

Estimacion y Propiedades del Modelo Lineal

Herramientas Econométricas - UCA

Ricardo Pasquini

Octubre 2022

Estimación del Modelo Lineal

- ▶ Previamente hemos discutido al Mejor Predictor Lineal de y dado x para un par de variables aleatorias $(y, x) \in \mathbb{R} \times \mathbb{R}^k$

$$\beta = (E(xx')^{-1})E(xy)$$

- ▶ Nos interesa ahora la estimación de β basada en datos observacionales.

Estimación del Modelo Lineal

- ▶ Un supuesto comúnmente utilizado para permitir la estimación del modelo es que las observaciones son *i.i.d.* (identicamente distribuídas). Esto garantiza que la esperanza de las observaciones sea estable.
- ▶ La media poblacional de una variable aleatoria está definida como:

$$\mu = E(y_i) = \int_{-\infty}^{+\infty} y dF(y)$$

- ▶ Comúnmente estimamos usando la media muestral:

$$\hat{\mu} = \frac{\sum y_i}{n}$$

Estimación del Modelo Lineal

- ▶ Para estimar una función de la media poblacional $g(\mu)$ también es estándar utilizar la media de la funciones $g(\hat{\mu}) = \frac{\sum g(y_i)}{n}$
- ▶ Una forma de obtener el estimador de

$$\beta = (E(xx')^{-1})E(xy)$$

Es reemplazando por los estimadores de los valores esperados respectivos

$$\hat{\beta} = ((\frac{\sum x_i x_i'}{n})^{-1})(\frac{\sum x_i x_i'}{n}) = (\sum x_i x_i')^{-1}(\sum x_i x_i')$$

Este método se conoce como *plug-in*.

Propiedades del Estimador de OLS

- ▶ Estudiamos las propiedades del estimador OLS
 - ▶ Para ello supondremos que el modelo es cierto en la población y analizamos el comportamiento del estimador.
- ▶ Podemos demostrar que
 - ▶ El estimador OLS es *insesgado*.
 - ▶ La varianza del estimador, en general, es

$$V_{\hat{\beta}} = \text{var}(\hat{\beta}|X) = (X'X)^{-1}(X'DX)(X'X)^{-1}$$

donde D es la matriz que tiene $E(e_i^2|x_i) = \sigma_i^2$ en la diagonal y $E(e_i e_j|X)$ fuera de la diagonal.

Propiedades del Estimador de OLS

- Notar que si las observaciones son *independientes* y estamos en el caso de *homocedasticidad*

$$V_{\hat{\beta}} = \text{var}(\hat{\beta}|X) = (X'X)^{-1}\sigma^2$$

- Donde σ^2 es la varianza (asumida constante) del error. Por esta razón, en general reducir el error del modelo (lo no-explicado) ayudará a mejorar la precisión de la estimación de los coeficientes.

El estimador de OLS es insesgado

- ▶ En promedio nuestro estimador obtiene el valor de β correcto.

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

- ▶ Proof:

$$E[\hat{\beta}] = E[(X'X)^{-1}X'Y] = E[(X'X)^{-1}X'(X\beta + e)] = \beta + E[X'e] = \beta$$

Varianza del error OLS

- ▶ La varianza del error es

$$\text{Var}(e|X) = E(ee'|X) \stackrel{\text{def}}{=} D$$

- ▶ Donde D es la matriz que tiene a $E(e_i^2|x_i) = \sigma_i^2$ en la diagonal, y $E[e_i e_j|X]=0$ fuera de la misma (debido a la independencia de las observaciones). También se puede denotar $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$
- ▶ Para el caso homocedástico:

$$\text{Var}(e|X) = I_n \sigma^2$$

Varianza del estimador de OLS

- En general la varianza de $\hat{\beta}$ es

$$V_{\hat{\beta}} = \text{var}(\hat{\beta}|X) = (X'X)^{-1}(X'DX)(X'X)^{-1}$$

donde $D = E(\mathbf{e}\mathbf{e}'|X)$ o dicho de otra forma la matriz que tiene $E(e_i^2|x_i) = \sigma_i^2$ en la diagonal y $E(e_i e_j|X)$ fuera de la diagonal.

$$(X'DX) = \sum_i^n x_i x_i' \sigma_i^2$$

Estimador de la Varianza del Error

- Recordemos que la varianza del error es equivalente al valor esperado del error al cuadrado. Queremos estimar $E(e_i^2|x_i)$. También podemos usar el estimador de momentos:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \hat{e}_i^2$$

- Resulta que el estimador es sesgado cuando cuando tengo pocas observaciones. Intuitivamente, para estimar el error se deben estimar previamente otro k parámetros, por lo que se pierden grados de libertad. Se ajusta de la siguiente forma:

$$s^2 = \left(\frac{1}{n-k}\right) \sum_{i=1}^n \hat{e}_i^2$$

Estimación de $\text{Var}(\hat{\beta})$ (bajo homocedasticidad)

- ▶ Bajo homocedasticidad, la matriz de covarianza toma la forma

$$\text{Var}(\hat{\beta}) = (X'X)^{-1}\sigma^2$$

- ▶ Podemos utilizar el estimador insesgado de σ^2 : s^2

$$\hat{\text{Var}}(\hat{\beta}) = (X'X)^{-1}s^2$$

- ▶ Este suele ser el estimador de default en muchos programas, incluso en STATA. Sin embargo hay que recordar que solo aplica bajo homocedasticidad (osea, condiciones bastante excepcionales).

Estimación de $\text{Var}(\hat{\beta})$ (bajo heterocedasticidad)

- Bajo heterocedasticidad, la matriz de covarianza toma la forma

$$V_{\hat{\beta}} = \text{var}(\hat{\beta}|X) = (X'X)^{-1}(X'DX)(X'X)^{-1}$$

- O también

$$V_{\hat{\beta}} = \text{var}(\hat{\beta}|X) = (X'X)^{-1} \sum_i^n x_i x_i' e_i^2 (X'X)^{-1}$$

Estimación de $\text{Var}(\hat{\beta})$ (bajo heterocedasticidad)

- ▶ Es fácil mostrar que si conociéramos los errores (quod-non), podríamos construir directamente un estimador insesgado.
- ▶ Un estimador *factible* podría reemplazar los errores estimados que surgen de OLS, simplemente \hat{e}_i :

$$V_{\hat{\beta}}^{White} = (X'X)^{-1} \left(\sum_i^n x_i x_i' E(\hat{e}_i^2 | X) \right) (X'X)^{-1}$$

- ▶ El estimador de *White* es muy utilizado y es la opción “robust” en Stata.

Estimación de $\text{Var}(\hat{\beta})$ (bajo heterocedasticidad)

- ▶ Como sabíamos que el estimador de la varianza del error es sesgado, una variante podría ser ajustar este estimador multiplicando por $n/(n - k)$

Scaled-White

$$V_{\hat{\beta}}^{\text{Scaled-White}} = \frac{n}{n - k} (X'X)^{-1} \left(\sum_i^n x_i x_i' E(\hat{e}_i^2 | X) \right) (X'X)^{-1}$$

- ▶ Otras especificaciones usualmente utilizadas incorporan los principios de utilizar estimaciones del error mediante predicción (forecast error) , o residuos estandarizados. (Ver opciones hc2 y hc3 en Stata).

Mínimos Cuadrados Generalizados

- Consideremos ahora una situación donde los errores, en general, podrían ser heterocedásticos, o incluso estar correlacionados (i.e., no ser independientes).

$$y = X\beta + e$$

$$E(e|X) = 0$$

$$\text{Var}(e|X) = \Omega$$

Mínimos Cuadrados Generalizados

- ▶ Bajo estos supuestos se puede mostrar que las propiedades de OLS en este caso son

$$E(\hat{\beta}|X) = \beta$$
$$\text{Var}(\hat{\beta}|X) = (X'X)^{-1}(X'\Omega X)(X'X)^{-1}$$

- ▶ Usando el teorema de Gauss-Markov se puede ver que OLS es ineficiente, pues la varianza no es la mínima posible (que para este caso sería $(X'\Omega^{-1}X)^{-1}$)

Mínimos Cuadrados Generalizados

- ▶ Suponiendo que conocemos Ω en una cierta escala (i.e., suponemos que $\Omega = c^2 * \Sigma$ con Σ conocido.) Es posible derivar un estimador alternativo que logra la mínima varianza.
- ▶ El procedimiento consiste en premultiplicar el modelo por $\Sigma^{-1/2}$, y obtener:

$$\tilde{y} = \tilde{X}\beta + \tilde{e}$$

donde $\tilde{y} = \Sigma^{-1/2}y$, etc.

- ▶ Se puede mostrar fácilmente que esto resulta en:

$$\tilde{\beta}_{GLS} = (X'\Sigma^{-1/2}X)^{-1}(X'\Sigma^{-1/2}y)$$

- ▶ Este estimador redundante en la menor varianza, pero como no conocemos Σ , no es factible. Pero en la práctica se utilizan varias formas de aproximar Σ .

Bondad de Ajuste

Partiendo de la relación:

$$Y_i = \hat{Y}_i + e_i$$

Se puede mostrar que se cumple:

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum e_i^2$$

Es decir:

$$SCTotales = SCExplicado + SCResiduos$$

Bondad de Ajuste

Por lo tanto, una medida de bondad de ajuste es:

$$R^2 = \frac{SC_{\text{Explicado}}}{SC_{\text{Totales}}} = 1 - \frac{SC_{\text{Residuos}}}{SC_{\text{Totales}}}$$

Bondad de Ajuste

Por definición:

$$R^2 = 1 - \frac{\sum \hat{e}_i^2}{\sum (y_i - \bar{y}_i)^2}$$

Pero cuando incluimos muchas variables esta medida pierde sentido.

Bondad de Ajuste R^2 ajustado

La arreglamos ajustando por grados de libertad:

$$\bar{R}^2 = 1 - \frac{\sum \hat{e}_i^2 / (n - k)}{\sum (y_i - \bar{y}_i)^2 / (n - 1)}$$

Test de Hipótesis

- ▶ Un test de hipótesis de especial interés será convalidar si (en la población) $\beta_j = 0$.
- ▶ Lo formalizamos con el test

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

- ▶ Podemos testearlo con:

$$t = \frac{b_j}{\text{std.error}(\beta)} = \frac{b_j}{\sqrt{\hat{V}ar(\beta)}} \sim T_{n-k-1}$$

Modelo de Regresión *Normal*

- Incorporando el supuesto de que los errores se distribuyen normalmente

$$\frac{\hat{\beta}_j - \beta_j}{s(\hat{\beta}_j)} = \frac{\hat{\beta}_j - \beta_j}{s \sqrt{[(X'X)^{-1}]_{jj}}} \sim \frac{N\left(0, \sigma^2 [(X'X)^{-1}]_{jj}\right)}{\sqrt{\frac{\sigma^2}{n-k} \chi_{n-k}^2} \sqrt{[(X'X)^{-1}]_{jj}}} = \frac{N(0, 1)}{\sqrt{\frac{\chi_{n-k}^2}{n-k}}} \sim t_{n-k}$$

a t distribution with $n - k$ degrees of freedom.

Figure 1: Test

Distribución T

Density of the t -distribution (red) for 1, 2, 3, 5, 10, and 30 degrees of freedom compared to the standard normal distribution (blue).

Previous plots shown in green.

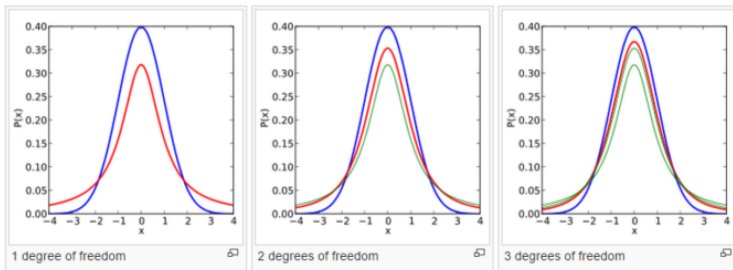


Figure 2: T-distribution

Referencias

- ▶ Hansen, B. 2018, Econometrics