Stat/For/Hort 571 | **Yandell and Zhu** | **19 December 2003**

Final

Name:____

For the section that you *attend* please indicate:

Instructor:(circle one) Yandell Zhu

TA: (circle one) Guang Cheng Mike Kozloski Ping Wang

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, and so you can work on part (b) or (c) separately from part (a).

For use by graders only:

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

1. (a) A food scientist wants to assay for a certain bacteria before and after making a change in pH. The response is log base 10 of the number of colony forming units (CFU). A sample is measured for bacteria (X), then the pH is changed, and after a period of time the sample is measured again (Y). Based on the data below, conduct a formal test for change in bacteria count.

	X	Y	$D = X - Y$
1	9.48	7.84	1.65
2	8.99	9.63	–0.64
3	7.61	4.88	2.73
4	8.93	5.99	2.95
5	8.53	6.72	1.82
	$\sum_{i=1}^{5} x_i = 43.55$	$\sum_{i=1}^{5} y_i = 35.05$	$\sum_{i=1}^{5} d_i = 8.50$
	$\sum_{i=1}^{5} x_i^2 = 381.36$	$\sum_{i=1}^{5} y_i^2 = 259.00$	$\sum_{i=1}^{5} d_i^2 = 22.56$

(b) The scientist would like to conduct a test for another bacteria using the same type of experiment. Assuming the variance of the difference is $\sigma_D^2 = 1$, how large a sample would the scientist need to have power 0.95 of detecting a difference of 1.5 units, with the type I error level set at 0.05?

2. (a) A food preference experiment with a certain beetle species (BS1) provided the beetles with two types of leaves, young and old. Did the beetles show a preference between young and old leaves? There were 60 beetles, 38 preferring young leaves and 22 preferring old leaves. Conduct a hypothesis test. Clearly state the hypotheses, the test statistic, and the results.

   (b) A second beetle species (BS2) was also studied against leaves in a similar manner. Here there were 48 beetles, 18 preferring young leaves and 30 preferring old leaves. Conduct a hypothesis test to see whether the two species (i.e., BS1 and BS2) differ in their preference between young and old leaves. Again, clearly state the hypotheses, the test statistic and the results.

3. Dairy scientists want to examine the effect of two dietary additives (A and B) on milk production. They have four treatment groups: 1 = control (no additives), 2 = additive A only, 3 = additive B only, 4 = additives A and B together. The following information is available:

Treatment	Sample size	Mean yield (lb/day)	Standard deviation
1	5	20	10.0
2	5	32	10.0
3	5	36	11.0
4	5	50	11.0

   (a) Conduct an overall test of treatment differences. That is, fill in the ANOVA table, state the hypotheses and conduct the test.

Source	df	SS	MS
Agent			
Error			
Total			

   (b) Contrast treatments 1 and 3 against treatments 2 and 4. Set up hypotheses and conduct a test for this contrast. Interpret your results in terms of the additives in 1 or 2 short sentences.

4. An agronomist wants to compare four cultivars of millet for yield, each with a sample size $n = 10$. The following information is available. The means for the four cultivars are $\bar{y}_1 = 32.0$, $\bar{y}_2 = 20.5$, $\bar{y}_3 = 48.2$, $\bar{y}_4 = 38.3$. The mean square error (MSE) from the one-way ANOVA table is 120.0 and the overall test of cultivar differences is significant at 5% level.

   (a) For all pairwise comparisons, compute the protected least significant distance (LSD) at an experiment-wise error level of 0.05. Order the cultivars based on the LSD. Indicate which cultivars are significantly different.

   (b) Repeat Part (a) of this problem by using the Bonferroni method on the same data. Then argue for using either Bonferroni or protected LSD method for this problem in 1 or 2 short sentences. [Note: Both approaches are "correct", you just need to justify your choice.]

5. Ecologists believe the size of a "patch" of habitat is critical to the survival of endangered species. Below are data on patch width X in km and number of butterflies Y found. The variable Z is the square root of the number of butterflies. Consider the regression model $Z_i = b_0 + b_1 x_i + e_i$, for $i = 1, \ldots, 6$.

patch ID:	1	2	3	4	5	6
x (km):	0.09	0.20	0.28	0.43	0.51	0.58
y (#):	27	39	51	63	87	102
$z = \sqrt{y}$:	5.20	6.24	7.14	7.94	9.33	10.10

In addition, the following summary statistics are available:
$\sum_{i=1}^{6} x_i = 2.0900$, $\sum_{i=1}^{6} z_i = 45.9500$, $\sum_{i=1}^{6} x_i^2 = 0.9079$, $\sum_{i=1}^{6} z_i^2 = 369.0597$, $\sum_{i=1}^{6} x_i z_i = 17.7457$.

(a) Compute the least squares estimates of the intercept (b_0) and the slope (b_1).

(b) Construct an ANOVA table and use the ANOVA table to test for a significant relationship (i.e., $H_0 : b_1 = 0$).

(c) A new patch of width 400m (i.e., 0.4km) is found. Construct a 95% prediction interval for the number of butterlies in such a patch. [Hint: You need to first construct a prediction interval based on the square-root data, and then transform back the interval limits to predict the number of butterflies.]