

## Simple algorithms for M-alternative forced-choice calculations

J. E. KEITH SMITH  
University of Michigan  
Ann Arbor, Michigan 48104

Recently, Hacker and Ratcliff (1979) published an accurate and extensive table for estimating  $d'$  from the percent correct in an M-alternative forced-choice experiment. Since tables are somewhat tedious to use with modern calculating equipment, I have developed two algorithms to calculate  $d'$  calibrated against the Hacker and Ratcliff table. The first, which can be used on any desk calculator having a natural logarithm key, is based on the Luce choice model and has a maximum probability error of less than .02 for M between 2 and 11 and  $d'$  between 0 and 3. The second algorithm is based on a normal approximation and, hence, requires a routine for computing the inverse normal function. Many desk calculators (such as, for example, the HP-25) either have this function wired in or are programmable. The second algorithm satisfies the 2% error bound for all M and all percentiles and, except for M = 3 or 4, satisfies a 1% error bound.

### Algorithm 1

If the effect of a stimulus presented in an interval increases the probability of choice of that interval by a factor  $\lambda$ , and if  $M < 12$ , the natural logarithm of  $\lambda$  is very nearly proportional to  $d'$ . Letting  $d^*$  be the estimate of  $d'$ ,

$$d^* = K_M \ln \frac{(M-1)P_c}{1-P_c}, \quad (1)$$

where

$$\hat{\lambda} = (M-1)P_c / (1-P_c). \quad (2)$$

Notice that  $P_c = 1/M$  implies  $d^* = 0$ . The factor  $K_M$  is a slowly decreasing function of M, and the formula

$$K_M = 0.86 - 0.085 \ln(M-1)$$

works very well over the range  $M = 2, 11$ . As M increases, however, the proportionality breaks down over the range of P, and for larger M another technique is necessary.

### Algorithm 2

The near normality of the entries in the Hacker-Ratcliff table for all values of M suggests an attempt

to find functions of M which describe the mean and standard deviation of the  $d'$  variation. The signal detection theory model has it that  $P_c$  is the probability that the observation in the signal interval is larger than the maximum of the observations in the other  $M-1$  intervals. Kendall and Stuart (1963) discuss the distribution of the maximum of  $n$  independent samples from a normal distribution and show that, for large  $n$ , the mode of this distribution increases as  $(\ln n)^{1/2}$ , while the standard deviation decreases as  $(\ln n)^{-1/2}$ . The limiting distribution is not normal, or even symmetric, being rather skewed to the right. Kendall and Stuart's asymptotic formula arises from taking only the leading terms in an exact formula they give (p. 334) for the mode and is quite inexact even for  $n$  as large as 1,000. Expanding their exact formula around a modal value of 2.0 to second order, one arrives at a formula for the mode,  $\chi_{on}$ ,

$$\chi_{on} = (-4 + \sqrt{16 + K \ln n})/3. \quad (3)$$

Matching this formula to the Hacker and Ratcliff values for  $P_c = .5$ , a value of  $K = 25$  fits the whole range (2-1,000) rather well and in fact fits  $M = 2$  exactly.

The variance of the difference between the normal deviates corresponding to the signal interval and the maximum noise interval should be roughly  $1 + f(\sqrt{\ln n})$ , and choosing a simple function that provides the right value for  $M = 2$ , one obtains  $V(\text{diff}) = [\ln(M-1) + 2] / [\ln(M-1) + 1]$ .

Finally, then, the second algorithm defines  $d^{**}$ , the estimate of  $d'$  to be

$$d^{**} = A + B N^{-1}(P_c), \quad (4)$$

where  $A = [-4 + \sqrt{16 + 25 \ln(M-1)}]/3$ ,  $B = \{[\ln(M-1) + 2] / [\ln(M-1) + 1]\}^{1/2}$ , and  $N^{-1}(P_c)$  is the unit normal deviate corresponding to a right tail area P, that is,  $N^{-1}(.5) = 0$ ,  $N^{-1}(0.1586) = +1$ .

### Accuracy

A number of indices of accuracy have been calculated using these algorithms, always using the Hacker and Ratcliff table as a criterion. For each of the three indices reported here, the tabled value is the maximum absolute error over that part of the table for which  $0 \leq d' \leq 3.0$ .

The two most obvious indices are the error in  $d'$  for a fixed p and the difference in p for a fixed  $d'$ . Both are imperfect. An error in  $d'$  between  $d' = 0$  and  $d' = .5$  would be quite serious, but the same error between  $d' = 3.0$  and  $d' = 3.5$  would usually be trivial, since, for 4AFC, the latter difference corresponds to a difference of less than .03 in probability, while

Table 1  
Error Analyses

M	a		b		c	
	$n_1^*(1)$	$n_2^*(1)$	$e_1(p)$	$e_2(p)$	$e_1(d')$	$e_2(d')$
2	118		.013	.000	.45	.00
3	152	772	.014	.015	.25	.06
4	278	1189	.013	.013	.18	.06
5	434	2066	.011	.009	.08	.06
6	865	2500	.009	.008	.08	.05
7	1890	2268	.010	.006	.05	.05
8	1736	2066	.012	.006	.04	.05
9	1276	2268	.013	.006	.05	.05
10	1189	2500	.013	.007	.05	.04
11	772	2500	.015	.007	.05	.04
12	625	2066	.019	.007	.07	.04
16	223	2770	.030	.007	.12	.04
24	80	3460	.051	.008	.20	.03
32		3086	.070	.008	.26	.03
256		4444		.005		.04
1000		1000		.004		.01

Note—Subscript “1” refers to the first algorithm; subscript “2” refers to the second. Irregularity in Column  $n_2^*(1)$  is due to the coarseness of the criterion.

the former corresponds to a difference of .14. The other criterion, the difference in  $p$  for a fixed  $d'$ , is much better, but seems to overcorrect; that is, the difference between  $p$  values of .90 and .95 “should” be more important than that between .47 and .52.

An intermediate index is proposed and tabled here:  $n^*(k)$ , which is the number of observations needed to make the two  $p$  values “ $k$ ” standard deviations apart. If  $p$  is the true  $p$  value and  $p'$  is the approximating one,

$$n^*(k) = k^2 p(1-p)/(p-p')^2. \quad (5)$$

Table 1 displays  $n^*(1)$ ,  $e(p)$ , the error in  $p$  for fixed  $d'$ , and  $e(d')$ , the error in  $d'$  for fixed  $p$ , respectively.

If one adopts the (arbitrary) criterion that the error should be less than one standard error for 100 observations, the simple algorithm is suitable for  $M$  up to 16. Even for the worst case,  $M=3$ , the second algo-

rithm, the matching  $p'$ , and hence  $d'$ , is in error less than one standard deviation for sample sizes less than 700. If one required merely that the error be insignificant at the 5% level ( $k=1.96$ ), the numbers in Table 1a would be quadrupled.

In Table 1b, if one adopts the Elliott (1977) criterion of a maximum error of .02 in  $p$ , the simple algorithm is appropriate for  $M=2$  to 11 or 12 and the second algorithm is for all tabled  $M$ .

In Table 1c, if one adopts the criterion of one decimal place of accuracy in  $d'$ , the simple algorithm is suitable for  $M=5$  to 12. The second algorithm is again suitable for the range of  $M$ . Indeed, the maximum  $d'$  error for the second algorithm over all of the Hacker and Ratcliff table is .11.

For  $M=2$ , the error in  $d'$  for the simple algorithm exceeds .1 for  $d' > 2.00$ . For  $M=3$ , this occurs for  $d' > 2.31$  and for  $M=4$ , for  $d' > 2.61$ .

### Recommendations

For experiments in which  $d'$  is calculated on less than 100-200 responses and  $M \leq 10$ , the error in using the simple algorithm is less than the inherent sampling variability of an experimental proportion.

The second algorithm is quite accurate over the whole range of  $M$  and  $p$ . In particular, it will generally be more accurate for nontabled values of  $M$  than interpolation between tabled values.

### REFERENCES

- ELLIOT, P. B. Tables of  $d'$ . In J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1977.
- HACKER, M. J., & RATCLIFF, R. A revised table of  $d'$  for  $M$ -alternative forced choice. *Perception & Psychophysics*, 1979, 26, 168-170.
- KENDALL, M. G., & STUART, A. *The advanced theory of statistics* (Vol. 1). New York: Hafner, 1963.

(Manuscript received August 13, 1981;  
accepted for publication September 30, 1981.)