Stat 145 (Multivariate Statistics)

Lesson 1.2 The Multivariate Normal Distribution

2025-08-21

Learning Outcomes

1. Describe the multivariate normal distribution.
2. Prove theorems related to the multivariate normal distribution.

1 The Multivariate Normal Distribution

A generalization of the familiar bell-shaped normal density to several dimensions plays a fundamental role in multivariate analysis. In fact, most of the techniques encountered in this course are based on the assumption that the data were generated from a multivariate normal distribution. While real data are never exactly multivariate normal, the normal density is often a useful approximation to the "true" population distribution. One advantage of the multivariate normal distribution stems from the fact that it is mathematically tractable and "nice" results can be obtained. This is frequently not the case for other data-generating distributions. Of course, mathematical attractiveness per se is of little use to the practitioner. It turns out, however, that normal distributions are useful in practice for two reasons: First, the normal distribution serves as a bona fide population model in some instances; second, the sampling distributions of many multivariate statistics are approximately normal, regardless of the form of the parent population, because of a central limit effect.

To summarize, many real-world problems fall naturally within the framework of normal theory. The importance of the normal distribution rests on its dual role as both population model for certain natural phenomena and approximate sampling distribution for many statistics.

Before we present the multivariate normal density, let us recall that if $y \sim N(\mu, \sigma^2)$ then

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \ -\infty < y < \infty$$

Some examples of univariate normal densities are shown in Figure 1.

The term

$$\frac{(y-\mu)^2}{2\sigma^2} = (y-\mu)(\sigma^2)^{-1}(y-\mu)$$

in the exponent of the univariate normal density function measures the squared distance from x to μ in standard deviation units. This can be generalized for a vector $\mathbf{y}$ of observations on several variables as

$$(\mathbf{y}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{y}-\mu)$$

The $p \times 1$ vector $\boldsymbol{\mu}$ represents the expected value of the random vector $\mathbf{y}$, and $p \times p$ matrix $\mathbf{\Sigma}$ is the variance–covariance matrix of $\mathbf{y}$. We shall assume that the symmetric matrix $\mathbf{\Sigma}$ is positive definite, so the expression above is the square of the generalized distance from $\mathbf{y}$ to $\boldsymbol{\mu}$.

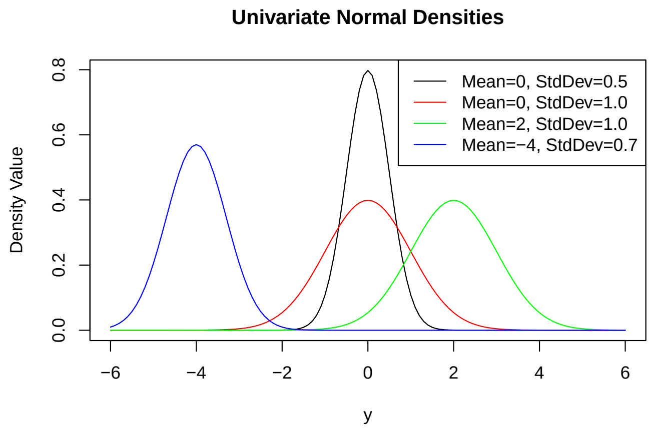


Figure 1: Univariate Normal Densities

Therefore, the multivariate normal density for a random vector $\mathbf{y}$ is given by

$$f(\mathbf{y}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{-1}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}$$

where $-\infty < y_i < \infty$, $i = 1, 2, \cdots, p$. We shall denote this p-dimensional normal density by $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ which is analogous to the normal density in the univariate case.

Example 1.2.1

When $p = 2$ we have the bivariate normal distribution and its density is

$$\begin{aligned} f(y_1, y_2) &= \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{-1}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})} \\ &= \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} exp\left\{ -\frac{1}{2}\left(\frac{1}{1-\rho_{12}^2}\right)\left[\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y_2-\mu_2}{\sigma_2}\right)^2 - 2\rho_{12}\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2\left(\frac{y_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \end{aligned}$$

where

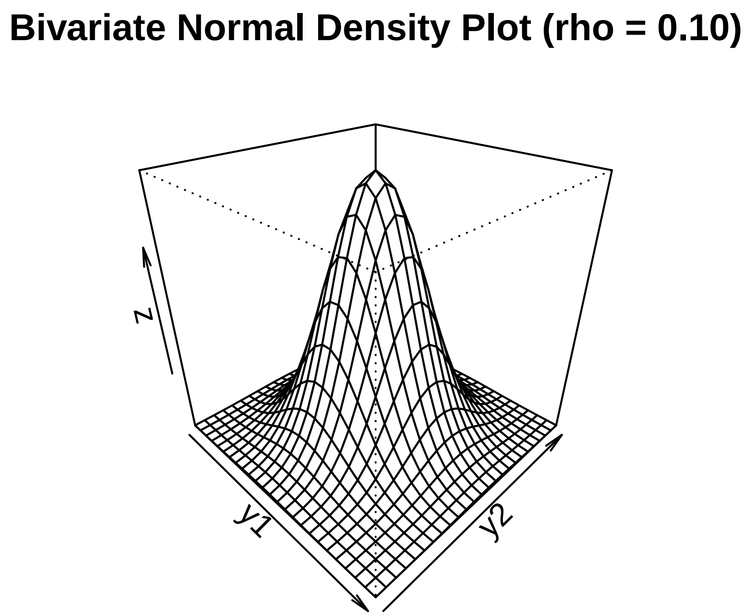
$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

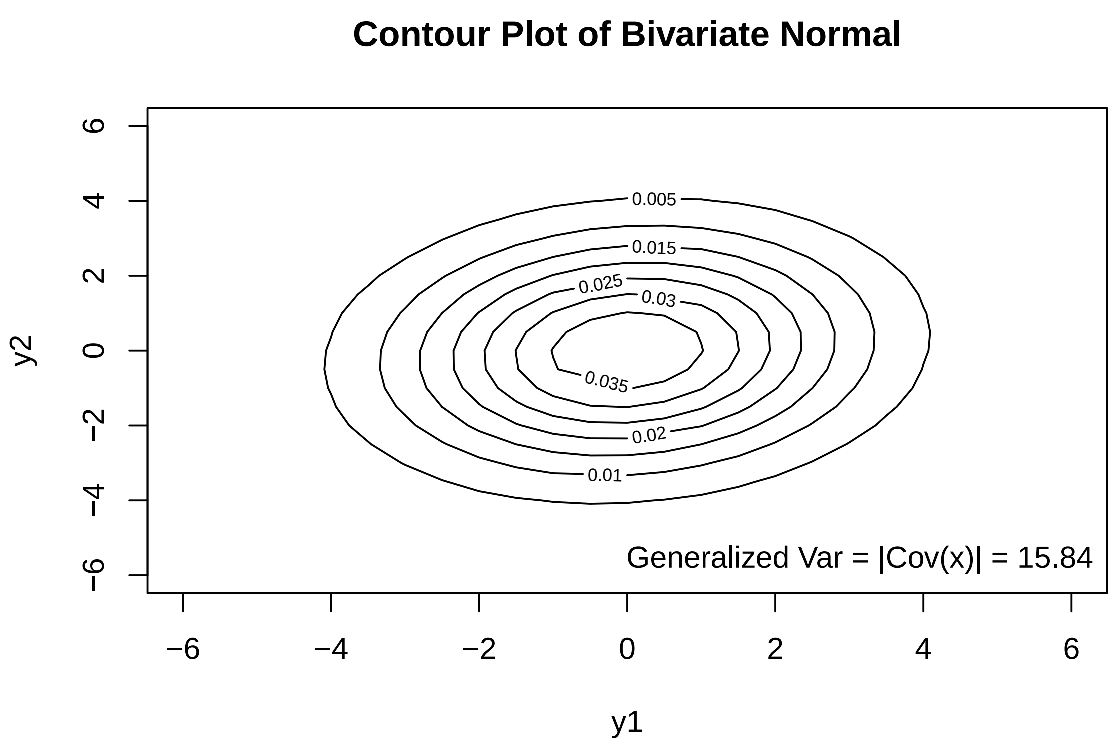
$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

and

$$\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$$


Bivariate Normal Density Plot (rho = 0.10)
z
y1
y2


Contour Plot of Bivariate Normal
0.005
0.015
0.025
0.03
0.035
0.02
0.01
Generalized Var = |Cov(x)| = 15.84
y2
y1
−6
−4
−2
0
2
4
6

2 Properties of the Multivariate Normal Random Variables

2.1 Distribution of Linear Combinations of Multivariate Normal Random Variables

If $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then a linear combination of y_i's is univariate normal, that is

$$\boldsymbol{a'y} = a_1 y_1 + a_2 y_2 + \cdots + a_p y_p \sim N(\boldsymbol{a'\mu}, \boldsymbol{a'\Sigma a})$$

Here we have

$$E(\boldsymbol{a'y}) = \boldsymbol{a'}E(\boldsymbol{y}) = \boldsymbol{a'\mu}$$

and

$$V(\boldsymbol{a'y}) = Cov(\boldsymbol{a'y}) = Cov(\boldsymbol{a'y}, \boldsymbol{a'y}) = \boldsymbol{a'}Cov(\boldsymbol{y}, \boldsymbol{y})\boldsymbol{a} = \boldsymbol{a'\Sigma a}$$

Remarks:

1. The covariance of two linear combinations $\boldsymbol{u = a'y}$ and $\boldsymbol{v = b'y}$ of multivariate normal random variable $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is given by

$$Cov(\boldsymbol{u}, \boldsymbol{v}) = Cov(\boldsymbol{a'y}, \boldsymbol{b'y}) = \boldsymbol{a'\Sigma b}$$

2. It can be shown that $\boldsymbol{a'\Sigma b} = \boldsymbol{b'\Sigma a}$.

Example 1.2.2

Consider a $p = 3$ random vector $\boldsymbol{y} = (y_1, y_2, y_3)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

a. Find the distribution of $W_1 = 2y_1 - y_2 - y_3$

b. Find the distribution of $W_2 = y_2 - y_3$

c. Find the covariance of W_1 and W_2.

Solution

a. We have $\boldsymbol{a'} = \begin{bmatrix} 2 & -1 & -1 \end{bmatrix}$. Hence,

$$\boldsymbol{a}'\boldsymbol{\mu} = \begin{bmatrix} 2 & -1 & -1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = 0$$

and

$$\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a} = \begin{bmatrix} 2 & -1 & -1 \end{bmatrix} \times \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = 8$$

Therefore, $W_1 \sim N(0, 8)$.

b. Left as a classroom exercise!

c. Left as a classroom exercise!

2.2 Joint Distribution of Linear Combinations of Multivariate Normal Random Variables

The joint distribution of q linear combinations of y_i's is multivariate normal, that is

$$\boldsymbol{A}_{q\times p}\boldsymbol{y}_{p\times 1} = \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_q' \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{y} \\ \boldsymbol{a}_2'\boldsymbol{y} \\ \vdots \\ \boldsymbol{a}_q'\boldsymbol{y} \end{bmatrix} \sim N_q(\boldsymbol{A}\boldsymbol{\mu}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}')$$

Example 1.2.3

Find the joint distribution of $W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}'$ in *Example 1.2.2.*

Solution

We have

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Thus,

$$\boldsymbol{A}\boldsymbol{\mu} = \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and

$$\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}' = \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 4 \end{bmatrix}$$

Therefore,

$$W \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}8 & 0\\0 & 4\end{bmatrix}\right)$$

<u>Example 1.2.4</u>

Suppose

$$\boldsymbol{y} = \begin{bmatrix}y_1\\y_2\end{bmatrix} \sim N_2\left(\begin{bmatrix}5\\1\end{bmatrix}, \begin{bmatrix}1 & 2\\2 & 4\end{bmatrix}\right)$$

Find the joint distribution of $W_1 = 2y_1 + 1$ and $W_2 = 2y_1 - y_2 + 3$.

<u>Solution</u>

We have

$$\boldsymbol{W} = \begin{bmatrix}W_1\\W_2\end{bmatrix} = \begin{bmatrix}2 & 0\\2 & -1\end{bmatrix} \times \begin{bmatrix}y_1\\y_2\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix}$$

Let

$$\boldsymbol{A} = \begin{bmatrix}2 & 0\\2 & -1\end{bmatrix}$$

and

$$\boldsymbol{c} = \begin{bmatrix}1\\3\end{bmatrix}$$

Thus,

$$\begin{aligned}
E[\boldsymbol{W}] &= \boldsymbol{A}E[\boldsymbol{y}] + \boldsymbol{c}\\
&= \begin{bmatrix}2 & 0\\2 & -1\end{bmatrix} \times E\begin{bmatrix}y_1\\y_2\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix}\\
&= \begin{bmatrix}2 & 0\\2 & -1\end{bmatrix} \times \begin{bmatrix}\mu_1\\\mu_2\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix}\\
&= \begin{bmatrix}2 & 0\\2 & -1\end{bmatrix} \times \begin{bmatrix}5\\1\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix}\\
&= \begin{bmatrix}11\\12\end{bmatrix}
\end{aligned}$$

and

$$\boldsymbol{A\Sigma A'} = \begin{bmatrix}2 & 0\\2 & -1\end{bmatrix} \times \begin{bmatrix}1 & 2\\2 & 4\end{bmatrix} \times \begin{bmatrix}2 & 2\\0 & -1\end{bmatrix} = \begin{bmatrix}4 & 0\\0 & 0\end{bmatrix}$$

Therefore,

$$\boldsymbol{W} \sim N_2\left(\begin{bmatrix}11\\12\end{bmatrix}, \begin{bmatrix}4 & 0\\0 & 0\end{bmatrix}\right)$$