Computer approximation of the inverse of the normal distribution function

PHILLIP L. EMERSON

Cleveland State University, Cleveland, Ohio 44118

Psychological research utilizing techniques based on signal detection theory and other scaling methods frequently requires the transformation of probabilities to their corresponding normal standardized deviates. For example, estimation of d' using data from the common yes-no method requires that the empirical estimates of the probabilities of "hits" and "false alarms" be transformed to the normal deviates, and the difference then taken. As there is no known closed algebraic expression for this transformation, it must be done numerically. It can be performed by hand by making use of an ordinary table of the normal distribution, entering in a reverse fashion and, perhaps, using linear interpolation. More convenient for paper-and-pencil methods are the special tables of Hochhaus (1972) and Appendix 1 in Swets (1964). With the increasing use of mini- and microcomputers in the laboratory, however, occasions frequently arise where one would like to avoid these manual operations. After the computer has controlled the experimental session and accumulated the data for the probability estimates, why not have it also compute and display the explicit numerical values of the statistics of main interest, such as d' ?

Computer methods of doing just that have appeared in this journal and elsewhere, and a crucial step is the inversion of the normal distribution function. Comparisons at this laboratory indicate that one method stands out well above the others for transforming empirical probability estimates to the corresponding normal deviates. A FORTRAN version¹ is listed in Figure 1. This algorithm is the same as that used in the APL function published by McGowan and Appel (1977) and attributed to Snyder (1973). Also, it is the approximation made available to IBM users for many years in the Scientific Subroutine Package. It is due to Hastings (1955), and the maximal error² is about .00045.

Instead of a presentation of details of the numerical tests that have been made in comparing this method with others, the following brief argument of its virtues should be sufficient: (1) Its accuracy³ (maximal error of .00045) is quite adequate, since empirical probabilities to be transformed are usually subject to larger sampling errors; (2) it is easy to code, requiring only a dozen or so FORTRAN or BASIC statements; (3) it is reasonably fast,⁴ involving only seven floating-point multiplications or divisions and two standard transcendental functions; (4) it is noniterative, so that the number of numerical

```

FUNCTION Z(P)
  SGN=-1
  Q=P
  IF (Q-.5) 2,2,1
1  SGN=1
  Q=1.-Q
2  Z=SQRT(-2.*ALOG(Q))
  W=Z*(.18926+Z*.001308)
  W=1.+Z*(1.432788+W)
  Q=.802853+Z*.010328
  W=(2.515517+Z*Q)/W
  Z=(Z-W)*SGN
  RETURN
END

```

Figure 1. Inverse normal distribution function using the Hastings (1955) formula.

operations does not depend on the value of the probability to be transformed; (5) it is compact, utilizing only six stored numerical constants.

Although such algorithms necessarily are compromises among the goals of compactness, speed, accuracy, and the ease of coding, one would have to search far and wide to find a more efficient or more balanced compromise than this one.

REFERENCES

- AHROON, W. A., JR., & PASTORE, R. E. Procedures for computing d' and β . *Behavior Research Methods & Instrumentation*, 1977, 9, 533-537.
- CUNNINGHAM, S. W. Algorithm AS-24: From normal integral to deviate. *Proceedings of the Royal Statistical Society, Series C, Applied Statistics*, 1969, 18, 290-300.
- HASTINGS, C. *Approximations for digital computers*. Princeton, N.J.: Princeton University Press, 1955.
- HOCHHAUS, L. A table for calculating d' and β . *Psychological Bulletin*, 1972, 77, 375-376.
- MCGOWAN, W. T., III, & APPEL, J. B. An APL program for calculating signal detection indices. *Behavior Research Methods & Instrumentation*, 1977, 9, 517-521.
- SNYDER, M. Interactive data analysis and non-parametric statistics. In P. Gierlow, H. J. Helms, & J. Nielson (Eds.), *APL Congress 73*. New York: American Elsevier, 1973.
- SWETS, J. A. (Ed.). *Signal detection and recognition by human observers*. New York: Wiley, 1964.

NOTES

1. For simplicity, the program listed here assumes that the value of the input argument is in the range for which the function is defined: namely, $0 < P < 1$. Violation will lead to an attempt to take the logarithm of zero or square root of a negative number. This differs from the "standard" version, in the IBM Scientific Subroutine Package, which returns z values of plus or minus 10^7 when the condition is violated.

2. This was the maximal error specified by Hastings (1955) and it has been verified numerically here by reference to published tables and by comparison with certain more accurate methods that have, however, disadvantages for general use.

3. If anyone can think of an application requiring more accuracy, the method of Cunningham (1969), which uses the Hastings (1955) algorithm to obtain a first approximation, is accurate to 10 decimal places. However, it is slower and less compact.

4. In some exceptional case where greater speed is required,

one could resort to the table look-up method of Arhoon and Pastore (1977), which, however, sacrifices much in the way of accuracy and compactness.

(Accepted for publication January 8, 1979.)