Pulling Two Correlated Normally-Distributed Random Variates
Part I - The Gaussian Copula and Positive Correlation

Gary Schurman, MBE, CFA

May, 2011

There are many instances especially with Monte Carlo simulations where we want to pull random numbers from normal distributions that are correlated. We will work through the process and develop the attendant mathematics applicable to pulling these correlated random numbers. We will follow this discussion with an example.

The Distributions of the Random Variates X and Y

The random variable X is normally-distributed with a mean equal to μ_x and a variance equal to σ_x^2. The distribution of X in equation form is...

$$X \approx N\left[\mu_x, \sigma_x^2\right] \tag{1}$$

The random variable Y is normally-distributed with a mean equal to μ_y and a variance equal to σ_y^2. The distribution of Y in equation form is...

$$Y \approx N\left[\mu_y, \sigma_y^2\right] \tag{2}$$

We will define ρ_{xy} to be the pairwise correlation between random variates pulled from the distribution of X and the distribution of Y as defined in equations (1) and (2) above. The equation for this correlation coefficient is...

$$\begin{aligned}\rho_{xy} &= \frac{Cov(xy)}{\sigma_x \sigma_y} \\ &= \frac{\mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]}{\sigma_x \sigma_y}\end{aligned} \tag{3}$$

Correlating the Random Variates X and Y

Step One - Pull three independent (i.e. uncorrelated) random variates Z_c, Z_x and Z_y from a normal distribution with mean zero and variance one. The distributions of these random variates and attendant expectations are...

$$Z_c \approx N\left[0, 1\right] \text{ ...where... } \mathbb{E}\left[Z_c\right] = 0 \text{ ...and... } \mathbb{E}\left[Z_c^2\right] = 1 \tag{4}$$

$$Z_x \approx N\left[0, 1\right] \text{ ...where... } \mathbb{E}\left[Z_x\right] = 0 \text{ ...and... } \mathbb{E}\left[Z_x^2\right] = 1 \tag{5}$$

$$Z_y \approx N\left[0, 1\right] \text{ ...where... } \mathbb{E}\left[Z_y\right] = 0 \text{ ...and... } \mathbb{E}\left[Z_y^2\right] = 1 \tag{6}$$

Because these random variates are independent the expectations of their products are...

$$\mathbb{E}\left[Z_c Z_x\right] = 0 \text{ ...and... } \mathbb{E}\left[Z_c Z_y\right] = 0 \text{ ...and... } \mathbb{E}\left[Z_x Z_y\right] = 0 \tag{7}$$

Step Two - Define two new random variates $\tilde{X}$ and $\tilde{Y}$ that are pulled from the distribution of X and the distribution of Y, respectively, and have a positive pairwise correlation equal to ρ_{xy}. These random variates are defined as follows...

$$\tilde{X} = \left[\sqrt{\rho_{xy}}\, Z_c + \sqrt{1 - \rho_{xy}}\, Z_x\right] \sigma_x + \mu_x \tag{8}$$

$$\tilde{Y} = \left[\sqrt{\rho_{xy}}\, Z_c + \sqrt{1 - \rho_{xy}}\, Z_y\right] \sigma_y + \mu_y \tag{9}$$

Proofs

The mean of $\tilde{X}$ is equal to μ_x as defined in equation (1) above. This proof requires the result of the expectation calculated in Appendix equation A.

$$\begin{aligned} mean &= \mathbb{E}\Bigg[\tilde{X}\Bigg] \\ &= \mu_x \end{aligned} \tag{10}$$

The mean of $\tilde{Y}$ is equal to μ_y as defined in equation (2) above. This proof requires the result of the expectation calculated in Appendix equation B.

$$\begin{aligned} mean &= \mathbb{E}\Bigg[\tilde{Y}\Bigg] \\ &= \mu_y \end{aligned} \tag{11}$$

The variance of $\tilde{X}$ is equal to σ_x^2 as defined in equation (1) above. This proof requires the results of the expectations calculated in Appendix equations A and C.

$$\begin{aligned} variance &= \mathbb{E}\Bigg[\tilde{X}^2\Bigg] - \Bigg[\mathbb{E}\Bigg[\tilde{X}\Bigg]\Bigg]^2 \\ &= \sigma_x^2 + \mu_x^2 - \mu_x^2 \\ &= \sigma_x^2 \end{aligned} \tag{12}$$

The variance of $\tilde{Y}$ is equal to σ_y^2 as defined in equation (2) above. This proof requires the results of the expectations calculated in Appendix equations B and D.

$$\begin{aligned} variance &= \mathbb{E}\Bigg[\tilde{Y}^2\Bigg] - \Bigg[\mathbb{E}\Bigg[\tilde{Y}\Bigg]\Bigg]^2 \\ &= \sigma_y^2 + \mu_y^2 - \mu_y^2 \\ &= \sigma_y^2 \end{aligned} \tag{13}$$

The correlation of $\tilde{X}$ and $\tilde{Y}$ is equal to ρ_{xy} as defined in equation (3) above. This proof requires the results of the expectations calculated in Appendix equations A, B and E.

$$\begin{aligned} Correl &= \frac{\mathbb{E}[\tilde{X}\tilde{Y}] - \mathbb{E}[\tilde{X}]\mathbb{E}[\tilde{Y}]}{\sigma_x \sigma_y} \\ &= \frac{\sigma_x \sigma_y \rho_{xy} + \mu_x \mu_y - \mu_x \mu_y}{\sigma_x \sigma_y} \\ &= \rho_{xy} \end{aligned} \tag{14}$$

An Example

Assume the following distribution of the random variates X and Y and the attendant correlation coefficient...

$$X \approx N\Bigg[3, 16\Bigg] \text{ ...and... } Y \approx N\Bigg[5, 36\Bigg] \text{ ...and... } \rho_{xy} = 0.80 \tag{15}$$

Step 1 - Pull three normally-distributed random variates from a distribution with mean zero and variance one which are...

$$Z_c = 1.00 \text{ ...and... } Z_x = 0.50 \text{ ...and... } Z_y = -01.25 \tag{16}$$

Step 2 - Calculate the two correlated random numbers which are...

$$\tilde{X} = \Bigg[\sqrt{0.80} \times 1.00 + \sqrt{1 - 0.80} \times 0.50\Bigg] \times \sqrt{16} + 3.00 = 7.47 \tag{17}$$

$$\tilde{Y} = \Bigg[\sqrt{0.80} \times 1.00 + \sqrt{1 - 0.80} \times -1.25\Bigg] \times \sqrt{36} + 5.00 = 7.01 \tag{18}$$

Appendix

A) The expected value of $\tilde{X}$ is...

$$\begin{aligned}\mathbb{E}\left[\tilde{X}\right] &= \mathbb{E}\left[\left\{\sqrt{\rho_{xy}}\,Z_c + \sqrt{1-\rho_{xy}}\,Z_x\right\}\sigma_x + \mu_x\right] \\ &= \mathbb{E}\left[\sigma_x\sqrt{\rho_{xy}}\,Z_c + \sigma_x\sqrt{1-\rho_{xy}}\,Z_x + \mu_x\right] \\ &= \mathbb{E}\left[\sigma_x\sqrt{\rho_{xy}}\,Z_c\right] + \mathbb{E}\left[\sigma_x\sqrt{1-\rho_{xy}}\,Z_x\right] + \mathbb{E}\left[\mu_x\right] \\ &= \sigma_x\sqrt{\rho_{xy}}\,\mathbb{E}\left[Z_c\right] + \sigma_x\sqrt{1-\rho_{xy}}\,\mathbb{E}\left[Z_x\right] + \mu_x \\ &= \mu_x \end{aligned} \tag{19}$$

B) Using the same equations as A above, the expected value of $\tilde{Y}$ is therefore...

$$\mathbb{E}\left[\tilde{Y}\right] = \mu_y \tag{20}$$

C) The expected value of the square of $\tilde{X}$ is...

$$\begin{aligned}\mathbb{E}\left[\tilde{X}^2\right] &= \mathbb{E}\left[\left\{\sigma_x\sqrt{\rho_{xy}}\,Z_c + \sigma_x\sqrt{1-\rho_{xy}}\,Z_x + \mu_x\right\}^2\right] \\ &= \mathbb{E}\left[\sigma_x^2\rho_{xy}Z_c^2 + \sigma_x^2(1-\rho_{xy})Z_x^2 + 2\sigma_x^2\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}Z_cZ_x + 2\mu_x\sigma_x\sqrt{\rho_{xy}}Z_c + 2\mu_x\sigma_x\sqrt{1-\rho_{xy}}Z_x + \mu_x^2\right] \\ &= \mathbb{E}\left[\sigma_x^2\rho_{xy}Z_c^2\right] + \mathbb{E}\left[\sigma_x^2(1-\rho_{xy})Z_x^2\right] + \mathbb{E}\left[2\sigma_x^2\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}Z_cZ_x\right] + \mathbb{E}\left[2\mu_x\sigma_x\sqrt{\rho_{xy}}Z_c\right] \\ &+ \mathbb{E}\left[2\mu_x\sigma_x\sqrt{1-\rho_{xy}}Z_x\right] + \mathbb{E}\left[\mu_x^2\right] \\ &= \sigma_x^2\rho_{xy}\,\mathbb{E}\left[Z_c^2\right] + \sigma_x^2(1-\rho_{xy})\,\mathbb{E}\left[Z_x^2\right] + 2\sigma_x^2\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,\mathbb{E}\left[Z_cZ_x\right] + 2\mu_x\sigma_x\sqrt{\rho_{xy}}\,\mathbb{E}\left[Z_c\right] \\ &+ 2\mu_x\sigma_x\sqrt{1-\rho_{xy}}\,\mathbb{E}\left[Z_x\right] + \mathbb{E}\left[\mu_x^2\right] \\ &= \sigma_x^2\rho_{xy} + \sigma_x^2(1-\rho_{xy}) + \mu_x^2 \\ &= \sigma_x^2 + \mu_x^2 \end{aligned} \tag{21}$$

D) Using the same equations as C above, the expected value of the square of $\tilde{Y}$ is therefore...

$$\mathbb{E}\left[\tilde{Y}^2\right] = \sigma_y^2 + \mu_y^2 \tag{22}$$

E) The expected value of the product of $\tilde{X}$ and $\tilde{Y}$ is...

$$\begin{aligned}
\mathbb{E}\Big[\tilde{X}\tilde{Y}\Big] &= \mathbb{E}\Big[\Big\{\sigma_x\sqrt{\rho_{xy}}\,Z_c + \sigma_x\sqrt{1-\rho_{xy}}\,Z_x + \mu_x\Big\}\Big\{\sigma_y\sqrt{\rho_{xy}}\,Z_c + \sigma_y\sqrt{1-\rho_{xy}}\,Z_y + \mu_y\Big\}\Big] \\
&= \mathbb{E}\Big[\sigma_x\sigma_y\rho_{xy}\,Z_c^2 + \sigma_x\sigma_y\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,Z_cZ_y + \mu_y\sigma_x\sqrt{\rho_{xy}}\,Z_c + \sigma_x\sigma_y\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,Z_cZ_x \\
&+ \sigma_x\sigma_y(1-\rho_{xy})\,Z_xZ_y + \mu_y\sigma_x\sqrt{1-\rho_{xy}}\,Z_x + \mu_x\sigma_y\sqrt{\rho_{xy}}\,Z_c + \mu_x\sigma_y\sqrt{1-\rho_{xy}}\,Z_c + \mu_x\mu_y\Big] \\
&= \mathbb{E}\Big[\sigma_x\sigma_y\rho_{xy}\,Z_c^2\Big] + \mathbb{E}\Big[\sigma_x\sigma_y\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,Z_cZ_y\Big] + \mathbb{E}\Big[\mu_y\sigma_x\sqrt{\rho_{xy}}\,Z_c\Big] + \mathbb{E}\Big[\sigma_x\sigma_y\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,Z_cZ_x\Big] \\
&+ \mathbb{E}\Big[\sigma_x\sigma_y(1-\rho_{xy})\,Z_xZ_y\Big] + \mathbb{E}\Big[\mu_y\sigma_x\sqrt{1-\rho_{xy}}\,Z_x\Big] + \mathbb{E}\Big[\mu_x\sigma_y\sqrt{\rho_{xy}}\,Z_c\Big] + \mathbb{E}\Big[\mu_x\sigma_y\sqrt{1-\rho_{xy}}\,Z_c\Big] + \mathbb{E}\Big[\mu_x\mu_y\Big] \\
&= \sigma_x\sigma_y\rho_{xy}\,\mathbb{E}\Big[Z_c^2\Big] + \sigma_x\sigma_y\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,\mathbb{E}\Big[Z_cZ_y\Big] + \mu_y\sigma_x\sqrt{\rho_{xy}}\,\mathbb{E}\Big[Z_c\Big] + \sigma_x\sigma_y\sqrt{\rho_{xy}}\sqrt{1-\rho_{xy}}\,\mathbb{E}\Big[Z_cZ_x\Big] \\
&+ \sigma_x\sigma_y(1-\rho_{xy})\,\mathbb{E}\Big[Z_xZ_y\Big] + \mu_y\sigma_x\sqrt{1-\rho_{xy}}\,\mathbb{E}\Big[Z_x\Big] + \mu_x\sigma_y\sqrt{\rho_{xy}}\,\mathbb{E}\Big[Z_c\Big] + \mu_x\sigma_y\sqrt{1-\rho_{xy}}\,\mathbb{E}\Big[Z_c\Big] + \mathbb{E}\Big[\mu_x\mu_y\Big] \\
&= \sigma_x\sigma_y\rho_{xy} + \mu_x\mu_y \qquad (23)
\end{aligned}$$