A Method for Directly Generating a Gaussian Distribution with Nonunit Variance and Nonzero Mean from Uniform Random Deviates

Marc A. Murison

Astronomical Applications Department
U.S. Naval Observatory
Washington, DC

murison@aa.usno.navy.mil
http://aa.usno.navy.mil/murison/

8 May, 2000

1. Introduction

It is universally the case that, when an algorithm is presented that takes uniform random deviates and produces a Gaussian distribution, the Gaussian distribution has zero mean and unit variance: $G(x) = \frac{e^{\left(-\frac{x^2}{2}\right)}}{\sqrt{2\pi}}$. For nonzero mean μ and nonunit variance σ^2, the Gaussian distribution is $G(x, \mu, \sigma) = \frac{e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}}{\sigma\sqrt{2\pi}}$, with normalization $\int_{-\infty}^{\infty} G(x, \mu, \sigma)\, dx = 1$. How does one obtain a distribution with a given variance from a distribution with unit variance? The Gaussian function has a useful invariant relation: $G(x, \mu, \sigma) = \frac{G(y, 0, 1)}{\sigma}$, where $y = \frac{x-\mu}{\sigma}$. Hence, if we have a set of random deviates y which are distributed as a Gaussian with unit variance and zero mean, we may convert those deviates to a Gaussian distribution with nonzero mean and nonunit variance by the simple trick of multiplying by σ and adding μ. However, one might still wish for an algorithm that takes uniform deviates and *directly* produces the desired nonunit variance Gaussian distribution.

2. A Modified Box-Muller Algorithm

The usual algorithm for producing a Gaussian distribution from uniform random deviates makes use of the Box-Muller transformation. We will modify the usual algorithm so that the Box-Mueller transformation produces the desired nonunit-variance Gaussian distribution. If z_1 and z_2 are independent and uniformly distributed in the range 0 to 1, then consider y_1 and y_2, given by

$$y_1 = \sqrt{-2\ln(z_1)}\cos(2\pi z_2)$$

$$y_2 = \sqrt{-2\ln(z_1)}\sin(2\pi z_2)$$

This is the Box-Muller transformation, which converts a 2D uniform distribution (z_1, z_2) to a bivariate Gaussian distribution (y_1, y_2) with zero mean and unit variance. (The proof that this transformation yields the desired Gaussian distribution is a special case of the proof of the modified transformation, which is given below.)

The trick to making the Box-Muller transformation produce a Gaussian distribution (x_1, x_2) with nonunit variance is to use the invariant relation mentioned above. Setting

$$x = y\sigma + \mu$$

we have

$$x_1 = \sqrt{-2\ln(z_1)}\cos(2\pi z_2)\sigma + \mu$$

$$x_2 = \sqrt{-2\ln(z_1)}\sin(2\pi z_2)\sigma + \mu$$

2.1. Verification

To show that this is indeed the desired distribution, we have

$$\sqrt{\frac{(x_1-\mu)^2}{\sigma^2} + \frac{(x_2-\mu)^2}{\sigma^2}} = \sqrt{2}\sqrt{-\ln(z_1)}$$

which we solve for z_1,

$$z_1 = \mathbf{e}^{\left(-1/2\frac{(x_2-\mu)^2}{\sigma^2}\right)}\mathbf{e}^{\left(-1/2\frac{(x_1-\mu)^2}{\sigma^2}\right)}$$

and we have

$$\frac{x_2}{x_1} = \frac{\sqrt{-2\ln(z_1)}\sin(2\pi z_2)\sigma + \mu}{\sqrt{-2\ln(z_1)}\cos(2\pi z_2)\sigma + \mu}$$

which we solve for z_2,

$$z_2 = -\frac{1}{2}\arctan\Bigg(\Bigg(\Big(\sqrt{-\ln(z_1)}\,x_1\,(-x_1+x_2)\,\mu - \sqrt{\ln(z_1)\,x_2^2\,(-x_1+x_2)^2\,\mu^2 + 2\ln(z_1)^2\,x_2^2\left(x_2^2+x_1^2\right)\sigma^2}\Big)x_2\Bigg) \Big/ \Big(\sqrt{-\ln(z_1)}\,x_2^2\,(-x_1+x_2)\,\mu + x_1\sqrt{\ln(z_1)\,x_2^2\,(-x_1+x_2)^2\,\mu^2 + 2\ln(z_1)^2\,x_2^2\left(x_2^2+x_1^2\right)\sigma^2}\Big)\Bigg)\Big/\pi$$

Substituting for z_1 and simplifying, we find

$$z_2 = \frac{1}{2}\frac{\arctan\left(\frac{-x_2+\mu}{-x_1+\mu}\right)}{\pi}$$

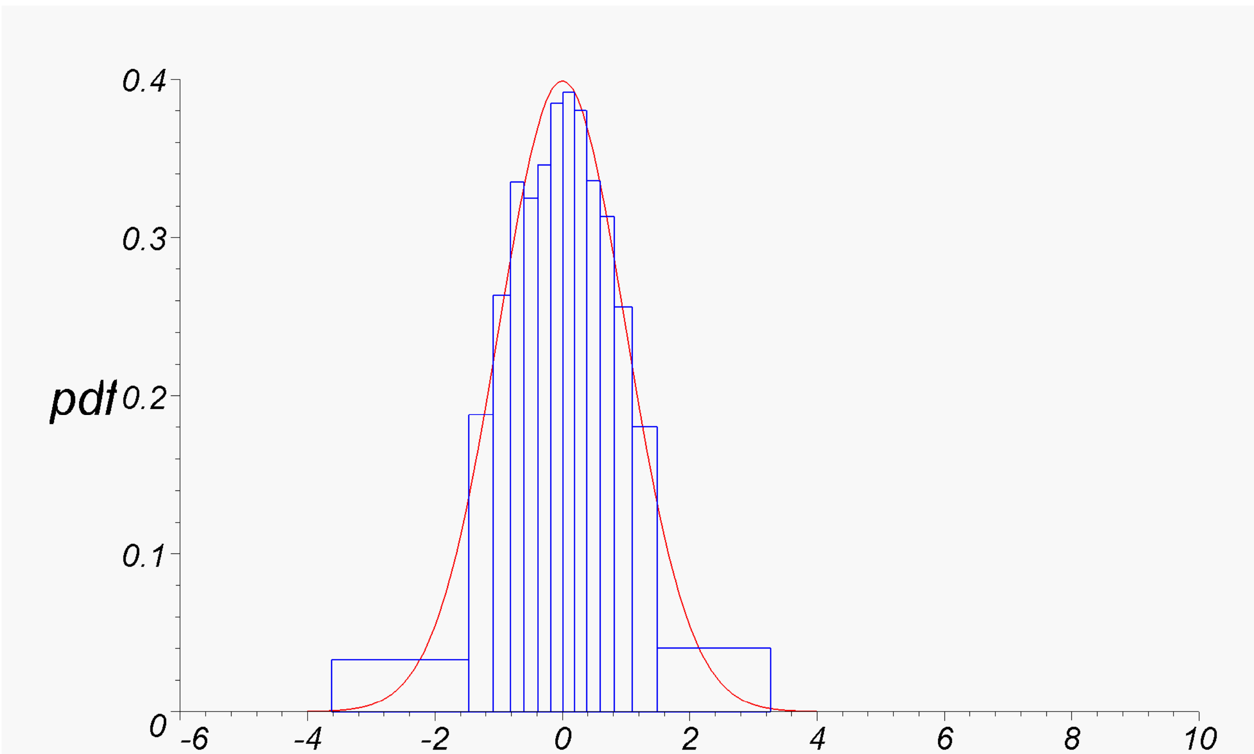
Now, the fundamental transformation law of probabilities states that $|\mathrm{p}(y)\,dy| = |\mathrm{p}(x)\,dx|$, or $\mathrm{p}(y) = \mathrm{p}(x)\left|\frac{dx}{dy}\right|$, where $\mathrm{p}(x)\,dx$ is the probability that x lies between x and $x + dx$, and $\mathrm{p}(y)\,dy$ is the probability that y lies between y and $y + dy$. The extension of this to n dimensions is $\mathrm{p}(Y)\,dY = \mathrm{p}(X)\,\mathrm{J}(X, Y)\,dY$, where $\mathrm{p}(X)$ is the joint probability distribution of X, $\mathrm{p}(Y)$ is the joint probability distribution of Y, $X = (x_1, x_2 .. x_n)$, $Y = (y_1, y_2 .. y_n)$, $dY = dy_1\,dy_2 .. dy_n$, and $\mathrm{J}(X, Y)$ is the Jacobi determinant. For two dimensions, $\mathrm{J}(X, Y) = \det\left(\begin{bmatrix} \frac{\partial}{\partial y_1}x_1 & \frac{\partial}{\partial y_2}x_1 \\ \frac{\partial}{\partial y_1}x_2 & \frac{\partial}{\partial y_2}x_2 \end{bmatrix}\right)$. Hence,

$$J(Z, X) = -\frac{1}{2} \frac{e^{\left(-1/2 \frac{(-x_1+\mu)^2}{\sigma^2}\right)} e^{\left(-1/2 \frac{(-x_2+\mu)^2}{\sigma^2}\right)}}{\pi \sigma^2}$$

This says that for a uniform distribution (z_1, z_2) the distribution of (x_1, x_2) is a symmetric bivariate Gaussian distribution with mean μ and variance σ^2. *Q.E.D.*

3. Numerical Tests

In the plot below, the red curve is a plot of $G(x, 0, 1)$, while the histogram is the result of running 5000 uniform random deviates through the modified Box-Muller transformation just described with $\mu = 0$ and $\sigma = 1$.

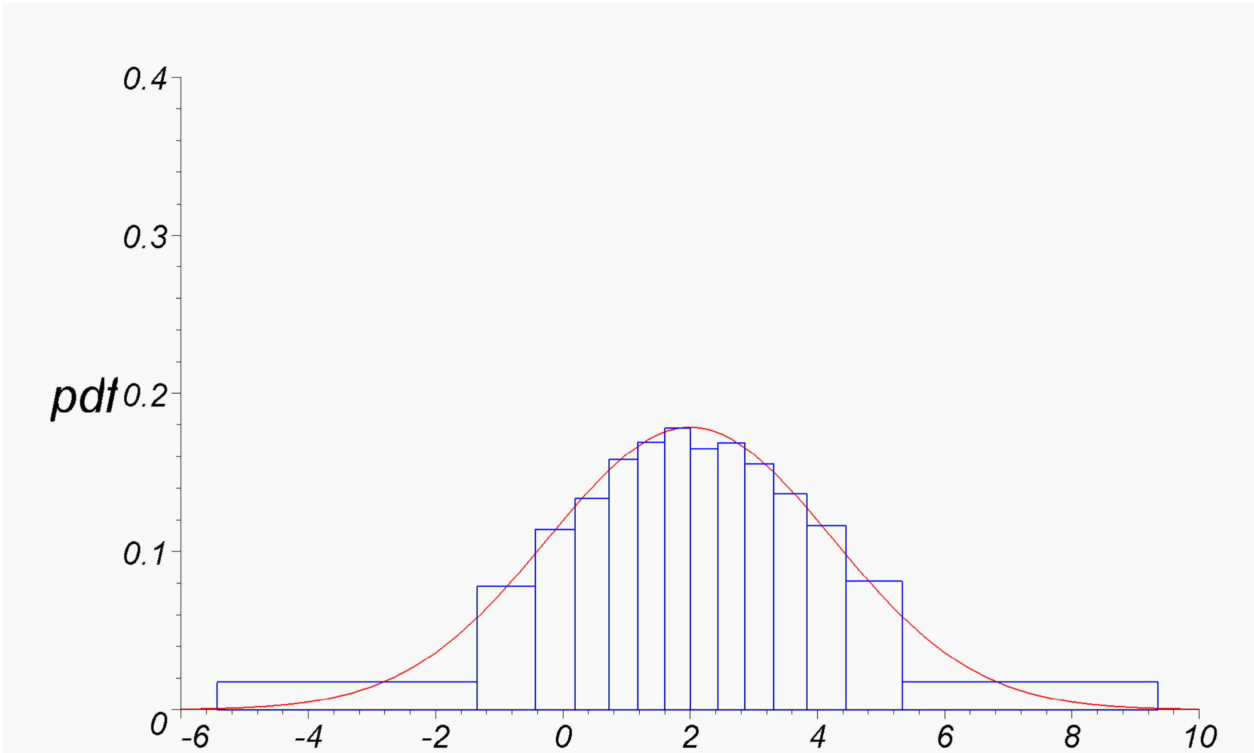


The peak of the curve should be

$$\frac{1}{\sqrt{2\pi}} = .3989422802$$

This shows that the modified numerical algorithm correctly reproduces the standard algorithm in the appropriate limits.

Now for a test of the modified algorithm. In the next plot, the red curve is a plot of $G(x, 2, \sqrt{5})$, while the histogram is the result of running 5000 uniform random deviates through the modified Box-Muller transformation with $\mu = 2$ and $\sigma = \sqrt{5}$.



The peak of the curve should be

$$\frac{1}{\sqrt{5}\,\sqrt{2\,\pi}} = .1784124116$$

This shows that the modified algorithm is indeed generating the correct Gaussian distribution with nonunit variance and nonzero mean.