Chapter 19: Two-Factor Studies (Equal Sample Sizes)

• Here we look at the simultaneous effect on the response of <u>two</u> factors.

<u>Example 1</u>: Response =
Factor A:

Factor B:

• With multiple factors, each <u>factor level combination</u> is called a "treatment".
• How many treatments in Example 1?

• The experimental units here are

<u>Example 2</u>: Response =
Factor A:

Factor B:

• Total of treatments.

Subjects:

• Unfortunately, some researchers (instead of planning a multifactor study) perform their study in stages, where, at each stage, one factor at a time is varied/explored.

This approach is inferior because it:
• may miss certain treatment combinations
• is more difficult to carry out logistically
• cannot be properly randomized
• is unable to properly assess interactions between factors
• is less efficient – requires more observations to get the same precision

• Pages 815-816 of the book discuss this in more detail.

Notation for the Two-Factor ANOVA Model (Two-Way ANOVA)

• Denote one factor by A (which has *a* levels) and the other factor by B (which has *b* levels).

• Let Y_{ijk} be the *k*-th observation from level *i* of factor A and level *j* of factor B.
(Here, $i = 1, \ldots, a$, and $j = 1, \ldots, b$.)

• If there are *n* observations per factor level combination, then $k = 1, \ldots, n$.

• The total number of observations in the study is

• The "Cell means" formulation of the Two-Way ANOVA model is:

• The μ_{ij} values are unknown parameters:

μ_{ij} is the population mean response at level *i* of A and level *j* of B.

• The random error term ε_{ijk} is assumed to have a normal distribution with mean zero and variance σ^2 (constant across all treatments):

• This model can be expressed in the form:

<u>Example</u>: (Suppose *a* = 2, *b* = 2)

Note

• The cell-means formulation is simple, but does not explicitly show the effects of each factor on the response, nor the interaction between factors.

Factor Effects Formulation of ANOVA Model

Here:

Interpretations of Main Effects:
α_i = difference between "mean response at level *i* of factor A" and "overall mean response averaged over the levels of both factors".

β_j = difference between "mean response at level *j* of factor B" and "overall mean response averaged over the levels of both factors".

• The interaction effects measure how the effects of one factor <u>vary</u> at <u>different levels</u> of the <u>other</u> factor.

• Significant interaction may or may not exist in a two-factor study (we need to check this with our data).

Example:

• See graphical examples of interaction in Figure 19.7 in book.

Notation in Two-Factor Model

• For each observation Y_{ijk}, the fitted value is

And the residual is

• These fitted values are the least squares estimates of μ_{ij}, found by minimizing the SSE subject to the restrictions:

Sums of Squares

- If SSA is large, a lot of variation in the response can be explained by factor A.
- If SSB is large, a lot of variation in the response can be explained by factor B.
- If SSAB is large, there is sizable interaction between factors A and B.

• **Dividing each SS by its associated degrees of freedom gives the Mean Square.**
• **This can be summarized in an ANOVA table:**

• **These F* statistics are obtained based on the <u>Expected Mean Squares</u>:**

• **For each F-test, values of F* much _______________ than 1 are evidence of significant effects.**

Example: (Castle Bakery)

Response: Sales of bread (in cases)

Factor A: Height of Shelf Display. Levels:

Factor B: Width of Shelf Display. Levels:

• There are

• Twelve experimental units (supermarkets) →
→

Results shown in Table 19.7, pg. 833 (6 cells, 2 observations per cell):

<u>SAS Example</u> (Castle Bakery data): PROC GLM gives the ANOVA table:

Checking Model Assumptions is again done by:
(1) Plotting Residuals vs. Fitted Values
(2) Normal Q-Q plots of residuals (separately by treatment if the treatment sample sizes are large, otherwise do one plot)
• See SAS plots:

• First we determine whether significant interaction exists.

Strategy: (1) Interaction Plots, and
(2) F-test about Interaction Effects

(1) Plot treatment sample means across levels of one factor, <u>separately</u> for <u>each</u> level of the other factor.

• Non-parallel lines indicate interaction.

(2) F-test:

SAS gives the F* and the P-value.

<u>Example</u> (Bakery data):

• What if significant interaction was found?

• In some cases, an interaction may be significant but <u>practically unimportant</u>. A judgment can be aided with <u>interaction plots</u>:
Picture:

• Often a transformation of *Y* can "remove" or mostly eliminate interaction effects and make them unimportant.
• See pages 826-827 for some mathematical examples of this.

• Common choices for such transformations include:
$Y^* = \ln(Y)$ $\quad$ $Y^* = Y^{1/2}$ $\quad$ $Y^* = Y^2$ $\quad$ $Y^* = 1/Y$

• If an interaction cannot be largely removed by a transformation, then it is called a <u>nontransformable interaction</u>.

• If there is NO significant / important interaction, then we may examine the main effects of each factor separately.

<u>Test about Factor A</u>

Example: (Castle Bakery data)

Test about Factor B

Example: (Castle Bakery data)

• **Note that these main-effects tests are typically done only when there is NO significant interaction.**