PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS SOCIALES
ESPECIALIDAD DE ECONOMÍA
SEMESTRE 2013-II

PUCP

ECONOMETRÍA 1: PRÁCTICA DIRIGIDA 4

Curso:	Econometría 1	
Profesor:	Gabriel Rodríguez	(gabriel.rodriguez@pucp.edu.pe)
Jefe de práctica:	Patricia Lengua Lafosse	(patricia.lengua@pucp.pe)

1 Conceptos

1. Función Score
2. Matriz Hessiana
3. Matriz de información de Fisher ?
4. Desigualdad de Cramer-Rao y frontera de Cramer-Rao
5. Algoritmo Newton-Raphson

2 Estimación por Máxima Verosimilitud

1. Se tiene el siguiente modelo de regresión multivariado:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\beta + \epsilon \\ \epsilon_t &\sim N(0, \sigma^2) \end{aligned}$$

Encontrar el estimador de β y σ^2 por el método de Máxima Verosimilitud.

2. Considere una muestra aleatoria $y_1, y_2, ..., y_T$ de T observaciones de una distribución Gamma (α, β) con función de densidad:

$$f(y_t \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y_t^{\alpha-1}.e^{-\beta y_t} \qquad \text{con } t = 1, ..., T.$$

Hallar los estimadores de Máxima Verosimilitud de α y β.

3. Dado el modelo de regresión multivariado:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\beta + \epsilon \\ \epsilon_t &\sim N(0, \sigma^2) \end{aligned}$$

Se tiene que $\widehat{\beta}_{MV}$ y $\widehat{\sigma}^2_{MV}$ son los estimadores de Máxima Verosimilitud de β y σ^2. Estimar la varianza asintótica de $\widehat{\beta}_{MV}$ y $\widehat{\sigma}^2_{MV}$. Además, encuentre la cota inferior de Cramer-Rao de $\widehat{\beta}_{MV}$ y $\widehat{\sigma}^2_{MV}$ e interprete.

3 Variables Dummy

1. Se cuenta con una muestra de profesionales y se propone el siguiente modelo:

$$ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 notas_i + \epsilon_i, \qquad i = 1, 2, ...N$$

donde $ingreso_i$ es el ingreso por hora en soles, $mujer_i$ es una variable dummy que toma el valor de 1 si es mujer y 0 en otro caso y $notas_i$ es la nota promedio obtenida en estudios superiores.

(a) Plantee la hipótesis nula para la relevancia del género para explicar el ingreso del individuo.

(b) Suponga el siguiente modelo alternativo:

$$ingreso_i = \alpha_0 Hombre_i + \alpha_1 mujer_i + \alpha_2 notas_i + u_i$$

hombre $\begin{cases} 1 & \text{si es homb} \\ 0 & \text{si es otro} \end{cases}$

Plantee la hipótesis nula para la relevancia del género para explicar el ingreso del individuo. ¿Es posible estimar este modelo?

(c) Si se desea permitir que el efecto de las notas sobre el ingreso sea distinto para hombres y mujeres, plantee el modelo adecuado. Interprete los coeficientes e ilustre gráficamente el modelo. Además, plantee la hipótesis nula para la relevancia del género.

(d) Si ahora en lugar de tomar en cuenta las notas, se desea tomar en cuenta la profesión utilizando la variable: *doctor* = 1, 0 en otro caso. Plantee el modelo adecuado, donde las variables *mujer* y *doctor* tengan efectos aditivos y multiplicativos. Interprete los coeficientes e ilustre gráficamente el modelo. Además, plantee la hipótesis nula para la relevancia del género y aquella para la relevancia de la profesión.

(e) Plantee el modelo que incluya las variables *mujer*, *notas* y *doctor*; y que además permita que el efecto del género varíe según profesión, así como que el efecto de las notas varíe según género y según profesión. Además, plantee la hipótesis nula para la relevancia del género.

(f) Suponga que se cuenta con dos variables adicionales: *abogado* y profesor. Si la muestra está compuesta sólo por doctores, abogados y profesores, responda:

(i) ¿Es posible estimar el siguiente modelo? Justifique su respuesta.

$$ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 abogado_i + \beta_4 profesor_i + \beta_5 notas_i + \epsilon_i$$

(ii) Interprete los coeficientes de los siguientes modelos alternativos:

- $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 abogado_i + \beta_4 notas_i + \epsilon_i$
- $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 profesor_i + \beta_4 notas_i + \epsilon_i$
- $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 abogado_i + \beta_3 profesor_i + \beta_4 notas_i + \epsilon_i$

Practica Dirigida 4

1. Estimación por Máxima Verosimilitud.

1. $Y = X\beta + \epsilon \qquad \epsilon \sim IN(0, \sigma^2)$

○ Función de verosimilitud.

$$\mathcal{L}(\beta, \sigma^2 / \underset{\sim}{y}) = p(\underset{\sim}{Y}; \beta, \sigma^2) = (2\pi\sigma)^{-T/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right\}$$

$$\ln(\mathcal{L}) = -\frac{T}{2}\log(\sigma^2) - \frac{T}{2}\log(2\pi) - \frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)$$

$$l = -\frac{T}{2}\log(\sigma^2) - \frac{T}{2}\log(2\pi) - \frac{1}{2\sigma^2}(Y'Y - 2\beta'X'y + \beta'X'X\beta)$$

CPO

○ $$\frac{\partial l}{\partial \beta} = -\frac{1}{2\sigma^2}\left[-2X'y + 2X'X\hat{\beta}\right] = 0 \quad \Rightarrow \quad \hat{\beta}_{MV} = (X'X)^{-1}(X'y)$$

○ $$\frac{\partial l}{\partial \sigma^2} = -\frac{T}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}(Y - X\beta)'(Y - X\beta) = 0 \quad \Rightarrow \quad \hat{\sigma}^2_{MV} = \frac{(Y - X\hat{\beta}_{MV})'(Y - X\hat{\beta}_{MV})}{T}$$

2.

$$f(y_t \mid \alpha, \beta) = \left[\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot y_t^{\alpha-1} \cdot e^{-\beta y_t}\right] \qquad \text{con } t = 1, \ldots, T$$

Función Verosimilitud: $\mathcal{L}(\alpha, \beta) = \prod_{t=1}^{T} f(y_t \mid \alpha, \beta) = \prod_{t=1}^{T}\left[\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot y_t^{\alpha-1} \cdot e^{-\beta y_t}\right]$

$$\mathcal{L} = \frac{\beta^\alpha}{\Gamma(\alpha)}\left(y_1^{\alpha-1} e^{-\beta y_1}\right) \cdot \left(\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot y_2^{\alpha-1} \cdot e^{-\beta y_2}\right) \cdot (\ldots) \cdot \left(\frac{\beta^\alpha}{\Gamma(\alpha)} y_T^{\alpha-1} e^{-\beta y_T}\right) = \left[\frac{\beta^\alpha}{\Gamma(\alpha)}\right]^T \cdot e^{-\beta \Sigma y_t} \cdot \prod_{t=1}^{T} y_t^{\alpha-1}$$

$$l = \ln(\alpha, \beta \mid \underset{\sim}{y}) = T \cdot \alpha \ln(\beta) - T\ln\{\Gamma(\alpha)\} - \beta \Sigma y_t + (\alpha - 1)\Sigma \ln(y_t)$$

CPO

○ $$\frac{\partial l}{\partial \beta} = \frac{\alpha T}{\beta} - \sum_{t=1}^{T} y_t = 0 \quad \Rightarrow \quad \hat{\beta}_{MV} = \frac{\hat{\alpha}_{MV}}{\bar{Y}} \quad (i)$$

○ $$\frac{\partial l}{\partial \alpha} = T\ln\beta - T \cdot \frac{\partial \ln(\Gamma(\alpha))}{\partial \alpha} + \sum_{t=1}^{T} \ln y_t = 0$$

Usamos la primera ecuación (i)

$$T.\ln(\alpha) - T.\ln(\bar{Y}) - T\psi(\alpha) + \sum \ln(y_t) = 0 \quad ; \quad \text{con: } \psi(\alpha) = \frac{\partial \ln\{\Gamma(\alpha)\}}{\partial \alpha}$$ {Fu. digam.}

I. Conceptos:

1. Funcion Score: Gradiente de la función de log-verosimilitud (vector de derivadas parciales)

$$\text{Score} = g(\underset{\sim}{\theta}) = \begin{bmatrix} \frac{\partial l(\underset{\sim}{\theta})}{\partial \theta_1} \\ \frac{\partial l(\underset{\sim}{\theta})}{\partial \theta_2} \\ \vdots \\ \frac{\partial l(\underset{\sim}{\theta})}{\partial \theta_p} \end{bmatrix}_{p\times 1}$$

2. Matriz Hessiana: Matriz de segundas derivadas de la Función de log-verosimili

$$H(\underset{\sim}{\theta}) = \begin{bmatrix} \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_1^2} & \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_1 \partial \theta_p} \\ \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_2^2} & \cdots & \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_2 \partial \theta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_p \partial \theta_1} & \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_p \partial \theta_2} & \cdots & \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_p^2} \end{bmatrix}$$

3). Matriz de Información de Fisher

$$I_F(\underset{\sim}{\theta}) = -E\left[H(\underset{\sim}{\theta})\right]$$

5) Algoritmo Newton-Raphson.

a) Caso con un parámetro:

$$\text{Func. Score} = g(\underset{\sim}{\theta}) = \frac{\partial l(\underset{\sim}{\theta})}{\partial \theta}$$

El objetivo es encontrar el valor de θ que haga que la "función score" sea igual a cero.

$$g(\underset{\sim}{\theta}) = 0$$

Si "θ^*" es una solución y se cumple que: $\left. \frac{\partial g(\underset{\sim}{\theta})}{\partial \theta} \right|_{\theta=\theta^*} < 0$

Entonces, θ^* es un máximo local; si es el único, será el global.

⇒ Consideremos la exposición de Taylor de $g(\underset{\sim}{\theta})$ en torno al punto θ_0

$$g(\theta) \approx g(\theta_0) + g'(\theta_0).[\theta - \theta_0]$$

Consideramos

Evaluando la exposición en $\theta = \theta^*$ se trae que:

$$\underbrace{g(\theta^*)}_{0} \simeq g(\theta_0) + g'(\theta_0)\cdot[\theta^* - \theta_0]$$

$$0 \simeq g(\theta_0) + g'(\theta_0)\cdot[\theta^* - \theta_0]$$

$$\boxed{\theta^* = \theta_0 - \frac{g(\theta_0)}{g'(\theta_0)}}$$

que induce al siguiente algoritmo:

$$\theta_1 = \theta_0 - \frac{g(\theta_0)}{g'(\theta_0)} \leadsto \theta_j = \theta_{j-1} - \frac{g(\theta_{j-1})}{g'(\theta_{j-1})} \Rightarrow$$

* y el algoritmo continuará hasta que

$$\left| \frac{\theta_j - \theta_{j-1}}{\theta_{j-1}} \right| < \epsilon$$

$$\begin{bmatrix}\hat{\beta}_{MV}\\ \hat{\sigma}^2_{MV}\end{bmatrix} = \left[\begin{array}{c|c}\frac{1}{\sigma^2}(X'X) & 0 \\ \hline 0 & T/2\sigma^4\end{array}\right]^{-1} = \left[\begin{array}{c|c}\sigma^2(X'X)^{-1} & 0 \\ \hline 0 & 2\sigma^4/T\end{array}\right]$$

* Sabemos que la cota inferior de Cramer Rao de $\hat{\beta}_{MV}$ y $\hat{\sigma}^2$ esta dada por $[I_F(\beta;\sigma^2)]^{-1}$
Los estimadores de MV tienen la menor varianza dentro de la clase de estimadores consistentes y asintóticamente normales. Incluso se llega a tener varianzas menores a MCO.

Ejemp: $var(\hat{\sigma}^2_{MV}) < var(s^2) \Rightarrow \frac{2\sigma^4}{T} < \frac{2\sigma^4}{T-k}$

III. Variables Dummy

$$ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 notas_i + \epsilon_i \qquad i = 1,2,\dots N$$

1: Ho de relevancia de genero: $H_0 \Rightarrow \beta_1 = 0$
$H_1: \beta_2 \neq 0$

2. Si $ingreso_i = \alpha_0 Hombre_i + \alpha_1 mujer_i + \alpha_2 notas_i + u_i$

$$Y = \begin{bmatrix} ingreso_1 \\ ingreso_2 \\ \vdots \\ ingreso_N \end{bmatrix} \quad X = \begin{bmatrix} 1 & 0 & notas_1 \\ 1 & 0 & notas_2 \\ \vdots & \vdots & \vdots \\ 1 & 0 & \\ 0 & 1 & \\ \vdots & \vdots & \\ 0 & 1 & notas_N \end{bmatrix}$$

$\nexists$ $y_i = \beta_0 + \alpha_1 hombre_i + \alpha_2 mujer_i + \alpha_3 notas_i + u_i$

$$X = \begin{bmatrix} 1 & 1 & 0 & notas_1 \\ 1 & 1 & 0 & notas_2 \\ 1 & 1 & 0 & \vdots \\ \vdots & 1 & 0 & \vdots \\ \vdots & 0 & 1 & \vdots \\ 1 & 0 & 1 & notas_N \end{bmatrix}$$

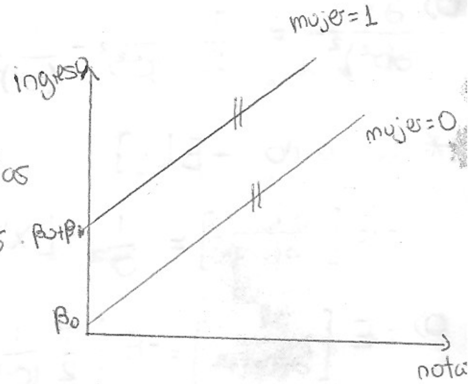
* Si sumamos las columnas de hombre y mujer obtendemos la columna del intercepto
→ Multicolinealidad

∴ Ho para relevancia de genero: $H_0: \alpha_0 = \alpha_1 = 0$
$H_1: \alpha_0$ y/o $\alpha_1 \neq 0$

3. Modelo Inicial: $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 notas + \epsilon_i$

Interpretación:

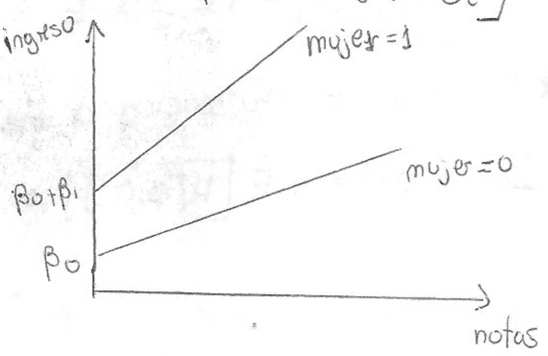
$E[ingreso \mid mujer = 1, notas] = (\beta_0 + \beta_1) + \beta_2 notas$
$E[ingreso \mid mujer = 0, notas] = \beta_0 + \beta_2 notas$



a) Modelo con influencia en notas: $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 notas + \beta_3 notas_i mujer_i + \epsilon_i$

Interpretación

$E[ingreso \mid mujer = 1, notas] = (\beta_0 + \beta_1) + (\beta_2 + \beta_3) notas$
$E[ingreso \mid mujer = 0, notas] = \beta_0 + \beta_2 notas$



4.) Desigualdad de Crámer - Rao Sea $\underset{\sim}{Y} = (y_1, y_2, \ldots, y_T)$ un vector de va[riables] aleatorios (no necesariamente independientes). La Función de densidad conjunta es[tá dada] por $f(\underset{\sim}{Y}; \underset{\sim}{\theta})$ donde $\underset{\sim}{\theta}$ es un vector p-dimensional de parametros, esto es:

$(\underset{\sim}{\theta}) = (\theta_1, \theta_2, \ldots, \theta_p)$.

Sea $L(\underset{\sim}{\theta}|\underset{\sim}{y})$ la Función de verosimilitud y sea $\hat{\theta}(y)$ un estimador insesgado de θ con Matriz de varianza - covarianzas Finita. Entonces, bajo ciertas condiciones de regularidad en $f(\underset{\sim}{Y}; \underset{\sim}{\theta})$

II. 3. $Y = X\beta + \epsilon \qquad \epsilon_t \sim N(0, \sigma^2)$

Teniendo en cuenta la propiedad de normalidad asintótica de los estimadores MV.
$\begin{bmatrix} \tilde{\beta}_{MV} \\ \hat{\sigma}^2_{MV} \end{bmatrix} \sim N\left[\begin{pmatrix} \beta \\ \sigma^2 \end{pmatrix}, I_F^{-1}(\beta, \sigma^2) \right]$

* Calculando las segundas derivadas de l.

o) $\dfrac{\partial l}{\partial\beta\partial\beta'} = \dfrac{-1}{\theta^2}[X'X]$

o) $\dfrac{\partial l}{\partial\beta\partial\sigma^2} = \dfrac{1}{2(\theta^2)^2}\cdot[-2X'y + 2X'X\beta] = \dfrac{-1}{(\theta^2)^2}\cdot X'[y - X\beta] = \dfrac{-1}{(\theta^2)^2}\cdot X'\cdot\epsilon$

o) $\dfrac{\partial l}{\partial(\sigma^2)^2} = \dfrac{T}{2}\,\dfrac{1}{(\sigma^2)^2} - \dfrac{1}{(\theta^2)^3}(Y - X\beta)'(Y - X\beta) = \dfrac{T}{2}\cdot\dfrac{1}{(\sigma^2)^2} - \dfrac{1}{(\sigma^2)^3}\epsilon'\epsilon$

* Tomando $-E[\cdot]$:

o) $-E\left[\dfrac{\partial l}{\partial\beta\partial\beta'}\right] = \dfrac{1}{\sigma^2}[X'X]$

o) $-E\left[\dfrac{\partial l}{\partial^2\sigma^2}\right] = -E\left[\dfrac{T}{2}\dfrac{1}{(\sigma^2)^2} - \dfrac{1}{(\sigma^2)^3}\epsilon'\epsilon\right] = -\dfrac{T}{2(\sigma^2)^2} + \dfrac{1}{(\sigma^2)^3}\underbrace{E(\epsilon'\epsilon)}_{T\sigma^2} = \dfrac{T}{2(\sigma^2)^2}$

o) $-E\left[\dfrac{\partial l}{\partial\beta\partial\sigma^2}\right] = +\dfrac{1}{(\sigma^2)^2}\cdot X'\cdot E(\epsilon) = 0$

* La Matriz de informacion de Fisher:

$$I_F(\beta, \sigma^2) = -E[H(\beta, \sigma^2)] = \begin{bmatrix} \frac{1}{\sigma^2}(X'X) & 0 \\ 0 & \frac{T}{2\sigma^4} \end{bmatrix}$$

$\Rightarrow$ La varianza asintótica es la inversa de la matriz de información de Fisher. $\Big\}$ $A\text{var}\begin{bmatrix} \hat{\beta}_{MV} \\ \hat{\sigma}^2_{MV} \end{bmatrix} = [I_F(\beta, \sigma^2)]^{-1}$

: Efecto aditivo de ser mujer sobre el ingreso: mide la diferencia de interceptos entre hombres y mujeres.

β_2: Efecto multiplicativo de ser mujer sobre el ingreso: mide la diferencia de los efectos de las notas sobre los ingresos entre hombres y mujeres (pendientes).

4. Relevancia de Profesión: $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 mujer_i doctor + \epsilon_i$

Interpretación:

- $E[ingreso \mid mujer=1, doctor=1] = (\beta_0+\beta_1)+(\beta_2+\beta_3)$ ⟶ mujer doctora
- $E[ingreso \mid mujer=1, doctor=0] = (\beta_0+\beta_1)$ ⟶ mujer no doctora
- $E[ingreso \mid mujer=0, doctor=1] = (\beta_0+\beta_2)$ ⟶ hombre doctor
- $E[ingreso \mid mujer=0, doctor=0] = \beta_0$ ⟶ hombre no doctor.

Relevancia de genero.

$H_0: \beta_1 = \beta_3 = 0$

Relevancia de profesión

$H_0: \beta_2 = \beta_3 = 0$

5. Relevancia genero, notas y profesión

$Ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 notas_i + \beta_4 doc.muje. + \beta_5 not.mujer + \beta_6 not.doc$

Interpretación:
- $E[ingreso / mujer=1, doctora=1, notas] = (\beta_0+\beta_1+\beta_2+\beta_4)+(\beta_3+\beta_5+\beta_6)$ notas.
- $E[ingreso / mujer=1, doctora=0, notas] = (\beta_0+\beta_1)+(\beta_3+\beta_5)$ notas.

6.
i) $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 abogado_i + \beta_4 profesor_i + \beta_5 notas_i + \epsilon_i$

No se puede estimar el modelo por la trampa de las variables dummies Entonces, tenemos que eliminar una categoria (base).

ii) $ingreso_i = \beta_0 + \beta_1 mujer_i + \beta_2 doctor_i + \beta_3 abogado_i + \beta_4 notas_i + \epsilon_i$

Interpretación

- $E[ingreso \mid mujer=1, doctor=1, abogado=0, notas] = (\beta_0+\beta_1+\beta_2) + \beta_4$ notas
- $E[ingreso \mid mujer=0, doctor=1, abogado=0, notas] = (\beta_0+\beta_2) + \beta_4$ notas
- $E[ingreso \mid mujer=0, doctor=0, abogado=1, notas] = (\beta_0+\beta_3) + \beta_4$ notas