Statistical testing of individual differences in sensory profiling

Per Bruun Brockhoff

Department of Mathematics and Physics, The Royal Veterinary and Agricultural University, ThorvaldSENSVej 40, 1871 Frederiksberg C, Denmark

Short summary
A generalization of the approach of Brockhoff and Skovgaard (1994) is given. The emphasis is on univariate assessor performance in sensory profiling. Statistical significance tests for difference between assessors of scaling, variability and discriminatory ability will be given. A test for disagreement effect is also presented. In addition, the approach will provide individual scaling, variability, disagreement and discriminatory ability values, that can be used for subsequent tabulation, plotting and statistical analysis. The method of maximum likelihood is used throughout and all computations are implemented in a SAS Macro PANMODEL that is available via the author's homepage: http://www.dina.kvl.dk/~per.

Introduction and Methods
The objective is reached by fitting a number of models around the ANOVA framework to the data. Formal statistical modeling of a scaling effect and variance heterogeneity leads to heterogeneous multiplicative interaction models, that are fitted to the data by alternating least squares algorithms. The approach does not provide tests for product differences, but can be seen as supplementary to the mixed model ANOVA approach one would typically take to obtain this (primary) information from the data set. Viewpoints on how to perform a mixed model ANOVA properly in the sensory context are given elsewhere. As pointed out, it is not the issue of the present paper. A formal combined approach would be preferable but has not yet been developed for sensory data. A step in that direction will be taken in an invited paper at the present meeting, Cullis (2002). A statistical test for differences in discriminatory ability among panel members is given, a new development in this context.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Variability</th>
<th>Disagreement</th>
<th>Scaling</th>
<th>Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pea odour</td>
<td>***</td>
<td></td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Pod odour</td>
<td>**</td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Sweet odour</td>
<td>***</td>
<td></td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Earthy odour</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Pea taste</td>
<td>***</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pod taste</td>
<td>**</td>
<td>*</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Sweet taste</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Bitterness</td>
<td>***</td>
<td></td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>Earthy taste</td>
<td>***</td>
<td>*</td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Crispiness</td>
<td>***</td>
<td>**</td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Juiciness</td>
<td>***</td>
<td></td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Hardness</td>
<td>***</td>
<td></td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Mealiness</td>
<td>***</td>
<td></td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Skin Viscosity</td>
<td>**</td>
<td></td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

Table 1: Summary of statistical tests of individual differences and disagreements.
***=significant on 0.1% level, **=significant on 1% level, *=significant on 5% level.
Results
The significance test results can be summarized as in Table 1. For this data on frozen pea products, individual differences in scaling, variability and discriminatory ability are significant for almost all the attributes. For only five attributes the panel exhibits significant disagreement. As output of the approach comes four assessor-by-attribute matrices - one for each of the four effect types. Principal Component Analysis was performed on the discriminatory abilities and on the disagreement contributions, see Figure 1. Overall, the most extreme assessors with respect to discriminatory ability are assessors 1, 9, 7 and 8: Assessor 9 is particularly sensitive to textural attributes like crispiness, juiciness and hardness, whereas assessor 1 and 7 are sensitive to the earthy taste and odor with assessor 8 being particularly insensitive in that direction. In the disagreement plot the first component expresses an overall disagreement behavior: Assessors to the right (no 6) disagree the most with the panel as a whole, whereas those to the left (no 4,8,11) are the mostly agreeing panel members.

Figure 1: Biplots of assessor-by-attribute matrices of individual discriminatory abilities (left) and disagreement contributions (right) including only those attributes that showed significance, cf. Table 1. Numbers refer to assessors.

References