Planned Comparisons and Post Hoc Tests

Planned: You define in advance a set of independent linear comparisons between the levels of a factor. This may reveal an internal difference even if there was no overall significance.
Post-hoc: After obtaining a significant effect for a factor, you carry out comparisons between specific levels to see which ones differ – you might consider all possible comparisons.

Planned **Analytical** **Orthogonal** **a priori**	**Unplanned** **post-hoc**
Do specified means differ?	*Which* means differ?
Comparisons test hypotheses	Comparisons are summaries
Ignore experimentwise error up to a limit	Control experimentwise error
Omnibus *F* need not be significant	Used after a significant *F*
Can be used instead of omnibus *F*	Used after a significant *F*
Comparisons usually < *d.f.*	Comparisons can be > *d.f.*

'EXPERIMENT-WISE' or FAMILY-WISE ERROR RATE

= probability of at least 1 Type I error

= p(family of comparisons will contain at least one Type I error)

= $1 - (1 - \alpha)^c$, where c = number of independent comparisons

Planned Comparisons

Orthogonal contrasts

- Partition the sum of squares for a factor A with k levels into a set of k-1 orthogonal contrasts, each with two levels (df=1) formed by grouping the levels in A.
- The two groups contrasted are assigned, respectively, positive and negative coefficients, with any level not included in the contrast assigned 0
- Coefficients of the contrast must add to 0, so use coefficients of 1/n where n is number in that group (or multiply through).
- Orthogonality: product of coefficients of any pair of contrasts sum to 0.

e.g. For a one-way ANOVA with 4 groups (A B C D), you could first compare AB with CD, then compare AC with BD, then AD with BC. The contrast coefficients would be:

Contrast	A	B	C	D
AB -CD	1	1	-1	-1

A-B	1	-1	1	-1
C-D	1	-1	-1	1

But perhaps it is more relevant to compare A against BCD (perhaps A is the control) and then look for difference of B from CD and C from D

Contrast	A	B	C	D
A-BCD	3	-1	-1	-1
B-CD	0	2	-1	-1
C-D	0	0	1	-1

There is a third possible set of orthogonal contrasts on the four levels, compare AB with CD, then A with B, and C with D. The contrast coefficients would be:

Contrast	A	B	C	D
AB -CD	1	1	-1	-1
A-B	1	-1	0	0
C-D	0	0	1	-1

Which set of contrasts you choose would be determined by the logic of your experiment. Can then calculate for each contrast:

$$\frac{\sum_{i=1}^{k} coeff_i * \bar{x}_i}{\sqrt{MS_R \sum_{i=1}^{k} \frac{coeff_i^2}{n_i}}}$$

Which is distributed as $t(df_R)$.

Post Hoc Tests

- Used when you discover an unforeseen effect in ANOVA (or had no prior expectation about what differences might be seen) – comparisons not planned in advance.
- Multiple *t*-tests with ***Bonferonni*** correction – adjust for familywise error rate
    - Divide α by number of comparisons made:
    - E.g. if making 5 comparisons, then $p = \alpha/5 = .05/5 = .01$
        - Increase chance of Type II error – being too conservative / reducing α, and the power of the experiment to find where the actual differences lie.
- Modified Bonferonni, Dunn - Sidak's correction:

$$\alpha_{\text{comparison}} = 1 - (1 - \alpha_{\text{overall}})^{1/c}$$

    - Still too conservative if comparisons not independent (e.g. if doing all possible pairwise comparisons.
- All use some variation of *t*-test
    - The top of the formula uses the difference between cell means
    - The denominator uses MS_R from the ANOVA table (i.e. estimates variance from all data, not just the means being compared).
- Other post hoc tests - some more conservative than others:

- Neuman-Keuls (S-N-K) – liberal on Type I and most likely to get a significant result
- Scheffé – strict on Type I – bad for Type II – less likely to show sig result
- Dunnett's – for comparing all treatments to a single control
- Games-Howell – doesn't assume equal variances
- Tukey HSD(Honestly Significant Difference) – preferred test; greatest power and readily available in many stats packages.
    - Power advantage of the Tukey test depends on the assumption that all possible pairwise comparisons are being made, which is usually the case for Post Hoc tests.