Notes on the Decision Model.*

Part I:
Introduction,
One and Two Interval Experiments.

Louis D. Braida
and
Nathaniel I. Durlach

Sensory Communication Group
Research Laboratory of Electronics
and
Department of Electrical Engineering and Computer Science
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge, MA 02139

February 4, 2009

*©L. D. Braida and N. I. Durlach
1 Introduction

The purpose of these notes is to introduce a decision model that is used extensively in psychophysics. This model is a special case of the “Law of Categorical Judgment” developed by the psychologist Thurstone, and is closely related to the theory of ideal processing considered in the field of Communications Theory. Some of its important properties are:

- It is probabilistic and takes account of the empirical fact that a listener will not always respond in the same way to the same stimulus.

- It explicitly incorporates the effects of judgmental factors (related to the instructions and information given the listener by the experimenter and to various types of personal bias).

- It can be applied to a wide variety of experimental paradigms and be used to relate a wide variety of experimental results.

- It has had a strong influence on the design of experiments.

- It is frequently used in the literature (so that one often has a difficult time understanding the literature if one does not understand it).

The reader is referred to Appendix A for an introduction to random variables, to Appendix ?? for an introduction to the statistics of sampling, and to Appendix A for an introduction to Gaussian random variables.
3 One-Interval, 2AFC, Paradigm

Consider an experiment in which

- There are two admissible signal sources $S_1$ and $S_2$;
- There are two admissible responses $R_1$ and $R_2$;
- On each trial, the experimenter presents $S_1$ or $S_2$ randomly with a priori probabilities $\Pr(S_1)$ and $\Pr(S_2) = 1 - \Pr(S_1)$;
- The subject is instructed to respond $R_1$ when the signal arises from source $S_1$ and $R_2$ when the signal arises from source $S_2$;
- The experimenter “pays off” the subject for responding $R_j$ when the stimulus is $S_i$ with payoff $V_{ij}$, where $V_{ii} > 0$ (a reward) for $i = 1, 2$ and $V_{ij} \leq 0$ (a punishment) for $i \neq j$.

We will refer to this type of experiment as a one-interval, two-alternative-forced choice (2AFC), experiment. The term “one-interval” refers to the fact that on each trial the subject is presented with $S_1$ or $S_2$, but not both [as would be the case for example, in a two-interval experiment where the subject is required to distinguish between the two temporally ordered presentations $(S_1, S_2)$ and $(S_2, S_1)$].

The term “two-alternative-forced-choice” refers to the fact that there are only two allowable responses and that the subject is forced to choose one of them. (If he is uncertain as to whether the presentation is $S_1$ or $S_2$, he must guess). In many cases, the subject is informed after making his response whether or not his response was correct (or, equivalently, which of $S_1$ and $S_2$ was actually presented). In these cases, the experimental method is referred to as one-interval, 2AFC, plus feedback.

In a simple discrimination experiment, the sources are $S_1$ and $S_2$. In an absolute detection experiment, $S_1$ will be identified with the no signal, and $S_2$ with the signal to be detected. In a noise-masked detection experiment (where the task is to detect a signal $S$ in a background of noise $N$), $S_1$ will be identified with the noise $N$ and $S_2$ with the signal-plus-noise $S + N$.

If we assume that the trials of the experiment are statistically independent\(^3\), and the only response property considered is the identity of the response (i.e., no attention is given to properties such as response latency), then the results of such an experiment can be summarized by a $2 \times 2$ matrix whose entries $f_{ij}$ are the relative frequencies of responding $R_i$ to $S_j$:

$$f_{ij} = \frac{N(R_j | S_i)}{N(R_1 | S_i) + N(R_2 | S_i)} = \frac{N_{ij}}{N_{i1} + N_{i2}}$$

(10)

where $N(R_i | S_j)$ is the number of times the subject responded $R_i$ to $S_j$, and $N(R_1 | S_j) + N(R_2 | S_j)$ is the number of times $S_j$ was presented. Furthermore, since $f_{1j} + f_{2j} = 1$, the

\(^3\)The assumption of statistical independence is almost certainly not correct during training. The extent to which it is true during data collection remains to be explored.
matrix can be specified by any one of the four pairs

\[
\begin{align*}
[f (R_2 | S_2) , f (R_2 | S_1)] \\
[f (R_1 | S_2) , f (R_1 | S_1)] \\
[f (R_2 | S_2) , f (R_1 | S_1)] \\
[f (R_1 | S_2) , f (R_2 | S_1)].
\end{align*}
\]

For the sake of uniformity, we shall consistently use the pair \([f (R_2 | S_2) , f (R_2 | S_1)]\). Also, as a carry-over from the special case in which the task is one of detection, we shall call \(f (R_2 | S_2)\) the relative frequency of detection and \(f (R_2 | S_1)\) the relative frequency of false alarm.

The extent to which \(f (R_2 | S_2)\) and \(f (R_2 | S_1)\) have different values measures the extent to which the subject has demonstrated a sensitivity to the differences between \(S_2\) and \(S_1\). The extent to which both \(f (R_2 | S_2)\) and \(f (R_2 | S_1)\) are close to unity measures the extent to which the subject has demonstrated a bias to respond \(R_2\) rather than \(R_1\). Note, also, that if \(f (R_2 | S_2)\) is significantly less than \(f (R_2 | S_1)\), then the subject has demonstrated a sensitivity to the differences between \(S_2\) and \(S_1\), but has employed the wrong response coding.

### 3.1 Decision Model for the One-Interval, 2AFC, Paradigm

The axioms of the decision model for the One-Interval 2AFC paradigm are:

1. There exists a real random variable \(X\) (the “decision axis”) with the property that each signal presentation determines a value \(X\).

2. There exists a fixed cut-off value \(C\) (the “criterion”) on the \(X\) axis.

3. The subject responds \(R_1\) if and only if \(X < C\), and \(R_2\) if and only if \(X \geq C\).

4. The statistics of \(X\) are independent of all aspects of the experiment except \(S_1\) and \(S_2\), and are described completely by the conditional probability density functions \(p_X (X_0 | S_1)\) and \(p_X (X_0 | S_2)\). In particular, the statistics are independent of the \(a\ priori\) probabilities and payoffs, and the trials of the experiment are statistically independent.

The model is illustrated schematically in Fig. 5.

The conditional response \(\Pr (R_1 | S_j)\) are given by:

\[
\Pr (R_1 | S_1) = \int_{-\infty}^{C} p_X (X_0 | S_1) dX_0
\]
Figure 5: The decision model for one-interval two-alternative forced-choice experiments.

\[
\begin{align*}
\Pr (R_2 | S_1) &= \int_{-\infty}^{+\infty} p_X (X_0 | S_1) dX_0 \\
\Pr (R_1 | S_2) &= \int_{-\infty}^{C} p_X (X_0 | S_2) dX_0 \\
\Pr (R_2 | S_2) &= \int_{C}^{+\infty} p_X (X_0 | S_2) dX_0
\end{align*}
\]

(11) (12)

and are related to each other by

\[
\Pr (R_1 | S_j) + \Pr (R_2 | S_j) = 1
\]

for \( j = 1, 2 \).

As with the response frequencies \( f (R_i | S_j) \), we shall specify response probabilities by the pair \([\Pr (R_2 | S_2), \Pr (R_2 | S_1)]\) and refer to \( \Pr (R_2 | S_2) \) as the probability of detection, denoted \( P_D \) and to \( \Pr (R_2 | S_1) \) as the probability of false alarm, denoted \( P_F \):

\[
\begin{align*}
P_D &= \Pr (R_2 | S_2) \\
P_F &= \Pr (R_2 | S_1)
\end{align*}
\]

(13) (14)
According to this model, the sensitivity (or resolution) of the system is determined by the extent to which the densities $p_X (X_0 | S_1)$ and $p_X (X_0 | S_2)$ are nonoverlapping, and the bias of the system by the criterion $C$. In general, the value of $C$ is presumed to depend on the \textit{a priori} probabilities $\Pr (S_j)$ and the payoffs $V (R_i | S_j)$. For example, if the subject knows that $\Pr (S_1) \gg \Pr (S_2)$ and therefore expects $S_1$ to occur much more frequently than $S_2$ he will probably be biased toward responding $R_1$ rather than $R_2$ and choose a relatively large value of $C$. Similarly, if $v_{22} \gg v_{11}$ i.e., the reward for responding $R_2$ to $S_2$ is much greater than for responding $R_1$ to $S_1$, and $-v_{21} \gg -v_{12}$ (i.e., the punishment for responding $R_1$ to $S_2$, is is much greater than for responding $R_2$ to $S_1$) then the subject will probably be biased toward responding $R_2$ rather than $R_1$ and choose a relatively small value of $C$.

As a carry-over from the detection case (in which $S_2$ contains the signal to be detected), a large value of $C$ is often referred to as a “conservative” criterion and a small value as a “liberal” criterion. A more liberal criterion produces larger values of both $P_D$ and $P_F$.

In addition to the general features of the model mentioned in the introduction (in particular, that it is probabilistic and that it incorporates the effects of judgmental factors on performance), the specific axioms stated in this section lead to three further important features.

1. The influence of the signals on performance (characterized by the properties of the densities $(p_X (X_0 | S_1)$ and $p_X (X_0 | S_2)$ and the influence of the \textit{a priori} probabilities and payoffs on performance (characterized by the value of $C$) are completely separable. A change in the \textit{a priori} probabilities or payoffs does not produce a change in the densities.

2. On each trial, the influence of the signal presentation on the subject’s response is completely summarized by the value of a single real number. In particular, the decision space $X$ is assumed to be unidimensional.

3. The decision sets on the $X$ axis are composed of single intervals $X < C$ and $C \leq X$ (which are fixed throughout the experiment). Thus, for example, the axioms rule out the possibility that there exist two criteria $C_1$ and $C_2$ with $R_1$ corresponding to $C_1 \leq X < C_2$ and $R_2$ corresponding to $(X < C_1) \cup (C_2 \leq X)$, or the possibility that the criterion $C$, as well as the decision variable $X$, is a random variable and changes from trial to trial.

Finally, it is important to note that in order to develop the decision model into a complete theory of performance, it is necessary to develop a model of sensitivity [for relating the densities $p_X (X_0 | S_1)$ and $p_X (X_0 | S_2)$ to the physical properties of the stimuli and the subject’s auditory system], and a model of bias (for relating the criterion $C$ to the \textit{a priori} probabilities, the payoffs, and the subject’s decision-making apparatus). In these notes, we shall be concerned primarily with sensitivity and will ignore the problem of constructing a model for bias. Also, at this stage, we will ignore the difference between the response probabilities $\Pr (R_i | S_j)$ and the relative response frequencies $f (R_i | S_j)$ [which are assumed to converge to $\Pr (R_i | S_j)$ as the number of trials becomes infinite].
3.2 ROC for the One-Interval, 2AFC, Paradigm

Figure 6: ROC for one-interval 2AFC experiments. Note that the ROC passes through the points (0, 0) and (1, 1) by definition.

Consider now the equations relating the response probabilities $P_D$ and $P_F$ to the densities $p_X (X_0 | S_2)$ and $p_X (X_0 | S_1)$ and the criterion $C$:

\[
P_D = \int_C^{+\infty} p_X (X_0 | S_2) \, dX_0 \quad (15)
\]

\[
P_F = \int_C^{+\infty} p_X (X_0 | S_1) \, dX_0 \quad (16)
\]

According to these equations, if we know the densities and the criterion, we can determine $P_D$ and $P_F$ merely by integration. Suppose, however, we estimate $P_D$ experimentally [from $f (R_2 | S_2)$] and $P_F$ experimentally [from $f (R_2 | S_1)$] and want to determine the densities $p_X (X_0 | S_2)$ and $p_X (X_0 | S_1)$. In general, this is an unsolvable problem because the constraints on the densities imposed by assigning values to $P_D$ and $P_F$ are too weak. For example, one could choose any density whatsoever for $p_X (X_0 | S_2)$, then choose $C$ to satisfy the equation for $P_D$, and then choose $p_X (X_0 | S_1)$ to be any density for which the integral from $C$ to $\infty$
equals $P_F$. Suppose, however, that we perform a series or experiments that are identical except for the values of the a priori probabilities and/or the payoffs, and that these factors are varied in such a way that the subject changes his criterion $C$ between experiments. We can then obtain estimates of a sequence of pairs

$$[P_D(C_1), P_F(C_1)], [P_D(C_2), P_F(C_2)], [P_D(C_3), P_F(C_3)],$$

etc. If we could determine the functions $P_D(C)$ and $P_F(C)$, then we could determine the densities merely by differentiating Eqs. 15 and 16.

$$\frac{dP_D}{dC} = -p_X(C|S_2)$$
$$\frac{dP_F}{dC} = -p_X(C|S_1)$$

Unfortunately, however, although we can vary $C$ and measure $P_D$ and $P_F$ as $C$ is varied, we cannot determine the functions $P_D(C)$ and $P_F(C)$ because we have no way of determining the value of $C$. On the other hand, since the value of $C$ is the same for $P_D$ and $P_F$ (whatever it is), we can eliminate $C$ and determine $P_D$ as a function of $P_F$. This function, which is obtained empirically for a fixed pair of densities (i.e., fixed $S_2$ and $S_1$) by causing the subject to vary the criterion between experiments [and which is referred to as the “iso-sensitivity curve” or “receiver operating characteristic” (ROC)], summarizes all the knowledge about the densities that can be determined from the response probabilities $P_D(C)$ and $P_F(C)$.

The question of what properties of the densities are determined by the ROC is considered in Sec. 9. A schematic illustration of an ROC is shown in Fig. 6.

Since

$$\frac{dP_D}{dP_F} = \frac{p_X(C|S_2)}{p_X(C|S_1)} \geq 0 \quad (17)$$

and

$$P_D(-\infty) = P_F(-\infty) = 1 \quad (18)$$
$$P_D(+\infty) = P_F(+\infty) = 0 \quad (19)$$

the model predicts that all ROC’s are nondecreasing functions which pass through the points (0,0) and (1,1).
4 Other Paradigms

In the previous paragraphs, our discussion of the decision process has been restricted to the one-interval, 2AFC, paradigm outlined in Sec. 3, and the model we have constructed is limited to this paradigm only. With minor modifications, however, the model may be extended to a variety of other paradigms. In fact, as suggested in the introduction, one of the main advantages of the ideas that we have been considering is that they can be used to relate results from widely different paradigms. We will now extend the model to a “confidence-rating” paradigm and to a “two-interval” paradigm.

4.1 Confidence-Rating Paradigm

In the confidence-rating paradigm,

- There are two admissible signal sources $S_1$ and $S_2$;
- There are $M$ admissible responses $R_m$, $m = 1, \ldots, M$, (where $M$ is an even integer);
- On each trial, the experimenter presents $S_1$ or $S_2$ randomly with a priori probabilities $\Pr(S_1)$ and $\Pr(S_2) = 1 - \Pr(S_1)$;

![Figure 7: Decision model for confidence-rating experiments.](image)
The subject is instructed to use the responses $R_m$ not only to indicate which of $S_1$ or $S_2$ occurred, but also to express his/her degree of confidence in his decision (i.e., $R_1$ for very confident that $S_1$ occurred, $R_2$ for moderately confident that $S_1$ occurred, ..., $R_{M/2}$ for very slightly confident that $S_1$ occurred, $R_{1+M/2}$ for very slightly confident that $S_2$ occurred, ..., $R_{M-1}$ for moderately confident that $S_2$ occurred, $R_M$ for very confident that $S_2$ occurred).

The subject is instructed to use all available responses, to use them in a consistent manner, and to maximize the number of correct responses (where any of $R_1, \ldots, R_{M/2}$ is correct when $S_1$ is presented, and any of $R_{1+M/2}, \ldots, R_M$ is correct when $S_2$ is presented.)

In terms of the decision model, this paradigm is essentially the same as the one-interval, 2AFC, paradigm, except for the number of criteria. Whereas in the one-interval paradigm, the subject is required to maintain only one criterion at a time (and the ROC is obtained by conducting a series of experiments in which the criterion is varied between the experiments by changing the a priori probabilities and/or payoffs), in this paradigm, the subject is required to maintain many criteria simultaneously (and the ROC is obtained in a single experiment). Also, since the experimenter has no objective means for evaluating the relative correctness of the responses in the classes $R_1, \ldots, R_{M/2}$ and $R_{M/2+1}, \ldots, R_M$, the feedback in this paradigm is restricted to telling the subject whether his response was in the correct class (i.e., which of $S_1$ or $S_2$ was presented).

The axioms of the decision model for the One-Interval Confidence-Rating paradigm are:

1. There exists a real random variable $X$ with the property that each signal presentation determines a value of $X$.

2. There exist $M + 1$ criteria

   $$C_0 = -\infty < C_1 < \cdots < C_{M-1} < C_M = +\infty$$

   on the decision axis;

3. The subject responds $R_m$ if and only if $C_{m-1} \leq X < C_m$ for $m = 1, \ldots, M$.

4. The statistics of $X$ are independent of all aspects of the experiment except $S_1$ and $S_2$, and are described completely by the conditional probability density functions $p_X (X_0 | S_1)$ and $p_X (X_0 | S_2)$. In particular, the statistics are independent of the a priori probabilities and payoffs, and the trials of the experiment are statistically independent.

Note that the decision model for the Confidence-Rating paradigm, differs from that for the 2AFC paradigm only with respect to axioms (2) and (3). A schematic illustration of the model for the Confidence-Rating paradigm is shown in Fig. 7.
As before, the statistics of $X$ are assumed to depend only on $S_1$ and $S_2$ (so that they
should be identical to those for the one-interval, 2AFC, paradigm, provided only that $S_1$ and
$S_2$ are the same), and the criteria are assumed to be constant throughout the experiment.

The conditional response probabilities $\Pr(R_m | S_j)$ are given by

$$
\Pr(R_m | S_j) = \int_{c_{m-1}}^{c_m} p_X(X_0 | S_j) dX_0.
$$

These probabilities are often summarized in a table, such as that shown in Tbl. 3: The

<table>
<thead>
<tr>
<th></th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$\cdots$</th>
<th>$R_{M-1}$</th>
<th>$R_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>$P_{11}$</td>
<td>$P_{12}$</td>
<td>$\cdots$</td>
<td>$P_{1M-1}$</td>
<td>$P_{1M}$</td>
</tr>
<tr>
<td>$S_2$</td>
<td>$P_{21}$</td>
<td>$P_{22}$</td>
<td>$\cdots$</td>
<td>$P_{2M-1}$</td>
<td>$P_{2M}$</td>
</tr>
</tbody>
</table>

Table 3: The Stimulus-Response matrix for the confidence rating experiment.

The cumulative response probabilities $\Pr(R \geq R_m | S_j)$ are given by

$$
\Pr(R \geq R_m | S_j) = \int_{c_{m-1}}^{+\infty} p_X(X_0 | S_j) dX_0 = \sum_{k=m}^{M} \Pr(R_k | S_j)
$$

The ROC is obtained by plotting $\Pr(R \geq R_m | S_2)$ as a function of $\Pr(R \geq R_m | S_1)$. A
single point on the ROC is obtained for each of the values $m = 1, 2, \ldots, M$. If the subject
makes use of all the available responses (as instructed), then the experiment will produce
$M - 1$ distinct nontrivial points on the ROC. [The case $m = 1$ leads to the trivial point
$(1,1)$]. Since the density functions are assumed to be the same in this paradigm as in the
one-interval, 2AFC, paradigm, the ROC obtained from this paradigm is predicted to be
identical to that obtained form the one-interval, 2AFC, paradigm.

### 4.2 Two-Interval, 2AFC, Paradigm

In the two-interval paradigm,

- There are two admissible presentations, $U_1 = (S_2, S_1)$ and $U_2 = (S_1, S_2)$, each of which
  is a temporally-ordered pair.
- There are two admissible responses $R_1$ and $R_2$.
- On each trial, the experimenter presents $U_1$ or $U_2$ randomly with *a priori* probabilities
  $\Pr(U_1)$ and $\Pr(U_2) = 1 - \Pr(U_1)$.
- The subject is instructed to respond $R_1$ when $U_1$ is presented and $R_2$ when $U_2$ is
  presented.
The experimenter “pays off” the subject for responding $R_j$ when the stimulus is $U_i$ with payoff $V_{ij}$, where $V_{ii} > 0$ (a reward) for $i = 1, 2$ and $V_{ij} \leq 0$ (a punishment) for $i \neq j$.

The axioms of the decision model for the Two-Interval 2AFC paradigm are:

1. There exists a real random variable $X$ with the property that each signal presentation determines a value of $X$.

2. Each stimulus presentation (of the form $U_1 = (S_2, S_1)$ or $U_2 = (S_1, S_2)$, determines an ordered pair of values $X_1, X_2$ where $X_1$ is the value of $X$ determined from the first member of the pair ($S_2$ in the case of $U_1$, $S_1$ in the case of $U_2$), $X_2$ is the value of $X$ determined from the second member of the pair ($S_1$ in the case of $U_1$ and $S_2$ in the case of $U_2$), and $X_1$ and $X_2$ are statistically independent.

3. The subject observes the pair $(X_1, X_2)$ and forms the decision variable $Y = X_2 - X_1$.

4. There exists a fixed cut-off value $C$ (the criterion) on the $Y$ axis.

5. The subject responds $R_1$ if and only if $Y < C$, and $R_2$ if and only if $Y \geq C$. 

Figure 8: The decision model for two-interval 2AFC experiments.
6. The statistics of $X$ are independent of all aspects of the experiment except $S_1$ and $S_2$, and are described completely by the conditional probability density functions $p_X(X_0|S_1)$ and $p_X(X_0|S_2)$. In particular, the statistics are independent of the a priori probabilities and payoffs, and the trials of the experiment are statistically independent.

Note that the decision model for the Two-Interval 2AFC paradigm, differs from that for the One-Interval 2AFC paradigm only with respect to axioms (2) and (3) of that model. A schematic illustration of the model for the Two-Interval paradigm is shown in Fig. 8. In particular, since the statistics of $X$ are assumed to depend only on $S_1$ and $S_2$, they should be identical to those for the one-interval, 2AFC, paradigm and the confidence-rating paradigm, provided only that $S_1$ and $S_2$ are the same.

If we ignore, for the moment, the way in which $Y$ is generated from $X_1$ and $X_2$, this model is identical to the model for the one-interval, 2AFC, paradigm, and all the statements for that paradigm can be made valid for the present paradigm merely by substituting $Y$ for $X$, $U_1$ for $S_1$, and $U_2$ for $S_2$. Thus, for example, if one makes these substitutions, Equations 11, and 12) provide the expressions for the conditional response probabilities $Pr(R_i|U_j)$. Similarly, the ROC is obtained by plotting $P_D(C) = Pr(R_2|U_2)$ as a function of $P_F(C) = Pr(R_2|U_1)$, and the previous comments on the ROC are all equally valid for the ROC obtained from this paradigm.

If, on the other hand, we make use of the assumptions that $Y = X_2 - X_1$, that $X_1$ and $X_2$ are statistically independent, and that the density functions $p_X(X_0|S_1)$ and $p_X(X_0|S_2)$ are the same in this paradigm as in the previous ones, we can derive a number of further results. For example, because $Y = X_2 - X_1$, and $X_1$ and $X_2$ are statistically independent, the densities $p_Y(Y_0|U_j)$ are given by the convolutions:

$$p_Y(Y_0|U_1) = \int_{-\infty}^{+\infty} p_X(Y_0 + X_0|S_1) p_X(X_0|S_2) dX_0$$

$$p_Y(Y_0|U_2) = \int_{-\infty}^{+\infty} p_X(Y_0 + X_0|S_2) p_X(X_0|S_1) dX_0$$

Second, by substituting $X_1 = X_0 + Y_0$ in the above integrals, one can show that

$$p_Y(Y_0|U_2) = p_Y(-Y_0|U_1)$$

Thus, the densities $p_Y(Y_0|U_j)$ are symmetric about $Y_0 = 0$ as shown schematically in Fig. 8. This symmetry, in turn, implies that

$$P_D(C) = 1 - P_F(-C)$$

$$P_D(-C) = 1 - P_F(C)$$

so that the ROC is symmetric about the negative diagonal (as illustrated in Fig. 9).
Finally, making use of the assumption that $p_X (X_0 | S_j)$ is the same in the various paradigms, one can show that if the a priori probabilities are equal and the subject is “unbiased” in the two-interval paradigm, i.e. that $\Pr (U_1) = \Pr (U_2) = 1/2$ and $C = 0$, then the probability of making a correct response $Q$ in the two-interval paradigm is equal to the area under the ROC for the one-interval paradigm. Specifically,

$$Q = \Pr (U_1) \Pr (R_1 | U_1) + \Pr (U_2) \Pr (R_2 | U_2)$$

(27)

Thus

$$Q = \Pr (R_2 | U_2) = \Pr (Y \geq 0 | (S_1, S_2))$$

but

$$\Pr (Y \geq 0 | (S_1, S_2)) = \Pr (X_2 \geq X_1 | X_1 \text{ from } S_1 \text{ and } X_2 \text{ from } S_2)$$

so that

$$\Pr (Y \geq 0 | (S_1, S_2)) = \int_{-\infty}^{+\infty} p_X (X_1 | S_1) dX_1 \int_{X_1}^{+\infty} p_X (X_2 | S_2) dX_2$$

Thus

$$Q = \int_{-\infty}^{+\infty} P_D (X_1) \frac{-dP_F (X_1)}{dX_1} dX_1$$
\[
\int_{-\infty}^{+\infty} P_D (X_1) \, dP_F (X_1) \, dX_1
\]

By changing the variable of integration from \( X_1 \) to \( P_F \) one then has

\[
Q = \int_{0}^{1} P_D (P_F) \, dP_F \quad (28)
\]

Thus the probability of being correct in a two-interval two-alternative forced-choice experiment is equal to the area under the ROC curve for the same stimuli in a one-interval two-alternative forced-choice experiment. This result, which is known as Green’s Theorem\(^4\) is one of the few general theorems, not tied to assumptions about the form of the decision densities, relating performance in different paradigms.

\(^4\)After David M. Green, who first proved it.