Notas de Clase: Econometría

- Clase: Tests de Hipótesis.
- Fecha: 17 Abril de 2023
- Profesor Ricardo Pasquini. FCE Universidad Austral

Qué sabemos sobre β poblacional?

- Nuestro interés es poder hacer decir algo sobre β (i.e., el valor del coeficiente poblacional) usando los valores de $\hat{\beta}$ (valores de los coeficientes estimados).
- Anteriormente mostramos que podíamos decir dos cosas. El estimador $\hat{\beta}$ tenía las siguientes propiedades:
    - $E[\hat{\beta}] = \beta$ (es *insesgado*)
    - $Var[\hat{\beta}] = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}$,
        - donde $\sigma^2 = Var(\epsilon)$
        - Nótese que la varianza del estimador depende de la varianza que tenga el error en la población (n.b., a priori desconocida), e inversamente de la variabilidad de la variable explicativa.
        - Para tener una estimación de la $Var[\hat{\beta}]$ podemos usar el valor estimado para la varianza del error $\hat{\sigma^2}$. Esa estimación es fácil de implementar, ya que:
            1. anteriormente vimos que una vez que el modelo estaba estimado, podíamos fácilmente conocer el *error estimado* de cada dato en mi muestra como la diferencia entre el valor de y y su predicción $\hat{y}$ (i.e., $\hat{\epsilon} = y - \hat{y}$), y
            2. podemos calcular la varianza muestral de los errores estimados: $\hat{\sigma^2} = \frac{\sum(e_i - \bar{e})^2}{n-1} = \frac{\sum e_i^2}{n-1}$
        - A la raíz de la varianza estimada ($\sqrt{\hat{Var}[\hat{\beta}]}$), se la conoce comúnmente como error estándar o **standard error**.
            - Seria algo así como el desvío estándar de coeficiente estimado, pero toma este nombre en la literatura.
            - Cuidado: Si bien se usa la palabra "error", no se refiere directamente al error del modelo, sino al desvío del coeficiente estimado.

Podemos decir algo más, por ejemplo, sobre un valor puntual de β en la población?

- Para ello tenemos el Test de Hipótesis, un ejercicio estadístico una lógica específica.
- Para entender el funcionamiento del Test, tendremos que saber que:
    1. No nos alcanza con conocer la esperanza y varianza del estimador, sino que necesitamos conocer toda la distribución de $\hat{\beta}$.

    - Vamos a explorar qué supuestos o resultados estadísticos nos permiten conocer esa distribución.
2. Tendremos que entender la lógica del test, incluyendo el paso a paso de su implementación.

- Veremos cada parte a continuación.

La distribución de $\hat{\beta}$

- Para contar con una distribución para $\hat{\beta}$, en la prctica econométrica se siguen varios caminos:
    1. Se hace un supuesto sobre la distribución del error.
    2. Se apela a resultados distribucionales que aplican cuando la muestra es grande (teoría asintótica)
    3. Se utilizan métodos computacionales que reutilizan los datos de la muestra múltiples veces (por ejemplo, el Bootstrap).
- Hoy veremos sólo el primer método y en la medida que haya tiempo en el curso, cubriremos el resto.

El supuesto de normalidad

- El primer método consiste en hacer un supuesto sobre la distribución del error. Supondremos que se distribuye normalmente.
- Si el error se distribuye normalmente entonces $\hat{\beta}$ se distribuye normalmente (no lo demostramos).

$$\hat{\beta} \sim N(\beta, Var(\beta)) \tag{1}$$

    - Es decir, se distribuye normalmente con centro en el valor esperado de beta, y varianza dada por la varianza de beta.
- Por las propiedades de la Normal, sabemos que el valor estandarizado del coeficiente, sigue la distribución *Normal Estándar*

$$Z = \frac{\hat{\beta} - \beta}{\sqrt{Var(\beta)}} \sim N(0,1) \tag{2}$$

- Es decir, podríamos usar la distribución normal estándar como base para entender el comportamiento del estimador.. pero hay un problema... hay dos cosas que necesitamos para estandarizar el coeficiente que no conocemos:
    - 1) el valor poblacional de beta, y
    - 2). σ^2 (la varianza del error), que como vimos más arriba es necesario para conocer la varianza de beta.

- Afortunadamente podemos solucionar los dos problemas:
  - 1) El valor poblacional de beta va a ser un supuesto que realizaremos usando la misma lógica del test de hipótesis (ver más adelante).
  - 2) Más arriba ya vimos que podíamos estimar σ^2 en base a los datos.
    - Este último procedimiento, sin embargo, no es inocuo. Se puede demostrar (no lo hacemos) que cuando usamos el valor estimado, el coeficiente de beta ahora se distribuye T-student con n-1 grados de libertad:
    - $$T = \frac{\hat{\beta} - \beta}{\sqrt{Va\hat{r}(\beta)}} \sim T_{n-1} \tag{3}$$

La lógica del Test de Hipótesis

- La lógica del Test de Hipótesis es la siguiente: 1) Voy a realizar una serie de supuestos que me permitirán arribar a una distribución para el estimador. Es decir, me permitirán decir con qué probabilidad espero observar cada valor del estimador. 2) Voy a ir a los datos, obtener el estimador, y ver si esa evidencia me permite rechazar los supuestos. Notar que:
  - Todo lo que voy a poder hacer con el test es rechazar (o no rechazar) un supuesto, en particular, un supuesto que me sea de interés de investigación. Es decir, nunca voy a concluir que el supuesto es válido, sino, a lo sumo, que no tengo evidencia para rechazarlo.
- En este contexto, el supuesto que voy a ser es un valor determinado para β (en la población).
  - Por ejemplo, típicamente de interés es probar el supuesto de que $\beta = 0$
    - La razón por la que este test es típicamente de interés es porque, si fuera válido, implicaría que, de acuerdo a nuestro modelo, X no tiene efecto sobre Y. Y en muchas circunstancias de investigación queremos saber si hay o no hay efecto!
    - Denotamos este test como $H_0 : \beta = 0$ versus $H_a : \beta \neq 0$.
  - Pero también me podrían interesar otros valores ($\beta = 1$ en el caso de que el modelo esté probando un efecto tipo elasticidad - ver ejemplos práctica-).
- Una vez que hice ese supuesto, casi que cuento con una distribución para el estimador $\hat{\beta}$. Estrictamente cuento con una distribución para el valor de $\hat{\beta}$ transformado, T, que me servirá para lo mismo. Veamos:
  - Para simplificar la exposición supongamos que queremos testear $H_0 : \beta = 0$. Notemos que:
  - Por lo que dijimos anteriormente sabemos que T tiene distribución T-student:

$$\hat{T} = \frac{\hat{\beta} - 0}{\sqrt{Va\hat{r}(\beta)}} \sim T_{n-1} \tag{4}$$

    - Notar que reemplacé $\beta = 0$.

        - El estadístico T no es más que el coeficiente estimado $\hat{\beta}$ divido en lo que anteriormente llamamos el **standard error** de beta.
        - Notar que cuando $\hat{\beta}$ es grande, también $\hat{T}$ lo es. Lo que ahora sé adicionalmente es con cuanta probabilidad va a ocurrir cada valor de $\hat{T}$.

- ¿Cuándo rechazar el supuesto? La idea que propone el Test es la siguiente: Vamos a rechazar el supuesto, si observamos un valor de $\hat{T}$ que ocurriría con muy poca probabilidad.
    - Veamos un ejemplo:
        - Supongamos que $\beta = 0$. Si $\beta = 0$ entonces lo que espero encontrar $\hat{\beta}$ muy cercano a 0 con alta probabilidad. Siguiendo la definición de $\hat{T}$ esto también implica que espero encontrar T cercano a 0.
        - Supongamos adicionalmente que encontramos un valor de $\hat{\beta} = 10$ y además medimos que $\sqrt{V\hat{ar}(\beta)} = 2$, con $n = 100$. Es decir $\hat{T} = 5$. ¿Cuán probable es encontrar $\hat{T} = 5$ en una distribución T-student con 99 grados de libertad? La respuesta rápida (ustedes pueden buscar ese valor) es que es muy improbable (ocurre con mucho menos del 1% de las chances). Y si es muy improbable entonces mejor rechazar el supuesto de partida (i.e. rechazar la hipótesis nula).
    - En general vamos a utilizar dos criterios para decidir cuándo un resultado es improbable:
        1. Estableciendo una región de rechazo. Vamos a identificar en la distribución de T cuales son los valores a partir de los cuales ocurren resultados con menos de un *nivel de significancia* (por ejemplo, un valor usual es rechazar si el valor ocurre con menos de un 5% de las chances. Bajo el supuesto de $\beta = 0$ rechazamos al 5% si encontramos un valor de T mayor que 1.96 o menor que -1.96)
        2. Midiendo la probabilidad de ocurrencia del valor que obtuvimos para el valor que efectivamente medimos ($\hat{T}$) en la distribución. A esto se lo conoce como **P-valor** o **P-value**.
            - Como la distribución de T es contínua lo que hacemos es medir la probabilidad de obtener un valor mayor a T (y si es un test a dos-colas miramos la probabilidad de obtener un valor mayor a $\hat{T}$ y menor a -$\hat{T}$).
                - Por esta razón a veces encontrarán el P-valor denotado como:
                    - $P(\hat{T} > T)$ o $P(|\hat{T} > T|)$.
            - Si el P-valor es muy bajo (por ejemplo menor al 5% o 1% de acuerdo a la convención de lo que consideremos bajo, entonces rechazamos el supuesto - la hipótesis nula-)
- Para ver ejemplos gráficos de la distribución T, de las regiones de rechazo y de las mediciones de P-valor, no dejen de leer el capítulo 4 del libro de Wooldridge.