



# Comparison of ANOVA with the Tobit-Model for Analysing Sensory Data

**M. Marin-Galiano**

University of Dortmund

**J. Kunert**

University of Dortmund



# Outlook

- Introduction
- Problems with ANOVA
- Alternative: The Tobit-Model
- Results from Previous Experiments
- Problems with the Tobit-Model
- Permutation Study
- Results
- Conclusions



# Introduction

- Data base: „Dijon data set“ (Danone 2001) as used at the Pangborn conference in Dijon, workshop „Data Sensory Analysis“.
- Sensory experiment:
  - 13 trained judges
  - 23 products (joghurts with strawberry flavor)
  - 24 sensory variables
  - judges mark intensity on an unstructured scale from 0 (no intensity) to 9 (maximum intensity)
  - order of products randomized for each judge
- Data characteristics:
  - high percentage of zeros
  - unbalanced data set (→ we used first replication only)



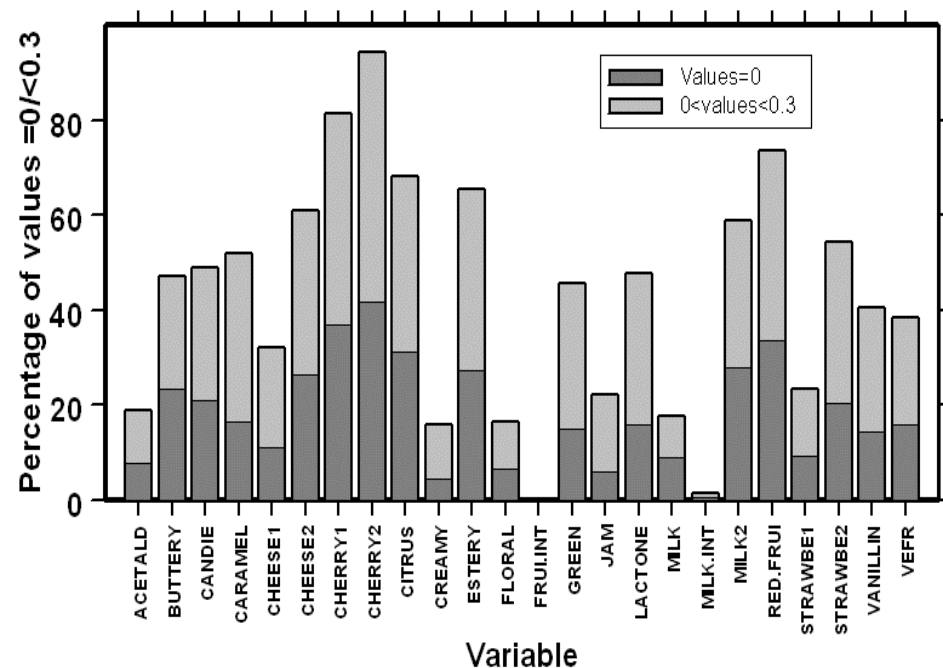
# ANOVA

- Standard technique for analyzing block models.
- Testing, whether treatment effects (products) have an impact on the dependent variable.
- Distribution of sum of squares:  
total SS = block SS + treatment SS + error SS
- Hypotheses testing via F-test:  
 $F = MSB/MSE$  or  $F = MST/MSE$ ,  
where MSB, MST and MSE denote the mean sum of squares for block effects, treatment effects and error, weighted by the degrees of freedom.

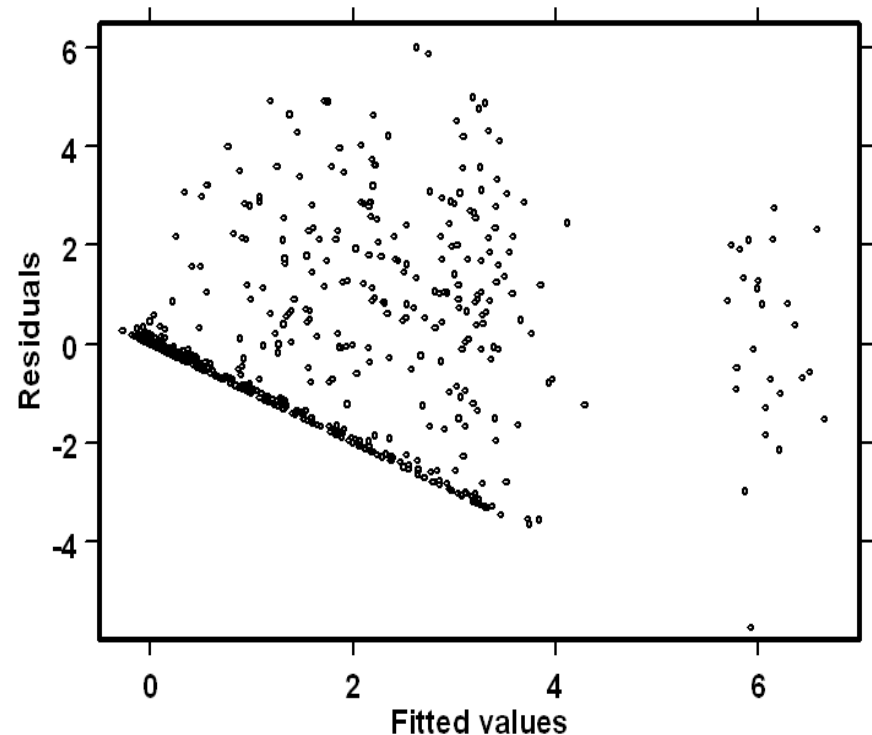


# Problems with ANOVA

Percentage of Zero-Values



Residuals of ANOVA - Variable MILK2



- High percentage of zero values violates normality assumption of ANOVA.
- Residuals do not look like a sample of iid observations, but contain a pattern.



# Alternative Proposal: Tobit-Model

- To avoid these problems, Guillet et al. 2001 proposed to use the tobit model
- Originated in the late 50's (Tobin 1958) in econometrics
- Original use: Analysis of household expenses
- For luxury goods, many households have no expenses at all → limited scale
  - linear regression models: poor fit (local regression more preferable)
  - probit analysis: loss of information
  - combining linear regression and probit analysis to overcome their disadvantages
- Further readings in econometric textbooks, e.g. Amemiya 1985.



# Tobit-Model: Theory (I)

- Let  $x_i \in \mathbb{R}^{k \times 1}$  be a vector of independent variables,  $y_i \in \mathbb{R}$  the observed value of the dependent variable and  $y_i^* \in \mathbb{R}$  an unobservable ideal version of the dependent variable. Then the Tobit-model is:

$$y_i^* = x_i' \beta + u_i, i = 1, \dots, n \quad (1)$$

$$y_i = \begin{cases} y_i^* & , \text{if } y_i^* > T \\ T & , \text{if } y_i^* \leq T \end{cases} \quad (2)$$

with  $u_i \sim_{\text{iid}} N(0, \sigma^2)$

- At first, the ideal version of the dependent variable is observed on a latent scale using a regression model (eq. (1), regression component).
- Then, all values below some „threshold limit“  $T$  are observed as  $T$  (eq. (2), probit component).



# Tobit-Model: Theory (II)

- Estimation of  $\beta$  and  $\sigma^2$  via ML-estimation.
- The Likelihood-function is:

$$L(\beta, \sigma^2) = \prod_{y_i^* \leq T} \Phi\left(\frac{T - x_i' \beta}{\sigma}\right) \cdot \prod_{y_i^* > T} \frac{1}{\sigma} \varphi\left(\frac{y_i - x_i' \beta}{\sigma}\right)$$

with  $\Phi$  and  $\varphi$  denoting the distribution and density function of  $N(0,1)$ .

- Minimizing  $\ln L$  with iterative methods, e.g. Newton-Raphson-Algorithm.
- The first two derivatives can be found in Amemiya 1985.
- R-function: `survreg`



## Tobit-Model: Theory (II)

- Let  $X \in \mathbb{R}^{n \times k}$  be the matrix, in which the  $i$ -th row is  $x'_i$ . Then, make the following assumptions:

(T1)  $\lim_{n \rightarrow \infty} n^{-1}X'X$  exists a.s.

(T2)  $\lim_{n \rightarrow \infty} n^{-1}X'X$  is p.d.

(T3) Let  $\theta = (\beta', \sigma^2)'$  and  $\theta_0 = (\beta'_0, \sigma_0^2)'$  be the true value of  $\theta$ . The parameter space  $\Theta$  is then compact, does not contain the region  $\sigma^2 \leq 0$ , but contains an open  $\varepsilon$ -neighbourhood around  $\theta_0$ .

- Theorem (Proofs in Amemiya 1973, Olsen 1978)**

Let  $L$  be the likelihood function given in (3). If the Jacobian Matrix of  $\ln L$  is not singular and (T1)-(T3) are fulfilled, then we have:

(1)  $\hat{\theta}_n$  is a strongly consistent estimator for the true  $\theta_0$ .

(2)  $\hat{\theta}_n$  is an unique maximum of  $\ln L$ .

(3) The asymptotic distribution of  $\hat{\theta}_n$  is gaussian with

$$\mathcal{L}(\sqrt{n}(\hat{\theta}_n - \theta_0)) \rightarrow_w N \left( 0, \left[ -\frac{1}{n} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'} \right]^{-1} \right).$$



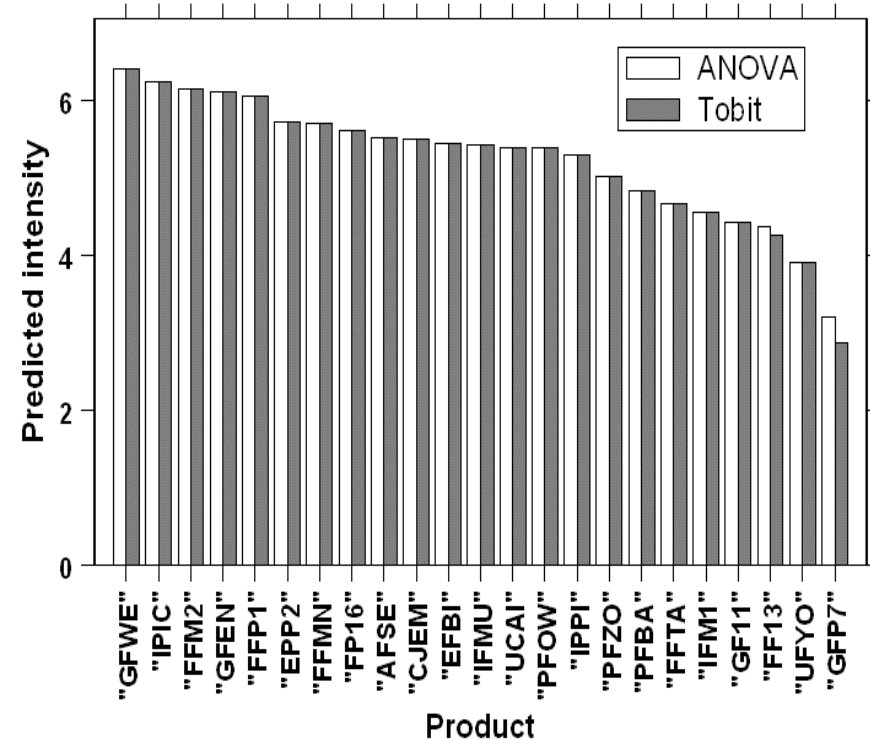
# Tobit-Model: Tests

- Testing hypothesis of the form  
 $H_0 : \{\text{Subset of } \beta_i \text{'s} = 0\}$  vs.  $H_1 : H_0 \text{ is not true.}$
- Using ML-Ratio test. Compute:  
 $L_{\text{full}}$ , the maximum value of the likelihood function, when fitting the full model and  $L_{\text{rest}}$ , the maximum value of the likelihood function, when fitting the restricted model under  $H_0$ .  
Test statistic:  $\lambda = -2 \ln (L_{\text{full}} / L_{\text{rest}}) \sim \chi^2_q$  with  $q = \# \{ i \mid \beta_i \equiv 0 \text{ (under } H_0) \}$ .
- Goal: Test for equality of all products  
= Test, whether any of the product coefficients differ from 0.  
→ Set corresponding  $\beta_i$  to 0.
- Goal: Test for equality of two products  
Use standard t-test with predicted values and variance estimation from the information matrix.

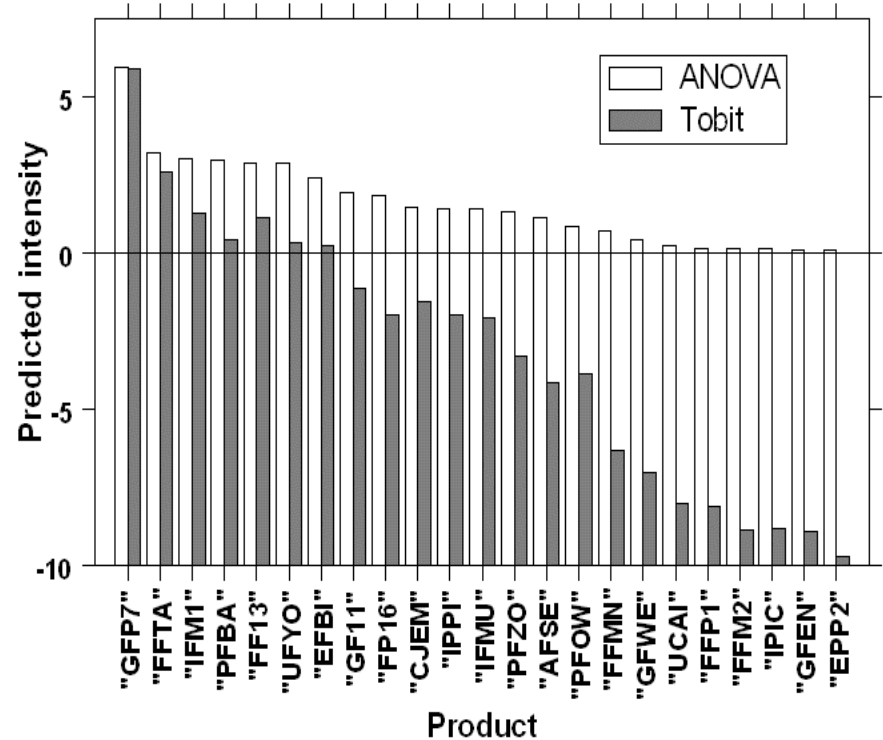


# Previous Results: Guillet et al. 2001

ANOVA vs. Tobit - Variable MILK.INT



ANOVA vs. Tobit - Variable MILK2



- Low amount of zeros → no differences between Tobit and ANOVA
  - High amount of zeros → predicted values for Tobit are on a wider scale.
- ➔ Tobit is a "prolongation" of ANOVA.



# Previous Results: Meyners 2001

- Well-known fact: ANOVA tends to be robust against violations of model assumptions (e.g. Conover et al. 1981)
  - Permutation study, whether the empiric distribution of the ANOVA test statistic matches the theoretically expected F-distribution.
    - Result: Distributions match each other very well.
    - ANOVA seems to be appropriate.
    - ANOVA keeps the theoretical test level.
- Resulting question: If ANOVA keeps the theoretical test level and Tobit detects more differences than ANOVA, will Tobit keep the test level as well?



# Problems of the Tobit-Model

- Asymptotic properties of the MLE only if (T1)-(T3) are true.
  - When drawing samples, (T1)-(T3) are true.
  - In a single block model no samples are drawn  $\rightarrow$   $n$  does not tend to infinity  $\rightarrow$  no asymptotic behavior.
  - Possibilities to let  $n$  grow:
    - $n$  can be increased by adding more treatments or blocks  $\rightarrow$   $\dim(X'X)$  tends to infinity  $\rightarrow$  limit in (T1) does not exist.
    - drawing samples from a pool of treatment-block-combinations  $\rightarrow$  very low likelihood that single block model will be randomly drawn.
- $\rightarrow$  (T1)-(T3) questionable
- $\rightarrow$  Good properties of the Tobit-model are questionable.



# Permutation Principle

- Purpose: Using the test statistic of a test for which the assumptions are not necessarily true in order to test a given pair of hypotheses.
- Procedure of a permutation test:
  - Make  $n$  permutations of the original sample according to the experimental setting of the data.
  - Calculate the test statistic for each of this permuted samples.
  - Null hypothesis can be rejected, if the statistic of the original sample lies among the top (bottom)  $\alpha \cdot n$  of the permuted test statistics.
- Allows to calculate p-values, e.g. number of permuted test statistics exceeding original test statistic divided by  $n$ .
- Allows to calculate an EDF for the test statistic.
- ➔ Possibility to examine, whether a given test is conservative or rejects the null hypothesis too often.



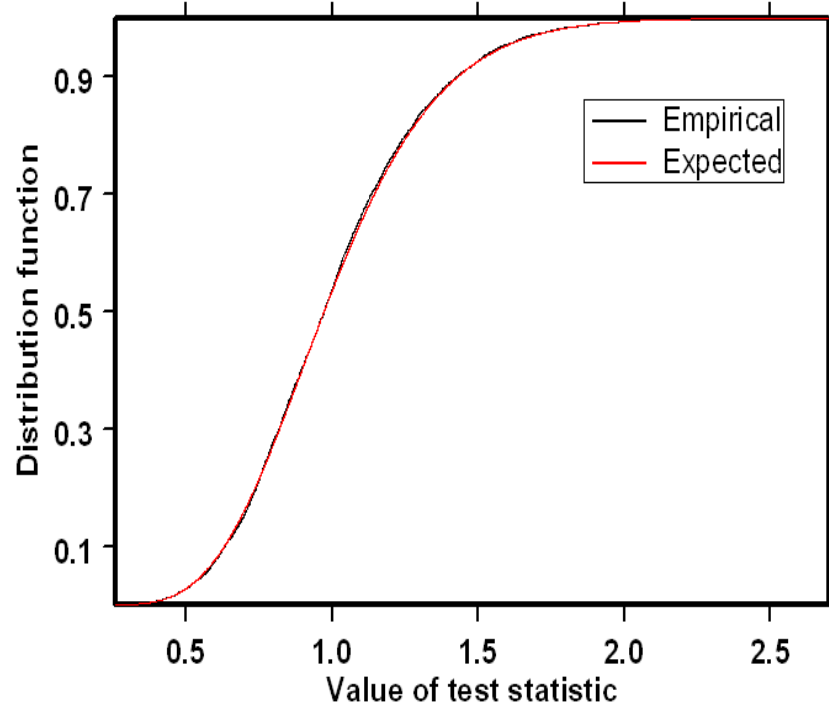
# Simulation Study

- Permutation tests for all 24 variables
  - Test statistics for equality of all products: F-test (ANOVA),  $\chi^2$ -test (Tobit)
  - Test statistics for equality of two products: t-test (ANOVA/Tobit)
- Permutation of intensity grades from each judge
- Usage of the first replication → no interaction term included
- $n = 4000$  for every test
- Performed analyses:
  - Comparison of EDF of the test statistic out of the permuted values with the theoretically expected distribution function.
  - Computation of empirical levels = p-values of the critical value (with  $p=0.01$  and  $p=0.05$ )

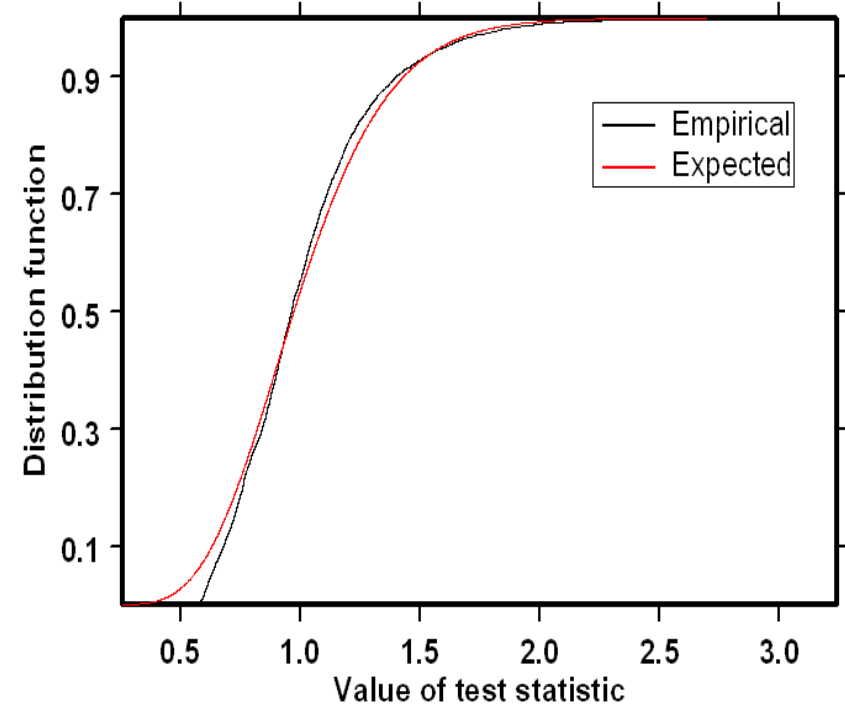


# Results: All Products (I)

Best ANOVA Fit - Variable STRAWBE1



Worst ANOVA Fit - Variable CHERRY2

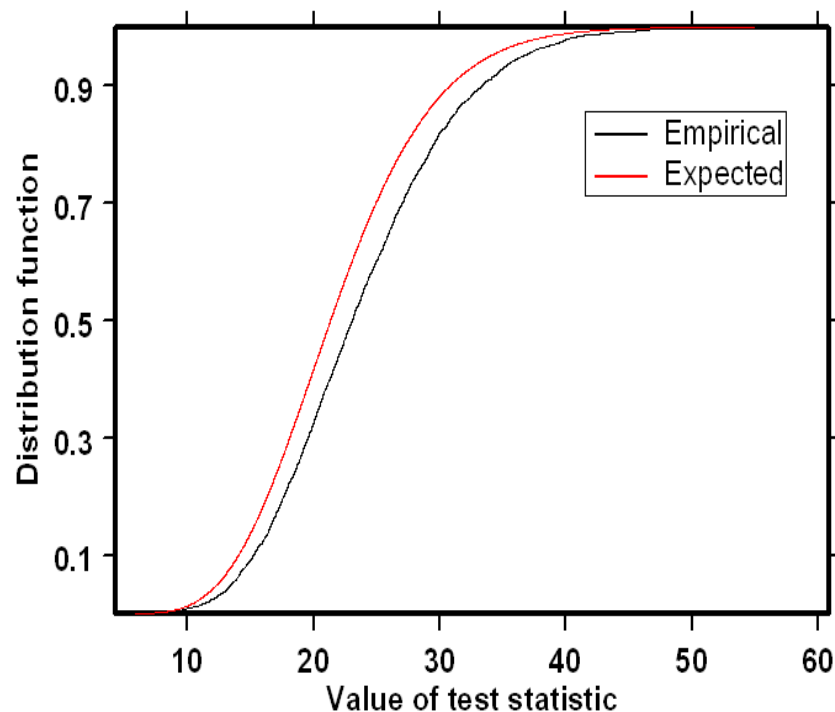


- No differences for most of the variables.
- Slight differences for variables with a high amount of zero values.

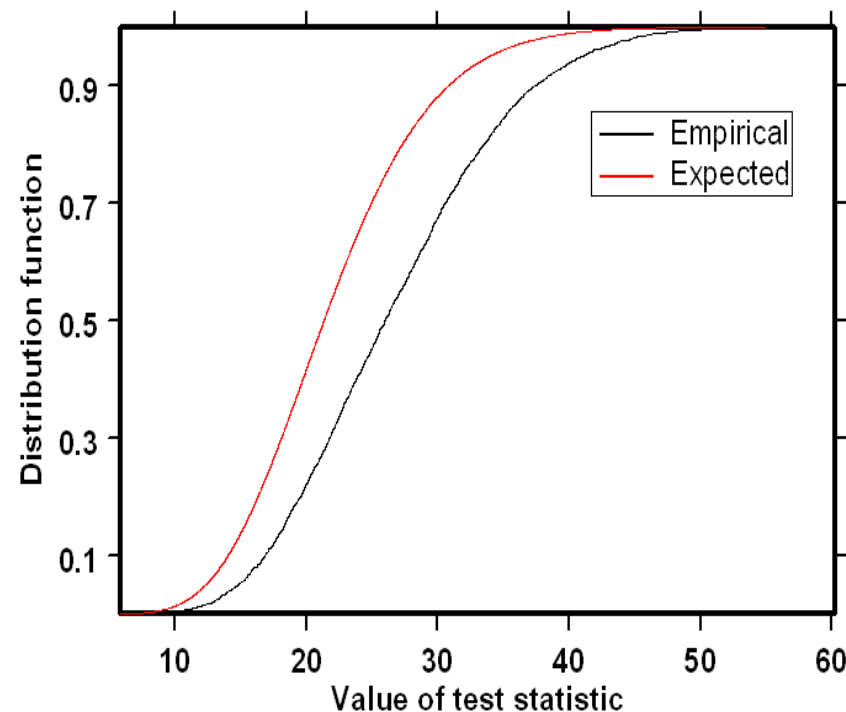


# Results: All Products (II)

Best Tobit Fit - Variable STRAWBE1



Worst Tobit Fit - Variable CHERRY1

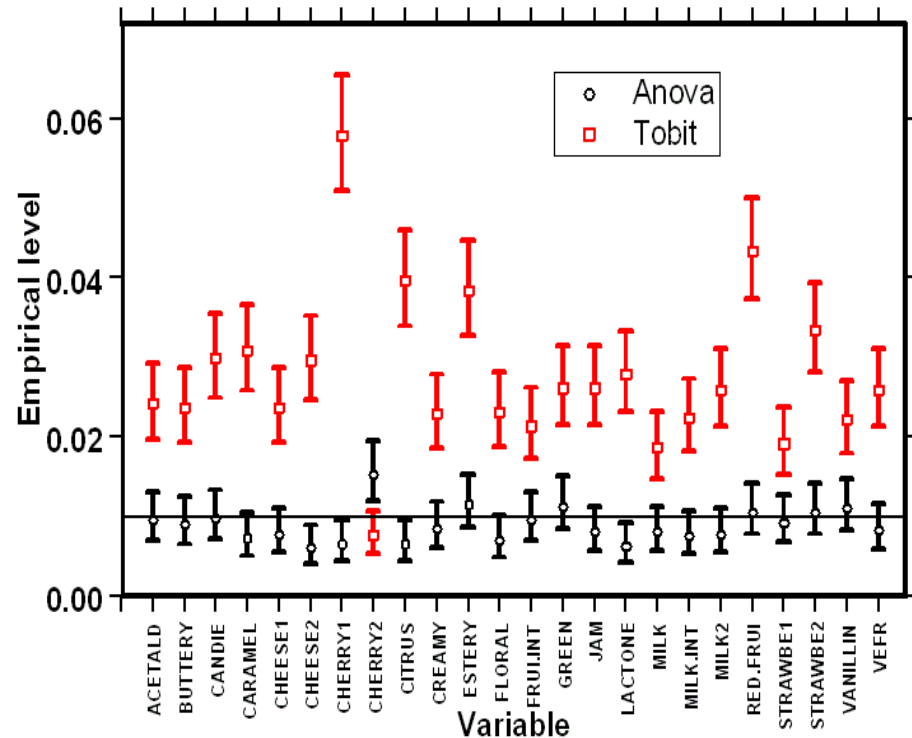


- Huge differences observable for every variable → questionable that Tobit keeps the test level.

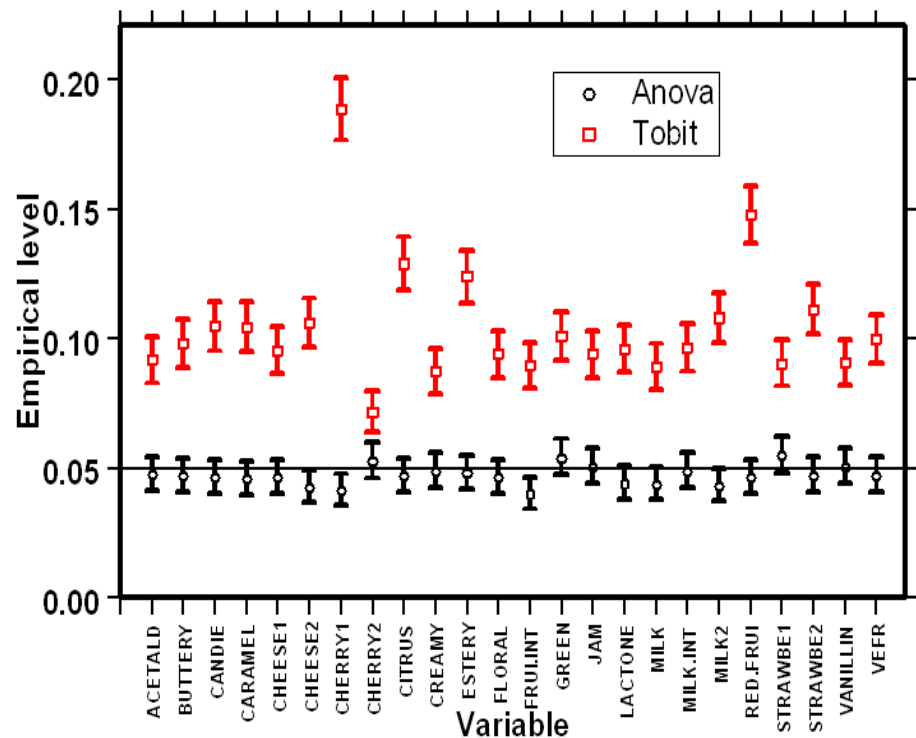


# Results: All Products (III)

Empirical Test Levels for Alpha=0.01



Empirical Test Levels for Alpha=0.05

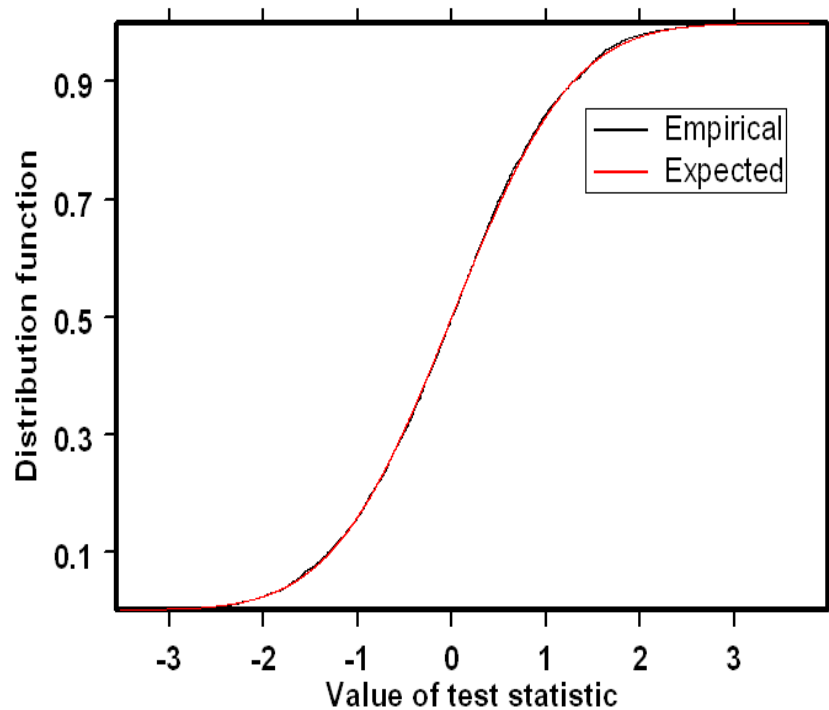


- ANOVA keeps the test level in most cases.
- The Tobit-test is anti-conservative.

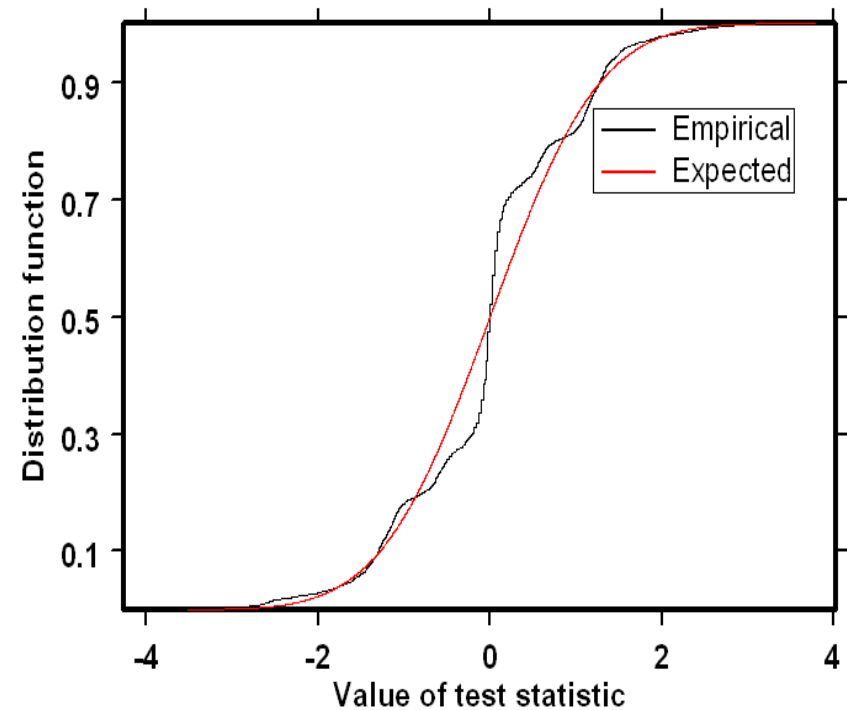


# Results: Two Products (I)

Best ANOVA t-test Fit - Variable LACTONE



Worst ANOVA t-test Fit - Variable CHERRY2

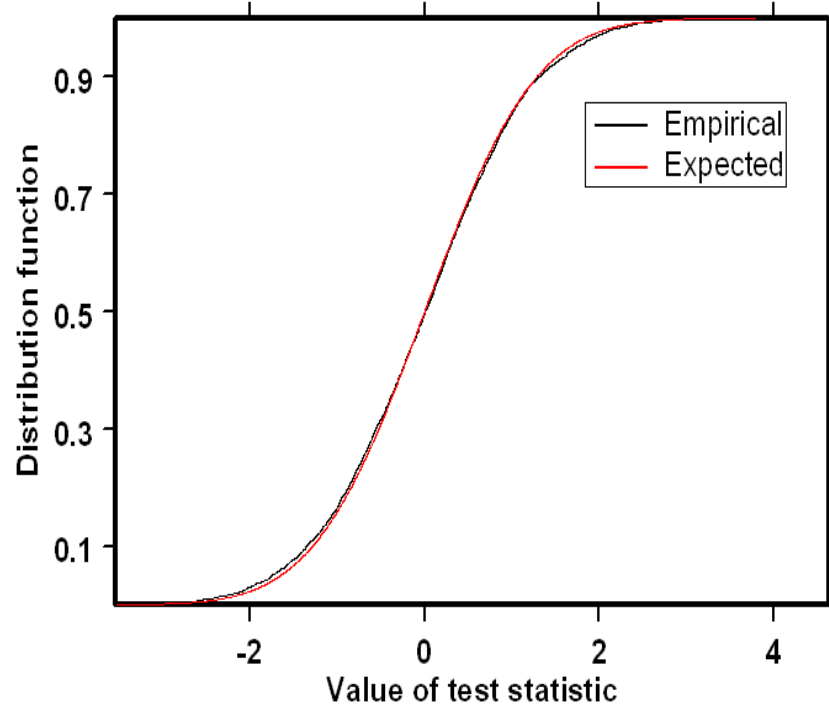


- No differences for most of the variables.
- Rather big differences for variables with a high amount of zero values, but only slight differences in the tails, where the level is built up.

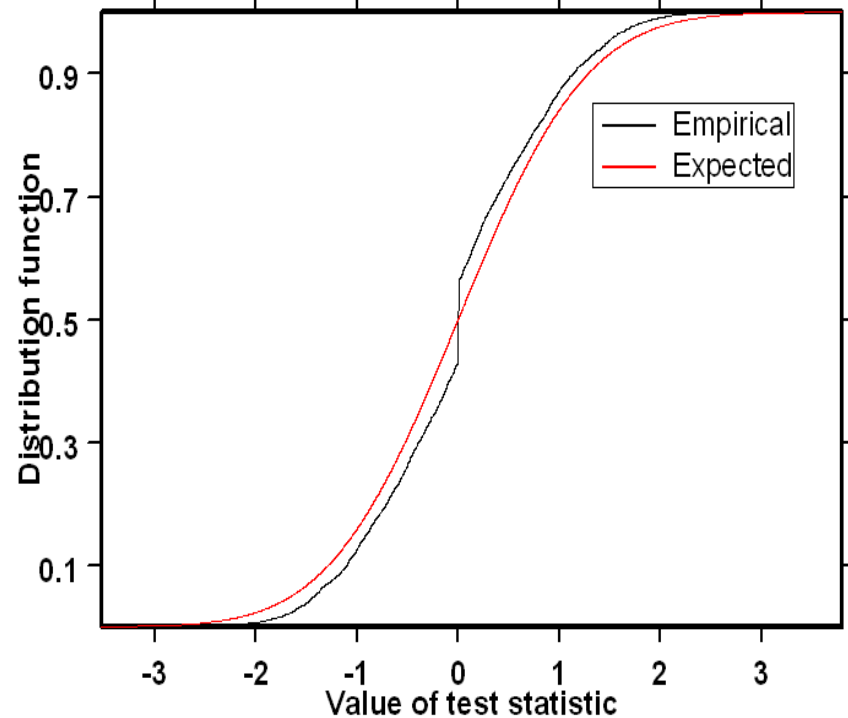


# Results: Two Products (II)

Best Tobit t-test Fit - Variable ACETALD



Worst Tobit t-test Fit - Variable CHERRY1

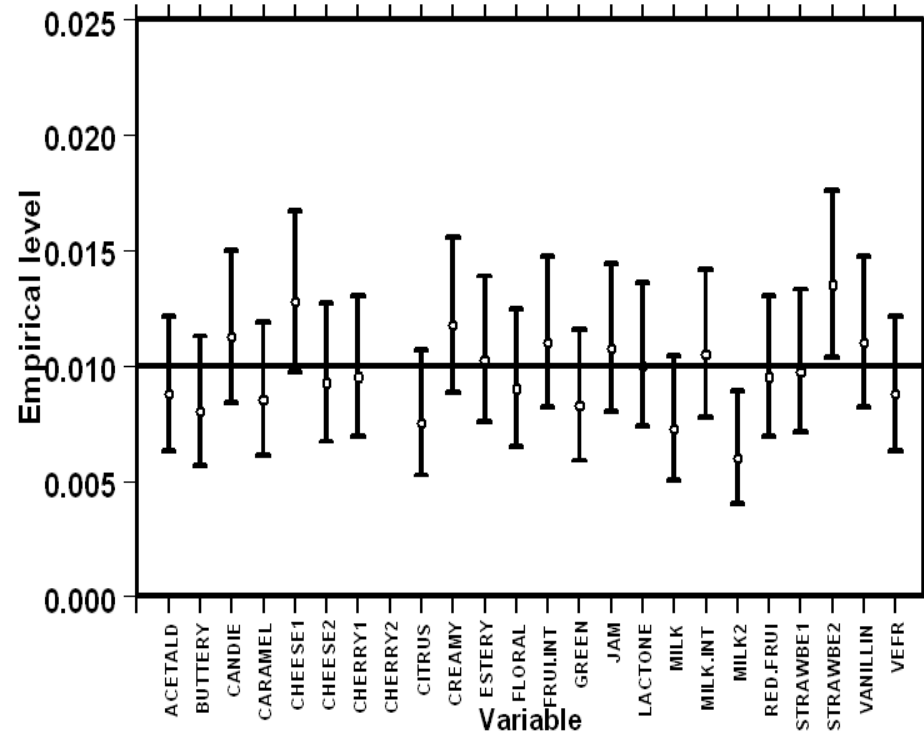


- Very slight differences observable in the tails of every variable  
→ Big enough to cause a violation of the level?

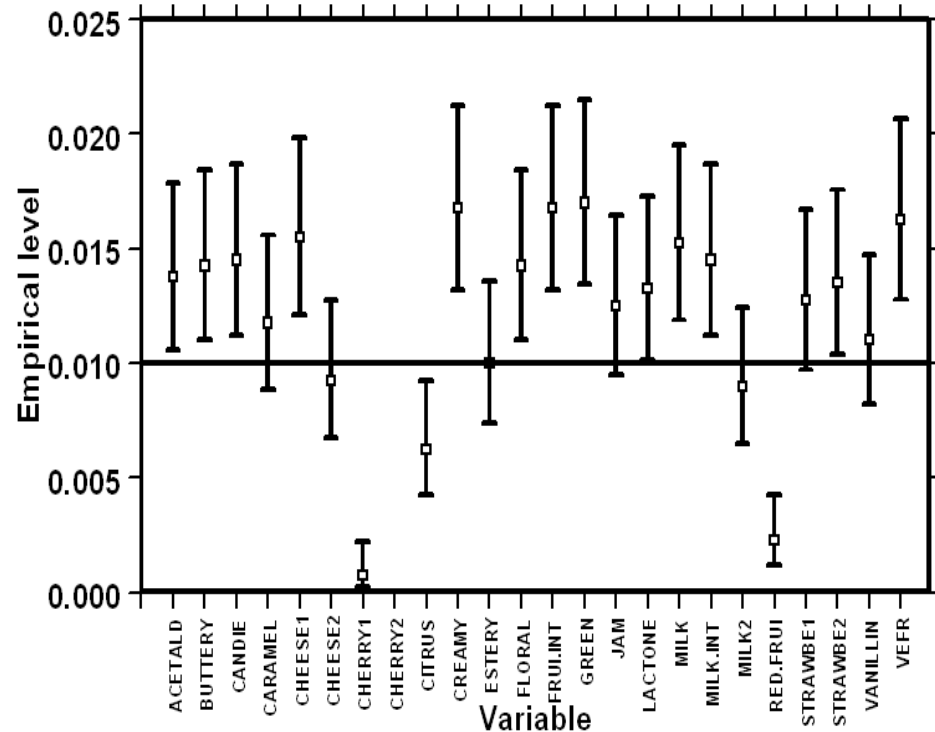


# Results: Two Products (III)

Empirical Test Levels for Alpha = 0.01 - ANOVA



Empirical Test Levels for Alpha = 0.01 - Tobit

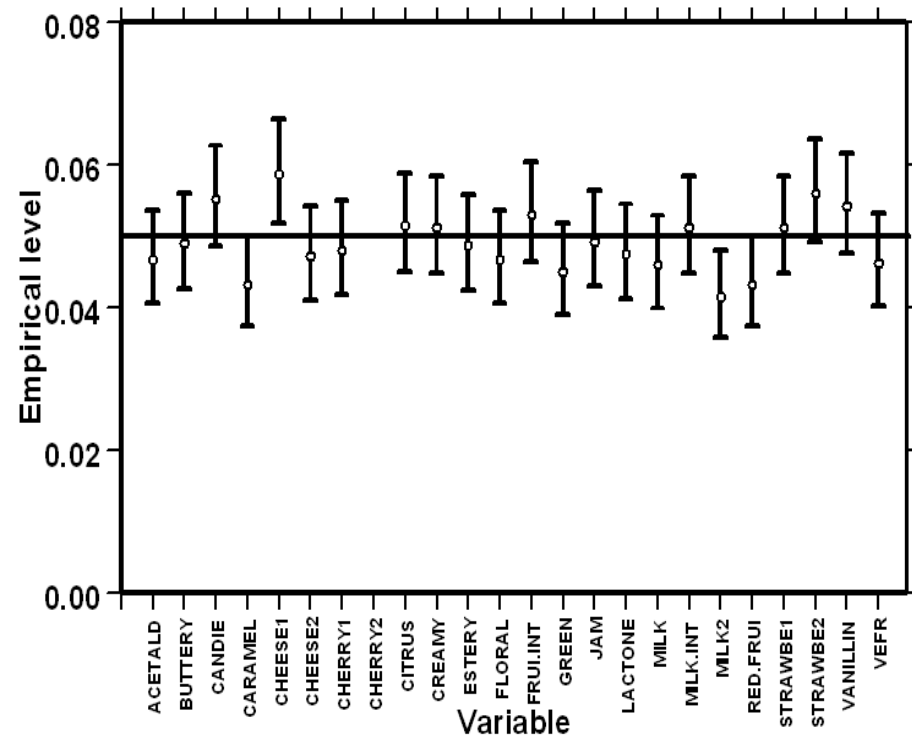


- ANOVA analysis guarantees to keep the level.
- When using tobit, levels are violated for variables with a low proportion of zeros. Tobit is conservative for variables with many zeros.

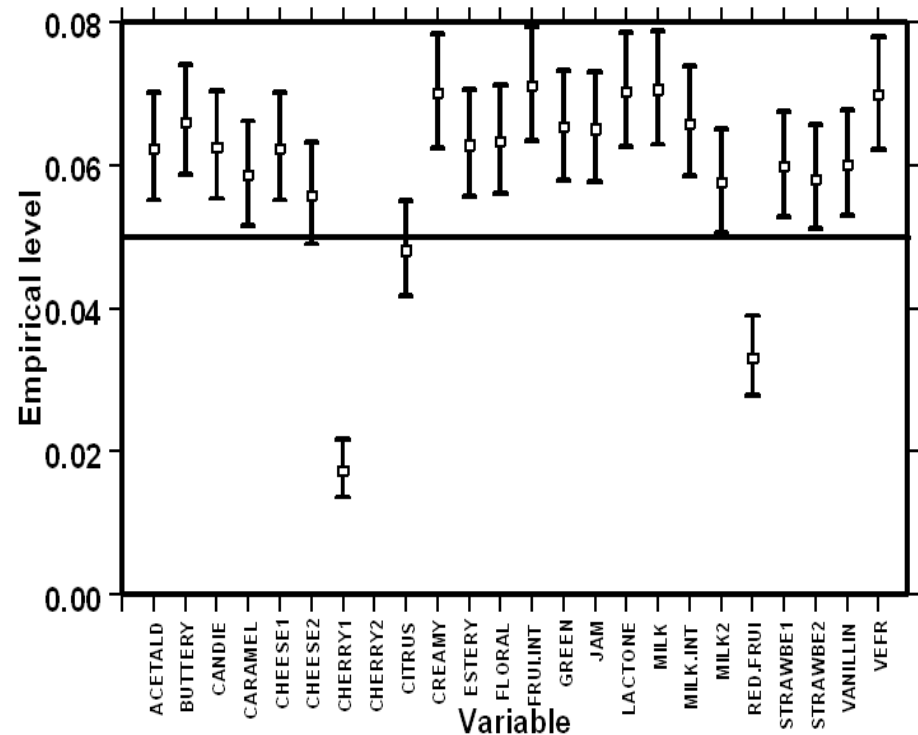


# Results: Two Products (IV)

Empirical Test Levels for Alpha = 0.05 - ANOVA



Empirical Test Levels for Alpha = 0.05 - Tobit



- ANOVA still keeps the level.
- Tobit is anti-conservative with the exception of the three variables with the highest amount of zeros.



# Conclusions

- Special context of sensory profiling
  - ANOVA is able to keep the level of the most tests although some of the model assumptions are not fulfilled.
  - The Tobit-model fails to keep the level most of the time.  
Exception: Test for differences in two products, if a variable contains a very high amount of zeros.
- Conclusion: ANOVA is better suited than the Tobit-model to analyze data in sensory profiling.
- Keeps the test level.
  - Easier to implement and better known.



# Literature

- Amemiya, T. (1973). Regression Analysis when the Dependent Variable is Truncated Normal, *Econometrica* 41, 997-1016.
- Amemiya, T. (1985). *Applied Econometrics*, Harvard University Press, Cambridge, Massachusetts.
- Conover, W.J., Johnson, M.E., Johnson, M.M. (1981). A Comparative Study of Tests for Homogeneity of Variances, with Applications to Outer Continental Shelf Bidding Data, *Technometrics* 23, 351-361.
- Danone (2001). Data set for the Workshop "Sensory Data Analysis". Data presented at the Pangborn conference in Dijon, to be found at <http://www.dijon.inra.fr/aromes/text/danone.rtf>.
- Guillet, M., Methot, S., Rodrigue, N. (2001). Application of Tobit Models to Handle Zero-Valued Attribute Intensities. Talk presented at the Pangborn conference in Dijon, to be found at <http://www.dijon.inra.fr/aromes/presentation/GuilletWS.pdf>
- Meyners, M. (2001). Anova with a Large Number of Zeros in Sensory Profiling. Poster presented at the Pangborn conference in Dijon, to be found at <http://www.dijon.inra.fr/aromes/presentation/MeynersWS.pdf>.
- Olsen, R.J. (1978). Note On The Uniqueness Of The Maximum Likelihood Estimator For The Tobit Model, *Econometrica* 46, 1211-1214.
- R Development Core Team (2004). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables, *Econometrica* 26, 24-36.