Jack Rodkey
STAT 642
Assignment 1
2/3/22

Read H.O.1 ; Chpts 1 & 2 in Design & Analysis of Experiments.

1.) A study is planned on the physiology of exercises w/ human subjects. The two treatments in the study are two methods of Aerobic exercise training (**Methods A & B**). At the end of a 10-week exercise period, each subject will undergo a treadmill test for standard respiratory and cardiovascular measurements. 19 subjects are listed in the following table by sex and age. All subjects are in good health and are in the normal weight range for their age, sex & height. The 19 subjects will be divided so that eight subjects will be evaluated in each of the exercise methods; that is, only 16 of the 19 subjects will participate. Each of the 16 subjects will be assigned to one and only one of the two exercise methods.

Individual	1	2	3	4	5	6	7	8	9
Sex	F	F	F	F	F	F	F	F	F
Age	18	18	19	21	38	39	41	44	68

Individual	10	11	12	13	14	15	16	17	18	19
Sex	M	M	M	M	M	M	M	M	M	M
Age	18	31	34	35	38	51	54	58	62	74

a.) How would you group the subjects prior to the assignment so that the experimental error variances would be as small as possible? Explain why you grouped the subjects in the manner that you provided.

~~• I would first get rid of the 68 year old female and the 18 year old male as they seem to be outliers w/ regards to their age within their gender.~~

• I would then group the 19 individuals by sex and age. Thus, the 19 remaining volunteers would be grouped into 4 groups consisting of individuals w/ similar age and the same gender. The reason I would group the subject in this manner is so I have relatively homogeneous EUs.

b.) Display your assignment of the 16 subjects to the two treatments.

	Method A	Method B
Group 1 (Male, Age ≤ 38)	10, 11, 12	13, 14
Group 2 (Male, Age > 38)	15, 16	17, 18, 19
Group 3 (Female, Age ≤ 38)	1, 2	3, 4, 5
Group 4 (Female, Age > 38)	6, 9	7, 8

2. 2) The EPA wants to investigate deposits on a filter in a cooling system. The factors of interest are:

- Factor 1: Flow rate: 5, 10 gps
- Factor 2: Filter Diameter: 0.5, 1, 2 cm
- Factor 3: Fluid Temperature: 75, 100, 125 °F

A run consists of switching out a filter to the correct diameter, changing the flow rate, heating the fluid to the selected temperature and then pressing the start button. After a specified period of time, the amount of deposit on the filter is measured. A maximum of 20 test runs can be made with each batch of cooling fluid mix of impurities that enter the cooling system. It is possible to have several batches of cooling fluid used in the experiment but the study's budget will only allow a maximum of 75 total test runs in the complete experiment. Provide a complete description of how you would design an experiment to meet the above specifications.

X 4 times for each combination of batches

- I would first note that there are $(2)(2)(3) = 18$ different treatments.

To run the experiment I would:

① Run through each combination of treatments for a batch

- I would randomly order the different treatment combinations and run them with random order e.g. we would have something like:

Run #	Flow rate	Filter diameter	Fluid temp
1	5 gps	1 cm	125°F
2	5 gps	0.5 cm	125°F
3	10 gps	0.5 cm	75°F
4	5 gps	2 cm	75°F
⋮			
18	10 gps	2 cm	100°F

For a total of 18 runs per batch.

② I would repeat this process 4 times (4 different batches)

For a total of 72 runs.

3.) An experiment is planned to compare 3 methods of instruction:

Method 1: Instructor lectures 3 times/week

Method 2: All materials provided over the web w/ a weekly Q&A session

Method 3: Students read materials before class and there is instructor-student discussion during the 3 class sessions/week

Each of the 3 methods is evaluated w/ a single classroom of 25 students. The three instructors selected for the study are randomly assigned to a single classroom. The researcher will use the results of four exams given to the 75 students over the 15 week semester to compare the three methods of instruction.

(a) Provide a short (100 words or less) critique of the proposed experiment.

The first problem I see is that the study will be affected by the instructors abilities. For example, if the professor assigned to method 1 is a worse instructor than the professor assigned to method 3, we might falsely conclude that Method 3 is a better teaching method than method 1. Also, we are not told that the students are randomly to the classrooms, thus, there we can't assume that the students w/in each classroom are homogeneous. For example, students w/ greater time constraints might choose method 2, but they might on average do worse than other students b/c of their other time commitments. Thus, we may falsely conclude that Method 2 is a worse method than the other two.

(b) Describe how the experiment can be improved.

(1) Randomly assign the 75 students to the 3 sections.

(2) either of the following two:

(i) Assign 1 professor to teach all three sections

(ii) If time/budget permits, conduct the experiment over 3 semesters and have the 3 professors rotate which section they teach each semester s.t. each professor teaches each section 1 time.

4.) An experiment involving 4 treatments (T_1, T_2, T_3, T_4) and 24 experimental units was conducted.

(a) How many different randomizations are possible if 6 EU's are randomly assigned to each of the 4 treatments.

$$\binom{24}{6} \cdot \binom{18}{6} \cdot \binom{12}{6} \cdot \binom{6}{6}$$

(b) How many different randomizations are possible if the 24 EU's are randomly assigned to the treatments w/ 6 EU's in T_1, 5 EU's in T_2, 7 EU's in T_3 & 6 EU's in T_4

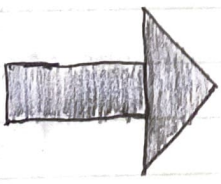
$$\binom{24}{6} \cdot \binom{18}{5} \cdot \binom{13}{7} \cdot \binom{6}{6}$$

more generally: Let $f(n,k)$ be the # of randomizations possible for n EUs & k treatments; $n \in \mathbb{N}$, t_i is the # of EU's assigned to T_i. Let $t_0 = 0$

$$f(n,k) = \binom{n}{t_1} \cdot \prod_{i=2}^{k} \binom{n - \sum_{j=1}^{i-1} t_j}{t_i} = \prod_{i=1}^{k} \binom{n - \sum_{j=0}^{i-1} t_i}{t_i}$$

5.) For the following experiment, identify the following components of the experimental design. (some may be absent in the experiment)

An experiment was run to assess the compressive strength of the cover of golf balls. The experimental plan includes random sampling of balls from four brands (B_1, B_2, B_3, B_4) of golf balls. Furthermore, each brand of golf ball has three thickness of covers (T_1, T_2, T_3) and two types of cover materials (M_1, M_2). The compressive strength of the golf ball was recorded at 5 randomly selected spots on each golf ball in the study. A total of 6 golf balls were evaluated for each combination of brand, thickness & material type. There are two major testing facilities so half of the balls are tested at one facility and the remaining balls at the second facility. The testing is done inside at a controlled temperature.



5.) (Contd) * [see H.O.1, pg 10-11] *

(1) Factors: F_1: Brand (could this also be a block?) F_2: Thickness of cover F_3: Cover material | Are the factors fixed? (see H.O.1 pg11) → NO b/c not measuring variability across the ball. i.e. in wafer example.

(2) Experimental Units: Golf balls

(3) Covariates: No covariates b/c testing done inside. If done outside things like the weather would be covariates.

(4) Factor levels: F_1: (4) brands, F_2: (3) thickness F_3: (2) cover material

(5) Measurement Units: A randomly selected spot on a golf ball.
[? Is it the case that any time the measurement units differ from the EUs that we have subsampling? ?]

(6) Blocking: Testing facilities
[? Not sure on this one, is it b/c we are uncertain if the testing conditions are the same for each facility? ?]

(7) Treatments: The treatments are the (4)(3)(2) = 24 combinations of the factors

(8) Replications: There are 3 replications for each treatment at each facility.

(9) Confounding: No confounding

(10) Response: Compressive strength of the cover on golf balls.

(11) Subsampling: There are 5 subsamples on each EU (the 5 different randomly selected spots on each golf ball)