An APL program for calculating signal detection indices

W. T. McGOWAN, III and J. B. APPEL
Behavioral Pharmacology Laboratory
University of South Carolina,
Columbia, South Carolina 29208

Since Green and Swets' *Signal Detection Theory and Psychophysics* (1966), numerous books and articles have developed statistical decision or signal detection theory (SDT), for example, Egan (1975), Egan and Clarke (1966), Green and Swets (1966/1974), McNicol (1972), and Swets (1973). In addition, the model has been applied to diverse areas of research: animal as well as human psychophysics, memory, personality, learning, medicine, and behavioral pharmacology (Aiken, 1969; Appel & Dykstra, 1977; Lloyd & Appel, 1976; Price, 1966; Swets, 1969; Terman, 1970).

In brief, SDT attempts to differentiate statistically between effects on discrimination of variables related to physical properties of the stimulus (sensitivity) and effects of variables related to characteristics of the response (criterion or bias). In this attempt, numerous parametric as well as nonparametric indices of both sensitivity and bias estimation have been developed (Dusoir, 1975; Green & Swets, 1966/1974; Grier, 1971; Hodos, 1970; Pollack, 1970; Pollack & Hsieh, 1969; Pollack & Norman, 1964). Of these, the most familiar parametric measures are d' (sensitivity) and β or $\ln\beta$ (bias). The most familiar nonparametric measures are A' or $P(I)$ (sensitivity) and $B'H$ (bias).

In applications of SDT, a "yes-no" procedure is often employed. In this paradigm a subject reports either that a stimulus has (R+) or has not (R-) occurred, given that the stimulus has (S+) or has not (S-) been presented. All possible outcomes can be represented as four relative frequencies, that is, "HIT" = $fr(R+/S+)$, "MISS" = $fr(R-/S+)$, "False Alarm" (FA) = $fr(R+/S-)$, and "Correct Rejection" (CR) = $fr(R-/S-)$. From these frequencies, conditioned probabilities, that is, $p(R+/S+)$ and $p(R+/S-)$, can be determined. When sample size is not too large or the number of observations is limited, indices of both sensitivity and bias can be calculated in a reasonable amount of time; moreover, tables of both d' and β are available for some, but not all, conditional probabilities (Freeman, 1973). A serious problem arises when parametric indices are used and the data are inconsistent with the assumptions of the model, that is, "Thurstone's case V of underlying normal distribution

curves of equal variance" (Pollack, 1970, p. 155). Unfortunately, it is only *after* all data have been collected and evaluated that a test of the underlying assumptions can be made (Green & Swets, 1966/1974). If the assumptions are not met, an alternate model based on less stringent assumptions must be used (e.g., non-parametric), in which case additional time and energy must be spent recalculating estimates of both sensitivity and bias.

The purpose of this paper is to describe a computer program (APL) that can be used to compute either d' , β , $\ln\beta$ (Green & Swets, 1966/1974), or A' , $P(I)$, $B'H$ (Grier, 1971; Pollack, 1970; Pollack & Hsieh, 1969; Pollack & Norman, 1964). The program is written interactively for the "yes-no" SDT design. However, with simple modification of input data, it can be used with other SDT procedures. The program is intended for users who are only vaguely familiar with APL. It is presented in Table 1 and explained in detail below.

Function 1: BUILDMTX. BUILDMTX is a general-purpose function (fns) for building data sets. This fns can be used alone or in combination with the calling fns (SDT). If used alone, all that is necessary is to type the name of the fns (BUILDMTX) and answer the questions generated by the computer.

The first question is whether the data set is old or new. "New" indicates that no data exist, "old" indicates that a data set already exists and the user intends to add to the existing data set. If the response is "old," the user must type in the name of the data set; if "new," the number of columns desired. The number of rows (when the data set is being built) is user determined, that is, a response of "stop" indicates that input of data is completed. If too few or too many values are typed in any one row, an error message occurs and the line may be retyped.

If fns BUILDMTX is used to build an SDT data set, the variables are typed as multiple row vectors, with each row typed as (1) HIT, (2) MISS, (3) FA, (4) CR (Table 2, Examples 1 and 2).

Function 2: SDT. The purpose of the SDT function is to pull together all required fns. In SDT the user must type the name of the function (SDT), followed by a space, and, finally, a P in single quotes ('P') for parametric statistics or an N in single quotes ('N') for non-parametric statistics. The user has then only to answer the questions appropriately (Table 2, Example 3).

Function 3: SIGDET. Fns SIGDET calls either the fns NDTRI (below) for parametric or the fns NONPAR (below) for nonparametric statistics and organizes the estimated values in formatted output (Table 2, Example 4). The output can be modified easily by any user familiar with APL. For those not familiar with APL, Polivka and Pakin (1975) offer an excellent and comprehensive presentation.

Supported by USPHS Research Grants MH 24,593 from the National Institute of Mental Health and 9 R01 DA 01799 from the National Institute on Drug Abuse. We thank Mr. Charles Poole of the Computer Services Division, University of South Carolina, for technical advice and assistance.

Table 1
Program Functions

FUNCTION 1

```

    ▽ BUILDMTX;CT;INC;P;STOP;Z
[1]  'MATRIX...OLD OR NEW?'
[2]  →DN1×1'0'€N
[3]  'COLUMNS?'
[4]  XVAR←(1,(P+[]))p(CT+INC+0)
[5]  →DN2
[6]  DN1:'VARIABLE DATA SET NAME'
[7]  P←(CT+1)+p(XVAR+[])
[8]  INC←1+pXVAR
[9]  DN2:STOP+P09999
[10] '1)TYPE DATA SET [BY ROW], 2)WHEN COMPLETED TYPE STOP'
[11] UP1:;(INC+INC+1)
[12] UP2:Z+[]
[13] →DN4×1(((+/Z)+P)=9999)
[14] →DN3×1((pZ)≠P)
[15] XVAR←XVAR,[1] Z
[16] →UP1
[17] DN3:'***ERROR***'
[18] →UP2
[19] DN4:±(CT=0)/'XVAR+1 0+XVAR'
[20] 'DATA STORED UNDER VARIABLE XVAR'
    ▽

```

FUNCTION 2

```

    ▽ SDT N
[1]  BUILDMTX
[2]  N SIGDET XVAR
    ▽

```

FUNCTION 3

```

    ▽ N SIGDET XVAR;A;Δ;C;FX;I;IA;IT
[1]  'LEFT ARGUMENT = NONPARAMETRIC ['N'],OR PARAMETRIC ['P']
[2]  'RIGHT ARGUMENT = DATA MATRIX [ORDER: 1)HIT, 2)MISS, 3)FA, 4)CR]
[3]  'OBSERVATION VECTOR: 3 DIGITS MAXIMUM FORMAT
[4]  A←'''
[5]  Δ←'
[6]  →DN7×1((I>4)∨((I+1+pXVAR)<4))
[7]  C← 3 0 ∨((IT,1)p1+pFX←((IT+1+pXVAR),4)p0)
[8]  P1←(IT,1)p(XVAR[(IA+1)]+(+/((IT,2)+XVAR)))
[9]  P2←(IT,1)p(XVAR[3]+(+/(IT,2)+XVAR)))
[10] PC←((IT,1)p((XVAR[1]+XVAR[4])+(+/XVAR)))×100
[11] →DN1×1'P'€N
[12] NONPAR
[13] L←'NONPARAMETRIC'
[14] →DN5
[15] DN1:L←'PARAMETRIC'
[16] PA←P1
[17] UP1:I←0
[18] UP2:NDTRI PA[(I-I+1);1]
[19] FX[I;IA]+R[1]
[20] FX[I;(IA+2)]+R[2]
[21] →UP2×1I<IT
[22] →DN3×1(IA+IA+1)>2
[23] PA←P2
[24] →UP1
[25] DN3:D←(IT,1)p(FX[1]-FX[2])
[26] LB←OB←(IT,1)p(FX[3]+FX[4])
[27] DN5:'OUTPUT DESIRED?'
[28] →0×1'N'€N
[29] (((37pΔ),'SIGNAL DETECTION ANALYSIS'),(36pΔ),[1](((59pΔ),L),(18pΔ))
[30] 98p'')
[31] →DN6×1'N'€N
[32] (42pΔ),'PERCENT',(4pΔ),'P[HIT]',(5pΔ),'P[FA]',(7pΔ),'D',A,(7pΔ),'BETA',(6pΔ),'LOG'
[33] 'OBS',(4pΔ),'HIT',(3pΔ),'MISS',(5pΔ),'FA',(5pΔ),'CR',(11pΔ),'CORRECT',(45pΔ),'BETA'
[34] 98p'')
[35] C,(7 0 ∨XVAR),(18 3 ∨PC),(10 3 ∨P1),(10 3 ∨P2),(10 3 ∨D),(10 3 ∨B),(9 3 ∨LB)
[36] →0
[37] DN6:(42pΔ),'PERCENT',(4pΔ),'P[HIT]',(5pΔ),'P[FA]',(7pΔ),'A',A,(7pΔ),'PIJ',(6pΔ),'B',A,'H'
[38] 'OBS',(4pΔ),'HIT',(3pΔ),'MISS',(5pΔ),'FA',(5pΔ),'CR',(11pΔ),'CORRECT'
[39] 98p'')
[40] C,(7 0 ∨XVAR),(18 3 ∨PC),(10 3 ∨P1),(10 3 ∨P2),(10 3 ∨AP),(10 3 ∨PI),(9 3 ∨BH)
[41] →0
[42] DN7:'***ERROR*** DATA MATRIX NOT FORMED PROPERLY'
    ▽

```

Table 1 Continued

FUNCTION 4

```

▽ Q←NDTRI P;W
[1] R← 0 0
[2] +((0.5≤|P-0.5),0=|P-0.5)/L1,L2
[3] W←(●+(P|1-P)*2)*0.5
[4] L6:R[1]←(W-((+/(2.515517 0.802853 0.010328)*W*~1+13)+/(1 1.432788 0.189269 0.001308)*W*~1+14))*P-
0.5
[5] R[2]←0.3989423*-0.5*R[1]*R[1]
[6] →0
[7] L1:R[1]←(×P-0.5)×10*74
[8] →0
[9] L2:R[2]←0.3989423
▽

```

FUNCTION 5

```

▽ NONPAR;N;S;SN
[1] BH←((N+1+ρP1),(I+1))ρ0
[2] PI+1←(3-(2×(AP+0.5+(((P1-P2)×((1+P1)-P2))+((4×P1)×(1-P2))))))
[3] SN←P1+P2
[4] UP1:←((S=0),(S<1),((S+SN[I;1])>1))/DN3,DN2,DN1
[5] DN2:BH[I;1]←1-((P2[I;1]×(1-P2[I;1]))+(P1[I;1]×(1-P1[I;1])))
[6] →DN3
[7] DN1:BH[I;1]←((P1[I;1]×(1-P1[I;1]))+(P2[I;1]×(1-P2[I;1])))←1
[8] DN3:←UP1×1(I+I+1)≤N
▽

```

Table 2
Examples

Example 1

```

BUILDMTX
MATRIX○○○OLD OR NEW?
N
COLUMNS?
□: 4
1)TYPE DATA SET [BY ROW], 2)WHEN COMPLETED TYPE STOP
1
□: 60 40 30 70
2
□: 15 85 30 70
3
□: 90 10 20
***ERROR***
□: 90 10 20 80
4
□: STOP
DATA STORED UNDER VARIABLE XVAR

```

Example 2

```

BUILDMTX
MATRIX○○○OLD OR NEW?
O
VARIABLE DATA SET NAME
□: XVAR
1)TYPE DATA SET [BY ROW], 2)WHEN COMPLETED TYPE STOP
4
□: 95 15 85 100
5
□: 30 70 15 95 100
***ERROR***
□: 30 70 15 95
6
□: STOP
DATA STORED UNDER VARIABLE XVAR
XVAR
60 40 30 70
15 85 30 70

```

Table 2 Continued

```

          90  10  20  80
          95  15  85 100
          30  70  15  95

                                Example 3

                                SDT 'P'
MATRIX○○○OLD OR NEW?
N
COLUMNS?
□:      4
1)TYPE DATA SET [BY ROW], 2)WHEN COMPLETED TYPE STOP
1
□:      60  40  20  80
2
□:      70  30  30  70
3
□:      75  25  25  75
4
□:      75  25  60  40
5
□:      95   5  15  85
6
□:      STOP
DATA STORED UNDER VARIABLE XVAR
OUTPUT DESIRED?
Y

```

SIGNAL DETECTION ANALYSIS PARAMETRIC

OBS	HIT	MISS	FA	CR	PERCENT CORRECT	P[HIT]	P[FA]	D'	BETA	LOG BETA
1	60	40	20	80	70.000	.600	.200	1.094	1.380	.322
2	70	30	30	70	70.000	.700	.300	1.048	1.000	.000
3	75	25	25	75	75.000	.750	.250	1.348	1.000	.000
4	75	25	60	40	57.500	.750	.600	.421	.823	-.195
5	95	5	15	85	90.000	.950	.150	2.682	.442	-.816

Example 4

```

                                'N' SIGDET XVAR
                                OUTPUT DESIRED?
                                Y

```

SIGNAL DETECTION ANALYSIS NONPARAMETRIC

OBS	HIT	MISS	FA	CR	PERCENT CORRECT	P[HIT]	P[FA]	A'	P[I]	B'H
1	60	40	20	80	70.000	.600	.200	.792	.706	.333
2	70	30	30	70	70.000	.700	.300	.786	.700	.000
3	75	25	25	75	75.000	.750	.250	.833	.750	.000
4	75	25	60	40	57.500	.750	.600	.644	.584	-.219
5	95	5	15	85	90.000	.950	.150	.946	.902	-.627

Function 4: NDTRI. Fns NDTRI, the "heart" of the parametric SDT analysis was written by Snyder (1973). The fns estimates the ordinate (height of normal curve) and Z score (point on the normal curve corresponding to a given probability) for a standard normal distribution (mean of 0 and a standard deviation of 1).

Function 5: NONPAR. Fns NONPAR is based on the computational formulas described by Grier (1971). The fns NONPAR estimates sensitivity [A' and $P(I)$] and bias ($B'H$).

All functions, with the possible exception of fns NDTRI, were developed on an IBM 370/168 under APL/SV.

REFERENCES

- AIKEN, E. G. Auditory discrimination learning: Prototype storage and distinctive features detection mechanisms. *Perception & Psychophysics*, 1969, 6, 95-96.
- APPEL, J. B., & DYKSTRA, L. A. Drugs, discrimination, and signal detection. In T. Thompson & P. B. Dews (Eds.), *Advances in behavioral pharmacology*. New York: Academic Press, 1977. Pp. 139-166.
- DUSOIR, A. E. Treatments of bias in detection and recognition models: A review. *Perception & Psychophysics*, 1975, 17, 167-178.
- EGAN, J. P. *Signal detection theory and ROC analysis*. New York: Academic Press, 1975.
- EGAN, J. P., & CLARKE, F. R. Psychophysics and signal detection. In J. B. Sidowski (Ed.), *Experimental methods*

- and instrumentation in psychology. New York: McGraw-Hill, 1966. Pp. 211-242.
- FREEMAN, P. R. *Tables of d' and β* . Cambridge: Cambridge University Press, 1973.
- GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966. (Reprinted, New York: Krieger, 1974.)
- GRIER, J. B. Nonparametric indices for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 1971, **75**, 424-429.
- HODOS, W. A. A nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, 1970, **74**, 351-354.
- LLOYD, M. A., & APPEL, J. B. Signal detection theory and the psychophysics of pain: An introduction and review. *Psychosomatic Medicine*, 1976, **38**, 79-93.
- McNICOL, D. *A primer of signal detection theory*. London: George Allen & Unwin, 1972.
- POLIVKA, R. P., & PAKIN, S. *APL: The language and its usage*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.
- POLLACK, I. A nonparametric procedure for evaluation of true and false positives. *Behavior Research Methods & Instrumentation*, 1970, **2**, 155-156.
- POLLACK, I., & HSIEH, R. Sampling variability of the area under the ROC curve and of d' . *Psychological Bulletin*, 1969, **71**, 161-173.
- POLLACK, I., & NORMAN, D. A. A nonparametric analysis of recognition experiments. *Psychonomic Science*, 1964, **1**, 125-126.
- PRICE, R. H. Signal detection methods in personality and perception. *Psychological Bulletin*, 1966, **66**, 55-62.
- SNYDER, M. Interactive data analysis and non-parametric statistics. In P. Gjerlov, H. J. Helms, & J. Nielsen (Eds.), *APL Congress 73*. New York: American Elsevier, 1973.
- SWETS, J. A. Signal detection as a model of information retrieval. In F. Bresson, & M. deMontmollin (Eds.), *The simulation of human behavior*. Paris: Dunod, 1969. Pp. 253-264.
- SWETS, J. A. The relative operating characteristic in psychology. *Science*, 1973, **182**, 990-1000.
- TERMAN, M. Discrimination of auditory intensities by rats. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 145-160.

(Accepted for publication April 12, 1977.)