PROBLEM CORNER

Md S. Warasi
Department of Mathematics and Statistics
Radford University, VA 24142
email: msarker@radford.edu

Problem 1

Random number generators are widely used in statistical applications for simulating data and validating statistical methods. Nowadays, nearly all programming languages are equipped with built-in programs for generating random numbers. The main goal of this problem is to introduce students with a basic algorithm for generating random numbers from a normal (bell-shaped) distribution without implementing any existing programs. This procedure works in two steps.

- Step 1: Generate two numbers U_1 and U_2 from a uniform distribution that is defined over $[0, 1]$. Note that a uniform distribution over $[0, 1]$ is a unit constant function, $f(y) = 1$ for $0 \leq y \leq 1$.
- Step 2: Calculate $X_1 = \sqrt{-2 \ln U_1} \cos(2\pi U_2)$ and $X_2 = \sqrt{-2 \ln U_1} \sin(2\pi U_2)$.

Upon completion, these steps will result in two independent observations, X_1 and X_2, that follow a normal distribution with mean = 0 and variance = 1. This algorithm was introduced by Box and Muller (1958).

The uniform numbers, U_1 and U_2, required for this algorithm can be generated as follows. Select two numbers at random from $0, 1, 2, ..., 200$ with replacement. If either 0 or 200 is selected, discard it and select another number. Divide both numbers by 200 (the maximum possible value). This will yield a pair of independent uniform numbers, U_1 and U_2, that fall between 0 and 1. Next, complete step 2 to obtain X_1 and X_2. Repeat the entire process to simulate multiple pairs of normal variates. For example, to simulate 100 numbers, one needs to repeat these steps 50 times. Note that 200 is a reasonable maximum value for the simple random sampling. However, the performance would be better if a larger value, such as 500, is used.

Do the following using the algorithm described above.

(a) Sample 100 random numbers from a normal distribution that has mean = 0 and variance = 1.

(b) Construct a histogram using the 100 sample points. Is the distribution bell shaped?

(c) Write your own program to implement the algorithm. Generate 1000 random numbers using your program.

Problem 2

Random numbers can be used to easily approximate areas under any (complicated) functions. For example, the area under the quadratic function $f(y) = y^2$ over $[0,1]$ is $\int_0^1 f(y)dy = \int_0^1 y^2 dy = 1/3$ (exact result). This can be approximated as follows. Generate 100 random numbers from a uniform distribution that is defined over $[0,1]$ by simple random sampling as described in Problem 1. Substitute each of the 100 numbers for y in $f(y) = y^2$ to get 100 values of the function and then take their average. This average should be a reasonable approximate for $\int_0^1 y^2 dy$ and should be close to 1/3. For a more accurate result, use at least 1000 uniform numbers rather than 100.

Use the numerical technique to approximate $\int_0^1 f(y)dy$ for each of the following functions.

(a) $f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$;

(b) $f(y) = y^2 e^{-y^2}$;

(c) $f(y) = 10^y e^{-e^{-y}}$.

The approximate answers for 2(a), 2(b), and 2(c) are 0.341, 0.189, and 2.336, respectively. Note that the approximation technique described above is valid only for the unit interval $[0,1]$. This can, however, be generalized easily.

References

Box, G. and Muller, M. (1958). A note on the generation of random normal variates. *Annals of Mathematical Statistics* **29**, 610–611.