Using Two-Way Analysis of Variance

Administrative Items

Getting help
– See me Monday 3-5:30 or Wednesday from 4-5:30.
– Send an e-mail to stine@wharton.
– Visit the StatLab/TAs, particularly for help using the computer.

Midterm Exam #2
Tuesday April 4, from 6-8 p.m. in Annenberg 110.

Assignment #4
Due Wednesday, March 29.

Questions about Multiple Comparions

Why do we need multiple comparisons?
The overall F test given in the anova table indicates only whether some difference occurs among the μ_i's. It does *not* indicate where a difference occurs. To tell which means differ, multiple comparison methods allow us to build *several* confidence intervals simultaneously while keeping α fixed.

Two procedures offer different types of comparisons. Hsu's method gives comparisons to the mean of the group with the largest (or smallest) average. Tukey's method allows *any* pairwise comparison.

If Hsu's method only compares to the extremes, why a table?
The folks at SAS blew this one. The output for Hsu's comparisons really should only have one row of the table that is shown. In the car example, the last three rows are not relevant and should be ignored.

Mean[i]-Mean[j]-LSD	E	B	S	X
E	-2.38	0.27	2.15	2.83
B	-5.03	-2.38	-0.50	0.17
S	-6.91	-4.26	-2.38	-1.71
X	-7.59	-4.93	-3.05	-2.38

If a column has any positive values, the mean is significantly less than the max.

Experimenting with Two-Way Anova

Gummy bears in space!

Review of Two-Way Anova

Model

y_{ijk} = (cell mean) + error

=(overall average) + (row effect) + (column effect) + (interaction) + error
where the data are assumed to be independent observations with constant variance, normally distributed about the cell mean value.

Questions in a two-way anova

(a) Are the data and experiment appropriate for a two-way anova?
(b) Is interaction present?
- If so, what is the nature of the interaction?
- If not, what differences are there among row and column averages?
(c) Are any means in the two-way layout significantly different?

We use the F-ratio from the anova summary to judge the amount of interaction (b), and look at the profile plot to understand the size of the effect. In the absence of interaction, we can use the other F-ratios to assess the significance of differences among the row averages and column averages.

To judge the size of differences among the individual cell averages (c), we can again use Hsu's and Tukey's intervals, thinking of the table as a one-way anova with (#rows)(#cols) groups.

Estimate error variance

Estimate the error variance using the error sum of squares from the anova table,

$$\hat{\sigma}^2 = \frac{Error\,SS}{n-IJ} = \frac{\sum(y_{ijk} - \bar{y}_{ij})^2}{n-IJ} = \frac{\sum(\hat{\varepsilon}_{ijk})^2}{n-IJ} = Error\,MS$$

JMP for Two-Way Anova

Fit model command

– Use the "Fit model" command.
– To form an interaction term:
    Add a factor twice, then with both it and the interacting term highlighted in the JMP dialog, use the "Cross" button to build the interaction term.
– Construct profile plots with the button near the graphic for the interaction term in the output of the *Fit model* command. (far right of output window)

Concatenate function

– Concatenate group labels (formula) to look at individual cell means.
– The concatenate function is a character formula in JMP's calculator.

Example: Two-Way Analysis of Variance

Web page experiment

Customers in a "focus group" represent the target audience of a web site. Sixty were randomized into six groups, with 10 in each group.

Each group was shown a web page with either high or low graphics layout, and with either one, two, or three+ clickout options. After seeing the page, customers rated the "convenience and esthetics" of the design on a 0-100 scale.

The design of the experiment thus has a "two-way" arrangement…

Layout-Clicks	One-click	Two-click	Three+	
Low graphics	10	10	10	30
High graphics	10	10	10	30
	20	20	20	60

Initial conceptual analysis

Can this experiment yield useful information for the business?

Initial graphical analysis

– Plots of data by marginal factors reveal no severe outliers, similar variances.
– The data for number of clicks shows a weak trend, but the differences in means are not significant.
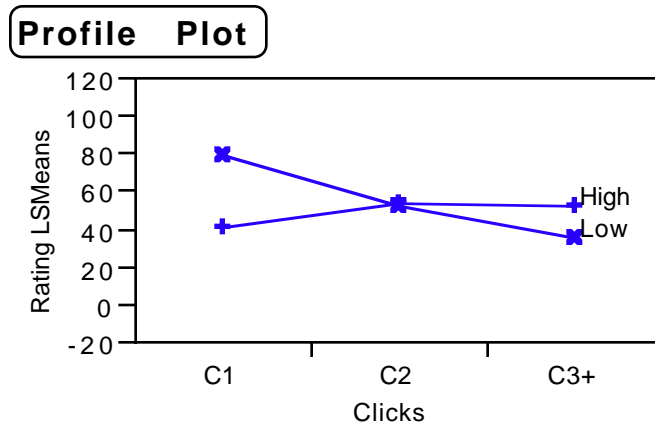
Anova table

– First check for interaction in the "Effect test" summary

Effect Test					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Clicks	2	2	2582.1	3.70	0.0312
Layout	1	1	653.4	1.87	0.1768
Layout*Clicks	2	2	7837.3	11.23	<.0001

– The interaction *is* significant. (F = 11.23 with p-value <.0001)
– Hence, you should *not* attempt to judge the differences due to the number of clicks and layout from the anova summary. They are not consistent; interaction implies that the rating for low/high graphics is not the same for the different check-out options.

Profile plot

– This plot of cell means is very useful (it graphs the averages of the six combinations of clicks and layout, joining those in the same row.)

Profile Plot

Rating LSMeans (y-axis: -20, 0, 20, 40, 60, 80, 100, 120)

Clicks (x-axis: C1, C2, C3+)

High, Low

– The lack of parallel lines indicates the presence of interaction. The layout with low graphics is preferred when featured with one-click checkout, whereas users are ambivalent or prefer high graphics with the more elaborate checkouts.

Multiple Comparisons

– Using JMP, reduce the problem to a one-way problem. Join the labels of the two factors using the concatenation function, then do a one-way anova.

Join the labels of "Clicks" and "Layout" as a new column, then use this one column of categories as "X" in a one-way anova.

– Here's a table of the cell means and SDs. These are the mean values plotted in the profile plot shown above.

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean
C1 - High	10	42.2	14.89	4.71
C1 - Low	10	79.6	17.67	5.59
C2 - High	10	53.7	24.51	7.75
C2 - Low	10	53.4	18.20	5.75
C3+ - High	10	53.5	14.95	4.73
C3+ - Low	10	36.2	20.08	6.35

Graphical analysis with comparison circles...

The one-click checkout option with low graphics layout is significantly different fro
the others using Tukey-Kramer.
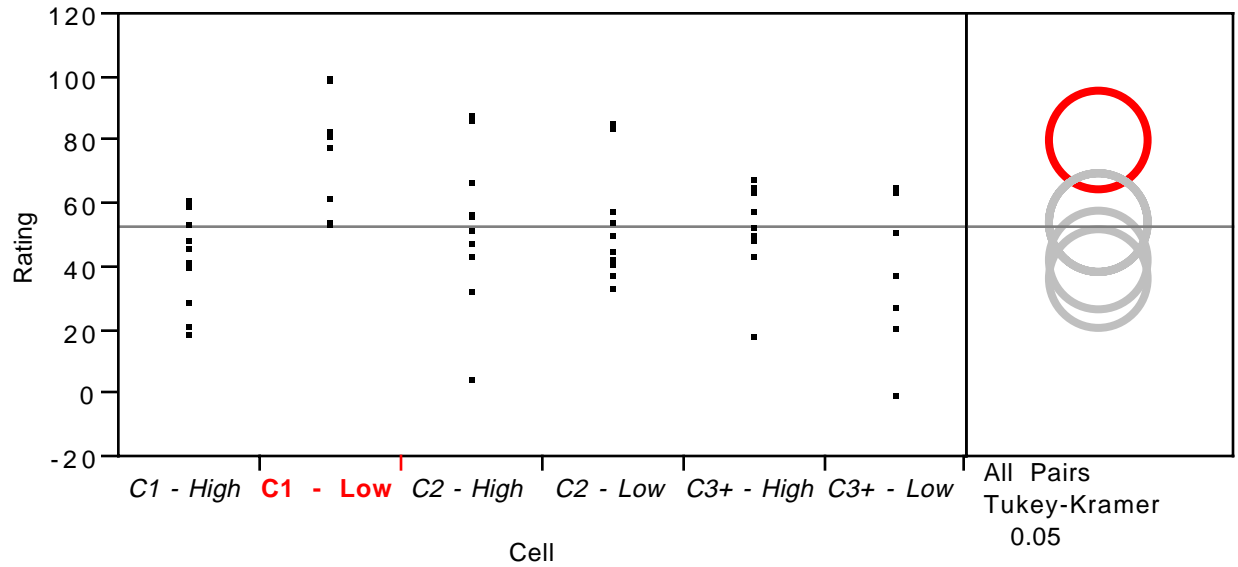
Tabular summary...

Comparisons for all pairs using Tukey-Kramer HSD

q *
2.95448

Abs(Dif)-LSD	C1 - Low	C2 - High	C3+ - High	C2 - Low	C1 - High	C3+ - L
C1 - Low	-24.68	1.22	1.42	1.52	12.72	18
C2 - High	1.22	-24.68	-24.48	-24.38	-13.18	-7
C3+ - High	1.42	-24.48	-24.68	-24.58	-13.38	-7
C2 - Low	1.52	-24.38	-24.58	-24.68	-13.48	-7
C1 - High	12.72	-13.18	-13.38	-13.48	-24.68	-18
C3+ - Low	18.72	-7.18	-7.38	-7.48	-18.68	-24

Positive values show pairs of means that are significantly different.

The Tukey intervals are constructed in the same manner as those in a one-way anova. For example, the profile plot shows a big difference between low and high graphics layouts when combined with a one-click checkout. Is this difference significant? The formula for the Tukey interval is

$$(\textit{difference in means}) \pm q_\alpha (\textit{number of means, error df}) \sqrt{\hat{\sigma}^2 / \textit{num per group}}$$

which in this case gives the interval

$$(79.6 - 42.2) \pm q_{.05}(6,54)\sqrt{\hat{\sigma}^2/10} = 37.4 \pm 4.16\sqrt{348.83/10} = 37.4 \pm 24.6$$

The lower endpoint of this interval (12.8) is within rounding of the value in JMP's more carefully computed table. I got the value for the estimated error variance from JMP's summary anova table (where it's labeled the mean square for error)

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	11072.800	2214.56	6.3486
Error	54	18836.600	348.83	**Prob>F**
C Total	59	29909.400		0.0001

and the value for q_α comes from Table 8 in the text (using the row for 60 df).

Summary

Effect and meaning of interaction.
    synergy of row and column factors
    marginal effects (clicks, layout) cannot be interpreted directly
    business implications

Tukey-Kramer intervals identify significant differences among cell means.

Challenge Question

Why do you think that the value labeled q* in JMP's Tukey Kramer output differs by the square root of 2 from that found in the table in the textbook?