

## **Trabajo Práctico 5**

### **Heterocedasticidad - Modelos Probit y Logit**

**Contenidos:** Heterocedasticidad. Variable dependiente binaria, modelo probit, modelo logit, efectos marginales.

**Fecha de entrega:** jueves 16 de noviembre. Ver [reglas de formato y presentación](https://econometria1unlp.weebly.com/praacutectica.html) en <https://econometria1unlp.weebly.com/praacutectica.html>.

---

#### **PARTE I: Heterocedasticidad**

##### **1. Consumo e ingreso**

En el archivo het\_2022.dta se presentan datos de consumo e ingresos de 44 familias. Se solicita lo siguiente:

- a) Construir un diagrama de dispersión para los datos (ingreso en el eje x, consumo en el eje y). Comentar.
- b) Estimar los parámetros de una función de consumo lineal por MCO. Interpretar los resultados obtenidos y discutir brevemente las consecuencias de utilizar MCO bajo heterocedasticidad.
- c) Graficar los residuos de la regresión contra los valores de ingreso. Comentar los resultados.
- d) Evaluar la presencia de heterocedasticidad con los siguientes tests: (d.1) Breusch-Pagan/Godfrey/Koenker. (d.2) Goldfeld y Quandt. (d.3) White.
- e) Estimar el modelo con las variables transformadas según el método de mínimos cuadrados generalizados. Interpretar cuidadosamente los coeficientes obtenidos.
- f) Evaluar la hipótesis nula de homocedasticidad con los residuos del modelo transformado en base al test de White.
- g) Estimar el modelo original por MCO utilizando el estimador consistente de White para la matriz de varianzas.

#### **PARTE II: Modelos de elección binaria**

##### **2. Modelo Logit**

Suponer el siguiente modelo no lineal donde  $Y_i$  es una variable binaria y  $F(.)$  es una función de distribución acumulada logística

$$\text{Prob}(Y_i = 1) = F(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i})$$

donde  $X_1$  y  $X_2$  son variables continuas y  $X_3$  es una variable binaria que toma únicamente los valores 0 y 1.

Se dispone de los siguientes resultados de la estimación del modelo Logit:

$$\hat{\beta}_0 = -19 \quad \hat{\beta}_1 = 0.6 \quad \hat{\beta}_2 = 0.1 \quad \hat{\beta}_3 = 2.3$$

Las medias muestrales de las variables explicativas son:

$$X_1 = 23.2 \quad X_2 = 3.3 \quad X_3 = 0.35$$

Recordar que en el caso del modelo Logit:

$$F(X_i'\beta) = \frac{e^{X_i'\beta}}{1+e^{X_i'\beta}} = F \quad \text{y su derivada es } f(X_i'\beta) = F(1-F)$$

Suponer que  $Y_i = 1$  implica pertenecer a un determinado sindicato, que  $X_1$  es la edad de un individuo,  $X_2$  son los años de experiencia en su ocupación principal y  $X_3$  es el sexo (=1 si hombre).

- ¿Cuál es la probabilidad de que un hombre de 30 años y con 10 años de experiencia esté afiliado al sindicato? Explicar el procedimiento.
- Estimar el efecto marginal de la edad sobre la probabilidad de ser miembro de un sindicato para un hombre de 30 años de edad y con 10 años de experiencia. Explicar el procedimiento e interpretar.
- Computar el efecto marginal de la variable  $X_3$  sobre la probabilidad de ser miembro de un sindicato para personas con 30 años de edad y 10 años de experiencia. Explicar el procedimiento e interpretar.
- Calcular el efecto marginal de los años de experiencia con las variables  $X_1$ ,  $X_2$  y  $X_3$  evaluadas en sus medias muestrales. ¿Tiene una interpretación económica este valor? Comentar.

### 3. Modelo de desempleo para el Gran Buenos Aires

En este ejercicio utilizaremos información proveniente de la Encuesta Permanente de Hogares (EPH) para investigar los factores determinantes de la probabilidad de desempleo. La base de datos *empleo2022.dta* contiene la información de la encuesta para el aglomerado urbano del Gran Buenos Aires en el segundo semestre de 2022, y para individuos de entre 25 y 65 años de edad<sup>1</sup>.

Se solicita lo siguiente:

- Estimar un modelo lineal de probabilidad utilizando como variable dependiente la situación laboral (desempleo) y las siguientes variables explicativas:
  - edad,

<sup>1</sup> Para una mayor comprensión de la base se recomienda consultar el “Diseño de registro y estructura para las bases preliminares de Hogares y Personas” publicado por INDEC, disponible en: [https://www.indec.gov.ar/ftp/cuadros/menusuperior/eph/EPH\\_registro\\_3T2021.pdf](https://www.indec.gov.ar/ftp/cuadros/menusuperior/eph/EPH_registro_3T2021.pdf)

- sexo (variable binaria),
  - estado civil soltero (variable binaria),
  - años de educación,
  - variable binaria que indique si el individuo es jefe del hogar (construir a partir de la variable ch03),
  - monto total de ingresos no laborales.
  - variable binaria que indique si el individuo vive en Capital Federal.
- b) Calcular los valores de probabilidad estimados por la regresión anterior para cada individuo y graficarlos. ¿Qué problema muestra el modelo lineal de probabilidad respecto de estos valores estimados?
- c) Estimar un modelo logit con las mismas variables que en el inciso a). ¿Qué interpretación tienen los valores de los coeficientes en este modelo? Explicar las diferencias con el modelo lineal.
- d) Estimar los efectos marginales de cada variable sobre la probabilidad del desempleo para un individuo de sexo femenino con educación primaria. Calcular lo mismo para un individuo masculino con educación universitaria. En los dos casos, evaluar las variables continuas en su valor promedio. Comparar los resultados de los efectos marginales y de la probabilidad de desempleo que se predice.