A sharper Bonferroni procedure for multiple tests of significance

By YOSEF HOCHBERG

*Department of Statistics, School of Mathematical Sciences, Tel Aviv University, Ramat-Aviv, 69978,
Israel*

SUMMARY

A simple procedure for multiple tests of significance based on individual p-values is derived. This simple procedure is sharper than Holm's (1979) sequentially rejective procedure. Both procedures contrast the ordered p-values with the same set of critical values. Holm's procedure rejects an hypothesis only if its p-value and each of the smaller p-values are less than their corresponding critical-values. The new procedure rejects all hypotheses with smaller or equal p-values to that of any one found less than its critical value.

Some key words: Familywise error rate; Hypotheses-free association; P-values; Strong and weak control.

The Bonferroni inequality is often used when conducting multiple tests of significance to set an upper bound on the familywise error rate; see, for example, Hochberg & Tamhane (1987, pp. 3, 363). Let $P_1, \ldots, P_m$ be p-values corresponding to m statistics for testing hypotheses $H_1, \ldots, H_m$. If a specific hypothesis H_i is rejected when $P_i \leq \alpha/m$, then the Bonferroni inequality,

$$\text{pr}\left\{\bigcup_{i=1}^{m}(P_i \leq \alpha/m)\right\} \leq \alpha \quad (0 \leq \alpha \leq 1), \tag{1}$$

ensures that the probability of rejecting at least one hypothesis when all are true is no greater than α. This procedure actually guarantees the following stronger property. Let $H = \{H_1, \ldots, H_m\}$ and let H' be a subset of H. Denote by H_0 and by H_0' the intersection of the hypotheses in H and in H', respectively. The probability of rejecting H_0' when it is true is not greater than α for any subset H'. The first mentioned property is referred to as weak control of the familywise error rate, i.e.

$$\text{pr}_{H_o}(\text{rejecting any } H_i) \leq \alpha, \tag{2}$$

where 'rejecting any H_i' is equivalent to 'rejecting H_0'. The second property is referred to as strong control of the familywise error rate and can be expressed as

$$\text{pr}_{H_0'}(\text{rejecting } H_0') \leq \alpha, \tag{3}$$

for all $H' \subseteq H$.

Holm (1979) gave the following improved Bonferroni procedure. Let $P_{(1)}, \ldots, P_{(m)}$ be the ordered p-values and $H_{(1)}, \ldots, H_{(m)}$ be the corresponding hypotheses. Reject $H_{(i)}$ when, for all $j = 1, \ldots, i$,

$$P_{(j)} \leq \alpha/(m-j+1). \tag{4}$$

Holm's procedure controls the familywise error-rate in the strong sense. It is described in the following section as a procedure for testing all subset intersection hypotheses.

Simes (1986) discussed another 'improved Bonferroni procedure'. His procedure, however, provides a test only for H_0. According to this procedure H_0 is rejected when, for any $j = 1, \ldots, m$,

$$P_{(j)} \leq j\alpha/m. \tag{5}$$

Simes proved that this procedure has level α under H_0 when the p-values are independent. He also showed by simulation that the level under H_0 does not exceed α for a variety of multivariate normal and gamma test statistics.

Simes raised the problem of making statements on individual hypotheses but did not provide a solution. Hommel (1988) employed the closure principle to extend Simes's procedure for making statements on individual hypotheses. In the following a similar approach is taken and a new procedure is derived after some simplification.

Assume a collection $H = \{H_1, \ldots, H_m\}$ of hypotheses satisfying the condition of free-association (Holm, 1979). For any subset $H' \subseteq H$ of $m' \leq m$ individual hypotheses, order the corresponding p-values into $P_{(i_1)}, \ldots, P_{(i_{m'})}$. The extended Simes procedure rejects any subset intersection hypothesis H'_0 when, for all $H'' \supseteq H'$,

$$P_{(i_j)} \leq j\alpha/m'', \tag{6}$$

for any $H_{(i_j)} \in H''$, where m'' is the number of hypotheses in H''.

When the original Simes procedure has an α level familywise error rate under H_0, the extended Simes procedure will control the familywise error rate in the strong sense, i.e. under any H'_0. That follows from the closure principle of Marcus, Peritz & Gabriel (1976).

To compare the modified Simes procedure with Holm's procedure we write Holm's test for any arbitrary intersection hypothesis H'_0 as follows. Reject H'_0 when, for all $H'' \supseteq H'$

$$P_{(i_1)} \leq \alpha/m'', \tag{7}$$

where (i_1) is the index of the smallest p-value in H''.

The equivalence between (7) and (4) is readily obvious when identifying the rejection of an intersection hypothesis with the rejection of at least one of its components and recognizing that (4) is a necessary and sufficient condition for any subset containing $H_{(j)}$ to be rejected according to (7). If (7) is satisfied then (6) obviously follows. Hence, the extended Simes procedure given here for the family of all subset intersection hypotheses is more powerful than the corresponding Holm procedure for that family.

Next we discuss a simplified version of the extended Simes procedure for making inferences on individual hypotheses.

LEMMA. *For any* $i = m, m - 1, \ldots, 1,$ *if*

$$P_{(i)} \leq \alpha/(m - i + 1) \tag{8}$$

then Simes's procedure rejects all $H_{(i')}$ $(i' \leq i)$.

Proof. Under (8), the extended Simes procedure obviously rejects any H''_0 such that the smallest p-value in H'' is $P_{(i)}$. Next consider a different set H'' with exactly k $(1 \leq k \leq i - 1)$ hypotheses $H_{(i_1)}, \ldots, H_{(i_k)}$ whose p-values are smaller than $P_{(i)}$. Then under condition (8), the extended Simes procedure (6) will reject H''_0 since $P_{(i)}$ will be compared with the value

$$(1 + k)/(m - i + 1 + k), \tag{9}$$

which is smallest when $k = 0$. Comparing (8) with (4), we see that the given procedure is more powerful and simpler than Holm's procedure. According to the new procedure, one starts by examining the largest p-value $P_{(m)}$. If $P_{(m)} \leq \alpha$ then all hypotheses are rejected. If not, then $H_{(m)}$ cannot be rejected and one goes on to compare $P_{(m-1)}$ with $\frac{1}{2}\alpha$. If smaller, then all $H_{(i)}$ $(i = m - 1, \ldots, 1)$ are rejected. If not, then $H_{(m-1)}$ cannot be rejected and one proceeds to compare $P_{(m-2)}$ with $\frac{1}{3}\alpha$, etc., according to (8).

REFERENCES

HOCHBERG, Y. & TAMHANE, A. (1987). *Multiple Comparison Procedures*. New York: Wiley.
HOLM, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **6**, 65–70.

HOMMEL, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika* **75**, 383-6.

MARCUS, R., PERITZ, E. & GABRIEL, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* **63**, 655-60.

SIMES, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* **73**, 751-4.