Chapter Twenty-One

Multidimensional Scaling and Conjoint Analysis

Chapter 21 - Multidimensional Scaling

4) Conducting Multidimensional Scaling

i. Formulating the Problem

ii. Obtaining Input Data

  a. Perception Data: Direct Approaches

  b. Perception Data: Derived Approaches

  c. Direct Vs. Derived Approaches

  d. Preference Data

iii. Selecting an MDS Procedure

iv. Deciding on the Number of Dimensions

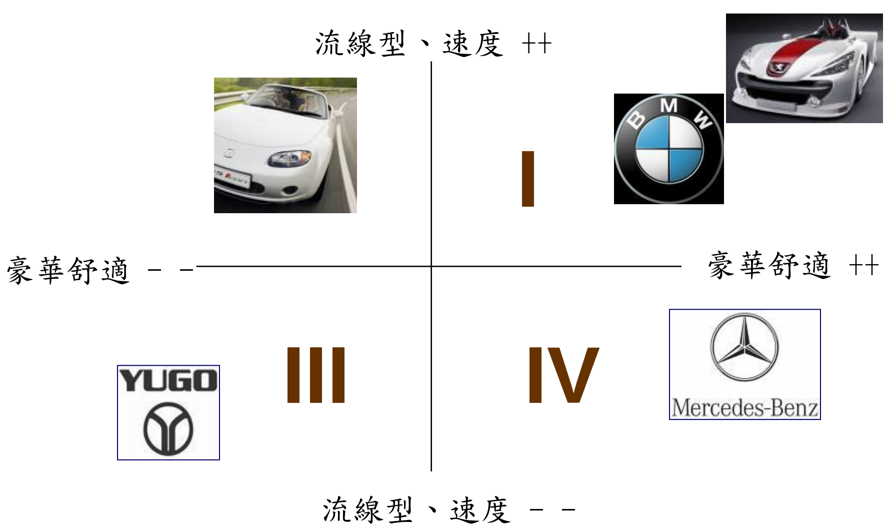
v. Labeling the Dimensions & Interpreting the Configuration

vi. Assessing Reliability and Validity

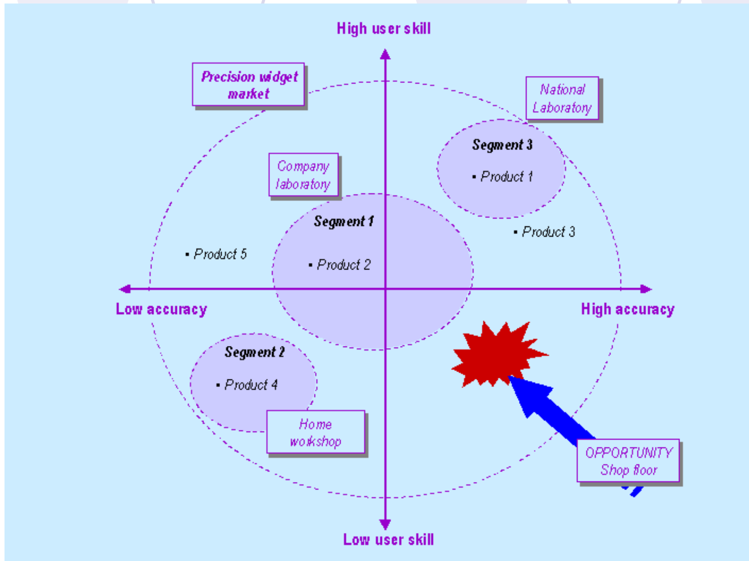
Multidimensional Scaling (MDS)

- **Multidimensional scaling (MDS)** is a class of procedures for representing *perceptions* and *preferences* of respondents spatially by means of a visual display.
- Perceived or psychological relationships among stimuli are represented as geometric relationships among points in a multidimensional space.
- These geometric representations are often called **spatial maps**. The axes of the spatial map are assumed to denote the psychological bases or underlying dimensions respondents use to form perceptions and preferences for stimuli.

Product Positioning產品定位



Marketing Applications



Marketing Applications

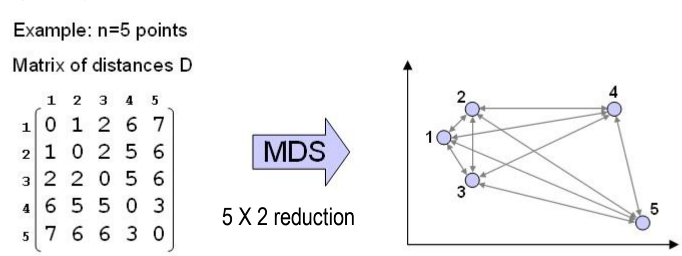
- Brand Image
- Market Segmentation (HTC)
- New product development
- Assessing advertising effectiveness (Markops, L'Oreal game)
- Pricing analysis (Luxgen, “coffee paradox,”)
- Channel decisions (compatible product sorting)
- Attitude scale construction (nostalgia underlying dimensions, cosmetics, e-Book)

Statistics and Terms Associated with MDS

- **Similarity judgments**. Similarity judgments are ratings on all possible pairs of brands or other stimuli in terms of their similarity using a Likert type scale.
- **Preference rankings**. Preference rankings are rank orderings of the brands or other stimuli from the most preferred to the least preferred. They are normally obtained from the respondents.
- **Stress**. This is a lack-of-fit measure; higher values of stress indicate poorer fits.
- **R-square**. R-square is a squared correlation index that indicates the proportion of variance of the optimally scaled data that can be accounted for by the MDS procedure. This is a goodness-of-fit measure.

Statistics and Terms Associated with MDS

- **Spatial map**. Perceived relationships among brands or other stimuli are represented as geometric relationships among points in a multidimensional space called a spatial map.
- **Coordinates**. Coordinates indicate the positioning of a brand or a stimulus in a spatial map.
- **Unfolding** . The representation of both brands and respondents as points in the same space is referred to as unfolding (psychometrics)



Conducting Multidimensional Scaling

Fig. 21.1

Formulate the Problem

↓

Obtain Input Data

↓

Select an MDS Procedure

↓

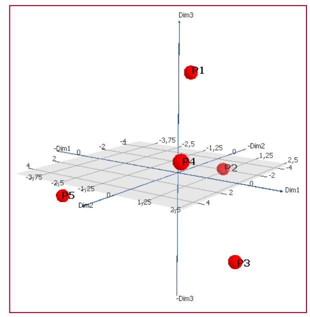
Decide on the Number of Dimensions

↓

Label the Dimensions and Interpret the Configuration

↓

Assess Reliability and Validity



Conducting Multidimensional Scaling

Formulate the Problem

- Specify the purpose for which the MDS results would be used.
- Select the brands or other stimuli to be included in the analysis. The number of brands or stimuli selected normally varies <u>between 8 and 25</u>.
- The choice of the number and specific brands or stimuli to be included should be based on the statement of the marketing research problem, theory, and the judgment of the researcher.

Input Data for Multidimensional Scaling
Fig. 21.2

MDS Input Data
- Perceptions
  - Direct (Similarity Judgments)
  - Derived (Attribute Ratings)
- Preferences

Input Data for Multidimensional Scaling

Similarity Judgments

Attribute Ratings

	Alternatives (Brands A, B, C)		Ideal Product
Attributes or Features			

Conducting Multidimensional Scaling
Obtain Input Data

- **Perception Data: Direct Approaches.** In direct approaches to gathering perception data, the respondents are asked to judge how similar or dissimilar the various brands or stimuli are, using their own criteria. These data are referred to as similarity judgments.

	Very Dissimilar						Very Similar
Crest vs. Colgate	1	2	3	4	5	6	7
Aqua-Fresh vs. Crest	1	2	3	4	5	6	7
Crest vs. Aim	1	2	3	4	5	6	7
.							
.							
.							
Colgate vs. Aqua-Fresh	1	2	3	4	5	6	7

- The number of pairs to be evaluated is $n(n-1)/2$, where n is the number of stimuli.

Similarity Rating Of Toothpaste Brands

Table 21.1

	Aqua-Fresh	Crest	Colgate	Aim	Gleem	Macleans	Ultra Brite	Close-Up	Pepsodent	Dentagard
Aqua-Fresh										
Crest	5									
Colgate	6	7								
Aim	4	6	6							
Gleem	2	3	4	5						
Macleans	3	3	4	4	5					
Ultra Brite	2	2	2	3	5	5				
Close-Up	2	2	2	2	6	5	6			
Pepsodent	2	2	2	2	6	6	7	6		
Dentagard	1	2	4	2	4	3	3	4	3	

Conducting Multidimensional Scaling
Obtain Input Data

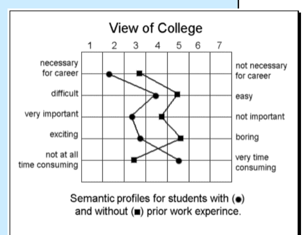
- **Perception Data: Derived Approaches.** Derived approaches to collecting perception data are attribute-based approaches requiring the respondents to rate the brands or stimuli on the identified attributes using semantic differential or Likert scales.

Whitens teeth __ __ __ __ __ __ __ __ __ __ **Does not whiten teeth**

Prevents tooth decay __ __ __ __ __ __ __ __ __ __ **Does not prevent tooth decay**

.
.
.
.

Pleasant tasting __ __ __ __ __ __ __ __ __ __ **Unpleasant tasting**



Semantic profiles for students with (●) and without (■) prior work experince.

- If attribute ratings are obtained, a similarity measure (such as Euclidean distance) is derived for each pair of brands.

Conducting Multidimensional Scaling
Obtain Input Data – Direct vs. Derived Approaches

The direct approach (brands-based) has the following advantages and disadvantages:

- The researcher does not have to identify a set of salient attributes.
- The disadvantages are that the criteria are influenced by the brands or stimuli being evaluated.
- Furthermore, it may be difficult to label the dimensions of the spatial map.

Conducting Multidimensional Scaling
Obtain Input Data – Direct vs. Derived Approaches

The attribute-based approach has the following advantages and disadvantages:

- It is easy to identify respondents with homogeneous perceptions.
- The respondents can be clustered based on the attribute ratings.
- It is also easier to label the dimensions.
- A disadvantage is that the researcher must identify all the salient attributes, a difficult task.
- The spatial map obtained depends upon the attributes identified.

It may be best to use both these approaches in a complementary way. Direct similarity judgments may be used for obtaining the spatial map, and attribute ratings may be used as an aid to interpreting the dimensions of the perceptual map.

Conducting Multidimensional Scaling
Preference Data

- Preference data order the brands or stimuli in terms of respondents' preference for some property.
- A common way in which such data are obtained is through preference rankings.
- Alternatively, respondents may be required to make paired comparisons and indicate which brand in a pair they prefer.
- Another method is to obtain preference ratings for the various brands.
- The configuration derived from preference data may differ greatly from that obtained from similarity data. Two brands may be perceived as different in a similarity map yet similar in a preference map, and vice versa...

Conducting Multidimensional Scaling
Select an MDS Procedure

Selection of a specific MDS procedure depends upon:

- Whether perception or preference data are being scaled, or whether the analysis requires both kinds of data.
- The nature of the input data is also a determining factor.
  - **Non-metric MDS** procedures assume that the input data are ordinal, but they result in metric output.
  - **Metric MDS** methods assume that input data are metric. Since the output is also metric, a stronger relationship between the output and input data is maintained, and the metric (interval or ratio) qualities of the input data are preserved.
  - The metric and non-metric methods produce similar results.
- Another factor influencing the selection of a procedure is whether the MDS analysis will be conducted at the individual respondent level or at an aggregate level.

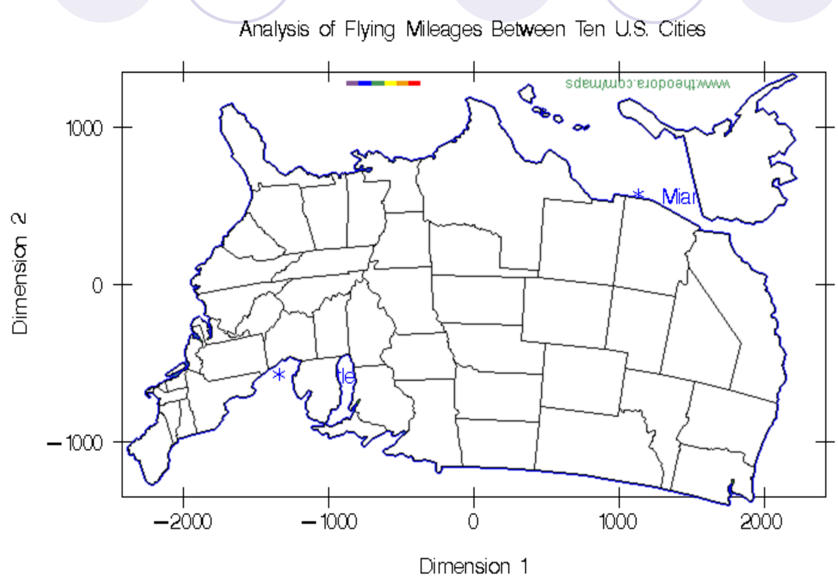
Conducting Multidimensional Scaling
Decide on the Number of Dimensions

- ***A priori* knowledge -** Theory or past research may suggest a particular number of dimensions.
- **Interpretability of the spatial map -** Generally, it is difficult to interpret configurations or maps derived in more than three dimensions.
- **Elbow criterion -** A plot of stress versus dimensionality should be examined.
- **Ease of use -** It is generally easier to work with two-dimensional maps or configurations than with those involving more dimensions.
- **Statistical approaches -** For the sophisticated user, statistical approaches are also available for determining the dimensionality.

Proc MDS in SAS

```
data   city;
   title 'Analysis of Flying Mileages Between Ten U.S. Cities';
   input  (atlanta chicago denver houston losangeles
           miami newyork sanfran seattle washdc) (5.)
           @56 city $15.;
   datalines;
    0                                                   Atlanta
  587    0                                              Chicago
 1212  920    0                                         Denver
  701  940  879    0                                    Houston
 1936 1745  831 1374    0                               Los Angeles
  604 1188 1726  968 2339    0                          Miami
  748  713 1631 1420 2451 1092    0                     New York
 2139 1858  949 1645  347 2594 2571    0                San Francisco
 2182 1737 1021 1891  959 2734 2408  678    0           Seattle
  543  597 1494 1220 2300  923  205 2442 2329    0      Washington D.C.
;
   proc mds data=city level=absolute out=out;
      id city;
   run;
```

Conducting Multidimensional Scaling

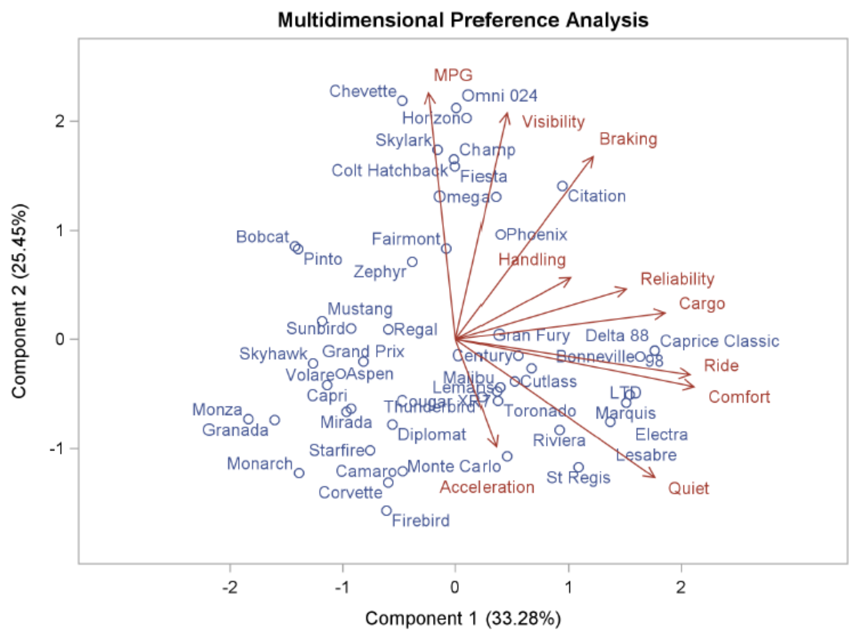


Relationship Among MDS, Factor Analysis, and Discriminant Analysis

- If the attribute-based approaches are used to obtain input data, spatial maps can also be obtained by using factor or discriminant analysis.
- By factor analyzing the data, one could derive for each respondent, factor scores for each brand. By plotting brand scores on the factors, a spatial map could be obtained for each respondent. The dimensions would be labeled by examining the factor loadings, which are estimates of the correlations between attribute ratings and underlying factors.
- To develop spatial maps by means of discriminant analysis, the dependent variable is the brand rated and the independent or predictor variables are the attribute ratings. A spatial map can be obtained by plotting the discriminant scores for the brands. The dimensions can be labeled by examining the discriminant weights, or the weightings of attributes that make up a discriminant function or dimension.

Perception Data: Derived Approaches by PCA

```
title 'Ratings for Automobiles Manufactured in 1980';

data cars;
   input Origin $ 1-8 Make $ 10-19 Model $ 21-36
         (MPG Reliability Acceleration Braking Handling Ride
          Visibility Comfort Quiet Cargo) (1.);
   datalines;
GMC      Buick     Century          3334444544
GMC      Buick     Electra          2434453555
GMC      Buick     Lesabre          2354353545

... more lines ...

GMC      Pontiac   Sunbird          3134533234
;

ods graphics on;

proc prinqual data=cars plots=all maxiter=100;
   transform monotone(mpg -- cargo);
   id model;
run;
```



Application of Dummy Regression – Conjoint Analysis

- **Conjoint analysis** attempts to determine the relative importance (weights) consumers attach to salient attributes and the utilities they attach to the levels of attributes.
- The respondents are presented with stimuli that consist of combinations of attribute levels and asked to evaluate these stimuli in terms of their desirability.
- Conjoint procedures attempt to assign values to the levels of each attribute, so that the resulting values or utilities attached to the stimuli match, as closely as possible, the input evaluations provided by the respondents.

Statistics and Terms Associated with Conjoint Analysis

- **Part-worth functions**. The part-worth functions, or *utility functions*, describe the utility consumers attach to the levels of each attribute.
- **Relative importance weights**. The relative importance weights are estimated and indicate which attributes are important in influencing consumer choice.
- **Attribute levels**. The attribute levels denote the values assumed by the attributes.
- **Full profiles**. Full profiles, or complete profiles of brands, are constructed in terms of all the attributes by using the attribute levels specified by the design.
- **Pairwise tables**. In pairwise tables, the respondents evaluate two attributes at a time until all the required pairs of attributes have been evaluated.

Statistics and Terms Associated with Conjoint Analysis

- **Cyclical designs**. Cyclical designs are designs employed to reduce the number of paired comparisons.
- **Fractional factorial designs**. Fractional factorial designs are designs employed to reduce the number of stimulus profiles to be evaluated in the full profile approach.
- **Orthogonal arrays**. Orthogonal arrays are a special class of fractional designs that enable the efficient estimation of all main effects.
- **Internal validity**. This involves correlations of the predicted evaluations for the holdout or validation stimuli with those obtained from the respondents.

Conducting Conjoint Analysis

Formulate the Problem
↓
Construct the Stimuli
↓
Decide the Form of Input Data
↓
Select a Conjoint Analysis Procedure
↓
Interpret the Results
↓
Assess Reliability and Validity

Conducting Conjoint Analysis Construct the Stimuli

- In the **pairwise approach**, also called *two-factor evaluations*, the respondents evaluate two attributes at a time until all the possible pairs of attributes have been evaluated.
- In the **full-profile approach**, also called *multiple-factor evaluations*, full or complete profiles of brands are constructed for all the attributes. Typically, each profile is described on a separate index card.
- In the pairwise approach, it is possible to reduce the number of paired comparisons by using cyclical designs. Likewise, in the full-profile approach, the number of stimulus profiles can be greatly reduced by means of fractional factorial designs.

Sneaker Attributes and Levels

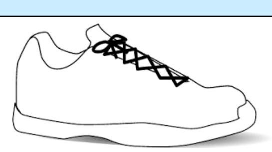
Table 21.2

Attribute	Level Number	Level Description
Sole	3	Rubber
	2	Polyurethane
	1	Plastic
Upper	3	Leather
	2	Canvas
	1	Nylon
Price	3	$30.00
	2	$60.00
	1	$90.00

Full-Profile **Approach to Collecting Conjoint Data**

Example of a Sneaker Product Profile

Sole	Made of rubber
Upper	Made of nylon
Price	$30.00



Conducting Conjoint Analysis Construct the Stimuli

- A special class of fractional designs, called orthogonal arrays, allow for the efficient estimation of all main effects. **Orthogonal arrays** permit the measurement of all main effects of interest on an uncorrelated basis. These designs assume that all interactions are negligible.
- Generally, two sets of data are obtained. One, the *estimation set*, is used to calculate the part-worth functions for the attribute levels. The other, the *holdout set*, is used to assess reliability and validity.

Conducting Conjoint Analysis Decide on the Form of Input Data

- For non-metric data, the respondents are typically required to provide rank-order evaluations.
- In the metric form, the respondents provide ratings, rather than rankings. In this case, the judgments are typically made independently.
- In recent years, the use of ratings has become increasingly common.
- The dependent variable is usually preference or intention to buy. However, the conjoint methodology is flexible and can accommodate a range of other dependent variables, including actual purchase or choice.
- In evaluating sneaker profiles, respondents were required to provide preference.

Sneaker Profiles & Ratings

	Attribute Levels [a]			
Profile No.	Sole	Upper	Price	Preference Rating (vs. Ranking)
1	1	1	1	9
2	1	2	2	7
3	1	3	3	5
4	2	1	2	6
5	2	2	3	5
6	2	3	1	6
7	3	1	3	5
8	3	2	1	7
9	3	3	2	6

[a] The attribute levels correspond to those in Table 21.2

Table 21.4 (**Fraction Factorial Design**)

Conducting Conjoint Analysis
Decide on the Form of Input Data

The basic **conjoint analysis model** may be represented by the following formula:

$$U(X) = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \alpha_{ij} x_{ij}$$

where

$U(X)$ = overall utility (attitude) of an alternative

α_{ij} = the part-worth contribution (weight) or utility associated with the j th level (j, j = 1, 2, . . . k_i) of the i th attribute (i, i = 1, 2, . . . m)

x_{jj} = 1 if the j th level of the i th attribute is present
= 0 otherwise

k_i = number of levels of attribute i

m = number of attributes

Conducting Conjoint Analysis
Decide on the Form of Input Data

The importance of an attribute, I_i, is defined in terms of the **range** of the part-worths, α_{ij}, across the levels of that attribute:

The attribute's importance is **normalized** to ascertain its importance relative to other attributes, W_i:

$$W_i = \frac{I_i}{\sum_{i=1}^{m} I_i}$$

So that $\sum_{i=1}^{m} W_i = 1$

The simplest estimation procedure, and one which is gaining in popularity, is dummy variable regression (see Chapter 17). If an attribute has k_i levels, it is coded in terms of k_i - 1 dummy variables (see Chapter 14).

Other procedures that are appropriate for non-metric data include **LINMAP**, **MONANOVA**, and the **LOGIT** model.

Conducting Conjoint Analysis
Decide on the Form of Input Data

The model estimated may be represented as:

$$U = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6$$

where

X_1, X_2 = dummy variables representing Sole
X_3, X_4 = dummy variables representing Upper
X_5, X_6 = dummy variables representing Price

For Sole the attribute levels were coded as follows:

	X1	*X2*
Level 1	1	0
Level 2	0	1
Level 3	0	0

Sneaker Data Coded for Dummy Variable Regression

Preference Ratings	Attributes					
	Sole		Upper		Price	
Y	X1	X2	X3	X4	X5	X6
9	1	0	1	0	1	0
7	1	0	0	1	0	1
5	1	0	0	0	0	0
6	0	1	1	0	0	1
5	0	1	0	1	0	0
6	0	1	0	0	1	0
5	0	0	1	0	0	0
7	0	0	0	1	1	0
6	0	0	0	0	0	1

Table 21.5

Conducting Conjoint Analysis
Decide on the Form of Input Data

The levels of the other attributes were coded similarly. The parameters were estimated as follows:

b_0 = 4.222
b_1 = 1.000
b_2 = -0.333
b_3 = 1.000
b_4 = 0.667
b_5 = 2.333
b_6 = 1.333

Given the dummy variable coding, in which level 3 is the base level, the coefficients may be related to the part-worths:

$$\alpha_{11} - \alpha_{13} = b_1$$
$$\alpha_{12} - \alpha_{13} = b_2$$

Conducting Conjoint Analysis
Decide on the Form of Input Data

To solve for the part-worths, an additional constraint is necessary.

$$\alpha_{11} + \alpha_{12} + \alpha_{13} = 0$$

These equations for the first attribute, Sole, are:

$$\alpha_{11} - \alpha_{13} = 1.000$$
$$\alpha_{12} - \alpha_{13} = -0.333$$
$$\alpha_{11} + \alpha_{12} + \alpha_{13} = 0$$

Solving these equations, we get,

$$\alpha_{11} = 0.778$$
$$\alpha_{12} = -0.556$$
$$\alpha_{13} = -0.222$$

Conducting Conjoint Analysis
Decide on the Form of Input Data

The part-worths for other attributes reported in Table 21.6 can be estimated similarly.

For Upper we have:

$$\alpha_{21} - \alpha_{23} = b_3$$
$$\alpha_{22} - \alpha_{23} = b_4$$
$$\alpha_{21} + \alpha_{22} + \alpha_{23} = 0$$

For the third attribute, Price, we have:

$$\alpha_{31} - \alpha_{33} = b_5$$
$$\alpha_{32} - \alpha_{33} = b_6$$
$$\alpha_{31} + \alpha_{32} + \alpha_{33} = 0$$

Conducting Conjoint Analysis
Decide on the Form of Input Data

The relative importance weights were calculated based on ranges of part-worths, as follows:

Sum of ranges of part-worths	= (0.778 - (-0.556)) + (0.445-(-0.556)) + (1.111-(-1.222))
	= 4.668

Relative importance of Sole	= 1.334/4.668 = 0.286
Relative importance of Upper	= 1.001/4.668 = 0.214
Relative importance of Price	= 2.333/4.668 = 0.500

Results of Conjoint Analysis

Table 21.6

	Level			
Attribute	No.	Description	Utility	Importance
Sole	3	Rubber	0.778	
	2	Polyurethane	-0.556	
	1	Plastic	-0.222	0.286
Upper	3	Leather	0.445	
	2	Canvas	0.111	
	1	Nylon	-0.556	0.214
Price	3	$30.00	1.111	
	2	$60.00	0.111	
	1	$90.00	-1.222	0.500

Conducting Conjoint Analysis
Interpret the Results

- For interpreting the results, it is helpful to plot the part-worth functions.
- The utility values have only interval scale properties, and their origin is arbitrary.
- The relative importance of attributes should be considered.

Conducting Conjoint Analysis
Assessing Reliability and Validity

- The goodness of fit of the estimated model should be evaluated. For example, if dummy variable regression is used, the value of R^2 will indicate the extent to which the model fits the data.
- Test-retest reliability can be assessed by obtaining a few replicated judgments later in data collection.
- The evaluations for the holdout or validation stimuli can be predicted by the estimated part-worth functions. The predicted evaluations can then be correlated with those obtained from the respondents to determine internal validity.
- If an aggregate-level analysis has been conducted, the estimation sample can be split in several ways and conjoint analysis conducted on each subsample. The results can be compared across subsamples to assess the stability of conjoint analysis solutions.

Part-Worth Functions

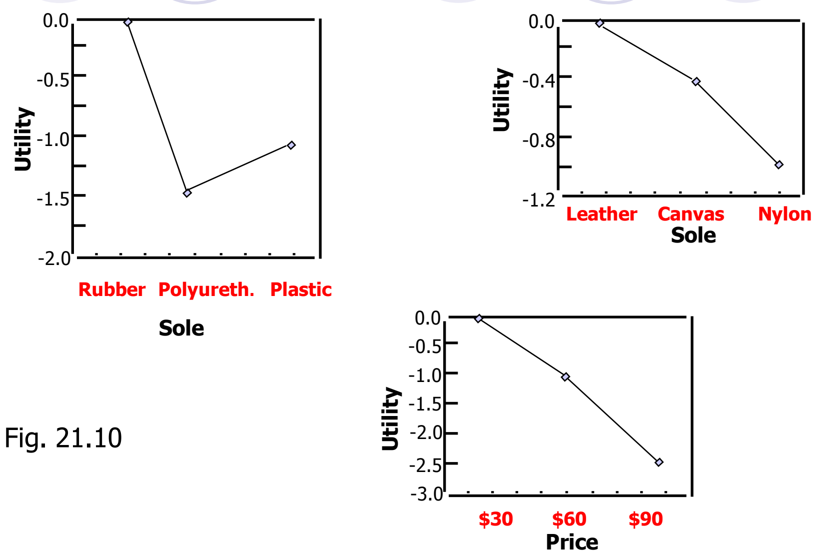


Fig. 21.10

Assumptions and Limitations of Conjoint Analysis

- Conjoint analysis assumes that the important attributes of a product can be identified.
- It assumes that consumers evaluate the choice alternatives in terms of these attributes and make tradeoffs.
- The tradeoff model may not be a good representation of the choice process.
- Another limitation is that data collection may be complex, particularly if a large number of attributes are involved and the model must be estimated at the individual level.
- The part-worth functions are not unique.

Hybrid Conjoint Analysis

- Hybrid models have been developed to serve two main purposes:
  1. Simplify the data collection task by imposing less of a burden on each respondent, and
  2. Permit the estimation of selected interactions (at the subgroup level) as well as all main (or simple) effects at the individual level.
- In the hybrid approach, the respondents evaluate a limited number, generally no more than nine, conjoint stimuli, such as full profiles.

Hybrid Conjoint Analysis

- These profiles are drawn from a large master design, and different respondents evaluate different sets of profiles, so that over a group of respondents, all the profiles of interest are evaluated.
- In addition, respondents directly evaluate the relative importance of each attribute and desirability of the levels of each attribute.
- By combining the direct evaluations with those derived from the evaluations of the conjoint stimuli, it is possible to estimate a model at the aggregate level and still retain some individual differences.

Choice-based Conjoint Analysis

- The fabric softener example is a small, somewhat more realistic example that discusses designing the choice experiment, randomization, generating the questionnaire, entering and processing the data, analysis, results, probability of choice, and custom questionnaires.
- The first vacation example is a larger, symmetric example that discusses designing the choice experiment, blocks, randomization, generating the questionnaire, entering and processing the data, coding, and alternative-specific effects.
- The second vacation example is a larger, asymmetric example that discusses designing the choice experiment, blocks, blocking an existing design, interactions, generating the questionnaire, generating artificial data, reading, processing, and analyzing the data, aggregating the data to save time and memory.
- The brand choice example is a small example that discusses the processing of aggregate data, the mother logit model, and the likelihood function.
- The food product example is a medium sized example that discusses asymmetry, coding, checking the design to ensure that all effects are estimable, availability cross effects, interactions, overnight design searches, modeling subject attributes, and designs when balance is of primary importance.
- The drug allocation example is a small example that discusses data processing for studies where respondents potentially make multiple choices.

Multinomial Logit, Discrete Choice Modeling

Candy Example

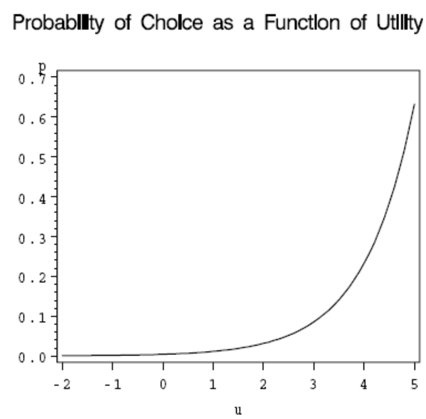
```
title 'Choice of Chocolate Candies';

data chocs;
   input Subj c Dark Soft Nuts @@;
   Set = 1;
   datalines;
 1 2 0 0 0    1 2 0 0 1    1 2 0 1 0    1 2 0 1 1
 1 1 1 0 0    1 2 1 0 1    1 2 1 1 0    1 2 1 1 1
 2 2 0 0 0    2 2 0 0 1    2 2 0 1 0    2 2 0 1 1
 2 2 1 0 0    2 1 1 0 1    2 2 1 1 0    2 2 1 1 1
 3 2 0 0 0    3 2 0 0 1    3 2 0 1 0    3 2 0 1 1
 3 2 1 0 0    3 2 1 0 1    3 1 1 1 0    3 2 1 1 1
```

Cox proportional hazards model

```
proc phreg data=chocs outest=betas;
   strata subj set;
   model c*c(2) = dark soft nuts / ties=breslow;
   label dark = 'Dark Chocolate' soft = 'Soft Center'
         nuts = 'With Nuts';
   run;
```

$$\lambda_i(t) = \lambda(t;\mathbf{Z}_i) = \lambda_0(t)\exp(\mathbf{Z}_i'\boldsymbol{\beta})$$

$$S(t;\mathbf{Z}_i) = [S_0(t)]^{\exp(\mathbf{Z}_i'\boldsymbol{\beta})}$$



where $\mathbf{x}_i$ is a vector of alternative attributes and β is a vector of unknown parameters. $U(c_i) = \mathbf{x}_i\beta$ is the utility for alternative c_i, which is a linear function of the attributes. The probability that an individual will choose one of the m alternatives, c_i, from choice set C is the exponential of the utility of the alternative divided by the sum of all of the exponentiated utilities.

There are $m = 8$ attribute vectors in this example, one for each alternative. Let **x** = (Dark/Milk, Soft/Chewy, Nuts/No Nuts) where Dark/Milk = (1 = Dark, 0 = Milk), Soft/Chewy = (1 = Soft, 0 = Chewy), Nuts/No Nuts = (1 = Nuts, 0 = No Nuts). The eight attribute vectors are

$\mathbf{x}_1 = (0\ 0\ 0)$ (Milk, Chewy, No Nuts)
$\mathbf{x}_2 = (0\ 0\ 1)$ (Milk, Chewy, Nuts)
$\mathbf{x}_3 = (0\ 1\ 0)$ (Milk, Soft, No Nuts)
$\mathbf{x}_4 = (0\ 1\ 1)$ (Milk, Soft, Nuts)
$\mathbf{x}_5 = (1\ 0\ 0)$ (Dark, Chewy, No Nuts)
$\mathbf{x}_6 = (1\ 0\ 1)$ (Dark, Chewy, Nuts)
$\mathbf{x}_7 = (1\ 1\ 0)$ (Dark, Soft, No Nuts)
$\mathbf{x}_8 = (1\ 1\ 1)$ (Dark, Soft, Nuts)

Say, hypothetically that $\beta' = (4 \quad -2 \quad 1)$. That is, the part-worth utility for dark chocolate is 4, the part-worth utility for soft center is -2, and the part-worth utility for nuts is 1. The utility for each of the combinations, $\mathbf{x}_i\beta$, would be as follows.

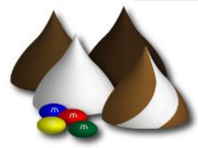
$$
\begin{array}{lcccccccc}
\text{U(Milk, Chewy, No Nuts)} & = & 0\times 4 & + & 0\times -2 & + & 0\times 1 & = & 0 \\
\text{U(Milk, Chewy, Nuts)} & = & 0\times 4 & + & 0\times -2 & + & 1\times 1 & = & 1 \\
\text{U(Milk, Soft, No Nuts)} & = & 0\times 4 & + & 1\times -2 & + & 0\times 1 & = & -2 \\
\text{U(Milk, Soft, Nuts)} & = & 0\times 4 & + & 1\times -2 & + & 1\times 1 & = & -1 \\
\text{U(Dark, Chewy, No Nuts)} & = & 1\times 4 & + & 0\times -2 & + & 0\times 1 & = & 4 \\
\text{U(Dark, Chewy, Nuts)} & = & 1\times 4 & + & 0\times -2 & + & 1\times 1 & = & 5 \\
\text{U(Dark, Soft, No Nuts)} & = & 1\times 4 & + & 1\times -2 & + & 0\times 1 & = & 2 \\
\text{U(Dark, Soft, Nuts)} & = & 1\times 4 & + & 1\times -2 & + & 1\times 1 & = & 3
\end{array}
$$

The denominator of the probability formula, $\sum_{j=1}^{m} \exp(\mathbf{x}_j\beta)$, is $\exp(0) + \exp(1) + \exp(-2) + \exp(-1) + \exp(4) + \exp(5) + \exp(2) + \exp(3) = 234.707$. The probability that each alternative is chosen, $\exp(\mathbf{x}_i\beta)/\sum_{j=1}^{m} \exp(\mathbf{x}_j\beta)$, is

$$
\begin{array}{lccccc}
\text{p(Milk, Chewy, No Nuts)} & = & \exp(0) & / \ 234.707 & = & 0.004 \\
\text{p(Milk, Chewy, Nuts)} & = & \exp(1) & / \ 234.707 & = & 0.012 \\
\text{p(Milk, Soft, No Nuts)} & = & \exp(-2) & / \ 234.707 & = & 0.001 \\
\text{p(Milk, Soft, Nuts)} & = & \exp(-1) & / \ 234.707 & = & 0.002 \\
\text{p(Dark, Chewy, No Nuts)} & = & \exp(4) & / \ 234.707 & = & 0.233 \\
\text{p(Dark, Chewy, Nuts)} & = & \exp(5) & / \ 234.707 & = & 0.632 \\
\text{p(Dark, Soft, No Nuts)} & = & \exp(2) & / \ 234.707 & = & 0.031 \\
\text{p(Dark, Soft, Nuts)} & = & \exp(3) & / \ 234.707 & = & 0.086
\end{array}
$$

Note that even combinations with a negative or zero utility have a nonzero probability of choice. Also note that adding a constant to the utilities will not change the probability of choice, however multiplying by a constant will.

One-to-one customization



Choice of Chocolate Candies

Obs	Dark	Soft	Nuts	p
1	Dark	Chewy	Nuts	0.50400
2	Dark	Chewy	No Nuts	0.21600
3	Milk	Chewy	Nuts	0.12600
4	Dark	Soft	Nuts	0.05600
5	Milk	Chewy	No Nuts	0.05400
6	Dark	Soft	No Nuts	0.02400
7	Milk	Soft	Nuts	0.01400
8	Milk	Soft	No Nuts	0.00600