Alpha Inflation and others

Matthias RAUTERBERG
Eindhoven University of Technology

2017

H0: Tested by k-groups ANOVA

- Regardless of the number of IV conditions, the H0 tested using ANOVA (F-test) is ...
  - “all the IV conditions represent populations that have the same mean on the DV”

- For only 2 IV conditions, the F-test of this H0 is sufficient
  - there are only three possible outcomes ...
    $T=C$ $T<C$ $T>C$ & only one comparison matches the Research Hypothesis (RH)

- With multiple IV conditions, the H0 is still that the IV conditions have the same mean DV...
  - $T_1 = T_2 = C$ but there are many possible patterns but only one pattern matches the RH

Omnibus F vs. Pairwise Comparisons

- **Omnibus F**
  - overall test of whether there are any mean DV differences among the multiple IV conditions
  - Tests H0 that all the means are equal
- **Pairwise Comparisons**
  - specific tests of whether or not each pair of IV conditions has a mean difference on the DV
- *How many Pairwise comparisons ??*
  - Formula, with k = # IV conditions

    # pairwise comparisons = [k * (k-1)] / 2
  - or just remember a few of them that are common...
    - 3 groups = 3 pairwise comparisons
    - 4 groups = 6 pairwise comparisons
    - 5 groups = 10 pairwise comparisons

How many Pairwise comparisons – revisited !!

There are two questions, often with different answers…

1. How many pairwise comparisons can be computed for this research design?
   - Answer → [k * (k-1)] / 2
   - But remember → if the design has only 2 conditions the Omnibus-F is sufficient; no pairwise comparison is needed.

2. How many pairwise comparisons are needed to test the RH?
   - Must look carefully at the RH to decide how many comparisons are needed.
   - E.g., The *T1* will outperform the *Control (C)*, but not do as well as the *T2*

     This requires only 2 comparisons

     T1 vs. C   T1 vs. T2

Example analysis of a multiple IV conditions design

T1	T2	C
50 mean	**40** mean	**35** mean

E.g., for this design, $F(2,27)=6.54$, $p=.005$ was obtained.

We would then compute the pairwise mean differences:

[T1 vs. T2] = 10 [T1 vs. C] = 15 [T2 vs. C] = 5

Say for this analysis the minimum mean difference is 7

Determine which pairs have significantly different means:

T1 vs. T2	T1 vs. C	T2 vs. C
<u>**Sig Diff**</u>	<u>**Sig Diff**</u>	**Not Diff**

Alpha Inflation

- Increasing chance of making a Type I error as more pairwise comparisons are conducted.

Alpha correction

- adjusting the set of tests of pairwise differences to "correct for" alpha inflation.
- so that the overall chance of committing a Type I error is held at 5%, no matter how many pairwise comparisons are made.

When does Alpha Inflation occur?

- As you know, whenever we reject H0, there is a chance of committing a Type I error (thinking there is a mean difference when there really isn't one in the population).
  - The chance of a Type I error = the p-value
  - If we reject H0 because $p < .05$, then there's about a 5% chance we have made a Type I error

- When we make multiple pairwise comparisons, the Type I error rate for each is about 5%, but that error rate "accumulates" across each comparison -- called "**alpha inflation**".
  - So, if we have 3 IV conditions and make 3 pairwise comparisons possible, we have about …

    $3 * .05 = .15$ or about a 15% chance of making at least one Type I error

Here are the pairwise comparisons most commonly used -- but there are several others…

Fisher's LSD (least significance difference)

- **no Omnibus-F – do a separate F- or t-test for each pair of conditions**
- **no alpha correction – use α = .05 for each comparison**

Fisher's "Protected tests"

- **"protected" by the omnibus-F -- only perform the pairwise comparisons IF there is an overall significant difference**
- **no alpha correction -- uses α = .05 for each comparison**

Scheffe's test

- emphasized importance of correction for Alpha Inflation
- pointed out there are "complex comparisons" as well as "pairwise" comparisons that might be examined

E.g., for 3 conditions you have…

3 simple comparisons T1 vs. T2 T1 vs. C T2 vs. C

3 complex comparisons – by combining conditions and comparing their average mean to the mean of other condition

(T1+T2) vs. C (T1+C) vs. T2 (T2+C) vs. T1

- developed formulas to control alpha for the total number of comparisons (simple and complex) available for the number of IV conditions

Bonferroni (Dunn's) correction

- pointed out that we don't always look at all possible comparisons
- developed a formula to control alpha inflation by "correcting for" the actual number of comparisons that are conducted
- the p-value for each comparison is set **α = .05 / #comparisons**

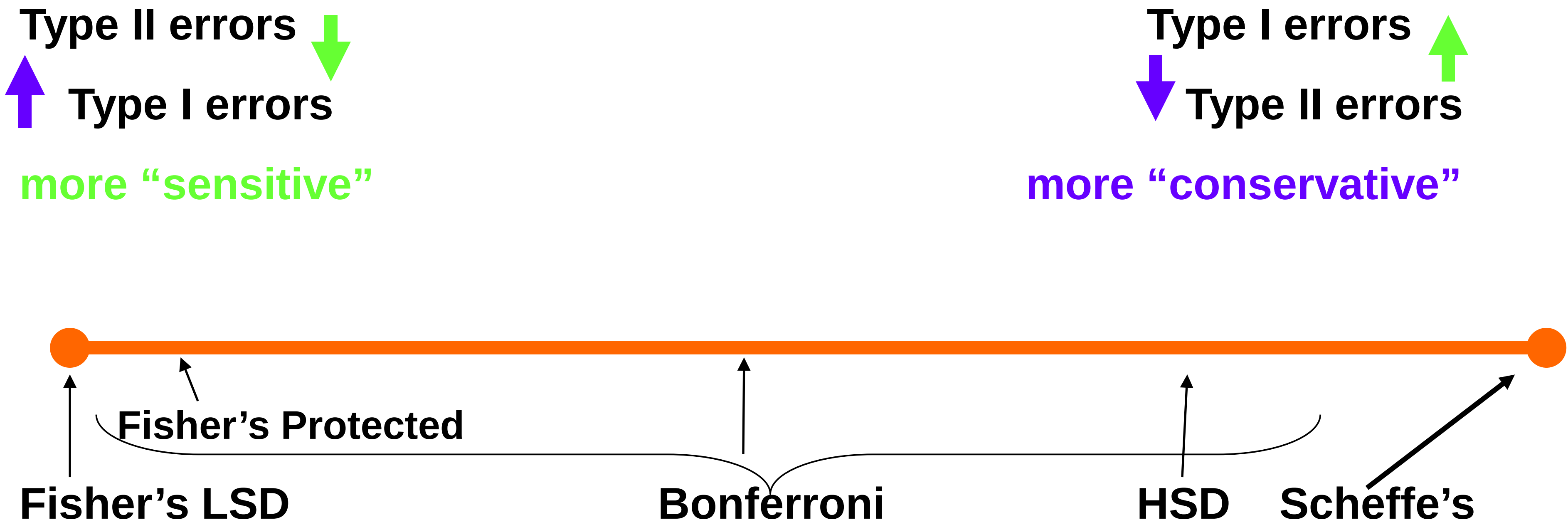
Tukey's HSD (honestly significant difference)

- pointed out the most common analysis was to look at all the simple comparisons – most RH are directly tested this way
- developed a formula to control alpha inflation by "correcting for" the number of pairwise comparisons available for the number of IV conditions

Dunnett's test

- used to compare one IV condition to all the others
- alpha correction considers non-independence of comparisons

The “tradeoff” or “continuum” among pairwise comparisons



- Bonferroni has a “range” on the continuum, depending upon the number of comparisons being “corrected for”
- Bonferroni is slightly more conservative than HSD when correcting for all possible comparisons

Prediction and Regression Analysis

- Correlational technique
- Simple prediction
  - Predicting an unknown score (Y) based on a single predictor variable (X)
  - $Y' = bX + c$
- Multiple prediction
  - Involves more than one predictor variable
  - $Y' = b_1X_1 + b_2X_2 + c$

Multiple Regression/Prediction

- a.k.a multiple correlation
- Determines the relationship between one dependent variable and 2 or more predictor variables
- Used to predict performance on one variable from another
  - $Y' = b_1X_1 + b_2X_2 + c$
- Standard Error (SE) of prediction is an index of accuracy of the prediction