Computing probabilities for F, t, chi-square, and z in BASIC

GERNOT von COLLANI

Technical University of Brunswick, Braunschweig, West Germany

Although quick and simple approximations have been suggested for the computation of probability distributions (see Abramowitz & Stegun, 1965), and computer programs for such computations are available (e.g., Jaspen, 1965; Ogasawara, 1982), the accuracy of these approximation formulas is rather poor in most cases. This is particularly true for extreme values of the distributions and for a small number of degrees of freedom.

In order to make some common probability distributions available on a microcomputer and to obtain a sufficient accuracy of computation over a wide range of arguments, series expansions are used here to compute values of chi-square, F, and t (Abramowitz & Stegun, 1965, pp. 941-946). For the z distribution, a sixth-degree polynomial approximation proposed by Hastings (1955) is chosen (see Abramowitz & Stegun, 1965, p. 932). The program is written in Level II BASIC for the TRS-80. A comparable program for larger machines in FORTRAN uses somewhat different algorithms from those employed here (O'Grady, 1981).

The following probability is determined for each of the distributions:

$$P(X > x) = \int_x^{+\infty} f(t) dt,$$

where x is a value from the domain of the random variable X (i.e., an F, chi-square, t, or z value) and $f(t)$ is the corresponding probability density function.

Accuracy and Computing Time. All significant numerical computations make use of double-precision variables and constants (about 16 decimal places on the TRS-80), with the exception of the exponentiation and the mathematical functions, which only return single-precision values on the TRS-80 (accuracy 6 decimal places). The resulting accuracy is at least 5 decimal places for the normal distribution and about 4 to 6 decimal places for the other distributions, depending on the kind of distribution and its arguments.

The accuracy was checked for a large number of extreme values of the distributions for both small

and large degrees of freedom. The correspondence of the computed values and those tabulated by Abramowitz and Stegun (1965) was satisfactory.

If an appropriate compiler that supports double-precision mathematical functions (e.g., the MICROSOFT BASIC compiler) is at hand, its use recommended because of the gain in computing time and accuracy.

It takes less than 2 sec to compute the probabilities for the z distribution. The typical computing time for the other distributions is a few seconds, depending upon the kind of distribution and the degrees of freedom. When the degrees of freedom are about 500, the largest computing time observed is 12 sec for the compiled program (MICROSOFT BASIC compiler) and 27 sec for interpreted BASIC.

Limitations. When the degrees of freedom exceed 500, the computing time may increase considerably. In order to speed up the computations in such cases, the common rough and simple approximation formulas, instead of the series expansions, are used. These are the Wilson-Hilferty approximations for the chi-square distribution and a related approximation for the F and the t distributions (see Abramowitz & Stegun, 1965, pp. 941; 947), which all follow a unit normal distribution.

To avoid an arithmetic overflow or underflow due to the internal representation of floating-point numbers on the TRS-80 (the admissible range of numbers is from $1.7 \cdot 10^{-38}$ to $1.7 \cdot 10^{+38}$, approximately), these simple approximations are also used for large values of F, t, and chi-square.

System Requirements. The BASIC program occupies 4.5 KB on the TRS-80. It should easily be adapted to other machines with MICROSOFT BASIC. If only single-precision variables are supported, nine significant digits are recommended. The compiled program (MICROSOFT BASIC compiler) takes up approximately 19 KB. Because of a copyright notice by the compiler manufacturer, the author of the present program is not free to distribute the compiled version.

Availability. Source code listings for the TRS-80 and for PET/CBM computers are available at no cost from G. von Collani, Department of Psychology, Technical University of Brunswick, D-3300 Braunschweig, West Germany.

REFERENCES

- ABRAMOWITZ, M., & STEGUN, I. A. (Eds.). *Handbook of mathematical functions*. New York: Dover, 1965.
- HASTINGS, C., JR. *Approximations for digital computers*. Princeton, N.J.: Princeton University Press, 1955.
- JASPEN, N. The calculation of probabilities corresponding to

The author's mailing address is: Department of Psychology, Technical University of Brunswick, D-3300 Braunschweig, West Germany.

- values of z, t, F, and chi-square. *Educational & Psychological Measurement*, 1965, **25**, 877-880.
- OGASAWARA, T. H. The calculation of the significance level of F, t, and r on the Apple II. *Behavior Research Methods & Instrumentation*, 1982, **14**, 492-493.
- O'GRADY, K. E. Probabilities and critical values for z, chi square, r, t, and F. *Behavior Research Methods & Instrumentation*, 1981, **13**, 55-56.

(Revision accepted for publication June 23, 1983.)