Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 138 (2008) 1521 – 1527
journal of statistical planning and inference
www.elsevier.com/locate/jspi

A two-step rejection procedure for testing multiple hypotheses

Jianjun (David) Li

Merck Research Laboratories, 351 N. Sumneytown Pike, North Wales, PA 19454, USA

Received 23 May 2006; accepted 2 April 2007
Available online 10 August 2007

Abstract

This paper considers p-value based step-wise rejection procedures for testing multiple hypotheses. The existing procedures have used constants as critical values at all steps. With the intention of incorporating the exact magnitude of the p-values at the earlier steps into the decisions at the later steps, this paper applies a different strategy that the critical values at the later steps are determined as functions of the p-values from the earlier steps. As a result, we have derived a new equality and developed a two-step rejection procedure following that. The new procedure is a short-cut of a step-up procedure, and it possesses great simplicity. In terms of power, the proposed procedure is generally comparable to the existing ones and exceptionally superior when the largest p-value is anticipated to be less than 0.5.

Keywords: Bonferroni inequality; Multiple tests; Strong control of family-wise error rate

1. Introduction

Consider testing m (null) hypotheses $H_1, \ldots, H_m$ simultaneously based on their p-values $P_1, \ldots, P_m$. Assume that $P_1, \ldots, P_m$ are obtained from independent and continuous test statistics. Several sequential rejection procedures with the property of strong control of the family wise error rate (FWER) have been available in the literature. These include the procedures developed by Holm (1979), Hommel (1988), Hochberg (1988), Rom (1990), and Liu (1997). These procedures use a set of critical values and carry out sequential rejections by comparing the ordered p-values and the corresponding critical values.

The procedure presented by Holm (1979) is a typical step-down procedure, which proceeds by comparing the smallest p-value to the corresponding critical value at the first step, and then the second smallest p-value to the corresponding critical value at the second step, and so on. At each step, reject the hypothesis associated with the ordered p-value if the ordered p-value is smaller than or equal to the critical value. Stop once an ordered p-value is larger than the corresponding critical value and then accept all the remaining hypotheses. The critical value used at the jth step in the Holm's procedure is $c_j = \alpha/(m - j + 1)$, where α is the pre-determined targeted FWER.

Hochberg (1988) proposed a step-up testing procedure that proceeds in a similar fashion as the Holm's procedure, but starts with the largest p-value towards the smallest p-value. Accept one hypothesis at a time, and stop once an ordered p-value is smaller than or equal to the corresponding critical value. At the time of stopping, reject the remaining hypotheses.

E-mail address: jianjun_li@merck.com.

0378-3758/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2007.04.032

Of course, all early ones have been accepted. In the Hochberg's procedure, the critical value is $c_j = \alpha/j$. The Rom's procedure (1990) is also a step-up procedure, but with different critical values of $c_1 = \alpha, \quad c_2 = \frac{\alpha}{2}, \quad c_3 = \frac{\alpha}{3} + \frac{\alpha^2}{12}$ etc.

The procedure proposed by Hommel (1988) is relatively complicated. To use this procedure, at each step, say the jth step, compare $P_{(m+i-j)}$ to $\alpha i/j$ to see if $P_{(m+i-j)} > \alpha i/j$, for $i = 1, \ldots, j$, where $P_{(1)} \leqslant \cdots \leqslant P_{(m)}$ are the ordered p-values. Stop at the first step where $P_{(m+i-j)} \leqslant \alpha i/j$ for some $1 \leqslant i \leqslant j$. At the time of stopping, reject all hypotheses with p-value being $\leqslant \alpha/j$.

Liu (1997) developed a general step-wise testing procedure which starts by comparing the j_0th smallest p-value $P_{(j_0)}$ to the corresponding critical value c_{j_0}, where j_0 is a pre-determined number. If $P_{(j_0)} \leqslant c_{j_0}$, reject all hypotheses with p-value being $\leqslant P_{j_0}$ and then continue in a step-down manner. Otherwise, accept all hypotheses with p-value being $\geqslant P_{j_0}$ and then continue in a step-up manner. As concluded by Liu (1997), when j_0 is determined to be m, the procedure becomes the Rom's procedure, and it is the most powerful one among the class of step-wise procedures considered for practical situations.

Among these procedures, the Holm's procedure is the least powerful, because it is based on the Bonferonni inequality. Both the Rom's procedure and Hommel's procedure are more powerful than the Hochberg's procedure due to the fact that sharp inequalities (or equalities) are used in both procedures, however, the power improvement is negligible compared to their complexities.

The aforementioned procedures, although different from one to another, all have the same characteristics: the critical values at the later steps are irrelevant to the exact value of the p-values used at the earlier steps. For example, if there are two null hypotheses to be tested, all step-up procedures compare the largest p-value $P_{(2)}$ to α first and then the smallest p-value $P_{(1)}$ to $\alpha/2$. If $P_{(2)} > \alpha$, the decision on $H_{(1)}$ is purely determined by whether or not the inequality $P_{(1)} > \alpha/2$ holds, and the exact magnitude of $P_{(2)}$ is ignored even though it provides valuable information for the decision on $H_{(1)}$. The analogy holds for the step-down procedure and other step-wise procedures.

In this paper, we consider step-up procedures but with a revised strategy that critical values at the later steps are determined as functions of p-values used at the earlier steps so that the magnitude of p-values at the earlier steps can be incorporated into decisions at the later steps. Towards this end, we have obtained a quite simple equality, and from there derived a two-step rejection procedure where the critical value at the first step is α, and at the second step is a simple linear function of the p-value used at the first step. The new procedure is a short-cut of a step-up procedure with all critical values from the second step and later steps being the same. Therefore it possesses great simplicity. In terms of power, simulations have shown that it is comparable to the existing ones in general. But in situations where the largest p-value is anticipated to be less than 0.5, which is often the case in confirmatory research, the proposed procedure is far superior to the others.

Hochberg and Tamhane (1987) discussed distinction of exploratory research and confirmatory research, two typical types of research in practice, in the context of multiple testing. In confirmatory research, where a strong control of FWER is usually required, the set of hypotheses is selected prior to conducting the research. Each hypothesis, as a requirement to be selected, is anticipated to be confirmed as “false” at the end of research. This implies that, in statistical language, the p-value of each hypothesis cannot be large. In this type of research, the proposed procedure outperforms the others because that it capitalizes on the fact that the research is confirmatory. Of course, the proposed procedure is not in a favorable position for exploratory research where the set of hypotheses is not pre-specified and/or the researchers do not have a good sense on whether the research itself by design has capability to conclude each hypothesis of interest as “false” at the end.

The details of the proposed procedure are included in Section 2, where the new equality is established and the proposed procedure is stated. The formula for computing adjusted p-value is also presented. In Section 3, we present two examples to illustrate the use of the proposed procedure, as contrast to other available procedures. A detailed numerical comparison between the propose procedure and the others is in Section 4. We have some final remarks in Section 5.

2. A two-step rejection procedure

2.1. A new equality

Recall that $P_1, \ldots, P_m$ are p-values and they are uniform $U(0, 1)$ variables under the global null hypothesis $\cap\{H_j : j = 1, \ldots, m\}$. For a subset $\{P_{1j}, \ldots, P_{jj}\} \subseteq \{P_1, \ldots, P_m\}$, let $P_{1:j}, \cdots, P_{j:j}$ be the ordered p-values of

$P_{1j}, \ldots, P_{jj}$. Assume that $P_1, \ldots, P_m$ are independent. To construct a step-up procedure, we need the following type of inequality (Liu, 1996):

$$\Pr(P_{i:j} > c_{j-i+1}, i = 1, \ldots, j) \geqslant 1 - \alpha \tag{1}$$

for $j = 1, \ldots, m$. Of course, $c_1 = \alpha$ (taking $j = 1$). For $c_i, i \geqslant 2$, we need the following result.

Lemma. *For $j \geqslant 2$, let $U_{(1)} \leqslant \cdots \leqslant U_{(j)}$ be the ordered values of $U_1, \ldots, U_j$, where $U_1, \ldots, U_j$ are uniform $U(0, 1)$ variables and independent from each other. Then*

$$\Pr\left(U_{(j)} > \alpha, U_{(i)} > \frac{1 - U_{(j)}}{1 - \alpha}\alpha,\ i = 1, \ldots, j - 1\right) = 1 - \alpha. \tag{2}$$

Proof. By definition, $U_{(1)} \leqslant \cdots \leqslant U_{(j)}$, and then

$$U_{(i)} > \frac{1 - U_{(j)}}{1 - \alpha}\alpha, \quad i = 1, \ldots, j - 1 \quad \Leftrightarrow \quad U_{(1)} > \frac{1 - U_{(j)}}{1 - \alpha}\alpha.$$

Therefore we just need to show

$$\Pr\left(U_{(j)} > \alpha, U_{(1)} > \frac{1 - U_{(j)}}{1 - \alpha}\alpha\right) = 1 - \alpha.$$

The joint density of $(U_{(1)}, U_{(j)})$ is $f(s, t) = j(j - 1)(s - t)^{j-2}, 0 < t \leqslant s < 1$. Therefore

$$\Pr\left(U_{(j)} > \alpha, U_{(1)} > \frac{1 - U_{(j)}}{1 - \alpha}\alpha\right) = \int_\alpha^1 \int_{(1-s)/(1-\alpha)\alpha}^{s} j(j - 1)(s - t)^{j-2}\,\mathrm{d}t\,\mathrm{d}s = 1 - \alpha.$$

The proof is complete. □

2.2. The two-step rejection procedure

Under the global null hypothesis $\cap\{H_i : i = 1, \ldots, m\}$, $P_{1j}, \ldots, P_{jj}$ are independent and uniform $U(0, 1)$ variables. Based on the lemma, we can take $c_i = \alpha(1 - P_{(m)})/(1 - \alpha), i \geqslant 2$, in (1) since $\alpha(1 - P_{(m)})/(1 - \alpha) \leqslant \alpha(1 - P_{j:j})/(1 - \alpha)$ for any $j \geqslant 2$ and $\{P_{1j}, \ldots, P_{jj}\} \subseteq \{P_1, \ldots, P_m\}$. Noting that $c_2 = \cdots = c_m$, the step-up procedure formed by (1) naturally becomes the following simple two-step procedure.

Procedure R. *Step* 1. Reject all H_i's if $P_{(m)} \leqslant \alpha$. Otherwise, accept the hypothesis associated with $P_{(m)}$ and go to Step 2.

Step 2. Reject any remaining H_i with $P_i \leqslant (1 - P_{(m)})/(1 - \alpha)\alpha$.

Theorem. *If $P_1, \ldots, P_m$ are independent, the procedure R provides a strong control of the FWER in that the probability of at least one erroneous rejection is less than or equal to α, regardless of which and how many of the H_i are true.*

The theorem can be established based on its derivation above using the well-known close principle (Marcus et al., 1976). An alternative and direct proof is stated below.

Proof of the Theorem. We need to show that

$$\text{FWER} = \Pr(\text{at least one erroneous rejection}) \leqslant \alpha$$

under any "true" and "false" configuration among m hypotheses. Consider a configuration with m_0 true null hypotheses and $m - m_0$ false hypotheses. If $m_0 = 0$, all null hypotheses are false and then FWER = 0. If $m_0 > 0$, note that

$$1 - \text{FWER} = \Pr(\text{all } m_0 \text{ true null hypotheses are accepted}).$$

Without loss of generality, assume that the m_0 true null hypotheses have p-values $Q_1, \ldots, Q_{m_0}$ and $Q_{(1)} \leqslant \cdots \leqslant Q_{(m_0)}$. All m_0 true null hypotheses are accepted by the procedure R if and only if $P_{(m)} > \alpha$ and $Q_{(1)} > \alpha(1 - P_{(m)})/(1 - \alpha)$. Note that $P_{(m)} \geqslant Q_{(m_0)}$,

$$1 - \text{FWER} = \Pr\left(P_{(m)} > \alpha,\ \ Q_{(1)} > \alpha\frac{1 - P_{(m)}}{1 - \alpha}\right) \geqslant \Pr\left(Q_{(m_0)} > \alpha,\ \ Q_{(1)} > \alpha\frac{1 - Q_{(m_0)}}{1 - \alpha}\right).$$

If $m_0 = 1$, $Q_1 > \alpha \Leftrightarrow Q_1 > \alpha(1 - Q_1)/(1 - \alpha)$ and $1 - \text{FWER} \geqslant \Pr(Q_1 > \alpha) = 1 - \alpha$. When $m_0 > 1$, $1 - \text{FWER} \geqslant 1 - \alpha$ following the lemma. The proof is complete. □

Remark 1. The proposed procedure can be thought of as a short-cut of the step-up procedure. For any step-up procedure, it is desired to have c_i, $i \geqslant 2$, as large as possible (the largest possible value for c_1 is α). In the proposed procedure, $c_i = \alpha$ if $P_{(m)} = \alpha$. As a result, c_i can be quite close to α if $P_{(m)}$ is larger than α but not too much. Therefore, in the cases where we nearly reject all hypotheses at the first step, we are able to reject more at the second step.

Remark 2. The critical values at the second step and later steps in the proposed procedure, if it is considered as a step-up procedure, are just a simple linear function of $P_{(m)}$. In theory, one can use different functional forms of $P_{(m)}$, for example, a quadratic form rather than linear, however, such form will not have simplicity as the proposed one has, and may not be as powerful as the proposed one is. Also it is natural to ask whether it will result in a more powerful procedure if c_3 is taken as a linear function of $P_{(m)}$ and $P_{(m-1)}$, c_4 is taken as a linear function of $P_{(m)}$, $P_{(m-1)}$ and $P_{(m-2)}$, etc. The answer is no. In fact, if one wants to use the linear form for critical values at the later steps, it will end up exactly with the proposed procedure, i.e., $c_i = \alpha(1 - P_{(m)})/(1 - \alpha)$ for $i = 2, \ldots, m$. In detail, with a desire of using the linear form of all earlier p-values, we need the following inequality (Liu, 1996):

$$\Pr(P_{j:j} > \alpha, P_{j-k+1:j} > a_{k,k}P_{j:j} + a_{k,k-1}P_{j-1:j} + \cdots + a_{k,2}P_{j-k+2:j} + a_{k,1}, k = 2, \ldots, j) \geqslant 1 - \alpha$$

for $j = 2, \ldots, m$. Under the constraint of having

$$a_{j,j}P_{j:j} + \cdots + a_{j,2}P_{2:j} + a_{j,1} = a_{j-1,j-1}P_{j:j} + \cdots + a_{j-1,2}P_{3:j} + a_{j-1,1}$$

if $P_{2:j} = a_{j-1,j-1}P_{j:j} + \cdots + a_{j-1,2}P_{3:j} + a_{j-1,1}$, for $j = 2, \ldots, m$, it is easy to show that $a_{j,j} = -a_{j,1} = -\alpha/(1 - \alpha)$ and $a_{j,j-1} = \cdots = a_{j,2} = 0$ for $j = 2, \ldots, m$.

2.3. Adjusted p-values

For single hypothesis testing problems, the p-value is the smallest significance level at which the hypothesis can be rejected. For multiple hypothesis testing problems, the (marginal) p-value for each hypothesis is not the smallest significance level at which that hypothesis can be rejected. In the literature, the smallest significance level at which a hypothesis can be rejected is called adjusted p-value (see Dunnett and Tamhane, 1992). Computing adjusted p-value can be complicated even for a relatively simple multiple hypothesis testing procedure. But for the step-up procedure considered in this paper, the adjusted p-value can be easily obtained as

$$\tilde{P}_i = \frac{P_i}{P_i + 1 - P_{(m)}}, \quad i = 1, \ldots, m.$$

Once the adjusted p-value $\tilde{P}_i$ is determined, the hypothesis H_i can be tested at any specified level α by simply rejecting H_i if $\tilde{P}_i \leqslant \alpha$. Using the adjusted p-value is equivalent to using the procedure R to test each hypothesis. The proof of such statement is trivial, hence omitted.

3. Examples

Example 1. In an example used by Hommel (1988), there were 10 statistical tests preformed and the associated ordered p-values were: $P_{(1)} = 0.0021$, $P_{(2)} = 0.0074$, $P_{(3)} = 0.0093$, $P_{(4)} = 0.0106$, $P_{(5)} = 0.0121$, $P_{(6)} = 0.0218$, $P_{(7)} = 0.0238$, $P_{(8)} = 0.0352$, $P_{(9)} = 0.0466$, and $P_{(10)} = 0.0605$. Using a typical FWER of $\alpha = 0.05$, the Hommel's procedure rejects the hypotheses corresponding to $P_{(1)}$, $P_{(2)}$, and $P_{(3)}$. Both the Hochberg's procedure and Rom's procedure only reject the hypothesis corresponding to $P_{(1)}$.

When the proposed procedure is applied, we accept the hypothesis associated with $P_{(10)}$ at the first step since $P_{(10)} = 0.0605 > 0.05$. By comparing the rest of p-values to $(1 - P_{(10)})/(1 - \alpha)\alpha = 0.049$ at the second step, we conclude that all remaining nine hypotheses are rejected.

Example 2. We consider an example used by Liu (1997). Six hypotheses were tested and corresponding ordered p-values were $P_{(1)} = 0.007$, $P_{(2)} = 0.011$, $P_{(3)} = 0.012$, $P_{(4)} = 0.020$, $P_{(5)} = 0.190$, and $P_{(6)} = 0.250$. Using $\alpha = 0.05$, the proposed procedure rejects the hypotheses corresponding to $P_{(1)}$, $P_{(2)}$, $P_{(3)}$ and $P_{(4)}$. The Hommel's procedure or Hochberg's procedure or Rom's procedure rejects the hypotheses corresponding to $P_{(1)}$, $P_{(2)}$, and $P_{(3)}$.

4. Power comparisons

A simple comparison is done by directly looking at the critical values in the proposed procedure and all other procedures. When $P_{(m)} \leqslant 0.5 + 0.5\alpha$, $c_i = (1 - P_{(m)})/(1 - \alpha)\alpha \geqslant \alpha/2$ for $i \geqslant 2$. Therefore, the critical value in the proposed procedure is larger than the corresponding one in the Hochberg's procedure or Rom's procedure at all steps except at the first step where the critical value is the same. In particular, when $P_{(m)}$ is close to α, all critical values in the proposed procedure are equal to or close to α and superiority of the proposed procedure becomes outstanding. The same conclusion can be reached when the proposed procedure is compared to the Hommel's procedure.

In situations where the magnitude of $P_{(m)}$ is uncertain, the performance of the proposed procedure is not clear compared to others. Therefore, we numerically evaluate the power. Consider a common normal model, where for $i = 1, \ldots, m$, the test statistic is $X_i \sim \mathrm{N}(d_i, 1)$ for the hypothesis $H_i\colon d_i = 0$. The two-sided p-value can then be computed as $2 * [1 - \Phi(|X_i|)]$, where $\Phi(x)$ is the distribution function of the standard normal distribution. For a given parameter configuration $(d_1, \ldots, d_m)$ or equivalently marginal power configuration $(q_1, \ldots, q_m)$, the probability of rejecting each hypothesis, rejecting at least one hypothesis, rejecting at least two hypotheses, etc., by different multiple testing procedures, can be calculated through simulations. For given d_i, the marginal power q_i is defined as the probability of rejecting H_i at level α marginally without taking other hypotheses into consideration, i.e.

$$q_i = \Pr\left(|X_i| \geqslant \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) = 1 - \Phi\left[\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - d_i\right)\right] + \Phi\left[-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - d_i\right)\right].$$

Table 1 presents the results for $m = 3$ and $\alpha = 0.05$. The results for other cases of m and α are similar, hence omitted.

The results in Table 1 show that the proposed procedure is generally more powerful than the Hochberg's procedure, Hommel's procedure, and Rom's procedure, when the marginal power of each hypothesis is not very low ($> 20\%$). Also note that the powers of the Hochberg's procedure, Hommel's procedure, and Rom's procedure are quite close, with the Hommel's procedure and Rom's procedure being more powerful, which is consistent with results obtained by Dunnett and Tamhane (1993).

5. Conclusions and discussions

In this paper, we consider step-up procedures with a new strategy that the critical values at the later steps in the procedures are determined as functions of p-values used at the earlier steps, with the intention of incorporating the magnitude of these p-values into decisions at the later steps. As a result, we have constructed a step-up procedure with the critical values at the second step and all later steps being equal to a linear function of $P_{(m)}$, hence, a two-step procedure. By comparing the critical values directly or by evaluating power through simulations, we demonstrate that the proposed procedure generally outperforms others if the marginal power of each hypothesis is not very low. In particular, the proposed procedure is remarkably superior when the largest p-value is anticipated to be less than 0.5. In confirmatory research (Hochberg and Tamhane, 1987), each hypothesis is pre-selected and expected to be confirmed as "false" through conducting the research. In these situations, the proposed procedure is particularly useful since it

Table 1
Probability (%) of rejecting each hypothesis, $\geqslant 1$ hypothesis, $\geqslant 2$ hypotheses by different testing procedures for given marginal power configuration ($\alpha = 0.05$)

Marginal power of (H1, H2, H3)		Hochberg's	Hommel's	Rom's	Proposed procedure
(20, 20, 20)	H1	11.1	11.3	11.2	12.7
	$\geqslant 1$	27.7	28.1	27.9	30.9
	$\geqslant 2$	4.9	4.9	4.9	6.3
(50, 50, 50)	H1	39.6	39.9	39.7	44.5
	$\geqslant 1$	71.2	71.9	71.4	78.5
	$\geqslant 2$	35.3	35.3	35.3	42.4
(80, 80, 80)	H1	76.4	76.5	76.4	78.8
	$\geqslant 1$	96.4	96.7	96.5	98.4
	$\geqslant 2$	81.6	81.6	81.6	87.1
(20, 20, 50)	H1	12.2	12.4	12.3	14.4
	H3	35.1	35.3	35.2	36.8
	$\geqslant 1$	46.5	47.1	46.8	50.3
	$\geqslant 2$	11.1	11.1	11.1	13.4
(20, 20, 80)	H1	13.5	13.6	13.5	15.2
	H3	67.5	67.7	67.7	67.7
	$\geqslant 1$	72.8	73.2	73.0	74.1
	$\geqslant 2$	18.4	18.4	18.4	20.8
(50, 50, 80)	H1	42.4	42.5	42.4	46.4
	H3	71.4	71.6	71.5	75.1
	$\geqslant 1$	85.5	86.0	85.6	90.1
	$\geqslant 2$	50.7	50.7	50.7	57.7
(20, 50, 50)	H1	13.8	13.9	13.8	16.8
	H2	37.0	37.3	37.1	40.1
	$\geqslant 1$	60.6	61.3	60.8	65.9
	$\geqslant 2$	20.2	22.2	22.2	26.0
(20, 80, 80)	H1	17.3	17.3	17.3	19.2
	H2	71.3	71.5	71.4	72.3
	$\geqslant 1$	90.0	90.4	90.1	91.5
	$\geqslant 2$	57.1	57.1	57.1	59.5
(50, 80, 80)	H1	45.5	45.6	45.6	48.6
	H2	73.8	74.0	73.9	76.8
	$\geqslant 1$	92.8	93.2	92.9	95.7
	$\geqslant 2$	68.4	68.4	68.4	74.7
(20,50, 80)	H1	15.3	15.4	15.3	17.9
	H2	39.2	39.4	39.3	41.5
	H3	69.2	69.4	69.3	70.9
	$\geqslant 1$	79.9	80.5	80.1	82.9
	$\geqslant 2$	35.8	35.8	35.8	39.4

capitalizes on the fact that the research is confirmatory and the p-value of each hypothesis is expected to be small at the end.

The proposed procedure is developed with the independent p-values in the spirit to the development of the Holm's, Hochberg's, Hommel's, and Rom's procedure. With some parametric assumptions on test statistics, the Hochberg's procedure and Hommel's procedure are still valid under dependence (Sarkar and Chang, 1997). It is unclear whether the lemma and the proposed procedure are valid under these dependent parametric models since the methods used by Sarkar and Chang (1997) are not applicable here. Further exploration on the issue is merited.

Acknowledgments

The author is very thankful to two referees and the associate editor for their thorough review and helpful comments.

References

Dunnett, C., Tamhane, A., 1992. A step-up multiple test procedure. J. Amer. Statist. Assoc. 87, 162–170.
Dunnett, C., Tamhane, A., 1993. Power comparisons of some step-up multiple test procedures. Statist. Probab. Lett. 16, 55–58.
Hochberg, Y., 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75, 800–803.
Hochberg, Y., Tamhane, A.C., 1987. Multiple Comparison Procedures. Wiley, New York.
Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Statisit. 6, 65–70.
Hommel, G., 1988. A stagewise rejective multiple test procedure based on a modified Bonferroni test. Biometrika 75, 383–386.
Liu, W., 1996. Multiple tests pf a non-hierarchical finite family of hypotheses. J. Roy. Statist. Soc. 58, 455–461.
Liu, W., 1997. On stepwise tests when test statistics are independent. Australian J. Statist. 39, 169–177.
Marcus, R., Peritz, E., Gabriel, K.R., 1976. On closed testing procedures with special reference to ordered analysis of variance. Biometrika 63, 655–660.
Rom, D.M., 1990. A sequentially rejective test procedure based on a modified Bonferroni inequality. Biometrika 77, 663–665.
Sarkar, Chang, 1997. The Simes method for multiple hypothesis testing with positively dependent test statistics. J. Amer. Statist. Assoc. 92, 1601–1608.