STAT 5200 Fall 2018 Midterm Exam

Do not open the exam until you are instructed to do so.

Directions: You have 70 minutes to complete the exam. You may use your calculator and a single page (both sides) of handwritten notes, but no laptops or wireless-capable devices are allowed. Be concise with all your responses (no more than 1-2 sentences are needed for each question). You may use 2 or 3 decimal places in all calculations. The point-worth of each question is given, and the total points sum to 100.

Student Name: Solutions

SAS Output: Partial SAS code, output, and graphics (all clearly identified by title or output number) from certain models are provided in a separate handout, and are necessary for some of the questions on the exam.

Statistical Significance: For all significance tests on this exam, use significance level $\alpha = 0.05$. Where multiple hypotheses are concerned, control the strong family-wise error rate at $\alpha = 0.05$.

Study: 96 elderly American male subjects with the same brand of hearing aid (and approximately the same degree of hearing loss) were recruited by the hearing aid manufacturer in a study of their product's performance under various conditions. There were 4 recordings of the same male speaker reading the same list of 50 words under various conditions (SAS variable '**`recording`**'), and 24 subjects were randomly assigned to each of the recordings. The recordings were arbitrarily numbered 1-4. While listening to their assigned recording, each subject repeated back each word as they heard it. An observer marked each word on the list as correctly or incorrectly identified by the subject, and the subject's score (SAS variable '**`score`**') was their total number of correctly identified words.

Question 1: (4 points) Identify the key feature of this study that makes it a good experiment.

Subjects were randomized to treatments

Question 2: Refer to the code and output for SAS Output 1.

a) (10 points) Write out an appropriate model corresponding to this SAS code, using the means model parameterization. Define any symbols (parameters or letters) you use, and specify the range of any subscripts you use.

+4 model: $Y_{ij} = \mu_i + \varepsilon_{ij}$

+4 define symbols: Y_{ij} = Score of Subject j in recording i; μ_i = mean of recording i; ε_{ij} = random error

+2 subscript ranges: $i = 1, \ldots, 4$; $j = 1, \ldots, 24$

b) (10 points) Identify two assumptions made by this model (in Question 2a above), referring to a specific component of the model. For each assumption, comment briefly on what the evidence in SAS Output 1 suggests regarding the appropriateness of the assumption. (The code and output for SAS Output 2 may also be helpful.)

+2 i) Assumption: ε's normally distributed

+1 Piece of evidence: Q-Q plot (or normal prob. plot, or resid-quantile plot, or residual histogram)

+2 What evidence suggests: assumption appears appropriate (or not grossly violated)

+2 ii) Assumption: ε's have constant variance

+1 Piece of evidence: residual-predicted plot (or residual plot)

+2 What evidence suggests: assumption appears appropriate (or not grossly violated)

Question 3: (6 points) Why must the assumptions you mention in Question 2b above be addressed before proceeding with statistical inference?

Inference not valid if assumptions not met

Question 4: Refer again to the code and output for SAS Output 1. Report the following:

a) (6 points) The numeric value of the test statistic for the null hypothesis 'mean score is the same for all four recordings.'

$F = \frac{MS_{model}}{MS_{error}} = \frac{334.49}{62.37} = 5.36$

+2 for any ratio (symbols not needed)

+4 correct numbers + arithmetic

b) (3 points) The sampling distribution for this test statistic.

$F_{3,92}$

(+1 F; +1 num. DF = 3; +1 denom. DF = 92)

c) (1 point) The p-value for the null in part (a) above.

.0019

d) (5 points) The conclusion of this test of significance, in the context of this application.

mean score is not the same for all four recordings

(+2 if only "reject null" or similar without context)
(+3 if conclude mean scores are all different)

e) (4 points) The percentage of variation in subject score explained by the differences among the four recordings.

$R^2 = \frac{SS_{model}}{SS_{total}} = \frac{1003.46}{6741.63} = .148$ or about 15%

+2 for any ratio (symbols not needed)

+2 calculation

Question 5: (12 points) Refer to the code and output for SAS Output 3. The LSMEANS statement (which does not depend on the MEANS statement) produces the following table of raw, unadjusted p-values. We are interested in all pairwise comparisons of recording means, and we want to control the strong family-wise error rate at $\alpha = 0.05$ while maintaining the highest statistical power. Based on this interest and using the appropriate table in SAS Output 3, circle which of these unadjusted p-values **would** be called statistically significant after appropriate adjustment. (Here, LSMEAN numbers 1-4 are the same as recording numbers 1-4.)

Least Squares Means for effect recording
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: score

i/j	1	2	3	4
1		0.3632	0.0014	0.0022
2			0.0196	0.0282
3				0.8841
4				

-3 for each missing or extra circle

Question 6: (12 points) Now you deserve some additional information about the four recordings. They are actually based on the combinations of the following two other factors: the accent used by the speaker in the recording (either American or British; SAS variable '**Accent**') and the type of background noise in the recording (either soft radio static or soft classical music; SAS variable '**Background**'), as summarized in this table:

recording	1	2	3	4
Accent	American	British	British	American
Background	Static	Music	Static	Music

Using parameters you defined in Question 2 above, define a contrast ψ, where H_0: $\psi = 0$ corresponds to a test of whether speaker accent has any effect on subject score.

H_0: American mean = British mean

H_0: $\frac{\mu_1 + \mu_4}{2} = \frac{\mu_2 + \mu_3}{2}$

H_0: $\mu_1 - \mu_2 - \mu_3 + \mu_4 = 0$

$\psi = \mu_1 - \mu_2 - \mu_3 + \mu_4$

(or any $k \cdot \psi$ for scalar $k \neq 0$)

(+5 for any $\sum w_i \mu_i$ s.t. $\sum w_i = 0$)

+4 correct understanding

+4 express as H_0

+4 write as contrast

Question 7: With the additional information about factors **Accent** and **Background** from Question 6 above, refer to the code and output for SAS Output 4, as well as the initial description of the study (on the first page of this exam).

a) (4 points) This is a 2 by 2 factorial design with 24 replicates at each factor level combination. (Fill in the blanks.)

(+1 each; +2)

b) (4 points) What is the experimental unit in this design?

Subject

(+1 for response variable, score or # correct words)

c) (4 points) The measurement unit in this design is not the same as the experimental unit. What is the measurement unit in this design?

word on list

(+2 for response variable, score or # correct words)

d) (7 points) The interaction term is significant here (P-value 0.0002). Without referring to the interaction plot, what does it mean to say there is an interaction here?

The effect of Accent (on Score) depends on Background

-or- The effect of Background (on Score) depends on Accent

-or- There is a synergistic effect of Accent and Background (on Score) that can not be explained by their additive effects alone

e) (7 points) An observer sees that the P-value for **Accent** (0.4554) is not significant, and says that you can conclude that speaker accent does not affect score. Referring to specific evidence in SAS Output 4, explain clearly why the observer's conclusion is wrong, and how speaker accent does affect score.

(+3 evidence; +4 discussion)

The interaction plot shows that for music background, British accent gives higher score, while for static background, American accent gives higher score.

(may also refer to p-value of interaction term)

Question 8: (1 point) What topic(s) did you study most that did not appear on this exam?

(anything)