Università degli Studi di Napoli Federico II
Scuola delle Scienze Umane e Sociali
Quaderni
11

ASMOD 2018
Proceedings of the International Conference on Advances in Statistical Modelling of Ordinal Data

Naples, 24-26 October 2018

Editors
Stefania Capecchi, Francesca Di Iorio, Rosaria Simone

Federico II University Press

fedOA

Università degli Studi di Napoli Federico II
Scuola delle Scienze Umane e Sociali
Quaderni
11

ASMOD 2018

Proceedings of the International Conference on Advances in Statistical Modelling of Ordinal Data

Naples, 24-26 October 2018

Editors

Stefania Capecchi, Francesca Di Iorio, Rosaria Simone

Federico II University Press

fedOA Press

ASMOD 2018 : Proceedings of the Advanced Statistical Modelling for Ordinal Data Conference : Naples, 24-26 October 2018 / editors Stefania Capecchi, Francesca Di Iorio, Rosaria Simone. – Napoli : FedOAPress, 2018. – (Scuola di Scienze Umane e Sociali. Quaderni ; 11).

Accesso alla versione elettronica:
http://www.fedoabooks.unina.it

ISBN: 978-88-6887-042-3
DOI: 10.6093/978-88-6887-042-3
ISSN Collana: 2499-4774

Università degli Studi di Napoli Federico II
Centro di Ateneo per le Biblioteche “Roberto Pettorino”
Piazza Bellini 59-60
80138 Napoli, Italy
http://www.fedoapress.unina.it/

Published in Italy

Modeling preferences: beyond the average effects

Cristina Davino*, Tormod Naes**, Rosaria Romano***,
Domenico Vistocco****

Abstract: Preference mapping are a collection of multivariate statistical techniques widely used by marketing and R&D divisions to understand which sensory characteristics drive consumer acceptance of goods. These techniques provide a perceptual map of the products based on the so-called sensory dimensions, on which the liking values for each consumer are regressed. This study proposes an innovative preference mapping based on the quantile regression. Using the quantile regression instead of the classical least squares regression allows to explore the whole distribution of the consumer preference. This permits to obtain additional information both at the individual consumer level, analyzing how the preference varies with respect to the different quantiles and at the general level, highlighting on the preference map consumers with homogeneous behaviors with respect to the different quantiles.

Keywords: Preference mapping, Rating data, Quantile regression.

1. Preference mapping

Preference mapping is a collection of multivariate statistical techniques that aim to analyze consumer acceptance of food and beverages products (Meilgaard, Civille, Carr, 2007). There are two different types of these methods, namely *internal preference mapping* and *external preference mapping* (Meullenet, Xiong, Findlay, 2008). Internal preference mapping uses consumer acceptance ratings to determine a multidimensional representation of products and consumers in a common space. External preference mapping (PREFMAP) uses sensory descriptive attribute ratings to obtain a multidimensional representation of products, sensory characteristics and consumers in a common space. PREFMAP is crucial to the food and beverages indus-

*University Federico II of Naples, cristina.davino@unina.it
**Nofima AS, tormod.naes@nofima.no
***University of Calabria, rosaria.romano@unical.it
****University of Cassino e del Lazio Meridionale, vistocco@unicas.it

tries to understand which sensory characteristics drive consumer acceptance of goods. This information is used by marketing and R&D divisions to adapt existing products or create new products that meet consumers' expectations.
The most common PREFMAP method consists of a two step procedure that combines principal component analysis (PCA) and least squares regression (LSR) (Naes, Brockhoff, Tomic, 2010). In the first step a *perceptual map* of the products is obtained through a PCA of the product-by-attribute sensory matrix, and the principal components obtained from the analysis are called *key sensory dimensions* (Meilgaard, Civille, Carr, 2007). In the second step, a regression model is used to fit each consumer in the perceptual space. The main assumption is that the preference of each consumer depends linearly on the sensory attributes. Furthermore, as the method is grounded on LSR, it focuses on the average effects of sensory dimensions.
In some situations it is also useful to study the whole distribution of the liking. At this aim, quantile regression (QR) (Koenker, 2005; Davino, Furno, Vistocco, 2013) can be used to provide an estimate of conditional quantiles of the dependent variable instead of conditional mean. QR was recently used in consumer study for relating liking to consumer factors (Davino, Romano, Naes, 2015), and for handling consumer heterogeneity (Davino, Romano, Vistocco, 2018). The aim of this study is to extend the use of QR to the PREFMAP in order to provide additional information, not only about how the sensory dimensions link to consumer preference on average. The classical approach to PREFMAP based on the LSR does not allow to distinguish consumers able to discriminate preferences among products from consumers with uniform liking. The use of QR is advisable to highlight precise consumers, that is consumers with a strong difference in the liking pattern. Applying the classical approach would obscure this information, treating uniform consumers alike the precise ones.
The study is structured as follows: i) the classical approach to PREFMAP based on LSR is presented in Section (2); ii) a new approach based on QR is described in Section (3); iii) results of the proposed method on a case study concerning consumer liking of apple juice are shown in Section (4).

2. External preference mapping by least squares regression

Let X be the sensory matrix ($I \times K$), where the entry x_{ik} is the measured value of product i and sensory attribute k ($i = 1, \ldots, I; k = 1, \ldots, K$). The PCA model to develop a perceptual map based on the sensory characteristics can be written as

$$X = TP^T + E \tag{1}$$

where T is the matrix ($I \times A$) of the *principal component scores* that are linear combinations of the original data X, and P is the matrix ($K \times A$) of the *loading values* that define the contribution of each of the original variables in the computation of the principal components. The matrix E represents random noise and A is the number of components included in the model.

Let Y be the liking matrix ($I \times J$), where the entry y_{ij} is the measured value of product i and consumer j ($j = 1, \ldots, J$). The liking values for each consumers are regressed onto the first *sensory dimensions*, generally the first two PC's (i.e., $A = 2$):

$$y_{ij} = \beta_{j1}t_{i1} + \beta_{j2}t_{i2} + \epsilon_{ij} \tag{2}$$

The final results of this two step procedure provide a perceptual map and a loadings plot. In the first, products are located on the basis of the sensory characteristics, while in the second consumers' preferences are visualized and the direction for their preferences are identified.

3. External preference mapping by quantile regression

QR can provide complementary information to the classical PREFMAP. At this aim, it is introduced in the second step of the previously described procedure, when liking for each consumer is related to the first sensory dimensions. As classical linear regression provides the estimation of the conditional mean of a response variable distribution as a function of a set of predictors, QR provides the estimation of the conditional quantiles of a response variable distribution as a function of a set of predictors. It results that Equation (2) can be generalized to the QR framework as:

$$y_{ij}(\theta) = \beta_{j1}(\theta)t_{i1} + \beta_{j2}(\theta)t_{i2} + \epsilon_{ij} \tag{3}$$

where $(0 < \theta < 1)$. The interpretation of the QR coefficients is analogous to LSR coefficients: they measure the rate of change of the θth quantile of the dependent variable distribution per unit change in the value of a given predictors, holding the others constant. It is potentially possible to estimate an infinite number of regression lines, but in practice a finite number is numerically distinct, which is known as the quantile process. In practice, it is quite common that each researcher defines the quantiles of interest which, in most cases, are the three quartiles. For each quantile of interest, a regression line is estimated and, consequently, a set of coefficients and a fitted response vector can be obtained.

With respect to each consumer, the introduction of QR in PREFMAP provides a set of coefficients for each quantile of interest. This information allows to measure what is the impact of a change in the sensory dimensions on the liking for the most and least preferred products. Note that consumers showing large differences between coefficients at two extremes quantiles correspond to consumers with a precise liking pattern. With respect to the whole panel of consumers, QR allow to obtain a consumer loading plot that visualizes groups of consumers who are similarly affected by a given change on the sensory dimensions. In the case study section, it is also suggested a conjoint representation able to simultaneously represent results related to two opposite quantiles (e.g $\theta = 0.25$ and $\theta = 0.75$).

4. Case study

The data used for this study have been obtained from the article by Rdbotten *et al.* (2009). Apple juice samples were selected according to an experimental design (a 2*3 factorial design) with two levels of acid concentration (H=high, L=low) and three levels of sugar concentration (H=high, M=medium, L=low). The 6 samples were tested by 125 consumers using the 9-point hedonic scale (Peryam and Pilgrim, 1957). Descriptive sensory analysis was also carried out, and details of the procedure are given in (Rdbotten *et al.*, 2009). Results from classical PREFMAP are given in Figure (1). A joint interpretation of the

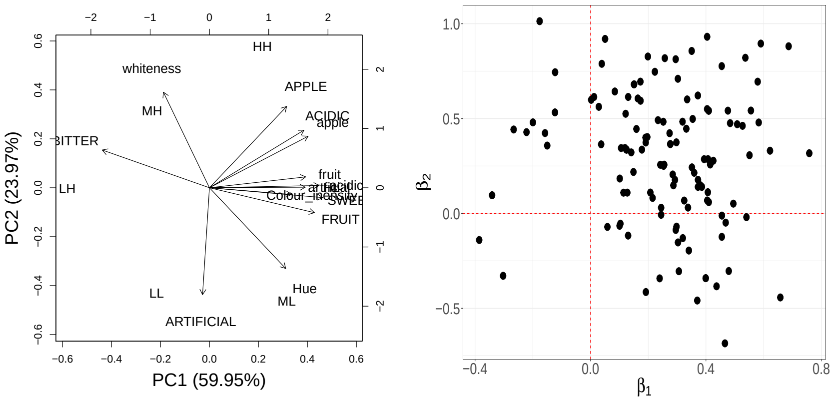

Figure 1. External preference mapping results

two plots shows that almost all consumers prefer sweet products, but some of them prefer products with high acid content, while the others prefer a low acid content. Note that products were evaluated both for flavor (descriptors labelled in capital letters) and smell (descriptors labelled in lower case).

4.1. Exploiting QR for single consumers

Consider estimating a QR model for each individual consumer and for a set of quantiles of interest, that is the three quartiles (θ = [0.25, 0.5, 0.75]). Three sets of coefficients are then estimated for each consumer. Figure 2 shows QR coefficients for two consumers, namely C27 (left-hand plot) and C49 (right-hand plot). For each plot, each panel represents a single regression coefficient (for sake of brevity, the the intercept is not shown). The horizontal axis displays the different quantiles, while the effect of each regressor holding the other constant is represented on the vertical axis. Standard errors and confidence intervals can also be added to the graph. For consumer C27 the two PCs have a different impact on the liking of C27: the coefficients β_1 are always higher than coefficients β_2 that are even negative. As discussed in Section (3), a regression coefficient at a specific quantile provides information of the effect of predictors on the selected conditional quantile of the liking distribution. For instance, C27 shows a $\beta(0.25)$ equal to 0.5 while the $\beta(0.75)$ is larger. This means that the effect of the predictors on the conditioned upper part of the liking distribution is stronger: increasing the level of sweetness (positive verse of PC1) increases more the preference for the most preferred products than for the less liked ones. Considering the lines related to the β_1 coefficients, it results that, modifying the sensory attributes explaining the first PC, has always positive impact on the liking of C27 but, moving from lower to higher quantiles, the effect of PC1 on liking increases showing that the most preferred products could take more advantage of an increment of the sensory attributes correlated to PC1. The opposite holds for β_2. Figure 2 (right-hand side) shows results for another consumer (C49). Here, the sensory dimensions not only have a different size compared to the different quantiles, but also a different sign. An increase in the level of sweetness (PC1) would increase the preference for the less preferred products ($\theta < 0.50$), while it would reduce the preference for the most preferred ones ($\theta = 0.75$).

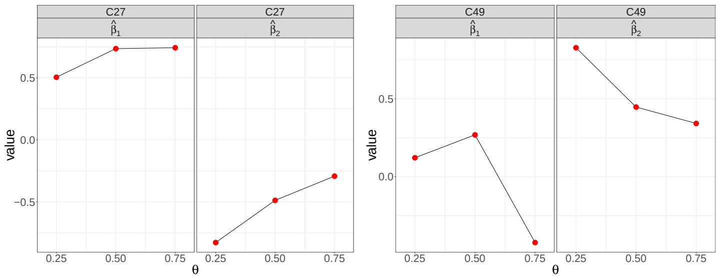

Figure 2. QR coefficients for single consumers

4.2. Exploiting QR for the whole panel

One of the main strengths of preference mapping is to suggest possible drivers to increase the liking, taking into account that whatever action will have different impacts on different groups of consumers.

Exploiting the proposed quantile approach, a further representation is proposed to simultaneously provide QR results at the two extreme considered quantiles. Figure 4 represents a group of 11 consumers. They have been selected as sample consumers because they shows different behaviors. Each

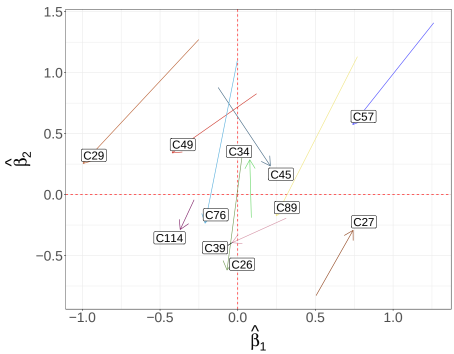

Figure 3. QR loading plot considering two extreme quantiles

consumer is represented according to the β_1 and β_2 coefficients estimated at the two quantiles. The two points representing each consumers are linked by an arrow depicted in the direction from $\theta = 0.25$ to $\theta = 0.75$. Considering consumer number 57 (from now C57), it is possible to appreciate that both coefficients related to PC1 and PC2 are positive but any action on the variables correlated to them will have a higher impact on the liking of the less preferred products. Arrows crossing two quadrants represent consumers with non-concordant signs at $\theta = 0.25$ and $\theta = 0.75$. It is the case of consumer C49, previously discussed. It is worth to note that consumers able to discriminate preferences among products are represented by a longer arrow, than consumers with uniform liking.

Finally, a plot that combines the results of the LSR and QR approach to PREFMAP is shown in the Figure (4). The different symbols correspond to all the different possible directions for the arrows in the previous plot. Specifically, the symbols corresponding to two equal numbers indicate consumers who are located in the same quadrant since coefficients of the two quantiles with respect to the two components have the same signs. For instance, consumer C27 included in the (4,4) group, has β_1 coefficients both positive for the two quantiles, and β_2 both negative. While consumer C49, included in the (1,2) group, has coefficients at $\theta = 0.25$ in the first quadrant and coefficients at $\theta = 0.75$ in the second quadrant. The information provided by this plot is very important because it allows to visualize the variability of preferences

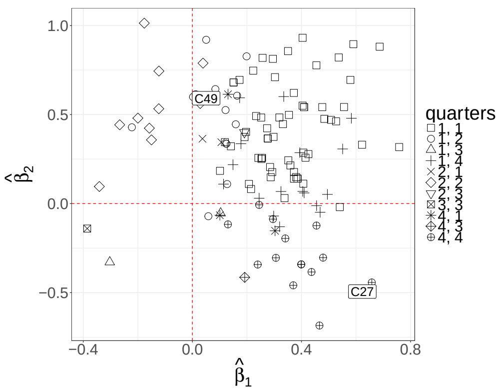

Figure 4. Loadings plot combining LSR and QR approach

with respect to preference directions. As an example, if we consider the direction of maximum preference, i.e. consumers in the first quadrant, we note that not all of them have coefficients consistent with the different quantiles. Consumers labeled with a cross show discrepancies. For a detailed analysis of these discrepancies we must then consider the arrows plot in Figure (4).

References

Davino C., Romano R., Naes T. (2015) The use of quantile regression in consumer studies, *Food quality and preference*, 40, 230-239.

Davino C., Furno M., Vistocco D. (2013) *Quantile regression: theory and applications*, John Wiley & Sons Ltd, United Kingdom.

Davino C., Romano R., Vistocco, D. (2018) Modelling drivers of consumer liking handling consumer and product effects, (forthcoming).

Koenker R. (2005) *Quantile regression*, Econometric Society Monograph, 38, Cambridge University Press, New York.

Meilgaard M.C., Carr B.T., Civille G.V. (1999) *Sensory evaluation techniques*, CRC press, Boca Raton.

Meullenet J.F., Xiong R., Findlay C.J. (2008) *Multivariate and probabilistic analyses of sensory science problems*, John Wiley & Sons Ltd, United Kingdom.

Naes T., Brockhoff P.B., Tomic O. (2010) *Statistics for Sensory and Consumer Science*, John Wiley & Sons Ltd, United Kingdom.

Peryam D.R., Pilgrim F.J. (1957) Hedonic scale method of measuring food preferences, *Food technology*, 11, 9-14.

Rdbotten M., Martinsen B.K., Borge G.I., Mortvedt H.S., Knutsen S.H., Lea P., Naes T. (2009) A cross-cultural study of preference for apple juice with different sugar and acid contents, *Food quality and preference*, 20, 277-284.