A THURSTONE-COOMBS MODEL OF CONCURRENT RATINGS WITH SENSORY AND LIKING DIMENSIONS

F. GREGORY ASHBY

University of California at Santa Barbara

 DANIEL M. ENNIS

The Institute for Perception
7629 Hull Street, Suite 200
Richmond, Virginia 23225
The Medical College of Virginia and The University of Illinois

Received for Publication June 11, 2000

ABSTRACT

A popular product testing procedure is to obtain sensory intensity and liking ratings from the same consumers. Consumers are instructed to attend to the sensory attribute, such as sweetness, when generating their liking response. We propose a new model of this concurrent ratings task that conjoins a unidimensional Thurstonian model of the ratings on the sensory dimension with a probabilistic version of Coombs’ (1964) unfolding model for the liking dimension. The model assumes that the sensory characteristic of the product has a normal distribution over consumers. An individual consumer selects a sensory rating by comparing the perceived value on the sensory dimension to a set of criteria that partitions the axis into intervals. Each value on the rating scale is associated with a unique interval. To rate liking, the consumer imagines an ideal product, then computes the discrepancy or distance between the product as perceived by the consumer and this imagined ideal. A set of criteria are constructed on this discrepancy dimension that partition the axis into intervals. Each interval is associated with a unique liking rating. The ideal product is assumed to have a univariate normal distribution over consumers on the sensory attribute evaluated. The model is shown to account for 94.2% of the variance in a set of sample data and to fit this data significantly better than a bivariate normal model of the data (concurrent ratings, Thurstonian scaling, Coombs' unfolding model, sensory and liking ratings)

1 Corresponding author. Department of Psychology, University of California, Santa Barbara, CA 93106. TEL: 805-893-2130; E-mail: ashby@psych.ucsb.edu
INTRODUCTION

In the concurrent ratings task, a subject rates a product or stimulus simultaneously on a number of sensory dimensions or aspects (Hirsch, Hylton, & Graham, 1982; Olzak, 1986). The observed ratings are then used to estimate the subject’s sensory, perceptual, or cognitive impressions of the stimulus. If the rated dimension is sensory, the resulting data often can be modeled accurately by assuming that 1) the unobservable perceived values have a trial-by-trial (or subject-by-subject) univariate normal distribution across the relevant sensory dimension, 2) the subject establishes a set of criteria or cut-points on the dimension that partitions the dimension into intervals, and 3) a different numerical rating is assigned to each interval (Ashby, 1988; Olzak, 1986; Wickens & Olzak, 1992). On each trial, the subject determines in which interval the percept is in and then selects the associated rating.

We propose that hedonic responses are based on or derived from sensory information, and also depend on the affective state of the subject. For example, consider a group of products that vary along one sensory dimension and suppose a market segment of consumers are asked to rate the sensory magnitude of one of these products along the relevant dimension (using a 1 to n scale) and also to rate the degree to which they like the product on that dimension. In this case, we expect the sensory ratings to increase monotonically with sensory magnitude but we expect the liking ratings to increase with sensory magnitude until some ideal level is reached and thereafter to decrease (assuming higher ratings reflect greater liking). This non-monotonicity, which is characteristic of single-peaked preference functions, is commonly expected in cases where an hedonic response is based on the sensory characteristics of the stimulus. An alternative to this behavioral model for liking ratings is to treat the intensity and liking ratings as random variables from a bivariate normal distribution. In this type of model, liking and sensory attributes are treated similarly and the implication is that there is an hedonic intensity continuum very much like a sweetness intensity continuum.

Sample data from a concurrent ratings task are shown in Table 1. These data are from a real consumer study. If the bivariate normal ratings model is correct, the response frequencies should have a unimodal distribution across the 49 data cells that decreases in a fairly symmetric and continuous fashion. When liking is high, this prediction holds (“7” is the highest value of liking), but when liking is low, the prediction is strongly violated. For example, a liking rating of “2” was assigned to small sensory magnitudes and to large sensory magnitudes, but never to intermediate sensory magnitudes. The bivariate normal ratings model cannot account for this result. Thus, these data require a fundamentally different sort of model. This article develops such a model by conjoining a univariate
normal ratings model for a sensory dimension with a probabilistic version of Coombs’ (1964) unfolding model for hedonic dimensions.

TABLE 1.
SENSORY SWEETNESS INTENSITY AND LIKING RATINGS FOR A CONSUMER PRODUCT

<table>
<thead>
<tr>
<th>Liking</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>28</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>14</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Models of Concurrent Ratings and Preferences

The univariate normal ratings model is rooted in signal detection theory (Peterson, Birdsall, & Fox, 1954; Swets, Tanner, & Birdsall, 1961; Green & Swets, 1966) and in the seminal work of L.L. Thurstone (1927a, 1927b, 1927c). The most important assumption of the model is that there is trial-by-trial (or subject-by-subject) variability in the sensory, perceptual, and cognitive information obtained from every object or event (Ashby & Lee, 1993). Thus, the psychological effect of a stimulus is represented by a univariate probability distribution. The variability is assumed to arise from many sources. First, physical stimuli are themselves intrinsically variable. For example, with consumer products there is variability in ingredients and manufacturing processes. Second, there is perireceptor noise in all modalities. This is especially true with the chemical senses. For example, variability in the amount of saliva and its chemical composition causes variability in the number of stimulus...
molecules reaching the gustatory receptors. Similarly, variability in sniff strength and amount of nasal mucus causes variability in the number of odorant molecules reaching the olfactory receptors. Third, there is spontaneous activity at all levels of the central nervous system that introduces more noise. Finally, variability in the subject’s levels of hunger, thirst, arousal, motivation, and affect can introduce even more variability.

A second assumption of the univariate normal model is that the rating task can be separated into two components: sensory and decision. The output of the sensory process is assumed to be the particular sensory value on that trial. The decision process is assumed to use the sensory value and the task instructions to select a response.

When applied to a concurrent ratings task, the univariate normal model assumes the subject perceives a value on the dimension along which a rating response is required. On each trial, this perceived value can be represented as a random sample from a univariate normal distribution. If the experiment uses an n-point rating scale, the subject is assumed to select n-1 criteria, which partition the dimension into n intervals. The n rating responses are assigned to these intervals in a left-to-right fashion. On each trial, the decision process notes which interval each perceived dimensional value is in, and then emits the associated rating. The value of the response criteria are assumed to be under the subject’s control, but the distribution of perceived values is determined by the subject’s sensory physiology and by certain slowly acting processes (e.g., sensory adaptation) that are not under the subject’s control.

Coombs (1964) developed his unfolding model to account for data from choice experiments in which the subject is presented with a set of related stimuli or products and is asked to choose the one he or she most desires. The unfolding model assumes that each subject imagines an ideal product of the type offered in the experiment and then compares each of the available products to the imagined deal. The product that is most similar to the ideal is selected by the subject as the most desired element in the choice set. In the original unfolding model, the ideal and each member of the choice set was represented as a point in a psychological space and similarity to the ideal was assumed to decrease with the distance from the ideal point. The unfolding model allowed for single-peaked preference functions but could not account for trial-by-trial variability in the choices made by a single subject. The model also assumes that choice is strictly transitive.

Probabilistic versions of the unfolding model were developed by Schonemann and Wang (1972), Zinnes and Griggs (1974), De Soete, Carroll, and DeSarbo (1986), and Ennis and Mullen (1992). In its most general form, the probabilistic unfolding model assumes trial-by-trial variability in the psychological representation of the stimuli and the ideal (Ennis & Mullen, 1992). Thus, rather than represent stimuli and the ideal as points, they are represented as multivariate probability distributions (i.e. multivariate normal distributions). On each trial, the psychological representation of each stimulus is
still a point in a psychological space, but now the point is assumed to be a random sample from the probability distribution associated with that stimulus. Likewise, the imagined ideal is also assumed to be a random sample from a probability distribution. From this point on, the probabilistic model is identical to the original unfolding model -- that is, the subject is assumed to choose the stimulus that is nearest the ideal point. For more details, see the reviews by either De Soete and Carroll (1992) or Ennis and Mullen (1992).

The Thurstone-Coombs Concurrent Ratings Model

**Theoretical Development.** Consider a set of stimuli that vary on at least one sensory dimension. Suppose a subject is asked to rate one of these stimuli on one of these sensory dimensions and on liking with respect to that attribute. Consider the case where both rating scales run from 1 to n, with n reflecting the greatest sensory strength and n reflecting the highest liking. Under these conditions, the Thurstone-Coombs concurrent ratings model assumes the subject's responses are based on the following steps.

1) The sensory magnitude experienced by the subject is represented by the random variable $X$. Because of stimulus noise, perceptual noise, and individual difference, $X$ varies probabilistically over trials and subjects. We assume $X$ is normally distributed with mean $\mu_X$ and variance $\sigma_X^2$.

2) The subject places $n - 1$ criteria on the sensory dimension. Let $C_i$ denote the $i$th criterion. Because of criterial noise (i.e., variability in the subject's memory of the criteria), $C_i$ is normally distributed with mean $c_i$ and variance $\sigma_{C_i}^2$. The subject gives the stimulus a sensory rating $k$ if and only if $C_{k-1} < X \leq C_k$ where it is understood that $C_0 = -\infty$ and $C_n = +\infty$ (and $C_0$ and $C_n$ are not random variables).

3) To generate a liking rating, the subject first imagines an ideal stimulus, which is represented by the random variable $Y$. Because of variability in the imagining process (e.g., due to variability in memory and affective state) and individual difference, $Y$ varies probabilistically over trials and subjects. We assume $Y$ has a normal distribution with mean $\mu_Y$ and variance $\sigma_Y^2$.

4) The subject compares the ideal value with the perceived value obtained during step 1. This is done by computing the distance (or dissimilarity) between the sensory value of the stimulus and the ideal value; that is, between $X$ and $Y$. Call this distance $D_{xy}$.

5) The subject places $n - 1$ criteria on the $D_{xy}$ dimension. Let $B_j$ denote the $j$th of these criteria. Because of criterial noise, $B_j$ is normally distributed with mean $b_j$ and variance $\sigma_B^2$. Smaller distances are associated with greater hedonic responses and so, with higher ratings. Thus the subject
gives the stimulus a liking rating of \( k \) if and only if \( B_{k-1} < D_{XY} \leq B_k \), where it is understood that \( B_0 = 0 \) and \( B_{\infty} = +\infty \) (and thus \( B_0 \) and \( B_{\infty} \) are not random variables).

These events are illustrated in Figure 1. The upper panel represents the distribution of sensory magnitudes. In this case, the mean sensory magnitude is 0 and the standard deviation is 1. The x at –1.8 represents the sensory value of the stimulus on one particular trial. This value is assumed to be a random sample from the underlying normal distribution. The subject has been told to respond on a seven point rating scale, and the vertical lines in the top panel of Figure 1 represent the six criteria selected by the subject to divide the sensory dimension into seven intervals. The perceived strength falls in region three, so the subject assigns the strength of the stimulus a numerical rating of 3. The bottom panel of Figure 1 depicts the events that lead to the subject’s liking rating. First, the subject imagines an ideal sensory magnitude. In our example, the distribution of all possible ideal values is normal with a mean of 1 and a standard deviation of 0.5. The y at 0.5 represents the ideal magnitude on this particular trial. Finally, the 12 vertical lines on the ideal magnitude dimension, six on each side of the ideal value, represent the subject’s criteria on the distance between sensory magnitude and ideal magnitude. The perceived strength of –1.8 falls in region 3, so the subject assigns the stimulus a liking rating of 3. Note that the model allows for substantial differences between subjects, or groups of subjects. In particular, different subjects could have very different ideal sensory magnitudes and they could use substantially different criteria for assigning their ratings.

Let \((i,j)\) denote the event that the subject assigns a rating of \(i\) to the sensory dimension and a rating of \(j\) to the liking dimension. Then the above assumptions imply that the probability that the subject responds \((i,j)\) is equal to

\[
P([i,j]) = P(C_{i-1} < X \leq C_i \text{ and } B_{j-1} < |Y - X| \leq B_j)
\]

Appendix 1 shows that this probability reduces to

\[
P([i,j]) = \int_{c_{i-1} - \mu_X}^{c_i - \mu_X} \int_{\mu_Y - \mu_X + b_j}^{\mu_Y - \mu_X + b_{j-1}} f(u_1, u_2) du_2 + \int_{\mu_Y - \mu_X + b_{j-1}}^{\mu_Y - \mu_X + b_j} f(u_1, u_2) du_2 \]

where \(f(u_1, u_2)\) is a bivariate normal density function with mean (0,0) and covariance matrix.
\[ \Sigma_U = \begin{bmatrix} \sigma_X^2 + \sigma_C^2 & \sigma_X^2 \\ \sigma_X^2 & \sigma_X^2 + \sigma_Y^2 + \sigma_B^2 \end{bmatrix} \]

Fig. 1. Hypothetical events on one trial of a concurrent ratings task as hypothesized by the Thurstone-Coombs concurrent ratings model.
Appendix 1 also describes an efficient method for evaluating this integral numerically.

Equation (3) makes it clear that $\sigma_Y^2$ and $\sigma_B^2$ are not identifiable parameters. Because they affect only the $U_2$ variance, all that can be estimated is their sum. This is analogous to the well know fact of signal detection theory that sensory noise variance and criterial noise variance are nonidentifiable (Wickelgren & Norman, 1996). On the other hand, note that $\sigma_X^2$ and $\sigma_C^2$ are identifiable. The sensory variance $\sigma_X^2$ equals the covariance between $U_1$ and $U_2$ and the criterial noise variance on the sensory dimension equals the difference between the variance of $U_1$ and the covariance between $U_1$ and $U_2$. Thus, in this model it is possible to get separate estimates of sensory noise and criterial noise on the sensory dimension.

According to step 3, the model assumes the subject uses the same sensory value to make both rating responses. On a particular trial, the actual sensory value perceived by the subject is denoted by $x$ and the observed discrepancy from the ideal is denoted by $|y - x|$. If both ratings use the same sensory value then the $x$ in $|y - x|$ and $x$ are identical, so the two random variables $X$ and $|Y - X|$ are correlated. This is the reason that $\Sigma_U$ is not diagonal. An alternative form of the model assumes the subject resamples the sensory dimension before making the liking rating. In this case $x$ and the $x$ in $|y - x|$ are different random samples from the same population and therefore, are statistically independent. As a consequence, in the resampling model, $\Sigma_U$ is a diagonal matrix. Otherwise, the two models are identical. Computationally, resampling makes the model much simpler because then each double integral in Eq. (2) reduces to the product of single integrals. On the other hand, if the subject resamples then the parameters $\sigma_X^2$ and $\sigma_C^2$ are no longer identifiable. The question of which version is more appropriate, however, is empirical and might depend on the particular stimulus being evaluated and on the specific instructions given the subject. One way to decide between the two versions is to fit the Eq. 2 model and then examine the estimated $\Sigma_U$ covariance. If the subject is resampling, the resulting estimate should be nearly zero.

Still another way to formulate the model is to assume that the subject compares the sensory value to the ideal by computing similarity rather than distance or dissimilarity. In models that assume a geometric stimulus representation, a popular measure of the similarity between two stimuli $i$ and $j$, denoted, $S_{ij}$ is

$$S_{ij} = \exp(-D_{ij}^\alpha)$$

(4)
where $D_{ij}$ is the distance between the sensory representations of the two stimuli and $\alpha$ is a nonnegative constant that defines the nature of the similarity function (Ennis & Johnson, 1993; Nosofsky, 1986; Shepard, 1957). The exponential similarity function results when $\alpha = 1$ and the Gaussian function when $\alpha = 2$. The most successful models have paired the exponential similarity function with either the city-block or Euclidean distance metrics or the Gaussian similarity function with the Euclidean metric (Ennis, 1988; Nosofsky, 1986; Shepard, 1987).

Because similarity decreases with distance, the strategy of placing criteria on a similarity dimension is closely related to the strategy of placing criteria on a distance dimension. For example, Eq. 4 implies

$$D_{ij} = (-\log S_{ij})^{1/\alpha} \quad (5)$$

A decision to give a liking rating of $k$ if and only if $B_{k-1} < D_{XY} \leq B_k$ is therefore equivalent to responding $k$ if and only if

$$B_{k-1} < (-\log S_{XY})^{1/\alpha} \leq B_k \quad (6)$$

or equivalently, if and only if

$$\exp(-B_k^{\alpha/1}) \leq S_{XY} < \exp(-B_{k-1}^{\alpha/1}) \quad (7)$$

Thus, for a set of criteria on the distance dimension there is a set of criteria on the similarity dimension that yield exactly the same set of ratings.

**Empirical Application.** As an empirical application, we fit the Thurstone-Coombs model to the data of Table 1. To evaluate its performance, we also fit the multivariate normal model by using the maximum likelihood algorithm developed by Wickens (1992). In the present application, the multivariate normal ratings model has 13 free parameters; 6 response criteria or cut-points on each dimension (i.e., sensory magnitude and liking) and the correlation between sensory magnitude and perceived liking. Without loss of generality, the mean perceived value on each dimension can be set to zero and the variances can be set to one (Wickens, 1992).

Our application of the Thurstone-Coombs ratings model assumed no criterial noise. Without loss of generality, the mean and variance of the distribution of sensory magnitudes can be set to zero and one, respectively. Thus, in this application the model has 14 free parameters: 6 response criteria on sensory magnitude, 6 on distance to the ideal, and a mean and variance of the ideal distribution. The parameters were estimated by using an iterative search routine that minimized the sum of squared deviations between the observed and predicted response frequencies.
The observed and predicted response frequencies are shown in Table 2. In each cell, the value in bold is the observed frequency. The top value is the frequency predicted by the multivariate normal model and the bottom value is the frequency predicted by the Thurstone-Coombs model. The multivariate normal model accounted for 82.4% of the variance in the data, whereas the Thurstone-Coombs model accounted for 94.2%.

Percentage of variance accounted for can be a misleading statistic if the models being compared have a different number of parameters. In this case, the Thurstone-Coombs model has one more free parameter than the multivariate normal model (i.e., 14 versus 13). Thus, it is important to ask whether the better fit of the Thurstone-Coombs model is because of its extra free parameter. A goodness-of-fit measure that penalizes a model for extra free parameters is AIC (Akaike, 1974; Takane & Shibayama, 1992). If a model has \( m \) free parameters, its AIC value is

\[
\text{AIC} = G^2 + 2m
\]  

(8)

where \( G^2 \) is the familiar maximum likelihood goodness-of-fit measure. AIC values from different models can be compared directly. The model with the smaller AIC provides the better fit, regardless of any differences in the number of free parameters. In the present application, the multivariate normal model has an AIC value of 119.1, whereas the AIC value of the Thurstone-Coombs model is 82.8. Thus, the fit of the Thurstone-Coombs model is substantially better than the fit of the multivariate normal model.

The superiority of the Thurstone-Coombs model is seen clearly in Table 2. First, the multivariate normal model badly under predicts the most frequent response (a sensory rating of 4 and a liking rating of 7), whereas the Thurstone-Coombs model perfectly accounts for this data cell. Second, the multivariate normal model predicts that a liking rating of “2” should be assigned most frequently to intermediate sensory magnitudes. However, subjects never followed this pattern. Rather, they assigned a liking rating of “2” only to small sensory magnitudes or to large sensory magnitudes. The Thurstone-Coombs model accurately accounts for this bimodal trend in the data.

Figure 2 shows the distributions of the sensory magnitudes and the ideal magnitudes as predicted by the best fitting version of the Thurstone-Coombs model, along with the response criteria on each dimension. Recall that the mean and standard deviation of the sensory magnitude distribution were set to 0 and 1, respectively. The estimated mean and standard deviation of the ideal distribution were 0.09 and 0.60, respectively. Thus, for these subjects, the mean ideal sensory magnitude was only slightly greater than the mean perceived magnitude of the product that was tested.
TABLE 2.
OBSERVED AND PREDICTED RESPONSE FREQUENCIES IN A CONCURRENT RATINGS TASK. OBSERVED FREQUENCIES ARE IN BOLD. THE VALUES ABOVE THE BOLD ARE THE FREQUENCIES PREDICTED BY THE MULTIVARIATE NORMAL RATINGS MODEL. THE VALUES BELOW THE BOLD ARE THE FREQUENCIES PREDICTED BY THE THURSTONE-COOMBS RATINGS MODEL.

<table>
<thead>
<tr>
<th>Sensory Magnitude</th>
<th>Liking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

The bounds on distance-to-the-ideal were constructed as if the subject perceived a sensory magnitude of 0. Overall, the spacing between bounds is fairly uniform, although a few significant response biases are evident. For example, when rating sensory magnitudes, subjects showed a strong bias in favor of a “6” response, relative to their use of the equally extreme “2” response. Thus, according to the model, the fact that, on the sensory magnitude dimension, there were many more ratings of “6” than “2” is primarily a function of how the subjects used the rating scale, rather than a function of the actual sensory strength of the product.
Although the Thurstone-Coombs model had one more free parameter than the multivariate normal model in this application, if judgements are collected on more than one product, this ordering will reverse. For example, with two products the multivariate normal model adds five free parameters (a product mean and variance on each dimension and a correlation between dimensions), whereas the Thurstone-Coombs model adds only two parameters (a product mean and variance on the sensory magnitude dimension).
CONCLUSIONS

In the concurrent ratings procedure, a market segment of consumers rate a product simultaneously on a sensory dimension and on how much they like the product with respect to that dimension. The Thurstone-Coombs model assumes that consumers use fundamentally different processes to rate the sensory dimensions than to rate liking. Past models have treated sensory and liking dimensions in the same fashion (e.g., the bivariate normal ratings model).

When fit to data, the Thurstone-Coombs model provides estimates of the mean perceived value on the sensory dimension and its variance. These estimates could be used for quality control and also to track developmental changes in sensory ability. The model also provides estimates of the distribution of the ideal product. This includes an ideal mean on the sensory dimension and the variance of the ideal on this dimension. The variance along a dimension should increase as the psychological importance of that dimension decreases.

ACKNOWLEDGEMENTS

This work was supported in part by National Science Foundation Grants DBS92-09411 and BCS99-75037 to the first author.

APPENDIX 1

To derive Eq. (2), first note that Eq. (1) can be rewritten as

\[ P[(i, j)] = P(C_{i-1} < X \leq C_i, \text{ and } B_{j-1} < |Y - X| \leq B_j) \]

(A.1)

\[ = P(C_{i-1} < X \leq C_i, B_{j-1} < Y - X \leq B_j) + P(C_{i-1} < X \leq C_i, B_{j-1} < X - Y \leq B_j) \]

\[ = P(C_{i-1} < X \leq C_i, Y - B_j < X \leq Y - B_{j-1}) + P(C_{i-1} < X \leq C_i, Y + B_{j-1} < X \leq Y + B_j) \]

Now \(X, Y, C_i,\) and \(B_j\) are all normally distributed and mutually independent with means, \(\mu_X, \mu_Y, c_i,\) and \(b_j\) respectively and variances \(\sigma_X^2, \sigma_Y^2, \sigma_C^2\) and \(\sigma_B^2\), respectively. Next, define four new random variables in the following manner:

\[ V_X = X - \mu_X, \quad V_Y = Y - \mu_Y, \quad V_C = C_i - c_i, \text{ and } V_B = B_j - b_j. \]
Note that all four are normally distributed, mutually independent, and have zero mean. Their variances are $\sigma^2_X, \sigma^2_Y, \sigma^2_C$, and $\sigma^2_B$, respectively. In terms of these new random variables, Eq. (A.1) can be rewritten as

$$P(i,j) = P(V_C + c_i - V_X < V_Y + \mu_Y - \mu_X - b_j < V_Y + \mu_Y - V_B - b_{j-1}) + P(V_C + c_i - V_X < V_Y + \mu_Y - V_B - b_{j-1} < V_X + \mu_X + \mu_Y + V_B + b_j$$

$$= P(c_i - V_X < V_Y - V_B - \mu_X - \mu_Y - b_j < V_Y - V_B - \mu_X - b_{j-1} - \mu_X + V_Y < V_X + \mu_X + \mu_Y + b_{j-1} - \mu_X$$

Finally, define the following three random variables:

$$U_1 = V_X - V_Y, \text{ is normally distributed with mean 0 and variance } \sigma^2_X = \sigma^2_C$$

$$U_2 = V_X - V_Y + V_Y, \text{ is normally distributed with mean 0 and variance } \sigma^2_X + \sigma^2_Y + \sigma^2_B$$

$$U_3 = V_X - V_Y - V_B, \text{ is normally distributed with mean 0 and variance } \sigma^2_X + \sigma^2_Y + \sigma^2_B$$

Note that $E(U_1 U_2) = E(U_1) = E(V_X^2) = \sigma^2_X$. Thus, $U_2$ and $U_3$ are identically distributed and they have identical correlations with $U_1$, so we can safely replace all occurrences of $U_3$, with $U_2$. Equation (1) can now be written as

$$P(i,j) = P(c_i - \mu_X < U_1 < \mu_X - \mu_Y - b_j, U_2 < \mu_Y - \mu_X - b_{j-1}) + P(c_i - \mu_X < U_1 < c_i - \mu_Y - \mu_X + b_{j-1} - \mu_X < U_2 < \mu_Y + b_j)$$

where $U = [U_1, U_2]$ has a bivariate normal distribution with mean vector $(0,0)$' and covariance matrix

$$\Sigma_U = \begin{bmatrix} \sigma^2_X + \sigma^2_C & \sigma^2_X \\ \sigma^2_X & \sigma^2_X + \sigma^2_Y + \sigma^2_B \end{bmatrix}$$

Equation 2 follows directly from this result.

The Eq. 2 integral can be evaluated numerically using a result given by Owen (1956). To begin, note that both probabilities on the right side of Eq. (A.2) are of the form $P(u_{11} < U_1 < u_{12}, u_{21} < U_2 < u_{22})$ for some constants $u_{11}, u_{12}, u_{21}$, and $u_{22}$. This rectangular probability is equal to

$$P(u_{11} < U_1 < u_{12}, u_{21} < U_2 < u_{22}) = P(U_1 > u_{11}, U_2 > u_{21}) - P(U_1 > u_{12}, U_2 > u_{21}) - P(U_1 > u_{11}, U_2 > u_{22}) + P(U_1 > u_{12}, U_2 > u_{22})$$
Owen’s (1956) result provides an efficient method for evaluating each term on the right. We illustrate the method for the first term. Note that

\[
P(U_1 > u_{11}, U_2 > u_{21}) = P \left( \frac{U_1 - u_{11}}{\sqrt{\sigma_X^2 + \sigma_C^2}} > 0, \frac{U_2 - u_{21}}{\sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_B^2}} > 0 \right)
\]

\[
= P(W_1 > 0, W_2 > 0)
\]

where \((W_1, W_2)\)' has a bivariate normal distribution with mean vector and covariance matrix

\[
\mu_W = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \frac{-u_{11}}{\sqrt{\sigma_X^2 + \sigma_C^2}} \\ \frac{-u_{21}}{\sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_B^2}} \end{bmatrix} \quad \text{and} \quad \Sigma_W = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}
\]

and where

\[
\rho = \frac{\sigma_Y}{\sqrt{\sigma_X^2 + \sigma_C^2} \sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_B^2}}
\]

Owen (1956) showed that Eq. (A.3) reduces to

\[
P(W_1 > 0, W_2 > 0) = \Phi(\mu_1) \Phi(\mu_2) + \int_0^\rho \frac{1}{2\pi \sqrt{1-t^2}} \exp \left( \frac{2t \mu_1 \mu_2 - \mu_1^2 - \mu_2^2}{2 - 2t^2} \right) dt
\]

This expression was derived independently by Palen and Ennis (1991), who also presented a derivation and discussed various methods for evaluating the integral numerically. A simple, yet fast alternative is to use Simpson’s rule. Ashby (1992, pp. 24-26) discusses other methods for the numerical evaluation of integrals over the multivariate normal density function.

REFERENCES


SHEPARD, R.N. 1957. Stimulus and Response Generalization: A Stochastic Model Relating Generalization to Distance in Psychological Space, Psychometrika, 22, 325-345.


