Mathematical Statistics 1

Ch.4 Bivariate Distributions

Jungsoon Choi

jungsoonchoi@hanyang.ac.kr

Table of Contents

5.1 Bivariate Normal distribution

Definition

$$\underset{\sim}{X} = \begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{BVN}\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix} \right), \quad |\rho| < 1$$

Then, the joint pdf of X and Y is

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-Q/2}$$

where $\mathrm{Corr}(X, Y) = \rho$ and

$$Q = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 \right].$$

· Bivariate Normal

$$\underset{\sim}{X} \sim BN\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} = \underset{\sim}{\mu},\ \underbrace{\begin{pmatrix} \sigma_X^2 & \overset{\overset{\rho_{XY} \cdot \sigma_X \sigma_Y}{\parallel}}{\mathrm{cov}(X,Y)} \\ \mathrm{cov}(X,Y) & \sigma_Y^2 \end{pmatrix}}_{\text{covariance matrix}} = \Sigma \right)$$

· Multivariate Normal

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \sim MN\left(\underset{\sim}{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix},\ \underset{n \times n}{\Sigma} = \begin{pmatrix} \sigma_1^2 & \cdots & \mathrm{cov}(X_1, X_n) \\ \vdots & \sigma_2^2 \ \ddots & \vdots \\ \mathrm{cov}(X_n, X_1) & \cdots & \sigma_n^2 \end{pmatrix} \right)$$

$X \sim N(\mu, \sigma^2)$

cov. mat.

$$\underset{\sim}{X} = \begin{pmatrix} X \\ Y \end{pmatrix} \sim BVN\left(\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right)$$

Joint pdf of X & Y

determinant

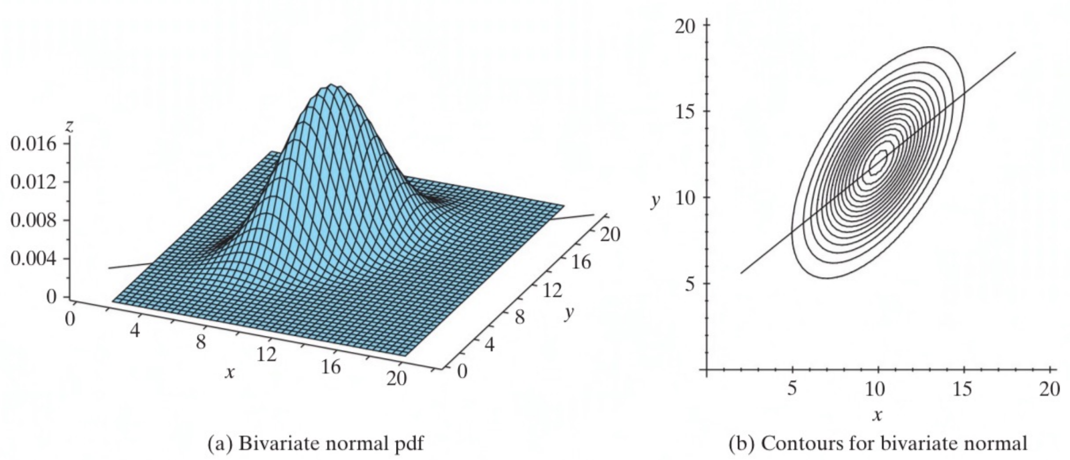
$$f_{XY}(x,y) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp\left\{-\frac{1}{2}\underbrace{(\underset{\sim}{x}-\underset{\sim}{\mu})^T}_{1\times 2}\underbrace{\Sigma^{-1}}_{2\times 2}\underbrace{(\underset{\sim}{x}-\underset{\sim}{\mu})}_{2\times 1}\right\}$$

$1\times n$ $\quad n\times n$ $\quad n\times 1$ $\Rightarrow 1\times 1$.

$$= \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})^T\Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\right\}$$

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(Z_X^2 - 2\rho Z_X Z_Y + Z_Y^2)\right\}$$

where $Z_X = \frac{x-\mu_X}{\sigma_X}$, $Z_Y = \frac{Y-\mu_Y}{\sigma_Y}$



(a) Bivariate normal pdf (b) Contours for bivariate normal

Figure 4.5-2 Bivariate normal, $\mu_X = 10$, $\sigma_X^2 = 9$, $\mu_Y = 12$, $\sigma_Y^2 = 16$, $\rho = 0.6$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad -\infty < x < \infty$$

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(Z_X^2 - 2\rho Z_X Z_Y + Z_Y^2)\right\}$$

where $Z_X = \frac{X-\mu_X}{\sigma_X}$ $\quad Z_Y = \frac{Y-\mu_Y}{\sigma_Y}$

1) $f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(Z_X^2 - 2\rho Z_X Z_Y + Z_Y^2)\right\}$

2) Joint mgf of X&Y

$$M_{XY}(t_1,t_2) = E[e^{t_1x+t_2y}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{(t_1x+t_2y)} f_{XY}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{(1-\rho)^2}} e^{t_1x+t_2y} \exp\left\{-\frac{(Z_X^2-2\rho Z_X Z_Y + Z_Y^2)}{2(1-\rho^2)}\right\} dx\,dy$$

① consider the simple situation $\mu_X = \mu_Y = 0$, $\sigma_X^2 = \sigma_Y^2 = 1$

$$M_{XY}(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{(t_1x+t_2y)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{(x^2-2\rho xy+y^2)}{2(1-\rho^2)}\right\} dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2-2\rho xy+y^2-2(1-\rho^2)t_1x}{2(1-\rho^2)} + t_2y\right\} dx\,dy$$

$$-(x^2 \quad -2x(\rho y + (1-\rho^2)t_1))$$

$$-\frac{(x-(\rho y+(1-\rho^2)t_1))^2}{2(1-\rho^2)}$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{y^2}{2(1-\rho^2)} + t_2y\right\} \times \exp\left\{\frac{\{\rho y+(1-\rho^2)t_1\}^2}{2(1-\rho^2)}\right\}$$

$$\times \sqrt{2\pi(1-\rho^2)} \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{\{x-(\rho y+(1-\rho^2)t_1)\}^2}{2(1-\rho^2)}\right\}}_{\text{pdf of } N(\rho y+(1-\rho^2)t_1,\,(1-\rho^2))} dx\,dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2(1-\rho^2)} + t_2y + \frac{\rho^2y^2}{2(1-\rho^2)} + \rho t_1 y + \frac{(1-\rho^2)t_1^2}{2}\right\} dy$$

✓ $$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2-2y(t_2+\rho t_1)+(t_2+\rho t_1)^2-(t_2+\rho t_1)^2}{2} + \frac{(1-\rho^2)t_1^2}{2}\right\} dy$$

$$= \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\{y-(t_2+\rho t_1)\}^2}{2}\right\}}_{\text{pdf of } N(t_2+\rho t_1,\,1)} dy \times \exp\left\{\frac{t_2^2+2\rho t_1t_2+\rho^2t_1^2+t_1^2-\rho^2t_1^2}{2}\right\}$$

$$M_{XY}(t_1,t_2) = \exp\left\{\frac{t_1^2+2\rho t_1t_2+t_2^2}{2}\right\}, \quad t_1,t_2 \in \mathbb{R}$$

② $X = \mu_X + \sigma_X Z_X$, $Y = \mu_Y + \sigma_Y Z_Y$

$$M_{XY}(t_1,t_2) = E[e^{t_1X+t_2Y}] = E[e^{t_1(\mu_X+\sigma_X Z_X)+t_2(\mu_Y+\sigma_Y Z_Y)}]$$

$$= E[e^{t_1\mu_X+t_2\mu_Y} \times e^{t_1\sigma_X Z_X + t_2\sigma_Y Z_Y}]$$

$$= \exp\{t_1\mu_X + t_2\mu_Y\} M_{Z_X Z_Y}(t_1\sigma_X, t_2\sigma_Y)$$

$$= \exp\{t_1\mu_X + t_2\mu_Y\} \exp\left\{\frac{t_1^2\sigma_X^2 + 2\rho\sigma_X\sigma_Y t_1t_2 + \sigma_Y^2 t_2^2}{2}\right\}$$

$$= \exp\left\{(\mu_X t_1 + \mu_Y t_2) + \frac{1}{2}(\sigma_X^2 t_1^2 + 2\rho\sigma_X\sigma_Y t_1t_2 + \sigma_Y^2 t_2^2)\right\},$$

$t_1, t_2 \in \mathbb{R}$

Marginal mgf of X, Y

$$M_X(t_1) = E[e^{t_1X}] = M_{X,Y}(t_1, 0)$$

$$= \exp\left\{\mu_X t_1 + \frac{\sigma_X^2}{2}t_1^2\right\}$$ mgf of $N(\mu_X, \sigma_X^2)$

$\therefore X \sim N(\mu_X, \sigma_X^2)$

$$M_Y(t_2) = E[e^{t_2Y}] = M_{XY}(0, t_2)$$

$$= \exp\left\{\mu_Y t_2 + \frac{\sigma_Y^2}{2}t_2^2\right\}$$ mgf of $N(\mu_Y, \sigma_Y^2)$

$\therefore Y \sim N(\mu_Y, \sigma_Y^2)$

The joint mgf of X and Y is

$X_1 \cdots X_n$

$$
\begin{aligned}
M(t_1, t_2) &= E[e^{t_1 X + t_2 Y}] \\
&= \exp\left[\mu_X t_1 + \mu_Y t_2 + \frac{1}{2}(\sigma_X^2 t_1^2 + 2\rho\sigma_X\sigma_Y t_1 t_2 + \sigma_Y^2 t_2^2)\right]
\end{aligned}
$$

$M(X_1, \cdots, X_n) = E[e^{\sum_{i=1}^{n} t_i X_i}]$

The mgf of X is

$$M_X(t_1) = M(t_1, 0) = \exp[\mu_X t_1 + \frac{1}{2}\sigma_X^2 t_1^2]$$

Similarly, the mgf of Y is

$$M_Y(t_2) = M(0, t_2) = \exp[\mu_Y t_2 + \frac{1}{2}\sigma_Y^2 t_2^2].$$

Thus, $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$.

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(z_X^2-2\rho z_X z_Y+z_Y^2\right)\right\}}{\frac{1}{\sqrt{2\pi\sigma_X^2}}\exp\left\{-\frac{1}{2}z_X^2\right\}}$$

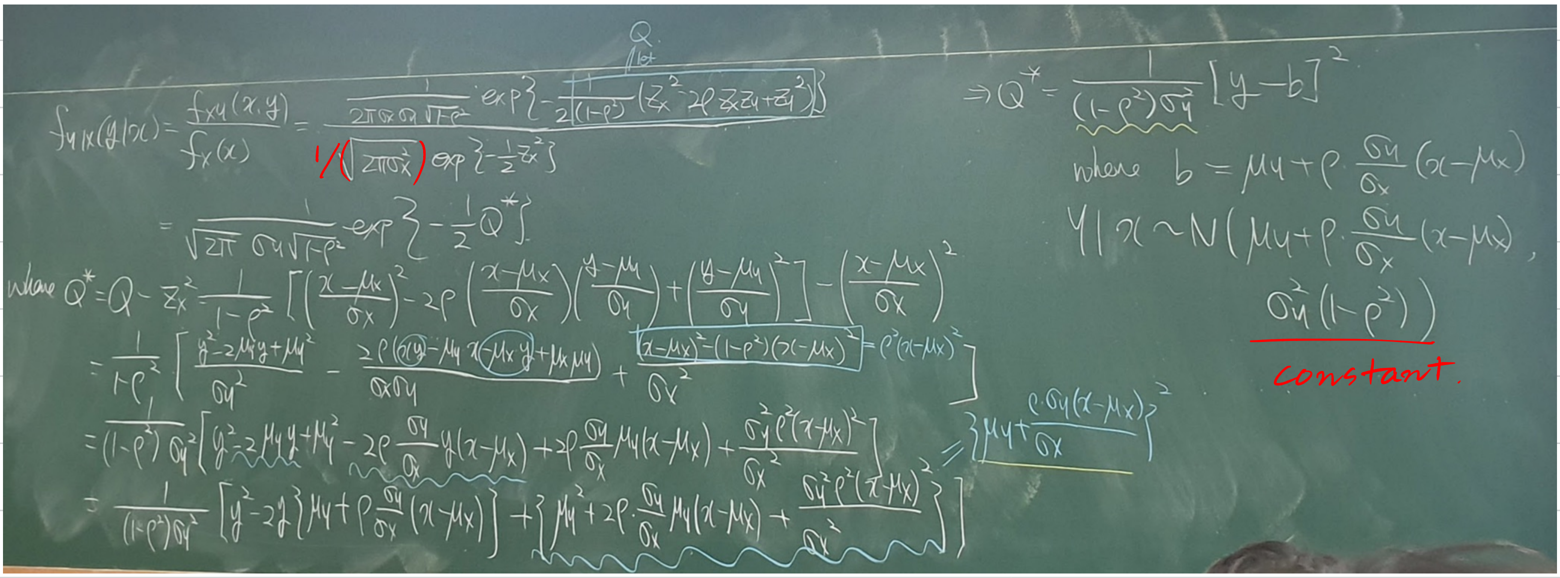
$$= \frac{1}{\sqrt{2\pi}\,\sigma_Y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2}Q^*\right\}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_Y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2}\frac{(y-b)^2}{(1-\rho^2)\sigma_Y^2}\right\}$$

where $Q^* = Q - z_X^2$

$$= \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right] - \left(\frac{x-\mu_X}{\sigma_X}\right)^2$$

$$= \frac{1}{1-\rho^2}\left[\frac{y^2-2\mu_Y y+\mu_Y^2}{\sigma_Y^2} - \frac{2\rho(xy-\mu_Y x-\mu_X y+\mu_X\mu_Y)}{\sigma_X\sigma_Y}\right.$$



$E[Y|x] = a + bx$

$Var[Y|x] = K(x)$ $\quad E[Var(Y|x)] = \sigma_Y^2(1-\rho^2)$

If $\rho = 0$, $\quad Y|x \sim N(\mu_Y, \sigma_Y^2)$

$f_{Y|X}(y|x) = f_Y(y)$.

$X \& Y$: indep $\Longrightarrow Cov(X,Y) = 0$

$(\Leftrightarrow f_{XY}(x,y) = f_X(x) \cdot f_Y(y))$ $\xleftarrow{\times}$ $Corr(X,Y) = \rho = 0$

$\xleftarrow{O}$ if $(X,Y) \sim BVN$

But In BN, if $\rho = 0$, $f_{XY}(x,y) = f_Y(y) = \dfrac{f_{XY}(x,y)}{f_X(x)}$

$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$

$X \& Y$: indep.

Conditional distribution

The conditional distribution of Y given $X = x$ is

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{(y-b)^2}{2\sigma_Y^2(1-\rho^2)}}$$

$$E[Y|x] = b = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

Thus,

$$Y|x \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X),\ \sigma_Y^2(1-\rho^2)\right),$$

$$X|y \sim N\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y),\ \sigma_X^2(1-\rho^2)\right).$$

Thm 4.5-1

Let (X, Y) have bivariate normal distribution with the correlation coefficient ρ.

$$X \text{ and } Y \text{ are independent.} \Leftrightarrow \rho = 0.$$