Part 1 : multivariable random vectors and the multivariate normal distribution

[L1]
10.01.2017

Random vector, $f(x)$, $F(x)$

Random vector = vector of random variables (RV)

$$\underset{p\times 1}{\mathbf{X}} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

EKS: height, weight, phys.act, diet

We will focus on continuous random variables.

1) $\mathbf{X}$ has a joint distribution (density) function (pdf)

$$f(x) = f(x_1, x_2, .., x_p)$$

Remember

$$f(x) \geq 0 \quad \forall x$$

$$\int_{-\infty}^{\infty}\int \cdots \int_{-\infty}^{\infty} f(x_1, x_2, .., x_p)\, dx_1 \cdots dx_p = 1$$

2) $\mathbf{X}$ has a joint cumulative distribution function (cdf)

$$F(x) = P(\mathbf{X} \leq x) = P(X_1 \leq x_1, X_2 \leq x_2, .. X_p \leq x_p)$$

$$\begin{tikzcd} f(x) \arrow[rr, bend left, "\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(t_1, t_2, .., t_p)\, dt_1 .. dt_p"] & & F(x) = F(x_1, x_2, .., x_p) \arrow[ll, bend left, "\frac{\partial^p F(x_1, .., x_p)}{\partial x_1 \partial x_2 .. \partial x_p}"] \end{tikzcd}$$

$f(x_1, .., x_p)$

Let $\mathbf{X} = \begin{bmatrix} \mathbf{X}_A \\ \mathbf{X}_B \end{bmatrix}$

$\mathbf{X}_A \in \mathbb{R}^k$
$\mathbf{X}_B \in \mathbb{R}^{p-k}$

eg. $\mathbf{X}_A = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \end{bmatrix}$ $\mathbf{X}_B = \begin{bmatrix} \mathbf{X}_4 \\ \vdots \\ \mathbf{X}_p \end{bmatrix}$

3) $\mathbf{X}_A$ has marginal distribution function

$$f_A(x_A) = \underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-k} f(x)\, dx_B$$

"integrating out the variables not of interest"

4) and conditional distribution of $\mathbf{X}_B$ given $\mathbf{X}_A$

$$f(x_B \mid x_A) = \frac{f(x_A, x_B)}{f_A(x_A)} \quad \text{for } f_A(x_A) > 0$$

Ex: $f(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}$
$-\infty < x_1, x_2 < \infty$

Hint: $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2}\, dx_1 = 1$

1) Find $f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}\, dx_2$$

$$= \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x_1^2}} \cdot \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2}} \, dx_2$$

$$= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x_1^2} \cdot \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} \, dx_2}_{1}$$

$$= \underline{\underline{\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x_1^2}}}$$

$$f_2(x_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2}$$

2) $$f(x_2 | x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{\cancel{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2}}{\cancel{\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x_1^2}}}$$

$$= \underline{\underline{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2}}} = f_2(x_2)$$

3) $$F(x) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(t_1, t_2) \, dt_1 \, dt_2 = \dots$$

$F(x_1, x_2)$

not on closed form, use R software

Independence

DEF: X_A and X_B are independent

$$\Updownarrow$$

$$f(x_A, x_B) = f_A(x_A) \cdot f_B(x_B)$$

Also: $f(x_A \mid x_B) = f_A(x_A)$, $f(x_B \mid x_A) = f_B(x_B)$

Ex: $f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x_1^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} = f_1(x_1) \cdot f_2(x_2)$

What about $F(x)$, $F_A(x_A)$, $F_B(x_B)$ when X_A and X_B are independent?

$$F(x) = \int_{-\infty}^{x_A} \int_{-\infty}^{x_B} \underset{\substack{\uparrow \\ f_A(t_A) \cdot f_B(t_B)}}{f(t)} \, dt_A \, dt_B$$

$$= \int_{-\infty}^{x_A} f_A(t_A)\, dt_A \cdot \int_{-\infty}^{x_B} f_B(t_B)\, dt_B = F_A(x_A) \cdot F_B(x_B)$$

From $f(x_1, x_2) \rightarrow f_2(x_2),\ f_1(x_1)$ by integrating

How can we get from $f_1(x_1)$ and $f_2(x_2)$ to $f(x_1, x_2)$?

$f_1(x_1)$ and $f_2(x_2) \Rightarrow f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$

x_1 and x_2 independent

But what if x_1 and x_2 are <u>not</u> independent?

$\downarrow$

solution is Copula $\Rightarrow$ see slides!