Estimation of signal detection theory parameters from rating-method data: A comparison of the method of scoring and direct search*

DONALD D. DORFMAN, LYNN L. BEAVERS,
and CARL SASLOW

University of Iowa, Iowa City, Iowa 52240

Signal detection theory parameters were obtained from rating-method data by the method of scoring and two direct search procedures. The method of scoring was much more efficient than direct-search procedures such as the Hooke-Jeeves direct search and Chandler's well-known variation, STEPIT. All were about equally accurate.

Maximum-likelihood estimates of signal detection theory parameters are easily obtained from rating-method data. Ogilvie & Creelman (1968) have provided the solution for the logistic distribution, and Dorfman & Alf (1968, 1969) have done so for the normal distribution. In both cases, the method of scoring was used (Rao, 1965). This method requires computation of both the first and expected second partial derivatives. A very common subroutine used by psychologists to estimate parameters is the subroutine STEPIT (Chandler, 1965). This subroutine uses a variation of the direct-search procedure of Hooke & Jeeves (1961). The major purpose of this paper is to compare the algorithm used by Dorfman & Alf (1969), called here RSCORE, with subroutine STEPIT. In addition, the Hooke and Jeeves direct-search procedure was programmed in our laboratory as a double-precision subroutine (MAX), and was also used in the comparisons. The unequal-variance Gaussian model was assumed.

The program used to evaluate the method of scoring was a revision of the program of Dorfman & Alf (1969). Note that, while their paper was conceptually correct, there were two errata (corrected by Grey & Morgan, 1972). In the Dorfman-Alf notation, the corrections are:

$$E\frac{\partial^2 \ln L}{\partial b \partial Z_j} = -n_2 f_{2j} b \left[\frac{f_{2j}Z_j - f_{2,j-1}Z_{j-1}}{F_{2j} - F_{2,j-1}} - \frac{f_{2,j+1}Z_{j+1} - f_{2j}Z_j}{F_{2,j+1} - F_{2j}} \right],$$

$$E\frac{\partial^2 \ln L}{\partial Z_j \partial Z_{j+r}} = \begin{cases} 0 \text{ if } r > 1 \\ \frac{n_2 f_{2j} f_{2,j+1} b^2}{F_{2,j+1} - F_{2j}} + \frac{n_1 f_{1j} f_{1,j+1}}{F_{1,j+1} - F_{1j}} \text{ if } r = 1, \\ r \neq 0 \end{cases}$$

where $r = 1$ to $n' - j$.

EXAMPLE

The data of the four Os presented in Swets, Tanner, & Birdsall (1955) were used. These data have already been analyzed in the yes-no case (Dorfman & Alf, 1968), and are here reanalyzed for the rating-method case (for the earlier analysis, see Dorfman & Alf, 1969). The experiment was a visual detection study in which the O placed each observation in one of six categories of a posteriori probability. There were 1,188 trials for each O with $p(s) = p(n) = 0.5$.

The program, RSCORE, computes least-squares estimates as initial parameter estimates for the method of scoring. These initial estimates were also used as the starting point for STEPIT and MAX. In addition, RSCORE computes the variance-covariance matrix. Search algorithms do not provide the variance-covariance matrix except by gross approximation. Both single- and double-precision versions of STEPIT were employed (called STEPIT and DSTEPIT, respectively). RSCORE was written in single precision and hence is most usefully compared to STEPIT with respect to efficiency. All the programs are of interest with respect to accuracy of maximization.

RESULTS AND DISCUSSION

Table 1 presents the maximized log likelihoods, χ^2s$(-2 \ln \lambda)$, and sensory-parameter estimates for the four programs $[a = (\mu_s - \mu_n)/\sigma_s; b = \sigma_n/\sigma_s]$. First, we find that none of the chi squares were statistically significant ($\chi^2 = 7.81$, $\alpha = .05$, $df = 3$). This provides very strong support for the Gaussian model when applied to rating-method data.

With respect to the major question of interest, we find that all four programs yield very similar results. STEPIT, however, was the least accurate, since it gave the smallest maximized log L on the average.

Table 2 presents the least-squares estimates of the sensory parameters, and the log Ls and chi squares based upon these estimates. For comparison purposes, the results from RSCORE are also presented. As can be seen, the method of least squares also provides fairly good estimates of the sensory parameters, although the chi squares are much too large, as one might expect.

Table 1
Maximized Log Likelihood, χ^2 (df = 3), and Sensory Parameters, a and b

0	RSCORE Ln L	χ^2	STEPIT Ln L	χ^2	DSTEPIT Ln L	χ^2	MAX Ln L	χ^2
1	−1920.985	1.49	−1920.988	1.49	−1920.985	1.48	−1920.985	1.48
2	−1802.133	6.78	−1802.138	6.79	−1802.134	6.78	−1802.134	6.78
3	−1577.664	2.82	−1577.664	2.82	−1577.664	2.82	−1577.664	2.82
4	−1285.527	2.05	−1285.566	2.13	−1285.528	2.05	−1285.528	2.05
	a	b	a	b	a	b	a	b
1	1.072	0.706	1.071	0.708	1.072	0.706	1.072	0.706
2	1.379	0.736	1.377	0.738	1.380	0.738	1.379	0.736
3	1.225	0.716	1.225	0.717	1.225	0.716	1.225	0.716
4	1.395	0.887	1.422	0.913	1.398	0.889	1.395	0.887

Table 2
Log Ls, χ^2s, and Sensory Parameter Estimates from Least Squares and Method of Scoring

0	Least Squares Ln L	χ^2	RSCORE Ln L	χ^2
1	−1922.40	4.32	−1920.98	1.49
2	−1813.26	29.04*	−1802.13	6.78
3	−1578.78	5.06	−1577.66	2.82
4	−1306.97	44.94*	−1285.52	2.05
	a	b	a	b
1	1.074	0.718	1.072	0.706
2	1.401	0.808	1.379	0.736
3	1.228	0.717	1.225	0.716
4	1.509	0.955	1.395	0.887

*$p < .05$

Table 3
Total Computer Costs and Other Relevant Information

	RSCORE*	STEPIT	DSTEPIT	MAX
Total Cost**	$0.90	$1.27	$1.48	$1.25
Compute Time (Sec)	5.11	8.94	13.27	14.79
Core† (Bytes)	44K	50K	54K	36K

*Includes computation of least-squares initial estimates and variance-covariance matrix.

**The rate for CPU time is $125 per hour, and the rate for core use in Link-Edit and Go Step is $0.25 per megabyte-second (IBM 360/65).

†High water mark.

Table 3 presents some information relevant to questions of efficiency.

With respect to the cost associated with this example, the method of scoring, RSCORE, was over 30% less expensive than MAX, the second least expensive program. It is much less expensive than the other programs in spite of the fact that RSCORE computed the least-squares estimates used as the starting point for STEPIT, DSTEPIT, and MAX, and also computes the variance-covariance matrix. The lower cost is primarily a result of the very rapid compute-time for the method of scoring. In summary, this investigation reveals that the method of scoring (Rao, 1965) is a very accurate and very efficient procedure for parameter estimation with rating-method data, and supports its use by Ogilvie & Creelman (1968) and Dorfman & Alf (1968, 1969). Finally, it is also of interest that the direct-search program, MAX, a double-precision program, was substantially less expensive than DSTEPIT, the double-precision STEPIT. MAX also required substantially less core storage than STEPIT.

REFERENCES

Chandler, J. P. Subroutine STEPIT: An algorithm that finds the values of the parameters which minimize a given continuous function. A copyrighted program. J. P. Chandler, Copyright, 1965.

Dorfman, D. D., & Alf, E., Jr. Maximum-likelihood estimation of parameters of signal-detection theory—A direct solution. Psychometrika, 1968, 33, 117-124.

Dorfman, D. D., & Alf, E., Jr. Maximum-likelihood estimation of parameters of signal-detection theory and determination of confidence intervals—Rating method data. Journal of Mathematical Psychology, 1969, 6, 487-496.

Grey, D. R., & Morgan, B. J. T. Some aspects of ROC curve fitting: Normal and logistic models. Journal of Mathematical Psychology, 1972, 9, 128-139.

Hooke, R., & Jeeves, T. A. Direct search solution of numerical and statistical problems. Journal of the Association for Computer Machinery, 1961, 8, 212-229.

Ogilvie, J. C., & Creelman, C. D. Maximum-likelihood estimation of receiver operating characteristic curve parameters. Journal of Mathematical Psychology, 1968, 5, 377-391.

Rao, C. R. Linear statistical inference and its applications. New York: Wiley, 1965.

Swets, J. A., Tanner, W. P., Jr., & Birdsall, T. G. The evidence for a decision-making theory of visual detection. Electronic Defense Group Technical Report No. 40, University of Michigan, Ann Arbor, 1955.