6.4 Multivariate Normal

Unit 6: Joint Distributions and Conditional Expectation

Adapted from Blitzstein-Hwang Chapters 7 and 9.

The Multivariate Normal (MVN) is a continuous multivariate distribution that generalizes the Normal distribution into higher dimensions. We will not work with the rather unwieldy joint PDF of the Multivariate Normal. Instead we define the Multivariate Normal by its relationship to the ordinary Normal.

Definition 6.4.1 (Multivariate Normal distribution).

A random vector $\mathbf{X} = (X_1, \ldots, X_k)$ is said to have a *Multivariate Normal* (MVN) distribution if every linear combination of the X_j has a Normal distribution. That is, we require

$$t_1 X_1 + \cdots + t_k X_k$$

to have a Normal distribution for any choice of constants $t_1, \ldots, t_k$. If $t_1 X_1 + \cdots + t_k X_k$ is a constant (such as when all $t_i = 0$), we consider it to have a Normal distribution, albeit a degenerate Normal with variance 0. An important special case is $k = 2$; this distribution is called the *Bivariate Normal* (BVN).

If $(X_1, \ldots, X_k)$ is MVN, then the marginal distribution of X_1 is Normal, since we can take t_1 to be 1 and all other t_j to be 0. Similarly, the marginal distribution of each X_j is Normal. However, the converse is false: it is possible to have Normally distributed r.v.s $X_1, \ldots, X_k$ such that $(X_1, \ldots, X_k)$ is not Multivariate Normal.

Example 6.4.2 (Non-example of MVN).

Here is an example of two r.v.s whose marginal distributions are Normal but whose joint distribution is not Bivariate Normal. Let $X \sim \mathcal{N}(0, 1)$, and let

$$S = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

be a *random sign* independent of X. Then $Y = SX$ is a standard Normal r.v., due to the symmetry of the Normal distribution. However, (X, Y) is not Bivariate Normal because $P(X + Y = 0) = P(S = -1) = 1/2$, which implies that $X + Y$ can't be Normal (or, for that matter, have any continuous distribution). Since $X + Y$ is a linear combination of X and Y that is not Normally distributed, (X, Y) is not Bivariate Normal.

Example 6.4.3 (Actual MVN).

For $Z, W \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$, (Z, W) is Bivariate Normal because the sum of independent Normals is Normal. Also, $(Z + 2W, 3Z + 5W)$ is Bivariate Normal, since an arbitrary linear combination

$$t_1(Z + 2W) + t_2(3Z + 5W)$$

can also be written as a linear combination of Z and W,

$$(t_1 + 3t_2)Z + (2t_1 + 5t_2)W,$$

which is Normal.

The above example showed that if we start with a Multivariate Normal and take linear combinations of the components, we form a new Multivariate Normal. The next two theorems state that we can also produce new MVNs from old MVNs with the operations of subsetting and concatenation.

> **Theorem 6.4.4.**
>
> If (X_1, X_2, X_3) is Multivariate Normal, then so is the subvector (X_1, X_2).

Proof

Any linear combination $t_1X_1 + t_2X_2$ can be thought of as a linear combination of X_1, X_2, X_3 where the coefficient of X_3 is 0. So $t_1X_1 + t_2X_2$ is Normal for all t_1, t_2, which shows that (X_1, X_2) is MVN.

> **Theorem 6.4.5.**
>
> If $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_m)$ are MVN vectors with $\mathbf{X}$ independent of $\mathbf{Y}$, then the concatenated random vector $\mathbf{W} = (X_1, \ldots, X_n, Y_1, \ldots, Y_m)$ is Multivariate Normal.

Proof

Any linear combination $s_1X_1 + \cdots + s_nX_n + t_1Y_1 + \cdots + t_mY_m$ is Normal since $s_1X_1 + \cdots + s_nX_n$ and $t_1Y_1 + \cdots + t_mY_m$ are Normal (by definition of MVN) and are independent, so their sum is Normal. A Multivariate Normal distribution is fully specified by knowing the mean of each component, the variance of each component, and the covariance or correlation between any two components. Another way to say this is that the parameters of an MVN random vector $(X_1, \dots, X_k)$ are as follows:

- the *mean vector* $(\mu_1, \dots, \mu_k)$, where $E(X_j) = \mu_j$;
- the *covariance matrix*, which is the $k \times k$ matrix of covariances between components, arranged so that the row i, column j entry is $\text{Cov}(X_i, X_j)$.

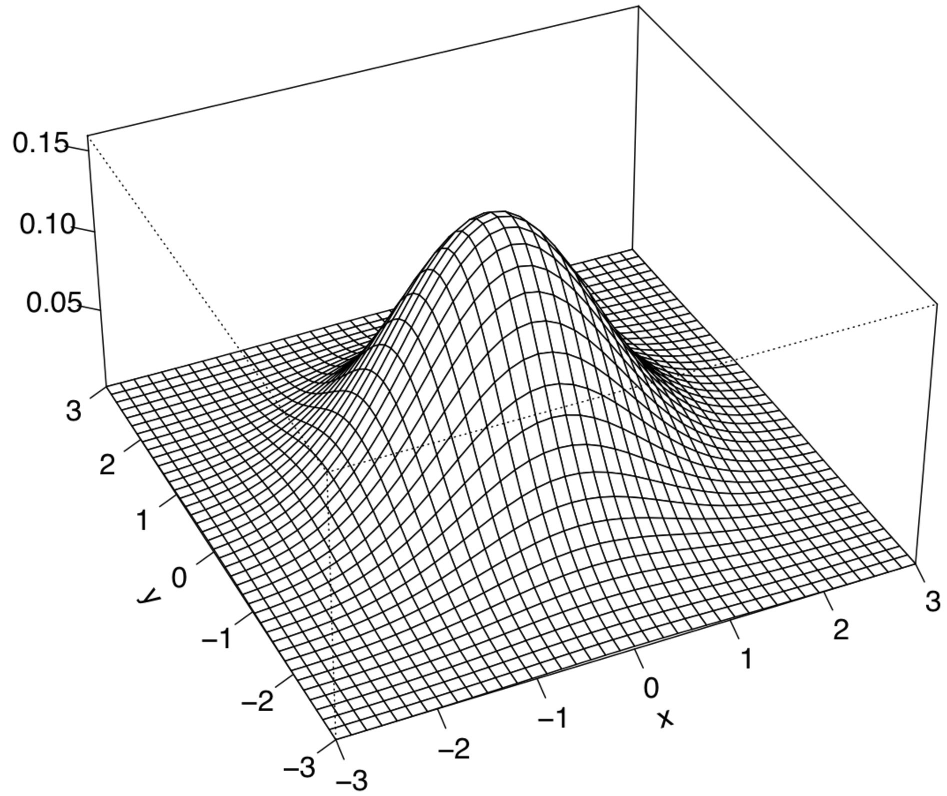
For example, in order to fully specify a Bivariate Normal distribution for (X, Y), we need to know five parameters:

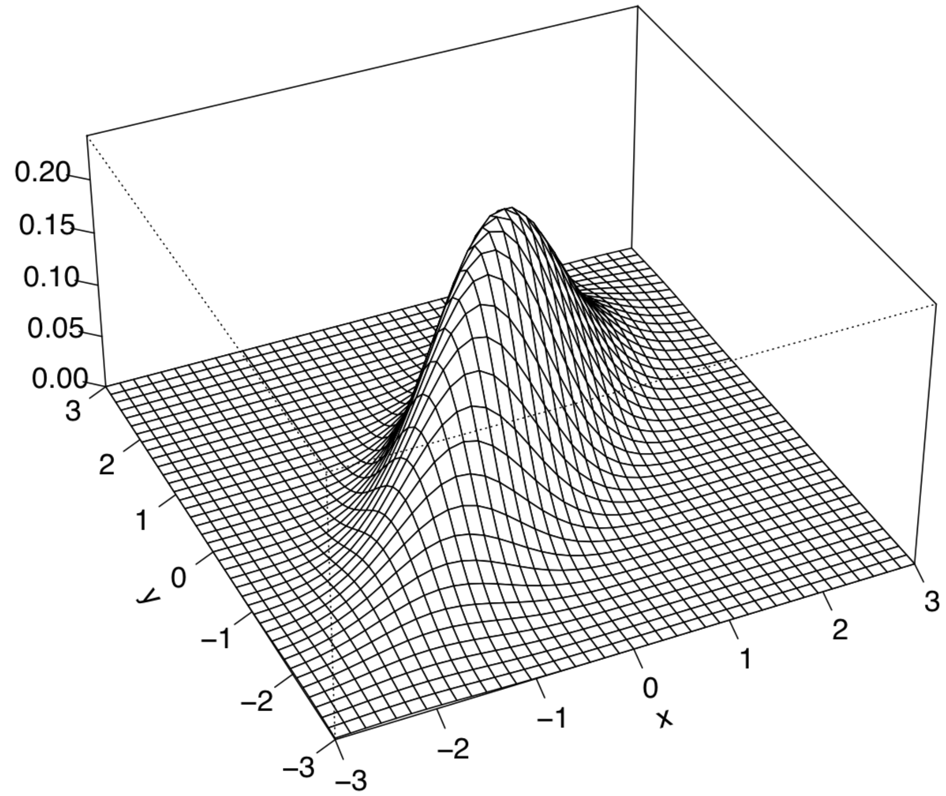
- the means $E(X)$, $E(Y)$;
- the variances $\text{Var}(X)$, $\text{Var}(Y)$;
- the correlation $\text{Corr}(X, Y)$.

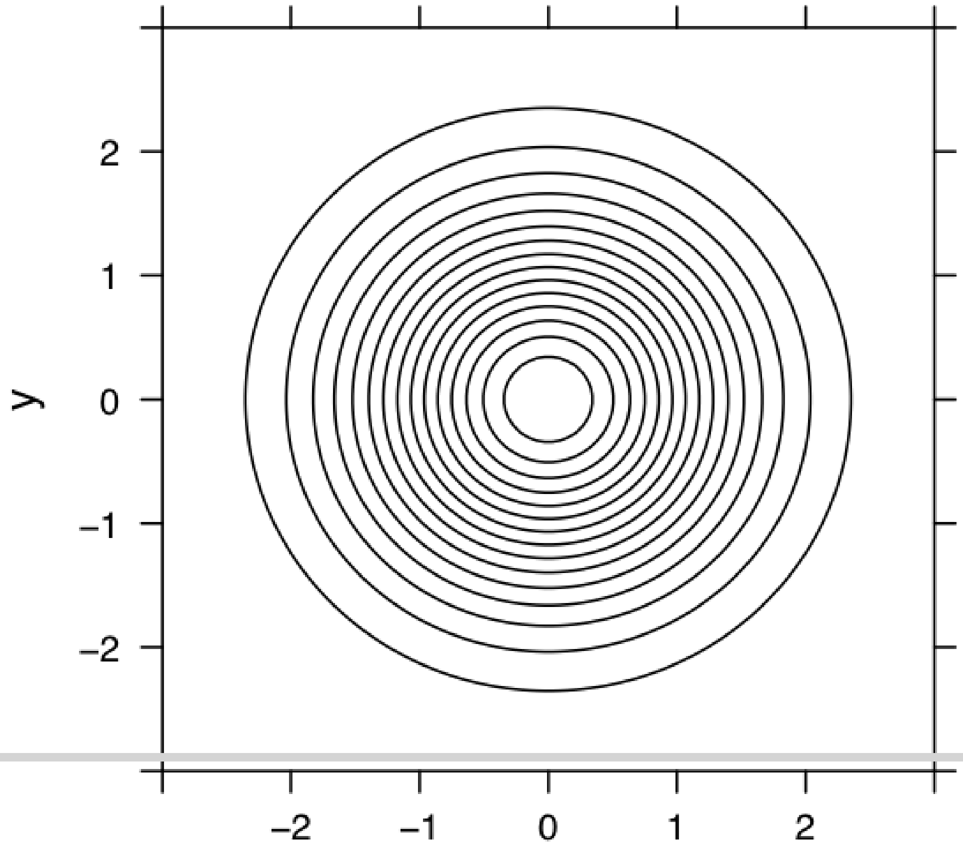
The joint PDF of a Bivariate Normal (X, Y) with $\mathcal{N}(0,1)$ marginal distributions and correlation $\rho \in (-1, 1)$ is

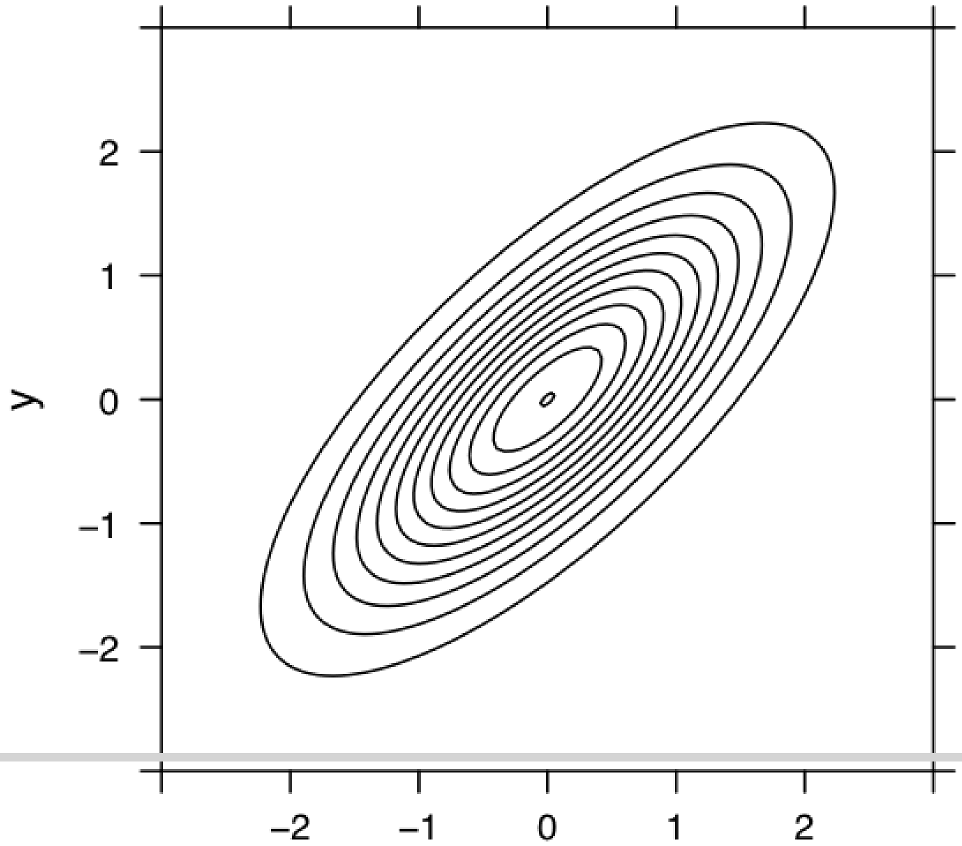
$$f_{X,Y}(x,y) = \frac{1}{2\pi\tau}\exp\left(-\frac{1}{2\tau^2}(x^2 + y^2 - 2\rho xy)\right),$$

with $\tau = \sqrt{1-\rho^2}$. Figure 6.4.6 plots the joint PDFs corresponding to two different Bivariate Normal distributions with $\mathcal{N}(0,1)$ marginals. On the left, X and Y are uncorrelated, so the contours of the joint PDF are shaped like circles. On the right, X and Y have a correlation of 0.75, so the level curves are ellipsoidal, reflecting the fact that X tends to be large when Y is large and vice versa.


0.15
0.10
0.05
y
x
3
2
1
0
−1
−2
−3


0.20
0.15
0.10
0.05
0.00
y
x
3
2
1
0
−1
−2
−3


y
2
1
0
−1
−2


y
2
1
0
−1
−2

x x

Figure 6.4.6: Joint PDFs of two Bivariate Normal distributions.
On the left, X and Y are marginally $\mathcal{N}(0,1)$ and have zero correlation.
Top: View Larger Image. Image Description.
Bottom: View Larger Image. Image Description.
On the right, X and Y are marginally $\mathcal{N}(0,1)$ and have correlation 0.75.
Top: View Larger Image. Image Description.
Bottom: View Larger Image. Image Description.

We know that in general, independence is a stronger condition than zero correlation; r.v.s can be uncorrelated but not independent. A special property of the Multivariate Normal distribution is that for r.v.s whose joint distribution is MVN, independence and zero correlation are equivalent conditions.

THEOREM 6.4.7.

Within an MVN random vector, uncorrelated implies independent. That is, if $\mathbf{X} \sim \text{MVN}$ can be written as $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are subvectors, and every component of $\mathbf{X}_1$ is uncorrelated with every component of $\mathbf{X}_2$, then $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent. In particular, if (X, Y) is Bivariate Normal and $\text{Corr}(X, Y) = 0$, then X and Y are independent.

Example 6.4.8 (Independence of sum and difference).

Let $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Find the joint distribution of $(X+Y, X-Y)$.

Solution

Since $(X+Y, X-Y)$ is Bivariate Normal and

$$\text{Cov}(X+Y, X-Y) = \text{Var}(X) - \text{Cov}(X,Y) + \text{Cov}(Y,X) - \text{Var}(Y) = 0,$$

$X+Y$ is independent of $X-Y$. Furthermore, they are i.i.d. $\mathcal{N}(0,2)$. By the same method, we have that if $X \sim \mathcal{N}(\mu_1, \sigma^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma^2)$ are independent (with the same variance), then $X+Y$ is independent of $X-Y$. It can be shown that the independence of the sum and difference is a unique characteristic of the Normal! That is, if X and Y are i.i.d. and $X+Y$ is independent of $X-Y$, then X and Y must have Normal distributions.

Learn About Verified Certificates