NIASRA

NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

National Institute for Applied Statistics Research Australia

University of Wollongong

Working Paper

03-18

QUADE'S TEST AND THE ANALYSIS OF ORDINAL CATEGORICAL DATA

D.J. BEST AND J.C.W. RAYNER

National Institute for Applied Statistics Research Australia, University of Wollongong,
Wollongong NSW 2522. Phone +61 2 4221 5435, Fax +61 2 4221 4845.
Email: anica@uow.edu.au

QUADE'S TEST AND THE ANALYSIS OF

ORDINAL CATEGORICAL DATA

D.J. BEST
School of Mathematical and Physical Sciences, University of Newcastle, Newcastle, Australia
Email: John.Best@newcastle.edu.au

and

J.C.W. RAYNER
National Institute for Applied Statistics Research Australia, University of Wollongong, Wollongong, Australia and
School of Mathematical and Physical Sciences, University of Newcastle, Newcastle, Australia
Email: John.Rayner@newcastle.edu.au

Summary

With ordered categorical data some researchers introduce arbitrary scores and analyse the data using the analysis of variance. This is problematic both because the scores are arbitrary and because the underlying ANOVA assumptions, such as normal residuals, may not be true. Alternative approaches include the Friedman and Quade tests. The following gives an introduction to the latter. A new Page test analogue is given.

Key Words: Friedman test; Page test; two-way ANOVA without replication

1. Introduction

We begin with an example. Twelve panellists made ratings on three wine colours using a categorical hedonic scale with ratings ‘don't like at all’ (1) up to ‘like very much’ (10). The order of tasting was random. The aim of the exercise was to determine whether there was a significant difference between colours as regards ‘liking’. Table 1 shows the data which is based on an example of Zaiontz (2018). The three colours compared were A = red, B = white and C = rose.

A common approach to statistical analysis of the Table 1 data is to assign scores 1, 2, ... , 10 to the ordered categories and to carry out a two-way analysis of variance (ANOVA) without replication. See, for example, Bower (2013, p.189). The choice of scores seems somewhat arbitrary and the ordinal categorical data are clearly not normally distributed. Normality is an assumption of ANOVA. A normality test on residuals may indicate the data are consistent with normality but this could be because the sample size is small or because residuals are a linear combination of values and so by the central limit theorem may appear normal. Given the possibly subjective choice of scores and lack of normality some researchers would apply a Friedman ranks test to the Table 1 data. Again, see for example, Bower (2013, p.189). Often the Friedman p-value is bigger than the ANOVA p-value.

TABLE 1

Hedonic ratings of wine colours

Panelist	A	B	C	Panelist	A	B	C
1	10	7	8	7	5	9	3
2	8	5	5	8	6	6	7
3	7	8	6	9	5	4	6
4	9	6	4	10	10	6	4
5	7	5	7	11	4	7	4
6	4	7	5	12	7	3	3

Intuitively a reason for this is clear from Table 1. Consider the tenth panellist with assigned scores of 10, 6 and 4 with a range of 6. When ranks are taken the range is 2. Such a difference in ranges can result in a loss of power for the Friedman test, especially in the present case when there are few ranks ($k \leq 5$ say) because there are a small number of products to compare. We can improve the Friedman test by weighting the within panellist ranks by the within panellist range and employing the Quade (1979) test. Conover (1999, p.375) gives a good explanation of the Quade test and the following section will define the test in a similar manner.

2. The Quade Test

We refer here to the Table 1 data. First we obtain the ranks within panellists, r_{ij}, for the ith panellist on the jth colour where mid-ranks are used for ties. Then for the ith panellist we find the range of the ranks. The b ranges are then ranked and denoted by q_i. Next, values of $s_{ij} = q_i\{r_{ij} - (k+1)/2\}$ are calculated. Finally carry out a two way ANOVA without replication on the s_{ij} values. The p-value for the Quade test is approximated by that for colours in the ANOVA table.

For the Table 1 data the Quade p-value for colours is 0.078 and for the Friedman test is 0.348 indicating as is often the case that the Friedman test is less sensitive than the Quade test. Here the Quade test is significant at the 10% level and the Friedman test is not. We suspect some researchers use ANOVA to find p-values as these are often smaller than p-values for the Friedman test. The Quade test may negate this possibly questionable approach.

Extensions to finding a Page test analogue for the s_{ij} values or for coping with missing values (including incomplete block designs) may be found in Best and Rayner (2015) and Best and Rayner (2017). The next section discusses the Page test analogue. We leave the reader to apply Best and Rayner (2017) to the s_{ij} values if any are missing.

3. A Page test analogue

Suppose we now return to the Table 1 data. Previous experience suggests there is a liking order A > B > C of the colours. Following Best and Rayner (2015) this ordering can be checked using an analogue F^* of the Page test statistic, given by

$$F^* = \frac{b}{Vd}\left(\sum_{j=1}^{k} l_j \; P_j\right)^2$$

in which the l_j are the linear trend coefficients in the Appendix, P_j is the jth mean of the s_{ij} values, V is the error mean square of the ANOVA on the s_{ij} values and d = sum of squares of the l_j.

For the colour data the ANOVA on the s_{ij} values gives V = 43.15. The colour means of the s_{ij} are 2.896, 0.563 and – 3.456. This gives F^* = 5.614 with p-vaue 0.027 using the F distribution with 1 and 22 degrees of freedom. The *a priori* liking order is confirmed.

4. Conclusion

In this short note we have questioned the common practice in sensory evaluation and other research areas of assigning possibly arbitrary scores to ordinal categorical data. An alternative is to use Friedman's ranking test but this is commonly not as good at picking significant differences between products. Ranks do not indicate the amount or degree of difference between products. We suggest use of the Quade test not the Friedman test. A new Page type analogue was given for *a priori* product ordering.

5. References

Best, D.J. and Rayner, J.C.W. (2015). An alternative to Page's test permitting tied and missing data. *Journal of Statistical Theory and Practice*, 9:3, 524-536. DOI: 10.1080/15598608.2014.940098.

Best, D.J. and Rayner, J.C.W. (2017). Analysis of ranked data in randomized blocks when there are missing values. *Journal of Applied Statistics*. DOI: 10.1080/02664763.2016.1158245. 44(1), 16-23.

Bower, J. (2013). *Statistical Methods for Food Science* (2nd ed.), Wiley, New York.

Conover, W. (1999). *Practical Nonparametric Statistics* (3rd ed.), Wiley, New York.

Quade, D. (1979). Using weighted rankings in the analysis of complete blocks with additive block effects. *Journal of the American Statistical Association,* 74, 680-683.

Zaiontz, C. (2018). *Real Statistics Using Excel.* www.real-statistics.com.

6. Appendix

Linear trend coefficients

k	$l_1, l_2, \ldots, l_k$	d
3	–1, 0, 1	2
4	–3, –1, 1, 3	20
5	–2, –1, 0, 1, 2	10
6	–5, –3, –1, 1, 3, 5	70
7	–3, –2, –1, 0, 1, 2, 3	28