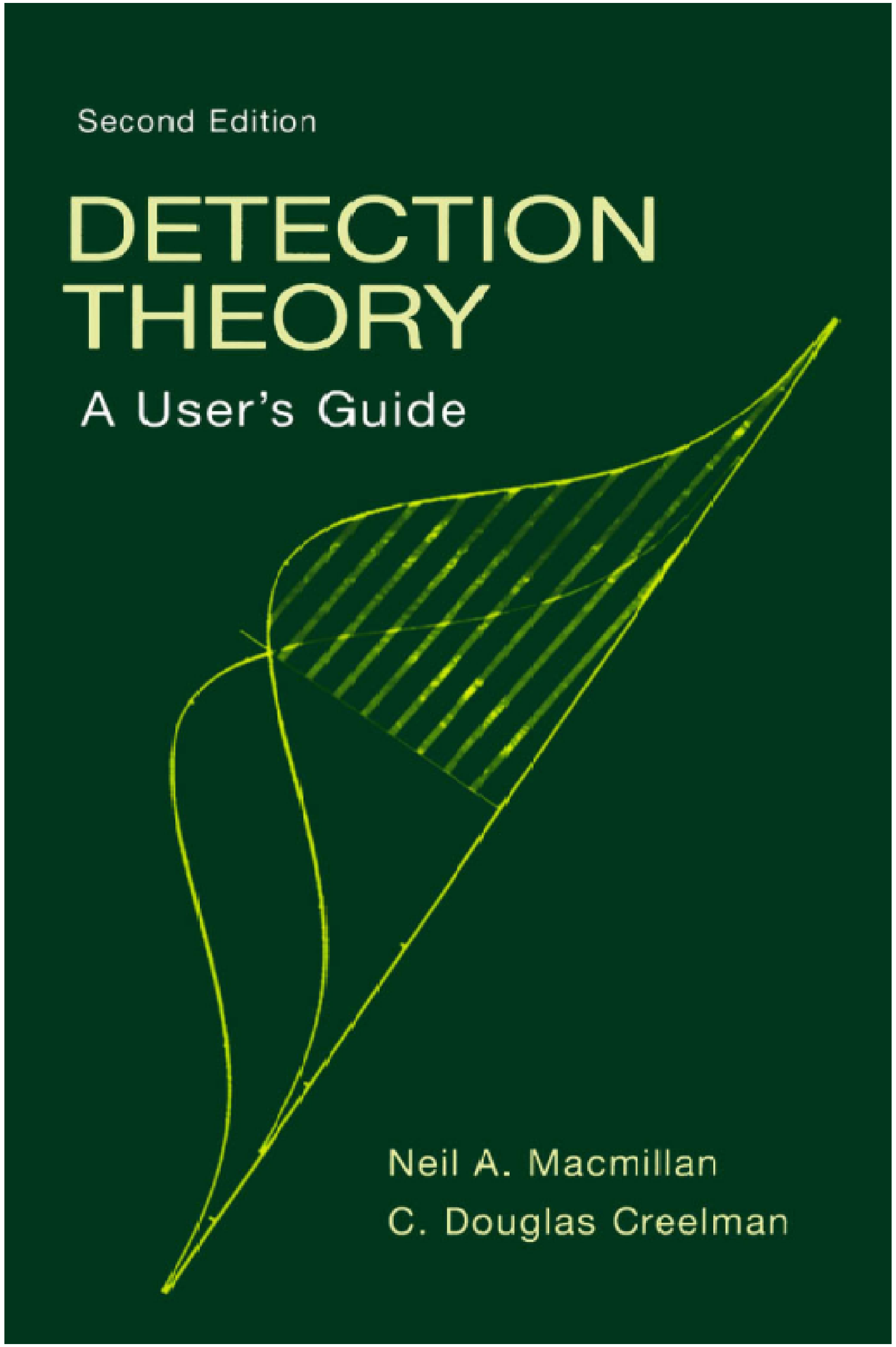

Implied ROCs

ROC Space and Isosensitivity Curves

What justifies the use of d' as a summary of discrimination? Why is this measure better, according to detection theory, than the more familiar $p(c)$? A good sensitivity measure should be invariant when factors other than sensitivity change. Participants are assumed by detection theory to have a fixed sensitivity when asked to discriminate a specific pair of stimulus classes. One aspect of responding that is up to them, however, is their willingness to respond "yes" rather than "no." If d' is an invariant measure of sensitivity, then a participant whose false-alarm and hit rates are (.4, .8) can also produce the performance pairs (.2, .6) and (.07, .35); all of these pairs indicate a d' of about 1.09, and differ only in response bias.

Excel

The locus of (false-alarm, hit) pairs yielding a constant d' is called an *isosensitivity* curve because all points on the curve have the same sensitivity.

This term was proposed by Luce (1963a) as more descriptive that the original engineering nomenclature *receiver operating characteristic* (ROC). Swets (1973) reinterpreted the acronym to mean *relative operating characteristic*. We use all these terms interchangeably.

Figure 1.1 shows ROCs implied by d'. The axes of the ROC are the false-alarm rate, on the horizontal axis, and the hit rate, plotted vertically. Because both H and F range from 0 to 1, the *ROC space*, the region in which ROCs must lie, is the unit square. For every value of the false-alarm rate, the plot shows the hit rate that would be obtained to yield a particular sensitivity level. Algebraically, these curves are calculated by solving Equation 1.5 for H; different curves represent different values of d'.

When performance is at chance ($d' = 0$), the ROC is the major diagonal, where the hit and false-alarm rates are equal. For this reason, the major diagonal is sometimes called the *chance line*. As sensitivity increases, the curves shift toward the upper left corner, where accuracy is perfect ($F = 0$ and $H = 1$). These ROC curves summarize the predictions of detection theory: If an observer in a discrimination experiment produces a (F, H) pair that lies on a particular implied ROC, that observer should be able to display any other (F, H) pair on the same curve.

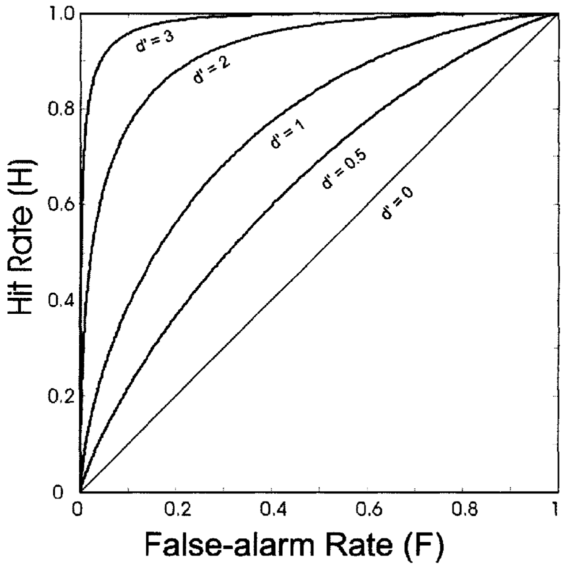

FIG. 1.1. ROCs for SDT on linear coordinates. Curves connect locations with constant d'.

The theoretical isosensitivity curves in Fig. 1.1 have two important characteristics. First, the price of complete success in recognizing one stimulus class is complete failure in recognizing the other. For example, to be perfectly correct with Old faces and have a hit rate of 1, it is also necessary to have a false-alarm rate of 1, indicating total failure to correctly reject New faces. Similarly, a false-alarm rate of 0 can be obtained only if the hit rate is 0. Isosensitivity curves that pass through (0, 0) and (1, 1) are called *regular* (Swets & Pickett, 1982).

Second, the slope of these curves decreases as the tendency to respond "yes" increases. The slope is the change in the hit rate, relative to the change in the false-alarm rate, that results from increasing response bias toward "yes." We shall see in a later section that this systematic slope change characterizes all ROCs.

ROCs in Transformed Coordinates

The features of regularity and decreasing slope are clear in Fig. 1.1, but other aspects of ROC shape are easier to see using a different representation of the ROC, one that takes advantage of our earlier description of a sensitivity measure as the difference between the transformed hit and false-alarm rates.

Look again at Equation 1.5, which describes the isosensitivity curve for d'. To find an algebraic expression for the ROC, we would need to solve this equation for H as a function of F. A simpler task is to solve for $z(H)$ as a function of $z(F)$:

$$z(H) = z(F) + d' \quad . \tag{1.8}$$

Equation 1.8 describes a *transformed ROC*, specifically a *z*ROC, in which both axes are marked off in equal z scores rather than in equal proportion units. The range of values in these new units is from minus to plus infinity, although scores of more than 2.5 (i.e., 2.5 standard deviations from the mean) are rarely encountered. In these coordinates, the ROC has a particularly simple shape: It is a straight line with unit slope, as shown in Fig. 1.2.

The linearity of *z*ROCs can be used to make a prediction about how much the false-alarm rate will go up if the hit rate increases (or vice versa). For example, suppose the false-alarm/hit pair (.2, .5) is on the ROC. Consulting Table A5.1, the z scores for F and H are –0.842 and 0. If we add the same number to each z score, the resulting scores correspond to another point on the ROC. Let us add 1.4, giving us the new z scores of 0.558 and 1.4. The table shows that the corresponding proportions are (.71, .92).

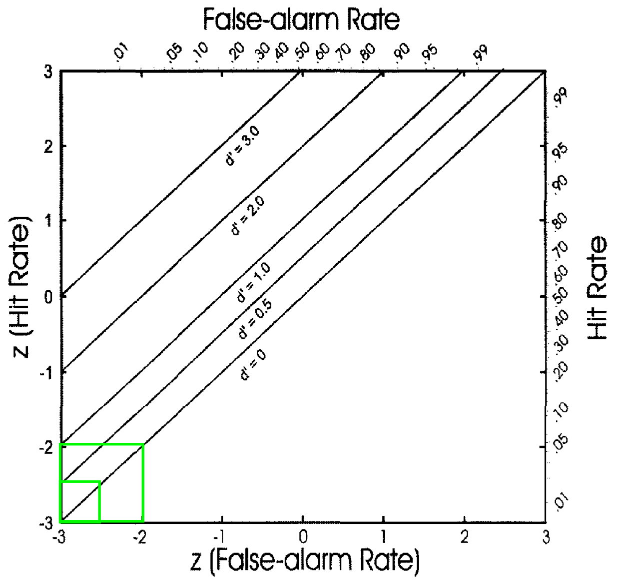

FIG. 1.2. ROCs for SDT on z coordinates.

The transformed ROC of Equation 1.8 provides a simple graphical interpretation of sensitivity: d' is the intercept of the straight-line ROC in Fig. 1.2, the vertical distance in z units from the ROC to the chance line at the point where $z(F) = 0$. In fact, because the ROC has slope 1, the distance between these two lines is the same no matter what the false-alarm rate is, and d' equals the vertical (or horizontal) distance between them at any point.

ROCs Implied by p(c)

Any sensitivity index has an implied ROC, that is, a curve in ROC space that connects points of equal sensitivity as measured by that index. To extend our comparison of d' with proportion correct, we now plot the ROC implied by $p(c)$. The trick is to take the definition of $p(c)$ in Equation 1.3 and solve it for H:

$$H = F + [2\,p(c) - 1]\ . \tag{1.9}$$

Equation 1.9 is a straight line of unit slope. Implied ROCs for $p(c)$ are shown in Fig. 1.3 for $p(c) = .5, .65$, and .8. The intercepts equal $2p(c) - 1$, that is, 0, .3, and .6.

H = 1*F + [2p(c)-1]; dove inclinazione = 1, intercetta = [2p(c)-1]

L'equazione lineare per p(c) si ha nelle coordinate ROC originali, non trasformate (e.g., Hit e False Alarm rates)

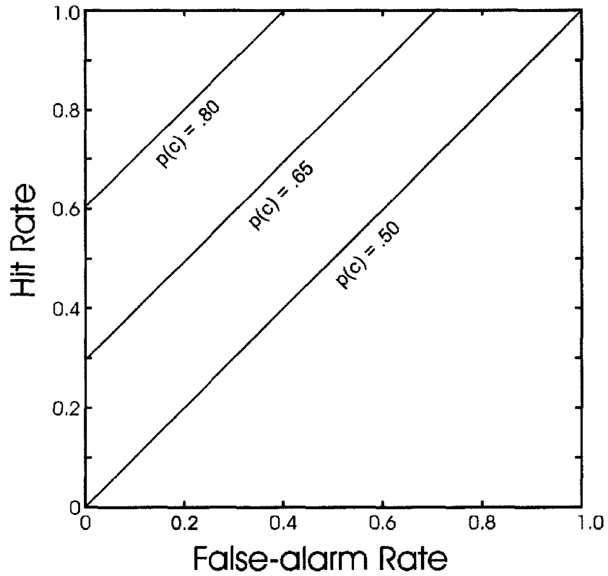

FIG. 1.3. ROCs implied by $p(c)$ on linear coordinates.

Not transformed

Consider again the false-alarm/hit pair (.2, .5). If we add the same number to each of these scores (without any transformation), the resulting scores correspond to another point on the ROC. Let us add .42, giving us the new hit and false-alarm proportions of (.62, .92). Simply using $p(c)$ as a measure of performance thus makes a prediction about how much the false-alarm rate will go up if the hit rate increases, and it is different from the prediction of detection theory.

Which Implied ROCs Are Correct?

The validity of detection theory clearly depends on whether the ROCs implied by d' describe the changes that occur in H and F when response bias is manipulated. Do empirical ROCs (the topic of chap. 3) look like those implied by d', those implied by $p(c)$, or something else entirely? It turns out that the detection theory curves do a much better job than those for $p(c)$. In early psychoacoustic research (Green & Swets, 1966) and subsequent work in many content areas (Swets, 1986a), ROCs were found to be regular, to have decreasing slope on linear coordinates, and to follow straight lines on z coordinates.

***Not transformed (%)

the perfect performance arises from statistical ("sampling") error. If, on the contrary, stimulus differences are so great that confusions are effectively impossible then the experiment suffers from a ceiling effect, and should be redesigned.

The Signal Detection Model

The question under discussion to this point has been how best to measure accuracy. We have defended d' on pragmatic grounds. It represents the difference between the transformed hit and false-alarm rates, and it provides a good description of the relation between H and F when response bias varies. Now we ask what our measures imply about the process by which discrimination (in our example, face recognition) takes place. How are items represented internally, and how does the participant make a decision about whether a particular item is Old or New?

Underlying Distributions and the Decision Space

Detection theory assumes that a participant in our memory experiment is judging a single attribute, which we call *familiarity*. Each stimulus presentation yields a value of this decision variable. Repeated presentations do not always lead to the same result, but generate a distribution of values. The first panel of Fig. 1.5 presents the probability distribution (or likelihood distribution, or probability density) of familiarity values for New faces (stimulus class S_1). Each value on the horizontal axis has some likelihood of arising from New stimuli, indicated on the ordinate. The probability that a value above the point k will occur is the proportion of area under the curve above k (see Appendix 1 for a review of probability concepts).

On the average, Old items are more familiar than New ones—otherwise, the participant would not be able to discriminate. Thus, the whole of the distribution of familiarity due to Old (S_2) stimuli, shown in the second panel, is displaced to the right of the New distribution. There must be at least some values of the decision variable that the participant finds ambiguous, that could have arisen either from an Old or a New face; otherwise performance would be perfect. The two distributions together comprise the *decision space*—the internal or *underlying* problem facing the observer. The participant can assess the familiarity value of the stimulus, but of course does not know which distribution led to that value. What is the best strategy for deciding on a response?

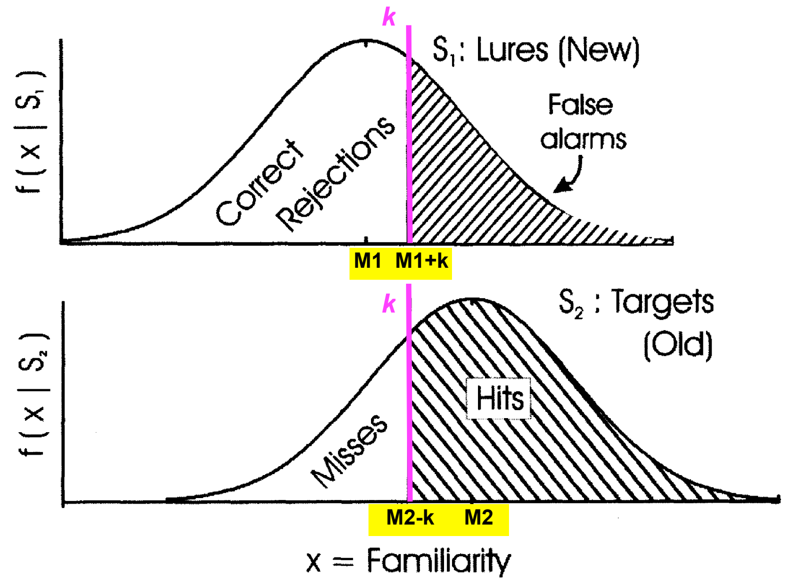

FIG. 1.5. Underlying distributions of familiarity for Old and New items. Top curve shows distribution due to New (S_1) items; values above the criterion k lead to false alarms, those below to correct rejections. Lower curve shows distribution due to Old (S_2) items; values above the criterion k lead to hits, those below to misses. The means of the distributions are M_1 and M_2. (In this and subsequent figures, the height of the probability density curve is denoted by f.)

Response Selection in the Decision Space

The optimal rule (see Green & Swets, 1966, ch. 1) is to establish a *criterion* that divides the familiarity dimension into two parts. Above the criterion, labeled k in Fig. 1.5, the participant responds "yes" (the face is familiar enough to be Old); below the criterion, a "no" is called for. The four possible stimulus–response events are represented in the figure. If a value above the criterion arises from the Old stimulus class, the participant responds "yes" and scores a hit. The hit rate H is the proportion of area under the Old curve that is above the criterion; the area to the left of the criterion is the proportion of misses. When New stimuli are presented (upper curve), a familiarity value above the criterion leads to a false alarm. The false-alarm rate is the proportion of area under the New curve to the right of the criterion, and the area to the left of the criterion equals the correct-rejection rate.

The decision space provides an interpretation of how ROCs are produced. The participant can change the proportion of "yes" responses, and generate different points on an ROC, by moving the criterion: If the criterion is raised, both H and F will decrease, whereas lowering the criterion will increase H and F.

We saw earlier that an important feature of ROCs is regularity: If $F = 0$, then $H = 0$; if $H = 1$, then $F = 1$. Examining Fig. 1.5, this implies that if the criterion is moved so far to the right as to be beyond the entire S_1 density (so that $F = 0$), it will be beyond the entire S_2 density as well (so that $H = 0$). The other half of the regularity condition is interpreted similarly. The distributions most often used satisfy this requirement by assuming that *any* value on the decision axis can arise from either distribution.

Sensitivity in the Decision Space

We have seen that k, the criterion value of familiarity, provides a natural interpretation of response bias. What aspect of the decision space reflects sensitivity? When sensitivity is high, Old and New items differ greatly in average familiarity, so the two distributions in the decision space have very different means. When sensitivity is low, the means of the two distributions are close together. Thus, the mean difference between the S_1 and S_2 distributions—the distance $M_2 - M_1$ in Fig. 1.5—is a measure of sensitivity. We shall soon see that this distance is in fact identical to d'.

Distance along a line, as in Fig. 1.5, can be measured from any zero point; so we measure mean distances relative to the criterion k. Thus expressed, the mean difference equals $(M_2 - k) - (M_1 - k)$: Sensitivity is the difference between these two distances, the distance from the S_1 mean to the criterion and the (negative, in this case) distance from the S_2 mean to the criterion. We now show that these two mean-to-criterion distances can be estimated using the z transformation discussed earlier in the chapter.

Underlying Distributions and Transformations

Figure 1.6 shows how the distances between the means of underlying distributions and the criterion are related to the response rates in our experiment. For each value of $M - k$, the figure shows the proportion of the area of an underlying distribution that is above the criterion. When $M - k = 0$, for example, the "yes" rate is 50%; large positive differences correspond to high "yes" rates and large negative differences to low ones. The curve in Fig. 1.6 is called a *(cumulative) distribution function*; in the language of calculus, it is the integral of the probability distributions shown in Fig. 1.5.

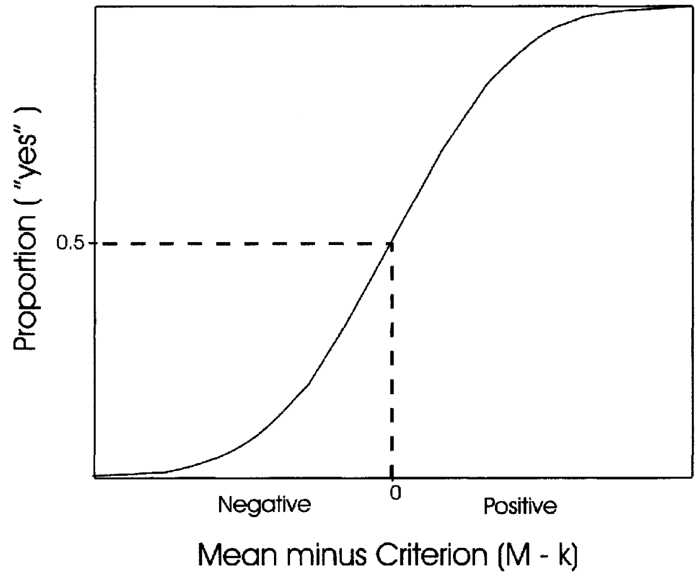

FIG. 1.6. A cumulative distribution function (the integral of one of the densities in Fig. 1.5) giving the proportion of "yes" responses as a function of the difference between the distribution mean and the criterion.

We can use the distribution function to translate any "yes" proportion into a value of $M - k$. This is the tie between the decision space and our sensitivity measures: For any hit rate and false-alarm rate (both "yes" proportions), we can use the distribution function to find two values of $M - k$ and subtract them to find the distance between the means. The distribution function transforms a distance into a proportion; we are interested in the inverse function, from proportions to distances, denoted z. In Fig. 1.7, the hit and false-alarm proportions from our face-recognition example are ordinate values, and the corresponding values $z(H)$ and $z(F)$ are abscissa values. The distance between these abscissa points, $z(H) - z(F)$, is the distance between the S_1 and S_2 means in Fig. 1.5. It is also, by Equation 1.5, equal to d'. Because z measures distance in standard deviation units, so does d'. Thus, the sensitivity measure d' is the distance between the means of the two underlying distributions in units of their common standard deviation.

The distance between the means of distributions is a congenial interpretation of d' because it is unchanged by response bias. No matter where the

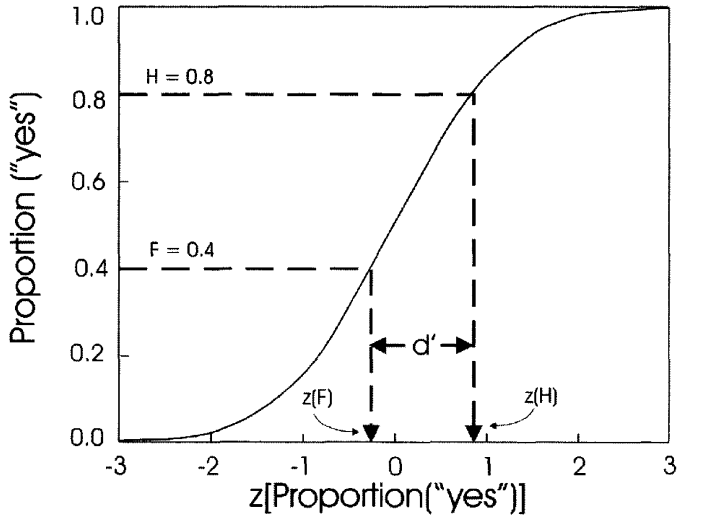

FIG. 1.7. A cumulative normal distribution function. The inverse function can be used to transform the proportions *H* and *F* into *z* scores, and sensitivity is the difference between *z*(*H*) and *z*(*F*).

participant locates the criterion, d' equals the same number. This relation is not specific to normal-distribution SDT: Any sensitivity measure obtained by subtracting transformed hit and false-alarm rates can be represented as the distance between the means of two distributions whose shape is given by the inverse of the transformation.

We can now venture a "definition" that will at least delimit the contents of this book. By *detection theory* we mean a theory relating choice behavior to a psychological decision space. An observer's choices are determined by the distances between distributions in the space due to different stimuli (sensitivities) and by the manner in which the space is partitioned to generate the possible responses (response biases).

Calculational Methods

Calculation of d' (and other statistics yet to be introduced) can be accomplished at several levels of technical sophistication. As we have seen, a table of the normal distribution is sufficient in principle. Computer programs have been developed specifically for this job and are much more convenient when the amount of data to be analyzed is large. Appendix 6 contains one

such program; it uses the most accurate algebraic approximation to z, according to Brophy (1985). A more complex program, which can also be used for the discrimination paradigms to be introduced later in the book, is *d'plus* (Macmillan & Creelman, 1997), which is available on the Internet.[3]

It is also easy to find d' using the "inverse normal" functions of spreadsheet programs; this is especially appealing for the many laboratories in which the data are collected or stored into spreadsheets. Basic calculations are illustrated in Table 1.1 for Excel, but are very similar in QuattroPro and other programs. The function z is written = NORMSINV. The indexes to be entered or computed are listed in Column A, and formulas are given that can be inserted in Rows 5 to 11 of Column B, then copied to subsequent columns. Sorkin (1999) explored the use of spreadsheets for SDT calculations in greater detail.

Detection theory procedures are also available as part of standard statistical packages such as Systat and SPSS. Because many users of detection theory make routine use of such packages, this is an attractive option. Data can be entered either as frequencies (number of hits, number of misses, etc.) or trial by trial, as they would be collected in an experiment. These packages can also be used when there are more than two response alternatives; we discuss them further in the context of rating designs (chap. 3).

TABLE 1.1 *Formulas for Spreadsheet (Excel) Calculation of SDT Statistics With Examples*

	A (Labels Only)	*Formula (for Column B; Then Copy to C and Other Columns)*	*B (Set 1)*	*C (Set 2)*
1	# hits		10	9
2	# misses		0	1
3	# false alarms		2	0
4	# correct rejections		8	10
5	*H* (hit rate)	= IF(B2>0, B1/(B1 + B2), (B1 – 0.5)/(B1 + B2))	.950	.900
6	*F* (false-alarm rate)	= IF(B3>0, B3/(B3+B4), 0.5/(B3+B4))	.200	.050
7	$z(H)$	= NORMSINV(B5)	1.645	1.282
8	$z(F)$	= NORMSINV(B6)	–0.842	–1.645
9	d'	= B7 – B8	2.486	2.926
10	c	= –0.5*(B7 + B8)	–0.402	0.182
11	β	=EXP(B9*B10)	0.368	1.702

[3]The site is http://psych.utoronto.ca/~creelman/.

1.6 (a). Suppose $d' = 1$. What is H if $F = .01, .1, .5$?
(b) Plot the ROC from these points on linear and z coordinates, and use the zROC to confirm the value of d'.

1.8. Are the points (.3, .9) and (.1, .7) on the same ROC according to detection theory (i.e., do they imply the same value of d')? Do they imply the same value of $p(c)$?

1.9. Suppose $(F, H) = (.2, .6)$. If F is unchanged, what would H have to be to double the participant's sensitivity, according to detection theory? If H is unchanged, what would F have to be?

1.11. Suppose a face-recognition experiment yields 20 hits and 10 false alarms in 45 trials. Can you compute d'? If not, is it possible to narrow down the possibilities? *Hint*: The stimulus–response matrix looks like this:

20		
10		
		45

What happens if there are 0 misses, or 0 correct rejections?