First Exam, Spring 1993

Question 1

Much of the statistical methodology used in psychology and the social sciences was borrowed from early work by statisticians analyzing agricultural experiments. So we've borrowed a test question from a classic textbook of that genre--Snedecor & Cochran. A study was done to compare the weight gain in rats as a function of level and source of protein in the diet. There were two levels of protein (high and low) and there were three types of sources (beef, cereal, and pork). Each diet was fed to separate groups of 10 rats each for a week. The dependent variable is weight gain (in grams) at the end of one week.

In analyzing the resulting data, weight GAIN was regressed on five contrast-coded predictors. LEVPROT codes level of protein (+1 if high, -1 if low); ANMLVEG codes whether the source of the protein is animal (+1 if beef, +1 if pork) or vegetable (-2 if cereal); X3 codes the two different sources of animal protein (+1 if beef, -1 if pork, 0 if cereal). LEVAMVEG is the interaction between LEVPROT and ANMLVEG and LEVBFPK is the interaction between LEVPROT and BEEFPORK. The following SAS output resulted from this analysis.

```
Model: MODEL1
Dependent Variable: GAIN       1-week weight gain in grams

Analysis of Variance

                          Sum of          Mean
Source           DF      Squares        Square      F Value       Prob>F

Model             5   4612.93333     922.58667        4.300       0.0023
Error            54  11586.00000     214.55556
C Total          59  16198.93333

    Root MSE       14.64772     R-square       0.2848
    Dep Mean       87.86667     Adj R-sq       0.2185
    C.V.           16.67039
```

[continued on next page]

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	87.866667	1.89101188	46.465	0.0001
LEVPROT	1	7.266667	1.89101188	3.843	0.0003
ANMLVEG	1	1.483333	1.33714732	1.109	0.2722
BEEFPORK	1	0.250000	2.31600710	0.108	0.9144
LEVAMVEG	1	3.133333	1.33714732	2.343	0.0228
LEVBFPK	1	0	2.31600710	0.000	1.0000

Variable	DF	Type II SS	Squared Partial Corr Type II
INTERCEP	1	463233	.
LEVPROT	1	3168.266667	0.21473562
ANMLVEG	1	???	0.02228123
BEEFPORK	1	2.500000	0.00021573
LEVAMVEG	1	1178.133333	???
LEVBFPK	1	0	0.00000000

Variable	DF	Variable Label
INTERCEP	1	Intercept
LEVPROT	1	hi protein vs lo prot contrast
ANMLVEG	1	source: animal vs vegetable
BEEFPORK	1	source: beef vs pork
LEVAMVEG	1	level by anim vs veg source
LEVBFPK	1	level by beef vs pork

Table of Means

Level of LEVEL	Level of SOURCE	N	GAIN Mean	GAIN SD
high	beef	10	100.0000000	15.1364167
high	cereal	10	85.9000000	15.0218360
high	pork	10	???	10.9163486
low	beef	10	79.2000000	13.8868443
low	cereal	10	83.9000000	15.7088086
low	pork	10	78.7000000	16.5465673

A. Write out the complete source table for a two way analysis of variance of these data, including the omnibus main effect test for source of protein and the omnibus interaction test. Make sure to include PRE values for each F* statistic. Notice that there are a few pieces of missing information in the printout, marked by ???, which you must calculate in other ways in order to complete the source table.

B. Compute the missing cell mean for the (high,pork) condition.

C. Draw a rough graph of the cell means and write a short paragraph giving a summary of the experimental results. In this paragraph you need only discuss the reliable differences that the analysis revealed.

D. The researcher now wants to conduct a post hoc simple effects test, examining whether the animal versus vegetable difference is reliable when the level of protein is high. Compute the SSR for this test, the resulting F*, and the value of PRE. Indicate whether the resulting F* exceeds the critical value.

E. Compute a 95% confidence interval for the difference between the animal and vegetable means. This question requires a bit of thought.

F. Another researcher plans on conducting a replication of this study, but this researcher only has enough rats for an n of 6 in each cell. Based on the analysis of the data reported above, what can we anticipate the power to be in this new test for testing the level of protein by source (animal versus vegetable) interaction (LEVAMVEG)?

Question 2

You read a journal article about the effect of early parental protection on disordered eating. The researchers administered the Eating Disorders Inventory (EDI) to 60 subjects. Twenty of these subjects came from families with high parental protection, twenty from families with moderate protection, and twenty from low parental protection families. The following table of means is presented in the article:

	Level of Protection		
	Low	Moderate	High
Mean EDI	70	35	50

In the text of the article, you are told that these three means reliably differ from each other (F (2,57) = 4.93; $p < .05$). The authors argue from these means and the reported F statistic that the mean level of EDI decreases as level of protection increases, although they fail to report whether the single degree of freedom contrast that would support their conclusion is reliable. You are concerned about the fact that they only report the omnibus test and so want to ask some further questions about their data.

A. What is the slope and SSR for the contrast that would test whether there is a decrease in EDI with higher levels of protection? Interpret the value of the slope. (Make sure you tell us the contrast weights that you use.)

B. Your own theoretical orientation leads you to expect that moderate levels of protection ought to produce the best (i.e., lowest) scores on the EDI and that too much (i.e., high) or too little (i.e., low) parental protection ought to produce higher EDI scores. What is the value of SSR associated with the contrast that would test your expectation? (Again, tell us what contrast weights you use.)

C. Given the two SSR's that you have just computed and the omnibus F statistic reported in the text of the article, write out the full source table for the analysis of variance, including your two single degree of freedom contrasts (You need not compute the PRE's. Just the sums of squares, degrees of freedom, mean squares, and F^*'s.)

D. Based on this source table, what do you think of the authors' conclusions from their data? Write a paragraph that you think should be the results section of their article, based on their data and your re-analysis.

Question 3

A cognitive psychologist is interested in factors affecting the transfer of skill from one task to another. He believes that if a subject trains on one task, the acquired skill should transfer to a second task only if the second task is similar to the first. Additionally, skill transfer should occur most efficiently if the subject is asked, during training on the first task, to verbalize the concrete steps involved in the acquired skill. Finally, the researcher believes that verbalization should make a difference on skill transfer only if the second task and the first are similar to each other. If the two tasks are dissimilar, then verbalization should not increase the amount of transfer as a function of training.

To test these ideas, he conducts an experiment in which subjects are trained on one task and then their skill on a second is assessed. The following three independent variables are crossed in a factorial design:

Amount of training on task 1:	none, 10 min, 20 min
Similarity of task 2 to task 1:	dissimilar, similar
Verbalization during task 1 training:	yes, no

Ten subjects are randomly assigned to each of the resulting 12 cells of the research design.

Assuming that the researcher's hypotheses are correct and that he conducts a traditional three-way analysis of variance (albeit with single degree of freedom contrasts) on the data he collects, answer the following questions.

A. Write out the rows of the resulting source table, indicating for each row only the source of the sum of squares (i.e., the effect that is tested) and the associated degrees of freedom. (Be clear about what the contrast is for each effect listed in the table, e.g., linear effect of amount of training.)

B. Which of the single degree of freedom tests in this source table do you expect to be reliable, assuming that the data confirm the researcher's hypotheses? Circle each row in the table where you expect a reliable F*.