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*

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.
© 2004 Elsevier B.V. All rights reserved.

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).

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

944   *D. Peña, V. Yohai / Journal of Statistical Planning and Inference 136 (2006) 942–961*

considered and the evaluation score reported includes some random measurement error which includes the rounding error and other computation errors made by the customer. Without loss of generality we assume that the data has been scaled so that the variables $y, x_1, \ldots, x_k$ and z are scores between 0 and 1. Suppose that a random sample of n customers has been surveyed, and let $(y_i, \mathbf{x}_i)$, where $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,k})'$, be the answer of customer i. We assume that

$$y_i = w_{i,1}x_{i,1} + \cdots + w_{i,k}x_{i,k} + w_{i,k+1}z_i + \varepsilon_i, \quad 1 \leqslant i \leqslant n, \tag{1}$$

where z_i is the unobserved random variable corresponding to the evaluation of the unspecified factors for customer i, $\mathbf{w}_i = (w_{i,1}, \ldots, w_{i,k+1})'$ is a random vector of weights measuring the relative importance that customer i gives to the different attributes x_j, $1 \leqslant i \leqslant k$, and to z in determining overall service quality y and ε_i is a measurement error. The variables $\mathbf{w}_i$, z_i and ε_i are not observed.

The error ε_i takes into account differences between the theoretical and the observed overall quality due to particular behavior of some of the respondents. We assume that the attribute evaluations are made without measurement error. In practice there will always be some measurement error which can be different for different attributes. However, we assume this hypothesis for simplicity, and in Section 7 we will comment on the implications of deleting it.

We make the following assumptions:

A1. The random variables $\mathbf{x}_i$, z_i, $\mathbf{w}_i$ and ε_i are independent. The justification that $\mathbf{x}_i$ and $\mathbf{w}_i$ are independent is that the evaluation of an attribute represents how the level of service in this attribute compares to an ideal or standard performance, whereas the weights represent the a prior wishes of the customer. The independence between $\mathbf{x}_i$, the evaluation of the known attributes and z_i the evaluation of the unknown attribute is made for simplicity and can be easily generalized by assuming for instance that $E(z_i|\mathbf{x}_i)$ is equal to the mean evaluation of the known attributes. In Section 7 we will comment on this possible generalization.

A2. The distribution of $\mathbf{w}_i$ is Dirichlet with parameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{k+1})'$, $(D(\boldsymbol{\alpha}))$, the distribution of z_i is beta with parameter $\mathbf{p} = (p_1, p_2)'$, $(B(\mathbf{p}))$, and the distribution of ε is Normal with mean 0 and variance σ^2, $(N(0, \sigma^2))$. Observe that the Dirichlet assumption for the weights is in agreement with the basic assumption of a linear quality indicator, that is, that $w_{i,j} \geqslant 0$ and that $\sum_{j=1}^{k+1} w_{i,j} = 1$, and therefore, according to (1), the score y_i is a weighted average of the scores $x_{i,j}$, $1 \leqslant j \leqslant k$, and z_i plus a measurement error. The Beta assumption for the distribution of z_i is in agreement with the values of this variable in the interval 0–1 and allows a reasonable flexibility in the form of the distribution. The Normal distribution for the noise is made for simplicity as a priori the value of σ^2 is expected to be small and therefore the values of the noise are not expected to move the evaluation score y out of the interval 0,1. In Section 7 we will comment on alternative ways to model the noise in this model.

This service quality index has the following advantages: (1) knowledge of the attribute weights allows the ordering of the attributes according to their relative importance to the customer, showing the key factors for improving quality. (2) Customer weights can be related to customer characteristics to make market segmentation directly linked to quality objectives. (3) Comparing the average value of the attributes in the service studied to the values of these attributes for other services will reveal their relative strengths and

weaknesses. (4) If the attributes can be related to some objective measures of performance it is possible to substitute the subjective evaluations of the attributes for objective measurements, allowing a simple monitoring of the quality index.

Let $\alpha_T = \sum_{j=1}^{k+1} \alpha_i$, $p_T = p_1 + p_2$, $\beta_j = \alpha_j/\alpha_T$, $i = 1, \ldots, k+1$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{k+1})'$. Then, by A2, we have

$$E(w_{i,j}) = \frac{\alpha_j}{\alpha_T} = \beta_j, \quad \text{var}(w_{i,j}) = \frac{\beta_j(1 - \beta_j)}{1 + \alpha_T}, \quad i, j = 1, \ldots, k+1, \tag{2}$$

$$\text{cov}(w_{i,j} w_{i,k}) = \frac{-\beta_j \beta_k}{1 + \alpha_T}, \quad j \neq k \tag{3}$$

and

$$E(z_i) = \frac{p_1}{p_T} = m, \quad \text{var}(z_i) = \frac{m(1 - m)}{1 + p_T}. \tag{4}$$

If there are many attributes, unless there is an exceptional one, it is very likely that all the β_i's are smaller than 0.5. In this case the variability implied by the Dirichlet distribution is such that important attributes with large mean weights will have larger variance than marginal attributes with small mean weight. Also the correlation between important attributes will be large, whereas marginal attributes will be almost uncorrelated. This agrees with a priori expectations.

Then, using A.1, (1), (2) and (4) we have

$$E(y_i|\mathbf{x}_i) = \gamma + \sum_{j=1}^{k} x_{i,j}\beta_j, \tag{5}$$

where

$$\gamma = m\beta_{k+1}. \tag{6}$$

In the Appendix we derive the conditional variance of y_i, which is given by

$$\text{var}(y_i|\mathbf{x}_i) = \mu_2 s_i^2 + \overline{x}_i \mu_1 + \mu_0, \tag{7}$$

where

$$s_i^2 = \sum_{j=1}^{k} \beta_j x_{i,j}^2 - \overline{x}_i^2 \tag{8}$$

is a pseudo weighted variance (note that $\sum_{j=1}^{k} \beta_j \leqslant 1$) of the attribute evaluations, and

$$\overline{x}_i = \sum_{j=1}^{k} x_{i,j}\beta_j \tag{9}$$

is a pseudo weighted mean,

$$\mu_0 = \text{var}(w_{i,k+1}z) + \sigma^2, \tag{10}$$

$$\mu_1 = -\frac{2\gamma}{(1 + \alpha_T)}, \tag{11}$$

and

$$\mu_2 = \frac{1}{(1 + \alpha_T)}. \tag{12}$$

Moreover, we show (see (32) in the Appendix) that

$$
\begin{aligned}
\text{var}(w_{i,k+1}z) &= \chi(\alpha_T, p_T, \boldsymbol{\beta}, m) \\
&= \frac{\gamma(1 - m)(m - \gamma)}{m(1 + \alpha_T)(1 + p_T)} \left(1 + \frac{\gamma(1 + \alpha_T)}{m - \gamma} + \frac{m(1 + p_T)}{1 - m}\right).
\end{aligned} \tag{13}
$$

Let τ be defined by

$$\tau = \frac{1}{n} \sum_{i=1}^{n} E\left(\left(\left(y_i - \gamma - \sum_{j=1}^{k} \beta_j x_{i,j}\right)^2 \bigg| \mathbf{x}_i\right)\right).$$

From (7) and (13) we get that

$$
\begin{aligned}
\tau &= \phi(\alpha_T, p_T, \sigma^2, \boldsymbol{\beta}, m, \gamma) \\
&= \sigma^2 + \chi(\alpha_T, p_T, \boldsymbol{\beta}, m, \gamma) + \frac{1}{n(1 + \alpha_T)} \sum_{i=1}^{n} (s_i^2 - 2\gamma \bar{x}_i).
\end{aligned} \tag{14}
$$

This expression will be used in the next section for estimating the parameters by maximum likelihood. Note that the estimation of τ and σ^2 provides a method to split the total variability of the residuals in the regression (5) between a term due to the measurement error, σ^2, and a term $\tau - \sigma^2$ due to the weights variability between customers and to the unknown attribute z.

3. Model estimation

This model is estimated in three stages:

1. Consistent estimates $\widehat{\boldsymbol{\beta}}$, $\widehat{m}$, $\widehat{\gamma}$, and $\widehat{\tau}$ of $\boldsymbol{\beta}$, m, γ, and τ are computed by iterative reweighted least squares (IRLS).
2. Let $\widehat{\boldsymbol{\beta}}$, $\widehat{m}$, $\widehat{\gamma}$ and $\widehat{\tau}$ be the estimates obtained in stage 1. In stage 2 the initial estimates of $\mathbf{p}$, $\boldsymbol{\alpha}$ and σ^2 are obtained by maximizing an approximate likelihood of $\mathbf{y}$ subject to the constraints

$$\boldsymbol{\alpha} = \alpha_T \widehat{\boldsymbol{\beta}}, \tag{15}$$

$$\mathbf{p} = p_T(\widehat{m}, 1 - \widehat{m})' \tag{16}$$

and

$$\phi(\alpha_T, p_T, \sigma^2, \widehat{\boldsymbol{\beta}}, \widehat{m}, \widehat{\gamma}) = \widehat{\tau}, \tag{17}$$

where the function ϕ is defined in (14). Observe that in this stage we have to maximize the approximate likelihood with respect to α_T and p_T. The other parameters are obtained from the constraints.

3. Final estimates of $\mathbf{p}$, $\boldsymbol{\alpha}$ and σ^2 are computed by approximate maximum likelihood starting from the solution found in stage 2 and using only the constraints (15) and (16). Observe that in this stage we have to maximize the approximate likelihood with respect to α_T, p_T and σ^2. The other parameters are obtained from the constraints.

3.1. Estimation by iterative reweighted least squares

By (5), (7), (11) and (12) we can write

$$y_i = \gamma + \beta_1 x_{i,1} + \cdots + \beta_k x_{i,k} + u_i, \quad i = 1, \ldots, n, \tag{18}$$

where

$$E(u_i|\mathbf{x}_i) = 0 \tag{19}$$

and

$$\mathrm{var}(u_i|\mathbf{x}_i) = \mu_2 t_i + \mu_0, \tag{20}$$

where

$$t_i = s_i^2 - 2\gamma \bar{x}_i. \tag{21}$$

Based on (18) we start by fitting by ordinary LS (18) obtaining coefficients estimates $\gamma^{(0)}, \widehat{\beta}_1^{(0)}, \ldots, \widehat{\beta}_k^{(0)}$ and residuals $\widehat{u}_i^{(0)}$, $1 \leqslant i \leqslant n$. Then, using (20), we can write

$$u_i^2 = \mu_2 t_i + \mu_0 + u_i^*, \tag{22}$$

where $E(u_i^*) = 0$. By replacing the u_i^2's by the $\widehat{u}_i^{(0)2}$'s, we obtain estimates $\widehat{\mu}_0$ and $\widehat{\mu}_2$ by fitting by ordinary LS (22). The t_i's are computed by (8), (9) (21) using the estimates $\widehat{\gamma}^{(0)}$, $\widehat{\beta}_1^{(0)}, \ldots, \widehat{\beta}_k^{(0)}$ as parameter values.

We estimate $\mathrm{var}(u_i|\mathbf{x}_i)$ by $\widehat{v}_i^2 = \widehat{\mu}_2 t_i + \widehat{\mu}_0$, and then go back to regression (18), which is now fitted by weighted LS with weights $(\widehat{v}_i^2)^{-1}$. This procedure is iterated until convergence, obtaining estimates $\widehat{\gamma}$, $\widehat{\boldsymbol{\beta}}$, where $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_k)$.

Since the $\beta_i \geqslant 0$, the values $\widehat{\beta}_i < 0$ are replaced by 0. Observe that since the $\widehat{\beta}_i$'s are consistent, if the model is correct, for all i such that $\beta_i > 0$, asymptotically $\widehat{\beta}_i > 0$ and thus this correction will not be necessary for large n. Then, for large n, the only $\widehat{\beta}_i$ that may require correction are those corresponding to $\beta_i = 0$. In this case, when $\widehat{\beta}_i$ is corrected, we will have $\widehat{\beta}_i = \beta_i$. Another possibility to avoid this correction is to define the $\widehat{\beta}_i$ using weighted LS with the constrain that all $\widehat{\beta}_i \geqslant 0$. An algorithm for non negative LS was proposed by Lawson and Hanson (1974). This algorithm is available as the function NNLS of MATLAB.

Since the sum of all β_i is 1, in the case that

$$\xi = \sum_{i=1}^{k} \widehat{\beta}_i > 1,$$

all the components of $\widehat{\boldsymbol{\beta}}$ are divided by ξ. Then we compute $\widehat{\beta}_{k+1} = 1 - \sum_{j=1}^{k} \widehat{\beta}_j$, and according to (6), we estimate m by $\widehat{m} = \widehat{\gamma}/\widehat{\beta}_{k+1}$. Again, if $\sum_{i=1}^{k} \beta_i < 1$, this correction is not necessary for large n. We can also define the $\widehat{\beta}_i$ using weighted LS with the constraints that $\widehat{\beta}_i \geqslant 0$ and $\sum_{i=1}^{n} \widehat{\beta}_i \leqslant 1$.

Finally, we compute the residuals

$$\widehat{u}_i = y_i - \widehat{\gamma} - \widehat{\beta}_1 x_{i,1} + \cdots + \widehat{\beta}_k x_{i,k}, \quad i = 1, \ldots, n \tag{23}$$

and we estimate τ by

$$\widehat{\tau} = \frac{1}{n} \sum \widehat{u}_i^2. \tag{24}$$

Remark. More efficient estimates of μ_0 an μ_2 in the simple regression (22) would be obtained by using weighted LS with weights $1/\text{var}(u_i^*|\mathbf{x}_i)$, $1 \leqslant i \leqslant n$. However, these variances depend on $\mathbf{p}$, and this parameter cannot be estimated at this stage. Since the choice of the weights is not crucial for consistency, we use ordinary LS.

3.2. Maximum likelihood estimation

The likelihood function of $y_1, \ldots, y_n$ is

$$l(\boldsymbol{\alpha}, \mathbf{p}, \sigma^2) = \prod_{i=1}^{n} f_i(y_i, \boldsymbol{\alpha}, \mathbf{p}, \sigma^2), \tag{25}$$

where

$$\begin{aligned} f_i(y_i, \boldsymbol{\alpha}, \mathbf{p}, \sigma^2) &= \int f(y_i|\mathbf{w}^*, z, \sigma^2) f_D(\mathbf{w}^*, \boldsymbol{\alpha}) f_B(z, \mathbf{p}) \, d\mathbf{w}^* \, dz \\ &= \int \varphi((y_i - \mathbf{x}_i'\mathbf{w}^* - zw_{k+1})/\sigma) f_D(\mathbf{w}^*, \boldsymbol{\alpha}) f_B(z, \mathbf{p}) \, d\mathbf{w}^* \, dz \end{aligned} \tag{26}$$

and where φ, f_D and f_B are the standard Normal, Dirichlet and Beta densities, respectively, and $\mathbf{w}^* = (w_1, \ldots, w_k)'$. This integral can be approximated by

$$f_i^A(y_i, \boldsymbol{\alpha}, \mathbf{p}, \sigma^2) = \frac{1}{N} \sum_{j=1}^{N} \varphi((y_i - \mathbf{x}_i'\mathbf{w}_j^* - z_j w_{j,k+1})/\sigma), \tag{27}$$

where $\mathbf{w}_j = (\mathbf{w}_j^{*\prime}, w_{j,k+1})'$ and z_j, $j = 1, \ldots, N$ are two independent random samples of the Dirichlet and Beta distribution with parameters $\boldsymbol{\alpha}$ and $\mathbf{p}$, respectively. Then the likelihood

(25) can be approximated by

$$l^A(\boldsymbol{\alpha}, \mathbf{p}, \sigma^2) = \prod_{i=1}^{n} f_i^A(y_i, \boldsymbol{\alpha}, \mathbf{p}, \boldsymbol{\sigma}^2), \tag{28}$$

where the integrals (26) are substituted by the sums (27). Since obtaining the value of this approximate likelihood is computationally very intensive, we use this likelihood to estimate the parameters, α_T, p_T and σ^2 fixing γ, $\boldsymbol{\beta}$, and m to the values obtained in the first stage. This is done in stages 2 and 3, as mentioned above. In stage 2, we take a grid of values for α_T and p_T and for each choice of values for these two parameters the corresponding value of σ^2 is obtained by the constraint (17). To define the grid we take into account that, according to (12), an initial estimate of α_T is given by $\widehat{\alpha}_T = \widehat{\mu}_2^{-1} - 1$, where $\widehat{\mu}_2$ is obtained in stage 1. Then we find the combination of the two parameters in the grid which maximizes (28). In stage 3 we keep the values of γ, $\boldsymbol{\beta}$, and m fixed on the values of stage 1 and maximize the likelihood on the three parameters α_T, p_T and σ^2 using as initial values those obtained in stage 2.

4. Model validation

The model can be validated by comparing the observed residuals after fitting the model to the residuals computed with artificial samples generated by using the estimated parameters. If the model is appropriate the observed residuals should have a similar distribution as the simulated residuals. We start by computing the residuals $\widehat{u}_i$ by (23) and its empirical distribution function F_n.

We generate N artificial samples of the form $y_1^*, \ldots, y_n^*$, where y_i^* is generated as in (1) with $\boldsymbol{\alpha}$, $\mathbf{p}$ and σ^2 replaced by the estimated parameters. Let F_{hn}^* be the empirical distribution function of the errors

$$u_i^* = y_i - \sum_{j=1}^{k} \mathbf{x}_{i,j}\widehat{\beta}_j - \widehat{\gamma}, \quad 1 \leqslant i \leqslant n$$

corresponding to the hth artificial sample and define

$$F_n^* = \frac{1}{N} \sum_{h=1}^{N} F_{nh}^*.$$

Then the Q–Q plot between F_n and F_n^* will be a diagnostic tool for detecting discrepancies between the model and the data.

In particular the Q–Q plot may detect outliers corresponding to respondents with atypical views or recording errors. However some groups of outliers may go undetected because of a masking effect, although this effect is not expected to be large, because the data must be between zero and one. In any case, it is safer to check that a robust estimate for regression is similar to LS.

The model assumes that the overall evaluation is a continuous variable when in fact in most applications this variable is measured on a discrete scale. Suppose that the observed

evaluation is made on a discrete rating scale, for example 0–10. Then the response variable in the model is not observed exactly but rounded off to the closest integer. In order to check the effect of this discrete scale we can simulate the model and, using the estimated values of the parameters, generate two types of samples (1) samples of continuous y values, and (2) samples of discrete y values obtained by rounding off the continuous values. Then we can estimate the model in both samples and compare the results. The average discrepancy found in many replications of this analysis will provide an estimate of the expected bias due to the discrete scale effect. We will give more details when analyzing an example with real data in Section 6.

5. Predicting the weights for each observation

Predicting the weights for each respondent is important if we want to relate these weights to the personal characteristics of the respondents (such as gender, income, education and so on). In this section we present two possible approaches to this problem. However, as we want to estimate the vector of weights (a dimension $k + 1$ vector) with the information of just one dependent variable this estimation will necessarily have a large error as there is an infinite number of weight vectors which produce the observed value of the response given the explanatory variables.

5.1. Linear prediction

Assuming that the parameters are known, the weights of each customer in the sample can be estimated by computing the best linear predictor of the random variables $w_{i,1}, \ldots, w_{i,k+1}$ given the observation y_i. In practice, since the parameters are unknown, this predictor is computed by using the estimated parameters as true parameters.

The vector of the optimal (minimum mean squared error) linear predictor of the weights $\mathbf{w}_i$ for the ith individual is given by

$$\widehat{\mathbf{w}}_i = \boldsymbol{\beta} + \frac{1}{\mathrm{var}(y_i|\mathbf{x}_i)} \, \mathrm{cov}(\mathbf{w}_i, y_i) \left(y_i - \sum_{j=1}^{k} \mathbf{x}_{i,j}\beta_j - \gamma \right).$$

It is easy to show that

$$\mathrm{cov}(\mathbf{w}, y_i) = \Omega \mathbf{x}_i^*,$$

where Ω is the covariance matrix of $\mathbf{w}$ given in the Appendix and $\mathbf{x}_i^* = (\mathbf{x}_i', m)'$. Therefore

$$\widehat{\mathbf{w}}_i = \boldsymbol{\beta} + \frac{\Omega \mathbf{x}_i^* (y_i - \sum_{j=1}^{k} \mathbf{x}_{i,j}\beta_j - \gamma)}{\mathrm{var}(y_i|\mathbf{x}_i)},$$

By (7), (11) and (12), we can estimate $\mathrm{var}(y_i|\mathbf{x}_i)$ by $\widehat{v}_i^2 = \widehat{\mu}_2 t_i + \widehat{\mu}_0$. Then if the $\widehat{u}_i$'s are defined as in (23), the weights can be estimated by

$$\widehat{\mathbf{w}}_i = \widehat{\boldsymbol{\beta}} + \frac{\widehat{u}_i}{\widehat{v}_i^2} \widehat{\Omega} \mathbf{x}_i^*.$$

A shortcoming of this procedure is that the predictions of the weights can be negative. In the next subsection we present an alternative nonlinear predictor giving nonnegative weights.

5.2. Nonparametric prediction

The vector of weights of the ith element of the sample $\mathbf{w}_i = (w_{i,1}, \ldots, w_{i,k}, w_{i,k+1})$ can be predicted by estimating $E(\mathbf{w}_i | y_i, \mathbf{x}_i)$. To this end, we generate samples $(\widetilde{\mathbf{w}}_{i1}, \ldots, \widetilde{\mathbf{w}}_{iN})$ and $(\widetilde{y}_{i1}, \ldots, \widetilde{y}_{iN})$ as follows. We generate three independent random samples $(\widetilde{z}_{i1}, \ldots, \widetilde{z}_{iN})$, $(\widetilde{\mathbf{w}}_{i1}, \ldots, \widetilde{\mathbf{w}}_{iN})$ and $(\widetilde{\varepsilon}_{i1}, \ldots, \widetilde{\varepsilon}_{iN})$ where each $\widetilde{z}_{i,j}$ is B($\widehat{\mathbf{p}}$), $\widetilde{\mathbf{w}}_{i,j} = (\widetilde{\mathbf{w}}_{i,j}^*, \widetilde{w}_{i,j,k+1})$ is D($\widehat{\boldsymbol{\alpha}}$) and $\widetilde{\varepsilon}_{ij}$ is N(0, σ^2). A sample of the distribution of y_i is generated as $\widetilde{y}_{ij} = \mathbf{x}_i' \widetilde{\mathbf{w}}_{ij}^* + \widetilde{w}_{ij,k+1} \widetilde{z}_j + \varepsilon_{ij}$, $1 \leqslant j \leqslant N$. Then the expectation of w_{ih} ($h = 1, \ldots, k$) given y_i can be computed by nonparametric regression between $(y_{i1}, \ldots, y_{iN})$ and $(w_{i1,h}, \ldots, w_{wN,h})$. For example, we can use a nearest neighborhood estimate. In this case, w_{ih} can be estimated by

$$\widehat{w}_{ih} = \frac{1}{\#A} \sum_{j \in A} \widetilde{w}_{ij,h},$$

where

$$A = \{j : |\widetilde{y}_{ij} - y_i| \leqslant a\},$$

and where the window size a should be conveniently chosen and $n_h = \#A$. An overview of nonparametric regression including the choice of the window size can be found for example in Härdle (1990).

5.3. Diagnostic checks

Predicted weights can be used as an additional diagnostic tool for checking the model: we can compare the distribution of the predicted weights of the observed data with those of artificial samples generated by using the proposed model with the estimated parameters.

Suppose that N samples are generated and let G_{hj}^* be the empirical distribution function of the weights for attribute j in the sample h and

$$G_j^* = \frac{1}{N} \sum_{h=1}^{N} G_{hj}^*.$$

We can use a Q–Q plot to compare G_j^* and the empirical distribution of the predicted weights for attribute j using the real data.

6. Measuring the quality of the Spanish railroad system

In this example we apply the previous model to build a quality index of the railroad system in Spain. The procedure used to build this index can be summarized as follows: (1) identifying the quality attributes, (2) taking a random sample of customers and obtaining

Table 1
Descriptive measure of the attributes and the overall quality evaluations for the Railroad data

Attribute	Mean	Median	St. Dev.
1. Information pre-journ.	5.1808	5	2.1884
2. Station security	4.3189	5	2.2868
3. Ticket sales	4.5330	5	2.3687
4. RR cleanness station	3.4803	4	2.3130
5. Information station	4.7298	5	2.0666
6. Train cleanness	4.6211	5	2.1876
7. Train security	4.7011	5	2.2462
8. Punctuality	3.7305	4	2.5145
9. Speed	3.8266	4	2.2576
10. Train cleanness	5.1568	5	2.0544
11. RR cleanness train	4.2548	5	2.2995
12. Comfort	4.9887	5	2.1017
13. Noise level	4.4777	5	2.1960
14. Information train	3.6631	4	2.3763
15. Person responsible	4.5991	5	2.3066
16. Claims information	4.0274	5	2.2033
17. Claims handling	3.5964	4	2.2095
Overall	5.2115	5	1.7313

their evaluations; (3) estimating the parameters of the model and predicting the customer weights.

Starting with the first step, the identification of the attributes was the result of several group sessions with customers of RENFE (the public railroad system in Spain). Initially 52 attributes were identified: 28 of these 52 attributes correspond to the pre-journey (information, ticket office, railroad station), 20 to the journey and 6 to the post-journey (claims and so on). A small random sample of 400 customers was taken to check these attributes and as a result of the statistical analysis of the questionnaire the number of attributes was reduced to 17: 6 for the pre-journey, 9 for the journey and 2 for the post-journey.

In order to obtain a representative sample of the railroad service, a stratified sample of 2000 people was selected including different types of train and days of the week. The interviews were held during the journey by train and the passengers were requested to evaluate on a 0–9 scale the observed quality of these attributes in RENFE and to evaluate the overall service quality. Of the 2000 questionnaires 51 have many missing values and were disregarded, so that the final sample size includes 1499 questionnaires. The data is available at http://halweb.uc3m.es/esp/personal/personas/dpena/articles/datarenfe JSPI2004.zip.

Table 1 lists the 17 attributes and presents some descriptive statistics of their distribution. The seventeen attributes are highly correlated, for instance, the first principal component of these attributes explains 93.4% of the variability. This principal component is a weighted average of all the attributes with similar weights, the smallest weight corresponds to attribute 3 (.0517) and the largest to attribute 13 (.0623). Table 1 shows that the mean value for overall quality is higher than the mean evaluation of all the attributes, suggesting that overall quality is not a weighted mean of the considered attributes and that some other unknown attributes are taken into consideration when evaluating the overall quality.
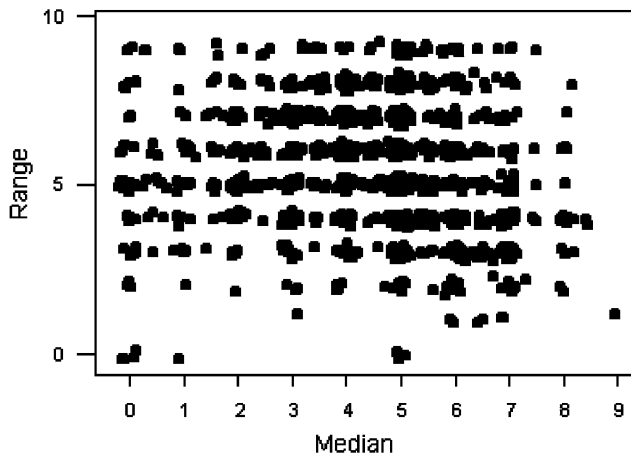
Fig. 1. Median and range of the evaluations of the attributes by each respondent for the RENFE data.

Rossi et al. (2001) have shown that when respondents vary in their use of the scale, for instance, some use only the middle of the scale or the upper or lower end, some biased are expected in the correlation inferences. This problem can appear in a plot of the range on the evaluation of the attributes for each customer with respect to the median of the attributes. For instance, if some respondents use only the top end of the scale they will produce points in the plot with large median and small range in the evaluation. Rossi et al. (2001) suggested a plot in which the points are slightly jittered so that the plot illustrates the number of respondents at any given combination of the variables. Fig. 1 shows this plot for our data. Some indication of scale usage heterogeneity is found in respondents with median response 9, which have a very small range in their evaluations, but this effect seems to be small since for the other median values the range has a similar variability.

The IRLS estimates was applied to the full model with the 17 attributes. Table 2 gives the results of applying the IRLS estimate to a model using only those attributes found significative at level 0.05 in the full model. The LSE column gives the coefficients estimated by ordinary LS and the IRLS column the ones obtained at the end of stage 1 as described in Section 3.1. When estimating the model all the evaluations have been divided by 10 to transform them to a 0–1 scale as assumed for our model, but we present the estimation results for the original data. The high value for the intercept is suggesting that some attributes considered by the customers are not taken into account in the questionnaire. Also, the sum of the regression coefficients is .8067 indicating that the overall quality is not a weighted average of the observed attributes. This is confirmed by the IRLS estimate. The IRLS stage leads to $\widehat{m} = 8.492$, $\widehat{\gamma} = 1.72$, $\widehat{\tau} = .136$ and $\widehat{\alpha}_T = 4.84$. The ML estimates obtained at the end of stage 3 are $\widehat{\alpha}_T = 5.9$, $\widehat{p}_T = 1.35$ and $\sigma = .52$. We conclude that an attribute with weight $\widehat{\beta}_{k+1} = \widehat{\gamma}/\widehat{m} = .2026$ and mean evaluation in the sample $\widehat{m} = 8.492$ is missing. Regarding the variability, the total residual variance, $\widehat{\tau} = 1.356$, can be split into the variance due to measurement error, $\sigma^2 = .274$, and the variance due to weights variability including the one due to the missing attribute, $1.356 - .274 = 1.082$. Thus the weight variability represents 79.7% of the observed variability.

Table 2
Regression coefficients of the attributes to explain the overall quality evaluations for the Railroad data

	Attribute		
	LSR	IRLS	*t*-values
Intercept	1.700	1.72	19.36
1. Information pre-journey	0.08422	0.0898	5.157
3. Ticket sales	0.05241	0.0527	3.47
6. Train cleanness	0.05852	0.0517	3.00
8. Punctuality	0.05273	0.0536	3.30
9. Speed	0.12650	0.1158	5.92
11. RR cleanness train	0.02534	0.0289	1.69
12. Comfort	0.09741	0.1032	4.59
13. Noise level	0.10720	0.1015	4.7
15. Person responsible	0.05623	0.0541	2.96
16. Claims information	0.06513	0.0661	2.89
17. Claims handling	0.08111	0.0798	3.58
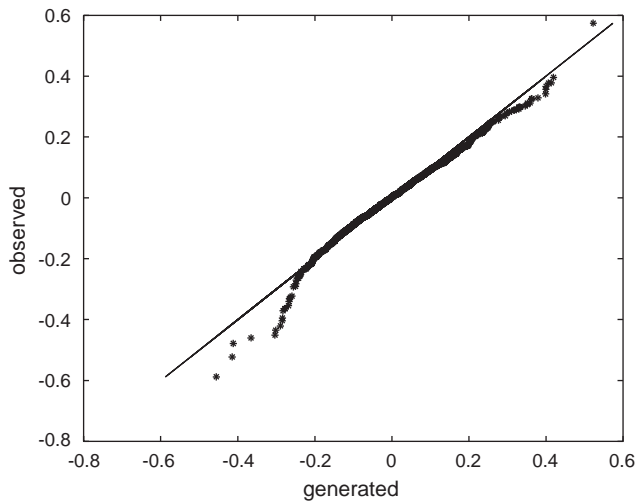
Fig. 2. Q–Q plot of the empirical distribution function of the observed residuals versus the theoretical distribution.

For model validation we proceeded as described in Section 4 and the Q–Q plot is shown in Fig. 2. It can be seen that agreement is very good apart from a few outliers in the tails. There are 19 residuals with absolutes values larger than three standard deviations and dropping these observations (1,2% of the sample data) we obtain the Q–Q plot indicated in Fig. 3. It can be seen that the agreement is now better.

In order to check the effect of the discreteness of the response variable we generated 500 samples of *y* using the estimated parameters as a true model and with the same *x* variables.
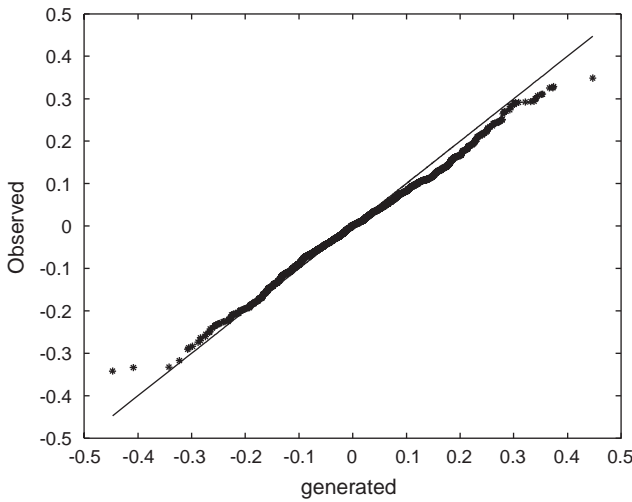
Fig. 3. Q–Q plot of the empirical distribution function of the observed residuals versus the theoretical distribution after deleting 19 outliers.

Table 3
Root mean squared error and bias of the unrounded (RMSEU, BIASU) and rounded estimators (RMSEU, BIASR) for the Railroad data and root mean squared error of the two methods for predicting the weights

	Attribute			
	RMSEU	RMSER	BIASU	BIASR
Intercept	0.0095	0.0101	0.0004	0.0006
1. Information pre-journey	0.0160	0.0166	−0.0012	−0.0010
3. Ticket sales	0.0148	0.0152	0.0004	0.0003
6. Train cleanness	0.0168	0.0173	0.0025	0.0025
8. Punctuality	0.0149	0.0156	−0.0007	−0.0006
9. Speed	0.0190	0.0197	0.0002	−0.0002
11. RR cleanness train	0.0162	0.0169	0.0009	0.0009
12. Comfort	0.0205	0.0207	0.0006	0.0007
13. Noise level	0.0201	0.0206	−0.0009	−0.0012
15. Person responsible	0.0170	0.0174	−0.0019	−0.0019
16. Claims information	0.0218	0.0224	−0.0005	−0.0002
17. Claims handling	0.0214	0.0219	0.0004	0.0003

We estimate the parameters β and γ in these samples first, with the generated y and second, by rounding off the values of y to the closest integer. The first two columns in Table 3 show the root mean squared error (RMSE) of estimation computed as

$$\mathrm{RMSE}(\beta_i) = \left[\frac{1}{N} \sum (\widehat{\beta}_i - \beta_i)^2 \right]^{1/2},$$
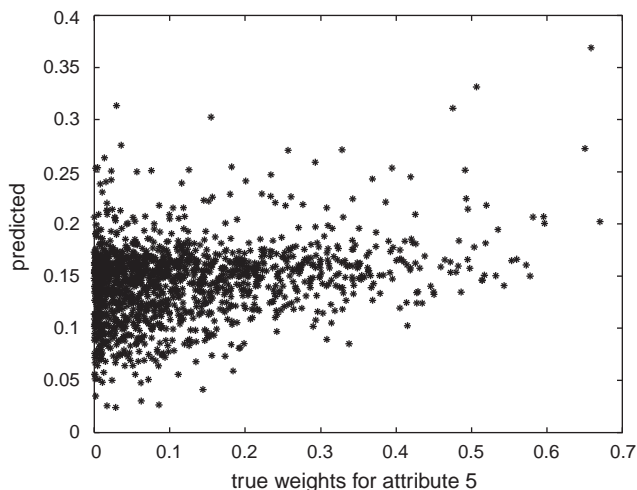
Fig. 4. Predicted and true weights for attribute speed in a generated sample.

The first column (RMSEU) corresponds to the exact one and the second (RMSER) to the rounding one. The next two columns correspond to the bias of the estimate computed as

$$\text{BIAS}(\beta_i) = \frac{1}{N} \sum (\widehat{\beta}_i - \beta_i),$$

for both estimates, the exact one (BIASU) and the rounding one (BIASR). The table shows that the effect of rounding on the estimation of the parameters is small. As expected there is a small increase in the RMSE when observations are rounded off but the average increase in RMSE of the rounding is just around 3%. The effect on the bias is negligible.

We have also compared the RMSE of the estimates of the weights for the two procedures presented in Section 4 applied to 500 generated samples. We have found that although as an average the nonparametric estimates have smaller error, the differences are small. Fig. 4 shows the true weights in one of these generated samples from the model and the nonparametric predicted weights for the attribute speed. It can be seen that estimation is not very good, especially for small values of the weights. This result is not surprising since we are estimating all the components of the vector of weights only with the information of the response at this point.

Fig. 5 presents the histogram of the nonparametric estimation of the weights for the most important attribute, speed in the railroad data. Note that this distribution is quite different from the one implied by the model which is a Beta (0.68, 5.22). This discrepancy does not imply that the model is wrong because when generating samples from the model and estimating the weights using these samples similar distributions were obtained. For instance, Fig. 6 presents the Q–Q plot of this distribution with respect to the nonparametric predicted weights in a generated artificial sample. This plot shows a high degree of agreement
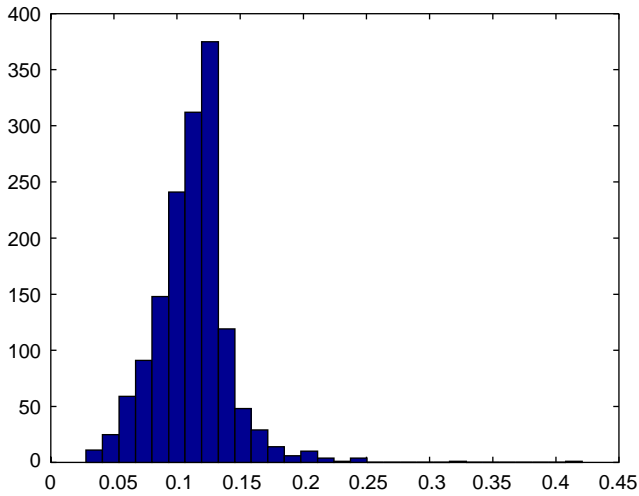
Fig. 5. Histogram of the distribution of weights for attribute speed estimated by nonparametric regression.
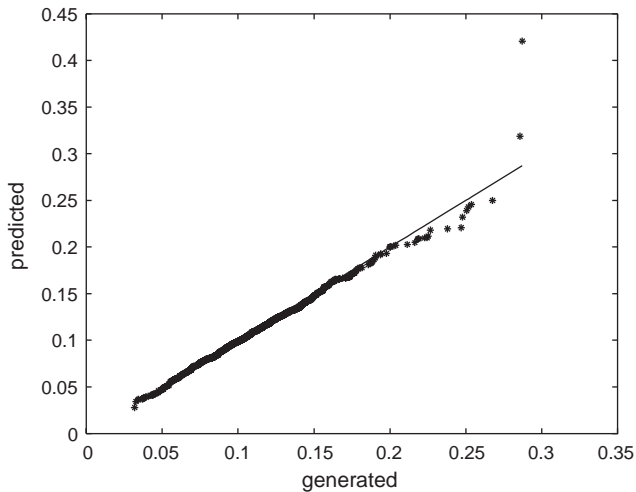

Fig. 6. Q–Q plot of the distributions of two nonparametric predicted weights: for the real sample and for a generated artificial sample. The selected attribute is speed.

between the two distributions except for a few outliers. Since the sample size is large, all the generated samples produce almost the same distribution. The same pattern was obtained with the other attributes. Thus we consider that the discrepancy between the theoretical and predicted distribution of the weights may be due to the limited information available for the prediction.

7. Discussion and conclusion

An alternative methodology that have been used for building quality indexes is by using linear structural relation models (LISREL). In this approach the unobserved latent variable quality, η, is related to a vector of p unobserved latent factors, ξ, by

$$\eta = \phi' \xi, \tag{29}$$

In order to estimate this model we have an observed variable y which is related to the latent variable quality by

$$y = \eta + \varepsilon, \tag{30}$$

where ε is a $N(0,\sigma^2)$ variable. We also have a set of $m > p$ observed $\mathbf{x}$ variables, which are related to the p factors ξ by the linear factor model equation

$$\mathbf{x} = \Lambda \xi + \upsilon, \tag{31}$$

where the vector υ has a $N_m(0, \Sigma)$ multivariate normal distribution. As the factors ξ will be estimated as linear function of the x variables, by using (29) and (30) we have that the relation between the observed variables is given by

$$y = \beta' \mathbf{x} + \varepsilon,$$

which is a linear regression model. From this point of view the model we are proposing can be seen as a reduced form of the structural model. However, the LISREL model usually assumes a fixed regression coefficient in the relation (29) among the latent variables, whereas our model allows for different weights among the customers, which we believe is a more realistic assumption. On the other hand, our model assumes that there is no measurement error in the explanatory variables. This possibility can be introduced into the model by using a equation similar to (31) with $\Lambda = \mathbf{I}$, the identity matrix, and assuming some error distribution for the measurement error and incorporating it into the model. Also, if a priori information on the mean of the attributes is available it can be included as prior information. Then the model can be set up in a Bayesian framework and estimated by Markov Chain Monte Carlo (MC^2) methods. Note that, as it has been used in the estimation of the model, the hierarchical structure of the model is well suited for Gibbs sampling estimation.

We have also assumed that the evaluation of the unknown attribute is independent of the evaluation of the known attributes. This assumption can be modified by assuming that $z_i | \mathbf{x}_i$ has a distribution with parameters which depend on $\mathbf{x}_i$. For instance, we may take $E(z_i|\mathbf{x}_i) = \mathbf{1}'\mathbf{x}_i/k$ and we can also relate $\text{var}(z_i|\mathbf{x}_i)$ to the observed variance among the components of $\mathbf{x}_i$. These assumptions by including additional information may make easier the estimation of the model but the problem is that they are hard to check with the observed data.

The assumption that the errors ε_i are normally distributed can be replaced by the more general assumption that they have a density of the form $\phi(u/\sigma)/\sigma$, where $\phi(u)$ is an arbitrary density with mean 0 and variance 1. For example ϕ may have compact support. In this case, the only difference in the estimation procedure would be to replace in (26) and (27) the normal density φ by ϕ. We can also consider different alternatives for ϕ and choose the one

giving the largest value of the likelihood function. Although these alternatives are worth exploring if we have evaluations close to the extremes of the scale, they are not expected to have a large effect on the conclusions of the model.

Another assumption we have made is that the observed variables can be approximated by continuous variables. An alternative approach would be to take into account that, in fact, they are measured as ordinal variables and to including this property into the model. For instance, Johnson (1996) has proposed to consider the evaluation as latent variables which are later discretised into the observed ordinal variables and use MC^2 to estimate the model. See Moreno and Rios-Insúa (1998) for an application of these ideas to Service Quality. This alternative will make the model more realistic, but also more complex and according to our experience will have a small effect in the conclusions.

A referee has suggested that in case of strong heterogeneity between subjects it would be useful to incorporate it into the main model. We agree with this suggestion and this can be done by assuming that the weights are generated by a mixture of Dirichlet distributions. That is, instead of assumption A2 we may assume that

$$w_i \backsim \sum_{j=1}^{J} \pi_j D(\boldsymbol{\alpha}_j)$$

where $\sum_{j=1}^{J} \pi_j = 1$ are the prior probabilities of the J populations of weights and $D(\boldsymbol{\alpha}_j)$ the Dirichlet distributions. The model can be estimated by the EM algorithm (see McLachlan and Krishnan, 1997) or by Markov Chain Monte Carlo methods.

In summary, knowledge of the relative importance of quality attributes for customers is crucial for any process of service quality improvement. The procedure presented in this paper seems to be a useful way to estimate the implicit weights used by the customers in their overall evaluation of service quality and to understand the relative importance of the different sources of variability in the overall evaluations. However, the model can still be improved in several ways and these extensions would be the subject of further research.

Acknowledgements

Appendix

From $y_i = \mathbf{x}_i' \mathbf{w}_i^* + w_{i,k+1} z_i + \varepsilon_i$, where $\mathbf{w}_i^* = (w_1, \ldots, w_k)'$, we have

$$\text{var}(y_i | \mathbf{x}_i') = \mathbf{x}_i' \Omega \mathbf{x}_i + \text{var}(w_{i,k+1} z) + 2 \, \text{cov}(\mathbf{x}_i' \mathbf{w}_i, w_{ik+1} z) + \sigma^2,$$

where Ω is the covariance matrix of $\mathbf{w}_i^*$. Denoting by $\boldsymbol{\beta}^* = (\beta_1, \ldots, \beta_k)'$, we have that $\Omega = (\text{diag}(\boldsymbol{\beta}^*) - \boldsymbol{\beta}^* \boldsymbol{\beta}^{*'})/(1 + \alpha_T)$, where $\text{diag}(\boldsymbol{\beta}^*)$ is a the $k \times k$ diagonal matrix with

diagonal terms the elements of $\boldsymbol{\beta}^*$. Then

$$\mathbf{x}_i'\Omega\mathbf{x}_i = \frac{\sum_{j=1}^k x_{i,j}^2\beta_j - (\sum_{j=1}^k x_{i,j}\beta_j)^2}{1+\alpha_T} = \frac{\sum \beta_j x_{i,j}^2 - \bar{x}_i^2}{1+\alpha_T} = \frac{s_i^2}{1+\alpha_T}.$$

Now

$$\mathrm{var}(w_{i,k+1}z) = \mathrm{var}(w_{i,k+1})\,\mathrm{var}(z) + E^2(w_{i,k+1})\,\mathrm{var}(z) + \mathrm{var}(w_{i,k+1})E^2(z).$$

This can also be written as

$$\mathrm{var}(w_{i,k+1}z) = \frac{m(1-m)\beta_{k+1}(1-\beta_{k+1})}{(1+p_T)(1+\alpha_T)}\left(1 + \frac{\beta_{k+1}(1+\alpha_T)}{1-\beta_{k+1}} + \frac{m(1+p_T)}{1-m}\right)$$

and since $\gamma = m\beta_{k+1}$ we have

$$\mathrm{var}(w_{ik+1}z) = \frac{\gamma(1-m)(m-\gamma)}{m(1+\alpha_T)(1+p_T)}\left(1 + \frac{\gamma(1+\alpha_T)}{m-\gamma} + \frac{m(1+p_T)}{1-m}\right). \tag{32}$$

Then, this term does not depend on the observations and is a constant function of the parameters of the model. Finally

$$\begin{aligned}
\mathrm{cov}(\mathbf{x}_i'\mathbf{w}, w_{ik+1}z) &= \mathbf{x}_i'E(\mathbf{w}_i^*w_{i,k+1})m - \mathbf{x}_i'\boldsymbol{\beta}^*\beta_{k+1}m = -\bar{x}_i\gamma\left(1 + \frac{1}{1+\alpha_T}\right)\\
&= \mathbf{x}_i'\,\mathrm{cov}(\mathbf{w}_i^*, w_{i,k+1})m + \mathbf{x}_i'E(\mathbf{w}_i^*)E(w_{i,k+1})m - \mathbf{x}_i'\boldsymbol{\beta}^*\beta_{k+1}m\\
&\quad - \frac{\mathbf{x}_i'\boldsymbol{\beta}^*\beta_{k+1}m}{1+\alpha_T} + \mathbf{x}_i'\boldsymbol{\beta}^*\beta_{k+1}m - \mathbf{x}_i'\boldsymbol{\beta}^*\beta_{k+1}m\\
&= -\frac{\mathbf{x}_i'\boldsymbol{\beta}^*\gamma}{1+\alpha_T} = -\frac{\bar{x}_i\gamma}{1+\alpha_T}.
\end{aligned}$$

where $\bar{x}_i = \sum_{j=1}^k \beta_j x_{ji}$. Therefore we can write $\mathrm{var}(y_i|\mathbf{x}_i) = \mu_2 s_i^2 + \bar{x}_i\mu_1 + \mu_0$, where μ_0, μ_1 and μ_2 are given by (10), (11) and (12), respectively.

References

Allenby, G.M., Rossi, P.E., 1999. Marketing models of consumer heterogeneity. J. Econometrics 89, 57–78.
Carroll, J.D., Green, P.E., 1995. Psychometric methods in marketing research: Part I, conjoint analysis. J. Marketing Res. XXXII, 385–391.
Cronin, J., Taylor, S., 1992. Measuring service quality: a reexamination and extension. J. Marketing 56 (July), 55–68.
Cronin, J., Taylor, S., 1994. SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality. J. Marketing 58 (January), 125–131.
Gumpertz, M.L., Pantula, S.G., 1998. Regression, random coefficient. In: Kotz, S., Read, C.B., Banks, D.L. (Eds.), Encyclopedia of Statistical Sciences, vol. 2. Wiley, New York, pp. 581–586.
Härdle, W., 1990. Applied nonparametric regression. Econometric Society Monographs, vol. 19. Cambridge University Press, Cambridge.
Johnson, V., 1996. On Bayesian analysis of multirated ordinal data: an application to automated essay grading. J. Amer. Statist. Assoc. 91, 42–51.
Lawson, C.L., Hanson, R.J., 1974. Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, NJ.

Lenk, P.J., DeSarbo, W.S., Green, P.E., Young, M.R., 1996. Hierarchical Bayesian analysis: recovery of partworth heterogeneity from reduced experimental. Marketing Sci. 28 (May), 215–222.

Luce, D.R., Tukey, J.W., 1964. Simultaneous conjoint measurement: a new type of fundamental measurement. J. Math. Psychol. 1, 1–27.

Lynch, J.G., Bouzas, T.E., y Berg, S.V., 1994. Regulatory measurement and evaluation of telephone service quality. Management Sci. 40 (2), 168–194.

Mallet, A., 1986. A maximum likelihood estimation method for random coefficient regression models. Biometrika 73, 645–646.

McLachlan, G.J., Krishnan, T., 1997. The EM Algorithm and Extensions, Wiley, New York.

Moreno, A., Rios-Insúa, D., 1998. Issues in Service Quality Modelling. In: Bernardo, J.M., et al. (Eds.), Bayesian Statistics, vol. 6. pp. 441–457.

Ostrom, A., Iacobucci, D., 1995. Consumer trade-offs and the evaluation of services. J. Marketing 59, 17–28.

Parasuraman, A., Zeithaml, V.A., Berry, L.L., 1988. SERVQUAL: A multiple item scale for measuring consumer perceptions of service quality. J. Retailing 64, 12–40.

Parasuraman, A., Zeithaml, V.A., Berry, L.L., 1991. Refinement and reassessment of the SERVQUAL scale. J. Retailing 67 (4), 420–450.

Parasuraman, A., Zeithaml, V.A., Berry, L.L., 1994. Reassessment of expectations as a comparison standard in measuring service quality: implications for further research. J. Marketing 58, 111–124.

Peña, D., 1997. Measuring service quality by linear indicators. In: Kunst, P., Lemmink, J. (Eds.), Managing Service Quality, vol. II. Chapman Publishing Ltd., London, pp. 35–51.

Rossi, P.E., Gilula, Z., Allenby, G.M., 2001. Overcoming scale usage heterogeneity: a Bayesian hierarchical approach. J. Amer. Statist. Assoc. 96, 20–31.

Teas, R.K., 1993. Expectations, performance evaluation, and consumers' perceptions of quality. J. Marketing 57, 18–34.

Teas, R.K., 1994. Expectations as a comparison standard in measuring service quality: an assessment of a reassessment. J. Marketing 58 (January), 132–139.

Wedel, M., DeSarbo, W., 1994. A review of recent developments in latent class regression models. In: Bagozzi, R. (Ed.), Advanced Methods of Marketing Research, vol. I–II. Blackwell Publishers, Massachusetts, pp. 353–388.