WILEY | Hindawi

Research Article

Ordinal Analysis of Variation of Sensory Responses in Combination with Multinomial Ordered Logistic Regression vs. Chemical Composition: A Case Study of the Quality of a Sausage from Different Producers

Tamar Gadrich [ID],[1] **Francesca R. Pennecchi** [ID],[2] **Ilya Kuselman** [ID],[3] **D. Brynn Hibbert** [ID],[4] **Anastasia A. Semenova** [ID],[5] **and Pui Sze Cheow** [ID][6]

[1]*Braude College of Engineering, Department of Industrial Engineering and Management, P.O. Box 78, 51 Snunit St. 2161002, Karmiel, Israel*
[2]*Istituto Nazionale di Ricerca Metrologica (INRIM), Strada Delle Cacce 91, Turin 10135, Italy*
[3]*Independent Consultant on Metrology, 4/6 Yarehim St., Modiin 7176419, Israel*
[4]*School of Chemistry, UNSW Sydney, Sydney, NSW 2052, Australia*
[5]*V. M. Gorbatov Federal Research Center for Food Systems, 26 Talalikhina St., Moscow 109316, Russia*
[6]*Health Science Authority, 1 Science Park Road #01-05/06, The Capricorn Singapore Science Park II, Singapore 117528*

Correspondence should be addressed to Ilya Kuselman; ilya.kuselman@gmail.com

The newly developed statistical technique of two-way ordinal analysis of variation (ORDANOVA) was applied for the first time to sensory responses in combination with multinomial ordered logistic regression of a response category vs. chemical composition. A corresponding tutorial is provided. As a case study, samples of a sausage from different producers, purchased at the same time from a market, were compared based on sensory responses of experienced experts. A decomposition of total variation of the ordinal data and simulation of the multinomial distribution of the relative frequencies of the responses in different categories showed a statistically significant difference between the producers' samples, and an insignificant difference between the experts' responses related to the same sample. The capabilities of experts were also evaluated. The influence of chemical composition of a sausage sample on the probability of a response category was modeled using multinomial ordered logistic regression of the response on mass fractions of the main sausage components. This statistical technique can be helpful for understanding sources of variation of sensory responses on food quality properties. It is also promising for a revision of specification limits for chemical composition, as well as for the prediction of sensory properties when the chemical composition of the product is subject to quality control.

1. Introduction

Despite the development of sophisticated chemical analytical methods and modern equipment, sensory analysis based on human organoleptic examination of food products, such as a sausage, will be always important for the people who are consumers of the products.

The international guidelines for sensory analysis [1] recommend the use of quantitative scales for human (assessor/expert) responses. For example, sensory responses on properties of a sausage quality can be classified by the following five quality categories: (1) very bad; (2) poor; (3) satisfactory; (4) good; and (5) excellent. Analysis of variance (ANOVA) is applied for interpretation of such property values [2–4]. Other statistical methods developed mainly for quantitative continuous values, including regression analysis, are also applied [5, 6]. An implementation of fuzzy logic for modification of sensory responses to obtain comparable

scores [7, 8] has attracted attention. Statistical methods for sensory analysis, in particular for meat and meat products, are regularly reviewed [9, 10].

However, responses on an ordinal property scale are categorical data for which a total ordering relation can be established according to the magnitude, with other quantities of the same kind, but no algebraic operations exist among those quantities [11, 12]. Their legitimate operations can be "equal/unequal" and "greater than/less than" [13]. The addition of categorical data is not a legitimate operation by definition, whereas one of the ANOVA assumptions is that the treated quantities are additive [14] and calculation of their mean and variance, is possible. Therefore, statistical techniques based on the ANOVA should not be applied directly to ordinal data.

Sensory responses may also depend on the chemical composition of a food product [15, 16]. Relationships between sensory evaluation and the chemical composition or physical properties of meat and meat products are a subject of research [17]. Regression analysis is the known tool for studying and modeling such relationships. However, as in the applications of ANOVA, the additivity assumption should not be violated in the regression analysis use [18].

The aim of the present paper is to overcome the "additivity" problem by implementation of the newly developed two-way ordinal analysis of variation (ORDANOVA) [19] combined with multinomial ordered logistic regression [20–22]. As a case study, samples of a kind of commercially available sausage from different producers were compared based on the ordinal data from the examination of quality properties by experienced experts. The influence of the chemical composition of a sausage sample on the probability of a response category was modeled.

2. Materials and Methods

2.1. Statistical Approach–A Tutorial

2.1.1. Two-Way ORDANOVA without Replication.
The approach of ORDANOVA is to calculate for each property the number of expert responses related to the same category, and then to analyze their relative frequency as a fraction of the total number of the responses for all categories.

A vector of expert responses for a given property as a random variable $\mathbf{Y} = (Y_1, Y_2, ., Y_K)$ on an ordinal scale with K categories (classes or levels $k = 1, 2, \ldots, K$) is characterized by a probability vector $\mathbf{p} = (p_1, p_2, \ldots, p_K)$, where p_k is the theoretical probability of responses related to the kth category and $\sum_{k=1}^{K} p_k = 1$. Let F_k denotes the cumulative theoretical probability up to the kth category, $F_k = \sum_{l=1}^{k} p_l$ and $F_K = 1$. The probability P of receiving a set of responses $\mathbf{n} = (n_1, n_2, \ldots, n_K)$, where n_k is the number of responses related to the kth category ($\sum_{k=1}^{K} n_k = N$), can be calculated based on the multinomial distribution with parameters (N, $\mathbf{p}$) as the probability mass function [23]:

$$P(n_1, n_2, \ldots, n_K) = \frac{N!}{n_1! \cdot n_2! \ldots n_K!} p_1^{n_1} \cdot p_2^{n_2} \cdots p_K^{n_K}. \quad (1)$$

Note that the probability of the event $Y_k = n_k$ is $P(Y_k = n_k) = p_k$ by definition. Variability in $\mathbf{Y}$ is explained in the present study by two independent factors as follows: sausage producers and experts as random factors. The examination of the samples of sausage was a properly conducted blind test, in which an expert has no opportunity to favor the sausage of a particular producer. No interaction between the two factors can be analyzed, since only one response at the specified levels of the factors (at each cell in the cross-over balanced design) was obtained, by analogy with laboratory proficiency testing [24]. The first factor $X1$ had I levels (sausage samples of I producers were tested) and the second factor $X2$ had J levels (J experts examined the samples' quality). There were N responses in total for a given property, each of them was of one of the K categories of variable $\mathbf{Y}$. Being a result of assessment of a sausage sample of producer i (one of the I levels of the first factor $X1$, $i = 1$, 2,..., I), each response was also associated with the expert j (one of the J levels of the second factor $X2$, $j = 1, 2, \ldots, J$). Thus, it was assumed that each one of the $I \times J = N$ cells (i, j) contains only one response belonging to one of the K categories k. In other words, a response at each cell (i, j) was of one category k chosen by expert j when examined a sample of producer i, and correspondingly, the number of responses of the chosen categories $n_{ijk} = 1$ and zero for all other categories in the same cell.

Treating N responses as a statistical sample and Y_{ijk} as a random variable, then, for the chosen k, $\widehat{p}_{ijk} = 1$ and zero for all other categories in the same cell. Thus, $\widehat{F}_{ijk} = \sum_{l=1}^{k} \widehat{p}_{ijl}$ denotes the sample cumulative frequency of responses of categories $l = 1, 2, \ldots, k$, i.e., up to the kth category in cell (i, j); $\widehat{F}_{i.k}$ is the sample cumulative relative frequency of responses up to the kth category at level i of factor $X1$; $\widehat{F}_{.jk}$ denotes the cumulative relative frequency at level j of factor $X2$; $\widehat{F}_{..k}$ denotes the cumulative relative frequency of all responses up to the kth category:

$$\widehat{F}_{i.k} = \frac{1}{J} \sum_{j=1}^{J} \widehat{F}_{ijk},$$

$$\widehat{F}_{.jk} = \frac{1}{I} \sum_{i=1}^{I} \widehat{F}_{ijk}, \quad (2)$$

$$\widehat{F}_{..k} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \widehat{F}_{ijk}.$$

The total sample variation $\widehat{V}_T$ of the response variable $\mathbf{Y}$, normalized to the [0, 1] interval, is defined in the two-way ORDANOVA model [19] as follows:

$$\widehat{V}_T = \frac{1}{(K-1)/4} \sum_{k=1}^{K-1} \widehat{F}_{..k} \left(1 - \widehat{F}_{..k}\right). \quad (3)$$

In the model without replication, $\widehat{V}_T$ is partitioned between the (inter) covariation component $\widehat{C}_B$ and the within (intra) residual variation $\widehat{V}_W$, i.e., $\widehat{V}_T = \widehat{C}_B + \widehat{V}_W$. For the present study, the variation

$$\widehat{C}_{\text{B}} = \frac{1}{(K-1)/4} \sum_{k=1}^{K-1} \left[\frac{1}{I} \sum_{i=1}^{I} \left(\widehat{F}_{i.k} - \widehat{F}_{..k} \right)^2 + \frac{1}{J} \sum_{j=1}^{J} \left(\widehat{F}_{.jk} - \widehat{F}_{..k} \right)^2 \right], \quad (4)$$

characterizes the between-producer variation of the responses, while the variation

$$\widehat{V}_{\text{W}} = \frac{1}{(K-1)/4} \sum_{k=1}^{K-1} \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \left(\widehat{F}_{i.k} + \widehat{F}_{.jk} - \widehat{F}_{..k} \right) \left(1 - \left[\widehat{F}_{i.k} + \widehat{F}_{.jk} - \widehat{F}_{..k} \right] \right), \quad (5)$$

is the within-producer variation.

The individual effects of factors $X1$ and $X2$ (producers and experts, respectively) can be evaluated using the $\widehat{C}_B$ decomposition $\widehat{C}_B = \widehat{C}_{X1}^B + \widehat{C}_{X2}^B$, where

$$\widehat{C}_{X1}^B = \frac{1}{(K-1)/4} \sum_{k=1}^{K-1} \frac{1}{I} \sum_{i=1}^{I} \left(\widehat{F}_{i.k} - \widehat{F}_{..k} \right)^2,$$

$$\widehat{C}_{X2}^B = \frac{1}{(K-1)/4} \sum_{k=1}^{K-1} \frac{1}{J} \sum_{j=1}^{J} \left(\widehat{F}_{.jk} - \widehat{F}_{..k} \right)^2. \quad (6)$$

Another $\widehat{C}_B$ decomposition, helpful for comparing the capability of the participating experts (as a group) to identify different categories, consists of evaluating the following parts of $\widehat{C}_B$ related to the responses of categories less than or equal to category k:

$$\widehat{C}_B \left(\sum_{l=1}^{k} Y_l \right) = \frac{1}{I} \sum_{i=1}^{I} \left(\widehat{F}_{i.k} - \widehat{F}_{..k} \right)^2 + \frac{1}{J} \sum_{j=1}^{J} \left(\widehat{F}_{.jk} - \widehat{F}_{..k} \right)^2. \quad (7)$$

For the first category $k=1$ of a sausage property by standard [25], the corresponding $\widehat{C}_B$ part is $\widehat{C}_B(Y_1)$ as a response of category less than 1 does not exist. For the second category, $\widehat{C}_B(Y_1 + Y_2)$ is the variation of the expert responses of categories $k=1$ and 2 (up to 2) because of the cumulative properties of the relative frequencies in (7). Therefore, the variation of responses of the participating experts for category $k=2$ is $\widehat{C}_B(Y_2) = \widehat{C}_B(Y_1 + Y_2) - \widehat{C}_B(Y_1)$. For the third category $k=3$, the variation is $\widehat{C}_B(\sum_{l=1}^{3} Y_l) = \widehat{C}_B(\sum_{l=1}^{3} Y_l) - \widehat{C}_B(\sum_{l=1}^{2} Y_l)$, and so on. A greater value of $\widehat{C}_B(\sum_{l=1}^{k} Y_l)$ indicates a weaker capability to identify category k. Note that the capability characterizing dispersion of the responses is analogous to measurement reproducibility [12]. When the cumulative relative frequencies in (7) achieve 1, the variation $\widehat{C}_B(\sum_{l=1}^{k} Y_l) = 0$. This is always the case for the last category K, but may also be for $(K-1)$ or even for $(K-2)$ [26], depending on the form of the cumulative distribution function.

The null hypothesis of homogeneity of the responses states that the probability of classifying the responses as belonging to the k-th category does not depend on the levels of the first factor (levels i) nor on those of the second factor (levels j), i.e., $p_{ijk} = p_k$ for all $i = 1, 2, ..., I$ and $j = 1, 2, ..., J$. Under this hypothesis, the following relations are applicable:

$$\frac{E\left(\widehat{V}_T\right)}{df_T} = \frac{E\left(\widehat{V}_W\right)}{df_W} = \frac{E\left(\widehat{C}_{X1}^B\right)}{df_{X1}} = \frac{E\left(\widehat{C}_{X2}^B\right)}{df_{X2}} = \frac{\{1/[(K-1)/4]\} \sum_{k=1}^{K-1} F_k \left(1 - F_k\right)}{N}, \quad (8)$$

where $E(\ldots)$ denotes the mathematical expectation of the variation in parentheses; $df_{X1} = I - 1$, $df_{X2} = J - 1$, $df_W = (I-1)(J-1)$, $df_T = N - 1$ are degrees of freedom. The numerator of the last term in equation (8) is equal to the population of the total ordinal variation corresponding to the probability vector $\mathbf{p} = (p_1, p_2, \ldots, p_K)$.

To check the statistical significance of both the factor effects, the following significance indices (test statistics) have been defined:

$$\widehat{SI}_{X1} = \frac{\widehat{C}_{X1}^B/df_{X1}}{\widehat{V}_T/df_T},$$

$$\widehat{SI}_{X2} = \frac{\widehat{C}_{X2}^B/df_{X2}}{\widehat{V}_T/df_T}. \quad (9)$$

Testing the null hypothesis H_0 on the effect significance requires knowledge of at least the asymptotical distribution of the index $\widehat{SI}$ for calculation of the critical values of the indices at a given level of confidence $(1 - \alpha) \cdot 100\%$.

A calculator tool for this purpose was proposed for the two-way ORDANOVA in reference [19]. The tool calculates from the empirical data the sample vector of relative frequencies $\widehat{\mathbf{p}} = (\widehat{p}_{..1}, \widehat{p}_{..2}, \ldots, \widehat{p}_{..K})$, as well as the variation components $(\widehat{C}_{X1}^B, \widehat{C}_{X2}^B, \widehat{V}_W, \widehat{V}_T)$ and the empirical significance indices $(\widehat{SI}_{X1}, \widehat{SI}_{X2})$. The critical values SI_{crit} for the indices in equation (9) are recovered through a Monte Carlo simulation based on at least 10000 trials. The null hypothesis H_0 is rejected when the significance index $\widehat{SI}$ exceeds the critical value SI_{crit} at the $(1 - \alpha) \cdot 100\%$ level of confidence, concluding that a statistically

significant effect on the response variable **Y** is detected. The calculator is freely available by following the link [27].

2.1.2. Multinomial Ordered Logistic Regression. Multinomial ordered logistic regression (ordered logit) is quite commonly applied in medicine [28], insurance, actuarial, and financial applications [29, 30], as well as to describe consumer purchasing behavior [31]. The ordered logit model can be considered as an extension of the logistic regression model for dichotomous dependent variables, when more than two ordinal response categories are used. It is applied here for analysis of ordinal quality sensory responses vs. chemical composition of a sausage. This model is based on the following concepts.

When **Y** is an ordinal outcome with K categories, the cumulative probability of responses of categories l less than or equal to a category k is $P(l \leq k)$. The odds of responses being less than or equal to a category k can be defined as $P(l \leq k)/P(l > k)$, for $k = 1,\ldots, K-1$. When $k = K$, the denominator is zero and the odds are not defined. The log odds, called logit, is defined as $\text{logit}(P(l \leq k)) = \log(P(l \leq k)/P(l > k))$. The ordinal logistic regression model is parameterized as follows:

$$\text{logit}(P(l \leq k)) = \beta_{k0} + \beta_1 c_1 + \cdots + \beta_m c_m, \quad (10)$$

where β_{k0} is the intercept (cutoff point) for category k; c_1 to c_m are the contents (mass fractions) of the main sausage components (continuous latent variables); β_1 to β_m are the corresponding regression coefficients (slopes), constant across categories. Note that this model is based on the parallel regression (proportional odds) assumption: the logit dependences on sausage compositions are parallel hyperplanes for different categories k, and hence, the intercepts are different for each category but the slopes are constant across categories. The odds of being less than or equal to category k are as follows:

$$P(l \leq k)/P(l > k) = \exp(\beta_{k0} + \beta_1 c_1 + \cdots + \beta_m c_m). \quad (11)$$

A computer program developed in the R environment for calculation of model parameters, including their confidence intervals and goodness-of-fit measures for the model, is described, for example, at the web page [32]. The following logit format is implemented in the R program as follows:

$$\text{logit}(P(l \leq k)) = \beta_{k0} - \eta_1 c_1 - \cdots - \eta_m c_m, \quad (12)$$

where $\eta = -\beta$ for all the regression coefficients of components' contents c_1 to c_m. Inverting equation (12), the probability of getting a response of a certain category k or below can be obtained [33] as follows:

$$P(l \leq k) = \text{logit}^{-1}(P(l \leq k)) = \frac{1}{[1 + 1/\exp(\beta_{k0} - \eta_1 c_1 - \cdots - \eta_m c_m)]}. \quad (13)$$

The "PoLR" function of the "MASS" package [34] was used to fit multinomial ordered logistic models to the experimental data, and function "predictorEffects" of the "effects" package [35] was used to calculate and plot probabilities of obtaining a response equal to a certain category k.

Goodness-of-fit of these models can be evaluated by calculating one of the several pseudo-R values [32], which estimates the variability in the outcome of the fitted model. For example, McFadden's pseudo-R^2 is defined as follows:

$$\text{Pseudo-}R^2 = 1 - \frac{\ln L(M_{\text{full}})}{\ln L(M_{\text{intercept}})}, \quad (14)$$

where M_{full} is a full model with predictors; $M_{\text{intercept}}$ is the model without predictors, i.e., containing the intercept only; and L is the estimated likelihood.

When the M_{full} model does not predict the outcome better than the $M_{\text{intercept}}$ model, its $\ln L(M_{\text{full}})$ is not much greater than $\ln L(M_{\text{intercept}})$; hence, the corresponding ratio is close to 1 and the McFadden's pseudo-R^2 is close to 0: the model has poor predictive value. Conversely, when the M_{full} model is good, its $\ln L(M_{\text{full}})$ is close to zero (since the likelihood value for each observation is close to 1), and McFadden's pseudo-R^2 is close to 1, indicating a successful predictive ability. If comparing two models on the same data, McFadden's pseudo-R^2 would be higher for the model with the greater likelihood.

The "PseudoR2" function of the R package "DescTools" [36] was applied for the corresponding calculations. Note that correlations between contents of the sausage main components may affect the regression coefficients and p values, but they do not influence the predictions, precision of predictions, or the goodness-of-fit statistics [22].

2.2. Experimental Methods. The comparative testing of a sausage as a consumer product [37] was organized by V.M. Gorbatov Federal Research Center for Food Systems, Russia. Samples of boiled-smoked sausage "Moscowskaya" [38] from $I = 16$ producers for comparison were purchased from a market in 2021, practically simultaneously. This sausage is prepared from beef and pork fat with addition of sodium nitrite curing salt (0.4–0.6 % mass fraction of $NaNO_2$ in $NaCl$) and spices. Its main chemical components are protein, fat, moisture, and salt.

All samples were examined before their expiration dates (set by the producers) by $J = 3$ assessors/experts [12] of the Research Center, women of 45–55 years old, each having more than 15 years' experience in sensory analysis of meat products. Examination of samples was performed in a test room [39] with individual testing cubicles.

Five quality sensory properties of the samples were assessed as follows: (1) appearance and packaging, named here "appearance"; (2) consistency; (3) color and appearance of cut

sausage, named here "color"; (4) taste; and (5) smell. An expert response related to each quality property was ordered by $K = 5$ categories from "very bad" to "excellent" ($k = 1, 2, \ldots, 5$). A total number $N = I \times J = 48$ responses was obtained for each property, and $48 \times 5 = 240$ responses for the five properties.

Contents (mass fractions) of the $m = 4$ main components were taken from the certificates of the producers. In total, $I \times m = 64$ continuous quantitative values were obtained.

The data set, including both qualitative and quantitative data (RawData Microsoft Excel file) is available at Zenodo [40].

2.2.1. Methods of Sensory Examination.
The experts who participated in the examination of the sausage samples were not informed about quantitative characteristics of the samples' chemical composition. The quality parameters of a sample, which was one whole packaged sausage, were examined by a standard protocol [25], in the following sequence: (a) appearance–by visual external examination of the packaged sausage; (b) consistency–by pressing with a spatula or fingers on the outside of the sausage, and then, after removing the casing from the sausage and cutting with a sharp knife into thin slices perpendicular to the surface of the product; (c) color–visually; (d) and (e) taste and smell–by pressing and chewing a slice.

2.2.2. Methods of Testing Chemical Composition.
The standard Kjeldahl method was used for measurement of protein content c_1 [41]. The standard method used for measurement of fat content c_2 [42] is based on multiple extractions from a dried sample with a solvent (hexane, diethyl ether, or petroleum ether) in a Soxhlet fat-extraction apparatus. Then, the solvent is removed and the fat dried to constant weight. The standard measurement method for moisture content c_3 [43] consists of drying a sample with sand to constant weight at a temperature of $(103 \pm 2)°C$. Salt content c_4 was measured by Mohr's standard titration method [44]. Measurement results are expressed as percent mass fractions. More details including estimation of measurement uncertainties in testing a sausage composition, its specification limits and conformity assessment, are described in the paper [45].

3. Results and Discussion

A flowchart of the data treatment is presented in Figure 1. ORDANOVA starts from calculation of frequencies of expert responses (of different categories) and evaluation of the total variation. The next step is decomposition of the total variation into components with the purpose to assess the effects of two factors influencing the variation: "producer" and "expert." Another kind of decomposition allows assessment of the abilities of the experts to determine different categories of the same quality property. The components of variation obtained are used for testing hypotheses on the homogeneity of the producers (i.e., the responses to their sausage quality properties) and homogeneity of the experts (their responses to a property of the same sausage sample). When responses of different experts are inhomogeneous, and/or the responses to different sausage

producers are homogeneous, the analysis is ended. Otherwise, it is assumed that the difference between responses to the sausage quality of different producers is caused by the differences in the sausage chemical composition, expressed as mass fractions of the main components. This hypothesis is tested with multinomial ordered regression analysis. If any of the regression coefficients are statistically significant (the hypothesis is not rejected), probabilities of obtaining a response related to a specific category for different component contents are calculated. The last step is plotting such probabilities for visualization of the results and their discussion.

3.1. Implementation of Two-Way ORDANOVA without Replication.
Frequencies of the responses from the data set are shown in Table 1 by categories (rows) and experts for each quality property of the sausage (columns). All purchased sausage samples were of higher quality than category 2 for any property. Two out of three experts rated the appearance of the whole sausage samples greater than category 3. Also, all experts reported that the sausage consistency was greater than category 3.

Total variation $\widehat{V}_T$ of the responses by equation (3), partitioned into the between-producer variation $\widehat{C}_B$ by equation (4) and the within-producer residual variation $\widehat{V}_W$ by equation (5) are presented in Table 2, which includes the individual effects $\widehat{C}_{X1}^B$ and $\widehat{C}_{X2}^B$ of factors $X1$ and $X2$ (producers and experts, respectively) evaluated using the $\widehat{C}_B$ decomposition by equation (6). To check the statistical significance of both the factor effects the significance indices $\widehat{SI}_{X1}$ and $\widehat{SI}_{X2}$ were also calculated by equation (9) at the degrees of freedom $df_{X1} = 15$ and $df_{X2} = 2$. The critical index values SI_{crit} at the 95% level of confidence in Table 2 were obtained using simulations of $\widehat{SI}$ distributions [27].

There is a statistically significant difference at 95 % level of confidence between the producers related to all the quality parameters of the sausage (appearance, consistency, color, taste, and smell). This difference is called the "inhomogeneity" of the producers. At the same time, the significance index values of the expert factor do not exceed its critical value at the 95 % level of confidence, i.e., the hypotheses on the "homogeneity" of expert responses with regard to each of the five sausage properties are not rejected.

Decomposition of the between producers' variation $\widehat{C}_B$ into kth parts $\widehat{C}_B(\sum_{l=1}^{k} Y_l)$ according to the response categories up to k by equation (7) and corresponding $\widehat{C}_B(\sum_{l=1}^{k-1} Y_l)$ values are shown in Table 3. Comparing the capability of the participating experts (as a group) to identify different categories, it is important to remember that a greater value of variation $\widehat{C}_B(\sum_{l=1}^{k} Y_l)$ indicates a weaker capability to identify the particular category k. This decomposition demonstrates that identification of very bad and poor sausage quality ($k = 1$ and 2, respectively) was the simplest task for the experts: $\widehat{C}_B(Y_1) = 0$ and $\widehat{C}_B(Y_1 + Y_2) = 0$ as no expert responses of these categories were obtained for any property. The most complicated part of their examination was to identify a difference between satisfactory and good quality ($k = 3$ and 4, respectively). The variation up to the last category $\widehat{C}_B(\sum_{k=1}^{5} Y_k)$ equals zero by definition of
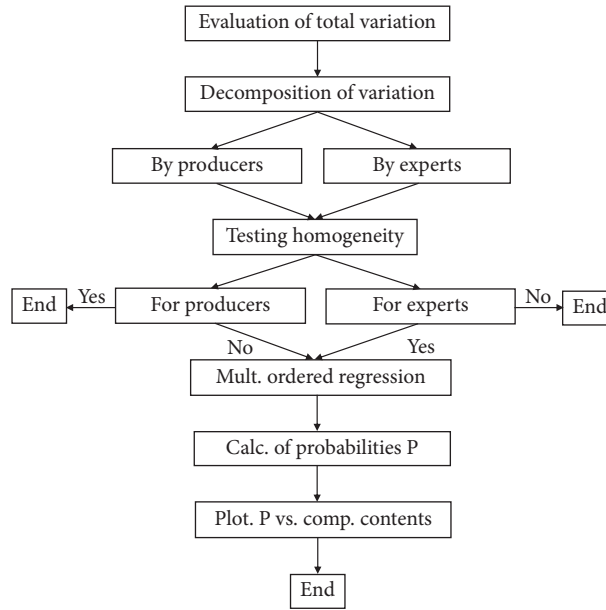
Figure 1: Flowchart of the data treatment.

Table 1: Frequencies of the responses on the quality properties of the sausage.

Category, k	Frequency, n_k														
	Appearance			Consistency			Color			Taste			Smell		
	Experts, j														
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	3	0	0	0	0	0	3	1	2	3	3	4	3	2	3
4	4	3	7	3	3	6	7	4	3	6	4	5	6	5	5
5	9	13	9	13	13	10	6	11	11	7	9	7	7	9	8

Table 2: Results of two-way ORDANOVA without replication for the sensory responses on the property of sausage.

Property	$\widehat{V}_T$	$\widehat{C}_B$	$\widehat{V}_W$	X1 and X2	$\widehat{C}_{X1}^B$ & $\widehat{C}_{X2}^B$	$\widehat{SI}_{X1}$ & $\widehat{SI}_{X2}$	df_{X1} & df_{X2}	SI_{crit}
Appearance	0.287	0.184	0.103	Producer	0.162	1.769	15	1.432
				Expert	0.022	1.801	2	2.686
Consistency	0.187	0.112	0.075	Producer	0.104	1.740	15	1.550
				Expert	0.008	0.980	2	2.892
Color	0.352	0.251	0.101	Producer	0.227	2.019	15	1.422
				Expert	0.024	1.621	2	2.554
Taste	0.414	0.293	0.121	Producer	0.289	2.186	15	1.403
				Expert	0.004	0.246	2	2.521
Smell	0.389	0.295	0.094	Producer	0.292	2.351	15	1.408
				Expert	0.004	0.210	2	2.510

equation (7). However, the number of responses of the category of perfect quality ($k = 5$) was maximum for each property shown in Table 1, and detection of this category was not a problem for the experts. It is also interesting that the maximum controversial responses here are on the sausage taste and especially on smell. Of course, the results shown in Table 3 do not mean in general that the examined ordinal variables have collapsed into a binary group.

3.2. Implementation of the Multinomial Ordered Logistic Regression. Intervals of the sausage main component contents in the data set, their means, and standard deviations are shown in Table 4.

The multinomial ordered logistic regression model by equation (12) was fitted to the component contents in order to predict appearance, color, taste, and smell, assessed by experts according to the three categories shown in Table 1,

Table 3: Decomposition of the between producers' variation component according to the response categories.

Property	$\widehat{C}_B (Y = 1)$	$\widehat{C}_B (Y = 2)$	$\widehat{C}_B (Y = 3)$	$\widehat{C}_B (Y_3 + Y_4)$	$\widehat{C}_B (Y = 4)$
Appearance	0	0	0.025	0.159	0.134
Consistency	0	0	0	0.112	0.112
Color	0	0	0.070	0.181	0.111
Taste	0	0	0.096	0.197	0.101
Smell	0	0	0.070	0.225	0.155

Table 4: Statistics of the chemical composition of the sausage samples.

Statistics	Mass fractions (%)			
	Protein	Fat	Moisture	Salt
Minimum	13.7	19.9	53.5	2.22
Maximum	19.5	26.4	59.5	3.53
Mean	15.79	23.00	56.01	2.55
Standard deviation	1.42	4.60	1.92	0.29

Table 5: Results of the multinomial ordered logistic regression analysis.

Property	Coefficient	Value	Standard error	2.5%	97.5%	Odds ratio	Pseudo-R^2
Appearance	$\beta_{k0}(3\|4)$	108.31	0.01	108.30	108.32	1.09×10^{47}	
	$\beta_{k0}(4\|5)$	110.71	0.61	109.51	111.91	1.21×10^{48}	
	η_1	1.65	0.30	1.06	2.24	5.20	0.13
	η_2	0.95	0.10	0.76	1.14	2.58	
	η_3	1.24	0.07	1.10	1.38	3.44	
	η_4	−2.22	1.31	−4.78	0.34	0.11	
Consistency	β_{k0}	−113.55	66.07	−243.05	15.95	4.85×10^{-50}	
	η_1	1.35	0.69	0.00	2.71	3.87	0.11
	η_2	1.08	0.56	−0.01	2.18	2.95	
	η_3	1.22	0.78	−0.30	2.74	3.40	
Color	$\beta_{k0}(3\|4)$	21.76	0.01	21.75	21.78	2.83×10^9	
	$\beta_{k0}(4\|5)$	23.56	0.44	22.69	24.42	1.70×10^{10}	
	η_1	0.56	0.25	0.07	1.05	1.75	0.09
	η_2	−0.03	0.10	−0.23	0.17	0.97	
	η_3	0.31	0.06	0.18	0.43	1.36	
	η_4	−0.51	1.28	−3.02	1.99	0.60	
Color*	$\beta_{k0}(3\|4)$	26.80	11.85	3.57	50.02	4.34×10^{11}	
	$\beta_{k0}(4\|5)$	28.59	11.93	5.21	51.97	2.60×10^{12}	
	η_1	0.56	0.27	0.07	1.05	1.75	0.09
	η_3	0.36	0.18	0.18	0.43	1.43	
Taste	$\beta_{k0}(3\|4)$	202.04	0.00	202.03	202.04	5.55×10^{87}	
	$\beta_{k0}(4\|5)$	204.47	0.55	203.39	205.55	6.31×10^{88}	
	η_1	2.87	0.30	2.28	3.46	17.61	0.28
	η_2	1.73	0.09	1.54	1.91	5.62	
	η_3	2.33	0.07	2.19	2.47	10.27	
	η_4	−4.42	1.51	−7.37	−1.47	0.01	
Smell	$\beta_{k0}(3\|4)$	218.42	0.00	218.41	218.42	7.19×10^{94}	
	$\beta_{k0}(4\|5)$	221.26	0.62	220.04	222.47	1.23×10^{96}	
	η_1	3.25	0.35	2.58	3.93	25.92	0.32
	η_2	1.99	0.10	1.80	2.18	7.33	
	η_3	2.41	0.08	2.26	2.57	11.18	
	η_4	−4.42	1.54	−7.43	−1.40	0.01	

*A shorted model of probability of the color category vs. contents of protein and moisture.

$k = 3$, 4, and 5. A logistic regression for dichotomous (binary) outcome variables was used for prediction of consistency, since the corresponding expert responses (Table 1) were only of two categories, $k = 4$ and 5. Since for each categorical variable, the responses were found to be homogeneous among the three experts in the ORDANOVA study, all their outcomes were taken together, constituting the set of values to be used in the regression. The calculated results are presented in Table 5, where the estimates for coefficients β_{k0} and η related to each component content are
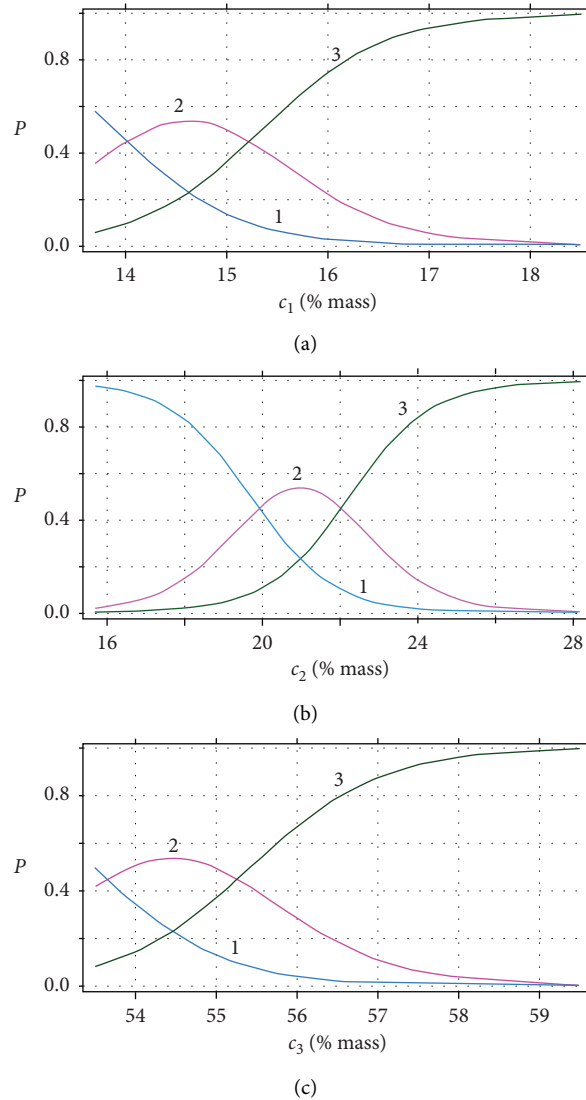
(a)

(b)

(c)

FIGURE 2: Probabilities P of responses of different appearance categories in dependence on content of (a) protein c_1, (b) fat c_2, and (c) moisture c_3, mass fractions, %. Each plot is calculated at contents of other main components equal to their observed mean values (Table 4). Line 1 is for the category "satisfactory," line 2 for the category "good," and line 3 for the category "excellent."

reported with their standard errors and 95 % confidence intervals (from the 2.5 % to the 97.5 % quantile). The η subscripts from 1 to 4 correspond to the subscripts of the component contents c.

The estimated odds ratios, derived by exponentiating the coefficients, and McFadden's pseudo-R^2 values by equation (14) are also shown in the table. For example, the model by equation (12) of category $k = 3$ for appearance is logit $(P(l \leq 3))$ $= 108.31 - 1.65\ c_1 - 0.95\ c_2 - 1.24\ c_3 + 2.22\ c_4$. A one-unit increase in the protein content c_1, for example, corresponds to the increase in the expected value of the logit $(P(l \leq 3))$ by 1.65 (on the log odds scale), when all the other variables in the model are held constant. The corresponding odds ratio $\exp(1.65) = 5.2$ indicates that, for every unit increase in the protein content, the odds of a sausage of having a better appearance outcome ($k = 4$ or 5, versus $k = 3$) is multiplied 5.2 times.

Note if a confidence interval does not cross zero, the parameter estimate is statistically significant. However, the

confidence interval for the estimate $\eta_4 = 2.22$ (of the regression coefficient of the salt content c_4) crosses zero and this means that η_4 is statistically not significant in this case. In other words, the salt content values in the interval shown in Table 5 do not influence the probability of the category appearance of a whole sausage.

Probabilities P of obtaining a response of category k by dependence on protein content c_1, calculated at the mean values of contents of other main components (Table 4), are shown in Figure 2(a). In the case of three categories of the observed responses ($k = 3$, 4 and 5), the probability $P(l = 3)$ $= P(l \leq 3)$, $P(l = 4) = P(l \leq 4) - P(l = 3)$, and $P(l = 5)$ $= 1 - P(l = 3) - P(l = 4)$, where $P(l \leq 3)$ and $P(l \leq 4)$ can be evaluated by equation (13). The P values for $k = 3$, 4, and 5 are shown by lines 1, 2, and 3 in Figure 2(a), respectively. The P dependences on fat content c_2 and on moisture content c_3 are shown in Figures 2(b) and 2(c), respectively. They were also calculated at contents of other main components equal
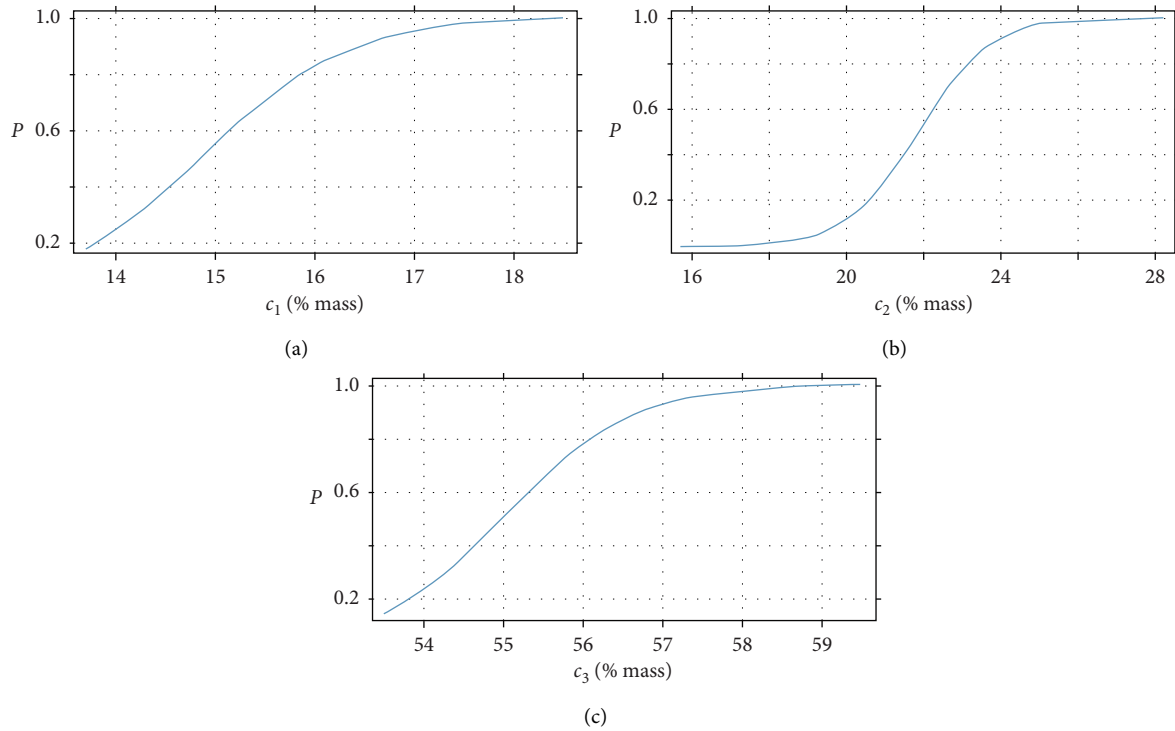
FIGURE 3: Probabilities $P(l = 5)$ of responses on consistency in dependence on content of (a) protein c_1, (b) fat c_2, and (c) moisture c_3, mass fractions, %. Each plot is calculated at contents of other main components equal to their observed mean values in Table 4.
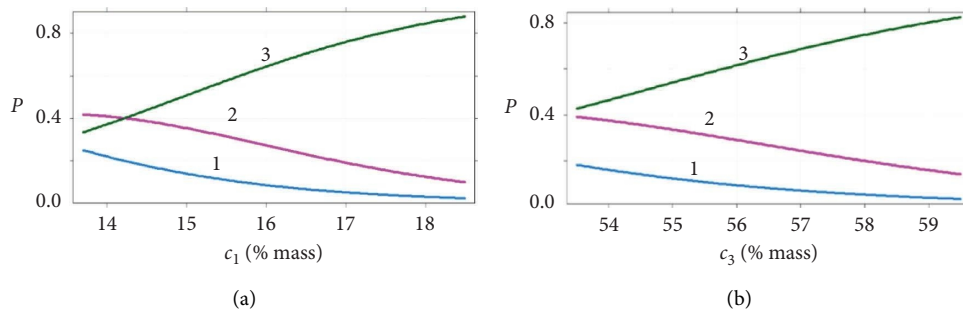

FIGURE 4: Probabilities P of responses of different color categories in dependence on content of (a) protein c_1 and (b) moisture c_3, mass fractions, %. Each plot is calculated at contents of other main components equal to their observed mean values in Table 4. The line numbers are as in Figure 2.

to their observed mean values in Table 4. The influence of all the component contents on the probability values is very similar, but the probability curves vs. contents of protein and moisture are cut at the lower limits of the content intervals (their minimum values observed in the comparison). The full picture is shown in Figure 2(b) for the probabilities vs. contents of fat, where the probability values of the appearance categories vary from zero to the maximum, or from 1 to 0 and vice versa. Note that increasing c_1, c_2, and c_3 leads to an increasing probability of the responses of the highest appearance category $k = 5$ (excellent quality).

The salt content c_4 does not influence probabilities of responses of consistency dichotomous categories $k = 4$ and 5; hence, it was removed from the list of regressors. The confidence interval of the intercept β_{k0} in the consistency

model in Table 5 crosses zero. However, the intercept closeness to zero does not reflect influence of a component content on the probabilities of consistency categories. Since responses of two categories were obtained, the probability of one of them is the complement to the other, i.e., $P(l = 5) = 1 - P(l = 4)$ and vice versa. Therefore, $P(l = 5)$ only is shown in Figure 3 to illustrate the probability dependence on component contents.

The probabilities of responses of different color categories do not depend on the contents of fat and salt, c_2 and c_4, in their observed intervals. Model "Color*" without these two variables has the same McFadden's pseudo-R^2 value 0.09 as full model "Color," and practically, the same regression coefficients for c_1 and c_3 in Table 5. Probabilities P of responses of different color categories in dependence on
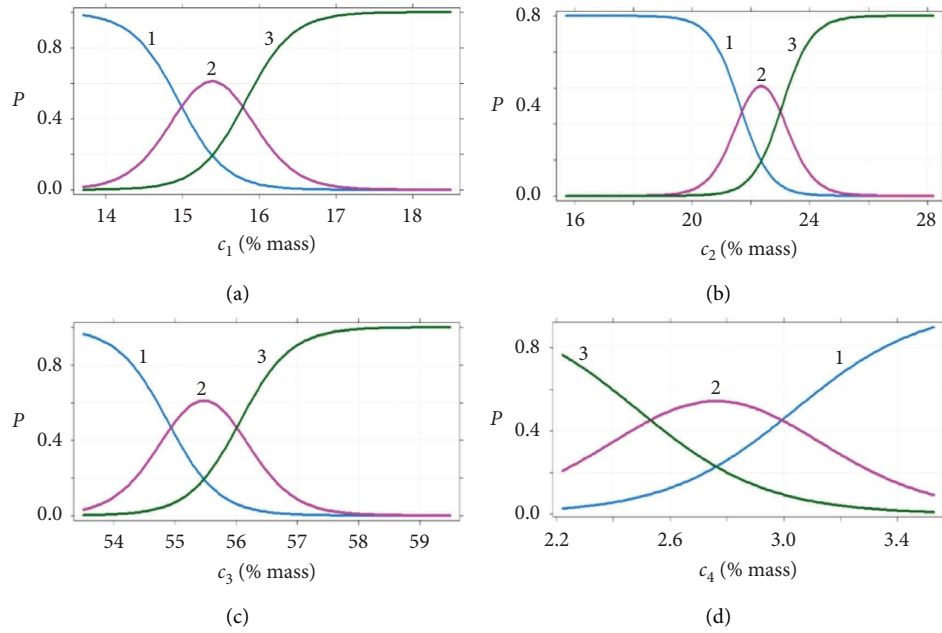
FIGURE 5: Probabilities P of responses of different taste categories in dependence on content of (a) protein c_1, (b) fat c_2, (c) moisture c_3, and (d) salt c_4, mass fractions, %. Each plot is calculated at contents of other main components equal to their observed mean values in Table 4. The line numbers are as in Figure 2.
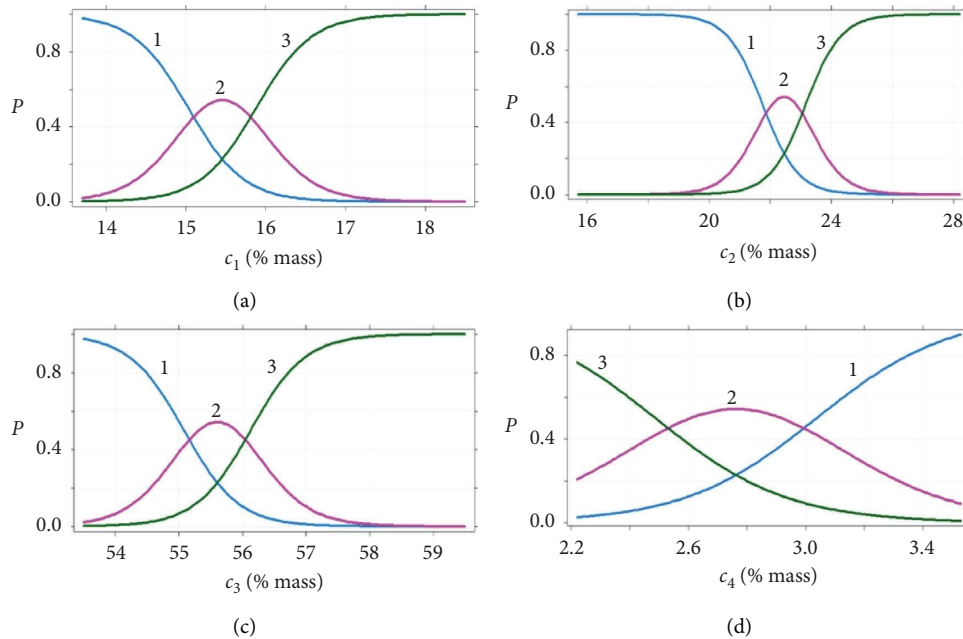

FIGURE 6: Probabilities P of responses of different smell categories in dependence on content of (a) protein c_1, (b) fat c_2, (c) moisture c_3, and (d) salt c_4, mass fractions, %. Each plot is calculated at contents of other main components equal to their observed mean values in Table 4. The line numbers are as in Figure 2.

content of protein c_1 are shown in Figure 4(a), and on moisture content c_3 in Figure 4(b).

The full models for taste and smell are the best fitting models among the qualitative sausage properties; their McFadden's pseudo-R^2 values in Table 5 are about two to three times greater than those for appearance, color, and consistency. Dependences of probabilities P of responses of

different taste and smell categories on the contents of main components, calculated at contents of other main components equal to their observed mean values, are shown on Figures 5 and 6, respectively.

In general, the maximum probability of responses of each category of taste and smell is reached at increasing contents of the influencing main components. Similar

effects are also observed in the plots in Figure 2 for appearance as follows: the first category reaching its maximum probability in the studied ranges of the component contents is 3, then 4, and finally 5, i.e., higher categories are more probable with greater contents of components. However, the salt contents in the interval considered in this study do not significantly influence responses on appearance, color, and consistency. At the same time, taste (Figure 5) and smell (Figure 6) are influenced by the salt contents in a reverse order than contents of other main components; the greater the salt content, the lower category is the more probable.

The probabilities of responses of the excellent quality category $P(l = 5)$ of both, taste and smell, increase with contents of protein up to c_1 of about 17 %; fat, c_2 of about 26 %; and moisture, c_3 of about 58 %, while the minimal salt content $c_4 = 2.2$ % is preferable. These estimates could be helpful for a revision of the specification limits of the sausage composition, necessarily taking into account the mass balance constraint; the sum of actual values of mass fractions of the four main components should be equal to 100 % or to a positive fraction less than 100 % [45].

4. Conclusions

A data set of ordinal responses of three experienced experts who assessed five quality properties of samples of a boiled, smoked sausage from sixteen producers was analyzed. The responses were ordered using five categories. Implementation of ORDANOVA allowed decomposition of total variation of the ordinal data and simulation of the multinomial distribution of the relative frequencies of the responses in different categories. A statistically significant difference in quality properties of the sausages from different producers was detected, while the difference between responses of the experts was insignificant.

Capabilities of the experts to identify different categories of the quality properties were also evaluated. It was shown that identification of "very bad" and "poor" quality, as well as "perfect" quality is the simplest task for the experts. The most complicated part of their examination was to identify a difference between "satisfactory" and "good" quality–the closest categories.

Influence of chemical composition of a sausage sample on the probability of a response category was modeled using the multinomial ordered logistic regression of the response category on mass fractions of four main sausage components. Obtained estimates could be helpful for a revision of the specification limits of the sausage composition, as well as for prediction of the product sensory properties when its chemical composition is under quality control.

Data Availability

All relevant data are included in the paper or the data set is available at Zenodo [40]. https://doi.org/10.5281/zenodo.7213008.

Conflicts of Interest

The authors declare that they have no conflicts of interest to disclose.

Acknowledgments

References

[1] ISO 4121, *Sensory Analysis. General Guidelines for the Use of Quantitative Response Scales*, International Organization for Standardization, Switzerland, 2003.

[2] ISO 8586, *Sensory Analysis. General Guidelines for the Selection, Training and Monitoring of the Selected Assessors and Expert Sensory Assessors*, International Organization for Standardization, Switzerland, 2014.

[3] C. Bord, L. Lenoir, R. Schmidt-Filgueras, J. Benoit, G. Dechambre, and C. Chassard, "Discrimination and sensory characterization of protected designation of origin Salers- and Cantal-type cheeses: an approach using descriptive analysis and consumer insights by check-all-that-apply questions," *Journal of Sensory Studies*, vol. 36, no. 6, Article ID e12698, 2021.

[4] S. Pati, T. Sarkar, H. I. Sheikh et al., "γ-Irradiated Chitosan from Carcinoscorpius rotundicauda (Latreille, 1802) Improves the shelf life of refrigerated aquatic products," *Frontiers in Marine Science*, vol. 8, Article ID 664961, 2021.

[5] D. B. Hibbert, "Chemometric analysis of sensory data," in *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, S. Brown, R. Tauler, and B. Walczak, Eds., Elsevier, Amsterdam, pp. 149–192, 2019.

[6] T. Sarkar, S. K. Saha, M. Salauddin, and R. Chakraborty, "Drying kinetics, Fourier-transform infrared spectroscopy analysis and sensory evaluation of sun, hot-air, microwave and freeze dried mango leather," *Journal of Microbiology, Biotechnology and Food Sciences*, vol. 10, no. 5, Article ID e3313, 2021.

[7] T. Sarkar, M. Salauddin, S. Pati et al., "The fuzzy cognitive map–based shelf-life modelling for food storage," *Food Analytical Methods*, vol. 15, no. 3, pp. 579–597, 2022.

[8] T. Sarkar, M. Salauddin, S. Pati et al., "Expert knowledge-–based system for shelf-life analysis of dairy cheese ball (Rasgulla)," *Food Analytical Methods*, vol. 15, no. 7, pp. 1945–1960, 2022.

[9] S. Ventanas, A. Gonzalez-Mohino, M. Estevez, and L. Carvalho, "Innovation in sensory assessment of meat and meat products," in *Meat Quality Analysis: Advanced Evaluation Methods, Techniques, and Technologies*, A. K. Biswas and P. K. Mandal, Eds., Academic Press – Elsevier, Amsterdam, pp. 393–418, 2020.

[10] C. Ruiz-Capillas, A. M. Herrero, T. Pintado, and G. Delgado-Pando, "Sensory analysis and consumer research in new meat products development," *Foods*, vol. 10, no. 2, Article ID 10020429, 2021.

[11] JCGM 200, *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM 3)*, 2012,http://www.bipm.org/en/publications/guides/.

[12] D. B. Hibbert, E.-H. Korte, and U. Örnemark, "Metrological and quality concepts in analytical chemistry (IUPAC Recommendations 2021)," *Pure and Applied Chemistry*, vol. 93, no. 9, pp. 997–1048, 2021.

[13] S. S. Stevens, "On the theory of scales of measurement," *Science*, vol. 103, no. 2684, pp. 677–680, 1946.

[14] H. Scheffé, *Analysis of Variance*, Willey Classics Library, New York, 1999.

[15] M. d. l. A. Vivar-Vera, A. Pérez-Silva, I. I. Ruiz-López et al., "Chemical, physical and sensory properties of Vienna sausages formulated with a starfruit dietary fiber concentrate," *Journal of Food Science & Technology*, vol. 55, no. 8, pp. 3303–3313, 2018.

[16] I. Brankovic Lazic, J. Jovanovic, S. Simunovic et al., "Evaluation of sensory and chemical parameters of fermented sausages," *Meat Technology*, vol. 60, no. 2, pp. 84–90, 2019.

[17] P. Kloucek and D. Bures, "Special issue sensory evaluation and quality analysis of meat and meat products," *Foods*, vol. 20, 2022.

[18] L. D. Schroeder, D. L. Sjoquist, and P. E. Stephan, *Understanding Regression Analysis: An Introductory Guide*, SAGE Publications, Los Angeles, 2nd ed edition, 2017.

[19] T. Gadrich and Y. N. Marmor, "Two-way ORDANOVA: analyzing ordinal variation in a cross-balanced design," *Journal of Statistical Planning and Inference*, vol. 215, pp. 330–343, 2021.

[20] P. McCullagh, "Regression models for ordinal data," *Journal of the Royal Statistical Society: Series B*, vol. 42, no. 2, pp. 109–127, 1980.

[21] J. A. Anderson and P. R. Philips, "Regression, discrimination and measurement models for ordered categorical variables," *Applied Statistics*, vol. 30, no. 1, pp. 22–31, 1981.

[22] J. M. Hilbe, *Logistic Regression Models*, Chapman & Hall/CRC Press, New York, 2009.

[23] NIST/Sematech, *E-Handbook of Statistical Methods*, New York China, 2021, https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/multpdf.htm.

[24] ISO/IEC 17043, *Conformity Assessment – General Requirements for Proficiency Testing*, International Organization for Standardization, Switzerland, 2010.

[25] GOST 9959, "Meat and meat products," *General Conditions of Organoleptic Assessment. Standardinform*, Russian Federation, 2015, https://gostperevod.com/catalogsearch/result/?q=gost+9959-2015.

[26] T. Gadrich, I. Kuselman, F. R. Pennecchi et al., "Interlaboratory comparison of the intensity of drinking water odor and taste by two-way ordinal analysis of variation without replication," *Journal of Water and Health*, vol. 20, no. 6, pp. 1005–1016, 2022.

[27] Y. N. Marmor and T. Gadrich, *Two-way ORDANOVA Tool. V1*, https://docs.google.com/spreadsheets/d/1fSF8ZpZf0pNRRm9l_dh4fN14nyLDROms/edit?usp=sharing&ouid=101198859611382942102&rtpof=true&sd=true, 2019.

[28] T. J. Hastie, J. L. Botha, and C. M. Schnitzler, "Regression with an ordered categorical response," *Statistics in Medicine*, vol. 8, no. 7, pp. 785–794, 1989.

[29] E. W. Frees, *Regression Modelling with Actuarial and Financial Applications*, Cambridge University Press, New York, 2010.

[30] M. Guillén, "Regression with categorical dependent variables," 2022, http://www.ub.edu/rfa/R/regression_with_categorical_dependent_variables.html.

[31] F. Kaygisiz, B. A. Bolat, and D. Bulut, "Determining factors affecting consumer's decision to purchase organic chicken meat," *Brazilian Journal of Poultry Science*, vol. 21, no. 4, pp. 1–8, 2019.

[32] UCLA, *Advanced Research Computing. Statistical Methods and Data Analysis*, https://stats.oarc.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/, 2022.

[33] M. K. W. Brown, "Evaluating an ordinal output using data modeling, algorithmic modeling, and numerical analysis," *Murray State Theses and Dissertations*, vol. 168, 2020.

[34] W. N. Venables and B. D. Ripley, *MASS: Modern Applied Statistics with S*, Springer, New York, 2002, http://www.stats.ox.ac.uk/pub/MASS4, 4th ed edition.

[35] J. Fox and S. Weisberg, *An R Companion to Applied Regression*, Sage, Los Angeles, 3rd ed edition, 2019.

[36] A. Signorell and R. Chen, *DescTools: Tools for Descriptive Statistics*, https://cran.r-project.org/package=DescTools, 2021.

[37] ISO/IEC Guide 46, *Comparative Testing of Cosumer Products and Relative Services – General Principles*, International Organization for Standardization, Switzerland, 2017.

[38] GOST R 55455, *Boiled-smoked Meat Sausages*, Specifications. Standardinform, Russian Federation, 2013, https://gostperevod.com/catalogsearch/result/?q=gost+55455-2013.

[39] ISO 8589, *Sensory Analysis. General Guidance for the Design of Test Rooms*, International Organization for Standardization, Switzerland, 2007.

[40] F. Pennecchi, T. Gadrich, I. Kuselman, B. Hibbert, A. A. Semenova, and P. S. Cheow, "Ordinal Analysis of a Sausage Quality - Dataset," *Zenodo*, vol. 1, 2022.

[41] GOST 25011, *Meat and Meat Products – Protein Determination Methods*, Standardinform, Russian Federation, 2017, https://gostperevod.com/catalogsearch/result/?q=gost+25011-2017.

[42] GOST 23042, *Meat and Meat Products – Methods of Fat Determination*, Standardinform, Russian Federation, 2015, https://gostperevod.com/catalogsearch/result/?q=GOST+23042-2015.

[43] GOST 9793, "Meat and meat products," *Methods for Determination of Moisture Content*, Standardinform, Russian Federation, 2016, https://gostperevod.com/catalogsearch/result/?q=GOST+9793-2016.

[44] GOST 9957, *Meat and Meat Products - Methods for Determination of Sodium Chloride*, Standardinform, Russian Federation, 2015, https://gostperevod.com/catalogsearch/result/?q=GOST+9957-2016.

[45] F. R. Pennecchi, I. Kuselman, A. Di Rocco, D. B. Hibbert, and A. A. Semenova, "Risks in a sausage conformity assessment due to measurement uncertainty, correlation and mass balance constraint," *Food Control*, vol. 125, Article ID 107949, 2021.