Statistics 5303 — Exam 2 NAME _____
November 29, 2004
Gary W. Oehlert ID# _____

This exam is open book, open notes; you may use a calculator. **Do your own work!** Use the back if more space is needed. There are seven questions, each worth 10 points. Please attach your data analysis notes for these data sets to your exam with the paper clips provided in the front of the room.

Questions 1 and 2 relate to the cheese viscoscity data. **Provide some justification for your answers!**

1. Do either temperature or rpm affect viscoscity? If so, how?

2. What model did you use for these data? Defend your choice.

3. Two common fish species in cold water streams are slimy sculpin and brown trout. These species tend to inhabit "riffles", which are shallow running stretches of the stream, sort of like miniature rapids. We are interested in whether the presence of the two species together inhibits or enhances total fish growth (combined across species). To study this, we place small cages called enclosures in riffles. Into each cage we can place either equal weights $x$ of slimy sculpin and brown trout, or a weight $2x$ of brown trout, or a weight $2x$ of slimy sculpin. After a month, we weigh the fish in each cage to assess total growth.

In our experiment there are five riffles. In each riffle we place three enclosures. The three treatments are randomized to the enclosures subject to the restriction that each treatment occurs once in each riffle. Describe the design that was used and give a skeleton ANOVA (sources and df only).

4. We are studying the Kraft pulping process for making paper. In this experiment, we look at the charge level (705, 853, or 1000), and which additive is used (control, DQ2016 at .1, DQ2016 at .2, AQ at .1, or DTPA at .2). We can make 10 batches of pulp per day and do the experiment over three days, producing two batches of pulp for each of the 15 combinations of charge level and additive. The fifteen factor/level combinations are randomly assigned to the 30 batches subject to the restriction that each combination is used twice. Describe the design that was used and give a skeleton ANOVA (sources and df only).

5. We wish to study "sensory specific satiety." This is the phenomenon wherein if you eat a lot of some food, then that food and similar foods become less liked. In our case we are investigating four kinds of potato chips: classic, sour cream, barbeque, and cheese. Each subject will participate in several sessions. At each session a subject will eat a load food (one of the four kinds of chips). After eating the load food, the subject will rate his or her liking of each of the four kinds of chips. We anticipate large subject to subject differences. We also anticipate that ratings could differ from session to session (for example, we suspect that first session ratings could be higher than last session ratings). Each subject will be available for two sessions, and we have 24 subjects. Choose an appropriate design for this experiment.

6. Poly-3-hydroxybutyrate (PHB) is a biologically produced polymer that is becoming popular because it is biodegradable. This experiment studies the lab method used for measuring the concentration of PHB in a sample. The overall method is to digest samples of known concentration and then measure the concentration via gas chromotography. The procedure involves internal standards, recalibration of instruments, various independent dilutions, and so on. In particular, at this stage we have to deal with the possibility that everything could interact with everything else.

We will do this on three randomly chosen days. On each day we will make up eight samples of PHB. Each sample is randomly assigned to one of the combinations of four concentrations and two digestion methods. Once we have made the sample, we will measure the concentration twice on the GC.

Draw the Hasse diagram for this experiment (including superscripts and subscripts).

7. We have a computer simulation model of traffic flow that we wish to use to test how the algorithm controlling freeway ramp meters should be set. The algorithm works by producing a "random" traffic pattern with different arrival times, different styles of driving, different destinations, etc. We combine this traffic with the simulator and the metering algorithm to measure system total travel time (how long it takes for all our drivers to get through the stretch of highway we are simulating). In this experiment we vary six factors, each at two levels. (Factors are things like metering initiation threshold, maximum wait time, lane capacity, etc; we'll just call them A through F.)

A single replication of the $2^6$ design was run, and here is a halfnormal plot of the results in standard order. Which factors and/or interactions are significant?