Lecture 22: Bivariate Normal Distribution

Statistics 104
Colin Rundel
April 11, 2012

6.5 Conditional Distributions

General Bivariate Normal

Let $Z_1, Z_2 \sim \mathcal{N}(0, 1)$, which we will use to build a general bivariate normal distribution.

$$f(z_1, z_2) = \frac{1}{2\pi} \exp \left[ -\frac{1}{2} (z_1^2 + z_2^2) \right]$$

We want to transform these unit normal distributions to have the follow arbitrary parameters: $\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho$

$$X = \sigma_X Z_1 + \mu_X$$

$$Y = \sigma_Y [\rho Z_1 + \sqrt{1 - \rho^2} Z_2] + \mu_Y$$

General Bivariate Normal - Marginals

First, let's examine the marginal distributions of $X$ and $Y$,

$$X = \sigma_X Z_1 + \mu_X$$

$$= \sigma_X \mathcal{N}(0, 1) + \mu_X$$

$$= \mathcal{N}(\mu_X, \sigma_X^2)$$

$$Y = \sigma_Y [\rho Z_1 + \sqrt{1 - \rho^2} Z_2] + \mu_Y$$

$$= \sigma_Y \mathcal{N}(0, 1) + \sqrt{1 - \rho^2} \mathcal{N}(0, 1) + \mu_Y$$

$$= \sigma_Y \mathcal{N}(0, 1) + \mathcal{N}(\mu_Y, \sigma_Y^2)$$

General Bivariate Normal - Cov/Corr

Second, we can find $\text{Cov}(X, Y)$ and $\rho(X, Y)$

$$\text{Cov}(X, Y) = E [(X - E(X))(Y - E(Y))]$$

$$= E \left[ (\sigma_X Z_1 + \mu_X - \mu_X)(\sigma_Y [\rho Z_1 + \sqrt{1 - \rho^2} Z_2] + \mu_Y - \mu_Y) \right]$$

$$= E \left[ \sigma_X \sigma_Y [\rho Z_1 + \sqrt{1 - \rho^2} Z_2] \right]$$

$$= \sigma_X \sigma_Y \rho E[Z_1^2]$$

$$= \sigma_X \sigma_Y \rho$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \rho$$
Consequently, if we want to generate a Bivariate Normal random variable with \( X \sim \mathcal{N}(\mu_X, \sigma_X^2) \) and \( Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2) \) where the correlation of \( X \) and \( Y \) is \( \rho \) we can generate two independent unit normals \( Z_1 \) and \( Z_2 \) and use the transformation:

\[
X = \sigma_X Z_1 + \mu_X \\
Y = \sigma_Y [\rho Z_1 + \sqrt{1-\rho^2} Z_2] + \mu_Y
\]

We can also use this result to find the joint density of the Bivariate Normal using a 2d change of variables.

Next we calculate the Jacobian,

\[
J = \det \left[ \begin{array}{ccc}
\frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial y} \\
\frac{\partial h_2}{\partial x} & \frac{\partial h_2}{\partial y}
\end{array} \right] = \det \left[ \begin{array}{cc}
\frac{1}{\sigma_X \sqrt{1-\rho^2}} & 0 \\
0 & \frac{1}{\sigma_Y \sqrt{1-\rho^2}}
\end{array} \right] = \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}}
\]

The joint density of \( X \) and \( Y \) is then given by

\[
f(x, y) = \frac{1}{2\pi} \exp \left[ -\frac{1}{2} (x^2 + y^2) \right] |J| = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \exp \left[ -\frac{1}{2} \left( \frac{\rho x - \mu_X}{\sigma_X} \right)^2 \right]
\]

Therefore,

\[
s_1(x, y) = \frac{x - \mu_X}{\sigma_X} \\
s_2(x, y) = \frac{1}{\sqrt{1-\rho^2}} \left[ \frac{y - \mu_Y}{\sigma_Y} - \rho x - \mu_X \right]
\]
Conditional Distributions

General Bivariate Normal - Density (Matrix Notation)

Obviously, the density for the Bivariate Normal is ugly, and it only gets worse when we consider higher dimensional joint densities of normals. We can write the density in a more compact form using matrix notation,

\[
\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix}
\]

\[
f(x) = \frac{1}{2\pi(\det \Sigma)^{-1/2}} \exp \left[ -\frac{1}{2} \left( \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \right)^T \Sigma^{-1} \left( \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \right) \right]
\]

We can confirm our results by checking the value of \((\det \Sigma)^{-1/2}\) and \((x - \mu)^T \Sigma^{-1} (x - \mu)\) for the bivariate case.

\[
(\det \Sigma)^{-1/2} = \left( \sigma_x^2 \sigma_y^2 - \rho^2 \sigma_x^2 \sigma_y^2 \right)^{-1/2} = \frac{1}{\sigma_x \sigma_y (1 - \rho^2)^{1/2}}
\]

Recall for a \(2 \times 2\) matrix,

\[
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}
\]

Then,

\[
(x - \mu)^T \Sigma^{-1} (x - \mu) = \frac{1}{\sigma_x^2 \sigma_y^2 (1 - \rho^2)} \left( \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \right)^T \left( \begin{pmatrix} \sigma_x^2 & -\rho \sigma_x \sigma_y \\ -\rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix} \right) \left( \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \right)
\]

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

\[
= \frac{1}{1 - \rho^2} \left( \frac{(x - \mu_x)^2}{\sigma_x^2} - 2 \rho \frac{(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y} + \frac{(y - \mu_y)^2}{\sigma_y^2} \right)
\]
Conditional Distributions

General Bivariate Normal - Examples

\[ X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 1) \]
\[ \rho = -0.25 \]
\[ X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 1) \]
\[ \rho = -0.5 \]
\[ X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 1) \]
\[ \rho = -0.75 \]
\[ X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 2) \]
\[ \rho = -0.5 \]
\[ X \sim \mathcal{N}(0, 2), \ Y \sim \mathcal{N}(0, 1) \]
\[ \rho = -0.75 \]
\[ X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 2) \]
\[ \rho = -0.75 \]

Multivariate Normal Distribution

Matrix notation allows us to easily express the density of the multivariate normal distribution for an arbitrary number of dimensions. We express the \( k \)-dimensional multivariate normal distribution as follows,

\[ X \sim \mathcal{N}_k(\mu, \Sigma) \]

where \( \mu \) is the \( k \times 1 \) column vector of means and \( \Sigma \) is the \( k \times k \) covariance matrix where \( \{\Sigma\}_{i,j} = \text{Cov}(X_i, X_j) \).

The density of the distribution is

\[
f(x) = \frac{1}{(2\pi)^{k/2}(\det \Sigma)^{-1/2}} \exp \left[ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right]
\]
**Conditional Distributions**

**Multivariate Normal Distribution - RNG**

Let $Z_1, \ldots, Z_k \sim \mathcal{N}(0, 1)$ and $Z = (Z_1, \ldots, Z_k)^T$ then

$$
\mu + \text{Chol}(\Sigma)Z \sim \mathcal{N}_k(\mu, \Sigma)
$$

This is offered without proof in the general $k$-dimensional case, but we can check that this results in the same transformation we started with in the bivariate case and should justify how we knew to use that particular transformation.

**Statistics 104 (Colin Rundel)  Lecture 22  April 11, 2012  16 / 22**

**Cholesky and the Bivariate Transformation**

We need to find the Cholesky decomposition of $\Sigma$ for the general bivariate case where

$$
\Sigma = \begin{pmatrix}
\sigma_X^2 & \rho \sigma_X \sigma_Y \\
\rho \sigma_X \sigma_Y & \sigma_Y^2
\end{pmatrix}
$$

We need to solve the following for $a, b, c$

$$
\begin{pmatrix}
a \\
b \\
c
\end{pmatrix} = \begin{pmatrix}
\sigma_X^2 & ab \\
ab & b^2 + c^2
\end{pmatrix} = \begin{pmatrix}
\sigma_X^2 & \rho \sigma_X \sigma_Y \\
\rho \sigma_X \sigma_Y & \sigma_Y^2
\end{pmatrix}
$$

This gives us three (unique) equations and three unknowns to solve for,

$$
\begin{align*}
a^2 &= \sigma_X^2 \\
b &= \rho \sigma_X \sigma_Y / a &= \rho \sigma_Y \\
c &= \sqrt{\sigma_Y^2 - b^2} &= \sigma_Y(1 - \rho^2)^{1/2}
\end{align*}
$$

**Conditional Expectation of the Bivariate Normal**

Using $X = \mu_X + \sigma_X Z_1$ and $Y = \mu_Y + \sigma_Y [\rho Z_1 + (1 - \rho^2)^{1/2} Z_2]$ where $Z_1, Z_2 \sim \mathcal{N}(0, 1)$ we can find $E(Y|X)$.

$$
E[Y|X = x] = E \left[ \mu_Y + \sigma_Y \left( \rho Z_1 + (1 - \rho^2)^{1/2} Z_2 \right) \bigg| X = x \right] \\
= E \left[ \mu_Y + \sigma_Y \left( \frac{x - \mu_X}{\sigma_X} \right) + (1 - \rho^2)^{1/2} Z_2 \bigg| X = x \right] \\
= \mu_Y + \sigma_Y \left( \frac{x - \mu_X}{\sigma_X} + (1 - \rho^2)^{1/2} E[Z_2|X = x] \right) \\
= \mu_Y + \sigma_Y \rho \left( \frac{x - \mu_X}{\sigma_X} \right)
$$

By symmetry,

$$
E[X|Y = y] = \mu_X + \sigma_X \rho \left( \frac{y - \mu_Y}{\sigma_Y} \right)
$$
6.5 Conditional Distributions

Conditional Variance of the Bivariate Normal

Using $X = \mu_X + \sigma_X Z_1$ and $Y = \mu_Y + \sigma_Y \left[ \rho Z_1 + (1 - \rho^2)^{1/2} Z_2 \right]$ where $Z_1, Z_2 \sim \mathcal{N}(0,1)$ we can find $\text{Var}(Y|X)$.

$$\text{Var}[Y|X = x] = \text{Var} \left[ \mu_Y + \sigma_Y \left( \rho Z_1 + (1 - \rho^2)^{1/2} Z_2 \right) \, | \, X = x \right]$$

$$= \text{Var} \left[ \mu_Y + \sigma_Y \left( \rho \frac{x - \mu_X}{\sigma_X} + (1 - \rho^2)^{1/2} Z_2 \right) \, | \, X = x \right]$$

$$= \text{Var}[\sigma_Y (1 - \rho^2) Z_2 | X = x]$$

$$= \sigma_Y^2 (1 - \rho^2)$$

By symmetry,

$$\text{Var}[X|Y = y] = \sigma_X^2 (1 - \rho^2)$$

Example - Husbands and Wives

Suppose that the heights of married couples can be explained by a bivariate normal distribution. If the wives have a mean height of 66.8 inches and a standard deviation of 2 inches while the heights of the husbands have a mean of 70 inches and a standard deviation of 2 inches. The correlation between the heights is 0.68. What is the probability that for a randomly selected couple the wife is taller than her husband?

Example - Conditionals

Suppose that $X_1$ and $X_2$ have a bivariate normal distribution where $E(X_1|X_2) = 3.7 - 0.15X_2$, $E(X_2|X_1) = 0.4 - 0.6X_1$, and $\text{Var}(X_2|X_1) = 3.64$.

Find $E(X_1)$, $\text{Var}(X_1)$, $E(X_2)$, $\text{Var}(X_2)$, and $\rho(X_1, X_2)$. 

Example - Conditionals, cont.