

# Supuestos del Modelo de Regresión Lineal Múltiple

---

$$Y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i, \text{ para todo } i = 1, \dots, n$$

- $u_i \sim N(0, \sigma)$ : El comportamiento normal del error algunas veces se llega a considerar innecesaria ya que no se requiere para varios de los resultados utilizados en el análisis de regresión múltiple, sin embargo, es útil al momento de construir intervalos de confianza y estadísticos de prueba. Además, con este supuesto, los estimadores de MCO  $\hat{\beta}_k$  son MEI (mejores estimadores insesgados).
- **No hay autocorrelación del error:** Para dos perturbaciones  $u_i$  y  $u_j$ , este supuesto menciona que dichos errores no están correlacionados. Si existiera correlación entre los errores, entonces  $Y_i$  no solo depende de las variables  $x_{ik}$ , también dependería de  $u_j$ , ya que  $u_i$  tiene cierta dependencia de  $u_j$ .
- **Homocedasticidad del error:** Sin importar el valor de  $x_{ik}$ , la varianza del término error  $u_i$  siempre es la misma, es decir, que para todas las combinaciones de valores de las variables explicativas, la varianza es la misma.
- **No multicolinealidad  $x$ 's:** Establece que no existe una relación lineal entre  $x_{ir}$  y  $x_{is}$  cuando  $r \neq s$ , en otras palabras, significa que ninguna de las regresoras se puede escribir como combinación lineal *exacta* del resto de las regresoras del modelo.
- **Estabilidad de los parámetros:** En el caso de las series de tiempo, al establecer un modelo de regresión múltiple para esta estructura de datos puede que haya un cambio estructural en la relación entre  $Y_i$  y las  $x_{ik}$ . Estabilidad en los parámetros significa que el valor de  $\beta_i$  no cambia en el tiempo, ni en los subconjuntos de datos, es decir, que a pesar de que podría haber cambios estructurales, el supuesto es que no se hacen presentes.
- **Forma funcional lineal:** Este supuesto considera que la linealidad del modelo de regresión lineal múltiple se presenta en los parámetros  $\beta_i$ , ya que  $Y_i$  y/o  $x_{ik}$  pueden ser funciones arbitrarias de las variables de interés.
- **Correcta especificación:** Implica que no se omiten ni se incluyen variables irrelevantes en el ajuste del modelo, ya que los dos casos mencionados traen consigo consecuencias.

*Omisión de una variable relevante*

Suponiendo un modelo verdadero

$$Y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i$$

pero se ajusta el modelo:

$$Y_i = \alpha_1 + \alpha_2 x_{i2} + \nu_i$$

las consecuencias son:

1. Si la variable excluida  $x_3$  está correlacionada con  $x_2$ , entonces  $\hat{\alpha}_1$  y  $\hat{\alpha}_2$  son sesgados

e inconsistentes. Es decir,  $E(\hat{\alpha}_1)$  no es igual a  $\beta_1$  y  $E(\hat{\alpha}_2)$  no es igual a  $\beta_2$ , y el sesgo no desaparece conforme aumenta el tamaño de la muestra.

2. Aunque las variables no estén correlacionadas,  $\hat{\alpha}_1$  es sesgado, a pesar de que  $\hat{\alpha}_2$  sea insesgado.
3. La varianza del error  $\sigma^2$  está estimada incorrectamente.
4. La varianza medida convencionalmente de  $\hat{\alpha}_2$  es un estimador sesgado de la varianza del verdadero estimador  $\hat{\beta}_2$ .
5. Es probable que el intervalo de confianza usual y los procedimientos de pruebas de hipótesis conduzcan a conclusiones equivocadas.
6. Los pronósticos basados en el modelo incorrecto y los intervalos de confianza del pronóstico no son confiables.

#### *Inclusión de una variable irrelevante*

Suponiendo que

$$Y_i = \beta_1 + \beta_2 x_{i2} + u_i$$

es el modelo verdadero, pero se ajusta:

$$Y_i = \alpha_1 + \alpha_2 x_{i2} + \alpha_3 x_{i3} + \nu_i$$

las consecuencias de este error de especificación son:

1. Todos los estimadores de MCO de los parámetros del modelo "incorrecto" son insesgados y consistentes, es decir,  $E(\alpha_1) = \beta_1$ ,  $E(\alpha_2) = \beta_2$  y  $E(\alpha_3) = \beta_3 = 0$ .
2. La varianza del error  $\sigma^2$  está correctamente estimada.
3. Los procedimientos usuales de intervalos de confianza y de pruebas de hipótesis conservan su validez.
4. Sin embargo, las  $\alpha$  estimadas por lo general serán ineficientes, sus varianzas generalmente serán más grandes que las del verdadero modelo.

## Referencias

---

- Gujarati, D. N. (2015). *Econometría* (5a ed.). México, D.F.: McGraw-Hill/Interamericana Editores S.A. de C.V.
- Greene, W. (2012). *Econometric Analysis* (7a ed.). Boston/Munich: Prentice Hall.
- Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach* (4a ed.). Mason, Ohio: South Western Cengage Learning.