Chapter 24: Three or More Factors

• If we have 3 or more factors, we have the possibility of higher-order interactions.

Example: Factors A, B, C:

Model:

• If the 3-way interaction is significant, this implies, for example, that the

• Having 3 or more factors implies having lots of "cells", often with small sample sizes per cell.

• If high-order interactions do not have a practical meaning, it is often better to assume they do not exist and not include them in the model.

• In that case, we can devote more degrees of freedom to estimating σ^2 (yields more power for the other F-tests).

• Analysis of 3-plus-factor experiments is done similarly to 2-factor experiments:

<u>Example</u> (Stress data – Table 24.4):

<u>Response</u>:

<u>Factors</u>:

• See SAS example for analyses:

Chapter 23: Unbalanced Data
(Situation when Cell Sample Sizes are Unequal)

Reasons for Unequal Sample Sizes:

- **In many observational studies, researcher has no control over which units have which treatments**
- **Subjects may leave a study before its conclusion**
- **Cost considerations**

Notation: Let n_{ij} denote the number of observations for treatment (i, j). (i-th level of A, j-th level of B)

Then:

Example (Growth Data):
Response:

Factor A:

Factor B:

Data as follows:

• Analysis of effects is <u>not</u> based on ANOVA table formulas (factor effect SS <u>do not</u> sum to SSTR).

• Instead, we use the regression approach to the ANOVA models with indicators defined as on p. 955-956

• F-tests for significant effects are based on the Full vs. Reduced Model approach.

• In SAS: We need to look at Type III SS in PROC GLM output (this does correct F-tests), not Type I SS.

• In PROC GLM, Type I SS are <u>sequential</u> SS (each SS gives the reduction in SSE associated with adding that term, <u>in sequence</u>, to the MODEL statement.
• Type III SS are <u>partial</u> SS (each SS gives the reduction in SSE associated with adding that term, <u>given all other terms</u> in the MODEL statement).
• Type I SS change if the order of the terms in the MODEL statement changes.
• In ANOVA with balanced data, Type I SS = Type III SS (not true with unbalanced data).

<u>Least Squares Means</u>

• The estimate of a factor level mean is the <u>least squares</u> mean (the <u>unweighted average</u> of the appropriate cell means).

Example:

• In SAS, the LSMEANS statement (rather than the MEANS statement) gives the least squares means and the correct Tukey CIs.

<u>SAS Example</u> (Growth data):

• Significant interaction between gender and bone development?

• Significant effect of gender on mean growth rate change?

• Significant effect of bone development on mean growth rate change?

Error df =

• Can use Tukey CIs or tests to investigate differences in mean growth rate change among the levels of bone development:

<u>Empty Cells in Two-Factor Studies</u>

• Suppose one or more treatment cells contain <u>no</u> observations.

<u>Previous example</u>: Suppose the observation in the (2,1) cell had been lost:

• Certain partial analyses can still be done.

• Analyses are typically done based on <u>contrasts</u> involving the non-empty cells.

• If we can find an <u>unbiased estimate</u> of a function of cell means using our observed data, then this is an <u>estimable function</u> of the cell means.

<u>Example</u>: Is there interaction between Gender and Bone Development?

• If there is no interaction, then

Look at the contrast:

• We can test

SAS Example: To get the correct ESTIMATE statement, write out the contrast in terms of the factor effects:

• Analysis using the ESTIMATE statement indicates

Error df =

• We can examine the effect of gender by comparing

• We can examine the effect of bone development by comparing

Note: These contrasts do not use

• We could compare all 3 levels of bone development, but only

• LSMEANS statement gives us estimates of