Introduction to bivariate normal distribution

Matthew Stephens
University of Chicago
February 10, 2026

See here for a PDF version of this vignette.

Prerequisites

You need to have basic familiarity with univariate normal distribution, and understand the basic property that linear combinations of normals are also normal.

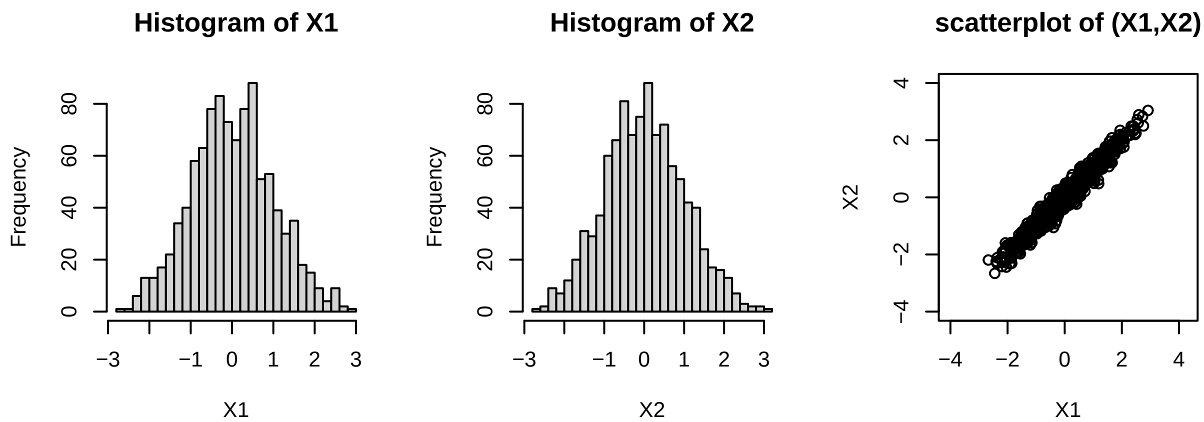
Motivating example

Suppose that Z_1, Z_2 are independent standard normal $N(0,1)$, and define $X_1 = Z_1 + 0.1Z_2$, $X_2 = Z_1 - 0.1Z_2$. What is the joint distribution of X_1, X_2?

We know from the basic property that X_1 will be univariate normal, and that X_2 will be univariate normal. However, they will not necessarily be independent because Z_1 and Z_2 were used to compute both. Indeed, you can see that X_1 and X_2 might both be expected to be close to Z_1 (because the 0.1 multiplier on Z_2 is "relatively small"). So when X_1 is X_2 should also be big, and when X_1 is small X_2 should also be small.

The following code illustrates this: the histograms illustrate both X_1 and X_2 are normal, and the scatterplot of X_1 and X_2 shows they are correlated (and the sample correlation is approximately 0.98).

```
Z1 <- rnorm(1000)
Z2 <- rnorm(1000)
X1 <- Z1 + 0.1*Z2
X2 <- Z1 - 0.1*Z2
par(mfcol = c(1,3))
hist(X1,breaks = 32)
hist(X2,breaks = 32)
plot(X1,X2,main = "scatterplot of (X1,X2)",ylim = c(-4,4),asp = 1)
```




```
cor(X1,X2)
# [1] 0.9792
```

The bivariate normal distribution

In fact the answer to the question "what is the joint distribution of X_1, X_2" is they have a "bivariate normal distribution". Thus the scatterplot shown above shows a scatterplot of 1,000 samples from a bivariate normal distribution. The prefix "bi" means two, referring to the fact that here we are looking at two variables, X_1 and X_2. The ideas here can be extended to more variables, and the resulting distribution is called the "multivariate normal" (multi = two or more). The bivariate normal is a special case of the multivariate normal.

Mean and covariance matrix

The bivariate normal distribution has 5 parameters: two means (for X_1 and X_2), two variances (for X_1 and X_2) and the covariance between X_1 and X_2. It is usual to write the mean parameter as a vector μ and the variance and covariance parameters as a 2×2 symmetric matrix Σ where the diagonal elements of Σ contain the variances and the off-diagonal elements contain the covariance. Σ is called the "covariance matrix" or sometimes the "variance-covariance matrix".

General construction

Suppose Z_1, Z_2 are independent random variables each with a standard normal distribution $N(0,1)$. Let Z denote the vector (Z_1, Z_2), let A be any 2×2 matrix, and μ be any r-vector. Then the vector $X = AZ + \mu$ has a bivariate normal distribution with mean μ and variance-covariance matrix $\Sigma = AA^T$. (Here, A^T means the transpose of the matrix A.) We write $X \sim N_2(\mu, \Sigma)$.

Example

We can redo the example above with vector and matrix notation, $\mu = (0,0)$ and $A = \begin{bmatrix} 1 & 0.1 \\ 1 & -0.1 \end{bmatrix}$. Here for clarity we just simulate a single sample instead of 1,000:

```r
mu <- c(0,0)
A <- rbind(c(1,0.1),c(1,-0.1))
A
#      [,1] [,2]
# [1,]    1  0.1
# [2,]    1 -0.1
z <- rnorm(2)
x <- mu + A %*% z
x
#       [,1]
# [1,] 1.246
# [2,] 1.541
```

It should be clear from the above that in our example the mean is $\mu = (0,0)$. What is the covariance matrix Σ? We can compute it from the formula $\Sigma = AA^T$:

```r
Sigma <- A %*% t(A)
Sigma
#      [,1] [,2]
# [1,] 1.01 0.99
# [2,] 0.99 1.01
```