Statistics 5303 — Exam 2 Sketched Solutions   NAME _____

November 21, 2003

Gary W. Oehlert   ID# _____

1. Describe the experimental design.

This is a randomized complete block design, with sheet as block, and position on the sheet as unit, with the twelve dough/temperature combinations randomly assigned to the units. At another level, the temperature treatments are randomly assigned to the balls of dough. It's a real stretch to conceptualize the random assignment of dough type to some notional, but as yet nonexistent, ball of dough, and I won't try to do so.

2. Summarize the effects of dough type and temperature on cookie thickness.

Dough type and temperature are both significant factors on thickness, but they don't interact. The thickest dough was with type 4 (box mix), although type 1 (home made with margarine) was not significantly different at .05 by HSD. The thickest dough was frozen, although refrigerated was not significantly different at .05 by HSD.

3. My family of four suffers from allergies, so we all take antihistamines of one sort or another. Our doctors have suggested four different drugs, but we would like to choose one drug for all four of us to use. We (I) want to run an experiment to choose that drug optimally. Some constraints on the design include (a) we should each try all the drugs, (b) the doctors say that we need to take a drug for a month or so to get a reasonable idea of how well it works, and (c) allergens change over time.

The four drugs are treatments. We need to block on subject and on month (time period), so a Latin Square seems appropriate. Some additional considerations are that we might want to balance for carry over effects, and/or we might want to include a wash out period between treatments.

4. Some trumpets sound better than others, and there are groups that claim that temperature treatments will improve the sound of a trumpet. Some groups advocate cryogenic freezing, whereas other groups advocate a heat treatment. We wish to compare the freezing treatment, the heat treatment, and a control of no treatment. A professional musician will play the instruments, which will be judged for sound by a panel of experts; the average of the experts scores will be the response for any unit.
    Without a doubt, different models of trumpet sound different. Some instrument manufacturers have loaned us twelve trumpets, two from each of six models. We also have the time constraint that we can only use each instrument once.

We need to block on model of instrument, but we have three treatments and blocks of size two. Thus we need an incomplete block design. In this case, we can run a BIBD, with each pair of treatments appearing twice.

5. Recent research suggests that a mixture of caffeine and alcohol injected into the blood after stroke can reduce stroke damage by 80% (my wife suggests prophylaxis via Irish coffee). We wish to replicate their experiment and study their mixture, caffeine alone, alcohol alone, and a control. We can use 80 inbred rats, in which we can artificially induce stroke.
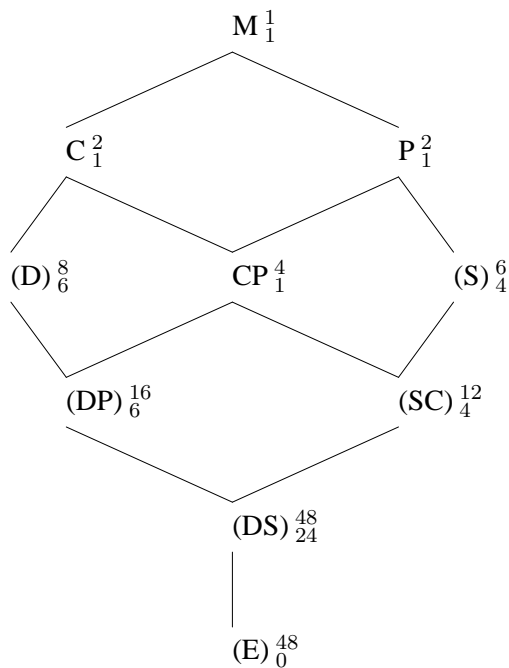
The rats are inbred and pretty exchangeable. I'd suggest a completely randomized design. If we have addition information, such as some rats will be kept in one room and others in another, or we'd have to do

the experiment in stages, etc, then we could consider some form of blocking, but we don't have that kine of information.

6. My daughters have supplied their Christmas wish lists (single spaced, double column, multipage—enough to bankrupt Bill Gates). These lists include many CDs and DVDs. You can buy these on-line or at "brick and mortar" stores. Being an impoverished academic, I'm always looking for good prices, so I collect some data. I randomly choose four each of CDs and DVDs from their combined wish list. From a list of retail and online stores, I randomly choose three brick and mortar stores and three on-line stores that sell digital media. I then price the eight selected items at the six selected stores.

  Construct a Hasse diagram for analyzing the collected prices.

Two types of store (P, fixed), with individual store (S) random and nested in type of store. Two categories of gift (C), with individual disk (D) random and nested in category of gift. Otherwise, all is crossed.

$$M_1^1$$

$$C_1^2 \qquad P_1^2$$

$$(D)_6^8 \qquad CP_1^4 \qquad (S)_4^6$$

$$(DP)_6^{16} \qquad (SC)_4^{12}$$

$$(DS)_{24}^{48}$$

$$(E)_0^{48}$$

7. Here is the output of three different ANOVAs on the same set of (unbalanced) data.

```
Cmd> anova("y=a*b",fstats:T)
Model used is y=a*b
WARNING: summaries are sequential
               DF          SS          MS           F       P-value
CONSTANT        1      480.19      480.19    67.98994    8.5625e-13
a               3      194.58       64.86     9.18355    2.1171e-05
b               3      40.143      13.381     1.89460       0.13565
a.b             9      48.572      5.3969     0.76414       0.64955
ERROR1         96      678.01      7.0626

Cmd> anova("y=b*a",fstats:T)
Model used is y=b*a
```

2

```
WARNING: summaries are sequential
                 DF           SS           MS            F      P-value
CONSTANT          1        480.19       480.19     67.98994   8.5625e-13
b                 3        217.67       72.555     10.27307   6.2893e-06
a                 3        17.058        5.686      0.80508      0.49406
b.a               9        48.572       5.3969      0.76414      0.64955
ERROR1           96        678.01       7.0626

Cmd> anova("y=a*b",fstats:T,marg:T)
Model used is y=a*b
WARNING: SS are Type III sums of squares
                 DF           SS           MS            F      P-value
CONSTANT          1        22.288       22.288      3.15575     0.078828
a                 3        19.816       6.6055      0.93527      0.42682
b                 3        21.207       7.0691      1.00091      0.39596
a.b               9        48.572       5.3969      0.76414      0.64955
ERROR1           96        678.01       7.0626
```

What do you conclude about the significance of the effects? (You may assume that all assumptions about normality, constant variance, etc are met.)

No main effects or interactions are significant using either type two or type three analysis. However, if you pool all 15 treatment degrees of freedom you get 283.22 SS with 15 df for a MS of 18.88. Testing this against error gives a p-value near 0. Thus the overall model does describe considerable variation, but we can't assign that reliably to any particular factor or interaction.