A Study of Statistical Tests Application to Conjoint Analysis

MA Kaijie * and Dai Min

Finland College of Computer and Computing Science
Zhejiang University City College, China

***Abstract* — In this paper, based on the theory of conjoint analysis, it introduces the basic theory and method of conjoint analysis. The article focuses on the application and theoretical analysis of conjoint analysis and complements from the perspective of statistical hypothesis testing based on the principles and methods of conjoint analysis. It also gives a certain test methods of statistical hypothesis testing. Results showed that the conjoint analysis is a better survey method. It also sum up the whole article on the joint analysis of the application in practice to make recommendations and to look forward the joint analysis of development prospects.**

Keywords - Consumer Preferences; Conjoint analysis; Statistical test

I. INTRODUCTION

The idea of conjoint analysis is introduced firstly by psychologist Luce and statistician Tukey in 1964. In the year of 1972, Green, Wind and Jain use its application to the commercial field and have achieved good results [1]. Then it received widespread attention in the field of consumer research and a wide range of applications as for its unique advantages. For 40 years, a large number of researchers continue to improve it, and a variety of analysis model matures, so it became the research one of the most popular tools in the market today. The method of conjoint analysis is to suppose the product which has some attributes, carry on the evaluation to the realistic product and these virtual products according to their own favorite, and adopt of the method of mathematics statistics to separate from the utility of the attribute and attribute level, thus make evaluation of quantification on the important degrees of every attribute and attribute levels. Conjoint analysis is a popular marketing research technique that marketers use to determine what features a new product should have and how it should be priced. Conjoint analysis became popular because it was a far less expensive and more flexible way to address these issues than concept testing.

Conjoint analysis was originated in mathematical psychology and was developed by marketing professor Paul Green at the Wharton School of the University of Pennsylvania and Data Chan. Other prominent conjoint analysis pioneers include professor V. “Seenu” Srinivasan of Stanford University who developed a linear programming (LINMAP) procedure for rank

ordered data as well as a self-explicated approach, Richard Johnson who developed the adaptive conjoint analysis technique in the 1980s [2, 3] and Jordan Louviere (University of Iowa) who invented and developed Choice-based approaches to conjoint analysis and related techniques such as maxdiff.

Conjoint analysis of the emergence is developed due to the rapid development of consumer behavior as the same of cluster analysis and multidimensional scaling analysis, thanks to [4]. Solving the problem of market segmentation cluster analysis is to determine the target audience, which relies on the data-similar needs, attitudes, and interests and so on. Multidimensional scaling approach to diagnosis, cognition and preference map is constructed. Its role is to describe rather than a prediction, researchers want to use it to study if a new product is unlikely, because it can get the individual properties of the product, and it’s unable to get trade-off between individual properties. The problem can be quantitatively solved by conjoint analysis which is a realistic simulation of the consumer purchase decision [5]. In the actual decision making process, due to some reasons such as price, consumer products multiple attributes are to be considered. In order to meet a number of requirements of the premise, it must sacrifice parts of its other properties and get the trade-off.

Conjoint analysis is a good method of evaluation of consumer preferences, it uses decomposition approach by allowing consumers to assign values to a series of product profiles to calculate the bias parameter which can be a value, weight and the ideal point [7]. Because there is a pervasive individual differences in consumer preference, conjoint analysis usually is used only on an individual level. [8]

II. THEORY OF CONJOINT ANALYSIS

Conjoint analysis is known to consumers on the basis of evaluation results for full-sized, after decomposition method to estimate the preference structure analysis. Analysis of conjoint analysis in the following purposes: (l) determining that consumers which are given a contribution and effectiveness of the predictor variables and the relative importance of property, (2) finding the best mix of products in market to the consumer this combination may not be initially evaluated by consumers. (3) Simulating market to estimate the market share and market share changes.

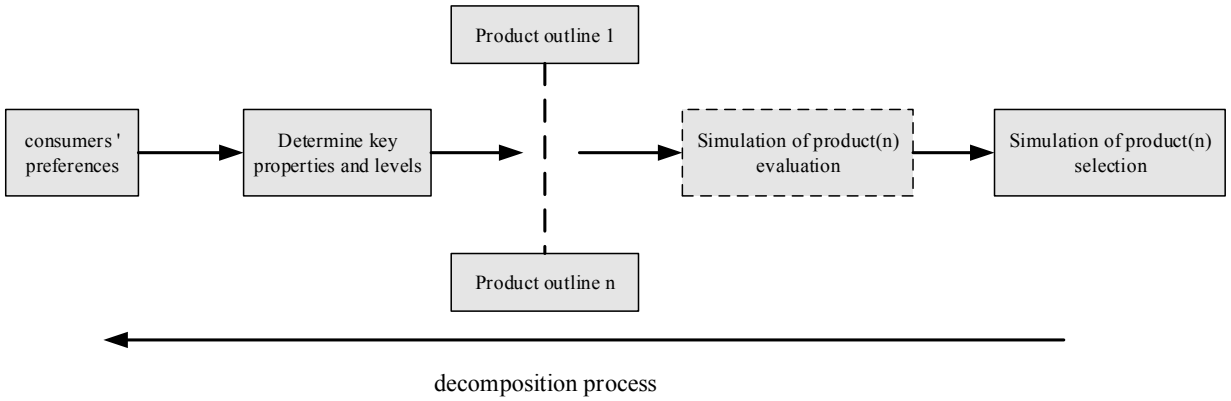

Fig. (1). General Procedure for Conjoint Analysis

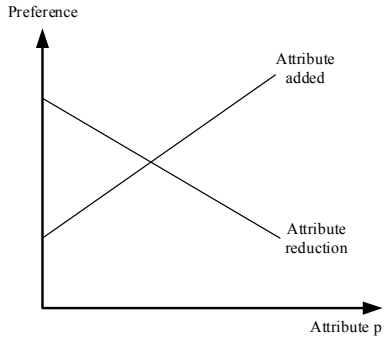

Fig. (2). Vector Model

To calculate the effectiveness of each product in the minds of consumers, it firstly calculate the composition of the product's effectiveness score for each attribute level. There are generally three ways: the vector model, the ideal point model and path-worth function model. Or you can mix them up sometimes that allow some of the properties function to calculate scores, while other properties calculated to be use vector model or ideal point model.

There are some of the indicators calculation method of conjoint analysis.

1) Horizontal utility. Horizontal utility describes the level of importance of consumer empowerment at all levels for each property. Each horizontal utility by conjoint analysis model estimation. Regular conjoint analysis model can be represented by the following formula:

$$Y = a = \sum vX \tag{1}$$

Where: Y indicates the full-sized preference score and a indicates intercept. v is estimation effectiveness and X is to specify different attribute level of dummy variables.

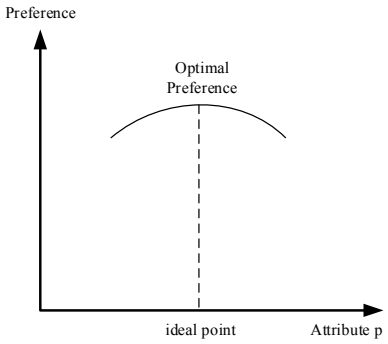

Fig. (3). Ideal Point Model

2) Properties of relative importance. It is based on such an assumption: higher margin indicates the property's importance in the overall outline of the higher margin as small as possible. Properties generally are expressed as a percentage of the relative importance. The formula is as follows:

$$W_j = \frac{Max(v_{ij}) - Min(v_{ij})}{\sum_{j=1}^{J}\left[Max(v_{ij}) - Min(v_{ij})\right]} \times 100\% \tag{2}$$

Where: W_j indicates the properties of importance in j, Max (v_{ij}) is the properties of maximum level of efficiency values in j, Min (v_{ij}) is the minimum level for the attribute utility values in j.

3) Profile utility. In generally the most common and most basic model to calculate the overall profile utility of conjoint analysis model is addition mode which believes that consumers just add up the value of each property to get a property portfolio's total value.

$$U(X) = \sum_{j}^{J}\sum^{I} v_{ij}X_{ij} \tag{3}$$

Where: $U(x)$ indicates the total utility, x_{ij} indicates a binary variable, and v_{ij} indicates the properties of the level i of effectiveness in j .

4) Simulation to estimate the market share. Many in the study of conjoint mission analysis which access to the utility value of the attribute levels often are not the ultimate goal of marketing research, and also look for the best mix of marketing of the product to simulate consumer market selection and estimation of the market share. Market share should be based on consumer preferences and market share, most commonly method used in the simulation market is the largest utility model.

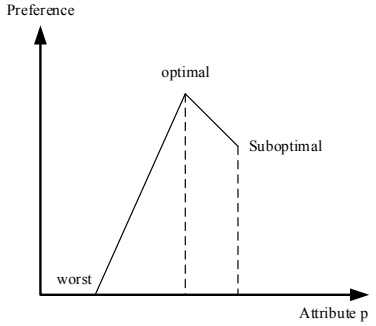

Fig. (4). Path-worth Function Model

Conjoint analysis is done by assuming that the product has certain characteristics to simulate the real-world products, and then lets the consumers to evaluate these virtual products according to their preferences, and uses statistical methods to separate the utility of these attributes and attribute levels which to estimate the importance of each attribute and attribute levels in a quantitative evaluation method. It usually consists of the following parts:

a) Determining the attributes and attribute level and determining the characteristics of a product or service attributes. The properties should be determined by the outstanding properties of influencing consumers' preferences. If there are many properties, it is heavier burden on respondents and reduce predictive accuracy of the model. If there is too little property, lost several key pieces of information may vary depending on model and reduce the predictive power of the model. The next step is to select the level, properties and property level to determine the number of conjoint analysis to estimate the number of parameters in the process which also influence the respondent to assess the number of outline of products or services. In order to guarantee the precision of parameter estimation and reduce the burden on respondents and test appropriate to attribute the number of levels.

b) Experimental design and product simulation. For number in three or more of the property of multi-attribute product simulations and under certain conditions, you can choose one of the property taken into account and do rigorous statistical analysis. Orthogonal experiment is applied to the conjoint analysis used to reduce the number of analog products. Orthogonal experiment is orthogonal to schedule and design experiments. Orthogonal is a normalized form which must have three characteristics: (1) each equal to the number of occurrences in each column; (2) different combinations of any two columns level repeat the same way; (3) elementary orthogonal transformation without changing its orthogonally. For instance, table 1 is the pilot experience of $L_9(3^4)$.

TABLE 1. ORTHOGONAL TABLE

Test number	Column number			
	1(A)	2(B)	3(C)	4(D)
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

c) Statistical test. We use statistical test to establish conjoint analysis statistical test system of practicality of the applicability of the models and reset the reliability and validity of the model of the scientific evaluation. We use specialized statistical analysis procedures to fit the model. Once the model is built, in order to evaluate at the level of individual consumers and consumer groups' level validity of conjoint analysis the model must be an analysis of the reliability and validity of the results.

III. HYPOTHESIS TESTING METHODS IN CONJOINT ANALYSIS

According to the different types of conjoint analysis, using different statistical methods such as ordinary least squares, weighted least squares and logit analysis to make the respondents' answers into the importance or utility. Using statistical methods to get actual numbers are not the most important, the more important thing is to associate with the various property values or the relationship between the properties. First of all, in order to examine practicability of its theoretical completeness fly method and scope as well as the results of scientific.

There are two aspects to hypothesis testing in conjoint analysis. This paper is intended to provide a starting point, the selection of joint analysis of object properties, the property relations, as well as combining analysis model method of hypothesis testing.

By examining the representativeness of each property and judging the importance of each attribute on clients, deleting or ignoring the judgment did not affect or influence a small property to reduce the actual amount of workload and client tasks in the survey.

The null hypothesis H0: "Assessment of the different attributes are different of the same number of points". And the other hypothesis H1: "their assessment as to the

individual properties are the same or largely consistent", we use collaborative coefficient as the variables investigated. Collaborative coefficient is defined as follows:

$$W = \frac{12S}{m^2(n^3 - 1)} \tag{4}$$

$$S = \sum_{i=1}^{n} (R_i - \frac{m(n+1)}{2})^2 \tag{5}$$

If the importance of property is higher, its proportion factors influencing consumer choice is greater and it can serve as a yardstick for measuring select properties. Based on levels of various properties of the utility, we can use the above formula (2) to determine the importance of property.

Testing the correlation coefficient between variables do not assume that all are equal or not equal. We consider if it is linear or nonlinear relationships between them and a deeper examination of the correlation matrix to them.

Original hypothesis H0: "Attributes have some orthogonally "; alternative hypothesisH1: “Attributes have no orthogonally". Orthogonal test of test statistics is as follows:

$$F = \frac{X / m}{Y / n} = \frac{nX}{mY} \tag{6}$$

When the quadratic regression model of an f-test is remarkable indicating regression equation at a level significantly, we can reject null hypothesis that experimental data and mathematical model are basically in line with.

By calculating the credibility and effectiveness of the model, conjoint analysis results are verified at the level of individual consumers and consumer groups at the level of accuracy. Verification validity include two parts: internal validity and external validity. Internal validity is to verify the goodness of fit of the model as well as the contour utility appropriate combination rule and external validity was verified on a population representative.

Original hypothesis H0: "the goodness of fit of the model is good"; the alternative hypothesis H1: "the goodness of fit of the model is very poor”. Measuring the goodness of fit test statistic is the coefficient of determination R2. In general, the goodness of fit which above 0.80 is appropriate.

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (y_i - \hat{Y}_i)^2} \tag{7}$$

For coefficient of determination or coefficient of determination, the range of values of the coefficient of determination is (0, 1). If it is closer to 1 that specifies the actual observations from sample line closer to high goodness of fit.

For sorting and rating data, the model can be used as a predictive value to product evaluation. Firstly considering if there are linear or nonlinear relationships between them and their correlation matrix for further inspection. Calculation of the predicted values with the actual consumers assess the degree of correlation between the values. Null hypothesis H0: "the predicted value and the actual value is relevant"; alternative hypothesis H1: "the predicted value and the actual value is not relevant or less”. There are three methods for testing: 1) Pearson correlation. The test statistic is as follows:

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} \tag{8}$$

If r is close to 0, it means that the linear relationship between x and y close lower; if r is close to 1, then x and y is a linear relationship between higher levels of close. 2) Spearman correlation. Regardless of the overall distribution of two variables, it can be studied using Spearman rank correlation, the test statistic is as follows:

$$r_p = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{9}$$

3) Kendall correlation. Adapted two categorical variables are ordinal classification of cases, non-parametric correlation of order related to the variable inspection, the test statistic is as follows:

$$T = \frac{4p}{n(n-1)} - 1 \tag{10}$$

In the consumer research, preference value of measurement is based on the theory of expectation model. We can use the model to evaluate the reliability and validity of the results in order to assess the individual and group levels of validity of conjoint analysis model. Conjoint analytical reliability usually include time reliability, properties, reliability, and attribute level reliability and methods of data collection reliability.

IV. APPLICATION AND PROSPECTS

Let's look at a simple example, a consumer want to buy a computer, so he went to the store and see the configuration of a computer which is like this: brand, Dell; processor, P4-3GHz; memory, 2MB; display 21-inch; hard disk, 80G; price 11,198 Yuan.

TABLE 2. INTERNAL EVALUATION OF PERSONAL COMPUTER

attribute	scores
Dell	20
Processor(P4-3GHz)	20
RAM(512M)	5
Display(21-inch)	15
Hard disk(80G)	10
Price(11198 Yuan)	30
Total utility	100

When we use the conjoint analysis to solve the problem, we can see:

TABLE 3. LEAST SQUARES REGRESSION MODEL

attribute	Level	utility	weights
Brand	IBM	3.4444	19.53%
	Dell	-4.8889	
	HP	1.4444	
Processor	2 GHz	-5.8889	21.87%
	3 GHz	2.4444	
	4 GHz	3.4444	
RAM	256 MB	-2.5556	16.41%
	512 MB	-1.8889	
	1 GB	4.4444	
Price	8798	10.4444	42.19%
	11198	-2.8889	
	13598	-7.5556	

Then we can get the result as follows:

Compared to traditional survey methods, conjoint analysis requires respondents to evaluation-only property portfolio into the overall evaluation of the product which is not required to evaluate individual properties, so it is close to the real situation. Conjoint analysis has distinct advantages:

(1) Objectivity, conjoint analysis is a method of directly requesting consumer products for various combinations of properties with different levels of preference ordering, analysis of the relative importance of each attribute, far more objective and realistic.

(2) Efficiency, conjoint analysis can be done by orthogonal design and draw all kinds of properties the best combination to minimize the number of combinations to get as much information as possible to greatly improve the efficiency/price ratio. Conjoint analysis can also appear in the sample properties of market share simulation to fully assess the potential competitiveness of new or improved products and the possible extent of the impact of the present market conditions.

(3) Adaptability, conjoint analysis can differ greatly in character properties for conjoint evaluation of these properties vary greatly as long as which can be represented as a different category of the property whether qualitative or quantitative indicators can co-exist.

(4) Biodegradable, evaluation of the respondents make preferences, the researcher will total preference level of decomposition for various attributes of utility value while traditional discriminant analysis and regression analysis is synthesis model where respondents were required to evaluate level of each attribute. And the establishment of prediction model can analyze the relationship between each attribute scores and preferences.

(5) Multiple attributes can be done individual level of analysis. Conjoint analysis analyze evaluation results for each respondent and produce separate models to calculate predictive accuracy of individual preference and also establish the total comprehensive information of each individual model. While most other multivariate analysis methods are observe a certain amount of individual reactions and comprehensive analysis of each individual model.

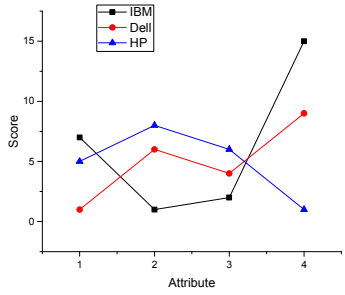

Fig. (5). Least Squares Regression Model

Business in the Simulator is based on conjoint analysis for the application of the principle of software. It is possible to predict a range of market issues in the simulator in the market. Firstly generation market simulator is very simple, it is limited to the product profile and enter a series of consumer score matrix. Market simulation results is a brand's market share and the sensitivity of this information and they are limited. After more than 40 years of development and application of the Simulator, it have become more precise now which fully taking into account the reaction of competitors and allowing participants to use. Application of conjoint analysis is more than just consumers, it has been widely used in tourism, recreation, health care, gaming, complex product and system, lawsuits, and many other areas.

ACKNOWLEDGMENT

This work is supported by Key Project of Guangxi Social Sciences, China (No.gxsk201424), the Education Science fund of the Education Department of Guangxi, China (No.2014JGA268), and Guangxi Office for Education Sciences Planning, China (No.2013C108).

REFERENCES

[1] Green, P. and Srinivasan, V, "Conjoint analysis in consumer research: Issues and outlook", Journal of Consumer Research, vol 5, pp 103–123 , September 1978.

[2] Carroll, Doug, Arabie P, Chaturvedi A. "Multidimensional Scaling and Clustering and their Applications". Advances in Marketing Research: Progress and Prospects, 2002.

[3] Arora, Neeraj, Huber J. "Improving Parameter Estimates and Model Prediction by Aggregate Customization in Choice Experiments". Journal of Consumer Research, 2001, 28: 273~283.

[4] Dijkstra J, Roelen W, Timmermans H. "Conjoint measurement in virtual environments: a framework". Design and Decision Support Systems in Architecture and Urban Planning Conference, 1996, 1: 59~71.

[5] John R, Vithala R. "Conjoint Analysis, Related Modeling,and Applications". Advances in Marketing Research: Progress and Prospects, 2002, 9~23.

[6] Hofstede T, Frenkel, Kim Y. "Bayesian Prediction in Hybrid Conjoint Analysis". Journal of Marketing Research, 2002, 36: 253~261.

[7] Bridges J F P, Hauber A B, Marshall D, et al. "Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force". Value in Health, 2011, 14(4): 403-413.

[8] Hainmueller J, Hopkins D J, Yamamoto T. “Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments”. Political Analysis, 2014, 22(1): 1-30.

[9] Gadioli I L, Pineli L L O, Rodrigues J D S Q, et al. “Evaluation of packing attributes of orange juice on consumers' intention to purchase by conjoint analysis and consumer attitudes expectation”. Journal of Sensory Studies, 2013, 28(1): 57-65.

[10] Claret A, Guerrero L, Aguirre E, et al. “Consumer preferences for sea fish using conjoint analysis: Exploratory study of the importance of country of origin, obtaining method, storage conditions and purchasing price”. Food Quality and Preference, 2012, 26(2): 259-266.