Midterm 1

October 3rd, 2023

Teaching sampling distributions in introductory statistics courses can be very challenging as it is a notoriously difficult topic. Current literature on this issue suggests two possible methods for teaching sampling distributions: computer simulations only (think: M&M applet) vs. hands-on simulations (think: counting M&M's) immediately followed by a computer simulation. We would like to perform a study investigating which method is best: (1) computer simulations alone or (2) hands-on simulations, then computer simulation. Note the reason for testing this is that performing both hands-on and computer simulations takes up a lot of class time.

1. In a paragraph, describe the type of study you would perform to compare teaching methods. Imagine you had all the resources and time to conduct your study. Make sure to discuss the following:

- Type of study: experiment or observational study and why you chose it
- Sampling method: how would you recruit students to be a part of your study (Note: you may used one of the named sampling methods we have learned in class or create your own sampling method)

Type of study: Experiment Study

Describe study: We can select a sample of classes, with the same contents of a lecture, we start teaching the classes by these two teaching methods: computer simulations only and hands-on simulations, then computer simulation. After that, we compare the results of variables and compare.

Sampling method: cluster sampling We can divide the classes into 2 cluster. One cluster is taught by computer simulations alone and the other is taught by hands-on simulations, then computer simulation.

Variables: Exam grades before applying the teaching method, exam grades after applying the teaching method , understandable level of the student, time spent to prepare for the teaching resources of each teaching methods, number of teaching resources needed , whether students easily use of apply the teaching resources in their works or exercises.

Type of variables: Exam grades before applying the teaching method, exam grades after applying the teaching method (quantitative), understandable level of the student(categorical), time spent to prepare for the teaching resources of each teaching methods (quantitative), number of teaching resources needed (quantitative), whether students easily use of apply the teaching resources in their works or exercises(categorical).

2. What would be your response variable, and how would you measure it?

Exam grades before and after teaching methods applied are response variable. Grades are seen to be the most effective measurement to measure outcome of student. They might be decided by teaching method element.

How to measure it: we can give the same exams for the students of each cluster to do for the same period of time and compare the results. One exam we do before applying the teaching method and the other we do after the teaching method ends.

3. List 2 other variables you would want to collect and (1) describe how they would be measured or obtained, and (2) explain why you chose them.

I would like to measure number of teaching resources needed (quantitative) and understandable level of the student(categorical).

Understandable level of the students for each teaching method: We can put this question on the survey to ask the students of each cluster. We can divided the understandable level in some rank: totally understand, somewhat understand, hard to understand, hardly understand. For this variable, we can listen to students' feedback to know how effective these teaching method applied for the students.

Number of teaching resources: we can collect the number of teaching resources needed for each teaching method. By this way, we can measure the support for each teaching method. It may take time for the lectures to prepare and research and to support the lecture, the lectures need a few or a lot of teaching resources.

4. Are there any potential biases in your study? Explain them below.

There are some potential biases. For example, time for understanding. Each student needs a certain period of time to apply the teaching meathod in their work.

Now that you have thought about your own ideal study, consider a pseudo-experimental study that was performed at UC Irvine in the Fall quarter of 2015 to investigate this exact research question.

At UC Irvine, the introductory statistics course was taught by **4 instructors across 8 sections consisting of approximately 55 students each**. **Half of the sections** were randomly assigned to learn sampling distributions using **only computer simulation methods (control group)**, while **the other half used a combination of hands-on and computer simulations (treatment group)**. In this study, we compared the students' average exam scores (2 midterms and 1 final) across the treatment and control groups. It should be noted that students did not learn sampling distributions until *after* the first midterm.

5. In 1-2 complete sentences, describe one ethical concern you might have regarding how the study was designed.

Maybe, the teaching method does not requires to learn sampling distribution so they did not learn it until after the first term.

6. What is a potential confounder in the the relationship between treatment/control group and exam score? Explain how the variable is both related to the explanatory variable (group) and the response variable (exam score).

It is possible that every student has different statistical knowledge base. So, knowledge baseline could be the confounder.

7. What type of plot would you use to visualize the relationship between the response and the explanatory variable? Explain how you would use this plot to compare the effectiveness of the treatment.

boxplot. The boxplot can show the difference between mean or sd of each teaching method.

8. The table below summarizes their scores:

	Control	Treatment
	mean (sd) or n (%)	mean (sd) or n (%)
Midterm 1	80 (11.9)	78 (14.2)
Midterm 2	77.9 (13.9)	76.2 (15.3)
Final Exam	75.3 (14.9)	77.2 (14.9)

Using the table, comment on the differences in exam scores across the treatment and control groups.

For Midterm 1 grades, we can see that the mean of Control group's score is higher than Treatment's one. This reflects that the average scores of students in Control group is better than the ones in Treatment group. Besides, the standard deviation of Control (11.9) is lower than that of Treatment(14.2), which means that Treatment group has greater variability in scores than Control. The gap is larger between the low and higher scores in Treatment Group.

For Midterm 2, the situation is simliar to the results of Midterm 1 between 2 groups. The average score of Control(77.9) is higher than Treatment score(76.2), but the standard deviation of Control (13.9) is lower than that of Treatment (14.2). However, the average scores of Midterm 2 are lower than Midterm 1 for both of groups.

For Final Exam, the average score of Control is lower than that of Treatment and the standard deviation of these groups is equal. This results reflects that the final exam result of Treatment is better than Control.

9. Knowing that students had not yet received the treatment, explain how the differences in average exam score for Midterm 1 should impact how we compare midterm 2 and the final exam scores.

I think the base knowledge of students might affect the results of Midterm 1. For the midterm 1, there might be a short period of teaching applied, so the previous knowledge of students can affect the Midterm 1 exam. We can see the slight gap between average score of each group, in particular the average score of Control is slightly higher than that of Treatment Group. For the results of Midterm 2, when the teaching method has applied for a longer period of time, this has an effect on the score of MIdterm 2. However, the final exam scores, the results are reverse compared the MIdterm 1 and 2. It reflects ability of students apply the teaching method for the whole knowledge of that class to do the final exam.

10. The exams covered sampling distributions, but also other topics in the course. Comment on why exams may still be a useful response variable.

Exam score is seen as a standardized assessment. Because exams are often designed to be standardized and it is a fair assessment tool to measure ability of student.

The distribution of scores on each exam across both groups was approximately normal. Let X represent the score of a randomly selected student in the control group and let Y represent the score of a randomly selected student in the treatment group. Using the `pnorm()` and `qnorm()` functions, calculate the following:

11. Calculate the probability that on the first exam a randomly selected student scored 75 or below:

- In the treatment group:

prob_treatment <- pnorm(75, 78, 14.2)

```
prob_treatment <- pnorm(75, 78, 14.2)
prob_treatment
```

```
## [1] 0.4163392
```

- In the control group:

prob_control <- pnorm(75, 80, 11.9)

```
prob_control = pnorm(75, 80, 11.9)
```

12. Interpret both probabilities that you calculated above in the context of the problem.

We can see the prob of Treatment is higher than the Control. It means that the higher probability of the students who get 75 scores in the Treatment compared to The COntrol GRoup.

13. For each group separately, students who scored in the top 10% on the final exam score above what value?

- In the treatment group:

```
qnorm(0.9, mean = 78, sd = 14.2)
```

```
## [1] 96.19803
```

- In the control group:

qnorm(0.9, mean = 80, sd = 11.9)

```
qnorm(0.9, mean = 78, sd = 14.2)
```

```
## [1] 96.19803
```

14. Speaking of sampling distributions... Using the M&M activity from class, in 2-3 sentences explain what a sampling distribution is in your own words.

A sampling distribution can show the students outcomes and assess the variability between students.

Read the description and use the image uploaded in the exam instructions on Canvas. Then, answer question 15 below.

15. Looking at the boxplots, comment on the efficacy (effectiveness) of Family Therapy in helping patients with anorexia gain weight.

How would you rate yourself on this exam? Choose one.

- E (excellent)
- M (meets expectations)
- R (revision requested)
- N (not assessible)

Explain your rating above in a sentence or two.