Fundamentos de Econometría

Ignacio Lobato

ITAM

Esperanzas en el Caso Bivariado

- Suponemos que hay un vector aleatorio (X,Y) que tiene una función de densidad/masa conjunta f(x,y). Luego la esperanza de una variable Z=h(X,Y) se calcula como:

$$E(Z) = \begin{cases} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dydx \text{ en el caso continuo,} \\ \sum_i \sum_j h(x_i,y_j)f(x_i,y_j) \text{ en el caso discreto} \end{cases}$$

- Si Z está en función de solo una de las variables, entonces su esperanza se calcula utilizando la función de densidad/masa marginal de dicha variable. Es decir, si h(x,y)=$h_1(x)$, *luego* :

$$E(Z) = \sum_i \sum_j h_1(x_i)f(x_i,y_j) = \sum_i h_1(x_i)[\sum_j f(x_i,y_j)] = \sum_i h_1(x_i)f_1(x_i)$$

Momentos de una Distribución Conjunta

- Los momentos de una distribución conjunta son esperanzas de ciertas funciones de (X,Y), o de (X^*, Y^*), donde $X^* = X - E(X)$, $Y^* = Y - E(Y)$. En particular, para valores no-negativos r,s:

.

- $E(X^r Y^s)$ es el momento (r,s) no centrado o centrado en cero,
- $E(X^{*r} Y^{*s})$ es el momento (r,s) centrado en la media.
  En particular:
- r=1, s=0: $E(X^1 Y^0) = E(X) = \mu_x$,
- r=2, s=0: $E(X^{*2} Y^{*0}) = E[X - E(X)]^2 = V(X) = \sigma_x^2$
- r=1, s=1: $E(X^* Y^*) = E[(X - E(X))(Y - E(Y))] = C(X, Y) = \sigma_{xy}$
- La última es la denominada covarianza de X y Y.

Teoremas con Esperanzas Bivariadas

- σ_x es denominada la desviación estándar de X
- $\rho = C(X,Y)/[\sqrt{V(X)}\sqrt{V(Y)}] = \sigma_{xy}/(\sigma_x\sigma_y)$ es denominado el coeficiente de correlación de X y Y.
- **T5: Funciones Lineales:** Supongamos que Z=a+bX+cY, donde a, b, c son constantes. Luego:

$$E(Z) = a + bE(X) + cE(Y),$$
$$V(Z) = b^2V(X) + c^2V(Y) + 2bcC(X,Y)$$

Teoremas con Esperanzas Bivariadas

Demostración: Por esperanzas:

$$E(Z) = \sum_i \sum_j (a + bx_i + cy_j) f(x_i, y_j)$$

$$= a \sum_i \sum_j f(x_i, y_j) + b \sum_i x_i \left[\sum_j f(x_i, y_j) \right] + c \sum_j y_j \left[\sum_i f(x_i, y_j) \right]$$

$$= 1 + b \sum_i x_i f_1(x_i) + c \sum_j y_j f_2(y_j)$$

Para el caso de las varianzas definimos $V(Z) = E[Z^{*2}]$, donde $Z^* = Z - E(Z) = bX^* + cY^*$ y reemplazando dicha expresión al cuadrado se obtiene lo enunciado en el teorema.

- **T6: Pares de funciones lineales:** Supongamos que: $Z_1 = a_1 + b_1 X + c_1 Y$ y $Z_2 = a_2 + b_2 X + c_2 Y$ donde a_i, b_i, c_i son constantes para $i = \{1, 2\}$. Luego:

$$C(Z_1, Z_2) = b_1 b_2 V(X) + c_1 c_2 V(Y) + (b_1 c_2 + b_2 c_1) C(X, Y)$$

Demostración:
La demostración es análoga a la demostración del T5 para el caso de la varianza.

- **T7: Varianza y Covarianza:**

C(X,Y)=E(XY)-E(X)E(Y)=C(Y,X)

V(X)=E(X^2) $- E^2(X) = C(X,X)$

Demostración:

$$C(X,Y) = E(X^*Y^*) = E\{[X - E(X)]Y^*\} = E(XY^*) - E(X)E(Y^*)$$
$$= E(XY^*) = E\{X[Y - E(Y)]\} = E(XY) - E(X)E(Y) = C(Y,X).$$

Esperanzas Condicionales

- Sea el vector aleatorio (X,Y) con distribución de densidad/masa conjunta $f(x,y) = g_2(y/x)f_1(x)$, y sea $Z = h(X,Y)$ una función de X y Y. Luego **la esperanza condicional de Z, dado X=x** es:

$$E(Z/x) = \int_{-\infty}^{\infty} h(x,y)g_2(y/x)dy$$

- Lo anterior es para el caso continuo, el caso discreto se calcularía de forma análoga.
- El símbolo $E(./x)$ denota la expectativa tomada en la distribución $g_2(y/x)$.

Resultados de las Expectativas Condicionales

(i) Sea $Z = h(X)$. Entonces $E(Z/x) = h(x)$.

(ii) Sea $Z = h_1(X)Y$. Entonces $E(Z/x) = h_1(x)E(Y/x)$.

(iii) Sea $Z = a + bX + cY$. Entonces $E(Z/x) = a + bx + cE(Y/x)$.

(iv) Sea $Z = Y$. Entonces $E(Z/x) = E(Y/x)$. En adelante se usará la notación $E(Y/x) = \mu_{y/x}$.

(v) Sea $Z = (Y - \mu_{y/x})$. Entonces $E(Z/x) = E(Y/x) - \mu_{y/x} = 0$.

(vi) Sea $Z = (Y - \mu_{y/x})^2$. Entonces $E(Z/x) = V(Y/x) = \sigma^2{}_{y/x}$, la varianza condicional de Y dado X=x.

(vii) Sea $Z = (Y - \mu_y)$. Entonces:

$$E(Z/x) = E(Y/x) - \mu_y = \mu_{y/x} - \mu_y$$

(viii) Sea $Z = (Y - \mu_y)^2$. Entonces $E(Z/x) = \sigma^2{}_{y/x} + (\mu_{y/x} - \mu_y)^2$.

- **T8: Ley de Esperanzas Iteradas:** La esperanza (marginal) de Z=h(X,Y) es la esperanza de su esperanza condicional:

$$E(Z) = E_x[E(Z/X)]$$

Demostración:

$$E(Z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dydx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)[g_2(y/x)f_1(x)]dydx$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} h(x,y)g_2(y/x)dy\right] f_1(x)dx = \int_{-\infty}^{\infty} E(Z/x)f_1(x)dx.$$

- **T9: Media Marginal y Condicional:** La esperanza incondicional o marginal de Y es igual a la esperanza de su esperanza condicional:

$$\mu_y = E(Y) = E_x[E(Y/X)] = E(\mu_{y/x})$$

- **T10: Análisis de Varianza:** La varianza de Y es igual a la esperanza de su varianza condicional más la varianza de su esperanza condicional.

$$\sigma_y^2 = V(Y) = E_x[V(Y/X)] + V_x[E(Y/X)] = E(\sigma_{y/x}^2) + V(\mu_{y/x})$$

Demostración:
Escribir $V(Y) = E(Z)$ donde $Z = (Y - \mu_y)^2$ y aplicar viii) utilizando que $E(Z/x) = \sigma^2{}_{y/x} + (\mu_{y/x} - \mu_y)^2$.

Resultados a partir de la LEI

- **T11: Esperanza de un Producto:** La esperanza del producto de X y Y es igual a la esperanza del producto de X y la esperanza condicional de Y dado X:

$$E(XY) = E_x[XE(Y/X)] = E(X\mu_{y/x})$$

- **T12: Covarianza:** La covarianza de X y Y es igual a la covarianza de X y la esperanza condicional de Y dado X:

$$C(X, \mu_{y/x}) = E(X\mu_{y/x}) - E(X)E(\mu_{y/x}) = E(XY) - E(X)E(Y)$$

Función de Esperanza Condicional

- Sabemos que la esperanza condicional de Y dado X=x es

$$E(Y/x) = \mu_{y/x} = \int_{-\infty}^{\infty} y g_2(y/x) dy.$$

- Si cambiamos x, es decir, si permitimos que X varíe, obtenemos $E(Y/X) = \mu_Y$, una función de X conocida como la función de expectativa condicional (CEF) o "función de regresión poblacional" de Y dado X.
- Similarmente $V(Y/x)$ es la varianza condicional de Y dado $X = x$ y $V(Y/X)$ es la función de varianza condicional de Y dado X.

Desviación respecto a la CEF

- La desviación de Y respecto a su CEF tiene ciertas propiedades. Así, sea $\epsilon = Y - E(Y/X)$, entonces como consecuencia de que ϵ es una desviación de la esperanza condicional, tenemos que:
- $E(\epsilon/X) = 0$
- $V(\epsilon/X) = \sigma^2_{Y/X}$
- $E(\epsilon) = 0$
- $C(X, \epsilon) = 0$
- $V(\epsilon) = E(\sigma^2_{Y/X})$
- si $Z = h(X)$ entonces $C(Z, \epsilon) = 0$
- La demostración de dichas propiedades se realiza utilizando el T8 y el T12. Además, nótese que en general se concluye que la desviación de Y respecto $E(Y/X)$ es una variable aleatoria con media cero y cuya covarianza con cualquier función de X también es cero.

Predicción

- En el tema anterior se vio que dada una variable aleatoria Y con cierta distribución de probabilidad conocida, tiene como mejor predictor constante a μ_y. Es decir, μ_y minimiza $E(U^2) = E(Y - c)^2$.
- Ahora consideremos el problema de predicción para un caso bivariado. Supongamos que tenemos un vector aleatorio (X,Y) con pdf o pmf conjunta conocida f(x,y) y quisiéramos encontrar la función de X que minimiza el error cuadrático medio dado por $E(U^2)$ donde $U = Y - h(X)$.
- **Proposición:** $E(Y/X) = argmin_{h(X)} E((Y - h(X))^2)$

Predicción

- **Demostración:** Sea h(x) cualquier función de X y sea $U = Y - h(X)$. Asimismo, sea $\epsilon = Y - E(Y/X)$ y $W = E(Y/X) - h(X)$. Note que $U = \epsilon + W$, con W estando solo en función de X.
- Para una $X = x$ particular, tenemos que $W = E(Y/x) - h(x) = w$. Entonces en $X = x$, tenemos $U = \epsilon + w$, de manera que $U^2 = \epsilon^2 + w^2 + 2w\epsilon$. Entonces:

$$E(U^2/x) = E(\epsilon^2/x) + w^2 + 2wE(\epsilon/x) = \sigma^2_{y/x} + w^2$$

- Luego tomando esperanza sobre X:

$$E(U^2) = E_X[E(U^2/X)] = E(\sigma^2_{y/x}) + E(W^2).$$

- Note que el último término es no negativo y se minimiza cuando es igual a cero, lo cual ocurre cuando $h(X) = E(Y/X)$. Por lo tanto, tendríamos que $E(Y/X)$ minimiza el error cuadrático medio cuando el mismo se minimiza a través de una función de X.

El Mejor Predictor Lineal (BLP)

- Por el T10, se puede ver que $E(Y/X)$ es un mejor predictor que E(Y). Ambos predictores son insesgados, pero en general, la información adicional que pueda proveer X mejora la predicción de Y.
- Ahora supongamos que queremos encontrar un mejor predictor a partir de alguna función lineal de X: $h(X) = a + bX$, en el sentido de que minimice $E(U^2)$, donde $U = Y - h(X)$.
- Dicha mejor función lineal estará dada por $E^*(Y/X) = \alpha + \beta X$, donde α y β están dados por:

$$\beta = \sigma_{xy}/\sigma_x^2$$

y

$$\alpha = \mu_y - \beta\mu_x$$

El Mejor Predictor Lineal (BLP)

- **Demostración:** Tenemos U=Y-(a+bX), y usando la linealidad de la esperanza obtenemos las siguientes condiciones de primer orden:

$$\partial E(U^2)/\partial a = E(\partial U^2/\partial a) = 2E(U\partial U/\partial a) = -2E(U)$$

$$\partial E(U^2)/\partial b = E(\partial U^2/\partial b) = 2E(U\partial U/\partial b) = -2E(XU)$$

- Es decir, tenemos que $E(U) = 0$ y $C(X, U) = 0$. Sustituyendo para U obtenemos que:

$$E(Y) = a + bE(X),$$

$$C(X, Y) = bV(X)$$

- De esta manera, obtenemos el mejor predictor lineal caracterizado por α y β.

El Mejor Predictor Lineal (BLP)

- La desviación de Y respecto a su mejor predictor lineal, dado por $U = Y - E^*(Y/X) = Y - (\alpha + \beta X)$, tiene las siguientes propiedades:

$$E(U) = 0$$

$$C(X, U) = 0$$

$$V(U) = V(Y) - \beta^2 V(X)$$

- Note que las dos primeras propiedades están dadas por las CPO. Luego utilizando que $E(U) = 0$, tenemos que $V(U) = E(U^2)$ y obtenemos la tercera propiedad.

Esperanzas Condicionales y Predictores Lineales

- Vimos que la esperanza condicional es el mejor predictor de Y, el BLP es el mejor predictor lineal y E(Y) es el mejor predictor constante.
- Esto significa que $E(Y/X)$, $E^*(Y/X)$ y E(Y) resuelven problemas de optimización cada vez más restringidos.
- En ese sentido, es claro que el BLP no es mejor predictor que la esperanza condicional, pero es mejor (o no peor) que la esperanza marginal o no-condicional.
- Asimismo, es importante notar que $U = Y - E^*(Y/X)$ y $\epsilon = Y - E(Y/X)$ difieren por su relación con X. Así, mientras U tiene covarianza cero con X, ϵ tiene covarianza cero con cualquier función de X.

- **T13: Aproximación Lineal a la CEF:** La mejor aproximación lineal a la esperanza condicional, en el sentido de minimizar $E(W^2)$, donde $W = E(Y/X) - (a + bX)$, es el BLP $E^*(Y/X) = \alpha + \beta X$ con $\beta = C(X, Y)/V(X)$ y $\alpha = E(Y) - \beta E(X)$. La demostración de dicho teorema es análoga al mismo problema de predicción lineal pero intercambiando W por U y $E(Y/X)$ por Y.

- **T14: CEF lineal:** Si la esperanza condicional es lineal, esta coincide con el BLP. Es decir, si $E(Y/X) = a + bX$, luego $b = C(X, Y)/V(X) = \beta$ y $a = E(Y) - \beta E(X) = \alpha$.