27/feb/2019

Estadística Multivariada & Datos Categóricos

Tarea #3

Alejandra Lelo de Larrea Ibarra 124433

1) Sea $X=(X_1,X_2)$ tal que $f(x_1,x_2)=(2\pi)^{-2/2}|\Sigma|^{-1/2}\exp\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\}$

con $\mu=(\mu_1,\mu_2)'$ vector de medias & $\Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}$ tal que $var(X_1)=\sigma_{11}$

$var(X_2)=\sigma_{22}$

$cov(X_1,X_2)=\sigma_{12}=\sigma_{21}$

Obtener la distribución condicional de X_1 a $X_2=x_2$

Preliminares:

i) p.d. $|\Sigma|^{1/2}=\sigma_{11}^{1/2}\sigma_{22}^{1/2}\sqrt{1-\rho^2}$ con $\rho=corr(X_1,X_2)$

Dem:

Por un lado tenemos que $|\Sigma|^{1/2}=\left|\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}\right|^{1/2}=(\sigma_{11}\sigma_{22}-\sigma_{12}\sigma_{21})^{1/2}=(\sigma_{11}\sigma_{22}-\sigma_{12}^2)^{1/2}$ (1)

Por otro lado sea $\rho=corr(X_1,X_2)=\dfrac{cov(X_1,X_2)}{\sqrt{var(X_1)}\sqrt{var(X_2)}}=\dfrac{\sigma_{12}}{\sigma_{11}^{1/2}\sigma_{22}^{1/2}}$ (2)

entonces:

$$\sigma_{11}^{1/2}\sigma_{22}^{1/2}\sqrt{1-\rho^2}=\sigma_{11}^{1/2}\sigma_{22}^{1/2}\sqrt{1-\left(\frac{\sigma_{12}}{\sigma_{11}^{1/2}\sigma_{22}^{1/2}}\right)^2}=\sigma_{11}^{1/2}\sigma_{22}^{1/2}\sqrt{\frac{\sigma_{11}\sigma_{22}-\sigma_{12}^2}{\sigma_{11}\sigma_{22}}}$$

$$=(\sigma_{11}\sigma_{22}-\sigma_{12}^2)^{1/2}\overset{(1)}{=}|\Sigma|^{1/2} \quad //$$

$\therefore$ $|\Sigma|^{1/2}=\sigma_{11}^{1/2}\sigma_{22}^{1/2}\sqrt{1-\rho^2}$ (3)

ii) Notar que $1-\rho^2=\dfrac{\sigma_{11}\sigma_{22}-\sigma_{12}^2}{\sigma_{11}\sigma_{22}}$ (4)

3) p.d. $f(x_1,x_2)$ se puede reescribir en términos de ρ como:

$$f(x_1,x_2)=(2\pi)^{-1}[\sigma_{11}\sigma_{22}(1-\rho^2)]^{-1/2}\exp\left\{-\frac{1}{2}(1-\rho)^{-1}\left[\left(\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right)^2-2\rho\left[\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right]\left[\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right]+\left[\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right]^2\right]\right\}$$

Dem:

$$f(x_1,x_2) = (2\pi)^{-2/2}|\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)'\,\Sigma^{-1}(x-\mu)\right\}$$

De (3)

$$= (2\pi)^{-1}\left(\sigma_{11}^{1/2}\sigma_{22}^{1/2}\sqrt{1-\rho^2}\right)^{-1} \exp\left\{-\frac{1}{2}\,(x_1-\mu_1 \quad x_2-\mu_2)\left(\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{12}^2}\right)\begin{pmatrix}\sigma_{22} & -\sigma_{21}\\ -\sigma_{22} & \sigma_{11}\end{pmatrix}\begin{pmatrix}x_1-\mu_1\\ x_2-\mu_2\end{pmatrix}\right\}$$

De (4)

$$= (2\pi)^{-1}\left(\sigma_{11}\sigma_{22}(1-\rho^2)\right)^{-1/2} \exp\left\{-\frac{1}{2}\,\frac{1}{\sigma_{11}\sigma_{22}(1-\rho^2)}\,(x_1-\mu_1 \quad x_2-\mu_2)\begin{pmatrix}\sigma_{22} & -\sigma_{21}\\ -\sigma_{12} & \sigma_{11}\end{pmatrix}\begin{pmatrix}x_1-\mu_1\\ x_2-\mu_2\end{pmatrix}\right\}$$

$$= (2\pi)^{-1}\left(\sigma_{11}\sigma_{22}(1-\rho^2)\right)^{-1/2} \exp\left\{-\frac{1}{2}\,\frac{1}{\sigma_{11}\sigma_{22}(1-\rho^2)}\left[\sigma_{22}(x_1-\mu_1)^2 - 2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2) + \sigma_{11}(x_2-\mu_2)^2\right]\right\}$$

$$= (2\pi)^{-1}\left(\sigma_{11}\sigma_{22}(1-\rho^2)\right)^{-1/2} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right)^2 - \frac{2\sigma_{12}}{\sigma_{11}^{1/2}\sigma_{22}^{1/2}}\left[\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right]\left[\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right] - \left[\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right]^2\right]\right\}$$

De (2)

$$= (2\pi)^{-1}\left[\sigma_{11}\sigma_{22}(1-\rho^2)\right]^{-1/2} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right)\left(\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right) - \left(\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right)^2\right]\right\}$$

//

Distribución Condicional de X_1 en $X_2 = x_2$

$$f(x_1|x_2=x_2) = \frac{f(x_1, x_2=x_2)}{f(x_2=x_2)}$$

$$= \frac{(2\pi)^{-1}\left[\sigma_{11}\sigma_{22}(1-\rho^2)\right]^{-1/2} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_{11}^{1/2}}\right)\left(\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right) - \left(\frac{x_2-\mu_2}{\sigma_{22}}\right)^2\right]\right\}}{(2\pi)^{-1/2}\,\sigma_{22}^{-1/2}\exp\left\{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right)^2\right\}}$$

$$\propto \exp\left\{-\frac{1}{2\sigma_{11}(1-\rho^2)}\left[x_1^2 - 2x_1\mu_1 - 2\rho\,\sigma_{11}^{1/2}x_1\left(\frac{x_2-\mu_2}{\sigma_{22}^{1/2}}\right)\right]\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma_{11}(1-\rho^2)}\left[x_1^2 - 2\left(\mu_1 - \rho\frac{\sigma_{11}^{1/2}}{\sigma_{22}^{1/2}}(x_2-\mu_2)\right)x_1\right]\right\}$$

$$\propto \exp\left\{-\frac{1}{2\sigma_{11}(1-\rho^2)}\left[x_1^2 - 2\left(\mu_1 - \frac{\rho\,\sigma_{11}^{1/2}}{\sigma_{22}^{1/2}}(x_2-\mu_2)\right)x_1 + \left(\mu_1 - \frac{\rho\,\sigma_{11}^{1/2}}{\sigma_{22}^{1/2}}(x_2-\mu_2)\right)^2\right]\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma_{11}(1-\rho^2)}\left[x_1 - \left(\mu_1 - \frac{\rho\,\sigma_{11}^{1/2}}{\sigma_{22}^{1/2}}(x_2-\mu_2)\right)\right]^2\right\}$$

que coincide con el kernel de una normal con media $\mu_{1|2} = \mu_1 - \frac{\rho\,\sigma_{11}^{1/2}}{\sigma_{22}^{1/2}}(x_2-\mu_2)$ y varianza $\Sigma_{1|2} = \sigma_{11}(1-\rho^2)$, que sustituyendo el valor de ρ es equivalente a tener

$$X_1|X_2=x_2 \sim N\left(\mu_1 - \sigma_{12}\sigma_{22}^{-1}(x_2-\mu_2);\ \sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{12}\right)$$

//

2) Sea $\underline{X} = (X_1, X_2)$ un vector de dimensión p tal que X_1 es de dimensión p_1 y X_2 es de dimensión p_2 y $X \sim N_p(x \mid \mu, \Sigma)$ con $p = p_1 + p_2$

Obtener la distribución condicional de X_1 en $X_2 = x_2$

Preliminares: (Recuperado de Graybill, F. (1961))

• Teorema: Si A es una matriz simétrica, positiva definida tal que $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ y si B es la inversa de A tal que $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ y si B_{ii} y A_{ii} son ambas de dimensión $m_i \times m_i$, se satisface que

$$A_{11}^{-1} = B_{11} - B_{12} B_{22}^{-1} B_{21}$$

$$\text{y } A_{11} B_{12} = -A_{12} B_{22}$$

Dem: Dado que $A = B^{-1}$ entonces $AB = I$ por lo tanto

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \quad (1)$$

De (1) se pueden obtener las siguientes ecuaciones matriciales:

$$A_{11} B_{11} + A_{12} B_{21} = I \; (2) \quad \text{y} \quad A_{11} B_{12} + A_{12} B_{22} = 0 \; (3)$$

De (3) obtenemos que $A_{12} = -A_{11} B_{12} B_{22}^{-1}$ (4) y también que $A_{11} B_{12} = -A_{12} B_{22}$ (5)

sustituyendo (4) en (2) y premultiplicando por A_{11}^{-1} tenemos

$$A_{11} B_{11} + \left[-A_{11} B_{12} B_{22}^{-1} B_{21} \right] = I$$

$$\Rightarrow A_{11}^{-1} A_{11} B_{11} - A_{11}^{-1} A_{11} B_{12} B_{22}^{-1} B_{21} = A_{11}^{-1} I$$

$$\Rightarrow B_{11} - B_{12} B_{22}^{-1} B_{21} = A_{11}^{-1}$$

Dado que A_{11} y B_{22} son matrices spd y son los menores principales de A y B respectivamente, se tiene que A_{11}^{-1} y B_{22}^{-1} existen. //

Distribución condicional de X_1 en $X_2 = x_2$.

Sea $\overline{X} = (X_1, X_2)$ donde X_1 es dimensión p_1 y X_2 es de dimensión p_2, tal que $X \sim N_{p_1+p_2}(x \mid \mu, \Sigma)$ con $\mu = (\mu_1, \mu_2)'$ y $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ la matriz de varianzas y covarianzas por bloques.

Sea $\Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} & \Sigma_{12}^{-1} \\ \Sigma_{21}^{-1} & \Sigma_{22}^{-1} \end{bmatrix}$ la matriz inversa de Σ.

Entonces, la distribución condicional de X_1 dado $X_2 = x_2$ está dada por:

$$f(x_1 | X_2 = x_2) = \frac{f(x_1, X_2 = x_2)}{f(X_2 = x_2)}$$

$$= \frac{(2\pi)^{-\frac{(p_1+p_2)}{2}} |\Sigma|^{-1/2} \exp\left\{ -\frac{1}{2} (x-\mu)' \Sigma^{-1} (x-\mu) \right\}}{(2\pi)^{-p_2/2} |\Sigma_{22}|^{-1/2} \exp\left\{ -\frac{1}{2} (x_2-\mu_2)' \Sigma_{22}^{-1} (x_2-\mu_2) \right\}}$$

$$\propto \exp\left\{ -\frac{1}{2} (x-\mu)' \Sigma^{-1} (x-\mu) + \frac{1}{2} (x_2-\mu_2)' \Sigma_{22}^{-1} (x_2-\mu_2) \right\}$$

$$= \exp\left\{ -\frac{1}{2} (x_1-\mu_1 \;\; x_2-\mu_2)' \begin{bmatrix} \Sigma_{11}^{-1} & \Sigma_{12}^{-1} \\ \Sigma_{21}^{-1} & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} + \frac{1}{2} (x_2-\mu_2)' \Sigma_{22}^{-1} (x_2-\mu_2) \right\}$$

$$= \exp\left\{ -\frac{1}{2} \left[(x_1-\mu_1)' \Sigma_{11}^{-1} (x_1-\mu_1) + 2(x_1-\mu_1)' \Sigma_{12}^{-1} (x_2-\mu_2) + \cancel{(x_2-\mu_2)' \Sigma_{22}^{-1} (x_2-\mu_2)} \right] + \cancel{\frac{1}{2} (x_2-\mu_2)' \Sigma_{22}^{-1} (x_2-\mu_2)} \right\}$$

$$= \exp\left\{ -\frac{1}{2} \left[(x_1-\mu_1)' \Sigma_{11}^{-1} (x_1-\mu_1) + 2(x_1-\mu_1)' \Sigma_{12}^{-1} (x_2-\mu_2) \right] \right\}$$

$$\propto \exp\left\{ -\frac{1}{2} \left[x_1' \Sigma_{11}^{-1} x_1 - 2x_1' \Sigma_{11}^{-1} \mu_1 + 2x_1' \Sigma_{12}^{-1} (x_2-\mu_2) \right] \right\}$$

$$= \exp\left\{ -\frac{1}{2} \left[x_1' \Sigma_{11}^{-1} x_1 - 2x_1' \Sigma_{11}^{-1} (\mu_1 - \Sigma_{11} \Sigma_{12}^{-1} (x_2-\mu_2)) \right] \right\}$$

Del teorema anterior, se sigue que $\Sigma_{11}^{-1} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ y que $\Sigma_{11}\Sigma_{12}^{-1} = -\Sigma_{12}\Sigma_{22}^{-1}$ entonces

$$f(x_1 | X_2 = x_2) \propto \exp\left\{ -\frac{1}{2} \left[x_1' (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) x_1 - 2x_1' (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) (\mu_1 + \Sigma_{12}\Sigma_{22}^{-1} (x_2-\mu_2)) \right] \right\}$$

$$\propto \exp\left\{ -\frac{1}{2} \left[\begin{array}{l} x_1' (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) x_1 - 2x_1' (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) (\mu_1 + \Sigma_{12}\Sigma_{22}^{-1} (x_2-\mu_2)) \\ + (\mu_1 + \Sigma_{12}\Sigma_{22}^{-1} (x_2-\mu_2))' (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) (\mu_1 + \Sigma_{12}\Sigma_{22}^{-1} (x_2-\mu_2)) \end{array} \right] \right\}$$

$$= \exp\left\{ -\frac{1}{2} \left[\left[x_1 - (\mu_1 + \Sigma_{12}\Sigma_{22}^{-1} (x_2-\mu_2)) \right]' (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) \left[x_1 - (\mu_1 + \Sigma_{12}\Sigma_{22}^{-1} (x_2-\mu_2)) \right] \right] \right\}$$

Que coincide con el kernel de una normal multivariada de orden p_1. Esto es

$X_1 | X_2 = x_2 \sim N_{p_1}(x_{1|2} | \mu_{1|2}, \Sigma_{1|2})$ donde $\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$

y $\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$