Solutions Problem Set 3 (*Bootstrap*)

Tatiana Rosá Alejo Eyzaguirre

PUC

17th November 2021

Outline

Pregunta 1: Bootstrap Paramétrico

Consideremos el modelo, donde las observaciones son *i.i.d.*:

$$y_i = x_i'\beta + \exp(x_i'\alpha)u_i \quad , \quad \forall i \in \{1, N\} \quad \text{y} \quad \mu_i | x_i \sim \mathcal{N}(0, 1)$$

En los 3 ejercicios x_i contiene K-1 regresores más una constante. Sea $\hat{\beta}$ la estimación por OLS de β y sea $\hat{\alpha}$ un estimador consistente de α. Queremos aproximar la distribución de β en muestras finitas. Para eso vamos a usar el *bootstrap paramétrico*.

Primero simulamos B muestras, $(u_1^{(s)}$, $u_2^{(s)}$, ..., $u_N^{(s)})$, $\forall$ $s \in \{1, ..., S\}$ mediante la extracción de variables aleatorias normales estándar independientes. Ahí computamos:

$$y_i^{(s)} = x_i'\hat{\beta} + \exp{(x_i'\hat{\alpha})}u_i^{(s)} \tag{1}$$

para todos los s = 1, ..., B. En cada muestra vamos a calcular $\hat{\beta^{(s)}}$.

Pregunta 1.1: Demostración A

Pruebe que:

$$\hat{\beta}^{(S)} = \hat{\beta} + \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i \exp\left(x_i'\hat{\alpha}\right) u_i^{(S)}$$

Solución: Note que por la fórmula vista en clases, y usando la Expresión 1 tenemos que:

$$\begin{aligned}
\hat{\beta}^{(S)} &= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i y_i^{(S)} \\
&= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i (x_i'\hat{\beta} + \exp\left(x_i'\hat{\alpha}\right) u_i^{(s)}) \\
&= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{N} x_i x_i'\right) \hat{\beta} + \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i \exp\left(x_i'\hat{\alpha}\right) u_i^{(s)}
\end{aligned}$$

Pregunta 1.2: Distribución de $\hat{\beta}^{(s)}$

Encuentre la distribución de $\hat{\beta}^{(S)}$ condicional en la muestra original M $= (y_1, x_1, y_2, x_2, ..., y_N, x_N)$. *Ayuda:* Note que condicionando en M también condiciona en $\hat{\beta}$ y $\hat{\alpha}$.

Solución: Usando el resultado anterior y teniendo en cuenta que $E(u_i|M) = 0$, tenemos que $\hat{\beta}^{(s)}|M \sim \mathcal{N}(\hat{\beta}, V)$ donde V corresponde a:

$$Var(\hat{\beta}^{(s)}|M) = Var\left[\hat{\beta} + \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i \exp\left(x_i'\hat{\alpha}\right) u_i^{(s)} \middle| M\right]$$

Como $Var(\hat{\beta}|M) = 0$ tenemos que:

$$Var(\hat{\beta}^{(s)}|M) = \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{N} x_i \exp\left(x_i'\hat{\alpha}\right)\right)^{2} Var\left[u_i^{(s)}|M\right] \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1}$$

Usando propiedades de las potencias y el hecho de que $Var(u_i|x_i) = 1$ tendríamos que: $Var(\hat{\beta}^{(s)}|M) = \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{N} x_i x_i' \exp\left(2\hat{\alpha} x_i'\right)\right) \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1}$

Pregunta 1.3: Distribución Asintótica de $(\hat{\beta}^{(s)} - \hat{\beta})$

Solución: Usando Teorema Central del Límite y *Slutsky* tenemos que:

$$\sqrt{N}(\hat{\beta}^{(s)} - \hat{\beta}) \xrightarrow{d} \mathcal{N}(0, Var(\hat{\beta}^{(s)} - \hat{\beta}))$$

Con
$$
\begin{aligned}
Var(\hat{\beta}^{(s)} - \hat{\beta}) &= Var\left[\left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} x_i \exp{(x_i'\hat{\alpha})} u_i^{(s)}\right] \\
&= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} Var\left[\sum_{i=1}^{N} x_i \exp{(x_i'\hat{\alpha})} u_i^{(s)}\right] \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \\
&= \left(\frac{1}{N}\sum_{i=1}^{N} x_i x_i'\right)^{-1} Var\left[\frac{1}{N}\sum_{i=1}^{N} x_i \exp{(x_i'\hat{\alpha})} u_i^{(s)}\right] \left(\frac{1}{N}\sum_{i=1}^{N} x_i x_i'\right)^{-1} \\
&\xrightarrow{p} \mathbb{E}(x_i x_i')^{-1} \cdot \mathbb{E}(x_i \exp{(x_i'\hat{\alpha})} u_i) \cdot \mathbb{E}(x_i x_i')^{-1}
\end{aligned}
$$

Esta corresponde a la fórmula de **White** en el modelo paramétrico particular que hemos considerado. Entonces el bootstrap paramétrico que proponemos entregará inferencia válida sobre $\hat{\beta}$.

Pregunta 2: Bootstrap Residual

Consideremos el modelo, donde las observaciones son *i.i.d.*:

$$y_i = x_i'\beta - u_i \ \ \forall i \in \{1, N\}$$

Donde las observaciones son *i.i.d.* y,

$$\mu_i | x_i \sim \mathcal{N}(0, 1)$$

Sea $\hat{\beta}$ la estimación por OLS de β. Queremos aproximar la distribución de $\hat{\beta}$ en muestras finitas. Para eso vamos a usar el *residual bootstrap*.

Primero estimamos $\hat{\beta}$ por OLS y calculamos $\hat{u}_i = y_i - x_i'\beta, \forall\, i \in \{1, N\}$. Sacamos B muestras con reemplazo ($u_1^{(s)}$, $u_2^{(s)}$, ..., $u_N^{(s)}$), $\forall\ s \in \{1, ..., S\}$ y calculamos $y_i^{(s)} = x_i'\hat{\beta} + \hat{u}_i^{(s)}$.

Finalmente calculamos $\hat{\beta}^{(s)}$ regresando y_i^s sobre x_i por OLS en cada una de las B muestras.

Pregunta 2.1: Demostración B

Muestre que:

$$\hat{\beta}^{(S)} = \hat{\beta} + \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i \hat{u}_i^{(S)}$$

Solución: Note que por la fórmula vista en clases, y reemplazando $y_i^{(s)}$ tenemos que:

$$\begin{aligned}
\hat{\beta}^{(S)} &= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i y_i^{(S)} \\
&= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i (x_i' \hat{\beta} + \hat{u}_i^{(s)}) \\
&= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{N} x_i x_i'\right) \hat{\beta} + \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i \hat{u}_i^{(s)}
\end{aligned}$$

Pregunta 2.2: Ortogonalidad

Muestre que $\mathbb{E}(x_i\hat{u}_i^{(s)}) = 0$. *Ayuda:* Primero condicione en la muestra original y luego use Ley de Esperanzas Iteradas.

Solución: Sea M la muestra original, entonces como x_i contiene una constante al condicionar por M:

$$\mathbb{E}\left(x_i\hat{u}_i^{(s)}|M\right) = x_i\frac{1}{N}\sum_{j=1}^{N}\hat{u}_j = 0$$

Entonces por Ley de Esperanzas Iteradas, tenemos que:

$$\mathbb{E}\left(x_i\hat{u}_i^{(s)}\right) = \mathbb{E}\left(\mathbb{E}(x_i\hat{u}_i^{(s)}|M)\right) = 0$$

Pregunta 2.3: Consistencia $\hat{\beta}^{(s)}$

Muestre que $\hat{\beta}^{(s)}$ es consistente.

Solución: Aplicamos *plim* al resultado de la sección 2.1.

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}^{(s)} = \operatorname*{plim}_{N\to\infty} \hat{\beta} + \operatorname*{plim}_{N\to\infty} \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \sum_{i=1}^{N} x_i \hat{u}_i^{(s)}$$

Entonces, si usamos el resultado de la sección anterior ($\mathbb{E}(x_i\hat{u}_i^{(s)}) = 0$) y tomando en cuenta que $\hat{\beta}$ es consistente (OLS):

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty} \hat{\beta}^{(s)} &= \beta + \left(\mathbb{E}(x_i x_i')\right)^{-1} \mathbb{E}\left(x_i \hat{u}_i^{(s)}\right) \\
&= \beta + \left(\mathbb{E}(x_i x_i')\right)^{-1} \cdot 0 \\
&= \beta
\end{aligned}$$

Q.E.D.

Pregunta 2.4: Demostración C

Muestre que $\mathbb{E}\left(x_i x_i' (\hat{u}_i^{(s)})^2\right) = \mathbb{E}\left(x_i x_i' \frac{1}{N} (\sum_{i=1}^{N} \hat{u}_i^2)\right)$.

Solución: Siguiendo la misma lógica que en 2.2 tenemos que:

$$\mathbb{E}\left(x_i x_i' (\hat{u}_i^{(s)})^2 | M\right) = x_i x_i' \frac{1}{N} \sum_{j=1}^{N} \hat{u}_j^2$$

Luego por Ley de Esperanzas Iteradas tenemos que:

$$\begin{aligned} \mathbb{E}\left(x_i x_i' (\hat{u}_i^{(s)})^2\right) &= \mathbb{E}\left(\mathbb{E}\left(x_i x_i' (\hat{u}_i^{(s)})^2 | M\right)\right) \\ &= \mathbb{E}\left(x_i x_i' \frac{1}{N} \sum_{j=1}^{N} \hat{u}_j^2\right) \end{aligned}$$

Pregunta 2.5: Distribución Asintótica

Para una s dada, encuentra la distribución asintótica de $\sqrt{N}(\hat{\beta}^{(s)} - \hat{\beta})$.

Solución: Siguiendo la misma lógica que en 1.3 tenemos que:

$$\sqrt{N}(\hat{\beta}^{(s)} - \hat{\beta}) \xrightarrow{d} \mathcal{N}(0, Var(\hat{\beta}^{(s)} - \hat{\beta}))$$

Donde, si usamos misma lógica de sándwich de las sección 1.3 y la ortogonalidad encontrada en 2.2:

$$\begin{aligned}
Var(\hat{\beta}^{(s)} - \hat{\beta}) &= \left(\frac{1}{N}\sum_{i=1}^{N} x_i x_i'\right)^{-1} Var\left[\frac{1}{N}\sum_{i=1}^{N} x_i \hat{u}_i^{(s)}\right]\left(\frac{1}{N}\sum_{i=1}^{N} x_i x_i'\right)^{-1} \\
&\xrightarrow{p} \mathbb{E}(x_i x_i')^{-1} \cdot \operatorname*{plim}_{N\to\infty}\left(\mathbb{E}\left(x_i x_i'(\hat{u}_i^{(s)})^2\right) - 0^2\right) \cdot \mathbb{E}(x_i x_i')^{-1} \\
&\xrightarrow{p} \mathbb{E}(x_i x_i')^{-1} \cdot \operatorname*{plim}_{N\to\infty} \mathbb{E}\left(x_i x_i'\left[\frac{1}{N}\sum_{i=1}^{N} \hat{u}_i^2\right]\right) \cdot \mathbb{E}(x_i x_i')^{-1}
\end{aligned}$$

Pregunta 3: Bootstrap No Paramétrico

Consideremos el modelo, donde las observaciones son *i.i.d.*:

$$y_i = x_i'\beta - u_i \ \ \forall i \in \{1, N\}$$

Donde las observaciones son *i.i.d.* y,

$$\mu_i | x_i \sim \mathcal{N}(0, 1)$$

Sea $\hat{\beta}$ la estimación por OLS de β. Queremos aproximar la distribución de $\hat{\beta}$ en muestras finitas. Para eso vamos a usar el *bootstrap no paramétrico*.

Empezando con la muestra original $M = (y_1, x_1, y_2, x_2, \ldots \ , y_N, x_N)$, sacamos B muestras con reeemplazo $M^{(s)} = (y_1^{(s)}, x_1^{(s)}, y_2^{(s)}, x_2^{(s)}, \ldots \ , y_N^{(s)}, x_N^{(s)})$, con $s \in \{1, 2, \ldots, B\}$. En cada muestra estimamos $\hat{\beta}^{(s)}$ por OLS regresando $y_i^{(s)}$ sobre $x_i^{(s)}$.

Pregunta 3.1: Demostración D

Muestre que:

$$\hat{\beta}^{(S)} = \hat{\beta} + \left(\sum_{i=1}^{N} x_i^{(s)} x_i'^{(s)}\right)^{-1} \sum_{i=1}^{N} x_i^{(s)} \hat{u}_i^{(s)} \quad \text{; Donde } u_i^{(s)} = y_i^{(s)} - x_i^{(s)} \beta$$

Solución: Siguiendo la misma lógica que en la sección 1.1 y 2.1 tenemos que:

$$\begin{aligned}
\hat{\beta}^{(S)} &= \left(\sum_{i=1}^{N} x_i^{(s)} x_i'^{(s)}\right)^{-1} \sum_{i=1}^{N} x_i^{(s)} y_i^{(S)} \\
&= \left(\sum_{i=1}^{N} x_i^{(s)} x_i'^{(s)}\right)^{-1} \sum_{i=1}^{N} x_i^{(s)} (x_i'^{(s)} \hat{\beta} + \hat{u}_i^{(s)}) \\
&= \left(\sum_{i=1}^{N} x_i x_i'\right)^{-1} \left(\sum_{i=1}^{N} x_i^{(s)} x_i'^{(s)}\right) \hat{\beta} + \left(\sum_{i=1}^{N} x_i^{(s)} x_i'^{(s)}\right)^{-1} \sum_{i=1}^{N} x_i^{(s)} \hat{u}_i^{(s)}
\end{aligned}$$

Pregunta 3.2: Demostración E

Muestre que:

$$\mathbb{E}(x_i^{(s)} x_i'^{(s)}|M) = \frac{1}{N}\sum_{i=1}^{N} x_i x_i'$$

$$\mathbb{E}(x_i^{(s)} u_i^{(s)}|M) = \frac{1}{N}\sum_{i=1}^{N} x_i y_i - \frac{1}{N}\sum_{i=1}^{N} x_i x_i' \hat{\beta} = 0$$

Solución: La primera es trivial dado que al condicionar en M $x_i^{(s)} x_i^{(s)\prime}$ es una constante, luego usando este resultado y tomando en cuenta que $u_i^{(s)} = y_i^{(s)} - x_i^{(s)}$ tenemos que:

$$\begin{aligned}\mathbb{E}\left(x_i^{(s)} \hat{u}_i^{(s)}|M\right) &= \mathbb{E}\left(x_i^{(s)} y_i^{(s)}|M\right) - \mathbb{E}\left(x_i^{(s)} x_i'^{(s)}|M\right)\hat{\beta} \\ &= \frac{1}{N}\sum_{i=1}^{N} x_i y_i - \frac{1}{N}\sum_{i=1}^{N} x_i x_i' \hat{\beta} = 0\end{aligned}$$

Lo cual es igual a 0 por propiedades de OLS.

Pregunta 3.3: Demostración F

Muestre que $\text{plim}_{N\to\infty} Var(x_i^{(s)}\hat{u}_i^{(s)}) = \mathbb{E}(u_i^2 x_i x_i')$

Solución: Primero, si usamos ley de esperanzas iteradas y el resultado de 3.2:

$$\mathbb{E}\left(x_i^{(s)}\hat{u}_i^{(s)}|M\right) = \mathbb{E}\left(\mathbb{E}(x_i^{(s)}\hat{u}_i^{(s)}|M)\right) = 0$$

Luego por la fórmula clásica de la varianza —$Var(a) = \mathbb{E}(a^2) - \mathbb{E}^2(a)$— tenemos que:

$$Var(x_i^{(s)}\hat{u}_i^{(s)}) = \mathbb{E}\left(x_i^{(s)}x_i'^{(s)}(y_i^{(s)} - x_i'^{(s)}\hat{\beta})^2\right) - 0^2$$

Pero como el $plim_{N\to\infty}\hat{\beta}$ es β entonces cuando N crece la varianza de $x_i^{(s)}\hat{u}_i^{(s)}$ tenderá a:

$$\mathbb{E}\left(x_i^{(s)}x_i'^{(s)}(y_i^{(s)} - x_i'^{(s)}\hat{\beta})^2\right) \xrightarrow{p} \mathbb{E}\left(x_i x_i'(y_i - x_i'\hat{\beta})^2\right) = \mathbb{E}(x_i x_i' u_i^2)$$

Pregunta 3.4: Distribución Asintótica

Para una s dada, encuentra la distribución asintótica de $\sqrt{N}(\hat{\beta}^{(s)} - \hat{\beta})$.

Solución: Siguiendo la misma lógica que en 1.3 tenemos que:

$$\sqrt{N}(\hat{\beta}^{(s)} - \hat{\beta}) \xrightarrow{d} \mathcal{N}(0, Var(\hat{\beta}^{(s)} - \hat{\beta}))$$

Donde, si usamos misma lógica de sándwich de las sección 1.3 y 2.5:

$$\begin{aligned} Var(\hat{\beta}^{(s)} - \hat{\beta}) &= \left(\frac{1}{N}\sum_{i=1}^{N} x_i^{(s)} x^{(s)}{}_i'\right)^{-1} Var\left[\frac{1}{N}\sum_{i=1}^{N} x_i^{(s)} u_i^{(s)}\right] \left(\frac{1}{N}\sum_{i=1}^{N} x_i^{(s)} x^{(s)}{}_i'\right)^{-1} \\ &\xrightarrow{p} \mathbb{E}(x_i x_i')^{-1} \cdot \operatorname*{plim}_{N\to\infty} Var(x_i^{(s)} \hat{u}_i^{(s)}) \cdot \mathbb{E}(x_i x_i')^{-1} \\ &\xrightarrow{p} \mathbb{E}(x_i x_i')^{-1} \cdot \mathbb{E}\left(u_i^2 x_i x_i'\right) \cdot \mathbb{E}(x_i x_i')^{-1} \end{aligned}$$

Esta corresponde a la fórmula de **White**. El bootstrap no paramétrico entrega inferencia válida sobre $\hat{\beta}$ la cual es robusta a heterocedasticidad condicional. Esto se debe a que “resampleamos” y_i y x_i al mismo tiempo.

Gracias!

jeeyzaguirre@uc.cl