PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS SOCIALES
ESPECIALIDAD DE ECONOMÍA
CICLO 2019-2

Curso: **Econometría 1**
Profesor: **Luis García (lgarcia@pucp.edu.pe)**
Jefe de práctica: **Yasmeen Destre (y.destre@pucp.pe)**

PRÁCTICA CALIFICADA 1

I. CONCEPTOS (7 pts.)

1. **¿En qué consiste el método de Mínimos Cuadrados Ordinarios? (1 pts.)**

Dado que la recta que mejor se ajusta a los datos será aquella que presente la menor suma de cuadrados de los residuos, el método de mínimos cuadrados ordinarios consiste en escoger los valores $\hat{\beta}_1$ y $\hat{\beta}_2$ tal que se minimice la SCR.

2. **Especifique el supuesto clásico bajo el cual el estimador de MCO es insesgado. Interprete la insesgadez de un estimador. (2 pts.)**

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i Y_i}{\sum x_i^2} = \frac{\sum x_i(\beta_1 + \beta_2 X_i + u_i)}{\sum x_i^2} = \frac{\beta_1 \sum x_i}{\sum x_i^2} + \frac{\beta_2 \sum x_i X_i}{\sum x_i^2} + \frac{\sum x_i u_i}{\sum x_i^2}$$

Dado que $\sum x_i = 0$ y que $\sum x_i X_i = \sum x_i^2$:

$$\hat{\beta}_2 = \beta_2 + \frac{\sum x_i u_i}{\sum x_i^2}$$

Tomando el valor esperado:

$$E[\hat{\beta}_2] = E\left[\beta_2 + \frac{\sum x_i u_i}{\sum x_i^2}\right] = \beta_2 + E\left[\frac{\sum x_i u_i}{\sum x_i^2}\right]$$

<u>El supuesto 4 indica que X es fija en muestras repetidas.</u> Por lo tanto:

$$E[\hat{\beta}_2] = \beta_2 + \frac{\sum x_i E[u_i]}{\sum x_i^2}$$

Por el supuesto 2, $E[u_i] = 0$. Así, se comprueba la insesgadez de $\hat{\beta}_2$:

$$E[\hat{\beta}_2] = \beta_2$$

3. **Qué busca explicar el R^2 y cuál es la importancia de este criterio. (2 pts.)**

$$R = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

El R2 toma valores entre 0 y 1.
Interpretación: Es una medida de bondad de ajuste, lo que las variaciones de X podrían explicar de las variaciones de Y. Cuando se acerca a 0 el ajuste es bajo (SCE tiende a cero) y cuando se acerca a 1 el ajuste es alto (SCR tiende a 0). Cuando el R2 es igual a 1, la relación entre X e Y es determinística y todos los puntos de la muestra caen en la recta de regresión, así la variable X explica perfectamente a Y.

4. **Sea *niños* la cantidad de hijos que ha tenido una pareja y *educación* los años de educación que tiene la pareja. Un modelo sencillo para relacionar fertilidad con años de educación es[1]:**

$$\boldsymbol{niños = \beta_0 + \beta_1 educación + u}$$

Donde *u* es el error no observado. ¿Qué tipo de factores son los contenidos en *u*? ¿Es posible que estos factores estén correlacionados con el nivel de educación? (2 pts.)

Los factores que pueden estar contenidos en el error podrían ser factores no observables que influyan en la educación como culturales, religiosos, entre otros. Ahora, bajo el modelo de MCO, recordemos que el error no observado no se encuentra en ninguna forma relacionado con las variables explicativas ni dependiente. Entonces, podemos concluir que no es posible que estos factores estén correlacionados con el nivel de educación. Escenario contrario ocurría si se levantan los supuestos básicos y principales del modelo MCO.

II. EJERCICIOS (13 pts)

5. **Carlos está interesado en estudiar el efecto de los fertilizantes (X) sobre las cosechas de maíz (Y). La variable "fertilizantes" solo cuenta con dos valores: 1 si se emplea abundante fertilizante en sus tierras en Cusco en el año 2019 y 0 si no emplea fertilizantes. Por su parte la variable "cosechas" también cuenta con dos valores: 30 toneladas si hubo una mala cosecha y 100 si hubo una buena cosecha. En la siguiente tabla contamos con las probabilidades conjuntas de estas variables resumidas[2]:**

	Sin fertilizantes	**Fertilizantes**	**Pr(Y)**
Mala cosecha (Y=30)	**0.35**	**0.08**	**0.43**
Buena cosecha (Y=100)	**0.15**	**0.42**	**0.57**
Pr(X)	**0.50**	**0.50**	**1**

a. **¿Cuáles son las probabilidades condicionales de Y dado X? (1 pto).**

[1] Wooldridge, J. M. (2006). *Introducción a la econometría: un enfoque moderno*. Editorial Paraninfo.
[2] García Nuñez, L. (2015). *Econometría 1.* Fondo Editorial de la Pontificia Universidad Católica del Perú.

$$Pr\left(\frac{Y}{X}\right) = \frac{P(Y \cap X)}{P(X)}$$
$$Pr(Y = 30 \mid X = 0) = 0.7$$
$$Pr(Y = 30 \mid X = 1) = 0.16$$
$$Pr(Y = 100 \mid X = 0) = 0.3$$
$$Pr(Y = 100 \mid X = 1) = 0.84$$

b. **Calcule las esperanzas condicionales de las cosechas de maíz dadas los fertilizantes (1 pto).**

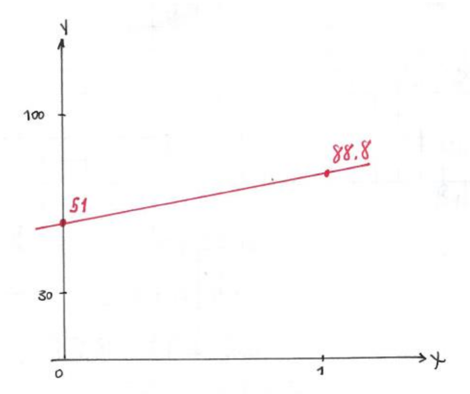
$$E(Y|X = 1) = Pr\ (Y = 30) + Pr\ (Y = 100)$$
$$= 30\ (0.16) + 100\ (0.84)$$
$$= 88.8$$

c. **Grafique estas esperanzas condicionales. ¿Podría deducir de su gráfico qué valores tomará los parámetros poblacionales $\boldsymbol{\beta_1}$ y $\boldsymbol{\beta_2}$, si modelamos esta relación entre variables como $\boldsymbol{Y_i = \beta_1 + \beta_2 X_i + u_i}$? (2 pts).**

$$E(Y|X = 0) = Pr\ (Y = 30) + Pr\ (Y = 100)$$
$$= 30\ (0.7) + 100\ (0.3)$$
$$= 51$$

$$y_i = \beta_1 + \beta_2 x_i + u_i$$
$$E[y_i|x = 0] = \beta_1 + \beta_2 x_i$$
$$Como\ x_i = 0$$
$$\beta_1 = 51$$

$$E[y_i|x = 1] = \beta_1 + \beta_2 x_i$$
$$Como\ x_i = 1$$
$$88.8 = 51 + \beta_2$$
$$\beta_2 = 37.8$$



d. **Calcule la varianza condicional y la covarianza entre X e Y. (2 pts).**

$$Var\ (Y|X=x_i)=\sum_{i=1}^{m}(y_i-E[Y|X=x_i])^2\ Pr\ (U=y_i|X=x_i)$$

$$(30x0.3)^2\ x\ 0.7+\ (100x0.7)^2\ x\ 0.3+\ (30x0.84)^2\ x\ 0.16+\ (100x0.16)^2\ x\ 0.84$$

$$=1843.34$$

$$Cov\ (Y|X=x_i)=\sum_{j=1}^{n}\ \sum_{i=1}^{m}\Big((y_i-E[Y])\big(x_j-E[x]\big)\Big)\ Pr\ (Y=y_i\ ;X=x_i)-\frac{\sum_i\sum_j x_i y_j n_{ij}}{N}-\bar{x}\,\bar{y}$$

$$Cov\ (Y|X=x_i)=9.45$$

6. **Dado el siguiente estimador de la pendiente del modelo $\boldsymbol{Y_i=\beta_1+\beta_2X_i+u_i}$**

$$\boldsymbol{b_2=\frac{\sum\left(\frac{y_i}{x_i}\right)}{n}}$$

en donde $\boldsymbol{x_i}$ y $\boldsymbol{y_i}$ están en desvíos respecto a la media, mostrar lo siguiente:

a. **$\boldsymbol{b_2}$ es insesgado (3 puntos)**

$$b_2=\frac{\sum(\frac{y_i}{x_i})}{n}=\frac{\sum(\frac{\beta_1+\beta_2X_i+u_i-\beta_1-\beta_2\bar{X}_i-\bar{u}}{x_i})}{n}=\frac{\sum(\frac{\beta_2(X_i-\bar{X}_i)+(u_i-\bar{u})}{x_i})}{n}$$

$$=\beta_2+\frac{1}{n}\sum\frac{(u_i-\bar{u})}{x_i}$$

$$E\ (b_2)=\beta_2\ +\frac{1}{n}E\ \left[\sum(\frac{u_i-\bar{u}}{(X_i-\bar{X}_i)})\right]=$$

$$=\beta_2+\frac{1}{n}\sum\left(\frac{E(u_i)-E(\bar{u})}{(X_i-\bar{X})}\right)=\beta_2$$

b. **$\boldsymbol{b_2}$ es lineal en $\boldsymbol{Y_i}$ (3 puntos)**

$$b_2=\frac{\sum_{i=1}^{n}(\frac{Y_i-\bar{Y}}{X_i-\bar{X}})}{n}$$

Hay que darle la forma de una combinación lineal de Y_i.

$$\mathrm{b}_2=\frac{1}{n}\sum\left(\frac{Y_i-\bar{Y}}{x_i}\right)=\frac{1}{n}\sum\left(\frac{Y_i}{x_i}\right)-\frac{1}{n}\sum\left(\frac{\bar{Y}}{x_i}\right)$$

$$= \frac{1}{n}\sum\left(\frac{1}{x_i}\right)Y_i - \frac{1}{n}\bar{Y}\sum\left(\frac{1}{x_i}\right) = \frac{1}{n}\sum\left(\frac{1}{x_i}\right)Y_i - \frac{1}{n^2}\left(\sum\left(\frac{1}{x_i}\right)\right)\sum Y_i$$

c. **Según el teorema de Gauss-Markov, ¿cuál es mayor, $Var(b_2)$ o $Var(\hat{\beta}_2)$?, donde $\hat{\beta}_2$ es el estimador MCO de la pendiente. (1 punto)**

El teorema de Gauss-Markov nos dice que los estimadores de MCO tienen la menor varianza dentro de la clase de los estimadores lineales e insesgados. Por ello, se dice que el estimador MCO es el mejor estimador lineal insesgado (MELI). Entonces, como $\hat{\beta}_2$ es el estimador MCO, su varianza será menor al de b_2 o ($Var\ (b_2) > Var\ (\hat{\beta}_2)$.