

Master's Comprehensive Exam

STATISTICS

September 9, 2006

Instructions:

Do 6 problems out of the 8 problems given below.

1. The director of a clinic wants to test for a difference across genders in the general severity ratings that graduate students assign to clients. Three male and three female clinical students (or trainees) are randomly selected to participate and each is randomly assigned four clients with whom to do an intake interview. Data for severity rating is shown below.

Male			Female		
Trainee 1	Trainee 2	Trainee 3	Trainee 4	Trainee 5	Trainee 6
49 40 31	42 48 52	42 46 50	54 60 64	44 54 54	57 62 66 71
40	58	54	70	64	

- (a) Write an appropriate statistical model for this experimental situation. State the model assumptions.
 - (b) Draw an ANOVA table and perform hypothesis tests for the gender and trainee effect.
 - (c) Compute the 95% simultaneous CI for comparing Male and Female with the Tukey method.
2. An experiment was performed to investigate the effects of training method (Factor A) and instructor (Factor B) on the mean improvement of student in the class at the end of the training program. Four different training method and five randomly selected instructors were selected. Four classes were assigned to each training-instructor combination. The calculated sums of squares are: $SSA=42$, $SSB=54$, $SSAB=47$ and $SSTO=269$.
 - (a) Write the statistical model with all specifications and give point estimates of all the parameters in the model.
 - (b) Draw an ANOVA table for the experiment and conduct hypothesis tests for each factor in the model at 0.05 level of significance.
 - (c) Now, perform an one-way ANOVA for the training method only. Is the test result compatible with the result in (b)?

3. Table below contains the results of the study by Apex Enterprises on the evaluation ratings of potential employees by its personnel officers. Five personnel officers were selected at random, and four prospective employee candidates were assigned at random to each selected officer.

<u>Officer</u>	<u>Candidate</u>				Mean
	1	2	3	4	
A	76	65	85	74	75.00
B	59	75	81	67	70.50
C	49	63	61	46	54.75
D	74	71	85	89	79.75
E	66	84	80	79	77.25

- (a) Write an appropriate statistical model for this experimental situation.
 (b) State the assumptions and estimate all the variances in the model.
 (c) Write the hypothesis for the personnel officer effect and draw a conclusion at 0.05 level.

4. A researcher conducted an experiment to compare the effects of three different insecticides on a variety of string beans. To obtain a sufficient amount of data, it was necessary to use four different plots of land. Since the plots had somewhat different soil fertility, drainage characteristics, and sheltering from winds, the researcher decided to conduct a randomized complete block design with the plots serving as the blocks. Each plot was subdivided into three rows. A suitable distance was maintained between rows within a plot so that the insecticides could be confined to a particular row. The insecticides were randomly assigned to the rows within a plot so that each insecticide appeared in one row within all four plots. The response of interest was the number of seedlings that emerged per row. The data and parts of SAS output for this study are given below.

Plot	Insecticide			Plot Mean
	1	2	3	
1	56	83	80	73
2	48	78	72	66
3	66	94	83	81
4	62	93	85	85
Insecticide Mean	58	87	80	Overall mean=75

General Linear Models Procedure
 Class Level Information

Class	Levels	Values
INS	3	1 2 3
PLOT	4	1 2 3 4

Number of observations in data set = 12

The SAS System

General Linear Models Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2270.000000	454.000000	104.77	0.0001
Error	6	26.000000	4.333333		
Corrected Total	11	2296.000000			
	R-Square	C.V.	Root MSE		Y Mean
	0.988676	2.775555	2.081666		75.000000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
INS	2	1832.000000	916.000000	211.38	0.0001
PLOT	3	438.000000	146.000000	33.69	0.0004

- Write an appropriate statistical model for this experimental situation and estimate all parameters in the model.
- Perform a hypothesis testing for the main effect.
- Write a statistical model for the completely randomized design with Insecticide effect only and draw a complete ANOVA table. What is the estimated error variance? Compare the estimate with that in (a).
- From the ANOVA table in (d), is the Insecticide effect significant? If so, provide 95% confidence interval for all possible pairwise comparisons using Bonferroni multiple comparison procedure.

5. A 2^{4-1} fractional factorial design was proposed with defining relation $I = ABCD$.

- What fraction of a 2^4 design will it be?
- What treatments are used for the design?
- Which effects can be estimated?
- What are the aliases for the main effects?
- When does one have to consider running a fractional factorial design rather than a full factorial design?

6. A mechanical engineer is studying the thrust force developed by a drill process. He suspects that the drilling speed and the feed rate of the material are the most important factors. He selects three feed rates and uses a high and low drill speed. The measurements are given in the table

	Feed Rate		
Drill Speed	0.01	0.025	0.03
Low	2.70	2.55	2.65

	3.75	2.45	2.70
High	2.85	2.85	2.90
	2.85	2.80	2.85

- (a) Write down the model for this design. Analyze the data. Which effects are significant?
- (b) Draw the main effect plots and the interaction plot. Comment.
- (c) Redo part (a) of this problem assuming that the feed rates were chosen randomly from a large pool of choices.

7. A state university developed a training program to teach general computer skills. Researchers conducted an experiment to evaluate the effects of three different incentive methods (labeled A, B, or C) on achievement during the program. The blocking variables were IQ and age of subject. The data below give the achievement scores for the participants in the experiment.

IQ	Age		
	Young	Middle	Old
High	19(B)	22(A)	25(C)
Normal	24(C)	16(B)	14(A)
Low	10(A)	12(C)	7(B)

- (a) Write down the statistical model for the design. Estimate the parameters.
- (b) Analysis the data. Which blocking factors are significant?
- (c) The program offers morning, day, and night classes. Does the time of the day matter? What design should be used to answer this question? Give an example of how this design might be run.

8. Particle-board is made from wood chips and resin. An experiment is conducted to study the effect of using slash chips (waste wood chips) along with standard chips. The researchers make eighteen boards by varying the amount of resin (factor A), the target density (factor B), and the fraction of slash (factor C). The response is the actual density of the boards produced.

Resin	Density=42		Density=48	
	0% Slash	50% Slash	0% Slash	50% Slash
6%	40.7	41.9	42.0	44.4
	45.4	45.0	46.2	47.4
9%	42.8	43.9	44.8	48.2
	46.6	50.7	44.2	49.9

- (a) Analyze these data to determine the effects of the factors on particle-board density.
- (b) Set up a design to run these observations in two blocks with ABC confounded. Don't do the analysis.
- (c) Set up a design to run these observations in two blocks with ABC confounded in Replicate I and AC confounded in Replicate II. Don't do the analysis.