Econometría I - 2009
Universidad de San Andrés

Trabajo Práctico 1
Modelo Lineal con dos variables

Contenidos: Modelo Lineal con 2 variables, estimadores Mínimo Cuadráticos (MC), propiedades algebraicas de los estimadores MC, supuestos clásicos, propiedades estadísticas de los estimadores MC, test "t", p-valores, coeficiente de determinación o R-cuadrado.

Parte I [1]

1. <u>Modelo lineal con dos variables</u>

Disponemos de los siguientes datos de ahorro *S* e ingreso *I* para 12 familias (en pesos).

Nº hogar	Ahorro (S)	Ingreso (I)
1	75	125
2	85	132
3	54	114
4	62	112
5	95	125
6	25	75
7	125	201
8	85	166
9	80	145
10	75	142
11	40	75
12	56	94

a) Estime la covarianza y la correlación entre el ahorro y el ingreso de las familias. Interprete.

b) Suponga que la relación entre el ahorro y el ingreso de las familias puede representarse mediante el siguiente modelo

$$S_i = \alpha + \beta I_i + \mu_i \qquad i=1,\dots, n$$

en donde μ_i es un termino aleatorio no observable. Estime los parámetros α y β de este modelo mediante el método de Mínimos Cuadrados (MC). Proporcione una interpretación económica para las estimaciones obtenidas.

---

[1] Para todos los ejercicios de la parte I de este Trabajo Práctico puede utilizar calculadora o planilla de cálculo, pero <u>no STATA</u>

c) *Propiedades algebraicas de los estimadores MC*. Usando los datos provistos, verifique que cuando se estima el modelo por el método de MC la suma de los errores es cero, y que la recta de regresión estimada pasa por el punto de las medias muestrales de ambas variables.

d) Construya un gráfico de dispersión de los datos (nube de puntos) ubicando en el eje de ordenadas los valores del ahorro y en el de abscisas el ingreso. En el mismo gráfico, dibuje la recta de regresión estimada en el inciso b. (Nota: no es un dibujo a mano alzada, sino uno que se corresponda con los datos proporcionados)

2. <u>Cambios en las unidades de medida en que se expresan las variables</u>

a) Estime el coeficiente β del modelo de ahorro e ingreso del ejercicio anterior, pero con el ahorro medido en centavos. Interprete el nuevo valor estimado y compare con el obtenido antes del cambio en la unidad de medida.

b) Muestre formalmente que el coeficiente de correlación muestral no cambia cuando la variable S (dependiente) se multiplica por una constante $k > 0$. Muestre formalmente cómo cambia el estimador MC de β ante esa misma transformación.

c) Repita (a) y (b) para el mismo cambio pero aplicado a los ingresos (variable explicativa).

d) Suponga que a todas las familias se les hace una transferencia de ingreso de T pesos. Mostrar formalmente que este cambio no afecta ni al coeficiente de correlación muestral ni al de regresión. ¿Qué es lo que cambia de la relación entre el ahorro y el ingreso ante un cambio como este?

3. <u>Modelos lineales con y sin ordenada al origen</u>

Siguiendo con el ejemplo de ahorro e ingresos de las familias, considere que un modelo alternativo al propuesto más arriba es el siguiente:

$$S_i = \delta\, I_i + \varepsilon_i \qquad i=1,\ldots, n.$$

Note que en este modelo *no hay un intercepto.*

a) Derive analíticamente un estimador mínimo cuadrático para δ.

b) Utilizando los datos provistos, estimar el coeficiente δ usando el estimador derivado en a). Compare con los resultados obtenidos en el ejercicio 1. En particular, verifique que la suma de los errores de estimación es distinta de cero y que la recta de regresión estimada no pasa por las medias muestrales de las variables.

c) Compute el coeficiente de determinación (R^2) para ambos modelos. Interprete y compare.

d) Haga un gráfico de dispersión de los datos (nube de puntos) y trace las dos rectas de regresión estimadas. Explique cómo se manifiesta gráficamente lo que observó en b).

4. Propiedades estadísticas de los estimadores MC

Considere el siguiente modelo lineal

$$Y_i = \alpha + \beta X_i + \mu_i \qquad i=1,\ldots, n,$$

en donde se cumplen todos los supuestos clásicos. Se pide:

a) Demostrar que el estimador MC de α es insesgado.

b) Demostrar que no es necesario que $E[\mu_i] = 0$ con $i=1,\ldots, n$ para que el estimador MC de β sea insesgado.

c) Obtener una expresión analítica para la varianza del estimador MC de α.

d) Obtener la expresión de la covarianza entre los estimadores MC de α y β. Explique intuitivamente por qué existe una covarianza entre ambos estimadores.

5. Test de hipótesis sobre los parámetros del modelo lineal

Nuevamente considere el modelo de ahorro e ingresos familiares presentado en el ejercicio 1. Suponga que todos los supuestos clásicos se cumplen y que, adicionalmente, los términos aleatorios μ_i con $i=1,\ldots, n$ se distribuyen en forma normal. Considere las siguientes hipótesis:

a) H_0: $\alpha = 0$ vs. H_A: $\alpha \neq 0$

b) H_0: $\beta = 0.5$ vs. H_A: $\beta \leq 0.5$

c) H_0: $\alpha + \beta = 0$ vs. H_A: $\alpha + \beta < 0$

Para cada caso se pide que proponga el estadístico de prueba relevante, indique su distribución bajo la hipótesis nula y evalúe las hipótesis considerando un nivel de significatividad del 5%.

Parte II

6. Problema empírico: estimación de una curva de Engel[2]

El objetivo de este ejercicio consiste en introducir el análisis de regresión bivariado a nivel empírico, es decir, a ejemplos concretos usando datos reales, y aprender los comandos básicos del paquete econométrico Stata.

Se estudiará la relación entre el gasto en alimentos y el ingreso de las familias. El gasto en alimentos es una aproximación al consumo de alimentos. Las relaciones entre el consumo de un determinado bien y el ingreso se conocen con el nombre de "curvas de Engel".

El trabajo seminal de Ernest Engel[3] (1857) investigó cómo cambiaba el consumo de alimentos ante cambios en el ingreso en Bélgica. El estudio encontró que si bien el consumo de alimentos tiende a crecer cuando aumenta el ingreso, lo hace en menor proporción. Por lo tanto, a medida que crece el ingreso, el porcentaje gastado en alimentos tiende a caer. Dicha proposición fue conocida luego como "la ley de Engel". Esta ley puede formularse de manera alternativa de la siguiente forma: cuanto mayor es la proporción del ingreso gastada en alimentos, menor es el bienestar de la familia. El análisis de curvas de Engel se extendió rápidamente al estudio de otros bienes. Mediante esta herramienta se clasificó a los bienes en normales e inferiores. Asimismo, los primeros se separaron en bienes de primera necesidad y bienes de lujo. Pero las aplicaciones son bastante más amplias. Entre otras, se destaca el uso de curvas de Engel de consumo de calorías para la estimación de líneas de pobreza.[4]

Este ejercicio consiste en la estimación de la curva de Engel del consumo de alimentos de Argentina en su versión más simple. La base de datos que se empleará (*gasto_ingreso.xls*) fue obtenida a partir de los resultados de la Encuesta de Impacto Social de la Crisis en Argentina (ISCA) realizada en 2002 por el Banco Mundial. La base contiene información sobre el ingreso per cápita familiar (*ipcf*) y el gasto en alimentos per cápita (*gasto_alimentos*) de 1748 hogares con ingresos y gastos en alimentos positivos en el año 2002. Las dos variables se encuentran medidas en pesos corrientes.

El modelo propuesto es:

$$G_i = \alpha + \beta Y_i + u_i \qquad\qquad i = 1,..., 1748$$

donde G, Y y u representan el gasto en alimentos per cápita familiar, el ingreso per cápita familiar y un término aleatorio no observable que suponemos cumple con los supuestos clásicos.

Se pide:

---

[2] A diferencia de los ejercicios de la primera parte, este ejercicio debe realizarse usando STATA.

[3] Ernest Engel fue un estadístico nacido en 1821 en Dresden, Alemania. En 1860 fue designado director de la la oficina de estadística de Prusia, cargo que ocupó hasta que se retiró en 1882. Murió en 1896.

[4] Algunas referencias que pueden consultarse sobre curvas de Engel son:
Lewbel, A. (2006), Engel Curves, Entry for The New Palgrave Dictionary of Economics, 2nd edition. Disponible en www2.bc.edu/~lewbel/palengel.pdf
Varian, H (1998) Microeconomía Intermedia, cap. 6, 4a edición, Antoni Bosch.

a) Realice un análisis estadístico descriptivo básico de cada variable (compute y comente los resultados para la media, la mediana, el rango, el coeficiente de variación y el desvío estándar).

b) Construya un diagrama de dispersión graficando el ingreso per cápita familiar en el eje de abscisas y el gasto en alimentos per cápita del hogar en el eje de ordenadas. Discuta.

c) Utilizando el método de Mínimos Cuadrados Ordinarios, estime los parámetros del modelo propuesto más arriba, donde el gasto en alimentos es función únicamente del ingreso del hogar.

d) Interprete estadística y económicamente los coeficientes obtenidos (considere especialmente las unidades de medida), errores estándar, estadísticos 't', 'valores p' y el estadístico F.

e) Analice la bondad de ajuste del modelo.

f) Evalúe la hipótesis de que el ingreso no es una variable relevante para explicar la demanda de alimentos de los hogares. Evalúe la significatividad estadística de la constante e interprete su significado económico.

g) Agrupe a los hogares según sus ingresos estén por encima o por debajo de la mediana de ingresos y estime un modelo lineal similar al anterior para cada grupo por separado. Compare los coeficientes obtenidos (considerando signo, magnitud, significatividad estadística, etc.). ¿Qué conclusiones se obtienen?

h) Como se mencionó anteriormente, las curvas de Engel permiten distinguir entre bienes inferiores, de primera necesidad y de lujo. Sin embargo, estos conceptos son esencialmente *locales*. En palabras de Hal Varian:

> El hecho de que un bien sea inferior o no, depende del nivel de renta que estemos examinando. Por ejemplo, es posible que las personas más pobres consuman más mortadela cuando aumenta su renta. Sin embargo, traspasado un determinado punto, probablemente consumirán menos.[5]

Comentarios semejantes pueden realizarse respecto del carácter de primera necesidad o de bien de lujo.

Empleando el modelo estimado, encuentre el cambio porcentual en el gasto en alimentos ante un cambio de un 1% en el ingreso per cápita del hogar para:

(i) un hogar con el **mínimo** ingreso per cápita de la muestra

(ii) un hogar con el **máximo** ingreso per cápita de la muestra.[6]

¿De qué tipo de bien se trata? ¿Resultan intuitivos los resultados encontrados? ¿A qué característica del modelo atribuye esto?

Ayuda: utilice la fórmula $\frac{\Delta\%G}{\Delta\%Y} = \beta.\frac{Y}{G}$

[5] Varian, H. *Microeconomía intermedia. Un Enfoque moderno.* Tercera edición. Antoni Bosch, editor.
[6] En ambos casos considere el gasto *esperado* en alimentos.

Notas computacionales para el ejercicio empírico de la Parte II

La siguiente descripción de comandos se basa en la versión 9 de Stata.

Para pasar datos xls a Stata.
Primero se debe guardar la base en Excel con extensión *csv* y asegurarse de que no contiene números con comas (deben reemplazarse por puntos).
Luego usar el siguiente comando en Stata:

```
insheet using “path\nombrebase.csv”, comma
```

Otra alternativa es guardar la base en Excel con extensión *txt* y usar el comando:

```
insheet using “path\nombrebase.txt”, tab
```

Para realizar un descripción estadística de las variables listadas:

```
summarize y x, detail
```

Para obtener algunos estadísticos descriptivos adicionales:

```
tabstat y x, s(mean median cv sd range)
```

Para obtener un gráfico de dispersión entre *x* e *y:*

```
graph twoway scatter y x
```

Para realizar una regresión por MCO de *y* en *x:*

```
regress y x
```

Si se quiere restringir la regresión a algunas observaciones que cumplen con una condición se puede utilizar la siguiente sintaxis:

```
regress y x if x>x0
regress y x if x≤x0
```

Para obtener la predicción lineal del modelo estimado (sólo después de ejecutar el comando regress):

```
predict yest
```