STATISTICS 642 - ASSIGNMENT 1 - SPRING 2022 - SOLUTIONS

1. (20 points)

   a. To reduce the size of the standard deviation associated with extraneous factors, σ_e, the selection of uniform experimental units is crucial. Thus, the 19 volunteers are divided into four groups of individuals constructed of individuals of similar age and the same gender. The two oldest males and the oldest female, individuals 9, 14, and 15, were excluded from the study. Treatments A and B are then randomly assigned within each group. By constructing groups of EU's as described above and randomly assigning the treatments within the groups, we can reduce the amount of difference in the individuals assigned to Treatment A and B prior to applying the treatments. Thus, we can reduce the experimental error, σ_e. When modeling the data, we generally would take into account the actual age (covariate) and not just the two groups of ages.

   b. The numbers under Method A and B identifies the Individual to received the treatment.

	Method A	Method B
Group 1 (Male, Age ≤ 35)	10, 11	17, 18
Group 2 (Male, $35 < \text{Age} < 60$)	1, 2	7, 13
Group 3 (Female,Age ≤ 35)	4, 16	5, 12
Group 4 (Female, $35 < \text{Age} < 60$)	3, 19	6, 8

- You could also use a paired analysis with 8 pairs of volunteers but it would yield fewer degrees of freedom for error in the analysis.

2. (20 points) There are a total of $18 = (2)(3)(3)$ treatments that the EPA wants to evaluate. The treatments consist of the 18 combinations of levels from the three factors: Flow Rate, Filter Diameter, and Fluid Temperature. The 18 treatments are presented in the following table.

Treatments

T1	T2	T3	T4	T5	T6	T7	T8	T9
R5D.5T75	R5D.5T100	R5D.5T125	R5D1T75	R5D1T100	R5D1T125	R5D2T75	R5D2T100	R5D2T125
T10	T11	T12	T13	T14	T15	T16	T17	T18
R10D.5T75	R10D.5T100	R10D.5T125	R10D1T75	R10D1T100	R10D1T125	R10D2T75	R10D2T100	R10D2T125

Because of potential differences in the types and quantities of impurities from batch to batch of cooling fluid, the experiment will use four batches of cooling fluid with 18 test runs per batch. This will allow all 18 treatments to be evaluated within each batch of cooling fluid. The experiment could be conducted as follows with the following assignment of treatments to test runs for each of the four batches. This would be a RCBD with Blocks being the Batch and 1 rep of each treatment per block.

	TEST RUN																	
BATCH	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	T9	T10	T13	T4	T16	T7	T18	T2	T12	T8	T15	T6	T17	T1	T11	T3	T5	T14
2	T1	T14	T6	T11	T13	T10	T12	T8	T2	T15	T5	T3	T9	T17	T16	T4	T18	T7
3	T11	T15	T8	T14	T17	T12	T2	T9	T13	T3	T5	T6	T7	T18	T1	T4	T10	T16
4	T6	T13	T7	T18	T16	T1	T8	T10	T9	T12	T2	T15	T5	T11	T4	T3	T17	T14

3. (20 points)

   a. The major problem with this experiment is that there is only one replication per treatment because the treatments were randomly assigned to the group of 25 students in the class and not to the individual students. This is a problem because the 25 students may not represent a random sample of students due to self-selection of the students to the class. There may be subgroups of the students which have very similar academic pursuits and study habits. Also, all 25 students in the class will have the same instructor, that is, the effect of instructor is confounded with the effect of method of instruction.

   b. There are numerous ways that this experiment could be redesigned to overcome some of the above problems. I will list two:

      - Design I: Randomized Block Design: Have three or more instructors, with each instructor randomly assigned to three classrooms. The three classrooms are then randomly assigned to a single instruction method. Thus, each instructor applies all three instruction methods but each classroom receives a single instruction method. The blocks are the instructors, the EU's are the classrooms, and the individual students are the measurement units.
      - Design II: Latin Square Design: Have three instructors and three classrooms. The three instruction methods would be applied by each instructor in each classroom. The order in which the classrooms receive the instruction would be one of three sequences. This design could be replicated for multiple sets of instructors and classrooms. The blocks are instructors and classrooms. The EU's are an instructor-classroom combination. The MU's are the individual students.

4. (20 points)

   a. $\binom{24}{6} \times \binom{18}{6} \times \binom{12}{6} \times \binom{6}{6} = (134596)(18564)(924)(1) = 2.308743 \times 10^{12}$
   b. $\binom{24}{6} \times \binom{18}{5} \times \binom{13}{7} \times \binom{6}{6} = (134596)(8568)(1716)(1) = 1.978923 \times 10^{12}$

   - Conducting a permutation would be time consuming. Thus, only a random sample of all possible combinations would be used to obtain the p-value using a permutation test.

5. (2 points each) From the description of the experiment we have

   1. Treatment Factors: F_1 : Brand; F_2 : Thickness of Covers; F_3 : Type of Cover Materials; Blocking Factor: F_4 : Testing Facility
   2. EU's: A golf ball
   3. Covariates: None in this experiment. However, if the testing would have been done outside at a golf course, temperature and humidity would be possible covariates.
   4. Levels: $F_1:\ B_1,\ B_2,\ B_3,\ B_4$; $F_2:\ T_1,\ T_2,\ T_3$; $F_3:\ M_1,\ M_2$; F_4: Fac_1, Fac_2
   5. MU's: A randomly selected spot on a golf ball
   6. Blocking: Two testing facilities
   7. Treatments: The $(4) \times (3) \times (2) = 24$ combinations of the three factor levels
   8. Replications: There are 3 replications with a given replication consisting of 3 golf balls of a given brand with a given Thickness and Type of Cover evaluated at a given testing facility
   9. Confounding: None were identified in this experiment.
   10. Response: Compressive strength of golf ball cover
   11. Subsampling: set of 5 measurements at random spots on a golf ball