Asymptotic (k-1)-mean significance levels of the multiple comparisons method based on Friedman's test

Citation for published version (APA):

Document status and date:
Published: 01/01/1977

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 11. Sep. 2020
Memorandum COSOR 77-22

Asymptotic \((k-1)\)-mean significance levels
of the multiple comparisons method
based on Friedman's test

by

J.H. Oude Voshaar

Eindhoven, October 1977

The Netherlands
Asymptotic \((k-1)\) mean significance levels of the multiple comparisons method based on Friedman's test

by

J.H. Oude Voshaar

1. Introduction

This paper should be regarded as a continuation of Oude Voshaar [2], but now we shall treat a multiple comparison procedure for a different model: Let \(\{x_{ij}; i = 1, \ldots, k; j = 1, \ldots, n\}\) be independent random variables, where \(x_{ij}\) has a continuous distribution function \(F_{ij}\) and there exist numbers \(\theta_1, \ldots, \theta_k, \beta_1, \ldots, \beta_n\) and a distribution function \(F\), such that

\[ F_{ij}(x) = F(x - \theta_i - \beta_j). \]

The \(\beta\)'s are called block parameters and the null hypothesis \(H_0: \theta_1 = \ldots = \theta_k\) is often tested by Friedman's test.

From this test a method for pairwise comparisons can be derived (see: Miller [1], page 172-178). Let \(\bar{r}_{ij}\) denote the rank of \(x_{ij}\) among \(x_{1j}, \ldots, x_{kj}\) and let \(\bar{r}_i\) be defined by

\[ \bar{r}_i := \frac{1}{n} \sum_{j=1}^{n} r_{ij}. \]

Then under the null hypothesis we have for \(n\) large (\(n \to \infty\)):

\[ P[|\bar{r}_i - \bar{r}_{i'}| < q_k \frac{\sqrt{k(k+1)}}{12n} \text{ for all } i, i' \in \{1, \ldots, k\}] = 1 - \alpha \]

where \(q_k^\alpha\) denotes the upper \(\alpha\) point of the distribution of the range of \(k\) independent standard normal variables.

However, if \(\theta_1 = \ldots = \theta_{k-1}\) and \(\theta_k = \theta_1 + c (c \neq 0)\), what will be in that case the value of \(a(F, c)\), defined by:

\[ a(F, c) := \lim_{n \to \infty} P[\max_{1 \leq i, i' \leq k-1} |\bar{r}_i - \bar{r}_{i'}| \geq q_k \frac{\sqrt{k(k+1)}}{12n}]. \]

In other words: What is the probability of concluding some of the \(\theta\)'s to be different, which in fact are equal; and our main question will be:

\[ a(F, c) \leq \alpha \text{ for all } F \text{ and } c? \]
2. The supremum of \( \alpha(F,c) \)

In order to answer this last question, we shall compute the supremum of \( \alpha(F,c) \) over \( F \) and \( c \).

We define \( p, q \) and \( r \) (which are functions of \( F \) and \( c \)) by:

\[
\begin{align*}
p &= \int (F(x - c)dF(x), \\
q &= \int F(x - c)dF(x), \\
r &= \int F(x)F(x - c)dF(x).
\end{align*}
\]

Then the vectors \( (\bar{r}_{1j}, \ldots, \bar{r}_{kj}) \) for \( j = 1, \ldots, n \) are independent and identically distributed, so we can conclude that \( (\bar{r}_{1}, \ldots, \bar{r}_{k}) \) has an asymptotically normal distribution for \( n \to \infty \).

From the formulas (2.5) and (2.6) of \([2]\) we can find \( \text{var} \bar{r}_{ij} \) and \( \text{cov}(\bar{r}_{ij}, \bar{r}_{ij}') \) for \( i, i' \in \{1, \ldots, k-1\} \) by substitution of \( n = 1 \).

Since \( \bar{r}_{ij} \) and \( \bar{r}_{ij}' \), are independent for \( j \neq j' \), we have for \( i, i' \in \{1, \ldots, k-1\} \) and \( i \neq i' \):

\[
\begin{align*}
\text{var} \bar{r}_{ij} &= \frac{1}{n}\left(\frac{1}{12}k^2 + (2r - p - \frac{1}{6})k + 3p - p^2 - 4r\right) \\
\text{cov}(\bar{r}_{ij}, \bar{r}_{ij}') &= \frac{1}{n}\left(-\frac{1}{12}k + 3p - p^2 - 2r\right).
\end{align*}
\]

Hence (the proof exactly parallels the derivation of (2.9) in \([2]\)):

\[
(2.1) \quad \alpha(F,c) = P\left[q_{k-1} > q_k\sqrt{\frac{1}{\frac{1}{12}(k^2 + k)}\frac{1}{\frac{1}{12}k^2 + (2r - p - \frac{1}{12})k}}\right].
\]

Since \( \frac{5}{24} \) is the supremum of \( 2r - p \) over all \( F \) and \( c \) (see \([2]\), section 4), we have:

\[
\sup_{F,c} \alpha(F,c) = P\left[q_{k-1} > q_k\sqrt{\frac{k^2 + k}{\frac{1}{2}k^2 + \frac{3}{2}k}}\right].
\]

So we find:

**Table 2.1:** \( \sup_{F,c} \alpha(F,c) \) for several values of \( \alpha \) and \( k \).

<table>
<thead>
<tr>
<th>( k )</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>12</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = .01 )</td>
<td>.0060</td>
<td>.0084</td>
<td>.0096</td>
<td>.0101</td>
<td>.0105</td>
<td>.0107</td>
<td>.0108</td>
<td>.0109</td>
<td>.0110</td>
<td>.0110</td>
<td>.0109</td>
</tr>
<tr>
<td>( .025 )</td>
<td>.0141</td>
<td>.0198</td>
<td>.0227</td>
<td>.0242</td>
<td>.0251</td>
<td>.0257</td>
<td>.0260</td>
<td>.0263</td>
<td>.0265</td>
<td>.0267</td>
<td>.0267</td>
</tr>
<tr>
<td>( .05 )</td>
<td>.0271</td>
<td>.0380</td>
<td>.0435</td>
<td>.0457</td>
<td>.0483</td>
<td>.0498</td>
<td>.0506</td>
<td>.0511</td>
<td>.0518</td>
<td>.0523</td>
<td>.0524</td>
</tr>
<tr>
<td>( .10 )</td>
<td>.0530</td>
<td>.0738</td>
<td>.0843</td>
<td>.0904</td>
<td>.0942</td>
<td>.0967</td>
<td>.0985</td>
<td>.0997</td>
<td>.1013</td>
<td>.1025</td>
<td>.1032</td>
</tr>
</tbody>
</table>
From table 2.1 we see that \(a(F,c)\) may be larger than \(\alpha\), but the exceedance will never be large. Once having this result, another question arises: If we define \(a(F)\) by:

\[
a(F) := \sup_{c} a(F,c),
\]

which conditions on \(F\) will be sufficient to guarantee \(a(F) \leq \alpha\)? In the next section we shall try to answer this question.

3. Conditions on \(F\) such that \(a(F) \leq \alpha\)

We shall use Van Zwet's convex order relation for distribution functions, defined by:

\[
F < G \Leftrightarrow \text{F convex on the support of F}
\]

(where we assume that \(F\) and \(G\) are elements of the class \(F\) defined in [2]). \(F < G\) should be interpreted as: \(G\) is more skewed to the right than \(F\).

If we denote furthermore \(F^*\), when \(F^*(x) = 1 - F(-x)\), than we can say \(F\) is less skewed than \(G\) if \(G^* \leq F < G\) or \(G < F \leq G^*\).

Now, since \(a(F,c)\) is an increasing function of \(2r-p\), from theorem (5.2) in [2] we can conclude:

**Theorem 3.1.** If \(F\) is less skewed than \(G\), then \(a(F) \leq a(G)\).

If we take for \(G\) the negative exponential distribution, then "\(F\) less skewed than \(G\)" is equivalent with: \(\log F\) and \(\log(1 - F)\) both concave. So we have the following application of theorem 3.1 (since \(2r-p \leq \frac{3}{16}\) for the negative exponential distribution):

**Theorem 3.2.** If \(\log F\) and \(\log(1 - F)\) both concave, then:

\[
a(F) \leq P[q_{k-1}] > q_k^\alpha \sqrt{\frac{k^2 + k}{k^2 + \frac{5}{4}k}}
\]

which is smaller than \(\alpha\) for the usual values of \(\alpha\) and \(k\), as shown in the following table.
Table 3.1: Supremum of $\alpha(F)$ when log $F$ and log$(1-F)$ both concave.

<table>
<thead>
<tr>
<th>$k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>12</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$ = .01</td>
<td>.0047</td>
<td>.0067</td>
<td>.0078</td>
<td>.0084</td>
<td>.0088</td>
<td>.0091</td>
<td>.0093</td>
<td>.0095</td>
<td>.0097</td>
<td>.0098</td>
<td>.0100</td>
</tr>
<tr>
<td>.025</td>
<td>.0115</td>
<td>.0165</td>
<td>.0191</td>
<td>.0207</td>
<td>.0217</td>
<td>.0224</td>
<td>.0229</td>
<td>.0233</td>
<td>.0239</td>
<td>.0243</td>
<td>.0247</td>
</tr>
<tr>
<td>.05</td>
<td>.0230</td>
<td>.0326</td>
<td>.0378</td>
<td>.0409</td>
<td>.0429</td>
<td>.0443</td>
<td>.0454</td>
<td>.0462</td>
<td>.0473</td>
<td>.0482</td>
<td>.0491</td>
</tr>
<tr>
<td>.10</td>
<td>.0465</td>
<td>.0653</td>
<td>.0752</td>
<td>.0813</td>
<td>.0853</td>
<td>.0881</td>
<td>.0901</td>
<td>.0917</td>
<td>.0939</td>
<td>.0959</td>
<td>.0976</td>
</tr>
</tbody>
</table>

Final note: As in table 2.1 sup $\alpha(F,c)$ does not exceed $\alpha$ very much, the results of section 2 do not appear to be alarming to a practical statistician, the more so as $\alpha(F)$ is smaller than $\alpha$ for a large class of distribution functions (see theorem 3.2).

However, a more serious disadvantage of the method, based on Friedman's test, is the fact that the distribution of $(\bar{r}_i, \bar{r}_j)$ (on which the comparison of $\theta_i$ and $\theta_j$ is based) depends not only on $\theta_i$ and $\theta_j$, but also on the other $\theta$'s.

References
