Behavior Research Methods & Instrumentation
1983, 15 (6), 604-605

PROGRAM ABSTRACTS/ALGORITHMS

Accuracy and speed of seven approximations of the normal distribution function

ALFRED L. BROPHY
Behavioral Science Associates,
West Chester, Pennsylvania

Table 1
Seven Approximations of the Normal Distribution Function

```
10 INPUT"Z "; Z1: Z=ABS(Z1)
20 INPUT"Which approximation (1-7) ", AP
30 IF Z>3.9 PRINT"P (One-tailed) < .00005": P=0: GOTO 60
40 ON AP GOSUB 110,210,310,410,510,610,710
50 PRINT"P (One-tailed) =" P
60 PC=100*P: IF Z1>0 THEN PC=100-PC
70 PRINT"Percentile =" PC
80 END
90 '
100 ' Zelen & Severo, Section 26.2.16
110 Y=.3989422804*EXP(-Z*Z/2): T=1/(1+.33267*Z)
120 P=Y*T*(.4361836-T*(.1201676-T*.937298)): RETURN
190 '
200 ' Zelen & Severo, Section 26.2.17
210 Y=.3989422804*EXP(-Z*Z/2): T=1/(1+.2316419*Z)
220 P=Y*T*(.31938153-T*(.356563782-T*(1.781477937-T*
    (1.821255978-T*1.330274429)))): RETURN
290 '
300 ' Zelen & Severo, Section 26.2.18
310 P=.5/(1+Z*(.196854+Z*(.115194+Z*(3.44E-4+Z*.019527))))[4:
    RETURN
390 '
400 ' Zelen & Severo, Section 26.2.19
410 P=.5/(1+Z*(.049867347+Z*(.0211410061+Z*(.0032776263
    +Z*(3.80036E-5+Z*(4.88906E-5+Z*5.383E-6))))))[16: RETURN
490 '
500 ' Cadwell (modified)
510 X=Z*Z: P=.5-SQR(1-EXP(-X*(.6366197724-X*(.009564223505
    -X*4E-4))))/2
590 '
600 ' Moran, Equation 4
610 DEFINT I: B=0: S=Z*.4714045208
620 FOR I=1 TO 12: A=EXP(-I*I/9)*SIN(I*S)/I: B=B+A: NEXT I
630 P=.5-.3183098862*(S/2+B): RETURN
690 '
700 ' Moran, Equation 5
710 B=0: S=Z*.4714045208
720 FOR H=.5 TO 12.5: A=EXP(-H*H/9)*SIN(H*S)/H: B=B+A: NEXT H
730 P=.5-.3183098862*B: RETURN
790 '
800 REMARKS
810 ' .3989422804=1/SQR(2*PI)    (lines 110 & 210)
820 ' .6366197724=2/PI,    .009564223505=2*(PI-3)/(3*PI*PI)
      (line 510)
830 ' .4714045208=SQR(2)/3       (lines 610 & 710)
840 ' .3183098862=1/PI           (lines 630 & 730)
```

In statistics programs, the evaluation of a normal distribution function is usually accomplished by an approximation algorithm, the most popular of which are those based on Hastings (1955) and adapted by Zelen and Severo (1964). This paper compares seven approximations of the normal distribution function with respect to accuracy and speed of execution on a microcomputer. The results can aid in choosing an appropriate procedure for estimating the probability of a normal deviate (z).

Table 1 shows the approximations, written in BASIC, together with a brief program to input a value of z, call a selected approximation, and print the results. Four of the approximations are from Zelen and Severo (1964, sections 26.2.16-26.2.19), one is based on Cadwell (1951), and two are newer approximations developed by Moran (1980) from Strecock's (1968) approximation of the error function. Approximations 26.2.16 and 26.2.17 from Zelen and Severo require the calculation of the probability density as a first step. Cadwell's approximation was modified slightly by substitution of a new correction term for two terms in the original formula. The modification not only simplifies the approximation but also improves its accuracy somewhat for many values of z. Moran's approximations, which have not yet received much attention, were programmed from Moran's Equations 4 and 5. Unlike Hastings-type approximations, Cadwell's and Moran's approximations do not rely on special constants, although they do use pi, and Moran's approximations use the sine function.

Each approximation returns the one-tailed probability of the value of z that is entered. A positive z is assumed as input for all the approximations except Cadwell's (1951), but Line 60 of the driver program, which calculates the percentile rank of z, illustrates how the sign of z can be taken into account.

Tests of the Approximations. The approximations were tested on a Radio Shack Color Computer. This machine was selected because it has nine-digit numeric precision and because it is a relatively popular, inexpensive computer that is useful in the psychology laboratory (Reed, 1982). Its level of precision, although comparable to that of Apple computers, is greater than the standard six-digit precision of many microcomputers. A computer with Extended Color BASIC was used because that language includes exponentiation and the natural exponential function required by some approximations.

The approximations were run as they appear in Table 1, except for the deletion of spaces and the omission of the DEFINT statement in Line 610, which is illegal in Color Computer BASIC. Each approximation was tested with 13 values of z (.02, .10, .40, .80, 1.28, 1.64, 1.96, 2.32, 2.58, 3.10, 3.30, 3.70, and 4.00). Correct values of the normal distribution function were considered to be those tabled by Zelen and Severo (1964, pp. 966-973), who gave cumulative probabilities to a minimum of 10 decimal places. The tail areas needed for comparison with the results of the approximations were calculated by subtracting the tabled values from unity.

Table 2 summarizes the results of the tests. Moran's (1980) approximations are the most accurate of those tested, and his Equation 5 should be used when maximal accuracy is required. However, Moran's approxi-

The author's mailing address is: Alfred L. Brophy, Behavioral Science Associates, P.O. Box 748, West Chester, Pennsylvania 19380.

Table 2
Accuracy, Execution Time (in Seconds), and Length (in Bytes) of Seven Approximations of the One-Tailed Probability of the Normal Deviate, Calculated on the Radio Shack Color Computer

Approximation	Range of Correct Decimal Places	Mean Time	Length
Zelen & Severo			
Section 26.2.16	4-6	.3	92
Section 26.2.17	6-7	.4	133
Section 26.2.18	2-4	.2	63
Section 26.2.19	5-7	.4	106
Cadwell (modified)	3-8	.3	71
Moran			
Equation 4	8-10	1.2	99
Equation 5	9-10	1.3	96

mations execute more slowly than the others, and the degree of accuracy they provide is greater than is ordinarily necessary in applied statistical work. Therefore, for many purposes, the other approximations offer a better combination of accuracy and speed, although Zelen and Severo's (1964) Approximation 26.2.18 is relatively inaccurate. Since the approximation routines differ little in size (see Table 2), memory usage will seldom be an important criterion for choosing among them.

The approximations also were tested on a TRS-80 Model I computer with Level II BASIC in its single-precision six-digit mode. For these trials, all constants with more than seven significant digits were rounded to seven significant digits. The ranges of correct decimal places attained by Zelen and Severo's (1964) approximations on the Model I were identical to those obtained with the Color Computer. The accuracy attained by Cadwell's (1951) and Moran's (1980) approximations on the Model I was lower than that obtained with the Color Computer: three-six decimal places for Cadwell's approximation, six-seven decimal places for Moran's Equation 4, and six-eight decimal places for Equation 5. For use on a computer with six-digit precision, it was found empirically that the sum of 10 terms in Line 620 and the sum of 11 terms in Line 720 yield results as accurate as the greater numbers of terms in the listing in Table 1.

Limited trials on the Color Computer with large absolute values of z found that Moran's (1980) equations lose accuracy when z is greater than approximately 6.8 and that Zelen and Severo's (1964) Approximation 26.2.19 overflows when z is greater than approximately 17.2. These problems occur on the Model I at somewhat lower values of z. Line 30 of the driver program avoids these situations, as well as extrapolation beyond the range of z values tested, by returning a predetermined probability when the absolute value of z exceeds 3.9.

Language. The approximations were run, with the minor modifications noted, on a Radio Shack Color Computer with Extended BASIC and on a TRS-80 Model I with Level II BASIC. They will run without change on manv computers with other BASIC interpreters. Changes necessary for some computers may include separation of multiple-statement lines and insertion of LET in assignment statements.

Availability. A listing of the approximations can be obtained without charge from Alfred L. Brophy, Behavioral Science Associates, P.O. Box 748, West Chester, Pennsylvania 19380.

REFERENCES

Cadwell, J. H. The bivariate normal integral. *Biometrika*, 1951, **38**, 475-479.

Hastings, C., Jr. *Approximations for digital computers.* Princeton, N.J: Princeton University Press, 1955.

Moran, P. A. P. Calculation of the normal distribution function. *Biometrika*, 1980, **67**, 675-676.

Reed, A. V. The Radio Shack Color Computer in the experimental psychology laboratory: An evaluation. *Behavior Research Methods & Instrumentation*, 1982, **14**, 109-112.

Strecock, A. J. On the calculation of the inverse of the error function. *Mathematics of Computation*, 1968, **22**, 144-158.

Zelen, M., & Severo, N. C. Probability functions. In M. Abramowitz & I. A. Stegun (Eds.), *Handbook of mathematical functions.* Washington, D.C: U.S. Government Printing Office, 1964.

(Manuscript accepted for publication August 4, 1983.)