Monte Carlo Simulations

Florida International University
Department of Physics

Nestor Viana

September 2021

Abstract

Several Monte Carlo techniques are performed. These are used to determine the volume of the unit-radius n-dimensional sphere using acceptance/rejection method and to compare with values from a analytically derived formula of the n-dimensional sphere for $n = 2,3,4,5$. We estimate the value of the mathematical constant $\pi = 3.142 \pm 0.001$ using the $n = 2$ case to two decimal places of accuracy. Lastly, we build random number generators drawn from common Gaussian and Poisson probability distribution functions using the Box-Muller and transformation methods, respectively, and obtain accurate results through non-linear least-squares fitting and χ^2 goodness-of-fit statistics.

1 Introduction

During World War II, a group of scientists and engineers at Alamogordo worked on building the first working electronic computer−the ENIAC. Almost upon completion, John Von Neumann, mathematics professor at the Institute for Advanced Study, sparked interest in using ENIAC for thermonuclear reactions, where him and other colleagues had a taste of its computational power. In 1946 in Los Alamos, a thorough review of this machine was presented, and among the audience stood Stan Ulam, professor of mathematics at the University of Southern California. Ulam had been weary that statistical tools for analysis were falling behind, so he discussed ideas for ENIAC with Von Neumann to prevent this field from becoming stagnant, and hence Monte Carlo emerged [3]. Monte Carlo is a method built on the use of random numbers primarily utilized in computation of multiple integrals to simulate experimental calculations and estimate possible outcomes of theoretical models in fields such as physics, engineering, finance, risk analysis, forecasting, etc. A Monte Carlo simulation carries out a process (lab experiment) a thousand, million, or billion times, giving information about the distribution of parameters and variables and their uncertainty. This method is particularly useful when lab conditions aren't suitable for performing experiments that are difficult or virtually impossible to replicate usually due to lack of resources or other means, or for predicting behaviors of variables that may be otherwise unethical to test and measure in practice.

The first part of this paper focuses in developing computational Monte Carlo techniques to obtain the volume of any n-dimensional sphere, thereby granting us access to measure the value of π to virtually any degree of accuracy desired as long as it's within reasonable computational margin through the acceptance/rejection method. This method helps determine the probability density function of a variable given an *auxiliary* random variable and its probability density function. The second half of this paper describes and utilizes the transformation methods used to generate random numerical deviates from Poisson and Gaussian distribution which implement in essence the same auxiliary mechanism of the acceptance/rejection method.

2 Estimating π

2.1 Method

Consider a circle of radius 1 inscribed in a square of sidelength 2 centered at the origin (Figure 1). Let

$(x,y) \in [-1,1]^2$ be a random point inside the square that may or may not be in the circle. The probability, p, that (x,y) is inside the circle is proportional to the ratio of the area of the circle, A_c, to the area of the square, A_s. If N_s is the total number of points that randomly lie in the square, then the (expected) number of points that also lie in the circle is

$$N_c = N_s\, p = N_s \frac{A_c}{A_s}. \tag{1}$$

In practice, the area of the circle is unknown as is the value of π. The Monte Carlo procedure is as follows: obtain N_s total points $\{(x_i,y_i)\}_{i=1}^{N_s}$, and if for any point (x_i,y_i) the condition

$$x_i^2 + y_i^2 \leq 1 \tag{2}$$

is true then (x_i,y_i) lies inside the circle, and the value of N_c increases by 1—this point is "accepted", otherwise it is "rejected" and N_c remains constant. Thus, the value of π is given by

$$A_s = \pi r^2 = A_s \frac{N_c}{N_s} \tag{3}$$

$$\pi = 2^2 \frac{N_c}{N_s} \tag{4}$$

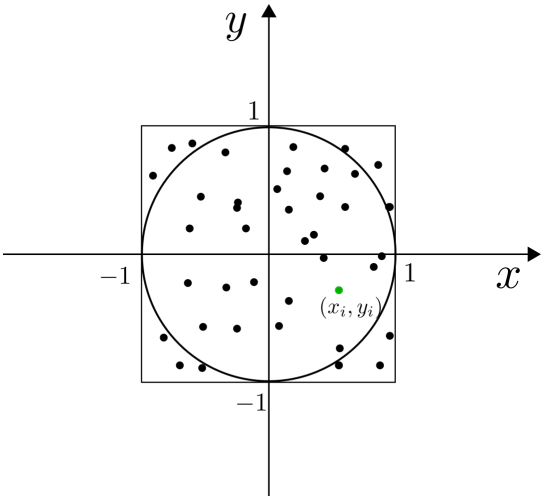

Figure 1: Sketch of the Monte Carlo process.

Values for the points $\{(x,y)\}$ will be drawn from a uniform random number generator on the interval [0,1]. (We need not extend the interval to [-1,1] as both x and y shall be squared).

2.2 Error Analysis

The number of points landing inside the circle N_c is a discrete random variable. As described, N_c models the number of "successes" of this Monte Carlo simulation, therefore it follows a Binomial Distribution with mean

$$\mu_{N_c} = N_s\, p \tag{5}$$

and variance

$$\sigma_{N_c}^2 = N_s\, p(1-p). \tag{6}$$

Thus, from (4), the expectation value and standard deviation of our estimates are

$$\mu_\pi = 4\frac{N_c}{N_s} \tag{7}$$

$$\sigma_\pi = \frac{4}{N_s}\sqrt{N_c\left(1-\frac{N_c}{N_s}\right)} \tag{8}$$

From (8), we can find the number of points N_s needed to achieve any desired relative uncertainty, δ_π, in our estimation of π. Using $p = \frac{\pi}{4}$ in (7),

$$\delta_\pi = \frac{\sigma_\pi}{\pi} = \frac{4}{\pi N_s}\sqrt{N_c\left(1-\frac{N_c}{N_s}\right)} \tag{9}$$

$$= \frac{4}{\pi N_s}\sqrt{N_s\frac{\pi}{4}\left(1-\frac{\pi}{4}\right)} \tag{10}$$

leads to

$$N_s = \left(\frac{4}{\pi}-1\right)\frac{1}{\delta_\pi^2} \approx \frac{0.273}{\delta_\pi^2} \tag{11}$$

This means that for every one more wanted decimal point of relative uncertainty, a factor of 10^2 more sample points is required.

If this Monte Carlo simulation is carried out M times with N_s sample points each, a number of M estimates for π will be produced. A convenient way to report these results, specially when M is large, is a histogram plot of the estimate distribution μ_π. In contrast to the variable N_c, the distribution of μ_π estimates behaves like a Normal Distribution as this variable is continuous and not discrete, and since its parent distribution N_c is a Binomial one which converges to a Normal distribution as $M \to \infty$ by the Central Limit Theorem [2]. It will be appropriate then to fit a Gaussian curve

$$G(x;\hat{\mu},\hat{\sigma},\hat{A}) \sim \hat{A}\exp\left[-\frac{(x-\hat{\mu})^2}{2\hat{\sigma}^2}\right] \tag{12}$$

to the μ_π histogram plot (Figure 2).

2.2.1 Reduced χ^2 Goodness-of-Fit Test

The absolute estimate of π and error reported will be the fit parameters $\hat{\mu}$ and $\hat{\sigma}$ of G using a least-squares non-linear fit. To quantify the performance of this fit, a reduced χ^2 goodness-of-fit test is carried out. Let H_i be the content (frequency) of each histogram bin of the μ_π distribution plot with error $\sigma_i = \sqrt{H_i}$ and x_i be the bin centers. The test statistic

$$\chi^2_\nu = \frac{1}{\nu}\sum_{i=1}^{M}\frac{\left[G(x_i;\hat{\mu},\hat{\sigma},\hat{A}) - H_i\right]^2}{\sigma_i^2} \tag{13}$$

is the ratio of the variance between the histogram data and the fitted curve to the variance of each histogram bin, where ν is the number of degrees of freedom. (In this case, ν = #bins - #parameters). One should see that if χ^2_ν is too small or too big, then the variances between these two quantities at each x_i do not correlate and therefore the fit is not satisfactory. Thus, a value of $\chi^2_\nu \approx 1$ is most desired for good fit.

2.3 Results

We performed $M = 1{,}000$ simulations with $N_s = 10{,}000$ sample points each. A histogram of the results is plotted below together with a Gaussian fit described in (12). The fitted parameters and their uncertainties are reported as well.

$$\begin{aligned}\hat{\mu} &= 3.1418 \pm 6\times10^{-4}\\ \hat{\sigma} &= 0.0170 \pm 6\times10^{-4}\\ \hat{A} &= 64.20 \pm 1.88\\ \chi^2_\nu &= 0.862\end{aligned}$$

The parameter $\hat{\mu}$ is our estimate for the constant π:

$$\pi = 3.142 \pm 0.001$$

The test statistic χ^2_ν being quite close to unity tells us our Gaussian fit resembles the histogram data quite well, which is also evidenced by a close π value.

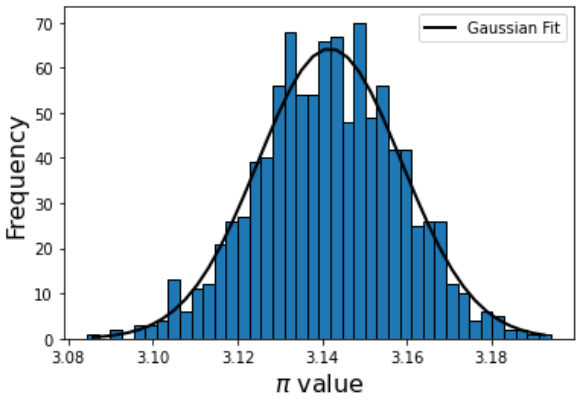

Figure 2: Histogram plot and Gaussian fit of the distribution of the estimate of π, μ_π.

3 Volume of an n-Dimensional Sphere

3.1 Derivation

Proposition: The measure (volume) of an n-dimensional sphere of radius R, $|V_n(R)|$, is proportional to the measure of the unit n-dimensional sphere by the relation

$$|V_n(R)| = R^n |V_n(1)| \tag{14}$$

Proof: The measure of an n-dimensional sphere of radius R is given by

$$|V_n(R)| = \int_{V_n(R)} dx_1 dx_2 \cdots dx_n \tag{15}$$

Make the change of variables $y_i = x_i/R$ to reduce the radius of the n-sphere to 1.

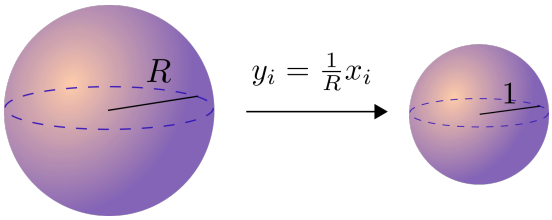

Figure 3: Change of variable transformation.

The Jacobian of this transformation is

$$\left|\frac{\partial(x_1,\ldots,x_n)}{\partial(y_1,\ldots,y_n)}\right| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} \tag{16}$$

$$= \begin{vmatrix} R & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R \end{vmatrix} = R^n \tag{17}$$

Hence (15) is equivalent to

$$|V_n(R)| = \int_{V_n(1)} R^n \, dy_1 dy_2 \cdots dy_n \tag{18}$$

$$= R^n \int_{V_n(1)} dy_1 dy_2 \cdots dy_n \tag{19}$$

$$= R^n |V_n(1)| . \tag{20}$$

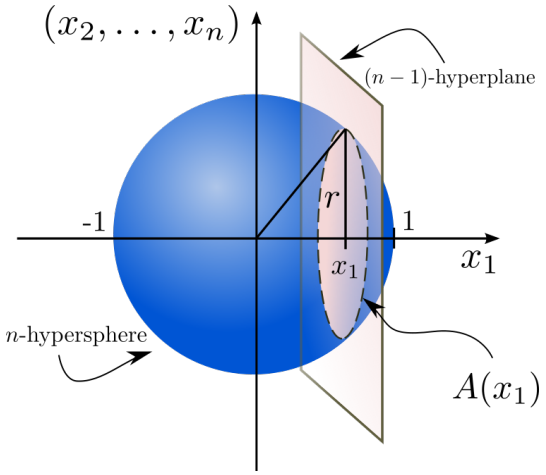

Figure 4: Hypersphere cross-section. The variable x_1 is plotted against in one axis for simplicity; the vertical axis represents the axes of the rest of the variables. $A(x_1)$ is the $(n-1)$-hypersphere that results from the intersection.

Theorem: The measure of an n-sphere of radius R is

$$|V_n(R)| = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} R^n \quad \text{for} \quad n \geq 2. \tag{21}$$

Before the proof of the theorem is presented, note that hyperplanes passing through the n-sphere are of dimension $n-1$. In particular, these hyperplane cross-sections are $(n-1)$-hyperspheres (Figure 4), just like in 3 dimensions, planes crossing spheres form 2-dimensional circles, and how lines intersecting a 2-dimensional circle form 1-dimensional intervals.

Proof: By the disk method, the volume of the unit n-hypersphere centered at the origin is the addition of the quantity $A(x_1)$ over $-1 \leq x_1 \leq 1$. Moreover, $A(x_1)$ is the measure of the $(n-1)$-hypersphere at any x_1 value.

$$|V_n(1)| = \int_{-1}^{1} A(x_1)\, dx_1 \tag{22}$$

$$= \int_{-1}^{1} |V_{n-1}(r)|\, dx_1 \tag{23}$$

$$= \int_{-1}^{1} \left| V_{n-1}\left(\sqrt{1-x_1^2}\right)\right| dx_1 \tag{24}$$

$$\underset{\text{proposition}}{\text{by}} = \int_{-1}^{1} \left(\sqrt{1-x_1^2}\right)^{n-1} |V_{n-1}(1)|\, dx_1 \tag{25}$$

$$= 2|V_{n-1}(1)| \int_0^1 (1-x^2)^{\frac{n-1}{2}}\, dx \tag{26}$$

Let $u = x^2$ so that

$$|V_n(1)| = 2|V_{n-1}(1)| \int_0^1 (1-u)^{\frac{n-1}{2}} u^{-\frac{1}{2}}\, du \tag{27}$$

$$= |V_{n-1}(1)| B\left(\frac{1}{2}, \frac{n-1}{2}\right) \tag{28}$$

where B is the Beta function

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{29}$$

and Γ is the Gamma function. Substituting this recursive relation into (28) itself gives

$$|V_n(1)| = |V_{n-1}(1)| \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} \tag{30}$$

$$= \sqrt{\pi}|V_{n-1}(1)| \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} \tag{31}$$

$$= \sqrt{\pi} \cancel{\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)}} \sqrt{\pi} \cancel{\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}} \cdots \sqrt{\pi} \cancel{\frac{\Gamma(1)}{\Gamma\left(\frac{3}{2}\right)}} \tag{32}$$

$$= \left(\sqrt{\pi}\right)^n \frac{\Gamma(1)}{\Gamma\left(\frac{n+2}{2}\right)} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}. \tag{33}$$

From that last result and the previous proposition,

$$|V_n(R)| = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} R^n . \tag{34}$$

3.2 Monte Carlo Method

A similar Monte Carlo procedure is implemented as in the previous section; instead, we check the slightly different acceptance condition

$$x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1 \tag{35}$$

is satisfied. That is, for any point $(x_1, x_2, ..., x_n) \in [-1,1]^n$ generated randomly that satisfies Eq. 7 will lie inside the n-dimensional hypersphere inscribed in the n-dimensional hypercube of sidelength 2 centered at the origin. For each of these success points, the variable N_c in (36) increases by 1. Values for the points $\{(x_1, x_2, ..., x_n)\}$ are drawn from a uniform random number generator on the interval [0,1] for a total of N_s points.

$$|V_n(1)| = 2^n \frac{N_c}{N_s} \tag{36}$$

3.3 Error Analysis

In this section we use a distinct approach to determine the measure of $V_n(1)$ and its uncertainty. For each run of this process, the standard deviation in $|V_n(1)|_i$ takes the form

$$\sigma_i = \frac{2^n}{N_s} \sqrt{N_c \left(1 - \frac{N_c}{N_s}\right)} \tag{37}$$

so that for a total of M runs, we have M estimates of $|V_n(1)|$. For large M, it's more convenient to summarize results with just a few statistics. Define the weighted mean and standard deviation of weighted mean respectively as

$$\mu = \frac{\sum \left(|V_n(1)|_i / \sigma_i^2 \right)}{\sum 1/\sigma_i^2} \tag{38}$$

$$\sigma_\mu^2 = \sum 1/\sigma_i^2 \tag{39}$$

for each value of n.

3.4 Results

For each $n = 2, \ldots, 5$ we perform $M = 100$ trials with $N_s = 10,000$ points each. Results of (38) and (39) are summarized in the table below together with the theoretical value of $|V_n(1)|$ from (34).

n	$\mu \pm \sigma_\mu$	$\lvert V_n(1)\rvert$
2	3.140 ± 0.001	3.1415
3	4.188 ± 0.004	4.1887
4	4.927 ± 0.007	4.9348
5	5.24 ± 0.01	5.2637

Table 1: n-dimensional unit sphere volumes. The middle column are the results of the Monte Carlo simulation. Right column are the theoretical values.

4 Poisson Deviates

4.1 General Tranformation Method

Numerical deviates are random numbers drawn from a particular probability distribution function $P(x)$. It is usual to start with a uniform probability distribution function $u(r)$ over the interval [0,1] and perform a transformation that allows us to obtain random numbers from $P(x)$. [1]

$$u(r) = \begin{cases} 1 & \text{for} \quad 0 \leq r < 1 \\ 0 & \text{everywhere else} \end{cases} \tag{40}$$

We would like to find a relation between r and x such that the probability measure of $u(r)$ and $P(x)$ are equal over a small region:

$$|u(r)\,\Delta r| = |P(x)\,\Delta x| \tag{41}$$

This is also known as conservation of probability. Since both u and P are nonnegative over all probability space we may remove the absolute values, and as both Δr and $\Delta x \to 0$,

$$\int_{-\infty}^{r} u(r)\,dr = \int_{-\infty}^{x} P(x)\,dx \tag{42}$$

$$r = \int_{-\infty}^{x} P(x)\,dx \tag{43}$$

Thus the corresponding deviate from the probability distribution P is the upper bound of (43) that satisfies such equation. For a Poisson distribution function with mean μ, we obtain deviates by finding integers n such that for any random deviate r from the distribution $u(r)$ the equation

$$r \leq \sum_{k=0}^{n} \frac{\mu^k}{k!} e^{-\mu} \tag{44}$$

is satisfied.

4.2 Results

Below we generate $N = 1,000$ Poisson deviates using the process described in section 4.1 for each value of $\mu = 1, 10.3, 102.1$ and make histogram plots of the results.

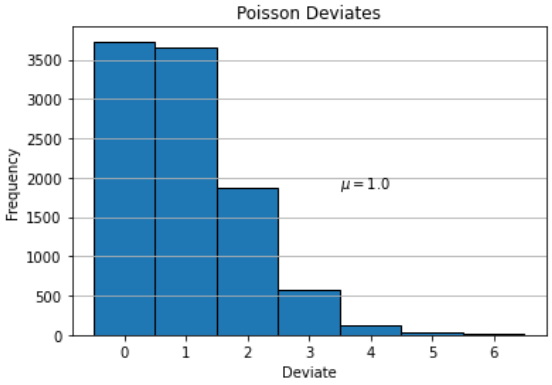

Figure 5: Poisson Monte Carlo deviates for $\mu = 1$

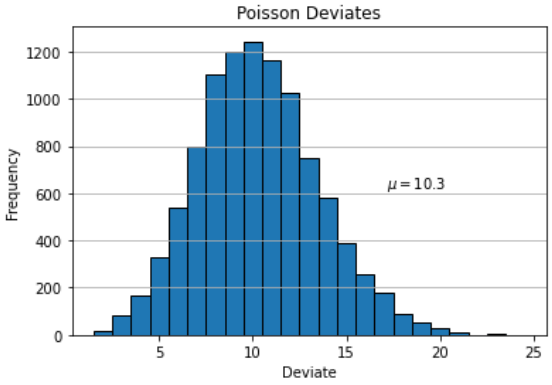

Figure 6: Poisson Monte Carlo deviates for $\mu = 10.3$; distribution behaves noticeably normal.

Note how the Central Limit Theorem begins to dictate the behavior of the distributions as μ gets larger. As the mean of the distributions strays away from 1, more deviates can generate to the left side of the histograms, making them adopt a Normal distribution shape.

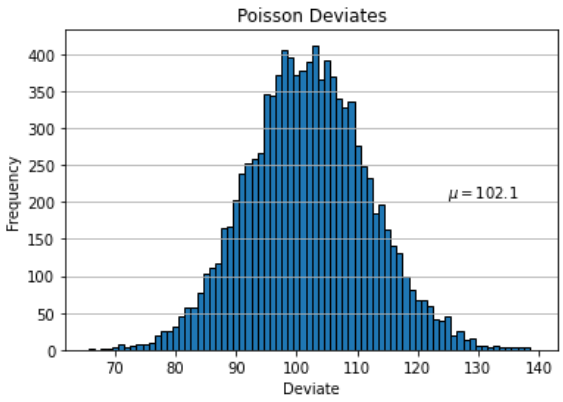

Figure 7: Poisson Monte Carlo deviates for $\mu = 10.3$; distribution is essentially Gaussian.

5 Gaussian Deviates

5.1 Box-Muller Transformation

We reference the same method used in section 4.1 and (43). In this case, the probability distribution function is a Gaussian function with mean of 0 standard deviation of 1.

$$P(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{45}$$

However, (45) has no antiderivative in terms of elementary functions, so solving for the upper bound in (43) is tricky. The Box-Muller transformation method outlined in [4] solves this issue. Let r_1 and r_2 be independently generated random numbers drawn from the unit square $(0,1) \times (0,1)$, meaning that any joint probability density function $f_{R_1,R_2}(r_1, r_2)$ of these two variables is equal to 1 over the unit square. Define

$$z_1 = \sqrt{-2 \ln r_1} \cos(2\pi r_2) \tag{46}$$

$$z_2 = \sqrt{-2 \ln r_1} \sin(2\pi r_2) \tag{47}$$

or equivalently,

$$r_1 = \exp\left[-\frac{1}{2}(z_1^2 + z_2^2)\right] \tag{48}$$

$$r_2 = \frac{1}{2\pi} \arctan\left(\frac{z_2}{z_1}\right) \tag{49}$$

We will show that the joint probability density function $f_{Z_1,Z_2}(z_1, z_2)$ is equal to the product of two functions $f_{Z_1}(z_1) f_{Z_2}(z_2)$ and that each f_{Z_1} and f_{Z_2} are Gaussian

of mean 0 and variance 1. Using a change of variables,

$$f_{Z_1,Z_2}(z_1,z_2) = f_{R_1,R_2}(z_1,z_2)\left|\frac{\partial(r_1,r_2)}{\partial(z_1,z_2)}\right| \tag{50}$$

where

$$\begin{aligned}\left|\frac{\partial(r_1,r_2)}{\partial(z_1,z_2)}\right| &= \begin{vmatrix}\frac{\partial r_1}{\partial z_1} & \frac{\partial r_2}{\partial z_1}\\ \frac{\partial r_1}{\partial z_2} & \frac{\partial r_2}{\partial z_2}\end{vmatrix}\\ &= \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{z_1^2}{2}}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{z_2^2}{2}}\right)\end{aligned}$$

Since $f_{R_1,R_2}(z_1,z_2) = 1$, it follows that the joint probability distribution function of z_1 and z_2 is a product of two independent functions, and each of these functions is a Gaussian. That is, z_1 and z_2 are Gaussian deviates.

5.2 Results

Results for the method discussed in previous section are presented in Figure 8. Additionally, we fit a Gaussian curve $G(x;\hat{\mu},\hat{\sigma},\hat{A})$ to these data, which come from (45) so the expected values for the mean and variance are $(\mu,\sigma) = (0,1)$. A similar method for fitting described in Section 2 is utilized, including the reduced χ^2 goodness-of-fit analysis.

$$\begin{aligned}\hat{\mu} &= 0.0074 \pm 0.0003\\ \hat{\sigma} &= 0.099 \pm 0.002\\ \hat{A} &= 1590 \pm 10\\ \chi^2_\nu &= 0.993\end{aligned}$$

The test statistic χ^2_ν being extremely close to 1 indicates that our Gaussian fit resembles the histogram data with little difference in variation between G and histogram data values, which is in addition suggested by high accuracy of the fitted parameters $\hat{\mu}$ and $\hat{\sigma}$ to the expected quantities, respectively.

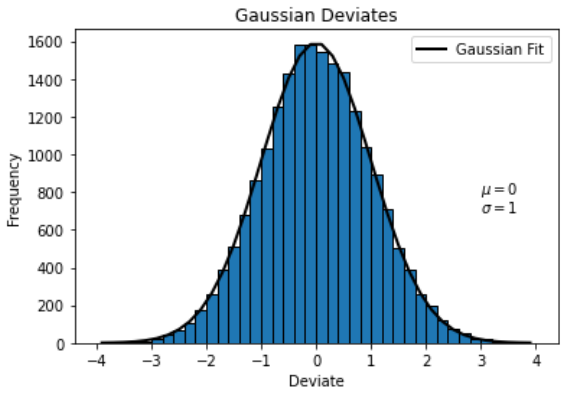

Figure 8: Gaussian deviates from Box-Muller transformation, together fitted with a Normal distribution curve.

References

[1] Philip R Bevington et al. "Data reduction and error analysis for the physical sciences". In: *Computers in Physics* 7.4 (1993), pp. 75–97.

[2] Sang Gyu Kwak and Jong Hae Kim. "Central limit theorem: the cornerstone of modern statistics". In: *Korean journal of anesthesiology* 70.2 (2017), p. 144.

[3] Nicholas Metropolis et al. "The beginning of the Monte Carlo method". In: *Los Alamos Science* 15.584 (1987), pp. 125–130.

[4] William Thistleton et al. "Generalized Box–MÜller Method for Generating -Gaussian Random Deviates". In: *Information Theory, IEEE Transactions on* 53 (Jan. 2008), pp. 4805–4810. DOI: 10.1109/TIT.2007.909173.