

Probabilistic unfolding models for sensory data

David B. MacKay*

Kelley School of Business, Indiana University, 1309 E Tenth Street, Bloomington, IN 47405-1701, USA

Received 11 July 2000; received in revised form 26 December 2000; accepted 5 February 2001

Abstract

Unfolding models are conceptually appealing for the analysis of consumers' hedonic evaluations of food products. The appeal of the unfolding model is three fold; its conceptual simplicity, spatial character, and assumption of satiety — more is not always better. Unfortunately, the success of unfolding models does not always match their appeal. Reformulating unfolding models in a probabilistic framework is shown to improve their success, extend their application and further enhance their conceptual attraction. Data provided by the organizers of The Fifth Sensometrics Meeting are used to illustrate the proposed reformulation. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Multidimensional unfolding; Hedonic ratings; Attribute ratings; Product maps; Multidimensional scaling

1. Introduction

In their simplest form, unfolding models (Coombs, 1964) start with consumers' hedonic liking ratings for real products and estimate coordinates in a multi-dimensional space to represent both these real products and ideal products. Coordinates are estimated so that the resulting distances between an ideal product and the real products inversely approximate the hedonic evaluations — large distances indicate minimal liking. Attribute information is not required. Identification of the dimensions of the space comes either from the analyst's knowledge of the objects being studied or from external information. Depending upon the analyst's objectives, ideal products may be estimated for individual consumers or for segments of consumers.

Differences in how traditional and the proposed probabilistic unfolding model conceptualize consumers' psychological processes are illustrated in Fig. 1. The traditional model, in the panel on the right, conceptualizes the real and ideal objects as points. In this hypothetical situation, the consumer subject prefers drink D2 to D1 since D2 is closer to the subject's ideal product *S*. If the dimensions are identified as measuring aroma and sweetness, then we conclude that the subject prefers moderate values of aroma and sweetness to

strong or weak values. The proposed model, in the panel on the left, conceptualizes the real and ideal objects as multivariate normal distributions. Each symbol represents a value sampled from these distributions. Even though the centroids of the distributions in the two panels are the same, drink D1 will be preferred more often than D2 with the proposed model since there is more overlap in the distributions of D1 and *S* than there is in the distributions of D2 and *S*. Stated differently, the expected distance between D1 and *S* is less than the expected distance between D2 and *S*. The ellipses in the panel on the left represent the unit standard deviational contours of the three objects.

In the following sections, we shall discuss selected properties of the proposed model, briefly describe how the model works, contrast the application of the proposed and traditional models with consumer data on beverage attributes and likings, and discuss the limitations and benefits of using probabilistic unfolding models with sensory data.

2. Selected properties

Both economics and psychology have contributed models for helping us better understand consumer behavior. Unlike many economic models, the unfolding model has the compelling property of satiety. It allows us to model situations in which preferred products are

* Tel.: +1-812-855-1009; fax: +1-812-855-6440.

E-mail address: mackay@indiana.edu

characterized by moderate as well as extreme values of their attributes.

Unfolding models assume that products exist in a common space, one that is usually multidimensional in nature. Differences in product preferences among consumers are assumed to be due to differences in their ideals — some consumers prefer sweet products and others prefer less sweet products. Consumers are assumed to perceive the products in a common fashion. With traditional deterministic models, the common space assumption is very severe. It requires, for example, that on each salient dimension, *all* consumers consistently view the coordinate or attribute values of the products in exactly the same order.

It does not take a lot of experience collecting perceived product attribute data to realize that these consistency assumptions do not hold, even for a single subject. Differences in product perception across subjects can be very high.

Choice consistency is also a property of traditional models. To comport with the model, if a consumer pre-

fers product *A* to product *B* and product *B* to product *C*, then product *A* must always be preferred to product *C*. Indeed, if a consumer prefers product *A* to product *B* once, product *B* will never be chosen. In Fig. 1, the right hand panel tells us that *S* always prefers D2 to D1. For the probabilistic model illustrated in the left hand panel, D1 is usually preferred to D2 but D2 is frequently chosen.

Violations of model assumptions are not new, they happen all the time. What makes these violations worth noting is (1) that they are the rule, not the exception, and (2) that failure to account for these violations will result in systematically biased results. Fig. 2 illustrates the nature of this bias. In the left panel of Fig. 2 are the mean and variance parameters (represented by the standard deviational ellipses) for eight real objects, *A-H*, and four ideal objects, *I-L*. The first four real objects and the four ideal objects have relatively small variances while the second four real objects have relatively large variances. (While these hypothetical simulation variances may seem large, in our experience, they are not that uncommon for actual empirical studies.)

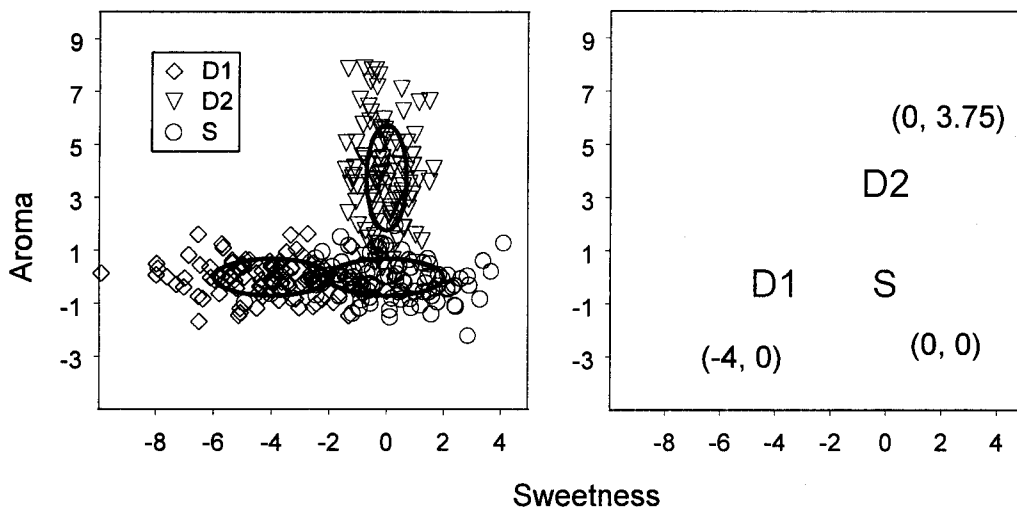


Fig. 1. Probabilistic and deterministic models of a subject *S*'s preference for drinks D1 and D2.

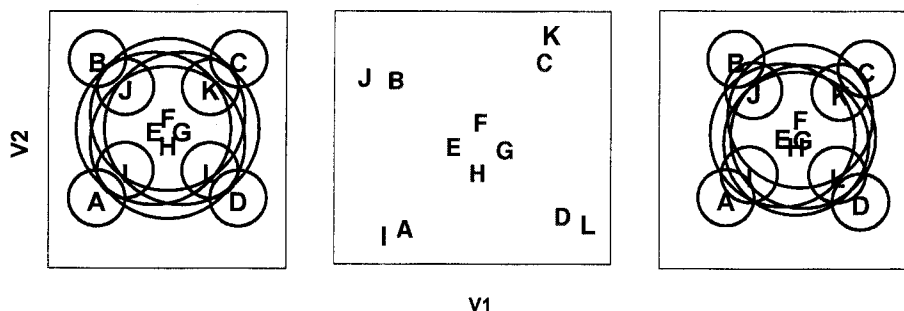


Fig. 2. Parameters, deterministic estimates and probabilistic estimates for eight real (*A-H*) and four ideal (*I-L*) objects. Estimates are from simulated hedonic judgments.

The middle panel of Fig. 2 illustrates the deterministic, nonmetric solution provided by KYST (Kruskal, Young, & Seery, 1977), a popular traditional unfolding program, when provided with the mean liking ratings for the eight real objects of 5600 simulated subjects, 1400 for each ideal object. It is observed that the traditional solution has inverted the locations of the first four real objects and the four ideal objects. This bias will not go away if the sample size increases; it will hold even for infinitely large samples. If dimensions V1 and V2 were identified, the traditional solution would tell product designers that all consumers preferred products with extreme values on both dimensions. This would clearly be bad advice.

The systematic bias shown in the middle panel has a very intuitive explanation. If two objects have fixed centroids, the expected distance and thus the average “disliking” between them will increase as one or more of the variance magnitudes increases. When judgments characterized by high variance are put into a deterministic model that does not admit variances, the effect of the variances is accommodated by estimating the points as being farther apart than the centroids actually are. Note that in this symmetric example, the ideal objects all move toward the periphery of the space. In so doing, the relatively short distances of the four ideal to the first four real products are maintained and the longer distance to the second four real products, the high variance products, is accommodated. In asymmetric situations, a common deterministic accommodation to differential variances is to produce degenerate solutions in which all the centroids of one set, the real or ideal objects, are estimated as having similar values.

The right panel of Fig. 2 shows the parameter estimates provided by PROSCAL (MacKay & Zinnes, 2000), a program for probabilistic unfolding and probabilistic mapping. The estimates are not perfect but they are very good. The correlation of inter-centroid distances in the first and last panel is 0.996. The actual variance values are 0.04 and 0.30 while the estimated variances are 0.04 and 0.28.

So far, we have seen that the inclusion of variance parameters provides a process that is more consistent with what we know of consumer behavior and that does

not have the estimation bias of deterministic models that exclude variance parameters. Another advantage is that the new models permit the estimation of the probability with which a real object will be chosen. Deterministic models have no basis for estimating choice probabilities and must rely on ad hoc heuristics. To estimate choice probabilities using a probabilistic model, the analyst must first specify whether independent or dependent sampling is assumed.

To illustrate independent and dependent sampling, consider the example of Fig. 3 where a single subject *S* is making a choice among three real products, *A*, *B* and *C*. Under independent sampling, when *S* estimates the distance from the ideal to a real product, values are sampled from the distribution about *S* and the distribution about the real product and a distance between the two is calculated. When a second product is evaluated, a separate independent sample from *S* and a sample from the new real product is undertaken again. Integration over the joint density function or simulation can be used to estimate the first choice probabilities, which, in this case, are 0.72 for product *A* and 0.14 for both products *B* and *C*. Under dependent sampling, a single sample is used to draw a value from the distribution about *S*. This value is then compared to the values sampled from the distributions about *A*, *B* and *C*. Here, the first choice probabilities are 0.72, 0.08 and 0.20. Under dependent sampling, a real product's first choice probability is thus contingent upon the locations of all the products, not just upon the relation of the real product to the ideal product. If the analyst believes that independent sampling occurs in the laboratory and dependent sampling occurs in the field, then it is possible to provide estimates under independence assumptions and make first choice predictions under dependence assumptions.

For those not used to using probabilistic unfolding models, the solutions of probabilistic unfolding models may present some surprises. It is, for example, possible for a real product whose centroid is close to the centroid of an ideal product to have a very low probability of being chosen. One cause for this may be that the real product has a very large variance. As a result, for any one decision, the chances are high that the real object is actually a good distance away from the ideal object.

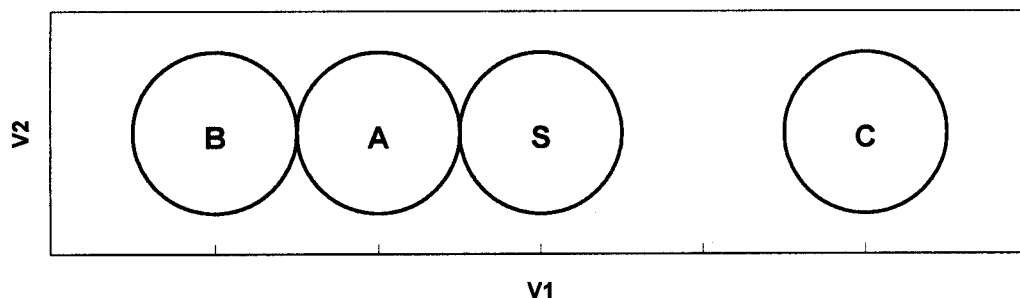


Fig. 3. Dependent and independent sampling among three objects (*A*–*C*) by one subject (*S*).

Stated differently, it is often true that expected distances are non-monotonic with distances among centroids.

The relation of variances to choice probabilities is complex. Very often, the reduction of a real product variance will increase the probability of the product being chosen. While smaller variances often lead to higher choice probabilities, a zero level of variance is not always optimal. A larger variance may extend a product's appeal to multiple segments and increase its overall first choice probability or "perceptual share." Optimizing a product's perceptual share depends both upon obtaining the optimal centroid and the optimal level of variation in perception.

The perceptual shares estimated by probabilistic unfolding should, of course, be distinguished from products' market shares. Even if the products being evaluated in an unfolding study are products on supermarket shelves, marketing and distribution variables will inevitably cause disparities between perceptual shares and market shares. Nevertheless, it is still meaningful to ask what attribute values will optimize perceptual share. Once the products and market segments have been established, it is possible to build "what-if" models from probabilistic unfolding model output to assess the perceptual shares of untested products.

Product variances are affected by many factors. When using unidentified products in the laboratory, variances are due to fluctuations in the subjects and the stimuli and to differences among the subjects. When using identified products in the marketplace, a host of marketing variables also becomes relevant. A common observation of advertising is that it will reduce the perceived variance in a product's perception. Sometimes, the introduction of a new product may draw attention to existing products and lower their variances as well. When this happens, regularity can be violated — the addition of a new element in the choice set may actually increase the choice probability of existing products.

Before leaving the topic of perceptual shares, it should be noted that the relationship of products' perceptual shares is almost never monotonic to their mean liking ratings. As will be seen later on, it is quite possible for a product with a modest liking rating to have a very high perceptual share, and vice versa. The lack of monotonicity is due to the forces exerted by different market segments and competitive products in the market place.

The last property to be discussed is the presence of an error theory. Traditional deterministic models have no error theory. To determine if a model should, for example, be estimated in one, two or more dimensions, heuristic rules of thumb must be employed. The variances of probabilistic models not only allow choice probabilities to be estimated, they also allow hypothesis tests to be made. Examples of hypothesis tests include the dimensionality of the space, equality of variances and equality of centroids. If a company is thinking of

changing the ingredients of a product, hypothesis tests can be used to answer the question of whether consumers can recognize the difference.

3. Model estimation

Maximum likelihood (ML) methods are used to estimate the probabilistic unfolding models. ML estimation proceeds by maximizing the log of the likelihood function, which is the sum of the logs of the probability density functions (PDFs) over all observations. To do this, we need to know the nature of the observation or judgment and its PDF.

The simplest judgment for unfolding analysis is a liking rating. If we are in a two dimensional space, with variances σ^2 that are the same for all products on all dimensions, and we assume that the liking rating is modeled as a Euclidean distance d_{ij} between an ideal object i and a real object j with centroids on dimension k of μ_{ik} and μ_{jk} , respectively, then the calculation of the PDF for the liking rating proceeds as follows:

Let

$$d_{ij}^2 = \sum_{k=1}^2 (x_{ik} - x_{jk})^2$$

where

$$x_{ik} \sim N(\mu_{ik}, \sigma^2).$$

Then,

$$d_{ijk} = x_{ik} - x_{jk} \sim N(\mu_{ijk}, 2\sigma^2); \mu_{ijk} = \mu_{ik} - \mu_{jk}$$

$$d_{ij}^2 = \sum_{k=1}^2 d_{ijk}^2$$

$$d_{ij}^2/2\sigma^2 \sim \chi_{v, \lambda_{ij}}^2; v = 2, \lambda_{ij} = \sum_{k=1}^2 (\mu_{ijk}^2/2\sigma^2)$$

$$d_{ij} \sim (d_{ij}/\sigma^2) \chi_{v, \lambda_{ij}}^2.$$

The PDF of d_{ij} is thus a function of the non-central chi-square distribution that has two parameters, a non-centrality parameter λ_{ij} and a degrees of freedom parameter v that is equal to the dimensionality of the space. Closed form expressions for the non-central chi-square distribution do not exist and approximations must be used (Zinnes & MacKay, 1983). Measurement models, described below, may be used to relate d_{ij} to the subject's judgment δ_{ij} .

Stated differently, it is often true that expected distances are non-monotonic with distances among centroids.

The relation of variances to choice probabilities is complex. Very often, the reduction of a real product variance will increase the probability of the product being chosen. While smaller variances often lead to higher choice probabilities, a zero level of variance is not always optimal. A larger variance may extend a product's appeal to multiple segments and increase its overall first choice probability or "perceptual share." Optimizing a product's perceptual share depends both upon obtaining the optimal centroid and the optimal level of variation in perception.

The perceptual shares estimated by probabilistic unfolding should, of course, be distinguished from products' market shares. Even if the products being evaluated in an unfolding study are products on supermarket shelves, marketing and distribution variables will inevitably cause disparities between perceptual shares and market shares. Nevertheless, it is still meaningful to ask what attribute values will optimize perceptual share. Once the products and market segments have been established, it is possible to build "what-if" models from probabilistic unfolding model output to assess the perceptual shares of untested products.

Product variances are affected by many factors. When using unidentified products in the laboratory, variances are due to fluctuations in the subjects and the stimuli and to differences among the subjects. When using identified products in the marketplace, a host of marketing variables also becomes relevant. A common observation of advertising is that it will reduce the perceived variance in a product's perception. Sometimes, the introduction of a new product may draw attention to existing products and lower their variances as well. When this happens, regularity can be violated — the addition of a new element in the choice set may actually increase the choice probability of existing products.

Before leaving the topic of perceptual shares, it should be noted that the relationship of products' perceptual shares is almost never monotonic to their mean liking ratings. As will be seen later on, it is quite possible for a product with a modest liking rating to have a very high perceptual share, and vice versa. The lack of monotonicity is due to the forces exerted by different market segments and competitive products in the market place.

The last property to be discussed is the presence of an error theory. Traditional deterministic models have no error theory. To determine if a model should, for example, be estimated in one, two or more dimensions, heuristic rules of thumb must be employed. The variances of probabilistic models not only allow choice probabilities to be estimated, they also allow hypothesis tests to be made. Examples of hypothesis tests include the dimensionality of the space, equality of variances and equality of centroids. If a company is thinking of

changing the ingredients of a product, hypothesis tests can be used to answer the question of whether consumers can recognize the difference.

3. Model estimation

Maximum likelihood (ML) methods are used to estimate the probabilistic unfolding models. ML estimation proceeds by maximizing the log of the likelihood function, which is the sum of the logs of the probability density functions (PDFs) over all observations. To do this, we need to know the nature of the observation or judgment and its PDF.

The simplest judgment for unfolding analysis is a liking rating. If we are in a two dimensional space, with variances σ^2 that are the same for all products on all dimensions, and we assume that the liking rating is modeled as a Euclidean distance d_{ij} between an ideal object i and a real object j with centroids on dimension k of μ_{ik} and μ_{jk} , respectively, then the calculation of the PDF for the liking rating proceeds as follows:

Let

$$d_{ij}^2 = \sum_{k=1}^2 (x_{ik} - x_{jk})^2$$

where

$$x_{ik} \sim N(\mu_{ik}, \sigma^2).$$

Then,

$$d_{ijk} = x_{ik} - x_{jk} \sim N(\mu_{ijk}, 2\sigma^2); \mu_{ijk} = \mu_{ik} - \mu_{jk}$$

$$d_{ij}^2 = \sum_{k=1}^2 d_{ijk}^2$$

$$d_{ij}^2/2\sigma^2 \sim \chi_{v, \lambda_{ij}}^2; v = 2, \lambda_{ij} = \sum_{k=1}^2 (\mu_{ij}^2/2\sigma^2)$$

$$d_{ij} \sim (d_{ij}/\sigma^2) \chi_{v, \lambda_{ij}}^2.$$

The PDF of d_{ij} is thus a function of the non-central chi-square distribution that has two parameters, a non-centrality parameter λ_{ij} and a degrees of freedom parameter v that is equal to the dimensionality of the space. Closed form expressions for the non-central chi-square distribution do not exist and approximations must be used (Zinnes & MacKay, 1983). Measurement models, described below, may be used to relate d_{ij} to the subject's judgment δ_{ij} .

Generalizing the above results to higher dimensions is trivial; generalizing to other variance structures, different types of judgments, and alternative metrics is more complicated. As an example of different variance structures, consider the four situations presented in Fig. 4. Following Thurstone's (1927) nomenclature, we distinguish between Case 5 models where, as above, the variances are the same for all objects and Case 3 models where the variances may differ from object to object. We also distinguish between isotropic models where the variances for any one object are the same on all dimensions and anisotropic models where the variances may differ from dimension to dimension. PDFs for anisotropic liking ratings, which follow the quadratic forms in normal variables distributions, are described in MacKay (1989).

The selection of an appropriate variance structure is up to the analyst. In the testing of unidentified food products, for example, it is often appropriate to use a Case 5a model. On the other hand, if the products are identified products with which the consumers are familiar, then Case 3a models may be a better choice. Likelihood ratio tests may be used to determine if the greater fit available by using a Case 3a model outweighs the loss in degrees of freedom. The modeling of variances is very flexible. It is a simple matter to add terms that account for things such as the order bias of objects and the distance magnitudes of judgments.

Instead of using liking ratings, other types of preference judgments may also be used. A type of judgment that is more powerful than liking ratings is the preference ratio. For preference ratios, subjects evaluate pairs of real objects and, for each pair, indicate the preferred item in the pair and state how many times it is preferred over the less preferred object. Graphic rating scales are frequently used to obtain these judgments. To

obtain the PDF of preference ratio judgments involves finding the distributions of ratios of distances (MacKay, 2001; MacKay & Zinnes, 1995; Zinnes & MacKay, 1987). The extra power in preference ratios comes from the fact that they are conjoint judgments — each judgment involves an evaluation of two real objects as well as one ideal object. Solutions obtained with preference ratio judgments tend to be more interpretable than solutions obtained from liking ratings, which are examples of disjoint judgments — each judgment involves the sampling of values from two objects that are in different sets.

Finally, another consideration is the type of metric one chooses to use. Euclidean metrics are used most often but density functions can also be formulated for the city-block metric (MacKay, 2001). It is commonly proposed that city-block metrics may be more appropriate for modeling the judgments of experts while Euclidean metrics are more appropriate for modeling the judgments of novices. Testing this proposition with nonmetric models is almost impossible, even using rule of thumb heuristics, since the presence of just a small amount of object variance quickly turns the selection of a metric into a situation that is little different than a coin toss. Probabilistic models, on the other hand, are very successful at testing metric properties.

Estimation of a probabilistic unfolding model is not restricted to one type of data. Thus, it is possible to combine both liking ratings and preference ratios in the estimation process. It is also possible to combine other types of data. A good candidate for inclusion is the dissimilarity judgment — a judgment where large values indicate that two real objects are very dissimilar and small values indicate that two real objects are very similar. Dissimilarities and similarities play a pivotal role in much theorizing in psychology (Goldstone, 1994). Dissimilarity judgments are conjoint judgments and can add a lot of explanatory power to an unfolding analysis. Dissimilarity judgments may be made directly by consumers or they may be derived from consumers' attribute evaluations. Dissimilarity judgments and liking ratings do not share the same scale. To simulta-

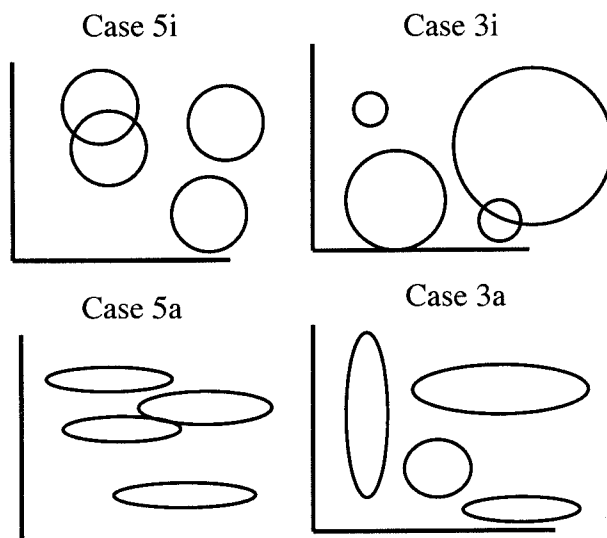


Fig. 4. Standard deviational ellipses for four types of variance structures.

Table 1
Probability density functions for different types of judgments, metrics and variance structures^a

	Liking ratings		Preference ratios	
	Euclidean	City-block	Euclidean	City-block
Case 5i	χ^2	FN	F'	RFN
Case 5a	Q	FN	Q	RFN
Case 3i	χ^2	FN	F'	RFN
Case 3a	Q	FN	Q	RFN

^a χ^2 , Non-central chi-square; F' , doubly non-central F distribution; Q , quadratic forms in normal variables distribution; FN, folded normal distribution; RFN, ratio of folded normal variables distribution.

