Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Statistical Planning and Inference 136 (2006) 942–961

journal of statistical planning and inference

www.elsevier.com/locate/jspi

A Dirichlet random coefficient regression model for quality indicators

Daniel Peña[a,*], Victor Yohai[b]

[a]Departments of Statistics, Universidad Carlos III de Madrid, Madrid, 126, Getafe (Madrid) 28903, Spain
[b]Universidad de Buenos Aires, Argentina

Received 6 November 2003; accepted 31 July 2004
Available online 28 September 2004

Abstract

We present a random coefficient regression model in which a response is linearly related to some explanatory variables with random coefficients following a Dirichlet distribution. These coefficients can be interpreted as weights because they are nonnegative and add up to one. The proposed estimation procedure combines iteratively reweighted least squares and the maximization on an approximated likelihood function. We also present a diagnostic tool based on a residual Q–Q plot and two procedures for estimating individual weights. The model is used to construct an index for measuring the quality of the railroad system in Spain.

MSC: 62J99; 62P25

Keywords: Random weights; Dirichlet distribution; Iterative least squares; Monte Carlo methods; Q–Q plot

1. Introduction

It is generally accepted that the quality of a service is usually a function of several quality factors, dimensions or attributes, (Parasuraman et al., 1988, 1991, 1994; Cronin and Taylor, 1992, 1994; Teas, 1993, 1994) and a key step in measuring service quality is determining the relative weight of each factor or attribute in overall satisfaction. Methods oriented

* Corresponding author. Tel.: +34 91 624 98 06; fax: +34 91 624 98 49.
E-mail address: dpena@est-econ.uc3m.es (D. Peña).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2004.07.012

to multidimensional quality measurements are usually based on Conjoint Analysis (Luce and Tukey, 1964). See Carroll and Green (1995) for a survey of the present state of this methodology and Lynch et al. (1994), Wedel and DeSarbo (1994) and Ostrom and Iacobucci (1995) for interesting applications to the evaluation of service quality. In this methodology customers are asked to provide quality evaluation on several hypothetical services defined by certain levels of the quality attributes. The method assumes that the quality attributes can be given an objective interpretation so that the levels of the attributes have, when presented to the customers for evaluation, a clear meaning to them.

Conjoint Analysis is less useful in situations in which the quality attributes do not have objective standards, and therefore it is very difficult to define a series of hypothetical quality situations for the customers to evaluate. An alternative procedure in these situations is to use hierarchical Bayesian methods that can be estimated by Markov Chain Monte Carlo method (MCMC), see Lenk et al. (1996), Allenby and Rossi (1999) and Rossi et al. (2001). A second alternative is to relate the evaluation of the attributes to the overall evaluation of service quality by using a random coefficients regression model. Peña (1997) proposed a model in which the weights of each customer are assumed to be random variables generated by a common multivariate normal distribution and show how to compute by generalized least squares (LS) the mean weights in the population imposing the restrictions that the weights must add up to one. This model was designed for the estimation of the mean weights in the population and the important problem of estimating the individual weights for each person, that is easily carried out in the hierarchical Bayesian approach, was not considered.

In this article we propose a random coefficient model in which the individual weights can be estimated. See Gumpertz and Pantula (1998) for a review of these models and their applications and Mallet (1986) for a non parametric approach to estimate the distribution of the coefficients. The model we propose in this article includes two features that generalize previous applications of random coefficient models for building quality indices. First, it incorporates the restriction that the weights must be positive and thus it avoids the problem of estimating some negative weights. Second, it allows that customers, in their evaluations of the overall quality, may be taking into account some attributes not considered in the model. This feature is formally incorporated in our model, and the distribution of the values of the unknown attribute can be obtained.

The rest of this article is organized as follows. In Section 2 the model is presented. In Section 3 its estimation is discussed and in Section 4 its validation is presented. In Section 5 two procedures for obtaining estimates of individual weights are presented. In Section 6 an application to evaluate the quality of the railroad system in Spain is discussed. In Section 7 some final remarks are included.

2. A model for linear quality indicators

Suppose that we have a population of potential customers. We assume that each customer has an evaluation score y of the perceived quality of a given service that is a weighted linear combination of several known attributes, factors or dimensions, $x_1, \ldots, x_k$ and, possibly, of a latent variable z depending on other unidentified factors. Thus the evaluation score is computed by the customer by giving weights to the different dimensions or attributes