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: ____

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

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

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

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?

(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.

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.

(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
P	**2**	12.2116	1.3345	12	9.15	<.0001
P	**3**	13.9981	1.3345	12	10.49	<.0001
P	**4**	18.0242	1.3345	12	13.51	<.0001

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.

(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.

(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.

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

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: ψ=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.)

ψ =

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.)

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?

(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?

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

Handout for STAT 5200 Fall 2018 Final Exam

<u>Study 1</u>: A commercial bakery has three new **recipes** (coded 1, 2, 3) for cakes, and two different **temperatures** (coded L, H) at which the cakes could be baked. The bakery is interested in identifying how recipe and temperature affect cake height. Three **days** (coded M, T, W) are set aside to study this. On each day, one large batch of cake batter is prepared from each recipe (so there are three batches each day; you can think of each batch being randomly assigned a recipe). From each batch that day, four large scoops of batter are extracted, and each scoop is poured into a separate pan. These are large pans, and each oven in the bakery can only hold one pan. Then each pan is randomly assigned a baking temperature in an oven, with both temperatures used twice for each recipe each day. Twelve pans (and twelve ovens) are used each day (two cakes at each of two temperatures for each of the three recipes). The **height** of each baked cake is measured after cooling. There is a total sample size of 36 in this study – 2 cakes from each of 3 recipes at each of 2 temperatures on each of 3 days. Based on past studies, the bakery is comfortable assuming that the exact pan and oven used does not affect anything.

<u>Study 2</u>: Workers at a call center make phone calls to potential customers, and each worker reads from a script to try to convince customers to purchase the product being sold. For a certain new product, the call center has created four possible **scripts** (coded A, B, C, D) and is interested in determining which of these scripts will result in the highest number of product sales. As a preliminary step in their selection of the best script to use, the call center identifies four **workers** (coded 1, 2, 3, 4) to participate in a pilot study. On each of four **days** (coded M, T, W, R) of the pilot study, each worker is assigned a script to use. This assignment is done in such a way that each script gets used by a different worker each day, and each worker eventually uses each script. The number of product **sales** made by each worker each day is recorded, with a total sample size of 16.

<u>Study 3</u>: Administrators in a large school district are interested in comparing two reading **programs** (coded X, W) for first-graders. They randomly choose four elementary **schools** (coded 1, 2, 3, 4) from the district, and randomly assign two of those schools to each of the programs. Within each school there are between three and five first-grade classes, comprised of between 18 and 25 students each. After six weeks on the reading program, each student in each class takes a standard reading comprehension test, and their **score** is recorded. (Scores are recorded for 350 individual students.)

Study 4: A study is conducted to examine how four pesticides (factor **P**, coded 1, 2, 3, 4) reduce beetle infestations in stored grain. There are many beetle types of interest, and the researchers choose five beetle types (factor **B**, coded 1, 2, 3, 4, 5) to use in their experiment. The five selected beetle types were not selected at random from all beetle types known to infest stored grain, but the researchers are confident these five types are representative of all beetle types known to infest stored grain. There are four application methods (factor **A**, coded 1, 2, 3, 4), i.e., there are four different ways each pesticide could be applied. At each of the 80 combinations of factors P, B, and A, a single large scoop of grain is stored for a certain amount of time in a large container. (Each scoop is placed in a container, and then 100 individual beetles of the assigned type are added to it. Then the assigned pesticide is applied using the assigned application method.) The response variable is **Y**, corresponding to the number of beetles surviving in the scoop. The total sample size is 80.

Source	Mean Square	Expected Mean Square	F Value	Pr > F
P	260.962182	Var(Residual) + 4 Var(P*B) + Q(P,P*A)	17.34	0.0001
A	19.304449	Var(Residual) + 4 Var(B*A) + Q(A,P*A)	1.53	0.2564
P*A	15.398121	Var(Residual) + Q(P*A)	1.89	0.0850
B	97.318627	Var(Residual) + 4 Var(B*A) + 4 Var(P*B) + 16 Var(B)	4.99	0.0149
P*B	15.053532	Var(Residual) + 4 Var(P*B)	1.85	0.0765
B*A	12.585604	Var(Residual) + 4 Var(B*A)	1.55	0.1528
Residual	8.137886	Var(Residual)	.	.

Differences of Least Squares Means									
Effect	**P**	**_P**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adjustment**	**Adj P**
P	**1**	**2**	-2.8106	1.2269	12	-2.29	0.0409	Tukey-Kramer	0.1548
P	**1**	**3**	-4.5971	1.2269	12	-3.75	0.0028	Tukey-Kramer	0.0128
P	**1**	**4**	-8.6231	1.2269	12	-7.03	<.0001	Tukey-Kramer	<.0001
P	**2**	**3**	-1.7865	1.2269	12	-1.46	0.1710	Tukey-Kramer	0.4912
P	**2**	**4**	-5.8125	1.2269	12	-4.74	0.0005	Tukey-Kramer	0.0023
P	**3**	**4**	-4.0261	1.2269	12	-3.28	0.0066	Tukey-Kramer	0.0290

<u>Study 5</u>: In an effort to understand how four bacteria strains of interest grow under various conditions, researchers prepare three cultures of each of the four **strains** (coded A, B, C, D). Each strain is randomly assigned to one of four **conditions** (coded L, M, H, X), so that all three cultures for a given strain are grown under the same condition. After 48 hours, the **amount** of bacteria in each of the twelve cultures is measured. (You may assume that this measurement is standardized in the sense that all cultures had the same initial amount.) The researchers want to know how **condition** and **strain** together affect **amount**. The sample size is 12 (see SAS code below).

```
data bacteria; input condition $ strain amount @@; cards;
   L A 4   L A 5   L A 6
   M C 6   M C 6   M C 5
   H D 3   H D 4   H D 3
   X B 8   X B 9   X B 3
;
proc glm data=bacteria;
  class condition strain;
  model amount = condition|strain;
run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	17.66666667	5.88888889	1.96	0.1982
Error	8	24.00000000	3.00000000		
Corrected Total	11	41.66666667			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
condition	0	0	.	.	.
strain	0	0	.	.	.
condition*strain	0	0	.	.	.