Methods Qualifying Exam

January 2001

Instructions:

1. Do not put your name on the exam. Place the number assigned to you on the upper left hand corner of each page of your exam.
2. Please start your answer to each question on a separate sheet of paper.
3. Answer all the questions.
4. Be sure to attempt all parts of every question. It may be possible to answer a later part of a question without having solved the earlier parts.
5. Be sure to hand in all of your exam. No additional material will be accepted once the exam has ended and you have left the exam room.

PROBLEM #1

Consider a set of observations Y_i, $i = 1, 2, \ldots 400$, which are assumed to be independent and identically distributed with a mean μ and variance σ^2. NOTE that due to the large sample size involved here, you may use normal-distribution tables (distributed with this examination) for any probability calculations or decision rules required below.

1. Consider the null hypothesis $\mathrm{H}_0 : \mu = 12$ and the two-sided alternative hypothesis $\mathrm{H}_1 : \mu \neq 12$. Use the following steps to present the customary t-test of this null hypothesis based on $Y_1, \ldots, Y_{400}$ and $\alpha = 0.05$.
   (a) Write down a general formula for the t test statistic commonly used for this hypothesis test.
   (b) Write down the decision rule for this hypothesis test. Use $\alpha = 0.05$.
2. **In the context of the hypothesis test presented in (1.)**, give clear, explicit definitions of the following terms.
   (a) Type I error
   (b) Type II error
   (c) Power of the test.
3. For parts (3.) and (4.) of this question, you may assume that σ is known and $\sigma = 1$.
   (a) Calculate the power of your test for the following six values of the true parameter μ: 11.9, 11.95, 11.975, 12.025, 12.05, 12.1.
   (b) Use your results from (3.(a)) to sketch a power curve for your test. Be sure to label your axes clearly.
4. An agonomist reviews your work from (1.) through (3.) and objects, "The power you have for $\mu = 12.025$ is lower than I was hoping to get. How can I increase it?" Answer your agronomist's question, paying careful attention to:
   (a) Your specific recommendation on how to increase the power; and
   (b) Explanation (based on the ideas from parts (1.) through (3.)) of *why* your recommendation will result in an increase in power. Hint: Use the formula for the power function in answering this part of the question.
5. The 400 observations considered above represent the weight (in grams) of pecans (a type of nut). However, (unknown to you) the 400 pecans were actually collected from 10 trees, with 40 pecans picked from each tree. Also, your agronomist admits that within a given tree, pecan weights cannot be considered independent, and will have a strong *positive* correlation, due to common genetic and environmental factors. Given this additional information, answer the following questions without carrying out additional calculations.
   (a) How will this positive correlation within trees affect the expectation of the variance estimator of the sample mean that you used in part (1.(a))?
   (b) Suppose you ignored the positive correlation within trees and proceeded to use the t-test you proposed in part (1.) Will the actual values of the power of the $t-$-test be larger or smaller than the values you calculated in part (3.)? Explain.
6. In light of your answer to part (5.), your agronomist says, "OK, I see that it's wrong to use the t-test from (1.) to test our null hypothesis. What should I do instead?" Answer your agronomist's question by presenting a standard testing method that will account appropriately for the nested design described in (5.). Be sure to give clear, explicit statements of both your test statistic formula and your decision rule.

PROBLEM #2

Consider the simple first-order autocorrelated error regression model

$$Y_i = \beta_0 + \epsilon_i,\ i = 1, \cdots, n,$$

where

$$\epsilon_i = \rho\epsilon_{i-1} + \delta_i,\ \mid \rho \mid < 1$$

and δ_i's are i.i.d $(0,\sigma^2)$ random variables for $i \in \{\cdots, -2, -1, 0, 1, 2, \cdots\}$.

(i) Show that $\epsilon_i = \rho^m \epsilon_{i-m} + \sum_{j=0}^{m-1} \rho^j \delta_{i-j}$ for any finite integer m. (You can use induction principle to prove it).

(ii) Heuristically we see that $\epsilon_i \to \Sigma_{j=0}^{\infty} \rho^j \delta_{i-j}$ in some sense as $m \to \infty$ in (i). Then

$$\epsilon_i = \Sigma_{j=0}^{\infty} \rho^j \delta_{i-j} \qquad (1)$$

with probability 1. Assuming that the interchange of the expectation and infinite summation operator is justified , use (1) to determine

(a) $E(\epsilon_i)$,
(b) $\mathrm{var}(\epsilon_i)$,
(c) Covariance(ϵ_i, ϵ_j).

(iii) Use the previous results to obtain the genaralized least-squares estimate of β_0 and its variance (you can assume $n = 2$ for this problem).

Question 3.

A researcher has approached you for help with designing her study. The study is to compare measurements made on individuals while seated in an office chair. The plan is to have different individuals, randomly selected from a population of individuals, sit in the chairs. There are a number of response variables, the principal one being a measure of pressure on the chair seat. The researcher may use any number of individuals, but would like to keep the number reasonable. In each of the situations described below, individuals will be asked to do multiple "sittings," each under different conditions. It is felt that a reasonable limit on the number of sittings per individual is in the range 6 to 10.

a. There are fifteen (15) different chairs to be compared. A simple experimental design would be to have each individual sit in each chair. However, this would mean that each individual will have to do 15 sittings. One approach would be to consider the individuals as "blocks" and the chairs as "treatments, and use a Balanced Incomplete Block Design. Balanced Incomplete Block Designs have the parameters t, b, r, k, and λ. Define each of these parameters and give the values for a suitable BIBD. How many individuals are needed? Discuss briefly the randomization process that you would use.

b. Unfortunately, the researcher changes her mind. She now wants to compare only eight different chairs. However, for each chair she wants to compare the chair with and without armrests (Factor A) and for two different backrest angles (Factor B). Hence, there are now 32 different "treatments" (8 chairs x 2 levels of A x 2 levels of B). Again, it is believed that for each individual to have 32 sittings is excessive. Again, considering individuals as blocks, suggest an appropriate design. (Hint: You can think of the 8 chairs as corresponding to combinations of factors C, D, and E, each at 2 levels.) Although a comparison of the chairs is of interest, the major interest is in the main effects for armrests (Factor A) and backrest angle (Factor B) and their possible interaction. How many individuals are needed? Show at least a portion of the design and discuss briefly the randomization process you would use. Remember, no individual should have more than ten sittings.

c. Still another change. There are now only six chairs, but the number of backrest angles is changed to three levels (factor B has 3 levels). Each chair will still be used with and without armrests (factor A still has 2 levels). Hence there are 6 "treatments" for each chair. For each individual to sit in each chair under each "treatment" would require 36 sittings for each individual. Again, this believed to be unreasonable. The researcher would like each individual to sit in each chair at least once and under each "treatment" at least once. Suggest an appropriate design. For simplicity you may use A, B, C, D, E, F to represent the six "treatments," i.e., the Armrest by Backrest combinations. How many individuals are needed? Show at least a portion of the design and discuss briefly the randomization process you would use