STAT 3350: Design of Experiments
Winter Semester 2006

Instructor:	Dr. Ron Hilburn
Office:	Chase 105
Telephone:	494-1981
E-mail:	rhilburn@mathstat.dal.ca

Lecture:	M.W.F. 2:35-3:25 pm
Class Room:	LSC 220
Office Hours:	M.W.F. 4:00-5:00 pm
Website:	WebCT

Course Text:	Design and Analysis of Experiments, 6th Edition by Douglas C. Montgomery

Course Description: The aim of this class is to develop the fundamental statistical concepts required for designing efficient experiments to answer real questions. Basic concepts of replication, blocking, and randomization are examined. A wide range of classical experimental designs, which include full and fractional factorial designs, randomized complete and incomplete block designs, and nested and split plot designs are introduced and evaluated. Both fixed and random factor designs are considered. The software package MINITAB, it's design of experiments module and general linear model capability, will be used througthout the course.

Method of Evaluation:	Assignments	25%
	Midterm exam	25%
	Final exam	50%

Grading Scheme:	A+	93 and above
	A	85 up to but not including 93
	A-	80 up to but not including 85
	B+	75 up to but not including 80
	B	70 up to but not including 75
	B-	65 up to but not including 70
	C+	60 up to but not including 65
	C	57 up to but not including 60
	C-	53 up to but not including 57
	D	50 up to but not including 53
	F	below 50

Note: There will be no supplemental examination.

Assignments: There will be eight assignments that provide an opportunity to use the techniques discussed in class to analyze data sets. Assignments will require hand computations using a calculator plus use of the MINITAB data analysis and graphics software package.

MINITAB Software
MINITAB release 14 is available on campus computer labs in LSC 200, LSC 2084, LSC 2087, MMASS 2018, MMASS 2019, MMASS 2020, Killam B400, Killam Learning Commons, and Chase 007.

MINITAB 14 is also available for installation on your own computer as a free download at:
http://ucis.dal.ca/services/downloads/minitab intro.html

Students with Disabilities: Students with permanent or temporary disabilities who would like to discuss classroom or exam accommodations are asked to come and see me as soon as possible. You can meet me after class or privately during office hours. For your information, the phone number for Student Accessibility Services is 4942836. Please call them and register.

STAT 3350: Design of Experiments
Winter, 2006

Week	Assignment	Topics	Chapters
1		Introduction to Experimental Design Simple Comparative Experiments	Cha 1 Cha 2-1 thru 2-3
2	1	Simple Comparative Experiments Single Factor Experiments: ANOVA	Cha 2-4 thru 2-6 Cha 3-1 thru 3-3
3		Single Factor Experiments: ANOVA	Cha 3-4 thru 3-8
4	2	Single Factor Experiments: ANOVA Randomized Block Designs	Cha 3-9 Cha 4-1 and 4-2
5	3	Randomized Block Designs Factorial Designs	Cha 4-4 Cha 5-1 thur 5-3.2
6	4	Factorial Designs	Cha 5-3.3 thur 5-5
7		Factorial Designs Midterm Exam (Cha 1-5) 2^k Factorial Design	Cha 5-6 Cha 6-1 and 6-2
8		Study Break	
9	5	2^k Factorial Design	Cha 6-3 thru 6-7
10		2^k Design:Blocking and Confounding	Cha 7-1 thru 7-8
11	6	Fractional Factorial Designs	Cha 8-1 thru 8-7
12		Designs with Random Factors	Cha 13-1 thru 13-5
13	7	Designs with Random Factors Nested and Split Plot Designs	Cha 13-6 and 13-7 Cha 14-1 thru 14-3
14	8	Nested and Split Plot Designs Review	Cha 14-4 and 14-5

DALHOUSIE UNIVERSITY FACULTY OF SCIENCE

Department of Mathematics and Statistics

STATISTICS 3350 (01) Final Examination

Date and Time: 7:00-10:00 PM, Tuesday 25 April, 2006

Name _____ **Student ID** _____

The entire exam is worth 100 points. It is closed book, but you may refer to two sheets (both sides) of notes. The exam consists of 8 problems, the number of points allocated to each part of a problem is shown in the left hand column. You may use a calculator and straight edge.

1. Consider a twice replicated 2^3 factorial design, used to investigate the effect of three factors (A,B, and C) on the process response. The data are shown below.

Run	A	B	C	Trt	Rep 1	Rep 2
1	−	−	−	(1)	550	604
2	+	−	−	a	669	650
3	−	+	−	b	633	601
4	+	+	−	ab	642	635
5	−	−	+	c	1037	1052
6	+	−	+	ac	749	868
7	−	+	+	bc	1075	1063
8	+	+	+	abc	860	1589

(3) (a) Compute values for main factor contrasts, C_A, C_B, and C_C.

(3) (b) Compute values for the main effects of factors A, B, and C. Which factor has the largest impact on the response?

(3) (c) Compute sums of squares for factors A, B, and C.

(3) (d) If $SS_E = 18,020.5$ with 8 degrees of freedom, compute the F-statistic and pValue for each factor. Which factors are significant at $\alpha = .05$?

2. An experiment was conducted to determine the effects of three pesticides (factor A) on the yield of fruit from three varieties (factor B) of citrus trees. Six trees from each variety were randomly selected from an orchard. The three pesticides were then randomly assigned to two trees of each variety and applications were made according to recommended levels. Yields of fruit (in bushels per tree) were obtained after the test period. Consider both factors to be fixed. The data are as follows:

	Pesticide		
Variety	1	2	3
1	49	50	43
	39	55	38
2	55	67	53
	41	58	42
3	66	85	69
	68	92	62

(2) (a) Write the linear statistical (effects) model for these data.

(6) (b) Compute SS_A and SS_B

(3) (c) Given that $SS_T = 3978.0$ and $SS_E = 325.0$, compute the F-statistic and pValue for the AB interaction.

(3) (d) Now, assume the AB interaction to not be significant and compute the F-statistic and pValue for testing the significance of factor A.

3. Consider a two-factor mixed model where factor A is fixed and factor B is random. The linear statistical model for this case is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

where τ_i is a fixed effect, β_j is a random effect, $(\tau\beta)_{ij}$ is assumed to be a random effect and ϵ_{ijk} is a random error. Also $i = 1, 2, ..., a$, $j = 1, 2, ..., b$, and $k = 1, 2, ..., n$.

(2) (a) Using the Rules for Expected Mean Squares outlined in the text, determine the factor A expected mean square, $E[MS_A]$, for the restricted case.

(2) (b) Again, using the Rules for Expected Mean Squares, determine the factor B expected mean square, $E[MS_B]$, for the restricted case.

(2) (c) Now determine the interaction AB expected mean square, $E[MS_{AB}]$, for the restricted case.

(2) (d) What is the proper form of the F-statistic to test for the fixed factor effects?

(2) (e) What is the proper form of the F-statistic to test for the significance of random factor B

(2) (f) What is the form of the F-statistic to test the interaction hypothesis, $H_0 : \sigma^2_{\tau\beta} = 0$?

4. Construct a design, using only 8 experimental runs, to investigate four factors (A, B, C, and D), each at two levels, effecting process yield.

(2) (a) What type of design is necessary here?

(2) (b) What is the defining relation for this design?

(2) (c) What is the resolution of this design?

(4) (d) Determine the alias structure for this design?

(1) (e) Will this design allow unambiguous, or at least acceptable determination of main effects? Explain.

(1) (f) Will this design allow unambiguous or at least acceptable determination two-factor interactions? Explain.

(2) (g) Considering only main effects and two-factor interactions, which terms (A,B,AB,CD, etc.) would you include in a regression model used to estimate effects and/or sums of squares for main factor and interaction F-tests?

5. Suppose that Nova Scotia Power purchases its raw coal from 2 suppliers and wants to determine if the sulfur content (% industry standard) of the coal is the same for each supplier. Four batches of coal are available from each supplier and three determinations of sulfur content were made from each batch. Data shown below are y_{ijk} =(% industry standard - 93). Here, suppliers (A) are fixed and batches (B) are random, so we have a mixed effects model.

	Supplier 1				Supplier 2			
Batch	1	2	3	4	1	2	3	4
	1	-2	-2	1	1	0	-1	0
	-1	-3	0	4	-2	4	-1	3
	2	-4	1	0	-3	2	-2	2
Batch totals, ($y_{ij.}$)	2	-9	-1	5	-4	6	-4	5
Supplier totals, ($y_{i..}$)	-3				3			

(2) (a) What kind of design is this?

(3) (b) What are the expected mean squares, E[MS_A], E[$MS_{B(A)}$], and E[MS_E] for the above design?

(2) (c) What is the form of the F-statistic needed to evaluate the significance of fixed factor A?

(2) (d) What is the form of the F-statistic needed to evaluate the significance of random factor B?

(3) (e) What is the form of the regression model you would use to compute sums of squares needed to calculate F-statistics and pValues for factors A and B. Identify all terms in the regression model.

6. Suppose we want to test the relative effects of four drugs (A,B,C, and D) on the reduction of fever. The drugs are prescribed to children aged 5-14 entering the outpatient ward of a hospital complaining of the flu, with fever 100°F to 100.9^{o}F. The parents are telephoned 4 hours after the administration of the drug and the reduction in fever is noted.

A	B	C	D
2.0	0.5	1.1	0.6
1.6	1.2	-1.0	1.3
2.1	0.3	-0.2	0.3

(3) (a) Compute the value of SS_{Tr}, the treatment sum of squares.

(b) Given that $SS_T = 9.1367$ for the above data, compute a value for the F statistic and pvalue. Describe the strength of evidence against H_0 of no difference between
(3) the drugs.

(6) (c) Use the Bonferroni method to find which drugs are significantly different at the $\alpha = .05$ level.

7. In a study to determine which sources of variation in an industrial process are important, two measurements are taken on yield for three operators chosen randomly. Considering both factors to be random, conduct the significance tests requested below at $\alpha = .05$. The data are as follows, with the response being percent by weight.

	Batches				
Operators	1	2	3	4	5
1	66.9	68.3	69.0	69.3	68.1
	67.4	69.8	70.9	67.2	67.7
2	66.3	68.1	69.7	69.4	65.4
	66.9	68.8	69.6	65.8	67.6
3	65.6	66.0	67.1	67.9	66.3
	66.9	66.2	68.4	65.2	67.3

```
MTB > indi c2 c11-c15
MTB > indi c3 c21-c23
MTB > let c31 = c12*c22
MTB > let c32 = c12*c23
MTB > let c33 = c13*c22
MTB > let c34 = c13*c23
MTB > let c35 = c14*c22
MTB > let c36 = c14*c23
MTB > let c37 = c15*c22
MTB > let c38 = c15*c23
MTB > regress c1 14 c12-c15 c22-c23 c31-c38

Regression Analysis: Y versus C12, C13, ...
...
Analysis of Variance
Source          DF      SS     MS     F      P
Regression      14  42.585  3.042  2.22  0.068
Residual Error  15  20.525  1.368
Total           29  63.110

Source  DF  Seq SS
C12      1   0.397
C13      1  18.911
C14      1   1.440
C15      1   0.480
C22      1   0.228
C23      1  15.665
C31      1   0.793
C32      1   1.740
C33      1   1.102
C34      1   0.701
C35      1   0.405
C36      1   0.327
C37      1   0.375
C38      1   0.020
```

(4) (a) Using the MINITAB output above, compute the F-statistic and pValue for the test of AB interaction.

.

(4) (b) Compute the the F-statistic and pValue for the test of factor A. Is the variability due to batches significant at $\alpha = .05$?

(4) (c) Compute the the F-statistic and pValue for the test of factor B. Is the variability due to operators significant at $\alpha = .05$?

8. Experimenters are interested in the effects of four different diets (A, B, C and D) on the weight gain of newborn hampsters. Three mother hampsters were identified that each had a litter of 4 newborn, so a total of 12 experimental units were used. After 6 months of treatment with one of the four diets, the weight gain (in grams) was measured for each of the twelve newborn hampsters. These data are shown in the table.

	Litter			
Diet	1	2	3	$y_{i.}$
A	15	17.9	17.5	50.4
B	14.6	17.4	14.8	46.8
C	17.7	16.0	14.2	47.9
D	10.4	12.2	14.8	37.4
$y_{.j}$	57.7	63.5	61.3	$y_{..} = 182.5$

(2) (a) What kind of design is this? What is it's purpose?

(6) (b) Assess the null hypothesis of no difference in mean weight gain for the different diets. Compute values for the test statistic and pvalue. Describe the strength of evidence against H_0.

(4) (c) Was blocking useful in this analysis? How can you tell?

```
MTB > print c1-c3
      c1       c2         c3
Row   Y      Batch  Operator
  1  66.9      1         1
  2  67.4      1         1
  3  66.3      1         2
  4  66.9      1         2
  5  65.6      1         3
  6  66.9      1         3
  7  68.3      2         1
  8  69.8      2         1
  9  68.1      2         2
 10  68.8      2         2
 11  66.0      2         3
 12  66.2      2         3
 13  69.0      3         1
 14  70.9      3         1
 15  69.7      3         2
 16  69.6      3         2
 17  67.1      3         3
 18  68.4      3         3
 19  69.3      4         1
 20  67.2      4         1
 21  69.4      4         2
 22  65.8      4         2
 23  67.9      4         3
 24  65.2      4         3
 25  68.1      5         1
 26  67.7      5         1
 27  65.4      5         2
 28  67.6      5         2
 29  66.3      5         3
 30  67.3      5         3
```