Perception & Psychophysics
1997, 59 (2), 313–316

Notes and Comment

Likelihood-ratio decision strategy for independent observations in the *same–different* task: An approximation to the detection-theoretic model

R. JOHN IRWIN and MICHAEL J. HAUTUS
University of Auckland, Auckland, New Zealand

An optimal decision strategy for deciding whether two things are the same or different is to adopt a likelihood-ratio criterion. The parametric equations for the receiver operating characteristic (ROC) based on the likelihood-ratio strategy when observations are independent are complicated; they require the numerical evaluation of a double integral. An approximation to the parametric equations for the likelihood-ratio strategy was developed. This approximation takes the form of a pair of equations that describe ROCs virtually indistinguishable from those of the full model.

The *same–different* psychophysical task has recently been the subject of considerable theoretical and experimental analysis (see, e.g., Dai, Versfeld, & Green, 1996; Hautus, Irwin, & Sutherland, 1994; Macmillan & Creelman, 1991). Part of the reason for this interest is that there are at least two different decision strategies that can be used to decide whether two stimuli are the same or different. One strategy is to use the difference between the two observations on a trial as the decision variable; Sorkin (1962) first provided a detection-theoretic model for the task on the basis of that decision variable. Later, however, Johnson (1980) and Noreen (1981) demonstrated that if the observations on a trial are assumed to be independent, a decision variable based on likelihood ratio provides optimal performance in the task.[1] Unfortunately, however, the parametric equations that define the receiver operating characteristic (ROC) for the likelihood-ratio strategy are not easy to evaluate because the equations for the hit and false-alarm rates entail a double integral with no closed-form solution. Here we provide a simple approximation to these equations, an approximation that contains only the well-known normal probability distribution function.

Assumptions and Terminology

We assume that the sensory effects of a stimulus can be represented by a normally distributed unidimensional random variable. The Gaussian assumption can be justified on several grounds (Green & Swets, 1966); it is especially appropriate for the standard experimental arrangement in the *same–different* task, in which two stimuli are presented separately (Laming, 1986). We also assume that the variances of the Gaussian distributions associated with the two stimuli are equal (again, see Laming, 1986, for the basis of this assumption), that their covariance is zero (in other words, that the two variables are independent), and that there is no systematic effect related to the two presentation intervals (Noreen, 1981). The discriminability of the two stimuli can be represented by the detection-theory index, d', the distance between the means of the Gaussian distributions in units of their common standard deviation.

We suppose that the observer adopts a strategy by implementing a decision rule. For the *same–different* task, only two responses are available ("same" and "different"), and the decision rule specifies how the information available from a particular observation should be classified in terms of those responses. In what follows, we adopt the notation of Macmillan, Kaplan, and Creelman (1977). In the *same–different* task, there are two classes of stimuli, A and B, and an instance from either class can be presented in two separate observation intervals that are contained in a trial.[2] If the instances are drawn from the same class in each interval, we use the notation <AA> or <BB> to represent this type of trial, and if they are drawn from different classes, we use the notation <AB> or <BA>. The stimulus configurations in these two types of trials are denoted S_1 for trials when the two stimuli are the same and S_2 for trials when they are different.

The Likelihood-Ratio Strategy

Noreen (1981) derived several decision rules in terms of likelihood ratio for the *same–different* task. For the standard experimental design in which S_1 and S_2 are presented with equal probability and each observation on a trial is independent, Noreen pointed out that the decision rule expressed in terms of likelihood ratio takes an intuitively plausible form: The observer should respond "same" if both observations on a trial were more likely to stem from A stimuli than from B stimuli, or if both were more likely to stem from B stimuli than from A stimuli; otherwise the observer should respond "different."

When the underlying probability densities are Gaussian, as assumed here, the likelihood ratio, $L(x)$, can be expressed in terms of the discriminability index, d', by making use of the relation $L(x) = e^{d'x}$ (see Green & Swets, 1966, Equation 3.3b). Noreen (1981) applied this relation to show that for the Gaussian case the likelihood ratio for the *same–different* task is given by

We thank Alan Lee for comments on early drafts of the manuscript, and Neil Macmillan, David Noreen, Donald Laming, and an anonymous reviewer for their advice and comments during the review process. Correspondence should be addressed to R. J. Irwin, Department of Psychology, University of Auckland, Private Bag 92019, Auckland, New Zealand (e-mail: j.irwin@auckland.ac.nz).

$$L(x) = \frac{e^{d'(x_1+x_2)} + 1}{e^{d'x_1} + e^{d'x_2}}. \tag{1}$$

A particular value of the likelihood ratio, β, can then be adopted to achieve some goal, such as maximizing the percentage of correct decisions. An optimal value of β, other things being equal, is equal to the ratio of the prior probability of presenting S_2 trials to the probability of presenting S_1 trials.

Figure 1 illustrates the decision space for the *same–different* task. Each axis represents the strength of evidence arising from each of the two observation intervals on a trial. The evidence over many trials for a particular stimulus sequence (e.g., <BA>) is bivariate-normal. The centroids of adjacent distributions are separated by d' and the concentric circles represent loci of equal probability density for each distribution. An example of a particular likelihood ratio is also shown. The equal likelihood-ratio contour for $\beta = 2$ is shown by the two solid curves, one in the upper right quadrant and one in the lower left quadrant.[3] An observer who adopted a criterion of $\beta = 2$ would call all pairs of observations falling beyond the solid contour in the upper right quadrant or below the solid contour in the lower left quadrant "same," and all other pairs of observations "different." The axes of Figure 1 represent yet another pair of likelihood ratios for the special case when $\beta = 1$. For this special case, any point falling in the upper right quadrant or the lower left quadrant would be judged as arising from the same stimuli, and any other point would be judged as arising from different stimuli.

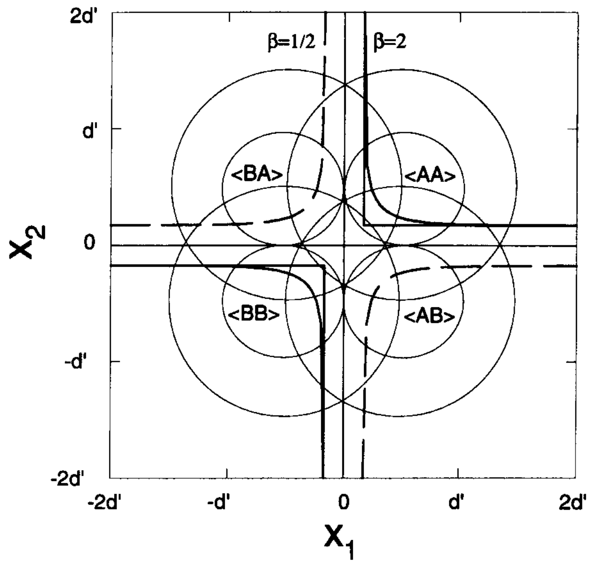

Figure 1. The centroid of a distribution is at the center of a set of concentric circles. The circles show loci of equal probability density for each distribution. The solid curves correspond to a likelihood ratio, β, of 2, and the broken curves to a likelihood ratio of ½. When $\beta = 2$, for example, observations falling beyond the criterion in the upper right quadrant or below the corresponding criterion in the lower left quadrant would be called "same"; other observations would be called "different." The solid lines represent an approximation to the likelihood ratio curves for $\beta = 2$.

In the example illustrated in Figure 1, for which $\beta > 1$, the hit rate for the likelihood-ratio criterion is the sum of two components: (1) the probability that an observation (x_1,x_2) falls beyond the criterion in the upper right quadrant, given an <AA> trial; and (2) the probability that an observation falls below the criterion in the lower left quadrant, given an <AA> trial.

If we denote the bivariate normal density function with unit variance for <AA> trials as $f_{AA}(x_1,x_2)$, then the hit rate (sum of component probabilities [1] and [2])[4] is given by the double integral of x_1 and x_2 over the region for which $L(x) > \beta$; this represents the volume under the <AA> density function that lies beyond the likelihood ratio, β (in both the upper right and lower left quadrants). To specify the limits of integration, substitute β for $L(x)$ in Equation 1 and solve for x_2. Next, note that the limit of x_2 as x_1 approaches infinity (or equivalently, for this case, of x_1 as x_2 approaches infinity) is

$$\lim_{x_1 \to \infty} x_2(x_1) = \lim_{x_1 \to \infty} \left(\frac{1}{d'} \ln\left(\frac{\beta e^{d'x_1} - 1}{e^{d'x_1} - \beta} \right) \right) = \frac{\ln(\beta)}{d'}, \tag{2}$$

so that $\ln(\beta)/d'$ is one limit of integration.[5] The other limit is given by $x_2(x_1)$; hence the hit rate is equal to

$$p(\text{"Same"} \mid \langle AA \rangle \vee \langle BB \rangle)$$
$$= \int_{\ln(\beta)/d'}^{\infty} \int_{x_2(x_1)}^{\infty} f_{AA}(x_1, x_2)\,dx_2\,dx_1$$
$$+ \int_{-\infty}^{-\ln(\beta)/d'} \int_{-\infty}^{x_2(x_1)} f_{AA}(x_1, x_2)\,dx_2\,dx_1. \tag{3}$$

Although Equation 3 can be simplified to some extent, a closed-form simplification is not available. Therefore, the hit rate for the likelihood-ratio criterion in the *same–different* task can only be evaluated numerically.

A further benefit of reflection symmetry is available for the calculation of the false-alarm rate: The volume beyond the contour in the upper right quadrant given <AB> is equal to that beyond the contour in the lower left quadrant given <AB>. Hence the false-alarm rate is given by

$$p(\text{"Same"} \mid \langle AB \rangle \vee \langle BA \rangle)$$
$$= 2 \int_{\ln(\beta)/d'}^{\infty} \int_{x_2(x_1)}^{\infty} f_{AB}(x_1, x_2)\,dx_2\,dx_1. \tag{4}$$

Again, however, a closed form simplification of Equation 4 is not available.

Although Equations 3 and 4 are unwieldy enough on their own, they are *still* insufficient to specify the hit rate and false-alarm rate for the likelihood-ratio decision strategy for any value of d' and β. This is because it is only when $\beta > 1$ that the contours of equal likelihood ratio lie in the upper right and lower left quadrants of Figure 1—the only case that we have considered so far. When $\beta < 1$, the contours lie in the lower right and upper left quadrants, as illustrated for $\beta = ½$ in Figure 1.

Fortunately, however, we can take advantage of some symmetrical properties of the relations among the various probabilities that need to be computed because likelihood ratios of β and $1/\beta$ are identical except for the quadrants in which they lie. Inspection of Figure 1 shows that the false-alarm rate for $\beta > 1$ is identical to the miss rate for a criterion of $1/\beta$, and so the hit rate for a criterion of $1/\beta$ (the complement of the miss rate for that criterion) is equal to the correct-rejection rate (the complement of the false-alarm rate) for a criterion of β. Similarly, the miss rate for $\beta > 1$ is equal to the false-alarm rate for a criterion of $1/\beta$.

An Approximation for the Likelihood-Ratio Strategy

We now present an approximation for the likelihood-ratio decision strategy. The approximation is analogous to one suggested by Nolte and Jaarsma (1967) for the optimal rule in the detection of one of m signals in Gaussian noise (see also Green & Birdsall, 1978). The approximation for $\beta = 2$ is illustrated by the solid straight lines in Figure 1: The approximation consists of two pairs of intersecting lines positioned at the asymptotic values of x_1 and x_2 assumed by the likelihood-ratio criterion.[6] This decision strategy, like that of the likelihood-ratio strategy, follows a *conjunctive* rule, in contrast to the disjunctive rule of Green and Birdsall. For this approximation to an optimal criterion, β, the observer accepts two sample presentations as the same if they both are less than $-\ln(\beta)/d'$ or if they both exceed $\ln(\beta)/d'$, which is the asymptotic value of x_1 as x_2 approaches infinity and of x_2 as x_1 approaches infinity for any value of β and d' (see Equation 2).

For the approximation, when $\beta > 1$ the optimal strategy for the observer is to favor the response "different" and, accordingly, respond "same" only if both stimuli are quite likely to be <A> (both x_1 and x_2 exceed $\ln(\beta)/d'$) or if both stimuli are quite likely to be <B> (both x_1 and x_2 are less than $-\ln(\beta)/d'$). For $\beta < 1$, the optimal strategy for the observer is to favor the response "same" and, accordingly, respond "different" only if one observation is quite likely to be a <B> while the other is quite likely to be an <A>. (Noreen, 1981, pointed out that it might be easier to think of the model as unidimensional with two criteria that set up three regions: "likely to be a <B>"; "ambiguous"; and "likely to be an <A>"; such models have been widely used in other information processing tasks, such as word recognition. From this point of view, for $\beta > 1$ the observer should respond "same" only to classifications of "AA" or "BB," and respond "different" otherwise. Conversely, when $\beta < 1$ the observer should respond "different" only to classifications of "AB" or "BA," and respond "same" otherwise.)

To compute the hit rate and the false-alarm rate for the approximation, we proceed as we did for the likelihood-ratio strategy, except that now the finite limit of integration is the same on each integral; that is, either $\ln(\beta)/d'$ or $-\ln(\beta)/d'$. The hit rate based on the approximation for $\beta > 1$ can be represented as

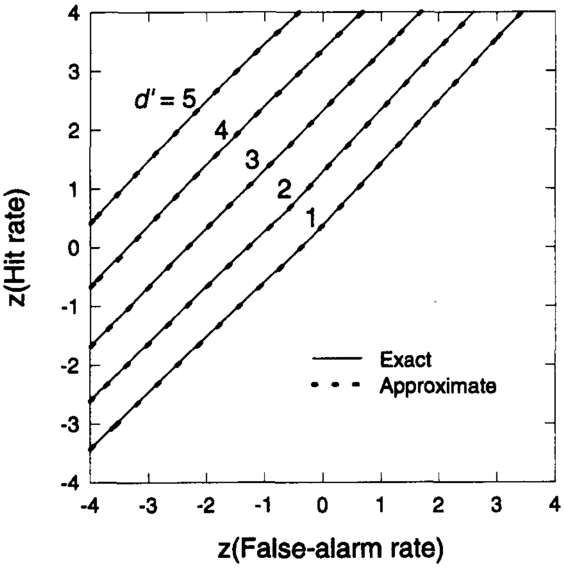

Figure 2. Receiver operating characteristics on z-coordinates for the likelihood-ratio decision strategy (solid lines) and its approximation (broken lines) for various values of d¢. The two sets of lines are barely discriminable in the figure for the values shown. Note that, although at a first glance the curves appear to be almost linear, close inspection reveals a degree of curvature that varies depending on the value of d¢.

$$p(\text{"Same"} \mid \langle AA\rangle \vee \langle BB\rangle) = \left[1 - \Phi\left(\frac{\ln(\beta)}{d'} - \frac{d'}{2}\right)\right]^2 + \left[1 - \Phi\left(\frac{\ln(\beta)}{d'} + \frac{d'}{2}\right)\right]^2, \tag{5}$$

where $\Phi(\cdot)$ is the normal probability distribution function.

Just as for the likelihood-ratio strategy, the total false-alarm rate for the approximation for the case when $\beta > 1$ is equal to twice the false-alarm rate for either stimulus pair <AB> or pair <BA>. The equation for the false-alarm rate for the approximation turns out to be

$$p(\text{"Same"} \mid \langle AB\rangle \vee \langle BA\rangle) = 2\left[1 - \Phi\left(\frac{\ln(\beta)}{d'} - \frac{d'}{2}\right)\right]\left[1 - \Phi\left(\frac{\ln(\beta)}{d'} + \frac{d'}{2}\right)\right]. \tag{6}$$

Again, Equations 5 and 6 are the parametric equations for half the ROC—that is, for the ROC when $\beta > 1$. Analogous reasoning to that used in determining the symmetry of the ROC for the likelihood-ratio strategy shows that, for the approximation, the ROC is also symmetric about the negative diagonal of the ROC square, so that a complete ROC can be constructed by using Equations 5 and 6 to calculate the ROC for $\beta > 1$. The obtained curve can then be reflected about the negative diagonal to obtain the curve for $\beta < 1$. Evaluating Equations 5 and 6 does not pose any special problems because the only complication in their numerical evaluation involves the well-known normal probability distribution function.

We are now in a position to compare the ROCs for the likelihood-ratio strategy with those of its approximation. For this comparison, we show the ROCs on z-transforms of the hit rate and false-alarm rate. As is well known, ROCs from Gaussian density functions are straight lines on these coordinates, and it turns out, as Macmillan and Creelman (1991) have stated and Dai et al. (1996) have confirmed, that *same–different* ROCs based on likelihood ratio are nearly straight lines with unit slope on these coordinates. Figure 2 depicts some ROCs for the likelihood-ratio strategy and its approximation for d' values of 1, 2, 3, 4, and 5. There are two ROCs illustrated in Figure 2 for each d' value. The approximation is uncannily close to the likelihood-ratio ROC—so much so that for the examples illustrated in Figure 2, it is virtually impossible to discern a difference between the two sets of ROCs. From this it can be concluded that the approximation is an excellent one.

Conclusion

Use of the approximation developed in this paper obviates the need to perform numerical evaluations of integrals (except the normal probability distribution function, for which there are efficient algorithms). Hence, the task of fitting the approximation to the model is fairly standard and not computationally intensive. The use of the approximation is justifiable because the approximation is an extremely good one over a wide range of parameter values.

REFERENCES

DAI, H., VERSFELD, N., & GREEN, D. (1996). The optimum decision rules in the *same–different* paradigm. *Perception & Psychophysics*, **58**, 1-9.

GREEN, D. M., & BIRDSALL, T. G. (1978). Detection and recognition. *Psychological Review*, **85**, 192-206.

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

HAUTUS, M. J., IRWIN, R. J., & SUTHERLAND, S. (1994). Asymmetry in the *same–different* ROC and the relativity of judgments about sound amplitude. *Quarterly Journal of Experimental Psychology*, **47A**, 1035-1045.

JOHNSON, K. O. (1980). Sensory discrimination: Decision process. *Journal of Neurophysiology*, **43**, 1771-1792.

LAMING, D. (1986). *Sensory analysis*. London: Academic Press.

MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

MACMILLAN, N. A., KAPLAN, H. L., & CREELMAN, C. D. (1977). The psychophysics of categorical perception. *Psychological Review*, **84**, 452-471.

NOLTE, L. W., & JAARSMA, D. (1967). More on the detection of one of M orthogonal signals. *Journal of the Acoustical Society of America*, **41**, 497-505.

NOREEN, D. L. (1981). Optimal decision rules for some common psychophysical paradigms. In S. Grossberg (Ed.), *Proceedings of the Symposium in Applied Mathematics of the American Mathematical Society and the Society for Industrial and Applied Mathematics* (Vol. 13, pp. 237-279). Providence: RI: American Mathematical Society.

SORKIN, R. D. (1962). Extension of the theory of signal detectability to matching procedures in psychoacoustics. *Journal of the Acoustical Society of America*, **34**, 1745-1751.

NOTES

1. More recently, Dai et al. (1996) have shown that the two previously suggested decision strategies for the *same–different* task lie at the extremes of a model, on the basis of likelihood ratio, that allows observations to be correlated. The model we discuss (and term the "likelihood-ratio model") corresponds to Dai et al.'s model for the extreme case when there is no correlation between observations.

2. The intervals can be separated in time or space, depending on the nature of the task.

3. The dashed curves show the likelihood-ratio contour for $\beta = \frac{1}{2}$, the reciprocal of $\beta = 2$. In general, the likelihood-ratio contour for $1/\beta$ is a reflection of that for β about the line $X_2 = 0$.

4. The hit rate for <AA> trials is equal to that for <BB> trials due to reflection symmetry about $X_2 = -X_1$ in the decision space. This permits the <BB> distribution to be neglected in the calculation of the hit rate.

5. The value $\ln(\beta)/d'$ is also the criterion adopted in the *yes–no* task when criterion location is defined in terms of the number of standard deviations from the equal-bias point (Noreen, 1981, p. 246, Decision Rule; Macmillan & Creelman, 1991, Equation 2.10).

6. This modification to the criterion has the effect of increasing both the false-alarm and hit rates by a small proportion when $\beta > 1$ and reducing these rates by a small proportion when $\beta < 1$. The effect is to move the point specified by a particular β on the receiver operating characteristic (ROC) closer to the minor diagonal. However, the point does not deviate appreciably from the ROC curve for the particular value of d'. When $\beta = 1$, the approximation and the likelihood-ratio model are equivalent.

(Manuscript received August 28, 1995;
revision accepted for publication April 22, 1996.)