STAT 5200 Fall 2018 Final Exam

Directions: You have 120 minutes to complete the exam. You may use your calculator and two pages (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). The point-worth of each question is given, and the total points sum to 100.

Student Name: Solutions

Separate Handout: Five studies (clearly numbered 1-5) are described in a separate handout for this exam. Variable names in resulting data sets are **bolded** in the handout. Total sample sizes reported in the handout refer to numbers of rows in the resulting data sets. For some of these studies, partial output using SAS procedures is provided in the handout. Each question on the exam clearly refers to a specific study by number.

Statistical Significance: For all significance tests on this exam, use significance level $\alpha = 0.05$. You may assume that relevant model assumptions are satisfied whenever a significance test is used on this exam. Where multiple hypotheses have been or are to be tested, control the strong family-wise error rate at $\alpha = 0.05$.

Question 1: (12 points) Refer to Study 1 (bakery) in the handout. Circle which of the following "named designs" below best describes this study, and briefly explain which features of the study lead you to this decision. Sketch out a table or diagram below [with sample randomization(s); not a Hasse diagram] to help explain your decision.

i. Randomized Complete Block Design
ii. Latin Square Design
iii. Split Plot Design
iv. Split Split Plot Design
v. Strip Plot Design
vi. Repeated Measures Design

+4 (circle) — iii. circled

[max +8/12 for (v), if miss the nested randomization]

+4 (explain) { one blocking factor (day), two treatment factors (recipe, temp), one nested randomization (temp) → KEY

+4 (table)

Day	M			T			W		
Recipe	1	2	3	1	3	2	3	2	1
Temp	L	H	L	H	H	H	L	L	H
	H	L	L	H	L	L	H	H	H
	L	H	H	L	H	L	L	L	L
	H	L	H	L	L	H	H	H	L

one batch; one scoop

(Several ways to sketch this design; okay if not "randomized")

[Must have each Temp twice within each batch]

Question 2: (3 points) Referring to Study 1 (bakery), is this design balanced? Why or why not?

+1 { Yes

+2 { Same # replicates (2) at each factor level combination

Question 3: (12 points) Refer to Study 2 (call center) in the handout. Circle which of the following "named designs" below best describes this study, and briefly explain which features of the study lead you to this decision. Sketch out a table or diagram below [with sample randomization(s); not a Hasse diagram] to help explain your decision.

i. Randomized Complete Block Design
ii. Latin Square Design
iii. Split Plot Design
iv. Split Split Plot Design
v. Strip Plot Design
vi. Repeated Measures Design

+4 (circle) — (ii) circled

[max +6/12 if (vi), ignoring sample size / 6]

+4 (explain)

two blocking factors (worker, day), one treatment factor (script), all with same # of levels (4)

+4 (table)

Worker

Day	1	2	3	4
M	A	B	C	D
T	B	C	D	A
W	C	D	A	B
R	D	A	B	C

[Several ways to sketch this design; okay if not "randomized", as here; must have each script (A-D) once within each Day and within each Worker]

Question 4: Refer to Study 3 (schools). Reading comprehension scores were recorded for 350 students. Administrators plan to perform a two-sample t-test comparing the scores for the 173 students from reading program X with the scores for the 177 students from reading program W.

(a) (3 points) Is this an appropriate analysis plan? NO

(b) (6 points) If so, explain clearly what assumptions must be satisfied. If not, explain clearly why it is not appropriate, and discuss briefly what a more appropriate analysis plan would be.

[If "yes" in (a), max +3/6 here for normal/constant variance]

+3 { why: Students are measurement units, not experimental units

+3 (one or the other okay) { Plan: Average student score within school, and compare (t-test or ANOVA okay) two X schools to two W schools

-or-

Fit a nested model accounting for random School term nested in Program

Question 5: Refer to Study 4 (grain) in the handout.

(a) (4 points) Explain clearly why beetle type should be considered a random effect even though its levels were not selected at random.

these five types of beetles can be considered representative of all types of interest

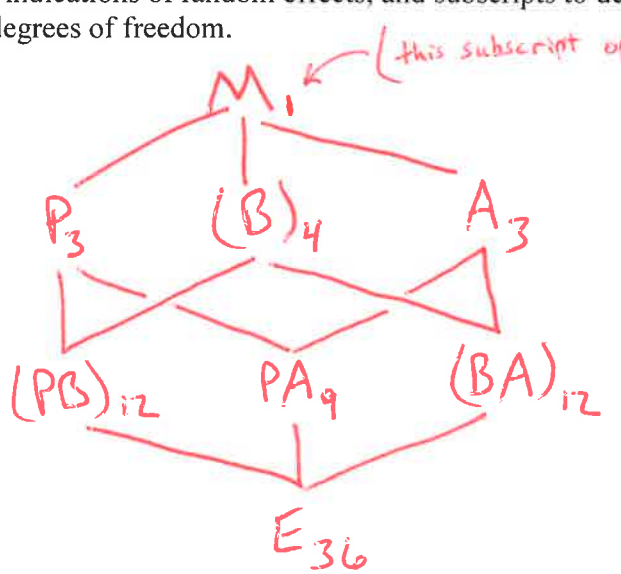
(b) (8 points) Add columns of letters to the table below to make true the statement, "LS-means with the same letter are not significantly different" while controlling a meaningful family-wise error rate at 0.05.

Least Squares Means							
Effect	**P**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	
P	**1**	9.4010	1.3345	12	7.04	<.0001	A
P	**2**	12.2116	1.3345	12	9.15	<.0001	A B
P	**3**	13.9981	1.3345	12	10.49	<.0001	B
P	**4**	18.0242	1.3345	12	13.51	<.0001	C

[Exact letters don't matter, but should have these same three groups clearly labeled; -2 for each missing or extraneous letter]

Question 6: Refer to Study 4 (grain) in the handout.

(a) (12 points) Draw a Hasse diagram for this design. Be sure to include all necessary lines, indications of random effects, and subscripts to denote corresponding degrees of freedom.



[-1 for each missing or extraneous piece, including term, DF, line, or parentheses; order of terms in each row doesn't matter]

(b) (4 points) What is the test statistic for the effect of pesticide, in terms of the mean squares (MS symbolically)? Also report its sampling distribution.

+2 ratio { $\frac{MS_P}{MS_{PB}}$ ~ $F_{3,12}$ } +2 dist'n

[Full points based on part (a)]

(c) (7 points) An observer sees the p-value for P in the partial SAS output and comments, "There is significant evidence that the four pesticides' effects are all different from each other." Explain clearly why the observer is wrong, and provide the correct conclusion (in context) based on this p-value.

+3 { Why: Some pesticides could have same effect (just not all; lots of ways to say this; may refer to groupings in #5b here)

+4 { Conclude: The pesticides' effects (may specify "on y") are not all the same (after accounting for other factors)
-or-
P has significant effect on y

(d) (4 points) Which of the effects (if any) tested in this experiment will require an approximate F test? How do you know?

+2 { B
+2 { no unique next random term below in Hasse Diagram

Question 7: Refer to Study 4 (grain).

(16 points) Using parameters in the effects model

$$Y_{ijk} = \mu + P_i + A_j + PA_{ij} + B_k + PB_{ik} + BA_{jk} + \varepsilon_{ijk},$$

construct a contrast ψ (as a linear combination of parameters; do not include any parameters with zero coefficients) such that "H_0: $\psi=0$" addresses the question of whether the difference between pesticides 1 and 3 is the same for both application methods 2 and 4. (Since this question implicitly averages over all beetle types, you can ignore B and its interactions in constructing this contrast.)

$$\mu_{ij} = \mu + P_i + A_j + PA_{ij}$$

+8 (interpret)

$$H_0: \underbrace{\mu_{12} - \mu_{32}}_{P_1 - P_3 \text{ at } A_2} = \underbrace{\mu_{14} - \mu_{34}}_{P_1 - P_3 \text{ at } A_4}$$

-5 if used "+" instead of "−" here

+7 (effects parameterization)

$$H_0: (\mu + P_1 + A_2 + PA_{12}) - (\mu + P_1 + A_4 + PA_{14})$$
$$= (\mu + P_1 + A_4 + PA_{14}) - (\mu + P_3 + A_4 + PA_{34})$$

$$H_0: PA_{12} - PA_{32} - PA_{14} + PA_{34} = 0$$

+1 (solution)

$$\psi = PA_{12} - PA_{32} - PA_{14} + PA_{34}$$

okay to have this multiplied by any nonzero constant; min. score +6/16 for any linear combination of parameters where coefficients sum to zero; do not need to show all steps given here

Question 8: (4 points) Refer to Study 5 (bacteria), including the SAS code and partial output provided in the separate handout. Note the lack of degrees of freedom and significance results in the Type III table. Using appropriate vocabulary from this course, explain briefly what went wrong with the design and/or analysis of this experiment. (There is no error in the SAS code, and the question is not asking how to fix anything – just diagnose the problem.)

condition and strain are confounded

(or equivalent, like they aren't crossed and same levels co-occur)

[+1/4 if only say lack of replication should preclude interaction]

Question 9 (These do not refer to any specific study or output on this exam.)

(a) (2 points) If a treatment effect is truly nonzero, what do we call the probability that a designed experiment will identify the estimated effect as statistically significant?

power

(b) (2 points) In a designed experiment, if there is no true treatment effect, what do we call the probability of observing (just by chance) a result (or test statistic) at least extreme as what is observed in the experiment?

p-value

[+1/2 for "significance level" since that's α]

Question 10: (1 point) What topic(s) did you study most that did not appear on this exam?

(anything)