DENOMINAN PERTURBACIONES ESFÉRICAS[3]

6. Covarianza entre las variables explicativas y los residuos es igual a cero: $Cov(X,\varepsilon) = Cov[X, E(\boldsymbol{\varepsilon}_i|X_{ik})] = 0$[4]

7. X es una matriz de rango completo, las columnas de X son linealmente independientes, y hay al menos k observaciones. CONDICIÓN DE IDENTIFICACIÓN $\boldsymbol{n > k}$

8. $Var(X_{ik}) \neq 0$

9. No hay sesgo o error de especificación.

10. No hay colinealidad perfecta entre las variables explicativas.

Adicional: los términos de error estocástico se distribuyen normales. $\varepsilon_i|X \sim N(0; \sigma^2 I)$

[3] El término describirá a una distribución normal multivariante. Si $\sum \sigma^2 I$ en la función de densidad normal multivariante, entonces la ecuación $f(x) = c$ es la fórmula de una esfera centrada en μ con radio σ en el espacio n dimensional. El nombre de esférica se usa si se trata de una distribución normal o si no; a veces se define explícitamente la distribución "NORMAL ESFÉRICA"

[4] El supuesto 3 implica que $Cov(X,\varepsilon) = 0$ ¿ES CIERTA LA INVERSA?

Estimación de los parámetros del modelo

Para el modelo a estimar:

$$\boldsymbol{Y_i = \widehat{\beta}_1 + \widehat{\beta}_2 X_{i2} + \widehat{\beta}_3 X_{i3} + \cdots + \widehat{\beta}_k X_{ik} + \hat{\varepsilon}_i}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{12} & X_{13} & \dots & X_{1k} \\ 1 & X_{22} & X_{23} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n2} & X_{n3} & \dots & X_{nk} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \hat{\varepsilon}_n \end{bmatrix}$$

n*1 n*k k*1 n*1

$$\boldsymbol{y = X\widehat{\beta} + \hat{\varepsilon}}$$

Lo que se busca (EL VECTOR DE COEFICIENTE DE MCO) es minimizar:

$$\boldsymbol{\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum \left(Y_i - \widehat{\beta}_1 + \widehat{\beta}_2 X_{i2} + \widehat{\beta}_3 X_{i3} + \cdots + \widehat{\beta}_k X_{ik}\right)^2 = \hat{\varepsilon}'\hat{\varepsilon}}$$

$$\hat{\boldsymbol{\varepsilon}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}$$

$$\hat{\varepsilon}'\hat{\varepsilon} = \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right)$$

$$= y'y - y'\hat{\beta}X - \hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}^{5}$$

$$= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

$$\frac{\partial(\hat{\varepsilon}'\hat{\varepsilon})}{\partial\hat{\beta}} = -2X'y + 2(X'X)\hat{\beta} = 0 \implies (X'X)\hat{\beta} = X'y$$

$\hat{\beta}$ satisface las ecuaciones normales de MCO[6]

$$(X'X)^{-1}(X'X)\hat{\beta} = (X'X)^{-1}X'y$$

$$I\hat{\beta} = (X'X)^{-1}X'y$$

$$\boldsymbol{\hat{\beta}_{MCO} = (X'X)^{-1}X'y}$$

Resultado fundamental de la teoría de MCO en notación matricial.

---

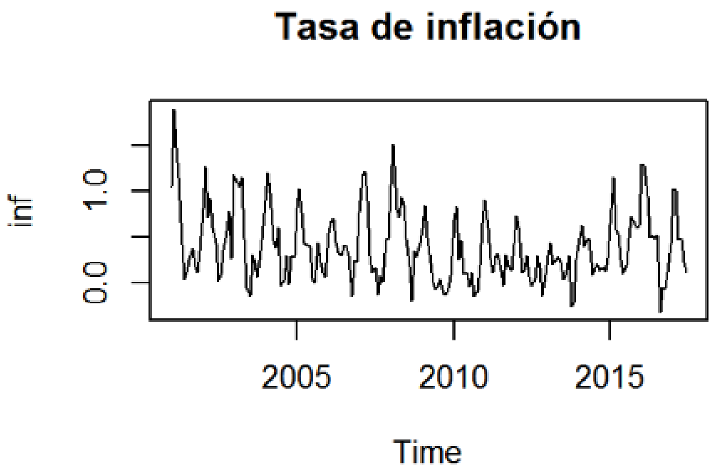
[5] Se utilizan las propiedades de la traspuesta de una matriz, a saber: $(X\hat{\beta})' = \hat{\beta}'X'$; y, como $\hat{\beta}'X'y$ es un escalar (un número real), es igual a su traspuesta $y'\hat{\beta}X$.

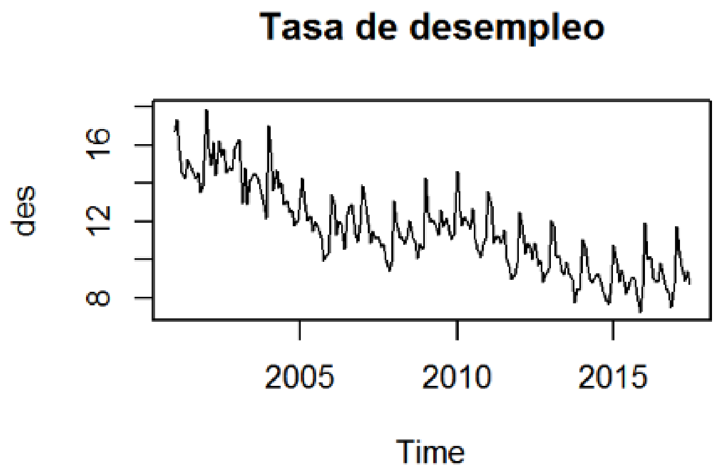
[6] **Respecto a *(X'X)*:**

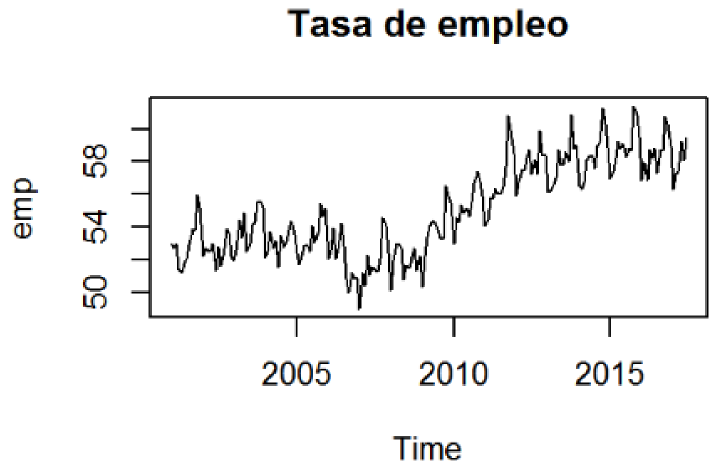
1. Proporciona las sumas simples de cuadrados y productos cruzados de las variables X, una de las cuales es el término del intercepto que toma el valor 1 para cada observación. Los elementos sobre la diagonal principal dan las sumas simples de cuadrados y los que no están en la diagonal principal dan las sumas simples de productos cruzados.
2. Es simétrica, pues el producto cruzado entre X_{i2} y X_{i3} es el mismo que entre X_{i3} y X_{i2}.
3. Es de orden ($k \times k$), es decir, tiene k renglones y k columnas.

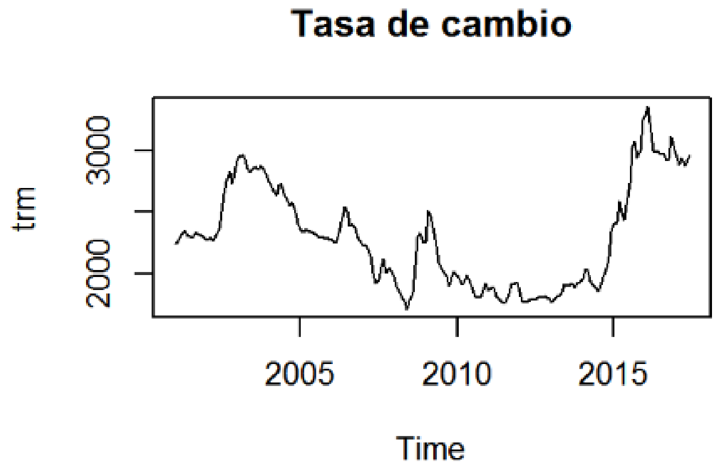
k*1 k*k (k*n)(n*1)

Ejemplo: ESTIMACIÓN DE LA TASA INFLACIÓN (inf) EN FUNCIÓN DEL DESEMPLEO (des), EL EMPLEO (emp) Y LA TASA DE CAMBIO (trm) , COLOMBIA 2001:01-2017:07.







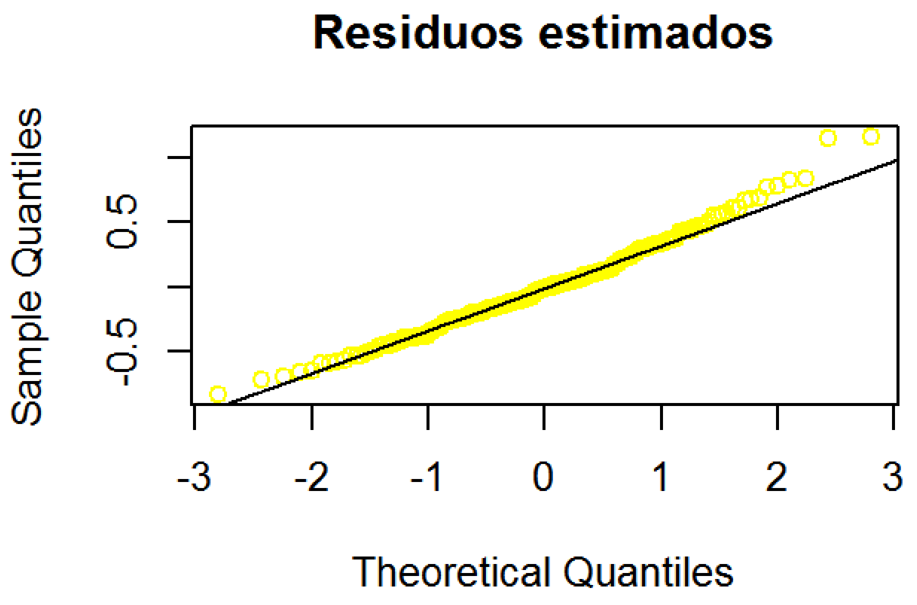


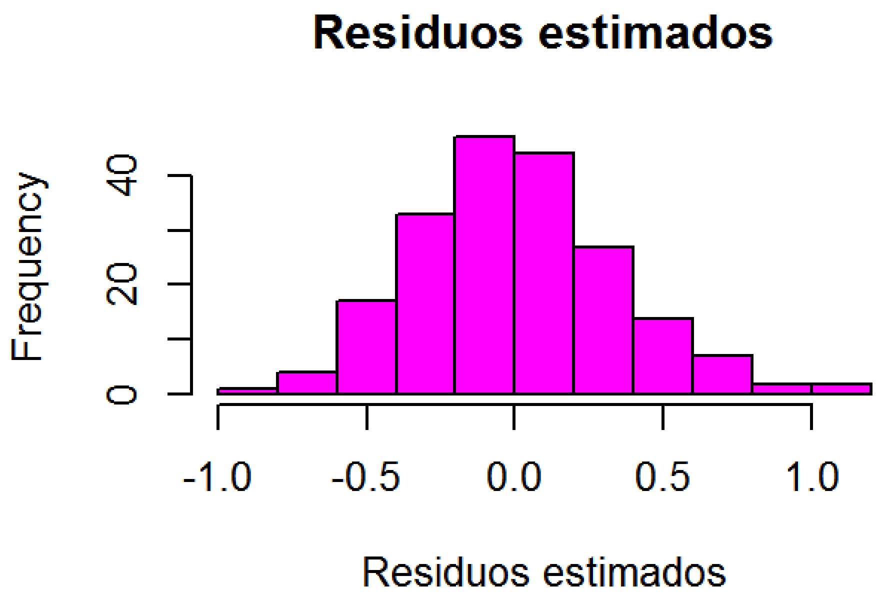
```
## Call:
## lm(formula = inf ~ des + trm + emp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.82788 -0.22876 -0.01029  0.21406  1.15809
## Coefficients:
##              Estimate  Std. Error   t value    Pr(>|t|)
## (Intercept) -4.891e-01   8.734e-01   -0.560  0.576125
## des          5.733e-02   1.715e-02    3.344  0.000992 ***
## trm          1.976e-04   6.239e-05    3.167  0.001792 **
## emp         -4.071e-03   1.330e-02   -0.306  0.759842
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.352 on 194 degrees of freedom
## Multiple R-squared:  0.1904, Adjusted R-squared:  0.1779
## F-statistic: 15.21 on 3 and 194 DF,  p-value: 6.272e-09
```

¿Qué puede usted concluir de ese modelo?

NOTA: trabaje con un nivel de significancia del 1%

```
##   Shapiro-Wilk normality test
##
## data:  res
## W = 0.98571, p-value = 0.04262
```





MATRIZ DE VARIANZAS Y COVARIANZAS DE $\hat{\boldsymbol{\beta}}_{MCO}$

Por definición, la matriz de varianza-covarianza de $\hat{\boldsymbol{\beta}}$ es:

$$Var-Cov(\hat{\beta}) = E\{[\hat{\beta}-E(\hat{\beta})][\hat{\beta}-E(\hat{\beta})]'\}$$

Así:

$$Var-Cov(\hat{\beta}) = \begin{bmatrix} var(\hat{\beta}_1) & Cov(\hat{\beta}_1,\hat{\beta}_2) & \dots & Cov(\hat{\beta}_1,\hat{\beta}_k) \\ Cov(\hat{\beta}_2,\hat{\beta}_1) & var(\hat{\beta}_2) & \dots & Cov(\hat{\beta}_2,\hat{\beta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\hat{\beta}_k,\hat{\beta}_1) & Cov(\hat{\beta}_k,\hat{\beta}_2) & \dots & var(\hat{\beta}_k) \end{bmatrix}$$

La cual se obtiene de: $Var-Cov(\hat{\beta}) = \sigma^2(X'X)^{-1}$

El estimador insesgado de σ^2 está dado por: $\hat{\sigma}^2 = \frac{1}{n-k}\hat{\varepsilon}'\hat{\varepsilon}$

Las sumas de cuadrados asociadas al modelo son:

$SCT = y'y - n\bar{Y}^2$ $SCR = \hat{\beta}'X'y - n\bar{Y}^2$ $SCE = y'y - \hat{\beta}'X'y$

PROPIEDADES ESTADÍSTICAS DEL ESTIMADOR $\hat{\beta}$

Los estimadores de MCO son los mejores estimadores lineales insesgados (MELI). Esta propiedad se extiende a todo el vector $\hat{\boldsymbol{\beta}}$.

1. **Linealidad:** Como $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ es una matriz de números fijos, $\hat{\boldsymbol{\beta}}$ es una función lineal de Y. Por tanto, por definición, es un estimador lineal.

$$\hat{\boldsymbol{\beta}}_{MCO} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}$$

2. **Insesgabilidad:**

$$\boldsymbol{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{E}[\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}] = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{E}(\boldsymbol{\varepsilon}) = \boldsymbol{\beta}$$

3. **Eficiencia:** Sea $\hat{\boldsymbol{\beta}}^*$ cualquier otro estimador lineal de **β**, el cual se escribe como $\hat{\boldsymbol{\beta}}^* = [(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' + \mathbf{C}]\mathbf{y}$ donde **C** es una matriz de constantes. Al sustituir **y, se obtiene**: $\hat{\beta}^* = [(X'X)^{-1}X' + \mathrm{C}](X\beta + \varepsilon)$, resulta

$$\hat{\boldsymbol{\beta}}^* = \boldsymbol{\beta} + \boldsymbol{C}\boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} + \boldsymbol{C}\boldsymbol{\varepsilon}$$

Como $\widehat{\boldsymbol{\beta}}^*$ es un estimador insesgado de $\boldsymbol{\beta}$, entonces:

$$\widehat{\boldsymbol{\beta}}^* - \boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} + \boldsymbol{C}\boldsymbol{\varepsilon}$$

Por definición, la matriz de varianza y covarianza de $\widehat{\boldsymbol{\beta}}^*$ resulta:

$$\boldsymbol{E}\left[\left(\widehat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}\right)\left(\widehat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}\right)'\right] = \boldsymbol{E}[(\,(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} + \boldsymbol{C}\boldsymbol{\varepsilon})(\,(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} + \boldsymbol{C}\boldsymbol{\varepsilon})']$$

$$\boldsymbol{Var} - \boldsymbol{Cov}\left(\widehat{\boldsymbol{\beta}}^*\right) = \boldsymbol{\sigma}^2\,(\boldsymbol{X}'\boldsymbol{X})^{-1} + \boldsymbol{\sigma}^2\boldsymbol{C}\boldsymbol{C}'$$

$$\boldsymbol{Var} - \boldsymbol{Cov}\left(\widehat{\boldsymbol{\beta}}^*\right) = \boldsymbol{Var} - \boldsymbol{Cov}\left(\widehat{\boldsymbol{\beta}}\right) + \boldsymbol{\sigma}^2\boldsymbol{C}\boldsymbol{C}'$$

La matriz de varianza-covarianza del estimador lineal e insesgado alterno $\widehat{\boldsymbol{\beta}}^*$ es igual a la matriz de varianza-covarianza del estimador MCO, $\widehat{\boldsymbol{\beta}}$ más $\boldsymbol{\sigma}^2$ veces $\boldsymbol{C}\boldsymbol{C}'$, que es una matriz semidefinida positiva. Por tanto, las varianzas de un elemento dado de $\widehat{\boldsymbol{\beta}}^*$ deben ser necesariamente iguales o mayores al elemento correspondiente de $\widehat{\boldsymbol{\beta}}$, lo cual demuestra que $\widehat{\boldsymbol{\beta}}$ es MELI. Por supuesto, si $\boldsymbol{C}$ es una matriz nula, es decir, $\boldsymbol{C}$= 0, entonces $\widehat{\boldsymbol{\beta}}^* = \widehat{\boldsymbol{\beta}}$, lo que equivale a decir que si encontramos un estimador MELI, éste debe ser el estimador de mínimos cuadrados $\widehat{\boldsymbol{\beta}}$.

En resumen, los estimadores de MCO son los mejores estimadores lineales insesgados (MELI). Esta propiedad se extiende a todo el vector $\widehat{\boldsymbol{\beta}}$; es decir, $\widehat{\boldsymbol{\beta}}$ es lineal (cada uno de sus elementos es una función lineal de Y, la variable dependiente). $E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$, es decir, el valor esperado de cada elemento de $\widehat{\boldsymbol{\beta}}$ es igual al elemento correspondiente de la verdadera $\boldsymbol{\beta}$, y en la clase de todos los estimadores lineales e insesgados de $\boldsymbol{\beta}$, el estimador de MCO, $\widehat{\boldsymbol{\beta}}$ tiene varianza mínima.

Esto significa que se cumple el TEOREMA DE GAUSS- MARKOV

BONDAD DE AJUSTE DEL MODELO DE REGRESIÓN MÚLTIPLE

Partiendo del modelo:

$$\boldsymbol{y} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} + \hat{\boldsymbol{\varepsilon}}$$

$$R^2 = \frac{\widehat{\beta}\prime X\prime y - n\bar{Y}^2}{y\prime y - n\bar{Y}^2} \qquad ó \qquad R^2 = 1 - \frac{y\prime y - \widehat{\beta}\prime X\prime y}{y\prime y - n\bar{Y}^2}$$

PRUEBA DE HIPÓTESIS DE SIGNIFICANCIA INDIVIDUAL

NOTA: $\varepsilon_i \sim N(0; \sigma^2 I)$

Con el supuesto de normalidad, se sabe que en los modelos de regresión lineal:

1. El estimador de MCO $\widehat{\boldsymbol{\beta}}$ y el estimador de MV $\widetilde{\boldsymbol{\beta}}$ son idénticos, pero el estimador de MV $\widetilde{\boldsymbol{\sigma}}^{\mathbf{2}}$es sesgado, aunque este sesgo se elimina mediante el estimador de MCO insesgado de $\widehat{\boldsymbol{\sigma}}^{\mathbf{2}}$**.**

2. Los estimadores de MCO $\widehat{\boldsymbol{\beta}}$ también están normalmente distribuidos. Para generalizar, en el caso de *k* variables es posible demostrar que:

$$\hat{\beta} \sim N(\beta; \sigma^2 \, (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}})$$

Para las perturbaciones distribuidas normalmente, $\widehat{\boldsymbol{\beta}}$ es un estimador insesgado que es una función de estadísticos suficientes. Lo que conduce a una forma más fuerte del Teorema de Gauss-Markov, como lo es el TEOREMA DE RAO – BLACKWELL[7].

[7] Los estimadores insesgados que son funciones lineales de estadísticos suficientes son eficientes entre la clase de todos los estimadores insesgados .

TEOREMA DE RAO – BLACKWELL: en el MCRL con perturbaciones distribuidas normalmente, el estimador MCO de $\widehat{\beta}$ tiene varianza mínima dentro de todos los estimadores insesgados.

La prueba de significancia individual de los parámetros se realiza igual que para un modelo de regresión simple.

$H_0: \beta_i = \beta$
$H_1: \beta_i \neq \beta$

Estimador puntual:

$t_c = \frac{(\widehat{\beta}_i - \beta_i)}{ee(\widehat{\beta}_i)} \sim t_{n-k}$ **RECHAZAR Ho SI $t_c > t_\alpha$**

Intervalo de confianza:

$\hat{\beta}_i \pm t_{\alpha/2} ee(\hat{\beta}_i)$

RECHAZAR Ho SI β NO ESTÁ CONTENIDO EN EL INTERVALO DE CONFIANZA

Ejemplo: Datos macroeconómicos del Economic Report Of The President, U.S Government Printing Office, 1968-1982. Y= Inversión real, GDP= Producto interno bruto real, R= tasa de interés promedio anual FED y P= tasa de inflación

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-5.091e-01	5.513e-02	-9.234	3.28e-06 ***
GDP	6.704e-01	5.500e-02	12.189	2.52e-07 ***
R	-2.326e-03	1.219e-03	-1.908	0.0854 .
P	-9.401e-05	1.347e-03	-0.070	0.9458
T	-1.658e-02	1.972e-03	-8.409	7.59e-06 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Los intervalos de confianza:

	2.5 %	97.5 %
(Intercept)	-0.631902939	-0.3862386432
GDP	0.547842006	0.7929248689
R	-0.005041735	0.0003898781
P	-0.003096384	0.0029083627
T	-0.020973752	-0.0121870369

PRUEBA DE SIGNIFICANCIA CONJUNTA O SIGNIFICANCIA GLOBAL

Busca probar si todos los coeficientes de regresión son en conjunto estadísticamente significativos, es decir, si Y está relacionada o no linealmente con las X_i a la vez.

$H_0: \beta_2 = \cdots = \beta_i = 0$
H_1**:** $al\ menos\ un\ \beta_i \neq 0$

Estadístico de prueba:

$$\boldsymbol{F_c = \frac{SCR/(k-1)}{SCE/(n-k)} \sim F_{k-1,n-k}}$$

Es también conocida como un test de significancia del R^2:

$$F_c = \frac{SCR/(k-1)}{SCE/(n-k)} = \frac{CM_R}{CM_E}$$

$$F_c = \frac{R^2}{(1-R^2)} * \frac{(n-k)}{(k-1)} \sim F_{k-1,n-k}$$

Equivalente a: $\boldsymbol{F_c = \frac{\hat{\beta}\prime X\prime y - n\bar{Y}^2/(k-1)}{y\prime y - n\bar{Y}^2/(n-k)} \sim F_{k-1,n-k}}$

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-5.091e-01	5.513e-02	-9.234	3.28e-06	***
GDP	6.704e-01	5.500e-02	12.189	2.52e-07	***
R	-2.326e-03	1.219e-03	-1.908	0.0854	.
P	-9.401e-05	1.347e-03	-0.070	0.9458	
T	-1.658e-02	1.972e-03	-8.409	7.59e-06	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.006714 on 10 degrees of freedom

Multiple R-squared: 0.9724, Adjusted R-squared: 0.9614

F-statistic: 88.19 on 4 and 10 DF, **p-value: 9.333e-08**

PREDICCIÓN MEDIANTE REGRESIÓN MÚLTIPLE

Después de la estimación el uso más habitual de la regresión consiste en la predicción. La regresión múltiple estimada predice:

1. La media

Sea $X_0 = \begin{bmatrix} 1 \\ X_{01} \\ X_{02} \\ \vdots \\ X_{0k} \end{bmatrix}$ el vector de variables de X para los cuales se quiere predecir $\widehat{Y}_0$, la predicción media de Y,

$\boldsymbol{\widehat{Y}_\iota = X\widehat{\beta}}$ **es la predicción media**

$$\left(\boldsymbol{\widehat{Y}_\iota} \middle| \boldsymbol{X'_0}\right) = \boldsymbol{X'_0\widehat{\beta}}$$

Los valores de X'_0 están especificados; la predicción media es insesgada

$$\boldsymbol{E(Y_i|X'_0) = E(X'_0\widehat{\beta}) = X'_0\beta}$$

La varianza de la predicción media es

$$Var\left(\widehat{Y_0}\middle|X'_0\right) = \sigma^2 X'_0\,(X'X)^{-1}X_0$$

2. Los valores individuales de Y, dados los valores de las regresoras X

La predicción individual de Y(= Y0) también está dada por $\widehat{Y}_\iota = X\widehat{\beta}$ La diferencia entre las predicciones de la media y la individual consiste en sus varianzas.

La varianza, $Var(Y_0|X_0) = \sigma^2[1 + X'_0\,(X'X)^{-1}X_0]$

Donde $Var(Y_0|X_0)$ representa a $E\left(Y_0 - \widehat{Y}_0\middle|X_0\right)^2$

Ejemplo: series de tiempo económicas de Estados Unidos de interés general. El periodo que abarcan estas cifras trimestrales es de I-1947 a IV-2007.

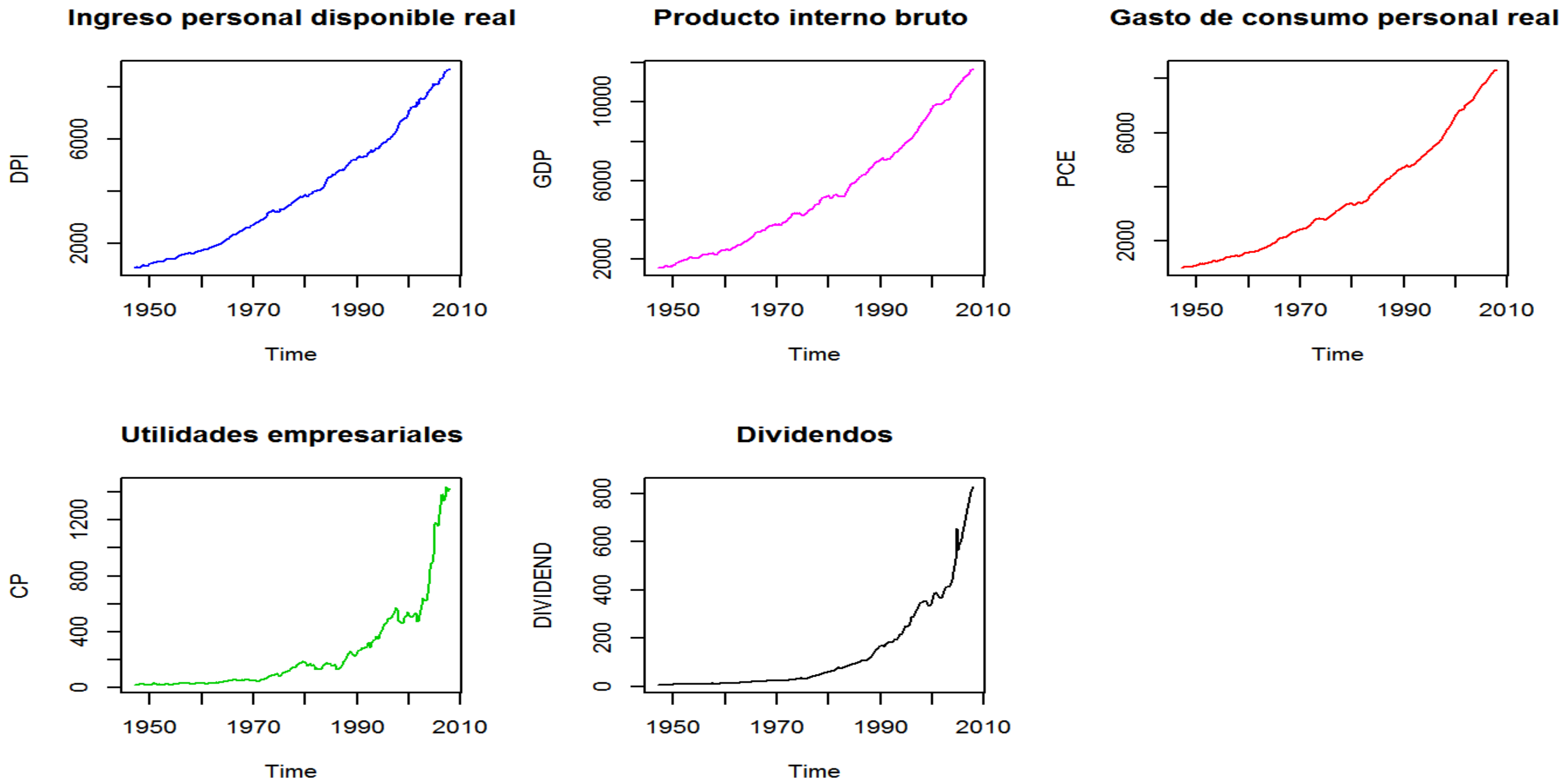
DPI = ingreso personal disponible real (miles de millones de dólares)

GDP = producto interno bruto (miles de millones de dólares)

PCE = gasto de consumo personal real (miles de millones de dólares)

CP = utilidades empresariales (miles de millones de dólares)

DIVIDEND = dividendos (miles de millones de dólares)



1. **Modelo Log-lineal o log-log:** son modelos intrínsecamente lineales en parámetros. Tienen la propiedad que miden elasticidades directamente, suponiendo que el coeficiente de elasticidad permanece constante a través del tiempo por ello son conocidos como ***modelos de elasticidad constante.***

$$Y_i = \beta_1 {X_{i1}}^{\beta_2} {X_{i2}}^{\beta_3} e^{\varepsilon_i}$$

$$lnY_i = ln\beta_1 + \beta_2 lnX_{i1} + \beta_3 lnX_{i2} + \varepsilon_i$$

$$lnY_i = \alpha_1 + \beta_2 lnX_{i1} + \beta_3 lnX_{i2} + \varepsilon_i$$

- Puede ser estimado por regresión MCO.
- $\hat{\alpha}_1$ y $\hat{\beta}_2$ son estimadores MELI. Sin embargo, α_1 es por naturaleza un estimador sesgado, pero recuérdese que en la mayor parte de los problemas el término intersección es de importancia secundaria.
- $\hat{\beta}_2$ mide $\frac{\Delta\%Y}{\Delta\%X}$
- La forma más sencilla para decidir si el modelo a utilizar el log-log es haciendo un gráfico de dispersión en logaritmo y observar si las observaciones se ajustan a una línea recta, cuando se trata de dos variables. Puede realizar una prueba de MWD.

```
## Call:
## lm(formula = log(DPI) ~ log(GDP) + log(PCE) + log(CP) + log(DIVIDEND))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.043846 -0.012105 -0.001805  0.008666  0.058197
##
## Coefficients:
##                Estimate Std. Error   t value   Pr(>|t|)
## (Intercept)   -1.155607   0.085728   -13.480   < 2e-16 ***
## log(GDP)       0.364006   0.061716     5.898   1.25e-08 ***
## log(PCE)       0.814895   0.063297    12.874   < 2e-16 ***
## log(CP)       -0.021092   0.007114    -2.965   0.00333 **
## log(DIVIDEND) -0.056681   0.007379    -7.681   4.04e-13 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.01885 on 239 degrees of freedom
## Multiple R-squared:  0.9991, Adjusted R-squared:  0.999
## F-statistic: 6.378e+04 on 4 and 239 DF,  p-value: < 2.2e-16
```

2. **Modelo Log-lin:** se caracterizan porque la variable dependiente se encuentra en logaritmo. Se utilizan preferiblemente para medir tasas de crecimiento de las variables.

$$lnY_t = \beta_1 + \beta_2 X_{i1} + \beta_3 X_{i2} + \beta_4 t + \varepsilon_t$$

- $\hat{\beta}_2$ mide el cambio relativo o proporcional constante en la regresada ante un cambio absoluto en la regresora.

- $\hat{\beta}_2$ multiplicado por 100 permite hallar el cambio porcentual en Y o la tasa de crecimiento que en la literatura se conoce como una ***semielasticidad.***

- $\hat{\beta}_2$ multiplicado por 100 mide la tasa de crecimiento instantánea de la variable dependiente. Para medir la tasa de crecimiento compuesta (TCC) se aplica el antilogaritmo a $\hat{\beta}_2$, se resta 1 y se multiplica por 100.

$$TCC = \left(antilog\hat{\beta}_2 - 1\right) * 100$$

- Para estimar un modelo de tendencia lineal se regresa la variable original sin logaritmo contra en tiempo.

```
## Call:
## lm(formula = log(DPI) ~ GDP + PCE + CP + DIVIDEND + Año)
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.079529 -0.026139  0.000867  0.024321  0.102753

## Coefficients:
##               Estimate   Std. Error   t value    Pr(>|t|)
## (Intercept) -7.040e+01   2.424e+00    -29.041    < 2e-16 ***
## GDP         -2.687e-05   2.958e-05    -0.908     0.365
## PCE          1.700e-05   3.826e-05    0.444      0.657
## CP           2.005e-04    4.975e-05    4.030      7.51e-05 ***
## DIVIDEND    -6.647e-04   1.281e-04    -5.189     4.52e-07 ***
## Año          3.978e-02   1.252e-03    31.781     < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.03563 on 238 degrees of freedom
## Multiple R-squared:  0.9967, Adjusted R-squared:  0.9966
## F-statistic: 1.425e+04 on 5 and 238 DF,  p-value: < 2.2e-16
```

3. Modelo lin-log: este tipo de modelo mide los cambios absolutos en Y producto de un cambio porcentual en X.

$$Y_t = \beta_1 + \beta_2 lnX_{i1} + \beta_3 lnX_{i2} + \varepsilon_t$$

- $\hat{\beta}_2$ divido entre 100 permite obtener una ***semielasticidad.***

```
## Call:
## lm(formula = DPI ~ log(GDP) + log(PCE) + log(CP) + log(DIVIDEND))
## Coefficients:
##                  Estimate   Std. Error  t value      Pr(>|t|)
## (Intercept)       4522.6      1206.1      3.750       0.000222 ***
## log(GDP)         -6498.4      868.3      -7.484      1.37e-12 ***
## log(PCE)          5922.4      890.5       6.651      1.97e-10 ***
## log(CP)            245.3      100.1       2.451      0.014972 *
## log(DIVIDEND)     1397.6      103.8       13.463     < 2e-16 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 265.2 on 239 degrees of freedom
## Multiple R-squared:  0.9857, Adjusted R-squared:  0.9855
## F-statistic:  4130 on 4 and 239 DF,  p-value: < 2.2e-16
```

4. Modelos Recíprocos:

$$Y_i = \beta_1 + \beta_2 \left(\frac{1}{X_i}\right) + u_i$$

- A medida que X aumenta indefinidamente, el término $\beta_2 \left(\frac{1}{X_i}\right)$ se acerca a cero y Y se aproxima al valor límite asintótico β_1

```
## Call:
## lm(formula = DPI ~ G + P + C + D)
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.131e+04  2.704e+02  41.829  < 2e-16 ***
## G           -6.942e+06  5.782e+06  -1.201    0.231
## P           -2.826e+07  3.676e+06  -7.687  3.9e-13 ***
## C            1.360e+04  1.363e+04   0.998    0.319
## D            1.382e+05  8.137e+03  16.988  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 668.1 on 239 degrees of freedom
## Multiple R-squared:  0.9095, Adjusted R-squared:  0.908
## F-statistic: 600.5 on 4 and 239 DF,  p-value: < 2.2e-16
```

5. Modelos de Regresión Polinomial: son aquellos modelos donde, por alguna razón teórica, una o varias variables explicativas están elevadas a alguna potencia superior o igual a dos.

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 {X_i}^2 + \varepsilon_i$$

```
## Call:
## lm(formula = CP ~ GDP + R)
##
## Coefficients:
##              Estimate  Std. Error    t value   Pr(>|t|)
## (Intercept)  1.759e+02  2.374e+01   7.408       2.16e-12 ***
## GDP         -9.502e-02  9.066e-03  -10.481      < 2e-16 ***
## R            1.547e-05  7.206e-07   21.468      < 2e-16 ***
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 87.32 on 241 degrees of freedom
## Multiple R-squared:  0.9237, Adjusted R-squared:  0.9231
## F-statistic:  1459 on 2 and 241 DF,  p-value: < 2.2e-16
```

Dependent Variable: COSTO				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	222.3833	23.48780	9.468037	0.0000
PROD	-8.025000	9.809494	-0.818085	0.4403
PROD^2	2.541667	0.869084	2.924534	0.0222
R-squared	0.928389	Mean dependent var		276.1000
Adjusted R-squared	0.907928	S.D. dependent var		65.81363
S.E. of regression	19.97004	Akaike info criterion		9.069668

Dependent Variable: COSTO				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	141.7667	6.375322	22.23678	0.0000
PROD	63.47766	4.778607	13.28372	0.0000
PROD^2	-12.96154	0.985665	-13.15005	0.0000
PROD^3	0.939588	0.059106	15.89677	0.0000
R-squared	0.998339	Mean dependent var		276.1000
Adjusted R-squared	0.997509	S.D. dependent var		65.81363
S.E. of regression	3.284911	Akaike info criterion		5.505730

Dependent Variable:W				
Method: Least Squares				
Included observations: 1500				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	586885.8	208927.2	2.809044	0.0050
EXPER	81919.2	8904.833	9.204911	0.0000
$EXPER^2$	-574.9295	91.36386	-6.292745	0.0000
R-squared		Mean dependent var		2215022.
Adjusted R-squared	0.451671	S.D. dependent var		551425.7
Sum cuadrado total	5.46E+14	Akaike info criterion		28.68030
Sum squared resid	2.99E+14	Schwarz criterion		28.69559
Log likelihood	-25749.91	F-statistic		370.6459
Durbin-Watson stat	2.052692	Prob(F-statistic)		0.000000

Lectura obligatoria:

- ✓ Novales, A. (1993). Econometría. 2da. Edición. McGraw Hill Interamericana.
- ✓ Gujarati, D. y Porter, D. (2010). Econometría. 5ta Edición. McGraw Hill