Quaderni di Statistica
Vol. 13, 2011

Analyzing and modelling rating data for sensory analysis in food industry

Marica Manisera
Department of Quantitative Methods, University of Brescia
E-mail: manisera@eco.unibs.it

Domenico Piccolo
Department of Theory and Methods of Human and Social Sciences
University of Napoli Federico II
E-mail: domenico.piccolo@unina.it

Paola Zuccolotto
Department of Quantitative Methods, University of Brescia
E-mail: zuk@eco.unibs.it

Summary: Consumers' and experts' preferences and perceptions of the sensory attributes of products are very important for manufacturers in the food industry, in order to avoid market disappointment and improve food quality. Indeed, appropriate sensory analyses combined with proper statistical methods allow to segment market, obtain positioning of products (brands, organizations, etc.) and identify the market acceptability. This finally has a great impact upon food quality and industrial competitiveness. In this paper, we use CUB models to analyze sensory data coming from a survey on the Italian espresso.

Keywords: Sensory analysis, CUB models, Italian coffee.

1. Introduction

Sensory evaluation is a scientific method where experimental results are collected on a set of sampled consumers who express preferences and reactions with respect to food and drink. Since samples are generally obtained according to standard statistical designs, this field attracts many approaches for a correct analysis of the results insofar as formal conditions for inferential procedures are respected.

On the other hand, consumer preferences result from complex interactions where subjective, objective and contextual factors are present with different roles. In fact, the expressed choice is the result of a human decision and we should assume that this process is a final act conditioned by personal history, environmental variables, subjective covariates and objects' characteristics, which all surely interact with the modality of the survey. As a consequences, it may be worth studying the stochastic structure of the choice process in order to adequately model the observed preferences.

Operationally, to collect sensory data, experts or untrained subjects are asked to rate or rank different products on the basis of some sensory descriptors (items), by expressing their perceptions on hedonic response scales (usually 9-point Likert scales). For example, consumers can be asked to evaluate quality attributes and express their preferences towards colour, smell, taste and mouth feel for a collection of coffee varieties, as we will pursue in this paper.

In this way, affective tests concern ordinal measurements. Such scales are substantially of qualitative nature although some numerical coding, as the integers $\{1, 2, \ldots, m\}$, is generally proposed. Then, a correct statistical analysis must be related to ordinal data modelling and current literature focuses on the models generated by cumulative probability in order to take the ordinal nature of sensory data into account (Agresti, 2010).

In this paper, following previous research in the area promoted by Piccolo (2003), we adopt a different structure by assuming that the response of each consumer is the combination of a *feeling* attitude towards the food being evaluated and an intrinsic *uncertainty* component surrounding the discrete choice.

This class of models have been successfully applied in several fields (D'Elia and Piccolo, 2005; Iannario, 2007) and sensory analysis is a favoured context (Piccolo and D'Elia, 2008; Piccolo and Iannario, 2010). In fact, these models allow to measure how the perception process is transformed into personal evaluations which are a mixture of several components: as already explained, the relevant ones are defined as feeling and uncertainty. Moreover, we will show that the added value of the proposal is mainly related to a sharp visualization of a huge amount of information by a graphical pattern of the estimated models represented in the parametric space.

This work shows how several varieties of Italian coffee (*espresso*) have been rated by a number of Italian and foreign tasters with respect to visual, olfactory and gustatory perceptions. The data set has been released without information on product and usage characteristics and the whole analysis will be concerned with the ability of the proposed models to cope with information derived by the frequency distributions of expressed preferences. The paper is organized as follows: in Section 2, we discuss the fundamentals of CUB modelling approach and in Section 3 we present the case study. Some final remarks conclude the paper.

2. CUB models

As mentioned in Section 1, the observed preferences result from the consumers' evaluation of food and drink, that is the expression of their preferences on a hedonic response scale. Perception and evaluation result from complex psychological mechanisms determined by many interacting factors of different nature (psychological, social, biological, physiological, etc.). Especially when eating and drinking behaviour is involved, human decision making occurs at a non-conscious level and sensory and consumer research should take psychological insights into account (Köster, 2009).

The philosophy of CUB models is perfectly in line with this, since feeling and uncertainty represent the latent components combined together in order to express the consumers' judgements (i.e., the observed discrete choices). The feeling component is the degree of agreement with a given item and results from subjective motivations. According with the latent variable approach, it can adequately be interpreted as a continuous latent random variable which is interpreted as discretized, since the consumers' ratings assigned to an item are discrete. On the other hand, the uncertainty component is the unstability intrinsically present in the human choices and resulting from factors related to the evaluation process (for example, the limited knowledge of the problem, the nature of the chosen questionnaire and response scale, the subjective interest towards items). Both components are explicitly considered in the CUB models, by means of a mixture of two random variables, as explained in Subsection 2.1.

2.1. Basic issues

Any model is strictly arbitrary; thus, the rationale of a modelling structure may be only derived from a blend of logical arguments and empirical facts. Overall, parsimony in parameters is a key issue. In line with these arguments, the class of models we are going to introduce aims at parametrically defining the behaviour of respondents as generated by two main latent components.

Specifically, *uncertainty* may be modelled with regard to the extreme choice of a person who assigns the same probability to each category, with a complete indifference. As a consequence, for the distribution related to uncertainty we introduce the discrete Uniform random variable U defined over the support $\{1, 2, \dots, m\}$, for a given m:

$$P_r\left(U = r\right) = \frac{1}{m} = U_r, \quad r = 1, 2, \dots, m\,.$$

This random variable maximizes the entropy, among all the discrete distributions with finite support $\{1, 2, \dots, m\}$, for a fixed m, and it is minimally informative about the choice (when one knows only the number m of modalities).

Instead, we model the *feeling* component by means of a shifted Binomial random variable V whose probability distribution is:

$$P_r\,(V=r \mid \xi) = \binom{m-1}{r-1}\,\xi^{m-r}\,(1-\xi)^{r-1} = b_r(\xi)\,, \quad r=1,2,\ldots,m\,,$$

The rationale for such distribution stems from heuristic and pragmatic point of views: the (shifted) Binomial distribution is able to cope with different shapes of sample data and just with a single parameter. In addition, from a statistical point of view, combinatorial and selective arguments confirm the convenience to adopt such distribution, as argued by Iannario (2012).

If we weight the components assumed for uncertainty and feeling, we are introducing a (convex) ***C***ombination of a discrete ***U***niform and a shifted ***B***inomial distributions, and this justifies the CUB acronym. Then, a CUB random variable R expressing the final choice of the respondent is defined by the probability mass function:

$$P_r\,(R=r \mid \boldsymbol{\theta}) = \pi\, b_r(\xi) \,+\, (1-\pi)\,U_r\,, \quad r=1,2,\ldots,m\,,$$

where $\boldsymbol{\theta} = (\pi,\xi)'$, $\pi \in (0,1]$ and $\xi \in [0,1]$. Then, the parametric space is the (left open) unit square:

$$\Omega(\boldsymbol{\theta}) = \Omega(\pi,\xi) = \{(\pi,\xi):\ 0<\pi\leq 1;\ 0\leq\xi\leq 1\}\,.$$

Iannario (2010) proved that CUB models are identifiable for any $m>3$.

The class of CUB models turns out to be a very flexible parametric family since the shape of the distribution largely varies over $\Omega(\pi,\xi)$, as shown by Piccolo (2003). This allows to fit data with positive or negative skewness, any intermediate modal value and also peaked or flat distributions.

Parameters are associated to the latent components of the responses and may be easily interpreted. The *feeling parameter* (ξ) is mostly related to location measures and strongly determined by the skewness of responses: it increases when respondents prefer low ratings. Usually, high values of the responses imply high consideration towards the food; then, in sensory analysis, the quantity $(1-\xi)$ increases with sensory satisfaction with the product. Instead, the *uncertainty parameter* (π) modifies the heterogeneity of the distribution and it is mostly related to the comparisons among probabilities. Then, uncertainty of the choice increases with $(1-\pi)$.

Since there is one-to-one correspondence among a CUB random variable and the parameter vector $\boldsymbol{\theta}=(\pi,\xi)'$, we represent each CUB model as a point in the unit square. This visualization is a focal issue of the approach since a single point summarizes any aspect of the probability distribution and allows for immediate comparison with respect to time, space and circumstances.

Since $1-\pi$ measures the *propensity* of respondents to behave in accordance to a completely random choice, and $1-\xi$ measures the *strength of feeling* of the subjects for a direct and positive evaluation of the food, hereafter we will consider the plot of CUB models as a point in $\Omega(\pi,\xi)$ with coordinates $1-\pi$ and $1-\xi$, respectively.

The expectation of R is given by: $\mathbb{E}\,(R) = \frac{(m+1)}{2} + \pi\ (m-1)\left(\frac{1}{2} - \xi\right)$. It confirms that the mean value moves towards the central value of the support on the basis of the value and sign of $(\frac{1}{2} - \xi)$ and this behaviour is related to the skewness of the distribution. In fact, a CUB random variable is symmetric if and only if $\xi = 1/2$.

A peculiar aspect of the last formula is that the expectation of R is constant for infinitely many values of the parameter vector $\boldsymbol{\theta} = (\pi,\, \xi)'$; as a consequence, we may obtain the same mean value for ratings distributions which are quite different. In addition, the expectation does not convey all the characteristics of a random phenomenon since these are explained by a sequence of higher moments instead.

CUB models have been extended in several directions as recently pointed out by Iannario and Piccolo (2012), and these generalizations concern the probability distribution of the components, the inclusion of subjects' and objects' covariates, the joint consideration of several objects/items in a multivariate context.

For example, if one considers that both uncertainty and feeling may be conditioned by subjects' characteristics, we can define CUB models with covariates by introducing a logistic link among parameters and covariates of the respondents. This extension is particulary noticeable since it allows for testing and measuring the effect of known characteristics on the responses and thus these extended models are especially valuable for marketing studies.

Another generalization stems from the circumstance that respondents may sometimes prefer a quick response instead to weigh up more demanding choices. This behaviour is frequent in sensory analysis and induces an anomalous value of the frequency of a given category. Since this component may imply both biases and inefficiencies in the statistical analysis, it can explicitly be modelled in CUB models with a *shelter effect* (Corduas *et al.*, 2009; Iannario, 2012).

2.2. Inferential issues

When sample data are available, the classical steps of the iterative cycle of specification, estimation and validation of a CUB model may be consistently pursued by maximum likelihood (ML) methods which lead to asymptotically efficient properties of the statistical procedures. Moreover, the involved mixture distribution suggests to introduce the EM procedure as an effective algorithm to reach convergence almost everywhere on $\Omega(\boldsymbol{\theta})$, as shown by Everitt and Hand (1981), McLachlan and Krishnan (2008), McLachlan and Peel (2000), among others.

For a general CUB model with covariates, the ML estimation has been derived by Piccolo (2006) and extended by Iannario (2012) to models with *shelter effect*. Several suggestions have been proposed for improving the convergence of the procedure by means of accurate preliminary estimators. In this context, the significance of the estimated parameters, the relevance of the covariates and the validation of the model are obtained by exploiting the asymptotic properties of the ML estimators.

A critical review of the fitting measures for ordinal models, and specifically for CUB models, is in Iannario (2009). When sample data are summarized by the observed frequencies $n_1, n_2, \ldots, n_m$, the log-likelihood function of the *saturated* CUB model is

$$\ell_{sat} = -n\,\log(n) + \sum_{r=1}^{m} n_r\,\log(n_r)\,,$$

whereas the log-likelihood function for a discrete Uniform distribution is:

$$\ell_0 = -n\,\log(m)\,.$$

In a sense, ℓ_0 is the worst achievable value for a likelihood function computed on ordinal data since it derives from a totally uninformative situation except for the number of modalities. On the contrary, ℓ_{sat} is the maximum achievable for a likelihood function given the observed data and it acts as a benchmark for comparing the effectiveness of parametric structures. Both quantities are easily computable on the basis of sample data and may be fruitfully used for fitting purposes.

Then, if $\ell(\hat{\boldsymbol{\theta}})$ is the log-likelihood of the estimated model, a convenient measure of fitting has been proposed as

$$\mathcal{I} = \frac{\ell(\hat{\boldsymbol{\theta}}) - \ell_0}{\ell_{sat} - \ell_0} = \frac{\ell(\hat{\boldsymbol{\theta}}) + n\,\log(m)}{\displaystyle\sum_{r=1}^{m} n_r\,\log(n_r) + n\,\log(m/n)}\,.$$

A further normalized fitting measure has been introduced for comparing observed f_r and expected $p_r(\hat{\boldsymbol{\theta}})$ relative frequencies:

$$\mathcal{F}^2 \;=\; 1 - \frac{1}{2}\sum_{r=1}^{m} \mid f_r - p_r(\hat{\boldsymbol{\theta}}) \mid\,.$$

It is related to a standard dissimilarity index and has an immediate interpretation as the proportion of correct predicted frequency responses.

A program in ***R*** –where the whole inferential procedure is effectively implemented with estimation, test results, statistical indexes and graphical displays– is freely available (Iannario and Piccolo, 2009). Work is currently in progress to release a standard ***R*** package.

3. Case study

In this Section we present the results of a case study dealing with sensory data about coffee tasting. Usually the coffee tasting method consists of three main evaluations (sensory attributes):

- the *visual analysis*, taking into account the colour (should not be either too light or too dark, but rather nutty-color with dark red streaks), the texture (should be dense, with a fine texture and without any gaps), and the persistence (quite long) of the cream;
- the *olfactory analysis*, taking into account the smell (should be pleasant and intense) and fragrances or aromas (toasted, chocolaty, floral, fruity, peanuts, spiced, ...);
- the *gustatory analysis*, taking into account flavour (sweet, acidic, bitter) and aftertaste (aroma, persistence).

The survey which produced the analyzed data was carried out by Centro Studi Assaggiatori (CSA, http://www.assaggiatori.com) of Brescia, Italy, along with the International Institute of Coffee Tasters (IIAC)[1] and was concerned with the sensory analysis of 43 different coffee varieties, evaluated by a number of experienced and non-experienced judges through the above described tasting method. For each coffee variety a set of judges (from a minimum of 6 to a maximum of 421) was selected from the 1650 judges involved in the survey, who formulated visual, olfactory, gustatory evaluations of the coffee on an 9-point Likert scale. After removing the coffee varieties evaluated by less than 60 judges, the data set turns out to be composed by 36 coffee varieties for which a total number of 7604 judgments on each sensory attribute are available. On the whole, each of the 1650 judges was asked to taste from a minimum of 1 to a maximum of 11 coffees, but more than 78% of judges tasted exactly 5 coffee varieties. For each judge some personal information is also available (gender, age, experience in tasting, consumption, etc.).

We fit CUB models separately to each of the 36 varieties of coffees with respect to visual, olfactory and gustatory perceptions. The estimated models are all significant and with good fitting measures ($\mathcal{F}^2$ varies in $(0.85,\ 0.97)$), as shown in Figure 1 (more detailed results are available from Authors).

We summarize results by plotting the estimated parameter vectors on the unit square. So, according to the estimated CUB models we locate the 36 coffee varieties on a map describing their relative positioning with respect to the selected sensory attributes, focusing attention on both the level of their evaluation and the degree of uncertainty of the judgements (Figure 2). It should be evident how the complex pattern of this experiment may be sharply simplified by CUB models in a unique representation. The ranking of preferences is not constant with respect to the three evaluations and this confirms that respondents react in different ways when faced to visual, olfactory and gustatory sensations. The close position of visual and olfactory perceptions is a further confirmation of well known results in sensometrics: as a matter of fact, sight and smell are senses which manifest themselves with high similarity.

[1] The authors thank Luigi Odello (director of CSA) and prof. Eugenio Brentari (University of Brescia) for making the data available.

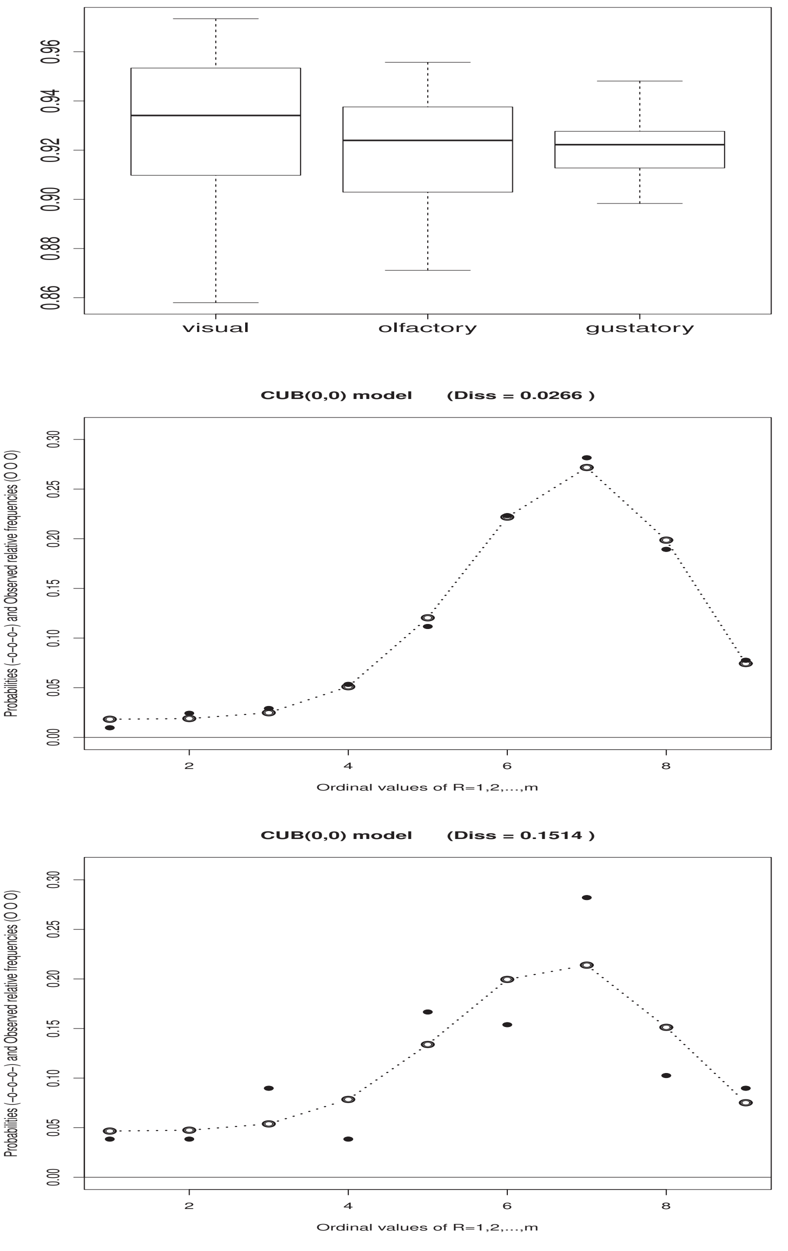

Figure 1. Boxplots of $\mathcal{F}^2$ separately for the three sensory attributes (left). Plot of estimated probabilities versus observed relative frequencies in the best ($\mathcal{F}^2 = 1 - Diss = 0.973$) and worst ($\mathcal{F}^2 = 1 - Diss = 0.849$) case (middle and right, respectively).

In addition, we notice that all evaluations (except for varieties 34 and 35) are expressed with a limited uncertainty, confirming that respondents are giving meditated preferences. However, the uncertainty generally increases when we pass from visual to olfactory and then to gustatory perceptions. Thus, gustatory perceptions are more related to subjectivity than olfactory perceptions, which, in turn, are more related to subjectivity than the visual ones. Thus, we conjecture that perceptions more heavily depend on the personal history, attitude, and habits when moving from visual to gustatory.

Figure 3 shows the estimated models of visual, olfactory and gustatory perceptions for each coffee variety separately: gustatory perceptions are generally more uncertain and also the atypical location of varieties 34 and 35 is confirmed.

None of the personal characteristics of the judges available from the questionnaire turned out to be a significant covariate for the estimated CUB models.

At this step, it is interesting to inspect the relationships among the judges' perceptions expressed through the visual, olfactory and gustatory ratings and the satisfaction about each coffee variety. Since the gustatory satisfaction towards coffee is significantly dependent by both visual and olfactory ratings but olfactory is much more relevant, we explore the relationship between the expressed level of olfactory rating and the gustatory satisfaction for the 36 coffee varieties, separately.

More specifically, using the expressed scores on the olfactory sensory attribute as covariate in the CUB models of gustatory satisfaction, we verify if gustatory satisfaction can be predicted by means of single judges' perceptions on the olfactory attribute. The logistic link $\xi_i = \dfrac{1}{1+e^{-\boldsymbol{w}_i \gamma}}$ is introduced in the CUB models, with ξ_i and $\boldsymbol{w}_i$ indicating the gustatory satisfaction and the olfactory rating of subject i, for $i = 1, 2, \ldots, n$, respectively.

Results are plotted in Figure 4. The coffee varieties show different reaction rates but it is insightful to observe that the shape is regularly homogeneous for all the coffees. This result may be usefully exploited by producers of coffees with poor olfactory perceptions, since an improvement of the consumer gustatory perception towards the product seems to be highly dependent upon a positive evaluation of the coffee's smell.

4. Concluding remarks

In this paper, CUB models have been studied for interpreting uncertainty and feeling of different brands of coffee. The experimental results obtained on a very large data set of different brand of coffees confirm that CUB models may be usefully exploited for comparing and summarizing several aspects of the data in an effective graphical display.

In addition, this class of models manifest themselves as useful also for measuring the predicting ability of gustatory responses given the olfactory ones.

The analysis so far proposed may be further deepened if we could insert product characteristics in the sensory analysis. It could allow, for example, to identify which coffee varieties show a peculiar behaviour, in order to better understand relations among

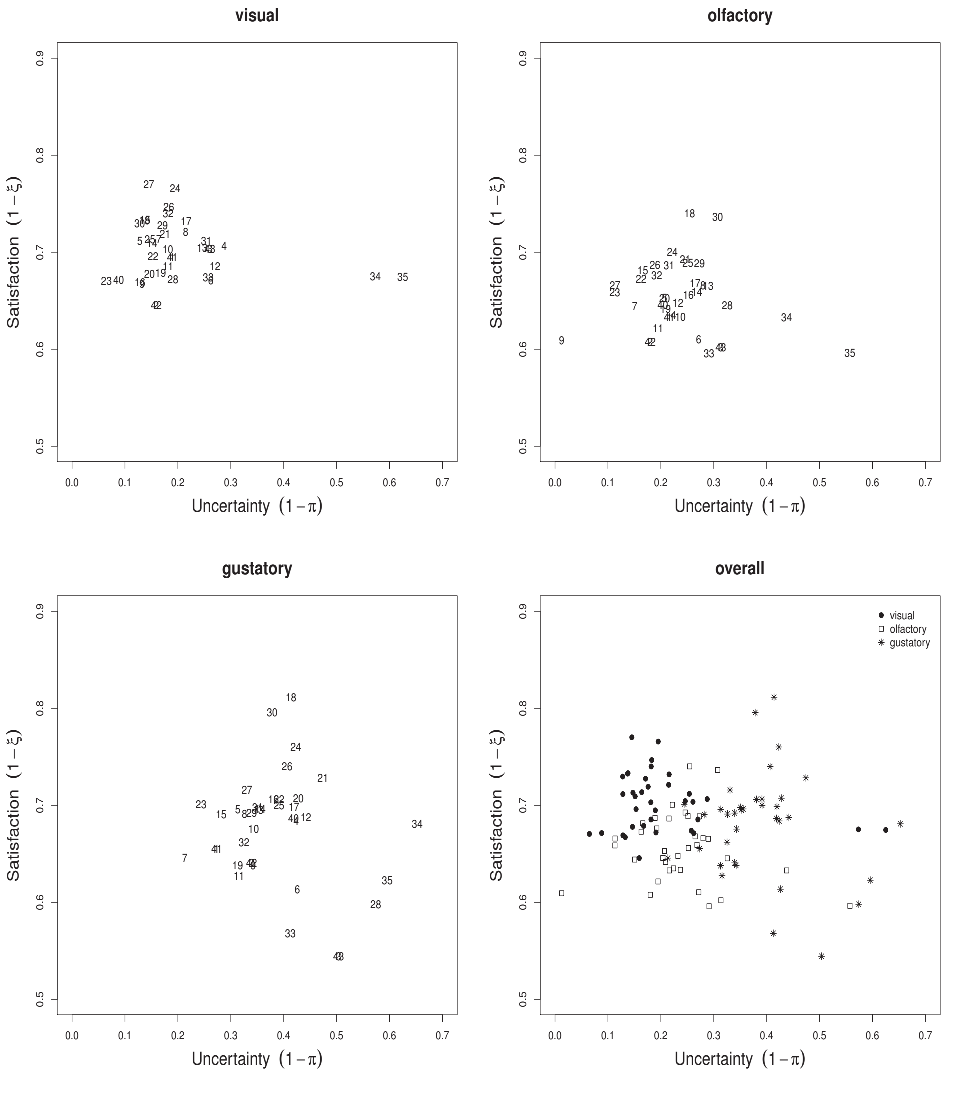

Figure 2. CUB *models visualization of visual, olfactory, gustatory perceptions of the 36 coffee varieties.*

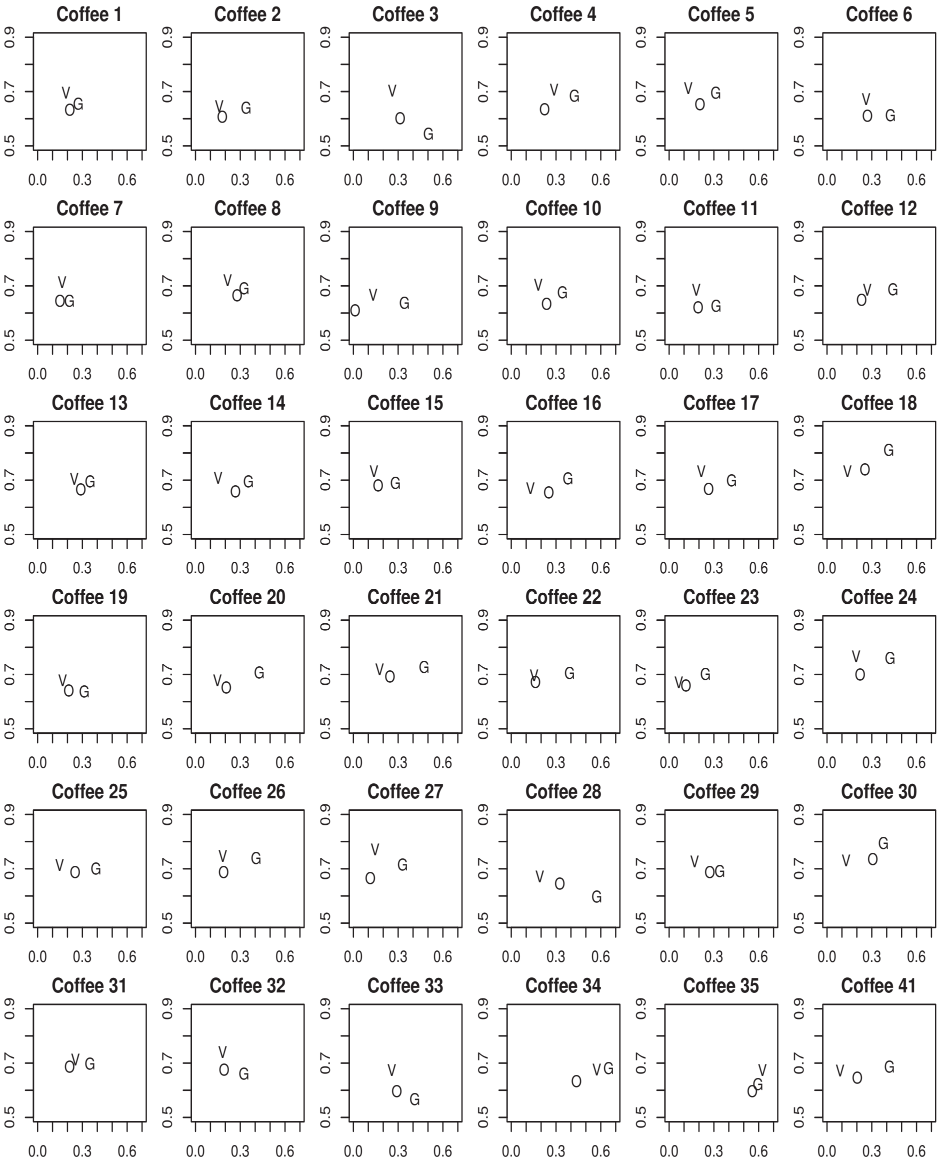

Figure 3. CUB *models visualization of visual, olfactory and gustatory perceptions for each coffee variety.*

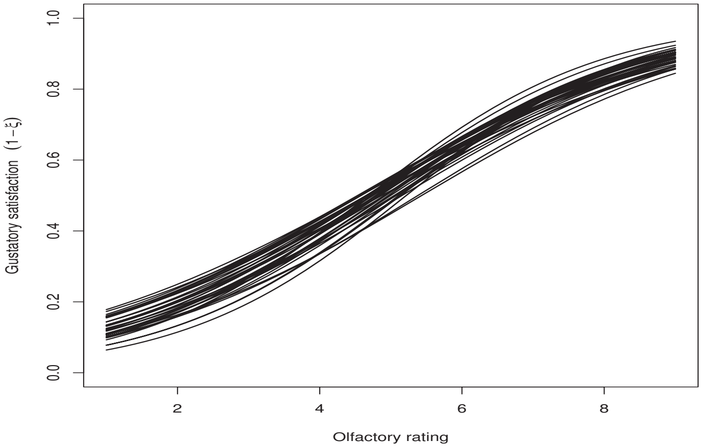

Figure 4. Prediction of gustatory satisfaction given the olfactory rating.

variety and perceptions and to finally direct the manufacturers' efforts to improve their competitiveness.

Results from CUB models could be integrated with other advanced statistical techniques useful for sensory analysis (among others, Brentari and Zuccolotto, 2011) in order to get a more complete picture of the phenomenon (see, for example, Iannario *et al*., 2011).

Acknowledgements: The second Author has been partly supported by MIUR projects PRIN2008: "Modelling latent variables for ordinal data: statistical methods and empirical evidence" (CUP E61J10000020001), and FARO 2011 project research, both at University of Naples Federico II. A preliminary version of the paper has been presented at AGROSTAT Meeting, Paris.

References

Agresti A (2010), *Analysis of Ordinal Categorical Data*, 2^{nd} ed., J.Wiley & Sons, Hoboken.

Brentari E., Levaggi R., Zuccolotto P. (2011), Pricing strategies for Italian red wine, *Food Quality and Preference*, 22, 725–732.

Brentari E., Zuccolotto P. (2011), The impact of chemical and sensory characteristics on the market price of Italian red wines, *Electronic Journal of Applied Statistical Analysis*, 4, 265–276.

Carpita M., Manisera M. (2011a), On the imputation of missing data in surveys with Likert-type scales, *Journal of Classification*, 28, 93–112.

Carpita M., Manisera M. (2011b), On the nonlinearity of homogeneous ordinal variables, in: S. Ingrassia *et al.* (eds.), *New Perspectives in Statistical Modeling and Data Analysis*, Springer, 489–496.

Corduas M., Iannario M., Piccolo D. (2009), A class of statistical models for evaluating services and performances, in: M. Bini *et al.* (eds.), *Statistical methods for the evaluation of educational services and quality of products*, Springer, 99–117.

D'Elia A., Piccolo, D. (2005), A mixture model for preference data analysis, *Computational Statistics & Data Analysis*, 49, 917–934.

Everitt, B.S., Hand D.J. (1981), *Finite mixture distributions*, Chapman & Hall, London. Iannario M. (2007). A statistical approach for modelling urban audit perception surveys, *Quaderni di Statistica*, 9, 149–172.

Iannario M. (2009), Fitting measures for ordinal data models, *Quaderni di Statistica*, 11, 39–72.

Iannario, M. (2010), On the identifiability of a mixture model for ordinal data, *Metron*, LXVIII, 87–94.

Iannario M. (2012), Modelling *shelter* choices in a class of mixture models for ordinal responses, *Statistical Methods and Applications*, 21, 1–22.

Iannario M., Manisera M., Piccolo D., Zuccolotto P. (2012), Sensory analysis in the food industry as a tool for marketing decisions, *Submitted for publication*.

Iannario M., Piccolo D. (2009), A program in R for CUB models inference, Version 2.0, available at `http://www.dipstat.unina.it/CUBmodels1/`

Iannario M., Piccolo D. (2012), CUB models: Statistical methods and empirical evidence, in: Kenett R. S. and Salini S. (eds.), *Modern Analysis of Customer Surveys: with applications using R*, J. Wiley & Sons, Chichester, 231–258.

Köster E. P. (2009), Diversity in the determinants of food choice: A psychological perspective, *Food Quality and Preference*, 20, 70–82.

Manisera M. (2011), A graphical tool to compare groups of subjects on categorical variables, *Electronic Journal of Applied Statistical Analysis*, 4, 1–22.

Manisera M., Van der Kooij A. J., Dusseldorp E. M. L. (2010), Identifying the component structure of satisfaction scales by nonlinear Principal Components Analysis, *Quality Technology and Quantitative Management*, 7, 97–115.

McLachlan G., Krishnan T. (2008), *The EM algorithm and extensions*, 2^{nd} ed., Wiley, NY.

McLachlan G., Peel G. J. (2000), *Finite mixture models*, J. Wiley & Sons, NY.

Piccolo D. (2003), On the moments of a mixture of uniform and shifted binomial random variables, *Quaderni di Statistica*, 5, 85–104.

Piccolo D. (2006), Observed information matrix for MUB models, *Quaderni di Statistica*, 8, 33–78.

Piccolo D., D'Elia A. (2008), A new approach for modelling consumers' preferences, *Food Quality and Preference*, 19, 247–259.

Piccolo D., Iannario M. (2010), A new approach for modelling consumers' preferences, *Proceedings of the* 11^{th} *European Symposium on Statistical Methods for the Food Industry*, University of Sannio, Benevento, Academy School, Afragola, 139–148.

Zuccolotto P. (2009), Missing values treatment with Interval Imputation in Satisfaction Measurement, *Quaderni di Statistica*, 11, 145–163.

Zuccolotto P. (2010), Measuring Job and Patient Satisfaction in presence of missing values: a case study with Interval Imputation, *Electronic Journal of Applied Statistics: Decision Support Systems and Services Evaluation*, 1, 79–94.