An APL function for computing distribution-free confidence limits on the median of a population

L. WOUTERS

Cardiovascular Department
Janssen Pharmaceutica, Beerse, Belgium B-2340

Depending upon the scale of measurement of a variable and the shape of its distribution, the mean, median, or mode are taken as a measure of its location. For data not normally distributed and measured at least on an ordinal scale, the median is recommended (Stevens, 1946). In order to indicate the uncertainty of the estimation, a confidence interval around the median value should be constructed.

Owen (1962, p. 362) gives tables, based on the sign test, for determining the distribution-free confidence limits on the median of a population. The tables consist of maximum values of A in the binomial sum $B(n, A, 0.5)$, such that:

$$B(n, A, 0.5) = \sum_{j=0}^A \frac{n!}{j!(n-j)!2^n} \leq \gamma \text{ and } 0 < A \leq n/2.$$

In this equation n is the sample size, γ the probability level, and A is the critical value of the sign test, or the smallest item lying within the confidence interval.

It is the purpose of this paper to describe a simple APL function (SCRITIC), which computes this binomial sum and returns values of A . Furthermore, the use of this function is demonstrated by its incorporation into a general program for descriptive statistics (DESTAT).

Computer and Language. The functions SCRITIC and DESTAT were developed and executed under VSAPL on an IBM 370/148 computer.

Input. The dyadic function SCRITIC has, as input variables, the probability level and the sample size. The monadic function DESTAT works on a vector containing the data.

Computational Procedures. Since the computation of a binomial sum often leads to overflow errors, a transformation of the equation to a logarithmic form is used. Figure 1 shows the routine SCRITIC, together with some test runs. The incorporation of SCRITIC into the general program DESTAT and a test run is illustrated in Figure 2.

Output. Let A represent the output of the function $C \text{ SCRITIC } N$, where C stands for the one-tailed probability level and N is the sample size. Then the 100% $(1 - 2C)$ confidence interval on the median of the population is from the A th largest to the $(1 + N - A)$ th largest value.

The output of DESTAT is shown in Figure 2. When-

The author would like to thank Ludo Gijpen for his suggestions and for reviewing this paper.

```

SCRITIC{C}V
V R←C SCRITIC N;J,LNF
[1] FUNCTION COMPUTES INDEX OF THE SMALLEST DATA-ITEM
[2] LYING WITHIN THE 100*(1-2*C) CONFIDENCE INTERVAL
[3] C : PROBABILITY LEVEL (ONE-TAILED)
[4] N : SAMPLE SIZE
[5] EXAMPLE :
[6] 95-PERCENT CONFIDENCE INTERVAL OF N ELEMENTS
[7] 0.025 SCRITIC N
[8]
[9] J←1+N*0.5
[10] LNF←(0, (+*1+N-J)-+*J)-N*0.2
[11] R←(0,J)*C>+*LNF
V

```

0.1 SCRITIC 10

2 0.025 SCRITIC 100

39 0.025 SCRITIC 6200

3022

Figure 1. The APL function SCRITIC for confidence limits on the median.

```

DESTAT{V}
V DESTAT X;R;MAX;MIN;N;MEAN;VAR;SD;ND;SEM;MED;CON
[1] FUNCTION PERFORMS DESCRIPTIVE STATISTICS ON A VECTOR X
[2]
[3] R←(MAX-X[N*0.5])-MIN-(X-X[N*0.5])[1]
[4] SEM←(SD*(VAR*(+/(X-NEAR-(+/X)*N)*2)+N-1)*0.5)+N*0.5
[5] ND←(+/X-NEAR)*N
[6] MED←0.5*(X[(N*0.5)+1]+X[N*0.5])
[7] CON←X[CON,1+N-CON-CON,10*CON-0.025 SCRITIC N]
[8] 'SAMPLE SIZE' 'N'
[9] 'MAXIMUM' 'MAX'
[10] 'MINIMUM' 'MIN'
[11] 'RANGE' 'R'
[12] 'MEAN' 'MEAN'
[13] 'VARIANCE' 'VAR'
[14] 'STANDARD DEVIATION' 'SD'
[15] 'S.E.M.' 'SEM'
[16] 'MEAN DEVIATION' 'ND'
[17] 'MEDIAN' 'MED'
[18] 'CONF LIMITS MEDIAN' 'CON'
V

```

DESTAT 60 88 63 85 70 81 75 96 87.7 70.2 89.5 98.6

SAMPLE SIZE 12

MAXIMUM 98.6

MINIMUM 60

RANGE 38.6

MEAN 80.33333333

VARIANCE 159.109697

STANDARD DEVIATION 12.61386923

S.E.M. 3.641310398

MEAN DEVIATION 10.57777778

MEDIAN 83

CONF LIMITS MEDIAN 63 96

Figure 2. The incorporation of SCRITIC into the program DESTAT.

ever DESTAT cannot compute the confidence limits, an empty vector is printed.

Limitations. The two variables (probability level and sample size) that are input to SCRITIC must be scalars. The argument of DESTAT must be a vector (only one dimension). The memory requirements of the two functions are mainly determined by the sample size of the data they operate upon.

Availability. Copies of this paper, test data, and listing may be obtained without charge from Luc Wouters, Cardiovascular Department 2744, Janssen Pharmaceutica N.V., Turnhoutsebaan, Beerse, Belgium B-2340.

REFERENCES

- STEVENS, S. S. On the theory of scales of measurement. *Science*, 1946, 103, 677-680.
- OWEN, D. B. *Handbook of statistical tables*. Reading, Mass: Addison-Wesley, 1962.

(Accepted for publication December 28, 1978.)