Statistics 5303 — Final Exam Sketched Solutions NAME ____________________
December 20, 2010
Gary W. Oehlert ID# ____________________

1. Every year at Christmas I make "thumbprint" cookies. I want to perfect them, and I need to find the right time and temperature for baking them. It should be about 375 degrees for about 11 minutes, but it may not be that exactly. I'm going to make 12 trays of these cookies this week, and from those trays I want to be able to estimate the best time and temperature for baking. Note, all the cookies on a tray are baked for the same time/temperature combination. The response is how good the cookies taste to me.

We should use a rotatable central composite design with four center points, with tray as experimental unit and cookie as measurement unit. Put the center points at 375 degrees and 11 minutes. Put the factorial points at (for example) plus or minus 10 degrees and plus or minus one minute.

2. The cookbook *Joy of Cooking* has a dynamite recipe for egg nog (and repeats an adage attributed to Mark Twain about how too much of anything is bad, except for whisky, for which too much is just right). We want to design an experiment that will compare the use of light rum, dark rum, and bourbon in this recipe; we make a batch using each liquor. A rater should drink at least four ounces of one of the egg nogs before giving his or her rating, and we expect substantial rater to rater variability. Furthermore, this recipe contains a lot of alcohol, and, to put it gently, ratings after the second cup are not reliable. We have twelve people at our party willing to dedicate a few moments to science and participate in the experiment.

We want a Youden Square, blocking on rater and whether it is the first or second cup tasted.

3. Our experiment will be done at UPS and will vary class (overnight, two-day, or ground), and size/weight of packages. There are four size/weight combinations obtained by crossing light and heavy with large and small. Each package contains only a recording accelerometer to measure the forces on the package and ballast to obtain the desired weight. On nine separate days, our colleague will bring four packages to the UPS store, one of each of the size/weight combinations. The packages will be presented at the counter in random order. On three randomly chosen days, the packages will be sent via overnight delivery; similarly, three randomly chosen days will be two-day delivery, and three days will be ground delivery. All packages are sent from Minneapolis to the same address in Connecticut. The recipient in Connecticut will open the package and send us the peak acceleration experienced by each package.

This is a split plot, with day as whole plot, package class as whole plot treatment, and the weight/size combinations as split plot treatments.

Source	df
Class	2
WP Error (day)	6
Size	1
Weight	1
Size:Weight	1
Class:Size	2
Class:Weight	2
Class:Size:Weight	2
SP Error	18

4. Our dog's feet get cold and packed with ice when they go for walks in the winter. We can buy little dog booties, but they always come off. What we would like to do is find the brand of booties that stay on the longest. We have purchased booties of six different brands. As our dog only has four feet, we can only test four of them at one time. There are 15 different sets of four brands that can be taken from the six brands. We randomly assign the 15 sets to 15 consecutive days. On each day, we randomly assign the four brands from that day's set to the four feet on the dog. When we go on the walk, we time how long it takes until each bootie comes off, and that is the response.

This is a BIBD, with day as block, brand as treatment, and dog foot within day as unit.

Source	df
Block (day)	14
Brand	5
Error	40

5. The Vikings are playing a football game tonight at TCF Bank Stadium, and traffic is going to be a nightmare around the university. Out of morbid curiosity, a group of twelve students decide to run an experiment to determine the fastest way to get from Minneapolis campus to St. Paul campus leaving around 5:30 PM. They break into three sets of four. One group of four will be randomly assigned to get in a car at the 4th Street parking ramp and take Como Avenue to St. Paul. One group of four will get in a car at the Washington Avenue parking ramp and take University and Raymond to St. Paul. The third group will pick up the Campus Connector at Coffman Union. All students leave from Morrill Hall at the same time, and the response is travel time to the St. Paul Student Center.

This is a completely randomized design, but the group is really the unit, not the individual, as they all must have exactly the same response. Of course, we wind up with zero degrees of freedom this way, so they cannot do inference.

Source	df
Method	2
Error	0

6. An opened bottle of wine will deteriorate and become undrinkable. Various products purport to help preserve opened bottles. In this experiment, we use three different varieties (Merlot, Cabernet, and Shiraz) from each of five randomly chosen vineyards (just labeled A through E). We obtain six bottles of each variety from each vinyard. These six are randomly assigned to three different wine preservation products, two bottles to each of the three products. The bottles are opened, one half of the wine is poured out, and then the bottles are closed with their assigned product. Four days later, a professional wine judge tastes wine from each of the bottles (fully blinded) and gives a rating on a 1–100 scale for the wine.

Below are two different Hasse diagrams. Tell me which one is correct and why the other is not correct.

Bottles are actually "units" nested in the combination of variety, vineyard, and product. Bottles do not cross with products.

7. We have run a 2^{6-3}_{III} fractional factorial with generators ABCD, ACE, and BCF.

(a) List all of the aliases of I.

I = ABCD = ACE = BDE = BCF = ADF = ABEF = CDEF

(b) In the Daniel plot, the main effects of B, D, and E looked large. What can you conclude?

There is still some ambiguity. Because BDE is one of the aliases of I, B = DE, D = BE, and E = BD. Thus we could have three main effects or different sets of two main effects and their interaction.

(c) Your colleague claims that the following factor/level combinations were run in the experiment: ace, cd, bcf, abcdef, be, def, af, abd. Is he correct? Explain why or why not.

He is not correct. The first four have an even number of ABCDs in them, and the last four have an odd number of ABCDs in them. Because ABCD is one of the aliases of I, they must all be even or all be odd if this is the true design.

8. Consider the following Type II analysis of variance table. Which terms would you include in your model for the data? (You may indicate them on the ANOVA table if you wish.)

```
          Sum Sq Df F value    Pr(>F)    
a          0.768  2  0.4329   0.65357    
b         23.812  1 26.8318 2.637e-05 ***
c          2.148  1  2.4202   0.13287    
d         58.236  1 65.6208 2.528e-08 ***
a:b        0.724  2  0.4079   0.66955    
a:c        2.459  2  1.3856   0.26948    
b:c        0.176  1  0.1982   0.66018    
a:d        7.103  2  4.0021   0.03163 *  
b:d       77.767  1 87.6290 1.756e-09 ***
c:d        0.064  1  0.0721   0.79057    
a:b:c      7.980  2  4.4958   0.02196 *  
a:b:d      3.782  2  2.1305   0.14069    
a:c:d      3.225  2  1.8171   0.18415    
b:c:d      3.261  1  3.6749   0.06722 .  
a:b:c:d    2.732  2  1.5395   0.23494    
Residuals 21.299 24                      
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Using a 5% cutoff, a:b:c, b:d, and a:d are the top of the hierarchy for significant terms. We thus need them and anything that they contain: a, b, c, d, a:b, a:c, a:d, b:c, b:d, and a:b:c.

9. A laboratory is trying to find compounds that will increase the “grip” of tires in the winter. The real test is to make tires containing the compound, but that is expensive, and the lab can only afford to do that when there is some indication that the compound might actually work. Before making actual tires, the lab screens a large number of compounds using a cheaper, but less precise, process. For each compound screened, the output of the cheaper process is a p-value for the null hypothesis that the compound does not affect winter grip. Because the lab is testing a lot of compounds, there could be a lot of false positives if they take all compounds with a p-value below 5%. The lab doesn’t mind have a small fraction of false positives that go on to the expensive test, but they don’t want the fraction of false positives to be larger than 10%.

Choose an error rate that gives the lab the control that they want while still selecting as many active compounds as possible. Defend your choice.

We should use a procedure to control the False Discovery Rate at 10%. It’s clear that we need some control beyond per comparison error rate, and the lab is willing to tolerate 10% false discoveries, so this will provide the most significant results while still controlling the error rate that the lab wishes to control.