Stat/For/Hort 571 **Clayton and Gangnon** **November 20, 2001**

Midterm II

Name:____

For the section that you *attend* please indicate:

Instructor:(circle one) Clayton Gangnon

TA: (circle one) Cheng Li Song

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. 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.
5. Do not dwell too long on any one question. Answer as many questions as you can.
6. 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	28	
2	20	
3	20	
4	16	
5	16	
Total	100	

1. (a) An experiment was conducted to study how different insecticides might influence the yield of corn plants which have been sprayed with the insecticides. For each insecticide there were several plots used; the assignment of the insecticides to the plots was done completely randomly. The following table summarizes the information available from this experiment.

Insecticide 1	$\bar{x}_{1\cdot} = 18.3$	$s_1^2 = 1.84$	$n_1 = 8$
Insecticide 2	$\bar{x}_{2\cdot} = 20.1$	$s_2^2 = 1.94$	$n_2 = 6$
Insecticide 3	$\bar{x}_{3\cdot} = 20.2$	$s_3^2 = 0.51$	$n_3 = 5$
Insecticide 4	$\bar{x}_{4\cdot} = 19.6$	$s_4^2 = 0.38$	$n_4 = 8$

Complete the following ANOVA Table and perform a test of the null hypothesis that the population mean yields corresponding to the four insecticides are all equal, versus the alternative that they are not all equal.

Source	df	SS	MS
Insecticide			
Error			
Total		43.47	

(b) (You can answer this question independently of part (a).) PCBs are chemical compounds that are used in a variety of manufacturing settings, but that also have toxic effects on fish. An experiment was conducted in which 140 fish eggs were available. Of the 140 eggs, 38 were chosen at random to be exposed to PCB Compound A, while 102 were chosen to be exposed to Compound B. The experiment focused on the number of eggs, for each compound, that eventually matured to become adult fish with deformities. For Compound A, 7 of the resulting adult fish had deformities; for Compound B, 32 of the adult fish had deformities. If p_A represents the proportion of adult fish that have deformities resulting from Compound A, and similarly for p_B, find a 99% confidence interval for $p_A - p_B$.

Imagine using the confidence interval to examine the hypothesis: $H_0 : p_A = p_B$. What conclusion might you draw regarding that hypothesis, and what limitations exist in terms of making that conclusion?

2. Gamma radiation has occasionally been suggested as a means by which foods could be rid of bacterial contaminents. To assess this, an experiment was conducted in which 17 petri plates were used. Each plate had a large colony of *E. coli* bacteria growing on it. Two different durations of radiation exposure were used: 30 seconds, or 2 minutes. Ten plates were randomly assigned to receive 30 seconds of exposure; the remainder received 2 minutes of exposure. After exposure, standard bacteriological methods were used to determine the number of surviving bacteria on the plates. The results are as follows:

30 seconds	0	0	1	5	6	8	13	15	29	64
2 minutes	0	2	3	4	7	9	12			

(a) The researchers believe that these data do not come from normal distributions and they do not want to transform the data. Based on that, conduct a test of the null hypothesis that there is no difference between the treatments in terms of numbers of bacteria surviving, against the alternative that there is a difference between the treatments in terms of the numbers of bacteria surviving.

(b) Below are normal scores plots for these two groups. Based on these plots, comment on the extent to which these data do or do not appear to come from normal distributions.

```
          -                                                     *
        60+
          -
30 sec    -
          -
          -
        40+
          -
          -
          -                                            *
          -
        20+
          -                                      *
          -                                 *
          -                         *   *
          -                     *
         0+       2        *
            ------+---------+---------+---------+---------+---------+
              -1.20     -0.60      0.00      0.60      1.20      1.80
```

```
          -
      12.0+                                                        *
          -
2 minutes -
          -
          -                                           *
       8.0+
          -                                   *
          -
          -
          -
       4.0+                            *
          -                     *
          -
          -              *
          -
       0.0+  *
            --------+---------+---------+---------+---------+--------
                -1.00     -0.50      0.00      0.50      1.00
```

3. This experiment focuses on the concentration of a particular steroidal compound in the leg muscles of rabbits. Nine rabbits were chosen at random from a large colony; 3 of the rabbits were randomly chosen and were exercised routinely; the remaining 6 rabbits were confined in cages and not exercised. After 3 weeks, the rabbits were euthenized, the muscle of interest was removed and assayed for the concentration of the steroidal compound.

   The data are presented below:

Exercise	14	15	18			
No exercise	22	14	15	8	12	10

   Here are some summary statistics:

Group	$\bar{x}$	s^2
Exercise	15.67	4.333
No exercise	13.50	23.9

   Based on the summary statistics, the researchers are concerned that the groups do not have the same variance. Perform a test of the null hypothesis that the two groups have equal population variances against the two-sided alternative.

4. A researcher will take 19 observations from a $N(\mu, \sigma^2)$ distribution. Of interest is the hypothesis $H_0 : \sigma^2 = \sigma_0^2$ versus the alternative $H_A : \sigma^2 < \sigma_0^2$ where σ_0^2 is some value to be determined. The researcher has decided to reject H_0 if the sample variance is less than or equal to 4. For what value of σ_0^2 will her rejection rule have $\alpha = 0.10$?

5. It is thought that the milk yield of cows will decrease if they are exposed to a stressful environment. To study this, 73 cows were randomly sampled from a number of farms. For each cow, her milk yield was measured, and then she was exposed to a stressful environment (nonstop loud music) for 7 days. At the end of the 7 days, her milk yield was measured again. For each cow, let X represent her milk yield before the stress; let Y represent her milk yield after the stress, and let D represent the difference $D = X - Y$. Here are some summaries of the data (this problem continues on the next page):

```
 Stem-and-leaf of X       N  = 73
 Leaf Unit = 1.0

          4 2333334
          4 5577779999
          5 1233
          5 5777789
          6 02334
          6 5567788999
          7 0000122233333444
          7 5666777789
          8 0023
------------------------------------------------------
 Stem-and-leaf of Y       N  = 73
 Leaf Unit = 1.0

          3 00111344
          3 56678
          4 00012334
          4 6888
          5 1222233334
          5 568999
          6 0000011122224
          6 555556666677789
          7 000
          7 8
------------------------------------------------------
 Stem-and-leaf of D       N  = 73
 Leaf Unit = 0.10

          5 0111245799
          6 01223346788
          7 113
          8 1244559
          9 22233467
         10 12999
         11 05667
         12 123344555679
         13 012337
         14 123368
------------------------------------------------------
               N     MEAN   MEDIAN   STDEV
 X            73    63.73    67.42   12.24
 D            73    9.719    9.456   3.030
 Y            73    54.01    56.98   12.57
```

Using this information, provide a 90% confidence interval for the change (decrease) in milk yield due to exposure to a stressful environment. State the assumptions that you needed to make and give a justification for those assumptions to the extent possible.