5. STATISTICAL ANALYSIS

Chaired by Drake R. Bradley, *Bates College*

SDT_SP, a program in Pascal for computing parameters and significance tests from several detection theory designs

JOSE M. REALES and SOLEDAD BALLESTEROS
Universidad Nacional de Educación a Distancia, Madrid, Spain

SDT_SP is a program written in Pascal that runs in IBM-compatible computers and calculates descriptive as well as inferential statistics for several experimental designs based on signal detection theory and Luce's choice theory (e.g., one-interval designs, rating designs, two-alternative forced choice [2AFC] designs and *m*AFC designs). The program also generates receiver operating characteristic (ROC) curves from individual as well as group *d*'s. ROC curve plottings can be obtained merely by pressing a key on the computer keyboard. The program is accurate, menu driven, fast, and easy to use. It is available in English and in Spanish.

Detection theory is a general, widely accepted approach to the study of human performance in many experimental, as well as clinical, settings. Researchers in perception (e.g., visual, haptic, auditory), recognition memory, and cognition often ask experimental subjects to discriminate among stimuli. Their performance in these experiments is then analyzed according to the detection theory approach.

Classical psychophysical methods confounded sensitivity parameters and response bias. An important advantage of signal detection theory (SDT; Green & Swets, 1966) and choice theory (Luce, 1963) is that they provide measures of sensitivity that are unaffected by bias toward a particular response (MacMillan & Kaplan, 1985).

SDT, as well as Luce's choice theory, allows computation and interpretation of sensitivity, as well as bias, parameters from a number of experimental designs (one-interval designs, rating designs, two-alternative forced choice [2AFC] designs and *m*AFC designs) in many areas of research. Students of perception and memory who often use detection theory methodology may find it useful to have a comprehensive software program that enables them to speed up their calculations.

This work was supported by a UNED Doctoral Fellowship to J.M.R. This computer program originated in the context of haptic perception and memory research supported by DGICYT Grant PB90-0003 and by UNED Research Grants (1991, 1992, 1993) to S.B. Correspondence should be addressed to: S. Ballesteros, Departamento de Psicología Básica II, Universidad Nacional de Educación a Distancia, 28040 Madrid, Spain.

Overview

SDT_SP is written in Pascal 6.0. The program not only computes a series of descriptive statistics from a number of different SDT designs and Luce's choice theory but also performs inferential statistical tests, providing strategies for generalizing from the experimental data. These tests can be used to answer questions about the goodness of an estimate and about whether the parameter values differ from 0 or from one another. Receiver operating characteristic (ROC) curve plottings from individual as well as from group *d*' can be obtained merely by pressing a key on the computer keyboard. Appendix A presents a listing of the interface section of a unit in SDT_SP.

The program is accurate, menu driven, easy to use, and quick. Table 1 (top) compares results from yes/no experiments obtained with the use of (1) SDT_SP (leftmost column), (2) an independent statistical program (provided by Boice and Gardner, 1988) that generates parametric as well as nonparametric signal detection parameters (middle column), and (3) MacMillan and Creelman's (1991, p. 8) textbook, *Signal detection theory: A user's guide* (rightmost column). The second part of Table 1 compares (1) parameters from a rating experiment design computed with SDT_SP with (2) results from examples in MacMillan and Creelman (1991, p. 60). The third part of Table 1 corresponds to a nonparametric analysis (MacMillan & Creelman, 1991, p. 364). Finally, the last part of Table 1 presents results from significance tests for two-parameter as well as for *k*-parameter signal detection hy-

Table 1
Comparison of Results from SDT_SP and Other Sources

Measure	SDT_SP	Boice & Gardner (1988) Software	MacMillan & Creelman (1991)	Marascuilo (1970, pp. 240-243)
Yes/No Experiments				
d'	1.095	1.09497	1.095	
β	0.725	0.27465		
Log_s		0.32207		
A'	0.792	0.79167		
B'_H	-0.333	-0.33333		
α	2.449		2.45	
$\log \alpha$	0.896		0.90	
Rating Experiments				
d'_1	0.905		0.916	
d'_2	0.875		0.874	
d'_3	0.914		0.913	
d'_4	0.904		0.902	
d'_5	1.041		1.046	
s	0.560		0.55	
	0.716		0.73	
A_s	0.894		0.90	
	0.773		0.76	
c_1	1.137		1.145	
	0.369		0.372	
	-0.811		-0.817	
	-1.455		-1.468	
c_2	0.637		0.630	
	0.207		0.204	
	-0.454		-0.449	
	-0.815		-0.807	
c_a	0.786		0.780	
	0.255		0.253	
	-0.560		-0.557	
	-1.005		-1.000	
c_e	0.817		0.813	
	0.265		0.264	
	-0.582		-0.580	
	-1.045		-1.042	
Nonparametric Analysis				
$P(c)$	0.70		0.70	
A'	0.792		0.79	
q	0.667		0.667	
B'_H	-0.571		-0.571	
B''	-0.400		-0.400	
Significance Tests for Two-Parameter Signal Detection Hypotheses				
Z	2.058		2.06	
d_1-d_2	0.500		0.50	
[,]	[0.024, 0.973]		[0.03, 0.97]	
Significance Tests for k -Parameter Signal Detection Hypotheses				
χ^2	7.81		7.82	
U_0	20.5361		20.54	
[,]	[-0.1801, 1.1801]		[-0.18, 1.18]	
[,]	[-0.900, 0.4601]		[-0.90, 0.46]	
[,]	[0.0117, 1.4283]		[0.01, 1.43]	
[,]	[-1.3461, -0.0939]		[-1.35, -0.09]	
[,]	[-0.4367, 0.8767]		[-0.44, 0.88]	
[,]	[0.2833, 1.5967]		[0.28, 1.60]	

Note—The different sections of this table compare results obtained with this program with results obtained by another software program (when applicable), in a textbook, and in an article on different experimental designs and theoretical assumptions.

potheses computed with SDT_SP, in comparison with results reported by Marascuilo (1970, pp. 240-243). As can be seen from all the comparisons, the program's computations are accurate.

Motivation to Prepare the Program

A number of computer programs have been developed in recent years to calculate several SDT parameters. Usually these programs compute SDT indexes from yes/no experiments (e.g., Gardner & Boice, 1986; MacMillan & Creelman, 1991, Appendix 6) and nonparametric indexes (Boice & Gardner, 1988). Other programs provide algorithms for maximum-likelihood ROC-fitting procedures for choice theory (Ogilvie & Creelman, 1968) and for SDT (Dorfman & Alf, 1969).

SDT_SP originated in our own research on haptic discrimination of bilateral symmetry (Ballesteros, 1993; Ballesteros & Reales, 1992; Ballesteros, Manga, & Reales, 1994)—we needed to compute a number of descriptive and inferential STD parameters fast and accurately. Our starting point was the listing provided by MacMillan and Creelman (1991, Appendix 6) for computing basic (d' , c , and β) STD parameters. Our main interest now is to provide researchers in the SDT area with a general program that accurately computes parameters from a broad number of experimental designs.

SDT_SP also permits the presentation of ROC curves from yes/no designs and rating experiments on the computer screen and plots the ROC curves without one's having to change to a drawing package. The program runs on IBM-compatible personal computers. It is menu driven and very fast, and it checks for errors in data input.

The outcome of a discrimination experiment can be expressed by either of two pairs of parameters: hits and false alarm rates, or sensitivity and bias. However, a new feature provided by SDT_SP is the ability to reverse the process of calculating hits and false alarms from SDT parameters as well as from choice theory.

Several menus can be selected from the computer screen by just entering the option number. Researchers in detection theory can choose among the options provided by the menus according to the experimental design and the theoretical assumptions.

Experimental Designs That Can Be Analyzed With This Program

SDT_SP computes parameters from the most common designs used in the field of detection theory. Figure 1 displays the main menu that appears on the computer screen when the program is called.

From this menu, one of the following options can be selected by pressing the appropriate option number: yes/no experiments, nonparametric analysis, rating experiments, or m AFC experiments.

Yes-No Designs

The first option of SDT_SP allows computation of parameters from yes/no designs (or one-interval designs).

```

Choose option number and <ENTER>

1.- Yes/No Experiments
2.- Nonparametric Analysis
3.- Rating Experiments
4.- mAFC Experiments
5.- Setup
6.- Exit

Enter option number:

```

Figure 1. A printout of the main SDT_SP menu, showing available options.

In these experiments, subjects are presented on each trial with a stimulus that can belong to one of two categories: noise (N) or signal + noise (SN). The subjects have to discriminate the category to which the stimulus belongs. A good way to characterize the observer's sensitivity is to compute a measure of the discrepancy between the hit rate and the false alarm rate. After entering the number of hits, false alarms, misses, and correct rejections obtained from the stimulus-response matrix of each subject in each experiment, the program generates accurately and immediately the following indexes from STD and choice theory.

1. d' , the most important SDT sensitivity measure defined in terms of the distance between the noise and the signal + noise distributions. The sensitivity measure d' is defined in terms of z :

$$d' = z(H) - z(F). \quad (1)$$

2. d' standard deviation and its confidence interval.

3. α .

4. $\text{Log } \alpha$ (from choice theory). $\text{Log } \alpha$ is calculated from

$$\alpha = [H(1-F)] / [(1-H)F]^{1/2}. \quad (2)$$

5. c (criterion), the basic SDT bias measure, defined as

$$c = -0.5[z(H) + z(F)]. \quad (3)$$

This index is recommended by MacMillan and Creelman (1991) as a better index of subject bias because it is statistically independent of sensitivity d' .

6. c 's standard deviation.

7. c' , the relative criterion location.

8. The likelihood ratio.

9. b , the choice theory's bias measure computed from

$$b = [(1-H)(1-F) / (HF)]^{1/2}. \quad (4)$$

10. b' , an analogue choice theory measure to c' .

11. β_L , the likelihood ratio, computed from

$$\beta_L = H(1-H) / [F(1-F)]. \quad (5)$$

The program also provides a second option that generates the type of information just discussed, using only the hit and false alarm rates. This gives a smaller number of indexes than the use of the whole confusion matrix as input does; however, we recommend the first option whenever that is possible.

Three options are available at this point: (1) to escape; (2) to present the ROC curve on the computer screen; or (3) to obtain an ASCII file of the results.

The ROC curve is a graph that represents the functional relation between the proportion of hits (to say "yes" when the signal + noise is present) and the proportion of false alarms (to say "yes" when the noise is present). This program allows the ROC graph corresponding to the input data to be displayed on the screen. In order to get ROC curve prints, a graphic manager, such as GRAPHYCS.COM from MS-DOS, should be loaded before the program. The researcher must also specify the type of printer to be used (see the DOS manual). When the curve appears on the screen, press ALT+PRINT SCREEN, or only PRINT, and a plotting of the ROC curve will be obtained.

Detection theory analysis of group data. SDT_SP allows the computation of d' and c averaged across subjects (d'_{avg} and c_{avg}). The formula provided by MacMillan and Kaplan (1985, Appendix A) was used to compare the variances of two statistics that estimate d' (or c) from the hit and false alarm rates of 2 subjects (subscripts a and b):

$$\begin{aligned} \text{var}(d'_{\text{avg}}) &= \text{var} [0.5(d'_a + d'_b)] \\ &= 0.25 [\text{var}(d'_a) + \text{var}(d'_b)], \end{aligned} \quad (6)$$

and, generalizing this formula to n number of cases,

$$= (1/n)^2 \sum \text{var}(d'_i). \quad (7)$$

Significance tests for one-parameter signal detection hypothesis. SDT_SP allows the significance test for hypothesis testing to be computed according to the procedure provided by Marascuilo (1970). This test can be used to estimate the following null hypotheses:

$$H_0 : d' = 0$$

$$H_0 : d'_1 = d'_2$$

$$H_{0k} : d'_1 = d'_2 = \dots = d'_k.$$

The same rationale was used to compute the significance test of c .

Nonparametric Analysis

The program provides indexes for each particular theory. From the single high-threshold theory, the program provides the adjusted hit rate q , where $[q = (H-F) / (1-F)]$, and the bias index u , which equals the false alarm rate.

From the high-threshold theory the program computes the proportion correct as a measure of sensitivity. The yes-rate index $[0.5(H+F)]$, and the error ratio (k) are computed as bias indices (see MacMillan & Creelman, 1991).

Several measures based on area under the ROC curve are computed. A' , an index of sensitivity for experiments in which only a single data point is obtained, is calculated with Grier's (1971) formula. Three bias indexes are provided: (1) B'_H , the Hodos index (Hodos, 1970); (2) B'' , the Grier index (Grier, 1971); and (3) B''_D , the Donaldson index (Donaldson, 1992), computed from

$$B''_D = \frac{(1-H)(1-FA) - HFA}{(1-H)(1-FA) + HFA} \quad (8)$$

Rating Experiments

In rating experiments, subjects are presented with just two events, and the number of possible responses to indicate the subject's level of confidence that one or the other stimuli is present are two or more.

Whenever this option is selected, the program asks for the number of possible responses. As an example, we entered the information shown in Table 2, which appeared in MacMillan and Creelman (1991, p. 60), from a complex learning experiment in which subjects were presented with a set of "grammatical" pseudowords and another set of "nongrammatical" words. In answer to the question, "How many possible responses?" we entered six as the number of possible responses. Then, one by one, from left to right and top to bottom, we introduced the data (49, 94, 75, . . . 113, 113, shown in Table 2). The program then produced the measures in Table 3. These indexes are (1) the $N-1$ d' estimated, (2) the $N-1$ criteria (both c and β), (3) the $N-1$ likelihood ratios, (4) the slope, and (5) the intercept. When the slope differs significantly from 1, SDT_SP produces the following statistics: (1) D_a , the length of the hypotenuse of the equilateral right triangle whose legs' length are D_{YN} ; (2) A_z , a performance index in terms of a proportion; and (3) d'_e , an index based on the arithmetic average of the standard deviations (see MacMillan & Creelman, 1991). The program does not accept an empty (zero) cell. When this occurs, users are prompted to pool data from adjacent cells (see McNicol, 1972).

A plotting of the ROC curve can be obtained with a keystroke. Figure 2 displays the ROC generated by the program corresponding to this example.

mAFC Experiments

The main feature of 2AFC designs is that both alternatives of response are always present in each trial sequentially or spatially. The mAFC designs are general cases

Table 2
Frequency of Each Response for Each Stimulus

Rules-Only Condition	Response					
	"Grammatical"			"Nongrammatical"		
	"3"	"2"	"1"	"1"	"2"	"3"
Grammatical	49	94	75	60	75	22
Nongrammatical	8	37	45	60	113	113

Note—From *Detection Theory: A User's Guide* (p. 60, Table 3.1), by N. A. MacMillan and C. D. Creelman, 1991, New York: Cambridge University Press. Copyright Cambridge University Press. Adapted by permission.

Table 3
Results Obtained From the SDT_SP Corresponding to the Data Presented in Table 2

d'	0.905	0.875	0.914	0.904	1.041
c	1.576	0.739	0.252	-0.196	-1.043
β	4.161	1.909	1.258	0.838	0.338
c_1	1.537	0.721	0.245	-0.191	-1.018
c_2	1.614	0.757	0.258	-0.200	-1.069
c_a	1.574	0.739	0.251	-0.195	-1.042
c_e	1.575	0.739	0.251	-0.196	-1.043

Note—Slope = 1.050; Intercept = 0.964.

of 2AFC experiments in which two stimulus classes are presented in each trial in a random order but the number of response alternatives is m . In this case, the geometrical space is m -dimensional, and the observer has to decide which distribution is closer to the observation just presented. This type of design is less prone to bias and allows measurement of smaller stimulus differences than do yes/no experiments.

Before running the mAFC option from the main menu (see Figure 1), one must generate an ASCII file. The first row of this file must have only a number corresponding to the confusion matrix order. The next rows correspond to the matrix data. Numbers in each row of the file must be separated by at least one empty space.

Conclusion

SDT_SP is a comprehensive program, written in Pascal Version 6.0, that runs in IBM-compatible computers. The program computes a number of indexes from SDT and choice theory. A CGA, an EGA, a VGA, or a Hercules-compatible monochrome adapter is necessary if one wishes to plot ROC curves. The program is available in English and in Spanish. It is menu driven, easy to use, quick, and accurate.

Figure 2. Receiver operating characteristic (ROC) curve provided by SDT_SP, corresponding to data from Table 2 on linear coordinates.

Future Development

We are currently developing an enhanced version of SDT_SP that will be still more comprehensive and functional. We plan to add computations for a number of designs, including *same-different*, ABX, and GOC (group operating characteristic) designs (Taylor, Boven, & Whitmore, 1991).

Availability

SDT_SP will be available from the Universidad Nacional de Educación a Distancia, 28040 Madrid, Spain.

REFERENCES

- BALLESTEROS, S. (May, 1993). *Haptic perception and types of memory for raised patterns and three-dimensional objects*. Paper presented at the International Conference on Object Representation in Visual and Haptic Systems, Madrid.
- BALLESTEROS, S., MANGA, D., & REALES, J. M. (1994). *Haptic discrimination of bilateral symmetry in raised unfamiliar displays*. Manuscript submitted for publication.
- BALLESTEROS, S., & REALES, J. M. (November, 1992). *The perception of symmetric and asymmetric patterns by touch and vision*. Paper presented at the meeting of the Psychonomic Society, St. Louis.
- BOICE, R., & GARDNER, R. M. (1988). A computer program to generate parametric and nonparametric signal-detection parameters. *Bulletin of the Psychonomic Society*, 26, 365-367.
- DONALDSON, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, 121, 275-277.
- DORFMAN, D. D., & ALF, E., JR. (1969). Maximum likelihood estimation of parameters of signal detection theory and determination of confidence interval-ratings-method data. *Journal of Mathematical Psychology*, 6, 487-496.
- GARDNER, R. M., & BOICE, R. (1986). A computer program to generate signal-detection theory values for sensitivity and response bias. *Behavior Research Methods, Instruments, & Computers*, 18, 54-56.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley. (Reprint 1974, Huntington, NY: Krieger)
- GRIER, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.
- HODOS, W. (1970). A nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, 75, 351-354.
- LUCE, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103-189). New York: Wiley.
- MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- MACMILLAN, N. A., & KAPLAN, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, 98, 185-199.
- MARASCUILO, L. A. (1970). Extensions of the significance test for one-parameter signal detection hypothesis. *Psychometrika*, 35, 237-243.
- McNICOL, D. (1972). *A primer of signal detection theory*. London: Allen & Unwin.
- OGILVIE, J. C., & CREELMAN, C. D. (1968). Maximum-likelihood estimation of receiver operating characteristic curve parameters. *Journal of Mathematical Psychology*, 5, 337-391.
- TAYLOR, A., BOVEN, R., & WHITMORE, J. (1991). Reduction of unique noise in the psychophysics of hearing by group operating characteristic analysis. *Psychological Bulletin*, 109, 133-146.

APPENDIX A

UNIT TDS; INTERFACE

```

PROCEDURE Inverse_Luce (alpha, beta          : real;
                        VAR hit, fa           : real);
PROCEDURE Inverse_TDS (alpha, beta          : real;
                        VAR hit, fa           : real);
FUNCTION FiMay( z1 : real)                   : real;
FUNCTION BsubR( F, Sensitivity: real)        : real;
FUNCTION Qf( H, F: real)                    : real;
FUNCTION Inverse_log_Odds( X: real)          : real;
FUNCTION Z2( p : real)                      : real;
FUNCTION Fi( p : real)                      : real;
FUNCTION Dprime( H, F: real)                : real;
FUNCTION Alpha ( H, F: real)                : real;
FUNCTION Log_Alpha ( H, F: real)            : real;
FUNCTION PCorrect ( Nhits, Nmisses, Nfa, Nrc: real) : real;
FUNCTION P_CorrectPrime ( Nhits, Nmisses, Nfa, Nrc: real) : real;
FUNCTION YesRate( H, F: real)              : real;
FUNCTION ErrorRatio ( H, F: real)          : real;
FUNCTION Aprime ( H, F: real)              : real;
FUNCTION BDoublePrimeD ( H, F: real)       : real;
FUNCTION BDoublePrime ( H, F: real)       : real;
FUNCTION BprimeH ( H, F: real)            : real;
FUNCTION Criterion_relative( H, F : real)   : real;
FUNCTION Likelihood_ratio_L ( H, F : real) : real;
FUNCTION Phi_Coefficient ( Nhits, Nmisses, Nfa, Nrc : real) : real;
FUNCTION Landa ( X, Mu : real)            : real;

```

IMPLEMENTATION

```

{ source code }
Begin
End.

```