

Analyzing preference rankings when there are too many alternatives

Kar Yin Lam* Alex J. Koning[†] Philip Hans Franses[‡]

Econometric Institute Report EI 2008-06

Abstract

Consumer preferences can be measured by rankings of alternatives. When there are too many alternatives, this consumer task becomes complex. One option is to have consumers rank only a subset of the available alternatives. This has an impact on subsequent statistical analysis, as now a large amount of ties is observed. We propose a simple methodology to perform proper statistical analysis in this case. It also allows to test whether (parts of the) rankings are random or not. An illustration shows its ease of application.

Key words: Rankings, Preferences, Ties

*ERIM & Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, kylam@few.eur.nl

[†]Econometric Institute, Erasmus University Rotterdam, The Netherlands

[‡]Econometric Institute, Erasmus University Rotterdam, The Netherlands

1 Introduction and motivation

Stated consumer preferences can be measured by rankings of alternatives. Rankings are easy to collect, they are easy to interpret, and various statistical tools to evaluate the randomness of observed rankings have been developed. Interestingly, such tools have not yet been fully developed and analyzed for the case where consumers would evaluate only a *subset* of the alternatives, and in the particular interesting case where not all consumers evaluate the *same subset* of alternatives.

This situation, which we address in the current paper, has become increasingly important these days in marketing and consumer research, as it has been widely recognized that individuals face difficulties, or even become dissatisfied, when having to evaluate and compare too many choice options. Iyengar and Lepper (2000) in their famous experiment showed that too much choice can be de-motivating for consumers. In a marketing context, Boatwright and Nunes (2001) demonstrated that a reduction of the assortment is in fact felt as beneficial to consumers, a finding which has been supported by Chernev (2003) and Gourville and Soman (2005), among others. In another context, Deshazo and Fermo (2002) showed that consumers experience task complexity when choice options are plenty, see also Sandor and Franses (2008). In sum, consumers may find it difficult to rank preferences across too large an amount of alternatives.

In the case of having many alternatives, a simple solution would now be to ask consumers only to evaluate a subset of these alternatives. A consequence of this liberty is that subsequent statistical analysis becomes more complicated as the potential number of ties becomes (much) larger. In this paper we advocate a simple methodology to deal with this consequence.

The outline of our paper is as follows. In Section 2 we give a few preliminaries to sketch the problem at hand. In Section 3 we outline the statistical methodology. In Section 4 we illustrate its relevance and ease of use on the observed rankings of ten blockbuster movies where respondents were asked only to rank four of these. Section 5 concludes with a few avenues for further research.

2 Preliminaries

Preference rankings are a common tool in consumer surveys. Various studies indicate that the task of comparing all k objects or products could be too difficult. The respondent may become frustrated if asked to rank many alternatives, and then may select only the most and least preferred and may ignore the rest. Even if the respondent would complete the task and ranks all products, the discriminating power may be doubtful as the ability of respondents to rank alternatives effectively and reliably is a function of the number of comparisons to be made. As a consequence, a statistical methodology to analyze observed rankings should better allow for ties as it is rather likely that respondents are indifferent between a subset of the alternatives.

The problem of task complexity can be alleviated by asking respondents to evaluate all k alternatives but only to give preference rankings for a subset s , with $s < k$, that is selected by each respondent. Hence, each respondent can have a different subset s .

A second issue of our interest is to statistically test observed preference rankings. The corresponding null hypothesis of interest is then

H_0 : There are no differences across the alternatives. Each arrangement of the k ranks is equally likely,

while the alternative is

H_1 : At least one alternative tends to yield a higher ranking than at least one other alternative.

The null hypothesis implies that the underlying distributions for each alternative for each respondent are the same. The null hypothesis is rejected at the significance level α if a test statistic T exceeds the $(1-\alpha)$ th quantile of a chi-square random variable with $k - 1$ degrees of freedom. In the most basic model, where we have one observation per respondent and each respondent ranks all alternatives, the Friedman test statistic (Friedman, 1937) is appropriate and most commonly used.

However, in our case where we ask respondents to evaluate k alternatives but only give preference rankings for a subset s , that is selected by each

respondent, matters are different. We assume that respondents are indifferent to alternatives outside their own subset and these alternatives thus receive equal rank. As a consequence, we must deal with a substantial number of ties.

We aim to propose a statistical method that can handle this situation and we propose an appropriate test statistic to examine if observed rankings imply statistically significant differences across the alternatives. If there is statistical evidence of such differences, the final question is of course which alternatives it concerns. To this end, we apply multiple comparison procedures to test which alternatives are significantly different from others and in this way we can construct homogeneous subsets with alternatives that have equal rankings.

3 Methodology

Suppose there are k alternatives, with $j = 1, \dots, k$, and that there are n respondents, with $i = 1, \dots, n$, who are asked to indicate their top s alternatives and to assign ranks to the alternatives in this subset, where the most preferred alternative gets rank value 1 and the least preferred gets rank s . We assume that respondents are indifferent between the $k - s$ alternatives outside this subset, which are thus observed as ties. Denote the observed rankings of respondent i for alternative j by x_{ij} . Let $R(x_{i1}, \dots, x_{ik})$ be the set of all possible rankings consistent with x_{i1}, \dots, x_{ik} given the ties for respondent i . The weighted rank \bar{r}_{ij} assigned to x_{ij} is defined as the average of all possible ranks within the set $R(x_{i1}, \dots, x_{ik})$. In Wittkowski (1988) it is shown that the weighted rank for tied ranks is equal to the average of the ranks "available" for those tied ranks. Adjusted (centered-weighted) ranks a_{ij} are obtained by subtracting the expected score $\frac{k+1}{2}$ under H_0 .

Insert Table 1 about here

In Table 1 we illustrate the computation of weighted and adjusted ranks for a respondent who has indicated her top $s = 3$ alternatives from a set of

$k = 6$. This respondent has indicated alternative C as most preferred, alternative A is next preferred and alternative D is least preferred. We assume that this respondent is indifferent to alternatives B, E and F . Ranks within subset (A, C, D) are fixed, however ranks assigned to alternatives outside this subset could be rank 4, 5 or 6. The set of all possible rankings is listed in Table 1. It then follows that the average of all possible ranks for an alternative equals the average of the "available" ranks for that alternative. Alternatives outside the subset thus receive tied rank $\frac{4+5+6}{3} = 5$. The adjusted ranks are obtained by subtracting the expected score under H_0 , which in this case is $\frac{6+1}{2} = 3.5$, from the weighted ranks, see the last column of Table 1.

The test statistic for H_0 is based on sums of adjusted ranks and the individual covariance matrix V_i , which is given in Wittkowski (1988) as

$$V_i = A_{0,i}^2 \left(\text{diag}(\iota) - \frac{\iota \iota^T}{k} \right) \quad (1)$$

where ι is a k -dimensional vector of ones, $\text{diag}(\iota)$ denotes a diagonal matrix with elements 1 on the diagonal and where $A_{0,i}^2$ denotes the (conditional) individual variance under H_0 with correction for ties, which is

$$A_{0,i}^2 = \frac{k(k+1)}{12} \left(1 - \frac{k-s-1}{k-1} \frac{k-s}{k} \frac{k-s+1}{k+1} \right) \quad (2)$$

This individual variance $A_{0,i}^2$ is the same for all i , as the sum of the (squared) adjusted ranks for each respondent is the same, which of course is a crucial aspect in our methodology. Since each respondent only has to assign ranks to s alternatives, and alternatives outside this subset all receive the same tied rank, this yields the same sum of ranks for each respondent. In our example in Table 1, $A_{0,i}^2 = 3.1$ for each respondent, see Table 2.

The individual variance $A_{0,i}^2$ is instrumental in computing the individual covariance matrix V_i . Note that (1) implies that the diagonal elements of V_i are given by

$$V_i = A_{0,i}^2 \frac{k-1}{k}$$

and the off-diagonal elements by

$$V_i = A_{0,i}^2 \frac{-1}{k}$$

Hence, the diagonal elements of (1) are $1 - k$ times those of the off-diagonal elements. One can observe that V_i only depends on k and $A_{0,i}^2$, and both are the same for each respondent. As a consequence, V_i is also the same for each respondent. In our example, V_i has $\frac{5}{6}$ as diagonal elements and $\frac{-1}{6}$ as off-diagonal elements, multiplied by $A_{0,i}^2$. The computation of $A_{0,i}^2$ and V_i in this example is given in Table 2.

Insert Table 2 about here

The adjusted ranks can be summarized in a vector a_+ by summing all a_i over the n respondents, that is, the j th element of vector a_+ is given by $a_{+j} = \sum_{i=1}^n a_{ij}$. The random vectors a_i are independent under H_0 and the covariance matrix V_+ of the vector a_+ is thus equal to the sum of the individual covariance matrices V_i . As the V_i is the same for all i , V_+ is obtained by multiplying V_i by the number of respondents n . Hence, the diagonal elements of V_+ are given by

$$V_{+jj} = nA_{0,i}^2 \frac{k-1}{k}$$

and the off-diagonal elements by

$$V_{+jj^*} = nA_{0,i}^2 \frac{-1}{k}, \quad j \neq j^*$$

As a_{ik} in each i can be expressed in terms of $a_{i1}, \dots, a_{i(k-1)}$, it follows that V_+ is not of full rank but has rank $k - 1$ instead.

For large n , a_+ approximately has a multivariate normal distribution with zero expectation vector and covariance matrix V_+ under the null hypothesis H_0 (Wittkowski, 1988).

Test statistic

The quadratic-form test statistic is now computed along standard lines as

$$W = a'_+ (V_+)^- a_+, \quad (3)$$

where $(V_+)^-$ denotes a generalized inverse of V_+ , that is, any matrix which satisfies

$$V_+ (V_+)^- V_+ = V_+$$

Below we will use the Moore-Penrose generalized inverse of V_+ .

Under H_0 and for large n , W in (3) has approximately a chi-square distribution with $k - 1$ degrees of freedom. Recall that $k - 1$ is the rank of V_+ . If the corresponding p -value is below the $\alpha\%$ significance level, the null hypothesis H_0 is rejected and there is sufficient evidence to conclude that there exists difference in preference rankings between the k alternatives.

Multiple comparisons

When the null hypothesis H_0 is not rejected, it is generally agreed that all hypotheses implied by that hypothesis (its "components") must also be considered as not rejected. When H_0 is rejected, we conclude that there exist differences between the alternatives. However, we do not know which alternatives differ in terms of preference rankings and so we resort to a multiple comparison procedure to make decisions about differences between all $\frac{k(k-1)}{2}$ pairs of alternatives. Note again that multiple comparisons are only of interest if the global null hypothesis H_0 is rejected.

Denote the average adjusted rank over the respondents for alternative j by

$$a_{.j} = \frac{1}{n} \sum_{i=1}^n a_{ij}$$

The null hypothesis H_{0,jj^*} for each pairwise comparison of no difference between alternative j and alternative j^* is rejected if

$$|a_{.j} - a_{.j^*}| \geq r_{k,n}^\alpha \quad (4)$$

where the critical value $r_{k,n}^\alpha$ is chosen to make the type I error rate equal to α . That is, $r_{k,n}^\alpha$ is the largest constant such that

$$P_{H_0} \left((\max a_{.j}) - (\min a_{.j}) \geq r_{k,n}^\alpha \right) \leq \alpha \quad (5)$$

This implies that when H_0 is true all $\frac{k(k-1)}{2}$ inequalities in (4) fail to exceed the critical constant with probability α .

When H_0 is true, we show in the Appendix that the covariance matrix of the difference $a_{.j} - a_{.j^*}$ is the same as the covariance matrix of the differences $Z_j - Z_{j^*}$ where Z_j, Z_{j^*} are independent random variables with mean zero and variance $\frac{A_{0,i}^2}{n}$, with $j, j^* = 1, \dots, k$. It then follows that when n is large, the critical value $r_{k,n}^\alpha$ can be approximated by

$$r_{k,n}^\alpha \approx q_k^\alpha \frac{A_{0,i}}{\sqrt{n}} \quad (6)$$

where q_k^α is the upper α percentile point of the range of k independent standard normal random variables. The critical points q_k^α for $k = 1, \dots, 20$ and $\alpha = 0.001, \dots, 0.1$ are given in Table 3.

Rank plots

To visualize the results of the multiple comparison procedures, we can construct for each j an interval Q_j centered at $a_{.j}$ with length $r_{k,n}^\alpha$ and endpoints

$$a_{.j} \pm \frac{r_{k,n}^\alpha}{2}$$

When we observe that interval Q_j and Q_{j^*} do not overlap, the distance between $a_{.j}$ and $a_{.j^*}$ exceeds $r_{k,n}^\alpha$ and hence (4) should be rejected, yielding the conclusion that there is a significant difference in rank between alternative j and alternative j^* . The rank plot simultaneously displays the intervals Q_1, \dots, Q_k . Moreover, a reference line can be drawn at the height of the upper boundary of the interval of the "most preferred" alternative, which naturally has the lowest upper boundary. This reference line in fact visualizes the "unconstrained multiple comparison procedure with the best, deducted

from all-pairwise comparisons” as described in Hsu (1996) (section 4.2.1.1). This implies that all alternatives with intervals above this reference line are rated significantly lower than the best ranked alternative.

Homogeneous subsets

Finally, based on the multiple comparisons results, we can form homogeneous groups of alternatives by performing a cluster analysis. The clustering algorithm we prefer is the complete linkage clustering (see for example Latin et al. (2003)), where the maximum distance between elements of each cluster is used. Because the cluster analysis should be based on the results of the multiple comparison procedure, we construct a distance matrix which summarizes the significance tests (4) by zero’s (non rejection) and ones (rejection). It is well known that such a distance matrix of zero’s and ones could lead to multiple solutions in complete linkage clustering. We prevent this by adding the distance matrix of $a_{.j}$ when alternatives are not significantly different. However, these distances may not dominate the multiple comparisons results and to avoid this, we multiply the distance matrix of $a_{.j}$ by ϵ , with ϵ small enough like 0.001.

4 Illustration

In this section we illustrate our statistical methodology to analyze preference rankings where individuals are asked to rank just s of the total k alternatives.

Data

We illustrate our proposed methodology with data of $n = 93$ individuals who are asked to evaluate a list of $k = 10$ blockbuster movies in Dutch cinema theatres in 2007. The movies are listed in Table 4.

Insert Table 4 about here

Respondents are asked to indicate their top $s = 4$ movies. Movies outside this subset are observed as ties and their weighted rank is the average of the

available ranks, that is $\frac{5+6+7+8+9+10}{2} = 7.5$. The adjusted ranks of individual i for each movie is obtained by subtracting the expected score, that is $\frac{10+1}{2} = 5.5$, from each weighted rank. As a_{ij} takes all values from 1 to s and 7.5 for tied ranks for each respondent, the sum of the adjusted scores is the same for each respondent and, as a consequence, the individual variance and individual covariance matrix are the same for all i . These are respectively $A_{0,i}^2 = \frac{65}{9}$ and V_i is given by a matrix with $A_{0,i}^2 \frac{9}{10}$ on the diagonal and $A_{0,i}^2 \frac{-1}{10}$ on the off-diagonal elements. To compute the test statistic (3) we sum the adjusted ranks over all 93 individuals and a_{+j} for each movie is given in the second column of Table 4. In the same way, we compute the conditional covariance matrix V_+ by summing all individual covariance matrices V_i . Recall that this V_+ has rank 9.

Results

Our null hypothesis H_0 of no differences between the movies is clearly rejected as the value of the test statistic (3) takes the value $W = 83.276$ with corresponding p -value is 0. We conclude that there is sufficient evidence that there exists a difference between the movies.

As H_0 is rejected, the question remains which movies differ in terms of preferences and so we perform multiple comparisons between all pairs of movies. According to Table 3 in the Appendix, for $\alpha = 0.05$ and $k = 10$, $q_k^\alpha = 4.474$. The critical value $r_{k,n}^\alpha$ can now be approximated by (6) and it takes the value $r_{k,n}^\alpha = 1.247$.

To visualize the results of the multiple comparison procedure, we construct rank plots. For convenience the $a_{.j}$ values are retranslated into the average weighted ranks $\bar{r}_{.j}$, that is $\frac{1}{n} \sum_{i=1}^n \bar{r}_{ij}$. As we have already calculated the sum of the adjusted ranks for each movie, we can also compute this average rank $\bar{r}_{.j}$ by adding the expected score to the average adjusted score of movie j , that is

$$\bar{r}_{.j} = a_{.j} + \frac{k+1}{2}$$

The average weighted ranks are given in the last column of Table 4. The interval Q_j for movie j is then centered at $\bar{r}_{.j}$ and has length $r_{k,n}^\alpha$. Hence,

the endpoints of Q_j are given by

$$\bar{r}_{.j} \pm \frac{r_{k,n}^\alpha}{2}$$

If the intervals Q_j and Q_{j^*} do not overlap, then the distance between $\bar{r}_{.j}$ and $\bar{r}_{.j^*}$ exceeds $r_{k,n}^\alpha$, and hence the corresponding null hypothesis (4) should be rejected. The corresponding rank plot is displayed in Figure 1.

Insert Figure 1 about here

We can observe in Figure 1 that the best ranked movie is $j = 10$, which corresponds to the movie "The Bourne Ultimatum". This interval has the lowest upper boundary and consequently all intervals above this upper boundary are ranked significantly lower than that movie. This holds for the intervals of the movies 2, 3, 4, 5, 6, 8 and 9, and hence, these movies are ranked significantly lower than "The Bourne Ultimatum". The intervals of movie 1, "Pirates of the Caribbean: At World's End", and 7, "Ocean's Thirteen", do have overlap with the interval of the movie "The Bourne Ultimatum" and hence are not significantly ranked lower.

Based on the multiple comparisons we now create homogeneous subsets of movies by performing a cluster analysis. The corresponding distance matrix is the sum of a matrix which summarizes the significance tests (4) and the distance matrix of $a_{.j}$ multiplied by $\epsilon = 0.001$. Homogeneous subsets of movies are then formed by the corresponding dendrogram, which can be found in Figure 2.

Insert Figure 2 about here

Figure 2 suggests three main clusters. Cluster 1 contains the best ranked movies (10, 1 and 7): "The Bourne Ultimatum", "Pirates of the Caribbean: At World's End" and "Ocean's Thirteen". Cluster 2 contains the worst ranked movie (5): "Mr. Bean's Holiday" and the last cluster contains all other movies: "Harry Potter and the order of the Phoenix", "Alles is Liefde", "Shrek the Third", "Ratatouille", "Spider-Man 3" and "Transformers". In sum, there seem to be just three clusters of movies with the same within-cluster rank.

5 Conclusion

Preference rankings are easy to perform and the outcomes are easy to understand. However, in practice we often encounter the problem that consumers find it difficult to rank preferences across a large amount of alternatives. Our solution is to ask consumers to rank only a subset of these alternatives and we have shown that it does not matter whether these individual subsets are all different. We have provided a methodology how to handle ties in this situation and how to analyze such data. We have given the test statistic to examine if observed rankings imply statistically significant differences between alternatives. Further, if there are differences across alternatives, we have explained how to apply multiple comparisons to determine which alternatives differs. Moreover, based on these multiple comparisons we propose a method to perform a cluster analysis to construct homogeneous groups of alternatives. We have illustrated our methodology with data of ten blockbuster movies in Dutch cinema theatres and we found that there are basically just three groups of movies with common ranks.

We envisage a range of practical applications of our methodology, for example in the area of conjoint analysis. Consumers can now face many alternatives, but when they are asked to rank just a few alternatives, the task will become less demanding.

References

- Boatwright, P., J.C. Nunes. 2001. Reducing assortment: An attribute-based approach. *Journal of Marketing* **65** 50–63.
- Chernev, A. 2003. When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *Journal of Consumer Research* **30** 170–183.
- Deshazo, J.R., G. Fermo. 2002. Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *Journal of Environmental Economics and Management* **43** 360–385.
- Friedman, M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* **32** 675–701.
- Gourville, J.T., D. Soman. 2005. Overchoice and assortment type: When and why variety backfires. *Marketing Science* **24** 382–395.
- Harter, H.L. 1960. Tables of range and studentized range. *The Annals of Mathematical Statistics* **31** 1122–1147.
- Hsu, J.C. 1996. *Multiple comparisons. Theory and methods*. Chapman & Hall, London.
- Iyengar, S.S., M. Lepper. 2000. When choice is de-motivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology* **79** 995–1006.
- Lattin, J., J.D. Carroll, P.E. Green. 2003. *Analyzing Multivariate Data*. Brooks/Cole, Canada.
- Sandor, Z., P.H. Franses. 2008. Experimental investigation of consumer price evaluations. *Journal of Applied Econometrics* **to appear**.
- Wittkowski, K.M. 1988. Friedman-type statistics and consistent multiple comparisons for unbalanced designs with missing data. *Journal of the American Statistical Association* **83** 1163–1170.

Appendix

Proof of (6) Let $a_{.j}$ denote the average adjusted rank for alternative j . It is shown that the moments of $a_{.j}$ are

$$\begin{aligned} E(a_{.j}) &= 0 \\ \text{Var}(a_{.j}) &= \frac{1}{n} \frac{k-1}{k} A_{0,i}^2 \\ \text{Cov}(a_{.j}, a_{.j^*}) &= \frac{1}{n} \frac{-1}{k} A_{0,i}^2 \\ \text{Cor}(a_{.j}, a_{.j^*}) &= \frac{-1}{k-1} \end{aligned}$$

and hence, the moments of the difference $a_{.j} - a_{.j^*}$ are

$$\begin{aligned} E(a_{.j} - a_{.j^*}) &= 0 \\ \text{Var}(a_{.j} - a_{.j^*}) &= \frac{2}{n} A_{0,i}^2 \\ \text{Cov}(a_{.j} - a_{.j^*}, a_{.j} - a_{.j^{**}}) &= \frac{1}{n} A_{0,i}^2 \\ \text{Cov}(a_{.j} - a_{.j^*}, a_{.j^{**}} - a_{.j^{***}}) &= 0 \\ \text{Cor}(a_{.j} - a_{.j^*}, a_{.j} - a_{.j^{**}}) &= \frac{1}{2} \\ \text{Cor}(a_{.j} - a_{.j^*}, a_{.j^{**}} - a_{.j^{***}}) &= 0 \end{aligned}$$

This means that the covariance matrix of the differences $a_{.j} - a_{.j^*}$ is the same as the covariance matrix of the differences $Z_j - Z_{j^*}$ where Z_j, Z_{j^*} are independent random variables with mean zero and variance $\frac{A_{0,i}^2}{n}$, with $j, j^* = 1, \dots, k$. Thus, the asymptotic distribution of

$$\frac{\max_{j,j^*} |a_{.j} - a_{.j^*}|}{A_{0,i}/\sqrt{n}}$$

coincides with the distribution of the range $Q_{k,\infty}$ of k independent standard normal random variables. This proves (6). \square

Table 1: Computation of adjusted ranks for a hypothetical dataset with $s = 3$ and $k = 6$ and one respondent.

Alternative	x_{ij}	$R(x_{i1}, \dots, x_{ik})$						Weighted rank	Adjusted rank
A	2	2	2	2	2	2	2	12/6=2	-1.5
B	-	4	4	5	5	6	6	30/6=5	1.5
C	1	1	1	1	1	1	1	6/6=1	-2.5
D	3	3	3	3	3	3	3	18/6=3	-0.5
E	-	5	6	4	4	6	5	30/6=5	1.5
F	-	6	5	6	5	4	4	30/6=5	1.5

Table 2: Example continued. Computation of conditional individual covariance matrix.

$$A_{0,i}^2 = \frac{6(6+1)}{12} \times \left(1 - \frac{2}{5} \times \frac{3}{6} \times \frac{4}{7}\right) = \frac{31}{10}$$

$$V_i = A_{0,i}^2 \times \begin{pmatrix} 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} \end{pmatrix}$$

$$V_+ = nA_{0,i}^2 \times \begin{pmatrix} 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} & \frac{-1}{6} \\ \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & \frac{-1}{6} & 1 - \frac{1}{6} \end{pmatrix}$$

Table 3: Selected percentile points of the range of k independent $N(0, 1)$ random variables.

$k \backslash \alpha$	0.1	0.05	0.025	0.01	0.005	0.001
2	2.326	2.772	3.170	3.643	3.970	4.654
3	2.902	3.314	3.682	4.120	4.424	5.063
4	3.240	3.633	3.984	4.403	4.694	5.309
5	3.478	3.858	4.197	4.603	4.886	5.484
6	3.661	4.030	4.361	4.757	5.033	5.619
7	3.808	4.170	4.494	4.882	5.154	5.730
8	3.931	4.286	4.605	4.987	5.255	5.823
9	4.037	4.387	4.700	5.078	5.341	5.903
10	4.129	4.474	4.784	5.157	5.418	5.973
11	4.211	4.552	4.858	5.227	5.485	6.036
12	4.285	4.622	4.925	5.290	5.546	6.092
13	4.351	4.685	4.985	5.348	5.602	6.144
14	4.412	4.743	5.041	5.400	5.652	6.191
15	4.468	4.796	5.092	5.448	5.699	6.234
16	4.519	4.845	5.139	5.493	5.742	6.274
17	4.568	4.891	5.183	5.535	5.783	6.312
18	4.612	4.934	5.224	5.574	5.820	6.347
19	4.654	4.974	5.262	5.611	5.856	6.380
20	4.694	5.012	5.299	5.645	5.889	6.411

Source: Table 3 in Harter (1960).

Table 4: List of $k = 10$ blockbuster movies in Dutch cinema theatres in 2007, ranked in the first column according to the total size of audience.

Name movie	Sum adjusted ranks	Average weighted ranks
1. Pirates of the Caribbean: At World's End	-79.5	4.645
2. Harry Potter and the order of the Phoenix	1.0	5.511
3. Alles is Liefde	40.5	5.936
4. Shrek the Third	-6.0	5.435
5. Mr. Bean's Holiday	150.0	7.113
6. Ratatouille	1.5	5.516
7. Ocean's Thirteen	-66.5	4.785
8. Spider-Man 3	76.5	6.323
9. Transformers	5.5	5.559
10. The Bourne Ultimatum	-123.0	4.177

Figure 1: Rank plot with $\alpha = 0.05$ of $k = 10$ blockbuster movies in Dutch cinema theatres in 2007.

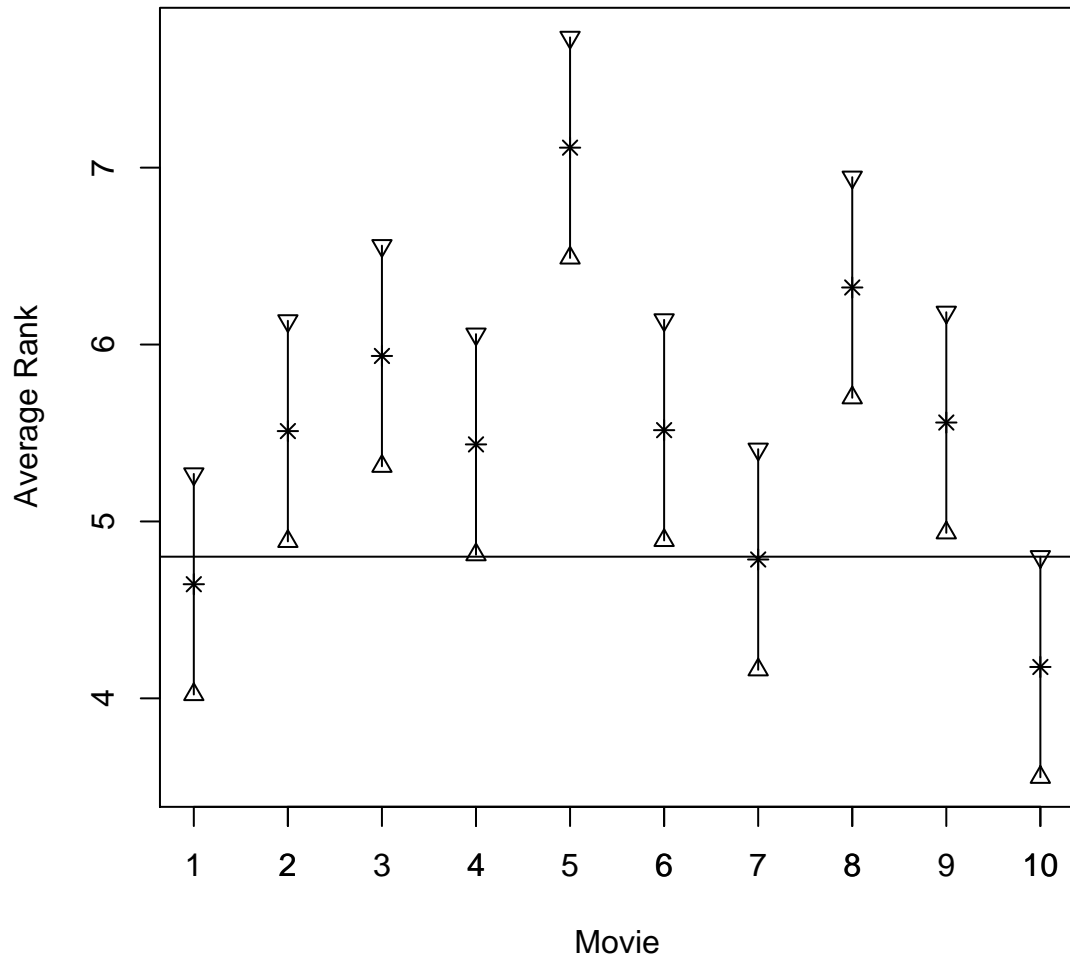


Figure 2: Dendrogram from cluster analysis based on multiple comparisons of $k = 10$ blockbuster movies in Dutch cinema theatres in 2007.

