POLAR NORMAL DISTRIBUTION

BY ALAIN LATOUR

Generate normal deviates and evaluate the cumulative probability of a normally distributed random variable

IN THE OCTOBER 1985 BYTE, Arthur G. Hansen suggested a way to generate normal deviates (Programming Insight: "Simulating the Normal Distribution"). Basically, a normal deviate is obtained by considering the value

$$X = (U_1 + U_2 + \ldots + U_n - n/2)/\sqrt{(n/12)}$$

where $U_1, U_2, \ldots, U_n$ are independent uniform numbers between 0 and 1. The value of X can be considered as a normal deviate with mean $\mu=0$ and standard deviation $\sigma = 1$, if n is sufficiently large. As reported by Abramowitz and Stegun (reference 1), "When $n=12$, the maximum errors made in the normal deviate are 9×10^{-3} for $|X| < 2$, 9×10^{-1} for $2 < |X| < 3$." But if n is too small, we have only an "approximate normal distribution."

THE POLAR METHOD

I would like to suggest a relatively simple approach, known as the "polar method," proposed by G. E. P. Box, M. E. Muller, and G. Marsaglia, which is proven in reference 2. Here is the pseudocode of the algorithm:

Repeat
1) Get 2 independent random variables, V_1 and V_2, uniformly distributed between -1 and 1.
2) Evaluate $S := (V_1)^2 + (V_2)^2$
until $S < 1$.
Evaluate $S := \sqrt{-2\ ln(S)/S}$
Set $X_1 := V_1S$ and $X_2 := V_2S$.

The two random numbers X_1 and X_2 are independent normal variables. If the uniform generator used returns a number U between 0 and 1, you just have to set $V = 2U-1$ to obtain a number between -1 and 1. The square root and the logarithm functions must be evaluated for each pair of normal deviates.

Although the algorithm is complex, its advantage is that you obtain numbers really distributed as normal deviates, and fewer uniform deviates are needed. The last point may be an important one. For example, the random-number generators available on some personal computers cannot produce a sequence of numbers longer than 65,536 without cycling.

CLASSIFYING THE OBSERVATIONS

Suppose you know that almost all the observations will fall between the limits YMin and YMax. That would mean that the data range is approximately YRange = YMax − YMin. To classify the values in Nb classes where Y is an observed value, you can use

Trunc((Y − YMin)/YRange × Nb)

For example, if Y = YMin, the above statement will produce 0. And you could use the extreme classes to collect extreme values.

I have written a short Pascal program called NORMAL.PAS that generates 10,000 normal deviates and sorts them into 30 classes as they are generated. The results could be used to plot the observed distribution of the numbers generated. [*Editor's note:* NORMAL.PAS *is available in a variety of formats; see page* 405 *for details.*]

THE CUMULATIVE PROBABILITY FUNCTION

If you're interested in comparing the simulated distribution with the theoretical distribution, you must

(continued)

Alain Latour (Département de Mathématiques et d'Informatique, C. P. 8888, Succ. "A," Université du Québec à Montréal, Montréal, Quebec, Canada H3C 3P8) teaches statistics and computer science.

evaluate the probability of each class. One way to do it is by tedious calculations using a table of probability. Another, more interesting way is to write a function that, given a number x, will give you the probability that a normal deviate will be less than or equal to x. We will denote this probability by $P|z \leq x|$. There is a polynomial approximation of degree 5 to this probability that is quite good (see reference 1). It requires the evaluation of the density function of a normal variable, that is:

$$f(x) = c\, e^{-x^2/2}$$

To compare the simulated distribution with the theoretical distribution, you must evaluate the probability of each class.

where c is a constant equal to $1/\sqrt{(2\pi)} = 0.3989422804$. So, for $x \geq 0$, we have

$$P|z \leq x| = 1 - f(x)\, t\, (b_0 + b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4) + \text{error}$$

where $t = 1/(1 + px)$ with $p = 0.2316419$ and where $b_0 = +0.319381530$, $b_1 = -0.356563782$, $b_2 = +1.781477937$, $b_3 = -1.821255978$, and $b_4 = +1.330274429$. The error is less than 7.5×10^{-8}. If $x < 0$, you can use the well-known property of the normal curve, $P|z \leq x| = 1 - P|z \leq -x|$. [*Editor's note: Both functions are contained in* DENSITY.PAS, *which is available in several formats; see page* 405.]

Figure 1 shows the empirical distribution obtained by the polar method. As you can see, the adjustment seems to be quite good.

Figure 2 shows the empirical distribution obtained by using the sum of three uniforms. Both cases simulated 10,000 deviates. Obviously, the adjustment is not as good as that in figure 1. You can also compare the two methods by evaluating the chi-square measure of adjustment. With 47 degrees of freedom, in the first case we have a chi-square equal to 36; in the second case, the chi-square is equal to 147. In other words, the method using the sum of three uniforms doesn't pass the chi-square test. ■

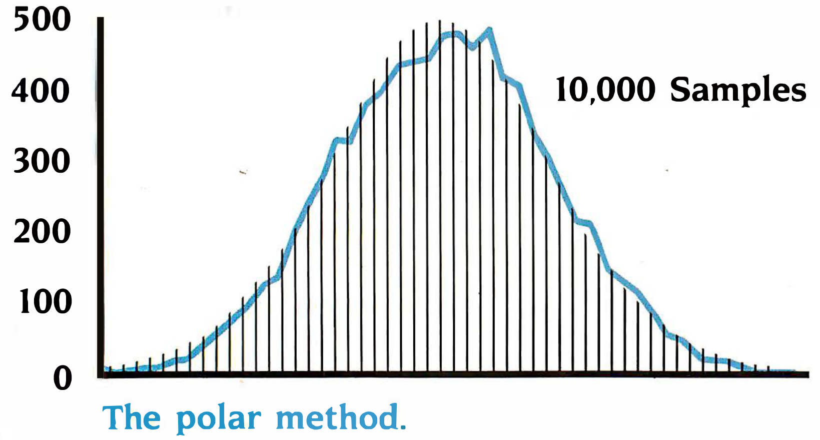

Figure 1: *The empirical distribution obtained by simulating 10,000 deviates by the polar method. The adjustment seems to be quite good.*

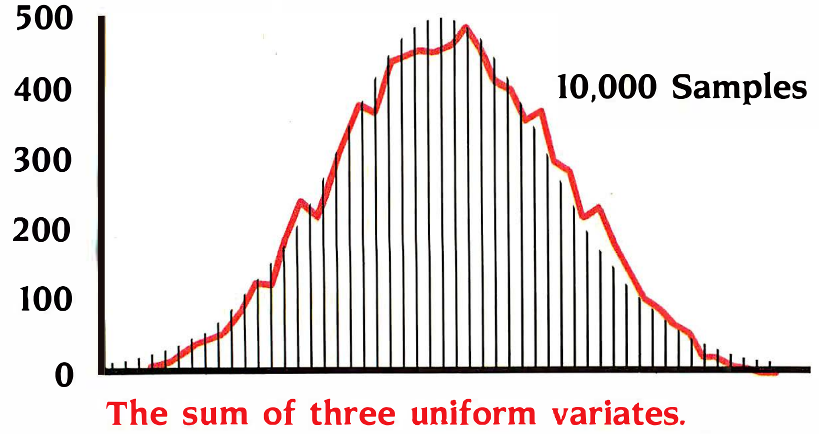

Figure 2: *The empirical distribution obtained by using the sum of three uniforms.*

REFERENCES

1. Abramowitz, M., and I. A. Stegun, eds. *Handbook of Mathematical Functions*. New York: Dover Books, 1964.
2. Knuth, D. *The Art of Computer Programming, Volume 2*. Reading, MA: Addison-Wesley, 1969.