Fundamentos de Econometría

Ignacio Lobato

ITAM

Independencia en una Distribución Bivariada

- En la sección anterior encontramos 3 posibles respuestas a la pregunta ¿cómo está Y relacionado con X en una distribución poblacional bivariada? Estas respuestas estaban dadas por $g_2(Y/X)$, $E(Y/X)$ y el BLP $E^*(Y/X)$.
- Ahora nos preguntamos ¿qué significa decir que Y no está relacionado con X en una distribución bivariada poblacional?
- Para responder a esta última pregunta, reconoceremos 3 posibles respuestas de forma análoga a la pregunta anterior.

Independencia Estocástica

- Habíamos visto antes que una distribución de probabilidad bivariada puede escribirse como el producto de la distribución condicional con la distribución marginal:

$$f(x, y) = g_2(y/x) f_1(x)$$

  Para todo $x \neq 0$

- Entonces decimos que Y es estocásticamente independiente de X sí y solo sí:

$$g_2(y/x) = f_2(y)$$

  Para todo (x,y) tal que $f_1(x) \neq 0$.

Independencia Estocástica

- Por lo tanto, Y es estocásticamente independiente de X sí y solo sí la distribución conjunta de ambas está dada por el producto de sus distribuciones marginales:

$$f(x, y) = f_1(x) f_2(y)$$

- Además note que la independencia estocástica es una relación simétrica. Por lo tanto, podemos decir que X es estocásticamente independiente de Y sí y solo sí Y es estocásticamente independiente de X.

Independencia Estocástica

- Si Y y X son estocásticamente independientes, tenemos las siguientes implicancias:

  I1: Si A es un evento definido solo en términos de X y B es un evento definido solo en términos de Y, luego $Pr(A \cap B) = Pr(A)Pr(B)$. Es decir, A y B son eventos independientes.

  I2: Si $Z = h(X)$ entonces Z y Y son estocásticamente independientes.

  I3: Sea $Z_1 = h_1(X)$ una función de X, y $Z_2 = h_2(Y)$ una función de Y. Luego Z_1 y Z_2 son independientes.

- La independencia estocástica será útil para 2 fines básicos: 1) para describir una ausencia de relación entre variables aleatorias y 2) para obtener la densidad conjunta de dos variables dadas sus densidades marginales.

Independencia en Media

- **Independencia en Media:** En cualquier distribución de probabilidad bivariada, tenemos que:

$$E(Y/X) = \int y g_2(y/x) dy$$

- Entonces decimos que Y es independiente en media de X sí y solo sí:

$$E(Y/X) = \mu_y$$

  Para todo x tal que $f(x) \neq 0$ y donde μ_y no depende de X

- Note entonces que a partir de la definición anterior y utilizando la LEI podemos obtener la siguiente relación:

$$\mu_y = E_x[E(Y/X)] = E(\mu_y) = \mu_y$$

Independencia en Media

- Otra implicación de la independencia en media es la siguiente:

  M1: Si Y es independiente en media de X, y $Z = h(X)$ es una función que depende solo de X, entonces Y es independiente en media de Z.

- ¿Cuál es la relación entre Independencia Estocástica y la Independencia en Media? Si dos distribuciones son iguales, deben tener la misma media. Por lo tanto, independencia estocástica implica independencia en media.

- Pero si dos distribuciones tienen la misma media, pueden tener distinta varianza o diferentes momentos. Por lo tanto, la independencia en media es más débil que la independencia estocástica.

- Note también que la independencia en media no es una relación simétrica como si lo es la independencia estocástica.

No-correlación

- Recordemos la definición de la covarianza en una distribución de probabilidad bivariada:

$$C(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

- Decimos que Y no está correlacionado con X sí y solo sí $C(X, Y) = 0$. Claramente esta relación es simétrica.
- ¿Cuál es la conexión entre independencia en media de X, y la ausencia de correlación entre X y Y? La respuesta a esta pregunta está dada por los siguientes resultados:

- **M2:** Si Y es independiente en media de X, luego Y no está correlacionado con X.

  Demostración: Por el T12, $C(X, Y) = C(X, E(Y/X))$. Si $E(Y/X) = \mu_y$ para todo X, luego $C(X, Y) = C(X, \mu_y) = 0$, pues μ_y es constante.

No-correlación

- **M3:** Si Y es independiente en media de X y Z=h(X) es una función que solo depende de X, entonces Y no está correlacionado con Z.

  Demostración: Por M1, Y es independiente en media de Z. Luego usamos M2 con Z jugando el rol de X.

- De los 2 resultados anteriores, se puede ver claramente que la no-correlación es más fuerte que la independencia en media.

- **M4:** Si Y no está correlacionado con $E(Y/X)$, entonces Y es independiente en media de X.

  Demostración: Sea $Z = E(Y/X)$, entonces $Y = Z + \epsilon$ con $C(Z, \epsilon) = 0$. Entonces $C(Y, Z) = C(Z + \epsilon, Z) = C(Z, Z) + C(Z, \epsilon) = V(Z) \geq 0$, con igualdad sí y solo sí Z es constante, lo cual significa que E(Y/X) es constante y por lo tanto no depende de X, lo cual es equivalente a que Y es independiente en media de X.

Tipos de Independencia

- Una forma útil de distinguir entre independencia estocástica, independencia en media y no-correlación es a través de los momentos de la distribución conjunta, llamados $E(X^r Y^s)$, donde r y s son enteros positivos. Así, tendremos que:
- Si Y no está correlacionado con X, entonces $E(XY) = E(X)E(Y)$.
- Si Y es independiente en media de X, entonces para todo r, se tiene que:

  $$E(X^r Y) = E(X^r)E(Y)$$

  Para ver este resultado, hay que usar la LEI ($h(X) = X^r$).
- Si Y es independiente estocásticamente de X, entonces:

  $$E(h(X)Y^s) = E(h(X))E(Y^s)$$

  para cualquier función $h(X)$. En particular, en este caso **sí** se cumple que para todo r y s:

  $$E(X^r Y^s) = E(X^r)E(Y^s)$$

Independencia y Predicción

- Sabemos que si Y es independiente en media de X, entonces $E(Y/X) = E(Y)$ para todo valor de X, luego la función de esperanza condicional de Y dado X es una línea horizontal, y el mejor predictor de Y dado X es E(Y).
- Si Y no está correlacionado con X, entonces $E^*(Y/X) = E(Y)$ para toda X, luego el mejor predictor lineal de Y dado X es una línea horizontal dada por E(Y).
- Note que $\epsilon = Y - E(Y/X)$ es independiente en media de X.
- Asimismo, $U = Y - E^*(Y/X)$ no está correlacionado con X.

Fuerza o Magnitud de una relación

- En algunos contextos resulta importante medir la dependencia entre X y Y en una población bivariada. Así, dado el coeficiente de correlación ρ, que se calcula como:

$$\rho = C(X,Y)/[\sqrt{V(X)}\sqrt{V(Y)}]$$

entonces: $0 \leq \rho^2 \leq 1$ **(Desigualdad de Cauchy-Schwartz)**

Demostración: Sabemos que $\beta = C(X,Y)/V(X)$. Por lo tanto:

$$\rho^2 = C^2(X,Y)/[V(X)V(Y)] = \beta^2 V(X)/V(Y)$$

$$= [V(Y) - V(U)]/V(Y)$$

Luego $\rho^2 = 1 - V(U)/V(Y)$. Pero $0 \leq V(U) \leq V(Y)$.

- Si $\rho^2 = 1$, diremos que X y Y están perfectamente correlacionados.

La Distribución Normal

- Recordemos que una variable aleatoria Z tiene una distribución normal estándar sí y solo sí su función de densidad es la siguiente:

$$\phi(z) = exp(-z^2/2)/\sqrt{2\pi}$$

- Se puede mostrar a partir de integración que dicha variable aleatoria Z presenta los siguientes momentos no centrados:

$$E(Z) = 0$$
$$E(Z^2) = 1$$
$$E(Z^3) = 0$$
$$E(Z^4) = 3$$

En general tendremos que:

$$E(Z^n) = (n-1)!! \quad \forall n \quad par \quad y \quad E(Z^n) = 0 \quad \forall n \quad impar$$

donde $(n-1)!! = (n-1)(n-3)(n-5)...(3)(1)$

La Distribución Normal

- Supongamos que $Z \sim$ normal estándar, y sea $X = a + bZ$, donde a y b son constantes con $b > 0$. Entonces dada la linealidad del operador esperanza y utilizando los momentos de dicha distribución tenemos que:

$$E(X) = a + bE(Z) = a,$$
$$V(X) = b^2 V(Z) = b^2$$

- De esta manera, podemos escribir la función lineal anterior de la siguiente manera:

$$X = \mu + \sigma Z$$

  con $\sigma > 0$

- De este modo, podemos obtener la función de densidad de X hallando primero su densidad acumulada:

$$G(x) = Pr(X \leq x) = Pr(\mu + \sigma Z \leq x) = Pr(Z \leq (x - \mu)/\sigma) = F(z)$$

La Distribución Normal

- De esta manera, $Z = (X - \mu)/\sigma$ sería la estandarización de la variable aleatoria X. Luego usando la densidad acumulada de la normal estándar $\Phi(.)$, tenemos que:

$$g(x) = \sigma^{-1}(2\pi)^{-1/2} exp(-z^2/2) = exp\left\{-[(x-\mu)/\sigma]^2/2\right\}/\sqrt{(2\pi\sigma^2)}$$

- Por lo tanto podemos decir que $X \sim N(\mu, \sigma^2)$. Es decir, las variables aleatorias con distribución normal están caracterizadas por 2 parámetros (μ y σ^2) y pueden escribirse en términos de una distribución normal estándar.

- Así, $X \sim N(\mu, \sigma^2)$ sí y solo sí $(X - \mu)/\sigma \sim N(0, 1)$.

La Distribución Normal

- Una implicación inmediata de lo anterior es que una función lineal de una variable que tiene distribución normal hereda dicha distribución. Es decir, si $X \sim N(\mu, \sigma^2)$ y $Y = a + bX$ con $b \neq 0$, entonces $Y \sim N(a + b\mu, b^2\sigma^2)$.

  Demostración:

  Es suficiente mostrar que $Y = a^* + b^* Z$ donde $Z \sim N(0, 1)$. Sea $Z = (X - \mu)/\sigma$ tal que $X = \mu + \sigma Z$. Luego
  $Y = a + bX = a + b(\mu + \sigma Z) = (a + b\mu) + b\sigma Z = a^* + b^* Z$

La Distribución Normal Bivariada

- Dado un vector aleatorio de N variables que se distribuyen conjuntamente como una $N(\mu_{Nx1}, \sum_{NxN})$, tendremos que su densidad conjunta es:

$$\Phi(Y) = \frac{1}{(\sqrt{2\pi})^n} \frac{1}{|\Sigma|^{1/2}} exp[-\frac{1}{2}(Y-\mu)'\Sigma^{-1}(Y-\mu)]$$

Note que en este caso $\mu_{Nx1} = [\mu_1, ..., \mu_N]'$ es un vector de medias.

- En esta sección nos centraremos en el caso N=2. De esta manera:

$$\mu_{2x1} = [\mu_1, \mu_2]' \quad \text{y} \quad \Sigma_{2x2} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

$$\Sigma_{2x2}^{-1} = \frac{1}{|\Sigma|} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

con

$$|\Sigma| = \sigma_1^2\sigma_2^2 - \sigma_{12}^2$$

Definamos $\rho^2 = \frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}$

La Distribución Normal Bivariada

- Sea $Y = [Y_1, Y_2]'$ un vector cuya distribución conjunta es normal bivariada. Entonces su densidad conjunta será:

$$\frac{1}{2\pi}(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)^{-0,5} exp\left\{-0,5\left[\frac{(Y_1-\mu_1)^2}{\sigma_1^2\sigma_2^2(1-\rho^2)}\sigma_2^2 + \frac{(Y_2-\mu_2)^2}{\sigma_1^2\sigma_2^2(1-\rho^2)}\sigma_1^2\right.\right.$$

$$\left.\left.-2\sigma_{12}\frac{(Y_1-\mu_1)(Y_2-\mu_2)}{\sigma_1^2\sigma_2^2(1-\rho^2)}\right]\right\} = f(Y_1, Y_2)$$

Sean $Z_1 = \frac{Y_1-\mu_1}{\sigma_1}$ y $Z_2 = \frac{Y_2-\mu_1}{\sigma_2}$. Entonces:

$$f(Z_1, Z_2) = \frac{1}{2\pi}(1-\rho^2)^{-0,5} exp\left[-\frac{1}{2}\frac{Z_1^2 + Z_2^2 - 2\rho Z_1 Z_2}{(1-\rho^2)}\right]$$

Ahora sea $Q = -\frac{1}{2}\frac{Z_1^2+Z_2^2-2\rho Z_1 Z_2}{(1-\rho^2)}$

La Distribución Normal Bivariada

$$
\begin{aligned}
Q &= -\frac{1}{2(1-\rho^2)}(Z_1^2 + \rho^2 Z_2^2 - 2\rho Z_1 Z_2 + (1-\rho^2)Z_2^2) \\
&= -\frac{1}{2(1-\rho^2)}((Z_1 - \rho Z_2)^2 + (1-\rho^2)Z_2^2) \\
&= -\frac{1}{2}\left[\frac{(Z_1-\rho Z_2)^2}{1-\rho^2} + Z_2^2\right]
\end{aligned}
$$

Entonces la densidad conjunta se puede expresar así:

$$
f(Y_1, Y_2) = \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\Big\{-\frac{(Z_1-\rho Z_2)^2}{2(1-\rho^2)}\Big\} exp\Big\{-\frac{1}{2}Z_2^2\Big\}
$$

Considere el caso en el cual $\rho = 0$, entonces:

$$
f(Y_1, Y_2) = \frac{1}{\sqrt{2\pi}\sigma_1} exp(-\frac{1}{2}Z_1^2) \frac{1}{\sqrt{2\pi}\sigma_2} exp(-\frac{1}{2}Z_2^2) = f(Y_1)f(Y_2)
$$

Por lo tanto, en el caso normal, la ausencia de correlación equivale a independencia estocástica.

Ahora recordemos que $f(Y_1/Y_2) = \frac{f(Y_1,Y_2)}{f(Y_2)}$. Entonces:

$$f(Y_1/Y_2) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} exp\left\{ -\frac{1}{2(1-\rho^2)}(\frac{Y_1-\mu_1}{\sigma_1} - \rho\frac{Y_2-\mu_2}{\sigma_2})^2 \right\}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} exp\left\{ -\frac{1}{2\sigma_1^2(1-\rho^2)}[Y_1-\mu_1-\rho\frac{\sigma_1}{\sigma_2}(Y_2-\mu_2)]^2 \right\}$$

Por lo tanto $f(Y_1/Y_2) \sim N(E(Y_1/Y_2), V(Y_1/Y_2))$ con:

$$E(Y_1/Y_2) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(Y_2-\mu_2)$$

La cual es lineal. Finalmente:

$$V(Y_1/Y_2) = \sigma_1^2(1-\rho^2)$$

La cual es constante.

La Distribución Normal Bivariada

- En un caso particular con $Y_1 = Y$ y $Y_2 = X$ se tendría que:

$$f(Y/X) \sim N(\mu_y + \rho \frac{\sigma_y}{\sigma_x}(X - \mu_x), \sigma_y^2(1 - \rho^2))$$

- Si $Y = E(Y/X) + \epsilon$ y $E(\epsilon/X) = 0$, entonces:

$$Y = \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x - \mu_x) + \epsilon = (\mu_y - \rho \frac{\sigma_y}{\sigma_x}\mu_x) + \rho \frac{\sigma_y}{\sigma_x}x + \epsilon = \alpha + \beta X + \epsilon$$

- Note que $\beta = \frac{\sigma_{xy}}{\sigma_x^2}$. Es decir, la función de regresión coincide con el BLP.
- Además $V(Y/X) = \sigma_y^2(1 - \rho^2) < \sigma_y^2 = V(Y)$
- Asimismo $C(\epsilon, X) = C(Y - \alpha - \beta X, X) = \sigma_{xy} - \beta \sigma_x^2 = 0$.

La Distribución Normal Bivariada

- De este modo, si $(Y_1, Y_2) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \sigma_{12})$ entonces:
- **P1:** Las esperanzas, varianzas y covarianza son: $E(Y_1) = \mu_1$, $E(Y_2) = \mu_2$, $V(Y_1) = \sigma_1^2$, $V(Y_2) = \sigma_2^2$ y $C(Y_1, Y_2) = \sigma_{12}$.
- **P2:** La distribución marginal de Y_1 es normal: $Y_1 \sim N(\mu_1, \sigma_1^2)$
- **P3:** La distribución condicional de Y_2 dado Y_1 es normal: $Y_2/Y_1 \sim N(\alpha + \beta Y_1, \sigma^2)$. Es decir, la función de regresión es lineal y la varianza condicional es constante.

- **P4:** No-correlación implica independencia estocástica: si $\sigma_{12} = 0$, entonces Y_1 y Y_2 son independientes.

- **P5:** Dado un par de funciones lineales de X_1 y X_2, tal que $(X_1, X_2) \sim BVN$: si $Y_1 = a_1 + b_1X_1 + c_1X_2$ y $Y_2 = a_2 + b_2X_1 + c_2X_2$ donde a_i, b_i, c_i son constantes para $i = 1, 2$ con $b_1c_2 - b_2c_1 \neq 0$ (lo último para descartar los casos en los cuales hay perfecta correlación o las variables son constantes), entonces $(Y_1, Y_2) \sim BVN$.

  La prueba de esta propiedad se realiza expresando Y_1 y Y_2 como funciones lineales de W_1 y W_2 tal que $(W_1, W_2) \sim SBVN(\rho)$, es decir, expresando las variables como funciones lineales de normales estándar.