Fundamentos de econometría

Formulario · Primavera 2021

Carlos Lezama EST - 21104

ITAM

1. Fundamentos estadísticos de econometría

Definición 1.1 (Error cuadrático medio). *Definimos el* ***error cuadrático medio*** *como sigue:*

$$\mathrm{ECM}(\hat{\theta}) = \mathbb{E}_{\theta}\left[\left(\hat{\theta} - \theta\right)^2\right] = \mathrm{Var}(\hat{\theta}) + \mathrm{Bias}^2(\hat{\theta}, \theta).$$

Este mide la diferencia en media cuadrada entre nuestros valores estimados y el real. Asimismo, podemos definir el ***error cuadrático medio*** *de una variable* X *respecto a alguna constante* c *como sigue:*

$$\mathbb{E}\left[(X - c)^2\right] = \sigma_{\mathrm{x}}^2 + (c - \mu_{\mathrm{x}})^2.$$

Teorema 1.1: Error cuadrático mínimo

El valor que minimiza $\mathbb{E}\left[(X-c)^2\right]$ es μ_{x}.

Corolario 1.1: Desigualdad de Markov I

Sean φ una función monótona creciente y no negativa para los reales no negativos, X una variable aleatoria, $a \geq 0$ y $\varphi(a) > 0$, entonces:

$$\mathbb{P}\left(|X| \geq a\right) \leq \frac{\mathbb{E}\left[\varphi\left(|X|\right)\right]}{\varphi(a)}.$$

Corolario 1.2: Desigualdad de Markov II

Sea X una variable aleatoria y $a > 0$, entonces:

$$\mathbb{P}\left(|X| \geq a\right) \leq \frac{\mathbb{E}\left[|X|^r\right]}{a^r}.$$

Teorema 1.2: Desigualdad de Chebyshev

Sean X una variable aleatoria con $\mathbb{E}(X) = \mu$ y $\mathrm{Var}(X) = \sigma^2$, y $k > 0$, entonces:

$$\mathbb{P}\left(\|X - \mu\|_{\alpha} \geq k\sigma\right) \leq \frac{1}{k^2},$$

o bien,

$$\mathbb{P}\left(\|X - \mu\|_{\alpha} \geq k\right) \leq \frac{\sigma^2}{k^2};$$

donde $\|\cdot\|_{\alpha}$ es la norma α.

Observación 1.1

Sean $X_1, X_2, \dots, X_n$ una colección de $n \in \mathbb{N}$ variables aleatorias, podemos expresar las funciones de densidad conjunta en términos de densidades condicionales de la siguiente manera:

$$\begin{aligned} f(x_1, x_2, \dots, x_n) &= f(x_1 \mid x_2, x_3, \dots, x_n) \cdot f(x_2, x_3, \dots, x_n) \\ &= f(x_1 \mid x_2, \dots, x_n) \cdot f(x_2 \mid x_3, \dots, x_n) \cdot \dots \cdot f(x_{n-1} \mid x_n) \cdot f(x_n). \end{aligned}$$

Teorema 1.3: Ley de esperanzas iteradas

Sean X y Y dos variables aleatorias tales que $\mathbb{E}\left[X\right]$ está definida y ambas están en el mismo espacio de probabilidad, entonces:

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]\right].$$

Proposición 1.1

Por la **ley de esperanzas iteradas**, tenemos:

$$\mathrm{Var}(X) = \mathbb{E}\left[\mathrm{Var}(X \mid Y)\right] + \mathrm{Var}\left(\mathbb{E}[Y \mid X]\right).$$

Proposición 1.2

$$\mathbb{E}[XY] = \mathbb{E}\left[X\mathbb{E}[Y \mid X]\right].$$

Proposición 1.3

Sea $\varepsilon = Y - \mathbb{E}[Y \mid X]$, entonces:

(I) $\mathbb{E}[\varepsilon \mid X] = 0$.

(II) $\mathrm{Var}(\varepsilon \mid X) = \sigma_{\mathrm{y|x}}^2$.

(III) $\mathbb{E}[\varepsilon] = 0$.

(IV) $\mathrm{Var}(\varepsilon) = 0$.

(V) $\mathrm{Cov}(h(X), \varepsilon) = 0$.

Proposición 1.4

$$\underset{h(X)}{\arg\min}\, \mathbb{E}\left[\left(Y - h(X)\right)^2\right] = \mathbb{E}\left[Y \mid X\right].$$

Definición 1.2 (Mejor predictor lineal). *Dado que* $\mathbb{E}[Y \mid X]$ *puede ser una función no lineal bastante complicada — excepto en el caso normal bivariado, consideramos* $\mathbb{E}^*[Y \mid X]$ *el* ***mejor predictor lineal*** *(BLP, por sus siglas en inglés) y lo definimos como sigue:*

$$\mathbb{E}^*\left[Y \mid X\right] = \alpha + \beta X,$$

tal que

$$\alpha = \mu_{\mathrm{y}} - \beta\mu_{\mathrm{x}} \qquad \beta = \frac{\sigma_{\mathrm{xy}}}{\sigma_x^2}.$$

Teorema 1.4

Si la esperanza condicional es lineal, esta coincide con el mejor predictor lineal.

Definición 1.3 (Independencia en distribución). *Decimos que, en un conjunto finito de* n *variables aleatorias* $\{X_1, \dots, X_n\}$*,* $X_1, \dots, X_n$ *son* ***mutuamente independientes*** *si:*

$$F_{\mathrm{x}_1, \dots, \mathrm{x}_n}(x_1, \dots, x_n) = \prod_{k=1}^{n} F_{\mathrm{x}_k}(x_k), \quad \forall x_1, \dots, x_n.$$

Definición 1.4 (Independencia en media). *Decimos que una variable aleatoria* Y *es* ***independiente en media*** *respecto a otra* X *si:*

$$\mathbb{E}[Y \mid X] = \mathbb{E}[Y].$$

Definición 1.5 (Independencia en covarianza). *Decimos que dos variables aleatorias* X *y* Y *son* ***independientes en covarianza*** *si*

$$\mathrm{Cov}(X, Y) = 0.$$

Teorema 1.5

Independencia en distribución implica independencia en medias, y esta, a su vez, implica independencia en covarianzas; pero no en el orden inverso.

Proposición 1.5

Si Y es independiente en media respecto a X, entonces:

$$\mathbb{E}[X^r Y] = \mathbb{E}[X^r]\mathbb{E}[Y], \quad \forall r \in \mathbb{R}.$$

Proposición 1.6

Si Y es independiente en distribución respecto a X, entonces:

$$\mathbb{E}[h(X)Y^s] = \mathbb{E}[h(X)]\mathbb{E}[Y^s]$$

y, por lo tanto, también se cumple

$$\mathbb{E}[X^r Y^s] = \mathbb{E}[X^r]\mathbb{E}[Y^s]$$

para cualquier función $h(X)$ y todo r y s.

Proposición 1.7

Sean $\varepsilon = Y - \mathbb{E}[Y \mid X]$ y $U = Y - \mathbb{E}^*[Y \mid X]$, entonces ε es independiente en media respecto a X y U no está correlacionado con X.

Teorema 1.6: Desigualdad Cauchy-Schwartz

Sea $\rho = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$, entonces $0 \leq \rho^2 \leq 1$.

Proposición 1.8

Sea $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces, para toda p entera no negativa, los momentos centrados son los siguientes:

$$\mathbb{E}[(X-\mu)^p] = \begin{cases} 0, & \text{si } p \text{ es impar;} \\ \sigma^p (p-1)!!, & \text{si } p \text{ es par.} \end{cases}$$

Proposición 1.9

En el caso normal, la ausencia de correlación equivale a independencia estocástica.

Proposición 1.10

Sea $g(Y \mid X) \sim \mathcal{N}(\mathbb{E}[Y \mid X], \mathrm{Var}(Y \mid X))$, entonces:

$$\mathbb{E}[Y \mid X] = \mu_{\mathrm{y}} + \rho \frac{\sigma_{\mathrm{y}}}{\sigma_{\mathrm{x}}}(X - \mu_{\mathrm{x}}),$$

y

$$\mathrm{Var}(Y \mid X) = \sigma_{\mathrm{y}}^2 (1-\rho)^2.$$

Nótese que la esperanza condicional es lineal y la varianza es constante.

Observación 1.2

En el caso normal, la función de regresión coincide con el mejor predictor lineal tal que, si $Y = \mathbb{E}[Y \mid X] + \varepsilon$ y $\mathbb{E}[\varepsilon \mid X] = 0$, $Y = \alpha + \beta X + \varepsilon$ con $\alpha = \mu_{\mathrm{y}} - \beta \mu_{\mathrm{x}}$ y $\beta = \dfrac{\sigma_{\mathrm{xy}}}{\sigma_{\mathrm{x}}^2}$.

Proposición 1.11

La distribución condicional de Y dado X es normal.

$$Y \mid X \sim \mathcal{N}(\alpha + \beta X, \sigma^2).$$

2. Estimación

Teorema 2.1

Sea X una muestra aleatoria finita con n observaciones, entonces, $x_1, x_2, \ldots, x_n$ son independientes e idénticamente distribuídas y, por lo tanto, la densidad conjunta de dicha muestra aleatoria es:

$$F_{\mathrm{x}_1,\ldots,\mathrm{x}_n}(x_1, \ldots, x_n) = \prod_{k=1}^{n} F_{\mathrm{x}_k}(x_k).$$

Definición 2.1 (Estadístico muestral). *Sea $T_n = h(X)$ una función escalar de una muestra aleatoria, entonces, este es un* ***estadístico muestral****.*

Observación 2.1

Todo estadístico muestral T_n es una variable aleatoria porque su valor es determinado por el resultado de un experimento. Asimismo, a la distribución de probabilidad de T_n se le conoce como *distribución muestral*, y está completamente determinada por $h(\cdot)$, $f(\mathrm{x})$ y n.

Proposición 2.1

La media muestral satisface las siguientes propiedades:

- Si $X \sim \text{Ber}(p)$, entonces, $n\bar{X} \sim \text{Bin}(n, p)$.
- Si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces, $\bar{X} \sim \mathcal{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$.
- Si $X \sim \text{Exp}(\lambda)$, entonces, $k\lambda\bar{X} \sim \chi^2(k)$ con $k = 2n$.

Teorema 2.2: de la media muestral

Dada un muestra aleatoria de tamaño n y cualquier población con $\mathbb{E}[X] = \mu$ y $\text{Var}(X) = \sigma^2$, la media muestral $\bar{X}$ tiene esperanza $\mathbb{E}[\bar{X}] = \mu$ y varianza $\text{Var}(\bar{X}) = \dfrac{\sigma^2}{n}$.

Definición 2.2 (Momento muestral centrado).

$$M_r = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^r.$$

Definición 2.3 (Momento muestral no centrado).

$$M_r' = \frac{1}{n}\sum_{i=1}^{n} x_i^r.$$

Definición 2.4 (Momento muestral centrado en la media poblacional).

$$M_r^* = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \mu\right)^r.$$

Nótese que este ***no*** *es un estimador muestral, pues, se requiere un parámetro poblacional. Sin embargo, les será útil para algunas demostraciones.*

3. Teoría asintótica

4. Inferencia en el modelo lineal