Middle-East Journal of Scientific Research 9 (3): 431-436, 2011
ISSN 1990-9233

Estimation of Markov Chains Transition Probabilities Using Conjoint Analysis (Expert Preference)

N. Akhondi, G. Yari, E. Pasha and R. Farnoosh

Department of Statistics, Science and Research Branch, Islamic Azad University, Teharan, Iran

Abstract: This paper proposes methodology to estimate transition probabilities on the base of judgments by experts that may be useful in situations of data absence. The Fractional Factorial Design (FFD) is used to cope with the curse of dimensionality. By means of Conjoint Analysis (CA) approach we finally reconstruct the complete Markov Chain transition probabilities. The experiment results show it is promising for us to use (CA) in estimating of the entropy rate of Markov Chains with a finite state space.

Key words: Markov Chain · Transition probabilities · Conjoint Analysis · Design of experiments · Efficient Design in Conjoint Analysis

INTRODUCTION

The present paper proposes a framework based on expert opinion elicitation, developed to estimate the transition probability matrix of an irreducible, discrete time, homogeneous Markov Chain with a finite state space. In this article we address the question of estimating the transition probability matrix of Markov Chain in situations of data absence. In general, the full probability distribution for a given stochastic problem is unknown. When data are available, the most objective estimation of them is the maximumlikelihood estimation of the transition probabilities (P_{ij}).

The difficulties grows when the aim is providing scenarios analysis involving future states perhaps never performed before. In this situation we need information gathered from experts and we cannot resort to past data [1]. Our methodology has the new idea of estimating transition probabilities using conjoint (FFD) methods that is useful in this conditions.

Conjoint analysis has as its roots the need to solve important academic and industry problems [2]. It is a popular marketing research technique. In order to respond to consumers' needs, makers have to research consumers' preferences of products, services and their selection criteria of products. The conjoint analysis measures the degree of importance which is given to particular aspects of a product or service [3]. The real genius is making appropriate tradeoffs so that real consumers in real market research settings are answering questions from which useful information can be inferred. In the thirty years since the original conjoint analysis articles, researchers in marketing and other disciplines, have explored these tradeoffs [2]. In conjoint experiments, each respondent receives a set of profiles to rate (or rank). Designing these experiments involves determining how many and which profiles each respondent has to rate (or rank) and how many respondents are needed [4]. Experimental design is a fundamental component of (CA). The complexity of experimental design arises from the exponential growth in the number of attributes, i.e. *the curse of dimensionality.* Use of a full factorial design (all profiles) will place an excessive burden on respondent for providing evaluations. Therefore, researchers utilize (FFD), i.e. fractional balanced orthogonal design, or a subset of all profiles [5]. The basic conjoint problem is to estimate the partworths that best explain the overall preference judgments made by respondents [2]. (CA) is a technique based on a main effects analysis-of-variance model that decomposes the judgment data into components, based on qualitative attribute of the products or services [6]. Most commonly used methods to acquire partworths are the Linear Programming Technique for Multidimensional Analysis of Preference (LINMAP), Hierarchical Bayes (HB) methods, Multivariate Analysis of Variance (MANOVA) *and* Ordinary Least Squares (OLS) Regression [7].

Corresponding Author: N. Akhondi, Department of Statistics, Science and Research Branch, Islamic Azad University, Teharan, Iran. E-mial: akhondinasrin@gmail.com.

Methods have been developed to take conjoint data to approach to optimal or near-optimal products and systems designs in tourism, entertainment, health maintenance, gambling and etc. We introduce a new application of (CA) in this article.

Leone and Fucili [1] used (CA) to estimate Markov Chains transition probabilities.They built Fractional Factorial Designs (FFD) on the starting states and in their method the experts are asked to identify the presumably destination states and quantify the probability of occurrence of the transitions towards each destination proposed scenarios(for each state included in a (FFD)). In their method psychological critics may be raised because of the respondents are not asked according to the well procedure of rating or ranking the (FFD) treatments by comparing each one of them. The difficulties may grow if the number of attributes grows. We overcome the difficulties in this paper by building 2 (FFDs) on the starting and destination states and we ask experts to give ratings on the likelihood of various states occurring in the future. we need two sets of states to get transition probabilities. The Fractional Factorial Design (FFD) can tackle the large number of states in an elegant way. We used (CA) approach and Logistic Regression to construct the complete Markov Chain transition probabilities (under the assumptions of independence of the attributes at an individual level for the respondents). The conjoint methods used for ratings data are now essentially dummy variable regression methods.

The remainder of this paper is organized as follows: First, we provide a review of (CA) and references to related substantial theoretical and empirical work, then we discuss our methodology and simulation data. Finally we conclude the most important strengths and weaknesses of the proposed methodology**.** We used Minitab (ver:15) for generating (P_{ij}), SPSS (ver:16) for generating 2 (FFDs) and Multivariate Analysis of Variance (MANOVA) for estimating parameters.

Conjoint Analysis: The essence of conjoint analysis is to identify and measure a mapping from more detailed descriptors of a product or service onto a overall measure of the customer’s evaluation of that product. Full-profile analysis remains the most common form of conjoint analysis and has the advantage that the respondent evaluates each profile holistically and in the context of all others profile. Its weakness is that the respondent’s burden grows dramatically with the number of profiles that must be ranked or rated. The respondent can be asked to rank order all stimuli (profiles) or to provide a metric rating of each stimulus [2]. When appropriate, efficient experimental designs, (FFD) are used so that the respondent need consider only a small fraction of all possible product profiles]8.[If the number of attributes is large often respondents can evaluate partial profiles (PP) in which some of the features are explicit and the other features are assumed constant [2]. (FFD) orthogonal arrays are categorized by their resolution. For example, resolution III designs enable the estimation of all main effects free of each other, but some of them are confounded with two-factor interactions. Higher resolution designs require larger number of profiles. Resolution III designs are most frequently used in marketing conjoint studies. Orthogonal arrays can be either balanced or unbalanced in terms of levels of attributes. An unbalanced design gives larger standard errors the parameter estimates for those attributes that are less frequently administered. The minimum standard error is attained when a full factorial design is used [5]. Various measures for discussing the efficiency of an experimental design can be described as follows for the linear model (Kuhfeld, Tobias and Garratt 1994),

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{1}$$

Where β is a px1 vector of parameters, X is an nxp design matrix, and e is random error. With the usual assumption on errors, the least squares estimate of β is given by $(X'X)^{-1}X'Y$. The variance-covariance matrix of the partworth (parameter) estimates of the attributes is proportional to $(X'X)^{-1}$. The efficiency of a design is based on the information matrix. An efficient design will have a smaller variance matrix.Two famous efficiency measures (all based on the eigenvalues of $(X'X)^{-1}$) are:

$$\text{A-efficiency}=1/n(\text{trace}\,((X'X)^{-1})\,/p);$$
$$\text{D-efficiency}=1/n|(X'X)^{-1}|^{1/p}$$

Orthogonal designs for linear models are generally considered to be efficient because their efficiency measure is close to 1 [5].

Jeng-Jong Lin [7] successfully presented an integrated product design model to be applied in clothing product design. His methodology focused not only on either expertise of designers or demands of consumer but on both of them. He used relationship matrix to combine both the (CA) data from the two individual groups (i.e., designer and consumer) to design product. Van Houtven *et al.* [9] applied (CA) to estimate health-related benefit-risk tradeoffs in a non-expected-utility framework. They demonstrate how this method can be used to test for and estimate nonlinear weighting of adverse-event probabilities. Jeremy J. Michalek *et al.* [10] presented a unified methodology for product line optimization that coordinates positioning and design models to achieve realizable firm-level optima. This method is demonstrated for a line of dial-readout scales, using physical models and conjoint-based consumer choice data. Hiromi Yamada *et al.* [3] administrated the study to estimate the structure of the variable to specify the quality requirement of the new product using the conjoint analysis and the entropy model. As a result, it was understood that the conjoint analysis and the entropy model are effective methods to estimate the quality requirement. Lekemoto and Yomaoka [11] proposesd a method of analysis by using (CA) that makes it possible to use a lower number of profile cards than that provided by the orthogonal design of experiment even when a large number of items is being surveyed. An Internet survey of 1,600 consumers using this method indicated that it generated identical analytical results to those produced when the orthogonal design of experiment was used. Byungun Yoon and Yongtae Park [12] applied a new hybrid approach that enhances the performance of morphology analysis (MA) by combining it with conjoint analysis (CA) and citation analysis of patent information. Alternatives for new technology development from among the emerging technologies are presented by combining the valuable levels of each attribute in a morphology matrix predefined by domain experts. The technological competitiveness of a company can be analyzed by a newly suggested index, "technology share," which is analogous to the concept of market share in traditional CA.

Proposed Methodology: Expert opinion is one of the key research areas in Probabilistic risk analysis (PRA) in engineering, public health, environment, program management, regulatory policy, finance and etc. The use of expert judgment is critical, and often inevitable, when there are no empirical data or information available on the variables of interest [13]. We illustrat our motivation for resorting to experts in this section.

Assume a dynamic system with components , a set of states S, a set of actions Á, a reward function. S x Á → R, a transition probability matrix that the full probability distribution for this stochastic problem is unknown. We restrict attention to time separable Markovian decision problems for dealing with the curse of dimensionality in dynamic systems [14]. An irreducible discrete time Markov chain with a finite state space has been studied previously by Papangelou [15], who establishes a Large Deviation Principle (LDP) for Markov chains whose order is unknown. Baris Tan and Kamil Yilmaz [16] presented a complete analytical framework for the testing procedure based on statistical theory of Markov chains. They studied the time dependence and time homogeneity properties of the Markov chains.

We assume an irreducible, discrete time, homogeneous Markov Chain with a finite state space that the starting and the destination states of the system are defined by combinations of "n" key attributes, each with $\mathbf{L_k}$ (K=1,…,n) levels (continuous variables are discretizing). For example, if n=3 the starting state **i** and the destination state **j** are given by $\mathbf{l_{i1}}, \mathbf{l_{i2}}, \mathbf{l_{i3}}, \mathbf{l_{j1}}, \mathbf{l_{j2}}, \mathbf{l_{j3}}$, respectively. The total number of states is $I=\prod_{k=1}^{n} L_K$ and the number of transition probabilities $\mathbf{P_{ij}=P(l_{j1}, l_{j2}, l_{j3} \mid l_{i1}, l_{i2}, l_{i3})}$ is equal to $\mathbf{I}^2$. Under the assumptions of independence of the attributes at an individual level for the respondents, P_{ij} is given by

$$\mathbf{P_{ij} = \prod_{k=1}^{n} P(l_{jK} \mid l_{i1}, l_{i2}, \ldots, l_{in})} \tag{2}$$

If transition probability matrix is unknown and data are available we use, p (the maximumlikelihood estimator of

$$\mathbf{P_{ij}) : p = N_{ij} / N_i} \tag{3}$$

Where $\mathbf{N_i}$is the number of times that the starting state "i " has occurred when the process has been observed, $\mathbf{N_{ij}}$ is the number of times that the process has been observed to go from the starting state "i" to the destination state " j " in one step.

When data are not available we need information gathered from experts.

Our initial purpose of estimating P_{ij} may be seen as the purpose of estimating $\mathbf{P(l_{jK} \mid l_{i1}, l_{i2}, \ldots, l_{in})}$.

We build 2 (FFD)s on the starting $\mathbf{I = \prod_{k=1}^{n} L_K}$ and the destination states . By using this solution we can start from **I** instead of $\mathbf{I}^2$. In our method the judges are asked to assign transition probabilities to the destination states in the (FFD's) for each state included in the (FFD) of starting states. The experts will be asked about a reduced number of starting states (the (FFD) ones) and under the assumptions of independence of the attributes at an individual level for the respondents the results will be generalized to the others non included in the (FFD) by means of Logistic Regression (LR). The (LR) model examines the relationships between the independent variables and the log-odds of the outcome variable;

$$\left(\mathbf{odds} = \frac{\mathbf{p}}{\mathbf{1-p}}\right). \tag{4}$$

The model on log-odds (Logit) scale is linear

$$\left(\mathbf{Logit} = \mathbf{log}\left(\frac{\mathbf{p}}{\mathbf{1-p}}\right) = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}\right) \tag{5}$$

As above mentioned we want to estimate $P(l_{j\hat{e}} \mid l_{i1}, l_{i2}, \dots, l_{in})$ by means of logistic regression. The estimated parameters are used to reconstruct the probability of arriving in l_{jK} also for starting states not included in the (FFD). Finally the probabilities of destination states are given by:

$$\mathbf{P_{ij}} = \prod_{k=1}^{n} \mathbf{P(l_{jK} \mid l_{i1}, l_{i2}, \dots, l_{in})}.$$

Application on Simulation Data: We propose an application on simulation data to illustrate our methodology. We suppose n=4 (The number of attributes) each with 3 levels and the number of experts is 8 in 2 groups (in each group 4 experts).We used SPSS (ver:16) for generating 2 resolution III designs (FFDs) for the starting and the destination states and Minitab (ver:15) for generating P_{ij}. All experts received the same (FFDs). We simulated transition probability matrix P **for each of respondents** as mentioned below:

In each of starting states in (FFD) we generated 4 independent bivariate normally distributed vectors: $\mathbf{z_1, z_2, z_3, z_4}$ from N ($\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$), μ=(0,0), Σ=[1,-.9; -.9,1]

Where:

$$\mathbf{z_j = (log\ (L_{jk} / 1 - L_{jk}))\ j=1,2,3,4,\ k=1,2,} \tag{6}$$

$$\mathbf{L_{jk} = P(l_{j\hat{e}} \mid l_{i1}, l_{i2}, l_{i3}, l_{i4})} \tag{7}$$

We computed $\mathbf{L_{jk}}$ based on $\mathbf{z_j}$. Under the assumptions of independence of the attributes at an individual level for the respondents we computed $P_{ij} = \prod_{k=1}^{n} P(l_{jK} \mid l_{i1}, l_{i2}, \dots, l_{in})$, n=4. We repeated this computation for 500 times and finally mean of them recorded as $P_{ij} = P(\mathbf{l_{j1}, l_{j2}, l_{j3}, l_{j4} \mid l_{i1}, l_{i2}, l_{i3}, l_{i4}})$. Then we updated $\mathbf{L_{jk}}$ based on P_{ij}. In the Table 1 we showed the partially of these computations for $\mathbf{L_{12}}$ for the first respondent. Finally we concerned eq. **(5)** to $\mathbf{z_j}$, **j=1,2,3,4**.

The Multivariate Analysis of Variance (MANOVA) is used for estimating of the parameters [17]. One of the levels for each factor, regarded as dummy variables, is eliminated. The estimated parameters are used to reconstruct the probability of arriving in $\mathbf{l_{jK}}$ also for starting states not included in the (FFD).

The Generalized Least Squares (GLS) Regression estimates of the model:

$$\mathbf{log(L_{jk} / 1 - L_{jk}) = \beta_0 + \beta_1 l_{11} + \beta_2 l_{12} + \beta_3 l_{21} + \beta_4 l_{22} + \beta_5 l_{31} + \beta_6 l_{32} + \beta_7 l_{41} + \beta_8 l_{42} + e,\ j=1,2,3,4, k=1,2} \tag{8}$$

The (LR) coefficients (significant at.05 level) for the first 4 respondents are given in the following:

$$\mathbf{log(L_{12} / 1 - L_{12}) = .465\ l_{21} + .518\ l_{22} + e}$$
$$\mathbf{(.183)\ (.183),} \tag{9}$$

$$\mathbf{log(L_{22} / 1 - L_{22}) = .497\ l_{22} + e}$$
$$\mathbf{(.161),} \tag{10}$$

$$\mathbf{log(L_{11} / 1 - L_{11}) = -1.244 - .409\ l_{22} + e}$$
$$\mathbf{(.169),} \tag{11}$$

$$\mathbf{log(L_{42} / 1 - L_{42}) = -.908 - .314 l_{31} + e}$$
$$\mathbf{(.139),} \tag{12}$$

$$\mathbf{log(L_{31} / 1 - L_{31}) = .273\ l_{42} + e}$$
$$\mathbf{(.132),} \tag{13}$$

Where the estimated coefficient standard errors are in parentheses.

Also

$$\mathbf{log(L_{21} / 1 - L_{21}) = -1.207} \tag{14}$$

$$\mathbf{log(L_{41} / 1 - L_{41}) = 0} \tag{15}$$

$$\mathbf{log(L_{32} / 1 - L_{32}) = -1.208} \tag{16}$$

The (LR) coefficients (significant at.05 level) for the seconds 4 respondents are given in the following:

Table 1: The FFDs(for starting and destination states) and disclosed transition probabilities(P_{ij}) for the first respondent

	Destination											
Starting	3231	3312	2132	2321	2213	1333	1111	3123	1222	L_{12}	Odds = $L_{12}/(1-L_{12})$	Log(Odds)
2231	.20	.12	.11	.08	.11	.21	.06	.01	.10	.31	.44	-.82
3213	.06	.12	.08	.11	.19	.20	.15	.08	.02	.37	.60	-.51
3132	.11	.18	.03	.22	.18	.19	.04	.01	.05	.43	.75	-.28
1111	.15	.14	.10	.16	.04	.14	.10	.10	.07	.30	.42	-.86
1333	.05	.07	.18	.15	.15	.16	.09	.07	.09	.48	.91	-.10
3321	.11	.15	.12	.15	.09	.18	.06	.10	.03	.36	.57	-.56
2123	.06	.14	.14	.14	.15	.00	.10	.13	.14	.42	.74	-.30
2312	.08	.18	.14	.13	.03	.05	.17	.14	.07	.30	.43	-.84
1222	.13	.17	.16	.11	.16	.06	.09	.01	.10	.43	.76	-.27

Note:2231 in FFD for **starting states** means that i= $l_{12}, l_{22}, l_{33}, l_{41}$ and 3231 in FFD for **destination states** means that j= $l_{13}, l_{22}, l_{33}, l_{41}$, P_{ij}=0.20 (obtained from simulation),L_{12} = P(2131 | 2231) + P(2321 | 2231) + P(2213 | 2231) etc.

$$\log(L_{21} / 1- L_{21}) = -1.409-.54 l_{31} + e \qquad (17)$$

$$\log(L_{32} / 1- L_{32}) = -.748 \qquad (18)$$

$$\log(L_{41} / 1- L_{41}) = -1.138 \qquad (19)$$

$$\log(L_{22} / 1- L_{22}) = \log (L_{31} / 1- L_{31})= \log (L_{11} / 1- L_{11})= \log(L_{12} / 1- L_{12}) = \log(L_{42} / 1- L_{42}) = 0 \qquad (20)$$

CONCLUSION

Strengths of the Proposed Methodology Are: Use of expert opinion in situations of data absence; overcome the curse of dimensionality by use of (FFD); reconstruct the transition probabilities not included in the orthogonal design definition; use (CA) in estimating of the entropy rate an irreducible, discrete time and homogeneous Markov chain with a finite state space.

Weaknesses of the Proposed Methodology Are: The logistic regression techniques encounter problems when the number of the attributes or the number of level for each attribute grow, it makes the problem using conjoint data more difficult especially when higher resolution designs were applied that required larger number of profiles.

REFERENCES

1. Leone, D. and M. Fucili, 2008. Estimation of markov chains transition probabilities by means of conjoint analysis approach. working paper.
2. John R. Hauser and Vithala R. Rao, 2002. Conjoint Analysis, Related Modeling, and Applications, Chapter prepared for Advanced in Marketing Research: Progress and Prospects September 23.
3. Hiromi Yamada, *et al.* 2004. The Entropy Model Using Attribute of Conjoint Analysis, Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference.
4. Roselinde, *et al.* 2008. Optimal designs for conjoint experiments, Computational Statistics & Data Analysis, 52(5): 2369-2387.
5. Vithala R. Rao, 2007. Developments in Conjoint Analysis.Cornell University Revised.
6. Warren F. Kuhfeld. Conjoint Analysis. Copies of this chapter (TS-722H) and all of the macros are available on the web http://support.sas.com.
7. Jeng-Jong Lin, 2008. An Optimal Design Search with Conjoint Analysis Using Genetic Algorithm. Tamkang J. Science and Engineering, 11(1): 73-84.
8. Warren F. Kuhfeld, Randall D. Tobias and Mark Garratt, 1994. Efficient Experimental Design with Marketing Research Applications. J. Marketing Res., pp: 545-557.
9. George Van Houtven, *et al.* 2011. Eliciting Benefit–Risk Preferences and Probability-Weighted Utility Using Choice-Format Conjoint Analysis, Med Decis Making., 31(3): 469-480.
10. Jeremy J. Michalek, *et al.* 2010. Enhancing Marketing with Engineering: Optimal Product Line Design for Heterogeneous Markets. Forthcoming in International J. Research in Marketing.
11. Hiroyuki Ikemoto and Toshiki Yamaoka, 2011. Conjoint Analysis Method That Minimizes the Number of Profile Cards, 173(1): 23-28.

12. Byungun Yoon and Yongtae Park, 2007. Development of New Technology Forecasting Algorithm: Hybrid Approach for Morphology Analysis and Conjoint Analysis of Patent Information, Ieee Transactions on Engineering Management, 54(3).
13. Fumika Ouchi, 2004. A Literature Review on the Use of Expert opinion in Probabilistic Risk Analysis. World Bank Policy Research Working Paper, pp: 3201.
14. John Rust, 2006. Dynamic Programmingentry for consideration by the New algrave Dictionary of Economics. University of Maryland.
15. Papangelou, F., 1996. Large deviations and the Bayesian estimation of higher order Markov transition functions, J. Appl. Prob., 33: 18-27.
16. Baris Tan and Kamil Yilmaz, 2002. Markov Chain Test for Time Dependence and Homogeneity: An Analytical and Empirical Evaluation. European J. Operational Res., Elsevier., 137(3): 524-543.
17. James Press. S., 1971. Applied Multivariate Analysis. New York: Holt, Rinehart.