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: $\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho$

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

General Bivariate Normal – Marginals

First, lets examine the marginal distributions of X and Y,

$$\begin{aligned} X &= \sigma_X Z_1 + \mu_X \\ &= \sigma_X \mathcal{N}(0,1) + \mu_X \\ &= \mathcal{N}(\mu_X, \sigma_X^2) \end{aligned}$$

$$\begin{aligned} Y &= \sigma_Y[\rho Z_1 + \sqrt{1-\rho^2} Z_2] + \mu_Y \\ &= \sigma_Y[\rho \mathcal{N}(0,1) + \sqrt{1-\rho^2}\mathcal{N}(0,1)] + \mu_Y \\ &= \sigma_Y[\mathcal{N}(0,\rho^2) + \mathcal{N}(0, 1-\rho^2)] + \mu_Y \\ &= \sigma_Y \mathcal{N}(0,1) + \mu_Y \\ &= \mathcal{N}(\mu_Y, \sigma_Y^2) \end{aligned}$$

General Bivariate Normal - Cov/Corr

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

$$
\begin{aligned}
Cov(X, Y) &= E\left[(X - E(X))(Y - E(Y))\right] \\
&= E\left[(\sigma_X Z_1 + \mu_X - \mu_X)(\sigma_Y[\rho Z_1 + \sqrt{1-\rho^2} Z_2] + \mu_Y - \mu_Y)\right] \\
&= E\left[(\sigma_X Z_1)(\sigma_Y[\rho Z_1 + \sqrt{1-\rho^2} Z_2])\right] \\
&= \sigma_X \sigma_Y E\left[\rho Z_1^2 + \sqrt{1-\rho^2} Z_1 Z_2\right] \\
&= \sigma_X \sigma_Y \rho E[Z_1^2] \\
&= \sigma_X \sigma_Y \rho
\end{aligned}
$$

$$
\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \rho
$$

General Bivariate Normal - RNG

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 ρ we can generate two independent unit normals Z_1 and Z_2 and use the transformation:

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

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

Multivariate Change of Variables

Let $X_1, \dots, X_n$ have a continuous joint distribution with pdf f defined of S. We can define n new random variables $Y_1, \dots, Y_n$ as follows:

$$Y_1 = r_1(X_1, \dots, X_n) \quad \cdots \quad Y_n = r_n(X_1, \dots, X_n)$$

If we assume that the n functions $r_1, \dots, r_n$ define a one-to-one differentiable transformation from S to T then let the inverse of this transformation be

$$x_1 = s_1(y_1, \dots, y_n) \quad \cdots \quad x_n = s_n(y_1, \dots, y_n)$$

Then the joint pdf g of $Y_1, \dots, Y_n$ is

$$g(y_1, \dots, y_n) = \begin{cases} f(s_1, \dots, s_n)|J| & \text{for } (y_1, \dots, y_n) \in T \\ 0 & \text{otherwise} \end{cases}$$

Where

$$J = \det \begin{bmatrix} \frac{\partial s_1}{\partial y_1} & \cdots & \frac{\partial s_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_n}{\partial y_1} & \cdots & \frac{\partial s_n}{\partial y_n} \end{bmatrix}$$

General Bivariate Normal - Density

The first thing we need to find are the inverses of the transformation. If $x = r_1(z_1, z_2)$ and $y = r_2(z_1, z_2)$ we need to find functions h_1 and h_2 such that $Z_1 = s_1(X, Y)$ and $Z_2 = s_2(X, Y)$.

$$\begin{aligned} X &= \sigma_X Z_1 + \mu_X \\ Z_1 &= \frac{X - \mu_X}{\sigma_X} \end{aligned}$$

$$\begin{aligned} Y &= \sigma_Y[\rho Z_1 + \sqrt{1-\rho^2} Z_2] + \mu_Y \\ \frac{Y - \mu_Y}{\sigma_Y} &= \rho \frac{X - \mu_X}{\sigma_X} + \sqrt{1-\rho^2} Z_2 \\ Z_2 &= \frac{1}{\sqrt{1-\rho^2}} \left[\frac{Y - \mu_Y}{\sigma_Y} - \rho \frac{X - \mu_X}{\sigma_X} \right] \end{aligned}$$

Therefore,

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

General Bivariate Normal - Density

Next we calculate the Jacobian,

$$J = \det \begin{bmatrix} \frac{\partial s_1}{\partial x} & \frac{\partial s_1}{\partial y} \\ \frac{\partial s_2}{\partial x} & \frac{\partial s_2}{\partial y} \end{bmatrix} = \det \begin{bmatrix} \frac{1}{\sigma_X} & 0 \\ \frac{-\rho}{\sigma_X\sqrt{1-\rho^2}} & \frac{1}{\sigma_Y\sqrt{1-\rho^2}} \end{bmatrix} = \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}}$$

The joint density of X and Y is then given by

$$\begin{aligned} f(x,y) &= f(z_1, z_2)|J| \\ &= \frac{1}{2\pi} \exp\left[-\frac{1}{2}(z_1^2 + z_2^2)\right] |J| = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}(z_1^2 + z_2^2)\right] \\ &= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \frac{1}{1-\rho^2}\left(\frac{y-\mu_Y}{\sigma_Y} - \rho\frac{x-\mu_X}{\sigma_X}\right)^2\right]\right] \\ &= \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}} \exp\left[\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}\right)\right] \end{aligned}$$

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,

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$$

$$f(\mathbf{x}) = \frac{1}{2\pi}(\det \boldsymbol{\Sigma})^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

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

$$(\det \boldsymbol{\Sigma})^{-1/2} = \left(\sigma_X^2\sigma_Y^2 - \rho^2\sigma_X^2\sigma_Y^2\right)^{-1/2} = \frac{1}{\sigma_X\sigma_Y(1-\rho^2)^{1/2}}$$

General Bivariate Normal - Density (Matrix Notation)

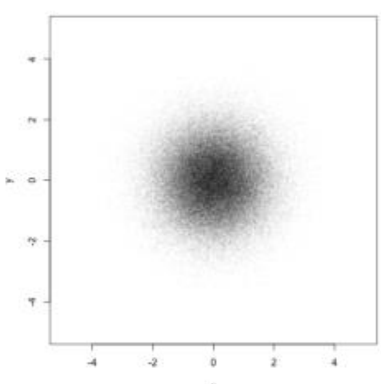
Recall for a 2×2 matrix,

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \mathbf{A}^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Then,

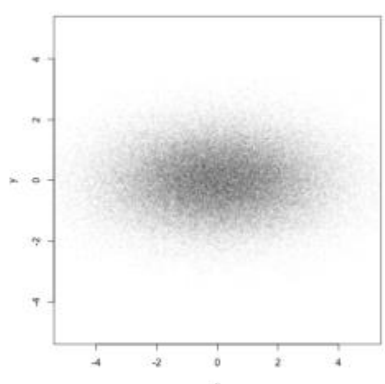
$$\begin{aligned}
&(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \\
&= \frac{1}{\sigma_X^2 \sigma_Y^2 (1 - \rho^2)} \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix}^T \begin{pmatrix} \sigma_Y^2 & -\rho\sigma_X\sigma_Y \\ -\rho\sigma_X\sigma_Y & \sigma_X^2 \end{pmatrix} \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \\
&= \frac{1}{\sigma_X^2 \sigma_Y^2 (1 - \rho^2)} \begin{pmatrix} \sigma_Y^2 (x - \mu_X) - \rho\sigma_X\sigma_Y (y - \mu_Y) \\ -\rho\sigma_X\sigma_Y (x - \mu_X) + \sigma_X^2 (y - \mu_Y) \end{pmatrix}^T \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \\
&= \frac{1}{\sigma_X^2 \sigma_Y^2 (1 - \rho^2)} \left(\sigma_Y^2 (x - \mu_X)^2 - 2\rho\sigma_X\sigma_Y (x - \mu_X)(y - \mu_Y) + \sigma_X^2 (y - \mu_Y)^2 \right) \\
&= \frac{1}{1 - \rho^2} \left(\frac{(x - \mu_X)^2}{\sigma_X^2} - 2\rho \frac{(x - \mu_X)(y - \mu_Y)}{\sigma_X \sigma_Y} + \frac{(y - \mu_Y)^2}{\sigma_Y^2} \right)
\end{aligned}$$

General Bivariate Normal - Examples



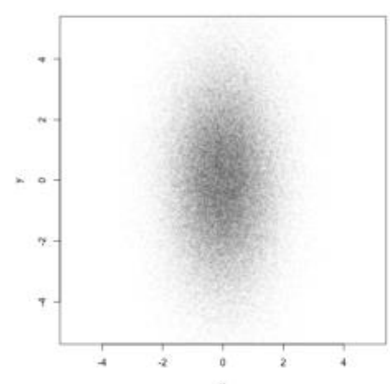
$X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 1)$

$\rho = 0$



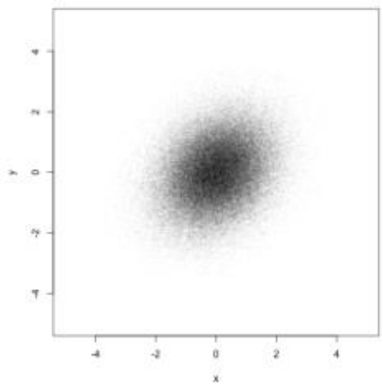
$X \sim \mathcal{N}(0, 2), \ Y \sim \mathcal{N}(0, 1)$

$\rho = 0$



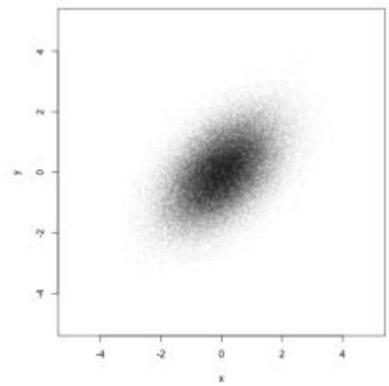
$X \sim \mathcal{N}(0, 1), \ Y \sim \mathcal{N}(0, 2)$

$\rho = 0$



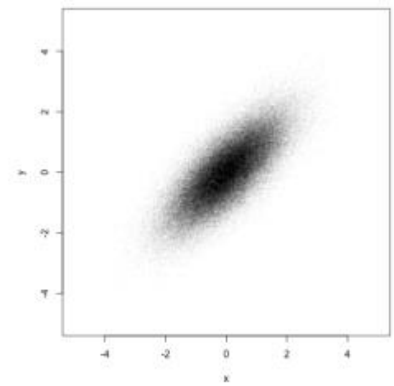
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = 0.25$



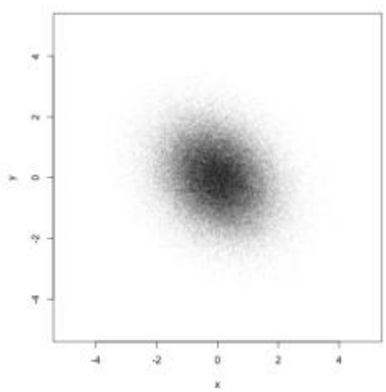
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = 0.5$



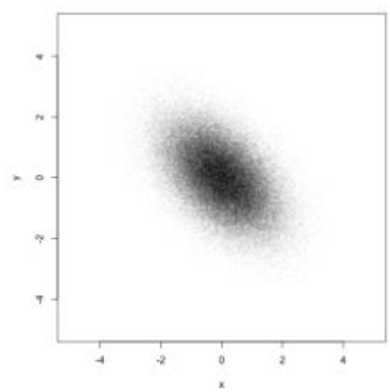
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = 0.75$



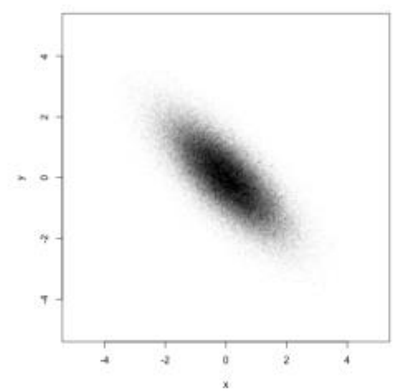
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = -0.25$



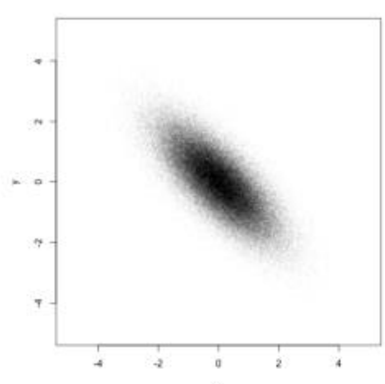
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = -0.5$



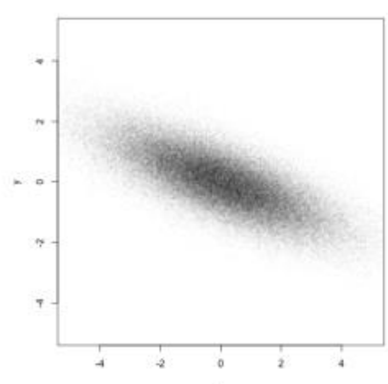
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = -0.75$



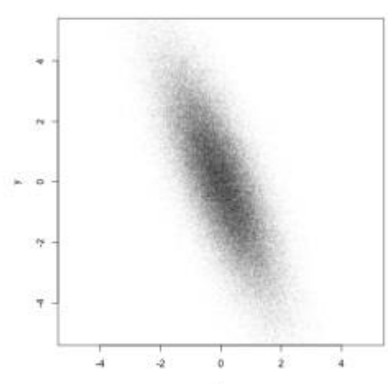
$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = -0.75$



$X \sim \mathcal{N}(0,2),\ \ Y \sim \mathcal{N}(0,1)$

$\rho = -0.75$



$X \sim \mathcal{N}(0,1),\ \ Y \sim \mathcal{N}(0,2)$

$\rho = -0.75$

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,

$$\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

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

The density of the distribution is

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

Multivariate Normal Distribution - Cholesky

In the bivariate case, we had a nice transformation such that we could generate two independent unit normal values and transform them into a sample from an arbitrary bivariate normal distribution.

There is a similar method for the multivariate normal distribution that takes advantage of the Cholesky decomposition of the covariance matrix.

The Cholesky decomposition is defined for a symmetric, positive definite matrix $\mathbf{X}$ as

$$\mathbf{L} = \text{Chol}(\mathbf{X})$$

where $\mathbf{L}$ is a lower triangular matrix such that $\mathbf{L}\mathbf{L}^T = \mathbf{X}$.

Multivariate Normal Distribution - RNG

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

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

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.

Cholesky and the Bivariate Transformation

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

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

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

$$\begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & b^2 + c^2 \end{pmatrix} = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$$

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

$$a^2 = \sigma_X^2 \qquad ab = \rho\sigma_X\sigma_Y \qquad b^2 + c^2 = \sigma_Y^2$$

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

Cholesky and the Bivariate Transformation

Let $Z_1, Z_2 \sim \mathcal{N}(0,1)$ then

$$\begin{aligned}
\begin{pmatrix} X \\ Y \end{pmatrix} &= \boldsymbol{\mu} + \text{Chol}(\boldsymbol{\Sigma})\mathbf{Z} \\
&= \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} + \begin{pmatrix} \sigma_X & 0 \\ \rho\sigma_Y & \sigma_Y(1-\rho^2)^{1/2} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \\
&= \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} + \begin{pmatrix} \sigma_X Z_1 \\ \rho\sigma_Y Z_1 + \sigma_Y(1-\rho^2)^{1/2} Z_2 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
X &= \mu_X + \sigma_X Z_1 \\
Y &= \mu_Y + \sigma_Y[\rho Z_1 + (1-\rho^2)^{1/2} Z_2]
\end{aligned}$$

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 $E(Y|X)$.

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

By symmetry,

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

Conditional Variance 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 $Var(Y|X)$.

$$\begin{aligned}
Var[Y|X=x] &= Var\left[\mu_Y + \sigma_Y\left(\rho Z_1 + (1-\rho^2)^{1/2} Z_2\right)\Big| X = x\right] \\
&= Var\left[\mu_Y + \sigma_Y\left(\rho \frac{x-\mu_X}{\sigma_X} + (1-\rho^2)^{1/2} Z_2\right)\Big| X = x\right] \\
&= Var[\sigma_Y(1-\rho^2)Z_2 | X = x] \\
&= \sigma_Y^2(1-\rho^2)
\end{aligned}$$

By symmetry,

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

Example - Husbands and Wives (Example 5.10.6, deGroot)

Suppose that the heights of married couples can be explained by a bivariate normal distribution. If the wives have a mean heigh 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 $Var(X_2|X_1) = 3.64$.

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