Pruebas de Hipótesis

Oscar Gálvez-Soriano

University of Houston
Department of Economics

October, 2021

Matriz de Varianza-Covarianza

- STATA asume que los errores son esféricos (homoscedasticidad y no correlación serial), a menos que le especifiquemos lo contrario.
- Comando "vce" para ver la matriz de varianzas y covarianzas.
- El error estándar (SE) del estimador MCO es la raíz cuadrada del elemento k-ésimo de la diagonal principal.

Pruebas de Hipótesis: t-test

Supongamos que hemos definido las siguientes hipótesis:

- Hipótesis nula. $H_0 : \beta_k = c$
- Hipótesis alternativa. $H_A : \beta_k \neq c$

Entonces el estadístico **t** es:

$$t_k = \frac{b_k - c}{SE(b_k)}$$

El cual tiene una distribución t con $(n - k)$ grados de libertad.

Pruebas de Hipótesis: t-test

¿De dónde viene el valor crítico? Tenemos que para una prueba de dos colas:

$$Pr\left(|t_k| \geq t_{critical}\right) = 2\Phi\left(-t_{critical}\right)$$

Por ejemplo, para una prueba al 5% de significancia (95% de confianza):

$$0.05 = 2\Phi\left(-t_{critical}\right) \;\Rightarrow\; 0.025 = \Phi\left(-t_{critical}\right)$$

$$\Rightarrow\; -\Phi^{-1}\left(0.025\right) = t_{critical} \;\Rightarrow\; t_{critical} = 1.96$$

Cuyo valor se puede consultar en una tabla de t.

Pruebas de Hipótesis: p-value

El p-value calcula el nivel de significancia al cual se rechazaría la hipótesis nula.

$$p = 2\Phi\left(-t_k\right)$$

Por ejemplo, suponga que una vez calculado el estimador de MCO, encuentra que el estadístico t es 1.72.

$$p = 2\Phi\left(-1.72\right) = 0.0854$$

Pruebas de Hipótesis: Intervalo de confianza

El intervlo de confianza para $\hat{\beta}_k$ está definido como:

$$\left[\hat{\beta}_k - t_{critical} * SE(\hat{\beta}_k), \hat{\beta}_k + t_{critical} * SE(\hat{\beta}_k)\right]$$

Por ejemplo, el intervalo al 95% de confinza:

$$\left[\hat{\beta}_k - 1.96 * SE(\hat{\beta}_k), \hat{\beta}_k + 1.96 * SE(\hat{\beta}_k)\right]$$

Si el valor hipotético del estimador MCO cae dentro del intervalo, entonces rechazamos la hipótesis nula al 5% de significancia.

Pruebas de Hipótesis: Coeficientes múltiples

Considere el siguiente modelo de salario como función de raza y educación:

$$ln(\omega_i) = \beta_0 + \beta_1 \cdot black_i + \beta_2 \cdot whiteduc_i + \beta_3 \cdot blackeduc_i + \varepsilon_i$$

Podríamos estar interesados en la hipótesis de que los retornos escolares son los mismos para negros y blancos.

- Hipótesis nula. $H_0 : \beta_2 = \beta_3$
- Hipótesis alterna. $H_A : \beta_2 \neq \beta_3$

Pruebas de Hipótesis: Coeficientes múltiples

Que también puede escribirse como:

- Hipótesis nula. $H_0 : \beta_2 - \beta_3 = 0$
- Hipótesis alterna. $H_A : \beta_2 - \beta_3 \neq 0$

Entonces:

$$t_k = \frac{\hat{\beta}_2 - \hat{\beta}_3}{SE(\hat{\beta}_2 - \hat{\beta}_3)}$$

Pero necesitamos encontrar $SE(\hat{\beta}_2 - \hat{\beta}_3)$.

Pruebas de Hipótesis: Coeficientes múltiples

Sabemos que la varianza de la diferencia de dos variables aleatorias es:

$$Var(\hat{\beta}_2 - \hat{\beta}_3) = Var(\hat{\beta}_2) + Var(\hat{\beta}_3) - 2Cov(\hat{\beta}_2, \hat{\beta}_3)$$

Por lo que tendríamos que:

$$t_k = \frac{\hat{\beta}_2 - \hat{\beta}_3}{\left[Var(\hat{\beta}_2) + Var(\hat{\beta}_3) - 2Cov(\hat{\beta}_2, \hat{\beta}_3)\right]^{1/2}}$$

Veamos este ejemplo en STATA.