ST512

Fall Quarter, 2005

Exam 2

Name:____

Directions: Answer questions as directed. Show work where appropriate. For true/false questions, circle either true or false.

1. (40 points) A horticultural experiment investigates the effects of $t = 4$ different herbicide formulations on weed growth in a completely randomized design involving a total of 20 crabgrass plants. Let y_{ij} denote the observed shoot dry weight (in cg) of the j^{th} plant 120 days after application of treatment i. Summary statistics are given below.

Treatment i	Herbicide Formulation	sample size	sample mean, $\bar{y}_{i+}$	sample std. dev, s_i	Grouping Label
1	Control	5	300	50	
2	A300	5	260	40	
3	A400	5	240	30	
4	Surflan	5	200	40	

Summary based on experiment conducted by Glenn Fain

(a) Fill in all 9 blanks in the ANOVA table below (or write "NP" if not possible). You may use the fact that $\sum_{i=1}^{4}\sum_{j=1}^{5}(\bar{y}_{i+} - \bar{y}_{++})^2 = 26000$.

Source	DF	Sum of squares	Mean Square	F
Herbicide				
Error				
Total				

(b) Consider a contrast, θ_1, quantifying the difference between the control mean and the average of the A300 and A400 means.

i. Report an estimate, $\hat{\theta}_1$.

ii. Report the standard error of the estimate.

iii. Report the sum of squares associated with the contrast

(c) Let θ_2 denote a contrast comparing the A300 and A400 means and θ_3 denote a contrast comparing Surflan mean with the average of the other non-Surflan herbicide means.

i. True/false: θ_1, θ_2 and θ_3 are mutually orthogonal.

ii. Report $SS(\theta_2) + SS(\theta_3)$. (You may want to use your answer to part (b)iii).

(d) Consider the family of all pairwise comparisons among the four means.

i. Using Tukey's procedure with familywise error rate $\alpha = 0.05$, obtain the honestly significant difference for any pairwise comparison.

ii. Use this honestly significant difference and the rightmost column in the table on the preceding page to label the treatment means with letters A,B,C,... in such a way that two means with the same letter do not differ significantly.

(e) Consider the model $Y_{ij} = \mu + \tau_i + E_{ij}$ where E_{ij} are iid $N(0, \sigma^2)$ and $\sum \tau_i = 0$.

i. Consider the hypothesis of equal mean dry weights for the four herbicides:

$$H_0 : \tau_i \equiv 0$$

Conduct an appropriate test at significance level $\alpha = 0.05$. You may use the fact that $F(0.05, 3, 12) = 3.49$.

ii. Report the least squares estimate of $\mu + \tau_4$ along with a standard error.

2. (30 points) An experiment investigates the growth of oysters. Four bags with ten oysters each are randomly placed at four underwater stations next to a power plant:

- Trt1: At the bottom of a discharge canal
- Trt2: At the top of a discharge canal
- Trt3: At the bottom of an intake canal
- Trt4: At the top of an intake canal

Average initial weight x and final weight y are measured for each of the 16 bags. (Bags serve as the experimental units.) Let $z = x - \bar{x}$ denote the difference from average of the initial weights. SAS code and output to fit an ANCOVA model appear at the end of the exam.

(a) Obtain the F-ratio for a test of equal final weights in a one-way ANOVA where initial weight z is ignored.

(b) Obtain the F-ratio for a test of equal final weights in a one-way ANOVA after controlling for initial weight z.

Treatment	Unadjusted Mean	Adjusted Mean	Std. Error
1			
–	–	–	—
4			

(c) Use the output to report the unadjusted means for treatments 1 and 4. (Use the 2^{nd} column in the table above.)

(d) For bag i, let $x_{i1}, \ldots, x_{i4}$ denote indicator variables for treatments 1-4, respectively. Consider the analysis of covariance model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 z_i + E_i$$

i. Use the fitted model to report the mean final weight, after adjustment to the average initial weight $\bar{x}$, for treatments 1 and 4. Fill in the table above, show any work below.

ii. Report the standard error for the adjusted means for locations 1 and 4. (Fill in the table, writing "NP" if it is not possible to give a number based on the provided output.)

(e) Consider the difference between mean final weights under treatments 1 and 4. Estimate this difference after controlling for initial weight. Report a standard error and a p-value for a test of no difference.

3. (30 points) An experiment measures "Ortho-P" reduction after running material through a centrifuge and adding either lime only (L) or an experimental floculant only (F), or both (LF) or neither (C). $N = 12$ total samples are randomized to the four treatment combinations and run through the centrifuge. The reductions are summarized below:

Analysis Variable : Ortho_P

treatment	N Obs	Mean	Std Dev	Variance	N
c	3	59.0000000	7.2111026	52.0000000	3
f	3	61.0000000	6.0827625	37.0000000	3
l	3	50.0000000	6.0000000	36.0000000	3
lf	3	82.0000000	6.2449980	39.0000000	3

(a) Complete the sum of squares column in the ANOVA table below. You may use the fact that the treatment sum of squares based on 3 degrees of freedom is $SS(Trt) = 1650$.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
lime	1	----------	----------		0.1432
f	1	867	----------		0.0018
lime*f	1	----------	----------		0.0036
Error	8	----------	----------		
Corrected Total	11	1978			

(b) Does adding lime improve Ortho-P reduction? Address this question using $\alpha = 0.05$. Briefly characterize the treatment effects on Ortho-P reduction. Some useful critical values are $F(0.05, 1, 8) = 5.32$, $F(0.05, 3, 8) = 4.07$, $t(.025, 8) = 2.31$.

```
proc glm;          /* OYSTER PROBLEM (#2) CODE AND OUTPUT  */
   class trt;
   model final=trt z  /solution;
   means trt;
run;  /* These data taken from Freund, Littell and Spector */
/* SAS output given below */
                              The SAS System                              1
                             The GLM Procedure

                     Class         Levels    Values
                     trt                4    1 2 3 4

                                        Sum of
 Source                      DF        Squares   Mean Square  F Value  Pr > F
 Model                        4    144.7553608    36.1888402   113.59  <.0001
 Error                       11      3.5046392     0.3186036
 Corrected Total             15    148.2600000

           R-Square     Coeff Var      Root MSE    final Mean
           0.976362      1.747523      0.564450      32.30000

 Source                      DF      Type I SS   Mean Square  F Value  Pr > F
 trt                          3     29.0450000     9.6816667    30.39  <.0001
 z                            1    115.7103608   115.7103608   363.18  <.0001

 Source                      DF    Type III SS   Mean Square  F Value  Pr > F
 trt                          3      7.6741036     2.5580345     8.03  0.0041
 z                            1    115.7103608   115.7103608   363.18  <.0001

                                            Standard
   Parameter           Estimate                Error    t Value    Pr > |t|
   Intercept        32.82685402 B         0.28398639     115.59      <.0001
   trt       1      -1.23028630 B         0.43892225      -2.80      0.0172
   trt       2      -1.36002698 B         0.40124637      -3.39      0.0060
   trt       3       0.48289720 B         0.41086022       1.18      0.2647
   trt       4       0.00000000 B          .                .         .
   z                 1.04670265           0.05492404      19.06      <.0001

NOTE: The X'X matrix has been found to be singular, and a generalized
      inverse was used to solve the normal equations.  Terms whose
      estimates are followed by the letter 'B' are not uniquely estimable.

 Level of      -----------final-----------   -------------z-------------
 trt         N          Mean       Std Dev          Mean        Std Dev
 1           4    34.4750000    3.18891309    2.75000000     3.20572405
 2           4    31.6500000    1.53731367    0.17500000     0.96046864
 3           4    30.8500000    2.95578529   -2.35000000     2.75862284
 4           4    32.2250000    4.29757684   -0.57500000     4.04917687
```