Statistics 5301 — Exam 2, Sketched Solution   NAME _____

February 27, 1998

Gary W. Oehlert   ID# _____


1. Describe the experimental design that was used. Tell me about treatments, units, responses, randomization, blocking (if any) and so on.

The is a randomized complete block. Enclosures are blocks, and tanks are experimental units. The treatments are the 4 factor/level combinations of ruffe and density. Measurement units would be the individual perch, and response is weight change.

2. Describe how you checked assumptions and what you decided. Tell me about nonnormality, nonconstant variance, outliers, and so on. Were there any problems that required fixing, and could you fix the problems?

Looking at the residuals vs predicted plot, we see a problem. There is a general pattern of a southwest to northeast blob, with isolated points in the northwest and southeast. This suggests that points 3 and 14 are outliers.
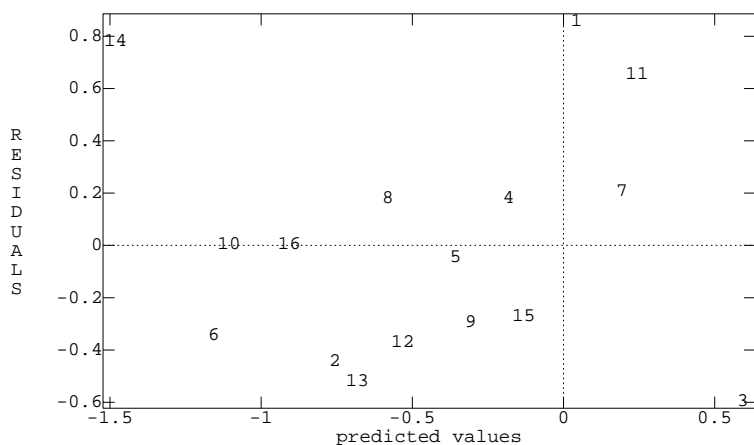
Refitting without those two points, the residual plot looks a little better, though there is still some curvature.
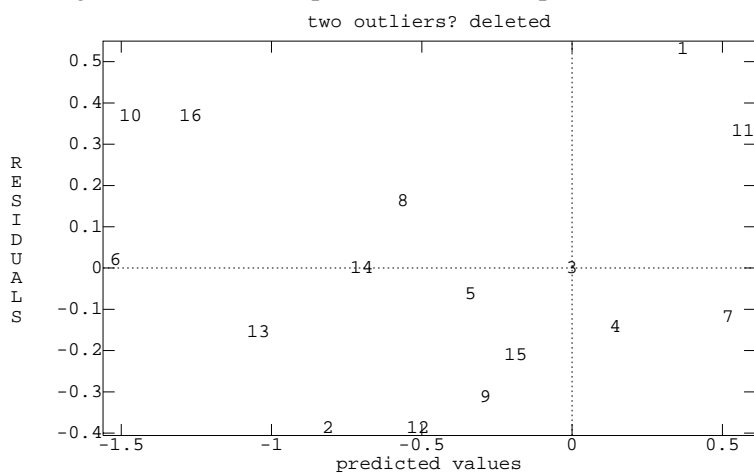
Analyzing the log of change+2 removes much of the curvature (which is more or less a Tukey type, since we can remove it by transformation). However, the two "outliers" are not very significant (p-values about .02 to .03), and certainly don't look very significant if we do a Bonferroni correction for searching for outliers.
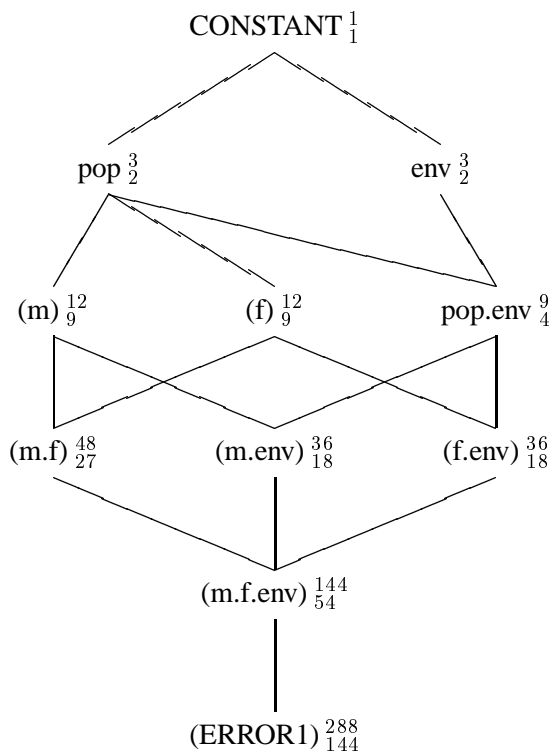
3. What did you conclude from the experiment?
If we remove the two outliers, the main effects of both ruffe and density are significant (p-values < .01). Presence of ruffe reduces the response by about 1.13g, and higher density reduces the response by about .9g; the two "outliers" are about 1.2g low and 1.5g high respectively. There is no evidence for interaction. If we keep all the data, there is still no evidence for interaction. The ruffe effect is significant with a p-value of .025, but the density effect is only marginally significant with a p-value of .088. The estimated ruffe effect from the full data set is a reduction of .78g when ruffe are present.

4. Nonstarter bacteria may affect the flavor of cheese.
   This is a randomized complete block design. Vats are experimental units and times of day are blocks. The four cultures are the treatments, and judged flavor is the response. The block of cheese is a measurement unit.

Source	DF
block	2
culture	3
error	6

CONSTANT $\frac{1}{1}$

pop $\frac{3}{2}$ env $\frac{3}{2}$

(m) $\frac{12}{9}$ (f) $\frac{12}{9}$ pop.env $\frac{9}{4}$

(m.f) $\frac{48}{27}$ (m.env) $\frac{36}{18}$ (f.env) $\frac{36}{18}$

(m.f.env) $\frac{144}{54}$

(ERROR1) $\frac{288}{144}$

5. An experiment was conducted to study the effects of genetics and the environment an annual prairie plant.
a) What is the expected mean square for environment?

$$EMS(env) = \frac{288}{288}\sigma^2 + \frac{288}{144}\sigma^2_{m.f.env} + \frac{288}{18}\sigma^2_{m.env} + \frac{288}{18}\sigma^2_{f.env} + \frac{288}{3}\sum_j \beta_j^2/2$$

$$= \sigma^2 + 2\sigma^2_{m.f.env} + 16\sigma^2_{m.env} + 16\sigma^2_{f.env} + 48\sum_j \beta_j^2$$

b) What is the denominator (error) to use when testing the environment by male interaction?
The denominator for enviroment by male is the environment by male by female interaction, which is random, below env.m, and the only additional factor it contains is random.

6. Many pages on the World Wide Web contain advertisements for other commercial sites.

Use a Latin square. Treatments are the three advertisements. Units are the 9 day/time-of-day combinations on Friday, Saturday, and Sunday. Block on day and time-of-day. Randomly assign the three ads to the 9 units subject to the restriction that each ad appears once each day and once in each time-of-day.

7. We wish to study how the hormone melatonin affects sleep habits.

This looks like a completely randomized design. There are four treatments (the 4 factor/level combinations of the melatonin factor (no/yes), and the light factor (no/yes)). There may be subject differences, but we have no information on which to construct blocks. Therefore we will randomly assign 8 subjects to each of the 4 factor/level combinations. change in sleep hours is response.

8. The EPA is trying to encourage volunteer monitors to check the health of streams by collecting and classifying macroinvertebrates in the streams.

I'd use a Balanced Incomplete Block Design. Treatments are the three subsampling methods. Collections are the blocks, and splits of the collection are the two units per block. Response is diversity. I'd randomly assign 10 blocks each to the treatment pairs (AB), (AC), and (BC), and then I'd randomly assign the treatments for a given block to the units in a given block.