Bivariate and Multivariate Normal Random Variables: Lecture XIII

Charles B. Moss

September 21, 2010

I. Bivariate Normal Random Variables

A. **Definition 5.3.1.** The bivariate normal density is defined by

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$$
$$\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2+\left(\frac{y-\mu_y}{\sigma_y}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]\right\} \tag{1}$$

B. **Theorem 5.3.1.** Let (X,Y) have the bivariate normal density. Then the marginal densities $f_X(X)$ and $f_Y(Y)$ and the conditional densities $f(Y|X)$ and $f(X|Y)$ are univariate normal densities, and we have $\mathrm{E}[X]=\mu_X$, $\mathrm{E}[Y]=\mu_Y$, $V[X]=\sigma_X^2$, $\mathrm{E}[Y]=\mu_Y$, $V[Y]=\mu_Y$, $\mathrm{Corr}(X,Y)=\rho$, and

$$\begin{aligned}\mathrm{E}[Y|X] &= \mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(X-\mu_X)\\ V[Y|X] &= \sigma_Y^2\left(1-\rho^2\right)\end{aligned} \tag{2}$$

Note that

$$f(x,y) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2\sigma_Y^2\left(1-\rho^2\right)}\left[y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right]^2\right\}$$
$$\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left[-\frac{1}{\sqrt{2\pi}\sigma_X}(x\mu_X)^2\right]$$
$$= f_2 f_1 \tag{3}$$

where f_1 is the density of $N\left(\mu_X,\sigma_X^2\right)$ and f_2 is the density function of $N\left(\mu_X+\rho\sigma_Y\sigma_X^{-1}(X-\mu_X),\sigma_Y^2\left(1-\rho^2\right)\right)$. The proof of the assertions from the theorem can then be seen by

$$f(x) = \int_{-\infty}^{\infty} f_1 f_2 dy$$
$$f_1 \textstyle\int_{-\infty}^{\infty} f_2 dy \tag{4}$$
$$= f_1$$

1. This gives us $x \sim N\left(\mu_X,\sigma_X^2\right)$. Next, we have

$$f(y|x) = \frac{f(x,y)}{f(x)} = \frac{f_2 f_1}{f_1} = f_2 \tag{5}$$

2 By Theorem 4.4.1 (Law of Iterated Means) $\mathrm{E}[\phi(X,Y)] = \mathrm{E}_X\mathrm{E}_{Y|X}[\phi(X,Y)]$ where E_X denotes the expectation with respect to X

$$\mathrm{E}[XY] = \mathrm{E}_X\mathrm{E}[XY|X] = \mathrm{E}_X[X\mathrm{E}[Y|X]]$$
$$= \mathrm{E}_X\left[X\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}X(X-\mu_X)\right] \tag{6}$$
$$= \mu_X\mu_Y+\rho\sigma_X\sigma_Y$$

C. **Theorem 5.3.2.** If X and Y are bivariate normal and α and β are constants, then $\alpha X+\beta Y$ is normal.

D. **Theorem 5.3.3.** Let $\{X_i\}$, $i = 1, 2, \ldots n$ be pairwise independent and identically distributed as $N(\mu, \sigma^2)$. Then $\bar{x} = 1/n \sum_{i=1}^{n}$ is $N(\mu, \sigma^2/n)$.

E. **Theorem 5.3.4.** If X and Y are bivariate normal and $\text{Cov}(X, Y) = 0$, then X and Y are independent.

II. Multivariate Normal Distribution

A. **Definition 5.4.1** We way X is multivariate normal with mean μ and variance-covariance matrix Σ, denoted $N(\mu, \Sigma)$, if its density is given by

$$f(x) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2}(x-\mu)' \Sigma^{-1} (x\mu)\right] \tag{7}$$

1. Note first that $|\Sigma|$ denotes the determinant of the variance matrix. The determinant is found by expansion down a row or column of a matrix

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} +$$

$$(-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{33} \end{vmatrix} \tag{8}$$

$$= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) +$$

$$a_{31}(a_{12}a_{33} - a_{22}a_{13})$$

2. Inverse of a matrix by row-reduction

$$\left[\begin{array}{cc} 1 & 0 \\ -\frac{\sigma_{21}}{\sigma_{11}} & 1 \end{array}\right]\left[\begin{array}{cc|cc} \sigma_{11} & \sigma_{12} & 1 & 0 \\ \sigma_{21} & \sigma_{22} & 0 & 1 \end{array}\right] =$$

$$\left[\begin{array}{cc|cc} \sigma_{11} & \sigma_{12} & 1 & 0 \\ 0 & \sigma_{22} - \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}} & -\frac{\sigma_{21}}{\sigma_{11}} & 1 \end{array}\right]$$

$$\left[\begin{array}{cc} 1 & 0 \\ 0 & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array}\right]\left[\begin{array}{cc|cc} \sigma_{11} & \sigma_{12} & 1 & 0 \\ 0 & \frac{\sigma_{12}\sigma_{22} - \sigma_{12}\sigma_{21}}{\sigma_{11}} & -\frac{\sigma_{21}}{\sigma_{11}} & 1 \end{array}\right] =$$

$$\left[\begin{array}{cc|cc} \sigma_{11} & \sigma_{12} & 1 & 0 \\ 0 & 1 & -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array}\right]$$

$$\left[\begin{array}{cc} 1 & -\frac{\sigma_{12}}{\sigma_{11}} \\ 0 & 1 \end{array}\right]\left[\begin{array}{cc|cc} 1 & \frac{\sigma_{12}}{\sigma_{11}} & \frac{1}{\sigma_{11}} & 0 \\ 0 & 1 & -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array}\right] =$$

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{\sigma_{11}} + \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}\left(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}\right)} & -\frac{\sigma_{12}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \\ 0 & 1 & -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array}\right] \tag{9}$$

Focusing on the first term in the inverse

$$\begin{aligned} \frac{1}{\sigma_{11}} + \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}\left(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}\right)} &= \frac{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21} + \sigma_{12}\sigma_{21}}{\sigma_{11}\left(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}\right)} \\ &= \frac{\sigma_{11}\sigma_{22}}{\sigma_{11}\left(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}\right)} \end{aligned} \tag{10}$$

3. Matrix Inverse of Matrices

$$\left[\begin{array}{cc|cc} \Sigma_{XX} & \Sigma_{XY} & I & 0 \\ \Sigma_{YX} & \Sigma_{YY} & 0 & I \end{array}\right] \quad \Sigma_{XX}^{-1} R_1$$

$$\left[\begin{array}{cc|cc} I & \Sigma_{XX}^{-1}\Sigma_{XY} & \Sigma_{XX}^{-1} & 0 \\ \Sigma_{YX} & \Sigma_{YY} & 0 & I \end{array}\right] \quad R_2 - \Sigma_{YX} R_1 \tag{11}$$

$$\left[\begin{array}{cc|cc} I & \Sigma_{XX}^{-1}\Sigma_{XY} & \Sigma_{XX}^{-1} & 0 \\ 0 & \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY} & -\Sigma_{YX}\Sigma_{XX}^{-1} & I \end{array}\right]$$

ultiplying the last row of Equation 11 by $\left(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right)^{-1}$ and then subtracting $\Sigma_{XX}^{-1}\Sigma_{XY}$ times the second row from the first row yields

$$\left[\begin{array}{cc} \Sigma_{XX}^{-1} + \Sigma_{XX}^{-1}\Sigma_{XY}\left(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right)^{-1}\Sigma_{YX}\Sigma_{XX}^{-1} & -\Sigma_{XX}^{-1}\Sigma_{XY}\left(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right)^{-1} \\ -\left(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma XY\right)^{-1}\Sigma_{YX}\Sigma_{XX}^{-1} & \left(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right)^{-1} \end{array}\right] \tag{12}$$

B. Conditional Formulations

1. Matrix Form of Conditional Variance: Let X be a vector of normally distributed random variables

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right] \tag{13}$$

2. Let Y composed of a vector of $\begin{pmatrix} Y_1' & Y_2' \end{pmatrix}$ be defined based on X by

$$\begin{aligned} Y_1 &= X_1 + BX_2 \\ Y_2 &= X_2 \end{aligned} \tag{14}$$

3. The matrix B is defined such that Y_1 is uncorrelated with Y_2. Mathematically

$$\begin{aligned}
0 &= \mathrm{E}\left[\left(Y_1-\mu_1\right)\left(Y_2-\mu_2\right)\right] \\
&= \mathrm{E}\left[\left(X_1+Bx_2-\mu_1-B\mu_2\right)\left(X_2-\mu_2\right)\right] \\
&= \mathrm{E}\left[\left\{\ \left(X_1-\mu_1\right)+B\left(X_2-\mu_2\right)\right\}\ \left(X_2-\mu_2\right)\right] \\
&= \mathrm{E}\left[\left(X_1-\mu_1\right)\left(X_2-\mu_2\right)+B\left(X_2-\mu_2\right)\left(X_2-\mu_2\right)\right] \\
\Sigma_{12}+B\Sigma_{22} = 0 \Rightarrow B &= -\Sigma_{12}\Sigma_{22}^{-1} \text{ and } Y_1 = X_1-\Sigma_{12}\Sigma_{22}^{-1}X_2
\end{aligned} \tag{15}$$

In matrix form

$$\begin{aligned}
\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} &= \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\
\mathrm{E}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} &= \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \mu_1-\Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix} \\
V\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} &= \begin{pmatrix} \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}
\end{aligned} \tag{16}$$