Stat/For/Hort 571 **Clayton and Lin** **December 16, 2002**

Final Exam

Name:____

For the lecture that you *attend* please indicate:
Instructor:(circle one) Clayton Lin

TA: (circle one) Gaffigan Tang Zheng

Instructions:

1. This exam is open book. You may use textbooks, notebooks, class notes, and a calculator.
2. You do not need to check the assumptions of the procedures that you use unless you are specifically directed to do so. In checking for normality it is sufficient to construct a stem and leaf display. It is *not* necessary to make a normal scores plot.
3. 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.
4. To get full credit, you must show your work. Partial credit will be awarded.
5. Some partial computations have been provided on some questions. You may find some *but not necessarily all* of these computations useful. You may assume that these computations are correct.
6. Do not dwell too long on any one question. Answer as many questions as you can.
7. Note that some questions have multiple parts. For some questions, these parts are independent; in such cases you can work, for example, on part (b) or (c) separately from part (a).

For graders' use:

Question	Possible Points	Score
1	25	
2	25	
3	15	
4	20	
5	15	
Total	100	

1. An experiment was conducted to evaluate the design of a cultivator in terms of the effect of the design on the yield of corn plants. Five different designs of cultivator were used; for our purposes we will refer to them by abbreviated codes: A, B, C, D, and E. In the experiment, 3 plots were randomly chosen for each cultivator, and the yield was measured for each plot. The data are summarized below:

Cultivator:	A	B	C	D	E
Sample mean yield:	14.06	11.26	13.59	19.44	17.68
Sample variance:	2.911	5.273	3.164	2.755	3.236

Let μ_A represent the population mean yield for corn that has received Treatment A; let μ_B represent the population mean yield for Treatment B, etc.

(a) Complete the ANOVA table for this experiment, perform the associated F test, and show that the p-value is less than 0.05.

Source	df	SS	MS
Treatments			
Error			
Total		165.278	

(b) State, in symbols, the null hypothesis that is being tested in part (a).

(c) Perform a test of $H_0 : \frac{1}{2}(\mu_B + \mu_C) = \frac{1}{3}(\mu_A + \mu_D + \mu_E)$ versus the two-sided alternative.

(d) By using Fisher's LSD method, perform a comparison of all pairs of means, and summarize your results using the display used in class.

2. A landscape ecology study investigated the characteristics of a number of plots of prairie land. For each plot they measured the "length of edge" and the number of species. The data are presented below, along with some summary statistics.

Length of edge (x):	18	10	13	29	16	13	22	9	32
Number of species (y):	7	10	14	28	12	9	10	11	25

$$\sum x_i = 162 \qquad \sum y_i = 126 \qquad \sum x_i y_i = 2648$$
$$\sum x_i^2 = 3448 \qquad \sum y_i^2 = 2200$$

(a) Consider the regression line relating y to x. Calculate the least squares estimates of the slope and intercept for the regression line relating number of species to length of edge (i.e. relating y to x).

(b) Construct the ANOVA Table for this regression problem. (Indicate Source, df, SS, and MS. You do not need to perform any tests.)

(c) Perform a test of the hypothesis $H_0 : b_0 = 10$ versus the two-sided alternative.

(d) The investigators plan to continue to examine additional plots of prairie land. If, tomorrow, they should find a plot whose length of edge is 13, give a prediction of the number of species in that plot, and give a 90% confidence interval for your prediction.

3. (a) Suppose we have a random sample of size 32 from a normal distribution with mean μ_1 and variance $\sigma_1^2 = 40$. Independent of that sample, we have another random sample of size 8 from a normal distribution with mean μ_2 and variance $\sigma_2^2 = 16$. Provide a test of $H_0 : \mu_1 - \mu_2 = 0$ versus the two sided alternative, if we observe a sample mean of 25 for the first sample, and a sample mean of 29 for the second sample.

(b) Repeat part (a), adding in the additional assumption that μ_1 is *known* to be 26.

4. This question concerns the comparison of two different rejection rules for evaluating a hypothesis. The data consist of measurements of the breaking strength of wooden boards. For each board, an increasing force is applied until the board breaks. The measurement is the amount of force required to break the board.

   A random sample of 100 observations of breaking force is available. Assume that these observations follow a $N(\mu, 800)$ distribution. Of interest is the null hypothesis $H_0 : \mu = 56$ versus the alternative $H_0 : \mu > 56$. Two different rejection rules are being considered; each has roughly the same value of α.

   Rule A Reject H_0 if $\bar{X} > 60$.

   Rule B For each observation, X_i, declare the observation to be "defective" if $X_i > 70$. Reject H_0 if the number of defective observations is ≥ 36.

   If, in fact, $\mu = 62$, find the power for each rule and determine which is more powerful.

5. Some researchers have collected a random sample of 120 observations from an unknown distribution. They would like to perform a test to see whether the data come from a standard normal distribution (their null hypothesis). To proceed, they have taken each observation and have recorded whether it was larger than 1.4, between -1.2 and 1.4, or smaller than -1.2. The observed data are:

	Number of observations
Less than -1.2	24
Between -1.2 and 1.4	85
Larger than 1.4	11

Based on this information, conduct a test of their null hypothesis.