Stat/For/Hort 571 **Chuang and Nordheim** **November 24, 1998**

Midterm II

Name:____

For the section that you *attend* please indicate:
Instructor:(circle one) Chuang Nordheim

TA: (circle one) Cong Li Zou

Instructions:

1. This exam is open book. You may use textbooks, notebooks, class notes, and a calculator.
2. Do all your work in the spaces provided. If you need additional space, use the back of the preceding page, indicating *clearly* that you have done so.
3. To get full credit, you must show your work. Partial credit will be awarded.
4. Do not dwell too long on any one question. Answer as many questions as you can.
5. Note that some questions have multiple parts. For some questions, these parts are independent, and so you can work on part (b) or (c) separately from part (a).

For graders' use:

Question	Possible Points	Score
1	18	
2	20	
3	18	
4	24	
5	20	
Total	100	

1. A test was undertaken to compare two different procedures for producing mutations in Arabadopsis (a plant). 358 randomly selected Arabadopsis seeds were treated with the first procedure and 9 mutations were produced. Similarly, 339 randomly selected Arabadopsis seeds were treated with the second procedure and 16 mutations were produced. Perform a test of the hypothesis that the rates of mutation for the two procedures are the same versus the two-sided alternative. Interpret the results.

2. A botanist conducted a study comparing the growth of several varieties of legumes. 6 plants were randomly selected from each of 5 varieties of legume. The total plant dry weight (without roots) was determined after 50 days of growth. The following data are available. The sample variances for the varieties are $s_1^2 = 0.258, s_2^2 = 0.755, s_3^2 = 0.346, s_4^2 = 0.608$ and $s_5^2 = 0.460$. SSTrt $= 17.093$.

   (a) Compute the ANOVA table. Carry out a test of the null hypothesis that the population means of all 5 varieties are equal. Interpret your results.

(b) Find a 95% confidence interval for the underlying error variance.

3. An animal scientist intends to conduct a study on the effect of a given diet on the muscle mass of young (4 month-old) chickens. It is known that muscle mass in a young chicken is distributed approximately normally with $\sigma^2 = 4^2$ gm^2. The hypotheses are $H_0 : \mu = 25 \quad$ vs. $\quad H_A : \mu > 25$. The scientist will take a random sample of n young chickens and will reject H_0 if the sample mean, $\bar{Y}$, is greater than 26.8 gm. How large must n be if the desired power is 0.90 when the true value of μ is 28.0 gm?

4. For each of the questions below, the *italicized* statement is either True or False. Indicate whether the statement is True or False and provide a justification for your response.

   (a) Suppose $Y \sim N(\mu, \sigma^2)$ and you perform a test of $H_0 : \mu = 10$ vs. $H_A : \mu \neq 10$ with a random sample of size 9. You observe $\bar{y}$ and s^2 and the subsequent t-test results in a p-value of 0.04. Suppose you now test the same hypotheses with a random sample of size 36. Suppose you observe the same values for $\bar{y}$ and s^2 as before. *You will certainly reject the null hypothesis at* $\alpha = 0.05$.

   (b) You conduct a paired comparison of the effect of a drug on the platelet counts in dogs' blood. You randomly select 7 dogs and determine the platelet counts both before and after administration of the drug on each dog. Assume that the distribution of platelet counts both 'before' and 'after' are normally distributed; thus, the distribution of the difference in platelet counts is also normally distributed. The resultant data are:

Dog	1	2	3	4	5	6	7
Before	63.2	78.2	57.5	64.4	73.6	59.7	66.6
After	58.4	78.0	55.1	64.7	70.1		

   The 'after' measurements on dogs 6 and 7 are missing since the vials containing the blood were both broken. *Due to the missing data, a good approach for conducting this test is to use a two-independent-sample t-test.* (Note you are not asked to actually perform the test here.)

(c) It is known that the lengths of mature leaves of oak trees are distributed approximately normally. You are interested in testing $H_0 : \mu = 12$ (cm) vs. $H_A : \mu \neq 12$. Within a large forest you randomly select 6 trees and from each tree you randomly select 15 mature leaves. From the 90 leaves you calculate $\bar{y}$ and s^2 and perform a t-test. The resultant value of the test statistic is $T = 2.50$. *You would feel comfortable rejecting H_0 at $\alpha = 0.05$.*

5. An investigator intends to conduct a study in which two treatments will be compared. The design is independent-sample. The distribution of observations from each treatment is normal and both distributions have the same variance. The goal of the study is to find a 95% confidence interval for $\mu_1 - \mu_2$. The investigator is considering three possible allocations of sample sizes.

(I)	$n_1 = 5$	$n_2 = 5$
(II)	$n_1 = 3$	$n_2 = 12$
(III)	$n_1 = 2$	$n_2 = 100$

   (a) Suppose σ^2 is known. Which of the three choices results in the narrowest confidence interval?

   (b) Suppose now that σ^2 is not known. Suppose that you compute the pooled sample variance (s_p^2) and that this value is the same for each possible allocation of sample sizes. Which of the three choices results in the narrowest confidence interval?