

Neuendorf

Post Hoc Tests—A Primer

(So many to choose from. . .)

Post hoc tests (“after the fact” tests) allow assessments of comparisons among IV groups, after the overall ANOVA-type test of a given effect (main effect or interaction) has been executed.

1. Cumulative Type I Error (α)
Acknowledges that Type I error can accumulate across tests within a single analysis plan.

PC (per comparison) error—traditional α , assumes tests are independent

vs.

FW (familywise) error:

$$\alpha_{FW} = 1 - (1 - \alpha)^C$$

where C = # of orthogonal comparisons
(the calculation for nonorthogonal comparisons is more complex)

2. Planned contrasts vs. Post hocs

Planned contrasts:

A priori

Asks “Is this particular difference significant?”

Part of the design

α_{FW} is typically ignored; α is used, sometimes with a Bonferroni adjustment

Maximum #? Keppel says there is no agreement on this. . . $\sim df_A$. . . theory

Post hocs:

A posteriori (actually another name for post hocs. . .)

Asks “Which differences are significant?”

Not part of the design—often unexpected findings

α_{FW} controlled via special evaluation procedures

Maximum #? See p. 167 Keppel for a simple formula.

3. Post hoc “Types”

- A. *Unrestricted comparisons* after screening via “omnibus F”
 If F is significant, then may proceed
 Criticized by many, because it doesn’t really control α_{FW}
 The other types at least try—Keppel calls them “alpha-adjusted techniques”

Example stat procedure—**Fisher’s LSD** (as described above—
 omnibus F followed by unrestricted comparisons)

- B. *All comparisons*, pairwise and complex

Example stat procedure—**Scheffe’s test**

Uses regular F, with “special” critical value:

$$F_S = (k - 1)F_{crit}$$

where k = # of group means

- C. *Total set of pairwise comparisons*
 Possible # = $(a)(a-1)/2$

Example stat procedure—**Tukey test** (one of Tukey’s many contributions)

Finds the minimum pairwise difference between means that must be exceeded to be significant with the Tukey test:

$$\bar{d}_T = q_T \sqrt{\frac{MS_{S/A}}{n}}$$

where q_T = an entry in the table of the studentized range statistic

$MS_{S/A}$ = the error term for the overall ANOVA

n = the sample size for each group

Example stat procedures—see also the **Newman-Keuls procedure** and the similar **Duncan procedure** as described by Winer (1971)

- D. *Comparing one group to all others*
e.g., Control group vs. several experimental groups

Example stat procedure–**Dunnett test**

Calculates the control-experimental mean differences and compares them against a critical mean difference (\bar{d}_D) that must be exceeded to be significant:

$$\bar{d}_D = q_D \sqrt{\frac{2(MS_{S/A})}{n}}$$

where q_D = an entry in the Dunnett table (see Keppel)
 $MS_{S/A}$ = the error term for the overall ANOVA
 n = the sample size for each group

4. An all-purpose test? An easy way out? The **Bonferroni technique**:

$$\alpha_{\text{BON}} = \alpha_{\text{FW}} / \# \text{ of tests}$$

Hair et al. apply this to k-groups as a post hoc.
Keppel applies this to multiple planned contrasts.

OR, Garson's online version:

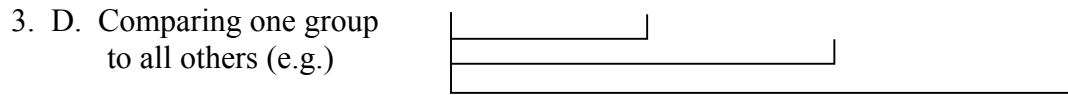
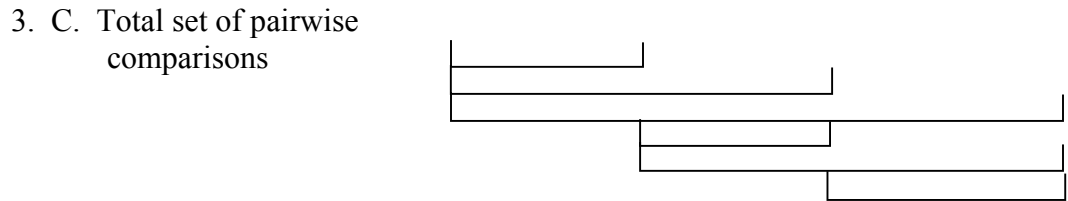
$$\alpha_{\text{BON}} = 1 - (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_3) \dots (1 - \alpha_n)$$

where α_1 to α_n are the set levels of alpha for a series of tests. For example, a series of 4 tests at $\alpha = .01$ would use a Bonferroni corrected alpha criterion of $1 - .99^4 = .039$. (This is equivalent to the FW (familywise) error formula on p. 1.)

Graphical representations for selected contrasts and post hoc families:

Assume this 4-group
IV:

Married	Separated/ Divorced	Widowed	Never Been Married
1	2	3	4



3. A. & B. –too complex to draw here

References

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill Book Company.