Correlated Gaussians

Angelo Gnazzo, Leonardo Pagani

1 Introduction

You are only equipped with a uniform sampler U(0,1), and with our beloved standard normal table .Devise a strategy to build two standard gaussian random variables $X \sim \mathcal{N}(0,1), Y \sim \mathcal{N}(0,1)$, such that their correlation coefficient $\rho_{xy} = \rho$.

2 Solution

First, we have to find a way to obtain Gaussian samples from our uniform samples. We can obtain the Gaussian samples with the inverse transform method. Let Φ_Z be the cdf of a standard normal random variable $Z \sim \mathcal{N}(0,1)$. Let $U \sim \mathcal{U}(0,1)$ be our uniform sampler random variable. We have that:

$$\mathbb{P}(\Phi_Z^{-1}(U) \leq z) = \mathbb{P}(\Phi_Z(\Phi_Z^{-1}(U)) \leq \Phi_Z(z)) = \mathbb{P}(U \leq \Phi_Z(z)) = \Phi_Z(z)$$

Thus, we have proved that the transformed variable $\Phi_Z^{-1}(U)$ has cdf $\Phi_Z(z)$ and it's therefore distributed as a standard Gaussian.

If u is a sample from $U \sim \mathcal{U}(0,1)$, then $\Phi_Z^{-1}(u)$ is a sample from $Z \sim \mathcal{N}(0,1)$. We can use the gaussian table to estimate Φ_Z^{-1}.

Now how do we build $X \sim \mathcal{N}(0,1), Y \sim \mathcal{N}(0,1)$, such that $\rho_{xy} = \rho$?

We can build X and Y as follows.

Let $U_1 \sim \mathcal{U}(0,1)$ and $U_2 \sim \mathcal{U}(0,1)$ be two uniform independent random variables. We have access to their pdf because we can repeatedly use the sampler. Let's set:

$$X = \Phi_Z^{-1}(U_1)$$
$$Y = \rho \cdot \Phi_Z^{-1}(U_1) + \sqrt{1-\rho^2} \cdot \Phi_Z^{-1}(U_2)$$

We have already shown that $\Phi_Z^{-1}(U_1)$ and $\Phi_Z^{-1}(U_2)$ are standard gaussians. We will respectively call them $Z_1 \sim \mathcal{N}(0,1)$ and $Z_2 \sim \mathcal{N}(0,1)$. They are also independent random variables as they are constructed from independent random variables (the samplers).

Then:

$X = Z_1$
$Y = \rho \cdot Z_1 + \sqrt{1-\rho^2} \cdot Z_2$

We now notice that X and Y are linear combinations of independent Gaussians and they are therefore Gaussians random variables as well. We only have to compute their mean and variance to have complete information on their pdf.

Trivially, for X:

$\mathbb{E}[X] = \mathbb{E}[Z_1] = 0$

$\mathrm{Var}[X] = \mathrm{Var}[Z_1] = 1$

as $Z_1 \sim \mathcal{N}(0,1)$.

For Y:

$$\text{E[Y]} \ = \mathbb{E}[\rho Z_1 + \sqrt{1-\rho^2} \cdot Z_2] = \mathbb{E}[\rho Z_1] + \mathbb{E}[\sqrt{1-\rho^2} \cdot Z_2] = \rho \mathbb{E}[Z_1] + \sqrt{1-\rho^2} \cdot \mathbb{E}[Z_2] = 0$$

$$\text{Var[Y]} \ = \mathrm{Var}[\rho Z_1 + \sqrt{1-\rho^2} \cdot Z_2] = \mathrm{Var}[\rho Z_1] + \mathrm{Var}[\sqrt{1-\rho^2} \cdot Z_2] = \rho^2 \cdot \mathrm{Var}[Z_1] + (1-\rho^2) \cdot \mathrm{Var}[Z_2] = 1$$

Here we have used the fact that $Z_1 \sim \mathcal{N}(0,1)$ and $Z_2 \sim \mathcal{N}(0,1)$ and are independent.
Finally, we have proved that: $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(0,1)$.

Now let's compute the covariance $\mathrm{Cov}(X,Y)$:

$$\begin{aligned}\text{Cov(X,Y)} &= \mathrm{Cov}(\mathrm{Z}_1, \rho \mathrm{Z}_1 + \sqrt{1-\rho^2} \cdot \mathrm{Z}_2) = \mathrm{Cov}(\mathrm{Z}_1, \rho \mathrm{Z}_1) + \mathrm{Cov}(\mathrm{Z}_1, \sqrt{1-\rho^2} \cdot \mathrm{Z}_2) = \\ &= \rho \cdot \mathrm{Cov}(\mathrm{Z}_1, \mathrm{Z}_1) + (\sqrt{1-\rho^2}) \cdot \mathrm{Cov}(\mathrm{Z}_1, \mathrm{Z}_2) = \rho \mathrm{Var}(\mathrm{Z}_1) = \rho\end{aligned}$$

Again we have used the fact that Z_1 and Z_2 are independent.

So, we can compute:

$\rho_{x,y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}[Z_1]\mathrm{Var}[Z_2]}} = \rho$

as requested.