

Bivariate and Multivariate Normal Random Variables: Lecture XV

Charles B. Moss

September 16, 2026

- 1 Bivariate Normal Random Variables
- 2 Multivariate Normal Distribution
- 3 Conditional Formulations

Bivariate Normal Random Variables

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

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

- Let us consider a slight change in the exponential terms – specifically let $z_x = (x - \mu_x) / \sigma_x$ and $z_y = (y - \mu_y) / \sigma_y$. With this change, the exponential term becomes

$$-\frac{1}{2(1-\rho^2)} [z_1^2 + z_2^2 - 2\rho z_1 z_2] \quad (2)$$

- Since $E[z_1] = E[z_2] = 0$ this implies that $\text{Cov}[x, y] = \rho$.
- Explaining the $(1 - \rho)^2$ term is a little harder.

- As a first point, let us consider the inverse of the variance matrix. Starting with uncorrelated normals

$$\begin{aligned} \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix}' \begin{bmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{bmatrix} \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix} \\ = \frac{(x - \mu_x)^2}{\sigma_x^2} + \frac{(y - \mu_y)^2}{\sigma_y^2} \end{aligned} \quad (3)$$

- The Σ matrix is referred to as the covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}, \Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{bmatrix} \quad (4)$$

- Next, consider that the variance matrix is typically more complex – it includes covariance terms

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \quad (5)$$

- The inverse of this matrix with covariance is somewhat more complicated.

Inverse of Matrix by Row-Reduction

$$\begin{aligned}
 & \left[\begin{array}{cc|cc} 1 & 0 & \sigma_{11} & \sigma_{12} & 1 & 0 \\ -\frac{\sigma_{21}}{\sigma_{11}} & 1 & \sigma_{21} & \sigma_{22} & 0 & 1 \end{array} \right] = \\
 & \left[\begin{array}{cc|cc} \sigma_{11} & \sigma_{12} & 1 & 0 \\ 0 & \sigma_{22} - \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}} & -\frac{\sigma_{21}}{\sigma_{11}} & 1 \end{array} \right] \\
 & \left[\begin{array}{cc|cc} 1 & 0 & \sigma_{11} & \sigma_{12} & 1 & 0 \\ 0 & \frac{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}}{\sigma_{11}} & 0 & \frac{\sigma_{12}\sigma_{22} - \sigma_{12}\sigma_{21}}{\sigma_{11}} & -\frac{\sigma_{21}}{\sigma_{11}} & 1 \end{array} \right] = \\
 & \left[\begin{array}{cc|cc} \sigma_{11} & \sigma_{12} & 1 & 0 \\ 0 & 1 & -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array} \right] \quad (6) \\
 & \left[\begin{array}{cc|cc} 1 & -\frac{\sigma_{12}}{\sigma_{11}} & 1 & \frac{\sigma_{12}}{\sigma_{11}} & \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & 0 \\ 0 & 1 & 0 & 1 & -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array} \right] = \\
 & \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{\sigma_{11}} + \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})} & -\frac{\sigma_{12}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \\ 0 & 1 & -\frac{\sigma_{21}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}} \end{array} \right]
 \end{aligned}$$

- Focusing on the first term in the inverse

$$\begin{aligned} \frac{1}{\sigma_{11}} + \frac{\sigma_{12}\sigma_{21}}{\sigma_{11}(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})} &= \frac{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21} + \sigma_{12}\sigma_{21}}{\sigma_{11}(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})} \\ &= \frac{\sigma_{11}\sigma_{22}}{\sigma_{11}(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})} \end{aligned} \quad (7)$$

$$\Sigma^{-1} = \begin{bmatrix} \frac{\sigma_{22}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{12}} & -\frac{\sigma_{12}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{12}} \\ -\frac{\sigma_{12}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{12}} & \frac{\sigma_{11}}{\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{12}} \end{bmatrix} \quad (8)$$

- Consider the denominator (i.e., the determinant of the matrix)

$$\begin{aligned}\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{12} &\Rightarrow \sigma_{11}\sigma_{22} \left(1 - \frac{\sigma_{12}\sigma_{12}}{\sigma_{11}\sigma_{22}}\right) \\ &= \sigma_{11}\sigma_{22} \left(1 - \left[\frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}}\right] \left[\frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}}\right]\right) \quad (9) \\ &= \sigma_{11}\sigma_{22} (1 - \rho^2)\end{aligned}$$

- 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 $E[X] = \mu_X$, $E[Y] = \mu_Y$, $V[X] = \sigma_X^2$, $V[Y] = \sigma_Y^2$, $\text{Corr}(X, Y) = \rho$, and

$$E[Y|X] = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X} (X - \mu_X) \tag{10}$$

$$V[Y|X] = \sigma_Y^2 (1 - \rho^2)$$

- As a starting point consider a part of Equation 1

$$\begin{aligned}\frac{1}{2(1-\rho^2)} \left(\frac{x-\mu_x}{\sigma_x} \right)^2 &= \frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x} \right)^2 + K \left(\frac{x-\mu_x}{\sigma_x} \right)^2 \\ \Rightarrow \frac{1}{2(1-\rho^2)} &= \frac{1}{2} + K \\ K &= \frac{1}{2(1-\rho^2)} - \frac{1-\rho^2}{2(1-\rho^2)} = \frac{\rho^2}{2(1-\rho^2)}\end{aligned}\tag{11}$$

- Therefore

$$\frac{1}{2(1-\rho^2)} \left(\frac{x-\mu_x}{\sigma_x} \right)^2 = \frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x} \right)^2 + \frac{\rho^2}{2(1-\rho^2)} \left(\frac{x-\mu_x}{\sigma_x} \right)^2\tag{12}$$

- The exponential part of Equation 1 can be rewritten as

$$\frac{1}{2(1-\rho^2)} \left[\rho^2 \left(\frac{x-\mu_x}{\sigma_x} \right)^2 + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) \right] - \frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x} \right)^2 \quad (13)$$

- Focusing on the term in the brackets

$$\begin{aligned} \rho^2 \left(\frac{x-\mu_x}{\sigma_x} \right)^2 + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) \\ = \left[\left(\frac{y-\mu_y}{\sigma_y} \right) - \rho \left(\frac{x-\mu_x}{\sigma_x} \right) \right]^2 \end{aligned} \quad (14)$$

- Further, Equation 14 can be further changed to

$$\frac{1}{\sigma_y^2} \left[y - \mu_y - \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x) \right]^2 \quad (15)$$

- Note that

$$\begin{aligned} f(x, y) &= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2\sigma_Y^2(1-\rho^2)} \left[y - \mu_Y - \rho \frac{\sigma_Y}{\sigma_X} (x - \mu_X) \right]^2 \right\} \\ &\quad \frac{1}{\sqrt{2\pi}\sigma_X} \exp \left[-\frac{1}{2\sigma_X^2} (x - \mu_X)^2 \right] \\ &= f_{21} f_1 \end{aligned} \quad (16)$$

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

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} f_1 f_{2|1} dy \\ &= f_1 \int_{-\infty}^{\infty} f_{2|1} dy \\ &= f_1 \end{aligned} \tag{17}$$

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

$$f(y|x) = \frac{f(x, y)}{f(x)} = \frac{f_{2|1} f_1}{f_1} = f_{2|1} \tag{18}$$

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

$$\begin{aligned}
 E[XY] &= E_X E[XY|X] = E_X [XE[Y|X]] \\
 &= E_X \left[X\mu_Y + \rho \frac{\sigma_Y}{\sigma_X} X(X - \mu_X) \right] \quad (19) \\
 &= \mu_X \mu_Y + \rho \sigma_X \sigma_Y
 \end{aligned}$$

- **Theorem 5.3.2.** If X and Y are bivariate normal and α and β are constants, then $\alpha X + \beta Y$ is normal.
- **Theorem 5.3.3.** Let $\{X_i\}$, $i = 1, 2, \dots, n$ be pairwise independent and identically distributed as $N(\mu, \sigma^2)$. Then $\bar{x} = 1/n \sum_{i=1}^n x_i$ is $N(\mu, \sigma^2/n)$.
- **Theorem 5.3.4.** If X and Y are bivariate normal and $\text{Cov}(X, Y) = 0$, then X and Y are independent.

Multivariate Normal Distribution

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

$$f(x) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left[-\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu) \right] \quad (20)$$

- 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

$$\begin{aligned}
 & \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \\
 & (-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \quad (21) \\
 & = a_{11} (a_{22}a_{33} - a_{32}a_{23}) - a_{21} (a_{12}a_{33} - a_{32}a_{13}) + \\
 & \quad a_{31} (a_{12}a_{23} - a_{22}a_{13})
 \end{aligned}$$

Matrix Inverse of Matrices

$$\begin{aligned}
 & \left[\begin{array}{cc|cc} \Sigma_{XX} & \Sigma_{XY} & I & 0 \\ \Sigma_{YX} & \Sigma_{YY} & 0 & I \end{array} \right] \Sigma_{XX}^{-1} R_1 \\
 & \left[\begin{array}{cc|cc} I & \Sigma_{XX}^{-1} \Sigma_{XY} & \Sigma_{XX}^{-1} & 0 \\ \Sigma_{YX} & \Sigma_{YY} & 0 & I \end{array} \right] R_2 - \Sigma_{YX} R_1 \\
 & \left[\begin{array}{cc|cc} I & \Sigma_{XX}^{-1} \Sigma_{XY} & \Sigma_{XX}^{-1} & 0 \\ 0 & \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} & -\Sigma_{YX} \Sigma_{XX}^{-1} & I \end{array} \right]
 \end{aligned} \tag{22}$$

- Multiplying the last row of Equation 22 by $(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY})^{-1}$ and then subtracting $\Sigma_{XX}^{-1}\Sigma_{XY}$ times the second row from the first row yields

$$\begin{bmatrix} \Sigma_{XX}^{-1} + \Sigma_{XX}^{-1}\Sigma_{XY}(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY})^{-1}\Sigma_{YX}\Sigma_{XX}^{-1} \\ -(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY})^{-1}\Sigma_{YX}\Sigma_{XX}^{-1} \\ -\Sigma_{XX}^{-1}\Sigma_{XY}(\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY})^{-1} \\ (\Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY})^{-1} \end{bmatrix} \quad (23)$$

Conditional Formulations

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

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right] \quad (24)$$

- Let Y composed of a vector of $(Y'_1 \ Y'_2)$ be defined based on X by

$$\begin{aligned} Y_1 &= X_1 + BX_2 \\ Y_2 &= X_2 \end{aligned} \quad (25)$$

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

$$\begin{aligned}
 0 &= E[(Y_1 - \mu_{Y_1})(Y_2 - \mu_{Y_2})] \\
 &= E[(X_1 + BX_2 - \mu_1 - B\mu_2)(X_2 - \mu_2)] \\
 &= E[\{(X_1 - \mu_1) + B(X_2 - \mu_2)\}(X_2 - \mu_2)] \\
 &= E[(X_1 - \mu_1)(X_2 - \mu_2) + B(X_2 - \mu_2)(X_2 - \mu_2)] \\
 \Sigma_{12} + B\Sigma_{22} &= 0 \Rightarrow B = -\Sigma_{12}\Sigma_{22}^{-1} \text{ and } Y_1 = X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2
 \end{aligned}
 \tag{26}$$

- In matrix form

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

$$E \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix}$$

$$V \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \quad (27)$$