1. **(1 point)** carroll_problib/statistics/InterStats/Preliminaries/bp.pg

A scientist wants to find out if eating more vegetables lowers blood pressure. She randomly samples 114 adults, randomly selects 44 of them and assigns them to eat at least 3 servings of vegetables per day, while she assigns the other 70 to eat fewer than 3 servings of vegetables per day. After three months of this diet, she finds that the 44 people who eat at least 3 servings of vegetables per day have blood pressures which are 27 mmHg lower than the 70 people who eat fewer than 3 servings of vegetables per day. She performs a statistical analysis on this difference and finds a p-value of 0.003 .

What is this p-value the probability of?

- A. This is the probability that we would get a difference in blood pressures between the two groups at least as large as what we got, just by random chance, if eating vegetables had an effect at all blood pressure.
- B. This is the probability that eating at least 3 servings of vegetables per day has an effect on blood pressure.
- C. This is the probability that we would get a difference in blood pressures between the two groups at least as large as what we got, just by random chance, if eating vegetables had no effect at all blood pressure.
- D. This is the probability that eating at least 3 servings of vegetables per day has no effect on blood pressure.

Do we have a significant difference at the 0.05 level?

- A. Yes
- B. No
- C. Maybe

Is this a controlled experiment?

- A. Yes, because we have randomly selected the sample.
- B. Yes, because she randomly decided what diet each person would have.
- C. Yes, because different people have different diets.
- D. No, because maybe people who have low blood pressure just randomly happen to eat more vegetables.
- E. No, because the sample was not randomly separated into different groups.
- F. No, because blood pressure is also affected by how much exercise people get, and she did not account for this factor.

Can we conclude that eating more vegetables lowers blood pressure?

- A. Yes, because if vegetables had no effect on blood pressure, then there wouldn't have been any difference in blood pressures.
- B. Yes, because she randomly assigned people to different diets, and they have significantly different blood pressures.
- C. Yes, because we randomly selected the people to be in this study.
- D. Yes, because eating more vegetables makes you healthier.
- E. No, because she did not control what diet each person would have.
- F. No, because what you eat does not necessarily determine your blood pressure.

What conclusions can we draw from this study?

- A. Eating at least 3 servings of vegetables per day probably causes lower blood pressure.
- B. People who eat at least 3 servings of vegetables per day tend to have lower blood pressure.
- C. People who change their diet to increase the amount of vegetables they eat will probably not see a drop in their blood pressure.
- D. We can draw no conclusions at all.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

2. **(1 point)** carroll_problib/statistics/InterStats/Preliminaries/ad.pg

A business wants to persuade consumers that their product "The uPhone" is cool. They put out advertisements for the uPhone on television, radio, the Internet, and in magazines. They even pay to have Agent 007 use a uPhone in the newest James Bond movie. Then, the business wants to know if these advertisements work, so they randomly survey 193 people across the country. First, they ask the people whether they have seen the advertisements for the uPhone, finding that 73 people have seen the ads, while 120 people have not. Then, they give all of the people a uPhone to try out for 15 minutes, before asking them to rate the coolness of the uPhone on a scale of 1 to 10, where 1 means not cool at all, and 10 means ultimate metaphysical coolness. They find that people who have not seen the ads give the uPhone an average coolness rating of 3.3, while people who have seen the ads give the uPhone an average coolness rating of 6.3. When they perform a statistical analysis of the difference between these results, they calculate a p-value of 0.005.

What is this p-value the probability of?

- A. This is the probability that we would get a difference in coolness ratings between the two groups at least as large as what we got, just by random chance, if the ads had an effect on the coolness ratings.
- B. This is the probability that the advertising campaign had no effect on the coolness ratings.
- C. This is the probability that we would get a difference in coolness ratings between the two groups at least as large as what we got, just by random chance, if the ads had no effect at all on the coolness ratings.
- D. This is the probability that the advertising campaign had an effect on the coolness ratings.

Do we have a significant difference at the 0.01 level?

- A. Yes
- B. No
- C. Maybe

Is this a controlled experiment?

- A. Yes, because we have randomly selected the sample.
- B. Yes, because we randomly chose who would see the ads.
- C. No, because people who haven't seen the ads would have no basis for rating the coolness of the uPhone.
- D. No, because it wasn't random who saw the ads and who didn't.

Can we conclude that these ads make people think the uPhone is more cool?

- A. Yes, because we randomly selected the people to be in this study.
- B. No, because this is not a controlled experiment.
- C. No, because maybe people think the uPhone is cool for reasons other than the ads.
- D. Yes, because advertisements affect the way people think.

What conclusions can we draw from this study?

- A. The advertisements are effective at making people think the uPhone is cool.
- B. People who see the advertisements give a higher coolness rating to the uPhone, on average.
- C. People who see the advertisements do not give a higher coolness rating to the uPhone, on average.
- D. The advertisements were not effective at making people think the uPhone is cool.
- E. We can draw no conclusions at all.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

3. (1 point) carroll_problib/statistics/InterStats/Preliminaries/headphones.pg

If something is marked with a higher price, do people believe it is of higher quality? We purchase a bunch of headphones of the same brand in two different colors, black and white. We gather a sample of 52 college students, and randomly divide the students into two groups of 26. All students listen to the "Friends" theme through two different pairs of headphones, one black and one white, and rate the sound quality of each set of headphones on a scale of 1 to 10. The first group of 26 students are told that both colors of headphones cost \$10. The second group of 26 students are told that the black headphones cost \$15 while the white headphones cost \$5. We find that the first group (told that both sets cost \$10) rated the black headphones an average of 0.16 points lower than the white headphones, while the second group (told that the black cost \$15 and white cost \$5) rated the black headphones an average of 1.0 points higher than the white headphones. So the difference between the ratings from the two groups was 1.16. When we perform a statistical analysis, we find a p-value of 0.003.

What is this p-value the probability of?

- A. This is the probability that we would get a difference in sound quality ratings between the two groups at least as large as 1.16, just by random chance, if the difference in prices had no effect on student quality ratings.
- B. This is the probability that price doesn't have an effect on sound quality ratings.
- C. This is the probability that we would get a difference in sound quality ratings between the two groups at least as large as 1.16, just by random chance, if the difference in prices had an effect on student quality ratings.
- D. This is the probability that price has an effect on sound quality ratings.
- E. This is the probability that the difference in sound quality ratings between the two groups happened by random chance.

Do we have a significant difference at the 0.025 level?

- A. Yes
- B. No
- C. Maybe

Is this a controlled experiment?

- A. Yes, because we randomly divided our sample into groups and we decided who was told there was a price difference and who was told the prices were the same.
- B. No, because there may have been a real difference between black headphones and white headphones.
- C. No, because some people may think that black headphones look better than white ones, an this may have affected their ratings of the sound quality.
- D. Yes, because exactly half of the people were told there was a price difference and half were told the prices were the same.

Does this mean that higher prices cause people give higher sound quality ratings to a set of headphones?

- A. Yes, because when people see a higher price marked, they look for higher quality.
- B. Yes, because we randomly separated people into two groups, and there was a significant difference between the results of the two groups.
- C. Yes, because if the marked price didn't make a difference, then everyone would have given the same rating to the headphones.
- D. No, because we can't say for sure that it was the difference in prices which caused the difference in sound quality ratings.
- E. No, because even if price had no effect, if we did this experiment many times, differences this size would happen often due to random variation.

What conclusions can we draw from this study?

- A. Marking a higher price on a set of headphones will make people give them a higher sound quality rating, on average.
- B. Headphones with a higher price do not receive higher sound quality ratings, on average.
- C. Headphones with a higher price tend to receive higher sound quality ratings, on average, but we cannot conclude that it is the higher price which causes this.
- D. We can draw no conclusions at all.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

4. **(1 point)** `carroll_problib/statistics/InterStats/Preliminaries/AttractiveTeachers.pg`

In **a Wall Street Journal article (linked here)** reporters outlined a study to determine if students learned better from attractive teachers. According to the study they claim to find a statistically significant difference in test scores when the lecturer was rated as being more attractive. They went on to say that "the researchers don't think that sexual interest explains the results" since the results appeared to hold up dispite the teacher's and student's genders.

Quoting from the article:
"Here's how the experiment worked. The researchers asked 131 college students to listen to a recording of a 20-minute introductory physics lecture. The students were randomly assigned to a male or female lecturer, each of whom read an identical text. While the lecture was playing, a computer displayed what the volunteers were told was a photo of the lecturer - who was highly attractive in some of the cases and not as fetching in others. (Earlier volunteers had rated some photos of possible 'lecturers' for attractiveness, enabling the researchers to pick the best and worst looking.) Taking notes was barred."

"After the lecture, participants got a 25-item quiz on the material. For those with the attractive instructor, the average score was 18.27; for those with an unattractive one, the average was 16.68. That gap isn't huge, but it is statistically significant, the researchers said."

(a) What was the null hypothesis for the researcher's study?

- A. The score from the group who had prior physics background will be the same as the score from the group who had no prior physics background.
- B. The score from group who had prior physics background will be higher than the score from the group who had no prior physics background.
- C. The score from the groups with the attractive teacher will be less than the score from the group with the unattractive teacher.
- D. There will be a difference between the scores of the groups with the attractive and the unattractive teachers.
- E. The score from the group with the attractive teacher will be the same as the group with the unattractive teacher.
- F. The score from the groups with the attractive teacher will be greater than the score from the group with the unattractive teacher.
- G. The appropriate null hypothesis is not listed here.

(b) What was the alternative hypothesis for the researcher's study?

- A. There will be a difference between the scores of the groups with the attractive and the unattractive teachers.
- B. The score from the groups with the attractive teacher will be the same as the score from the group with the unattractive teacher.
- C. The score from group who had prior physics background will be higher than the score from the group who had no prior physics background.
- D. The score from the group who had prior physics background will be the same as the score from the group who had no prior physics background.
- E. The score from the groups with the attractive teacher will be less than the score from the group with the unattractive teacher.
- F. The score from the group with the attractive teacher will be greater than the group with the unattractive teacher.
- G. The appropriate alternative hypothesis is not listed here.

(c) In order to claim statistical significance the researchers would have had to find a p-value for the hypothesis test you indicated in parts (a) and (b). What is this p-value the probability of?

- A. The probability of finding a score of 18.27 in attractive teacher group and 16.68 in the unattractive teacher group.
- B. The probability of finding scores between 16.68 and 18.27 on a 25 question test.
- C. The probability of finding a difference in scores as great as this or larger assuming that the two groups should have received the same score.
- D. The probability of finding a difference in scores as great as this or smaller assuming that the two groups should have received the same score.
- E. The probability of finding a difference in scores as great as this or smaller assuming that the two groups should have received different scores.
- F. The probability of finding a difference in scores as great as this or larger assuming that the two groups should have received different scores.

(d) Given a decision level of 0.025 and the test scores listed in the quote above, which of the following is the most plausible p-value for the researcher's claim?

- A. $p = 0.025$
- B. $p = 0.058$
- C. $p = 0.022$
- D. $p = 0.0001$
- E. $p = 0$
- F. $p = 1$
- G. None of the p-values listed are plausible for this study.

Answer(s) submitted:

-
-
-
-

(incorrect)

---

5. **(1 point)** `carroll_problib/statistics/InterStats/Preliminaries/DelTacoExperiment.pg`

An experiment investigated the effect of length and repetition of TV ads on students choosing to eat at Del Taco. All 60 students watched a 40-minute television program that included ads for Del Taco. Some students saw a 30-second commercial; others a 90-second commercial. The same commercial was shown either 1, 3, or 5 times during the program. After the viewing, each student was asked to rate their craving for Del Taco on a scale of 0 to 10.

(a) How many treatments are there? _____

(b) What is the response variable?

- A. 60 students
- B. craving for Del Taco on a scale of 0 to 10
- C. 1, 3, or 5 commercials during the 40-minute television program
- D. 30-second and 90-second commercials
- E. 40-minute television program

(c) What are the levels of the length factor?

- A. 60 students
- B. 40-minute television program
- C. 30-second and 90-second commercials
- D. craving for Del Taco on a scale of 0 to 10
- E. 1, 3, or 5 commercials during the 40-minute television program

(d) What are the subjects of this experiment?

- A. 60 students
- B. 40-minute television program
- C. craving for Del Taco on a scale of 0 to 10
- D. 1, 3, or 5 commercials during the 40-minute television program
- E. effect of length and repetion of TV ads

Answer(s) submitted:

-
-
-
-

(incorrect)

---

6. **(1 point)** `carroll_problib/statistics/InterStats/Preliminaries/strawberries.pg`

We want to know whether a new brand of fertilizer will improve the productivity of our strawberry fields, so we put the fertilizer to the test. In a particular field, we apply this fertilizer to 29 randomly selected rows, leaving the remaining 29 rows without fertilizer. At the end of the season, we add up the productivity of each row, finding that the rows with fertilizer produced an average of 10.2 pounds of strawberries, while the rows without fertilizer produced an average of 8.5 pounds of strawberries, so the fertilized rows produced an average of 1.7 more pounds. We perform a statistical analysis of the difference between these results, and calculate that the p-value is 0.004.

What is this p-value the probability of?

- A. This is the probability that the fertilizer does not improve the productivity of strawberry plants.
- B. This is the probability that we would get a difference in productivity of exactly 1.7 pounds, just by random chance, if the fertilizer had an effect on the plants.
- C. This is the probability that we would get a difference in productivity of 1.7 or fewer pounds, just by random change, if the fertilizer had an effect on the plants.
- D. This is the probability that we would get a difference in productivity of 1.7 or more pounds, just by random chance, if the fertilizer had an effect on the plants.
- E. This is the probability that we would get a difference in productivity of 1.7 or more pounds, just by random chance, if the fertilizer had no effect on the plants.
- F. This is the probability that we would get a difference in productivity of exactly 1.7 pounds, just by random chance, if the fertilizer had no effect on the plants.
- G. This is the probability that the fertilizer improves the productivity of strawberry plants.
- H. This is the probability that we would get a difference in productivity of 1.7 or fewer pounds, just by random chance, if the fertilizer had no effect on the plants.

Is this result statistically significant at the 0.05 level?

- A. Yes
- B. No
- C. Maybe

Is this a controlled experiment?

- A. Yes, because we randomly chose which rows would receive the fertilizer.
- B. Yes, because half of the rows got fertilizer and half did not.
- C. No, because it wasn't random which rows received the fertilizer.
- D. No, because there are many other factors which affect strawberry production, like irrigation.

Does this mean that fertilizer increases strawberry production?

- A. No, because the amount of water that the plants receive has a big impact on the number of pounds produced, and this factor was not accounted for.
- B. Yes, because we randomly chose which rows would receive fertilizer, and those had significantly higher productivity.
- C. No, because birds may eat a significant number of the berries, and birds eat randomly.
- D. Yes, because strawberry plants don't know whether they're receiving fertilizer or not.
- E. Yes, because a difference of 1.7 pounds is a lot, and would make a huge difference if we scaled this up to many acres.

What are the strongest conclusions can we draw from this study?

- A. This study tells us nothing about whether or not the fertilizer works.
- B. This fertilizer does not have an effect on strawberry productivity.
- C. This fertilizer causes the plants to produce more pounds of strawberries.
- D. The plants which received the fertilizer produced significantly more pounds of strawberries, but we're not sure if this was really the cause.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

7. (1 point) `carroll_problib/statistics/InterStats/Preliminaries/bball.pg`

We want to know whether it helps people make a free throw in basketball if they first take a second and visualize making the basket. We get 90 volunteers to participate in our study. We randomly choose half of our volunteers to take 10 free throws just doing their best, while the other half first visualize making the basket before they take each of 10 free throws. We keep track of the number of shots made, and find our volunteers who practiced visualization shot 10 percent more baskets. We perform a statistical analysis on this difference and calculate that the p-value is 0.012.

What is this p-value the probability of?

- A. This is the probability that our volunteers who did the visualization would shoot at least 10 percent better than the other group, just randomly, even if visualization doesn't make any difference.
- B. This is the probability that our volunteers who did the visualization would shoot at least 10 percent better than the other group, just randomly, if visualization improved shooting accuracy.
- C. The is the probability that visualization doesn't make any difference.
- D. This is the probability that visualization improves shooting accuracy.

Is this difference statistically significant at the 0.1 level?

- A. Yes
- B. No
- C. Maybe

Is this a controlled experiment?

- A. Yes, because we randomly chose which people would use visualization and which would not.
- B. No, because we didn't randomly separate our sample into different groups.
- C. Yes, because it's random whether a particular person is going to make a particular shot or not.

- D. No, because some people are better at making free throws than others.

Does this mean that visualization improves free throw accuracy?

- A. Yes, because we randomly chose which volunteers would use visualization, and those people made significantly more free throws.
- B. No, because some people have practiced making free throws a lot more than others, and so this isn't really random at all.
- C. Yes, because visualizing a successful free throw helps to guide your hands and muscles to do things correctly.
- D. No, because whether or not you make a free throw has to do with your height, which is a random factor.
- E. Yes, because when you see it happen in your mind, then you are expecting to succeed, and this positive attitude can have a big impact.

What are the strongest conclusions can we draw from this study?

- A. If you visualize a successful shot before taking a free throw, you have a better chance of success.
- B. People make significantly more free throws when they practiced visualization, but we can't tell if visualization is what caused their success.
- C. This study doesn't tell us anything about whether or not visualization helps people make free throws.
- D. If you visualize a successful shot before taking a free throw, it has no impact on your chances of success.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

8. **(1 point)** `carroll_problib/statistics/InterStats/Preliminaries/river2.pg`

A farmer is trying to determine which of two fertilization application schedules, A or B, is best. There are forests to the north and south of his land, a river to the west, and mountains to the east. The farmer divides his land into 8 plots with plot 1 in the north-west corner, plot 2 immediately south of plot 1, plot 4 in the south-west corner, plot 5 in the north-east corner, and plot 8 in the south-east corner. (Start by drawing a picture).

Assume that $\alpha = 0.025$ for each of the following scenarios.

(A) Presume that the farmer used fertilizer schedule A in plots 1-4 (on the west of his land) and schedule B on plots 5-8 (on the east of his land). After the growing season he observes that the plots that received fertilizer schedule A yielded 68 bushels and those receiving schedule B yielded 37 bushels. He performs a statistical analysis on the difference between the yields and finds a p-value of 0.004. Which of the following is the most appropriate conclusion?

- A. No conclusion can be made since the farmer didn't account for the influence of the river on the growth of his crops.
- B. Based on the p-value we can conclude that there is a statistical difference between the two fertilizer schedules
- C. No conclusion can be made since the farmer failed to account for the influence of the forests on the growth of his crops.
- D. No conclusion can be made since the farmer failed to include randomization in his experimental design.
- E. Based on the p-value we cannot conclude that there is a statistical difference between the two fertilizer schedules.

(B) Presume that the farmer used fertilizer schedule A in plots 1, 3, 6, and 7 and schedule B on the remaining plots. After the growing season he observes that the plots that received fertilizer schedule A yielded 51 bushels and those receiving schedule B yielded 36 bushels. He performs a statistical analysis on the difference between the yields and finds a p-value of 0.018. Which of the following is the most appropriate conclusion?

- A. No conclusion can be made since the farmer didn't account for the influence of the river on the growth of his crops.
- B. No conclusion can be made since the farmer failed to include randomization in his experimental design.
- C. Based on the p-value we cannot conclude that there is a statistical difference between the two fertilizer schedules.
- D. Based on the p-value we can conclude that there is a statistical difference between the two fertilizer schedules
- E. No conclusion can be made since the farmer failed to account for the influence of the forests on the growth of his crops.

(C) Now presume that the farmer used fertilizer schedule A on four randomly selected plots and schedule B on the remaining four plots. After the growing season he observes that the plots that received fertilizer schedule A yielded 51 bushels and those receiving schedule B yielded 36 bushels. He performs a statistical analysis on the difference between the yields and finds a p-value of 0.018. Which of the following is the most appropriate conclusion?

- A. No conclusion can be made since the farmer should have made sure that the plots were evenly scattered
- B. Based on the p-value we can conclude that there is a statistical difference between the two fertilizer schedules
- C. No conclusion can be made since the farmer should have controlled for the influence of the river.

- D. No conclusion can be made since the farmer failed to control for the geography.
- E. Based on the p-value we cannot conclude that there is a statistical ules.

(D) Next presum on the following plots: a randomly selected plot from plots 1 and 4, a randomly selected plot from plots 2 and 3, a randomly selected plot from p from plots 6 and 7. ing four plots. Afte plots that received fertilizer schedule A yielded 53 bushels and those receiving schedule B yielded 45 bushels. He performs a statistical analysis on the difference between the yields and finds a p-value of 0.019. Which of the following is the most appropriate conclusion?

- A. Based on the p-value we cannot conclude that there is a statistical difference between the fertilizer schedules.
- B. No conclusion can be made since he should have controlled for the influence of the river.
- C. No conclusion can be made since he should have controlled for the influence of the mountains
- D. No conclusion can be made since the farmer should have controlled for the influence of the forests.
- E. No conclusion can be made since the farmer should have made sure that the plots were evenly scattered
- F. Based on the p-value we can conclude that there is a statistical difference between the two fertilizer schedules.

Answer(s) submitted:

-
-
-
-

(incorrect)

---

9. **(1 point)** `carroll_problib/statistics/InterStats/Preliminaries/SimulationsAndPValues.pg`

A gambler is trying to decide if a coin is weighted so he flips it 40 times and finds that only 24% turn up heads. From this data it looks like the coin is weighted to turn up tails more often, but his sample could have been random chance. Being well-initiated in the practice of statistics he decides to run a hypothesis test.

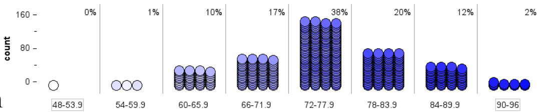
(a) Complete the mathematical statement of the null hypothesis where we are taking p to be the proportion of the 40 flips that turn up heads:

H_0 : p [?/=/¡/¿] ____

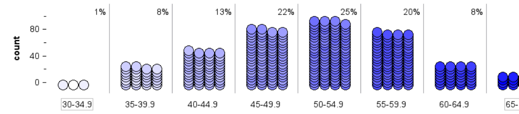
(b) Complete the mathematical statement of the alternative hypothesis:

H_A : p [?/=/¡/¿] ____

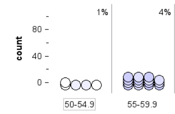
(c) Which of the histograms A - E below most closely shows 400 simulations of 40 coin flips under our null hypothesis? [?/A/B/C/D/E]

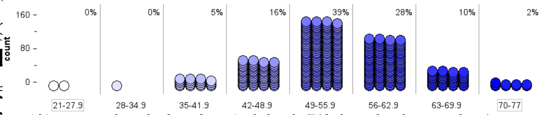


A

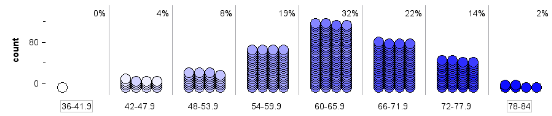


B





D



E

(Click on a graph to enlarge it.)

(d) Based on the simulation chosen in part (c), what is the approximate p-value for finding 24% heads under your null hypothesis?

- ?
- 0 percent (no chance)
- less than 1 percent
- about 1 percent
- less than 4 percent
- less than 5 percent
- about 5 percent
- about 25 percent
- about 40 percent
- about 50 percent
- about 100 percent

(e) Do you have statistical evidence to state that the coin is likely weighted?

- A. Yes, we have evidence to reject the alternative hypothesis, and hence we conclude that the coin is likely weighted.
- B. No, we do not have evidence to reject the alternative hypothesis, and hence we cannot conclude that the coin is weighted.
- C. No, we do not have evidence to reject the null hypothesis, and hence we cannot conclude that the coin is weighted.
- D. Yes, we have evidence to reject the null hypothesis, and hence we conclude that the coin is likely weighted.

Answer(s) submitted:

-
-
-
-
-

-
-

(incorrect)

---

10. **(1 point)** `carroll_problib/statistics/InterStats/Preliminaries/SimulationsAndPValues2.pg`

According to an advertisement, "three out of four dentists in Montana recommend using AquaBrush tooth brushes." Being the skeptic that your grandmother is, she obtains a list of all of the dentists in Montana and randomly samples 40 of them. In her sample she finds that only 22 of them actually recommended AquaBrush tooth brushes. Your grandmother wants to know if the advertisement is correct based on her evidence.

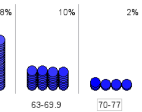
(a) Complete the mathematical statement of the null hypothesis where we are taking p to be the proportion of the 40 doctors that recommend using AquaBrush tooth brushes:

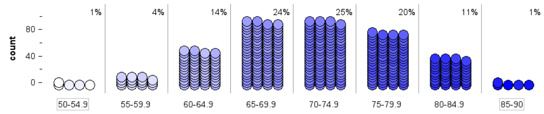
H_0 : p [?/=/¡/¿] _____

(b) Complete the mathematical statement of the alternative hypothesis:

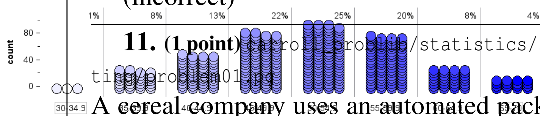
H_A : p [?/=/¡/¿] _____

(c) Which of the histograms A - E below most closely shows 400 simulations of selecting 40 dentist's opinions under our null hypothesis? [?/A/B/C/D/E]

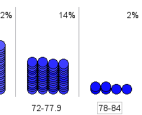


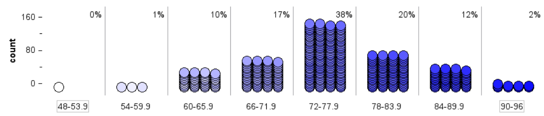


B



C





E

(Click on a graph to enlarge it.)

(d) Based on the simulation chosen in part (c), what is the approximate p-value for finding 22 out of 40 doctors who recommend AquaBrush tooth brushes under your null hypothesis?

- ?
- 0 percent (no chance)
- about 1 percent
- less than 4 percent
- less than 5 percent
- about 5 percent
- about 25 percent
- about 40 percent
- about 50 percent
- about 100 percent

(e) Does your grandmother have statistical evidence to state that the advertisement is incorrect?

- A. Yes, we have evidence to reject the null hypothesis, and hence we conclude that the advertisement is likely incorrect.
- B. Yes, we have evidence to reject the alternative hypothesis, and hence we conclude that the advertisement is likely incorrect.
- C. No, we do not have evidence to reject the null hypothesis, and hence we cannot conclude that the advertisement is incorrect.
- D. No, we do not have evidence to reject the alternative hypothesis, and hence we cannot conclude that the advertisement is incorrect.

Answer(s) submitted:

-
-
-
-
-
-
-

(incorrect)

---

11. **(1 point)** `carroll_problib/statistics/simple_hypothesis_testing/problem01.pg`

A cereal company uses an automated packaging system so that their cereal boxes consistently contain the correct amount of cereal. After the machine was verified to be running correctly, they allowed their automated system to fill 10,000 boxes of cereal and then weighed them to verify the content.

The table below sums up the results of their measurements.

Observed Weights	¡17	17.0-17.2	17.2-17.4	17.4-17.6	17.6-17.
Frequency	133	462	637	1002	1548

This morning, an inspector tests a box of cereal off of the packaging line and finds that it contains 17.2 ounces of cereal.

Let's test the hypothesis that the actual weight of the cereal boxes is 18 versus the alternative that the actual weight is *lower* than that (i.e. that the machine is not functioning properly).

Null Hypothesis: The mean weight of cereal boxes is 18. Any sample that suggests otherwise is simply a chance event.

Alternate Hypothesis: The mean weight of cereal boxes is less than 18.

According to the company's data, what is the probability of getting a sample that weighs 17.2 *or less* when the machine is functioning correctly.

_____

Let's agree that if the observed outcome has a probability of *less than* 5% under the tested hypothesis, we will reject the hypothesis.

What should we conclude regarding the hypothesis?

- A. We should reject the null hypothesis. It is unlikely that the inspector would have found a box this light simply by chance.
- B. We cannot reject the null hypothesis. It is not unlikely that the inspector would have found a box this light by chance.

Assuming that the data collection was not flawed, which of the following statements is most appropriate regarding our conclusion?

- A. The machine is definitely working properly.
- B. We do not have sufficient evidence to suggest that the machine is broken.
- C. We have sufficient evidence to suggest that the machine is broken.
- D. The machine is definitely not working properly.
- E. We have sufficient evidence to suggest that the machine is working properly.
- F. None of the statements is an appropriate conclusion.

Answer(s) submitted:

-
-
-

(incorrect)

---

12. (1 point) `carroll_problib/statistics/simple_hypothesis_testing/problem02.pg`

If we flip a fair coin 10 times and record the number of times "Heads" comes up, there are eleven possible outcomes. (We could record 0 heads, 1 head, 2 heads, etc.).

The table below gives the probability of each possible outcome. (We will study how to determine these probabilities later in this course.)

Number of Heads	0	1	2	3	4	5	6	7	8	9	10
Probability	0.00098	0.00977	0.04395	0.11719	0.20508	0.24609	0.20508	0.11719	0.04395	0.00977	0.00098

If the coin is truly fair (and the flipper is not lying), then the most likely outcome is to get 5 heads. Suppose that a friend of yours claims to have flipped a coin 10 times and gotten 9 heads.

Let's test the hypothesis that the coin is fair (i.e. that the expected number of heads is 5) versus the alternative that the coin is biased towards heads (i.e. that the expected number of heads is *higher* than that).

According to the table below, what is the probability of getting 9 heads *or more* with a fair coin?

_____

Let's agree that if the observed outcome has a probability of *less than* 5

What should we conclude regarding the hypothesis?

- A. We should reject the hypothesis.
- B. We cannot reject the hypothesis.

Assuming that our friend is telling the truth about the experiment, which of the following statements is most appropriate regarding our conclusion?

- A. Our friend is definitely using a fair coin.
- B. We do not have sufficient evidence to suggest that our friend is using a biased coin.
- C. Our friend is definitely using a biased coin.
- D. We have sufficient evidence to suggest that our friend is using a biased coin.
- E. We have sufficient evidence to suggest that our friend is using a fair coin.
- F. None of the statements is an appropriate conclusion.

Answer(s) submitted:

-
-
-

(incorrect)

---

13. (1 point) `carroll_problib/statistics/simple_hypothesis_testing/problem03.pg`

All Carroll statistics students are assigned to randomly select a sample of 40 people from the Helena community and measure how much time it takes them to solve a word search puzzle. Each student takes the average of the 40 times and submits this average.

The table below gives the distribution of these averages. The first bin contains all averages less than or equal to 15. The second bin contains all averages greater than 15 and less than or equal to 30, and so forth.

Average Time to Solve Puzzle (sec)	≤ 15	15-30	30-45	45-60	6
Number of Statistics Students	1	3	7	9	

At the same time that the students are gathering their data, a professor randomly selects a sample of 40 Helena resident and has them do the same puzzle. However, before the 40 people do the puzzle, the professor first has these 40 people spend one minute doing mental relaxation exercises, to see if this helps people to do the puzzle faster. The professor then calculates that the average of these 40 times is 75 seconds.

What is the null hypothesis?

- A. Doing mental relaxation exercises makes people solve the puzzle slower.
- B. Doing mental relaxation exercises has no effect on the time it takes to solve the puzzle.
- C. Doing mental relaxation exercises makes people solve the puzzle faster.

How many statistics students submitted data?

_____

Using the table of data, estimate the p-value of the professor's result.

_____

For this analysis we will use a significance level of $\alpha = 0.05$. What should we conclude regarding the null hypothesis?

- A. We should reject the null hypothesis.
- B. We cannot reject the null hypothesis.

Which of the following statements is most appropriate regarding our conclusion?

- A. Doing mental relaxation exercises makes people solve the puzzle faster.
- B. We cannot draw any conclusions about whether or not doing mental relaxation exercises has any effect on the time it takes to solve the puzzle.
- C. Doing mental relaxation exercises makes people solve the puzzle slower.
- D. Doing mental relaxation exercises has no effect on the time it takes to solve the puzzle.
- E. None of the statements is an appropriate conclusion.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

14. **(1 point)** `carroll_problib/statistics/simple_hypothesis_testing/problem04.pg`

At the Helena Roasting Company, the empty coffee bags are filed down a conveyer belt and filled by a machine. Due to natural variability of the machine as well as the weight of the beans, no bag is filled to exactly the same weight. The table below shows the average weight of all of the bags on a given day for many days. Assume that the number of bags filled is the same each day.

The table below gives the distribution of the daily averages sampled. The first bin contains all daily averages less than or equal to 200 grams. The second bin contains all daily averages greater than 200 grams and less than or equal to 220 grams. The third bin contains all of the daily averages greater than 220 grams and less than or equal to 240 grams, and so forth.

Average Weight (grams)	$\leq 200g$	200 g to 220 g	220 g to 240 g	240 g to 260 g	260 g to 280 g	$\geq 280g$
Number Days	20	85	147	129	49	19

A maintenance person does some routine repairs on the machinery and we are wondering if these repairs have altered the average daily weight of the bags of coffee. We take one more day's average and find that it is 261 grams.

What is the null hypothesis?

- A. The average daily weight of the bags of coffee is less than it was before.
- B. The average daily weight of the bags of coffee has changed.
- C. The average daily weight of the bags of coffee has remained unchanged.
- D. The average daily weight of the bags of coffee is greater than it was before.

What is the alternative hypothesis?

- A. The average daily weight of the bags of coffee is less than it was before.
- B. The average daily weight of the bags of coffee has changed.
- C. The average daily weight of the bags of coffee is greater than it was before.
- D. The average daily weight of the bags of coffee has remained unchanged.

How many days worth of information did we have before the repairs occured?

____

Using the table of simulated data, estimate the probability of finding the new average weight of 261 grams or more assuming that the repairs had no impact on the machines.

____

For this analysis we will use a significance level of $\alpha = 0.035$.

What should we conclude regarding the null hypothesis?

- A. We should reject the null hypothesis.
- B. We cannot reject the null hypothesis.

Which of the following statements is most appropriate regarding our conclusion?

- A. It appears that the repairs have decreased the average weight of the bags filled by the machines.
- B. It appears that the repairs have increased the average weight of the bags filled by the machines.
- C. We cannot say that the machine is operating the same as it did before the repairs.
- D. We cannot draw any conclusions about whether the repairs increased the average weight of the bags filled by the machines.
- E. None of the statements is an appropriate conclusion.

Answer(s) submitted:

-
-
-
-
-
-

(incorrect)

---

15. **(1 point)** `carroll_problib/statistics/simple_hypothesis_testing/problem05.pg`

In a **matched pairs study** experimenters have one group that performs the same (or similar) tasks under two different conditions. For example, assume that a mathematics education researcher has students perform problem solving tasks with music and without music. Participants are randomly assigned to which treatment they are assigned to first, and the problem solving tasks are different enough that the researchers assume that any learning bias between each participant's two problem sets is negligible.

Each of the problem solving tasks is scored on a 0 to 100 scale, and the difference between the music score and the non-music score is recorded for each participant. If a difference for a participant was negative then the ”no music” score was greater. If a difference for a participant was positive then the ”music” score was greater. A difference of zero indicates that the participant scored the same on the ”music” and ”no music” problem solving sets.

The mathematics education researchers took 125 students through this matched pairs study and found an average difference of -22. They are curious if the scores without music will be better than the scores with music.

The table below shows the result of a simulation built on the following null hypothesis:

Null Hypothesis: The scores on the problem solving tasks with and without music will have the same average. Hence, we assume that the average difference in scores is zero.

Each simulation took 125 pairs of numbers from the same distribution and found their difference. The simulation was repeated many times and the results were tabulated in the table. The first bin in the table contains all average differences (music - no music) that were less than or equal to -20, the second bin in the table contains all average differences (music - no music) that were greater than -20 and less than or equal to -10, and so on.

Average of 125 Paired Differences (music - no music)	≤ -20	-20 to -10	-10 to 0	0 to 10	10 to 20	≥ 20
Count of Simulated Results	22	95	91	115	98	18

What is the alternative hypothesis?

- A. The average score from the problem solving task with music will be the same as the score from the problem solving task without music.
- B. The average score from the problem solving task with music will be different than the score from the problem solving task without music.
- C. The average score from the problem solving task with music will be less than the score from the problem solving task without music.
- D. The average score from the problem solving task with music will be greater than the score from the problem solving task without music.

How many times did the researchers simulate the scenario to get the data in the table?

____

Using the simulated data that we have, estimate the probability of finding a difference of -22 or less.

____

For this analysis we will use a significance level of $\alpha = 0.035$.

What should we conclude regarding the null hypothesis?

- A. We should reject the null hypothesis.
- B. We cannot reject the null hypothesis.

Which of the following statements is most appropriate regarding our conclusion?

- A. We cannot that there is a difference between doing the problem solving tasks with and without music.
- B. We cannot conclude that doing the problem solving tasks without music will result in a better score.
- C. It appears that we have evidence to suggest that doing the problem solving tasks with music will result in a better score.
- D. It appears that we have evidence to suggest that there is a difference between doing the problem solving tasks with and without music.
- E. We cannot conclude that doing the problem solving tasks with music will result in a better score.
- F. It appears that we have evidence to suggest that doing the problem solving tasks without music will result in a better score.
- G. None of the statements is an appropriate conclusion.

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

16. (1 point) carroll_problib/statistics/simple_hypothesis_testing/problem06.pg

A ”social experiment” conducted by a TV program questioned what people do when they see a very obviusly bruised woman getting picked on by her boyfriend. On two different occasions at the same restaurant, the same couple was depicted. In one scenario the woman was dressed ”provocatively” and in the other scenario the woman was dressed ”conservatively”. The table below shows how many restaurant diners were present under each scenario, and whether or not they intervened.

	Provocatively Dressed	Conservatively Dressed	To
Intervened	3	16	
Did Not Intervene	13	12	
Total	16	28	

A simulation was conducted to test if people react differently under the two scenarios. 878 simulated differences were generated to construct the null distribution shown in the table below. The value $\hat{p}_{pr,sim}$ represents the proportion of diners who intervened in the simulation for the provocatively dressed woman, and $\hat{p}_{con,sim}$ is the proportion for the conservatively dressed woman. The first bin contains the number of simulations found with difference less than or equal to -0.45. The second bin contains the number of simulations found with difference greater than -0.45 and less than or equal to -0.35, and so on.

$\hat{p}_{pr,sim} - \hat{p}_{con,sim}$	≤ -0.45	-0.45 to -0.35	-0.35 to -0.25
Count of Simulated Results	6	14	42

(a) What was the null hypothesis used to create the simulated results?

- A. The proportion of people that intervened for the provocatively dressed woman is different than it is for the conservatively dressed woman.
- B. The proportion of people that intervened for the provocatively dressed woman is greater than it is for the conservatively dressed woman.
- C. The proportion of people that intervened for the provocatively dressed woman is less than it is for the conservatively dressed woman.
- D. The proportion of people that intervened is the same for the provocatively dressed woman as it is for the conservatively dressed woman.

(b) What is the alternative hypothesis?

- A. The proportion of people that intervened for the provocatively dressed woman is less than it is for the conservatively dressed woman.
- B. The proportion of people that intervened for the provocatively dressed woman is different than it is for the conservatively dressed woman.
- C. The proportion of people that intervened is the same for the provocatively dressed woman as it is for the conservatively dressed woman.
- D. The proportion of people that intervened for the provocatively dressed woman is greater than it is for the conservatively dressed woman.

(c) Calcuate the observed difference between the rates of intervention under the provocative and conservative scenarios:
$\hat{p}_{pr} - \hat{p}_{con} =$ _____

(c) Estimate the p-value for the hypothesis test based on the simulated results.

p-value $\approx$ _____

For this analysis we will use a significance level of $\alpha = 0.035$.

What should we conclude regarding the null hypothesis?

- A. We should reject the null hypothesis.
- B. We cannot reject the null hypothesis.

Which of the following statements is most appropriate regarding our conclusion?

- A. It appears that we have evidence to suggest that there is no difference between the proportion of people that will intervene based on the way the woman is dressed.
- B. We cannot conclude that fewer people will intervene if the woman is dressed conservatively.
- C. It appears that we have evidence to suggest that fewer people will intervene if the woman is dressed provocatively.
- D. It appears that we have evidence to suggest fewer people will intervene if the woman is dressed conservatively.
- E. We cannot conclude that fewer people will intervene if the woman is dressed provocatively.
- F. We cannot conclude that more people will intervene if the woman is dressed conservatively.
- G. None of the statements is an appropriate conclusion.

Answer(s) submitted:

-
-
-
-
-
-

(incorrect)

---

17. (1 point) `carroll_problib/statistics/simple_hypothesis_testing/problem08.pg`

The Stanford University Heart Transplant Study was conducted to determine whether an experimental heart transplant program increased lifespan. Each patient entering the program was designated an official heart transplant candidate, meaning that he was gravely ill and would most likely benefit from a new heart. Some patients got a transplant and some did not (based on random selection). The variable "transplant" indicates which group the patients were in, patients in the treatment group got a transplant and those in the control group did not. Another variable called "survived" was used to indicate whether or not the patient was alive at the end of the study.

Of the 37 patients in the control group, 4 were alive at the end of the study. Of the 65 patients in the treatment group, 25 were alive at the end of the study. The table below summarizes these results.

	Control	Treatment	Total
Alive	4	25	29
Dead	33	40	73
Total	37	65	102

(a) What proportion of the patients in the treatment group died? _____

(b) What proportion of the patients in the control group died? _____

One approach for investigating whether or not the treatment is effective is to use a randomization technique. Recall that any randomization technique is based on an appropriate null hypothesis.

(c) The null hypothesis in this case examines the difference between the proportion who died in the treatment group ($p_{dead,trmnt}$) vs the proportion who died in the control group ($p_{dead,ctrl}$). Fill in the null hypothesis:
$H_0 : (p_{dead,trmnt} - p_{dead,ctrl})$ [?/¡/¿/=] ___

(d) The doctors clearly did the study to determine if the treatment caused a reduction in the death rate. With this in mind, fill in the alternative hypothesis.
$H_A : (p_{dead,trmnt} - p_{dead,ctrl})$ [?/¡/¿/=] ___

(e) The paragraph below describes the set up for such a randomization approach, if we were to do it without statistical software. Fill in the blanks with a number and choose the correct

drop-down choices.
We write down "alive" on ____ cards representing patients who were alive at the end of the study, and "dead" on ____ cards representing patients who were not. Then, we shuffle these cards and split them into two groups: one group of size ____ representing the treatment, and another group of size ____ representing control. We calculate the difference between the proportion of "dead" cards in the treatment and the control groups (treatment - control) and record this value. We repeat this many many times to build a sampling distribution centered at ____. Lastly, we calculate the fraction of simulation where the simulated difference in proportions are

- ?
- less than or equal to
- greater than or equal to
- equal to
- not equal to

our measured difference in proportions of ____. If this fraction is low, we conclude that it is unlikely to have observed such an outcome by chance and that the null hypothesis should be rejected in favor of the alternative.

(f) Finally, let's assume that the proportion of simulated results as extreme or more so as our measured difference is 0.049. What should we conclude regarding the null hypothesis? (use $\alpha = 0.05$)

- A. We should reject the null hypothesis.
- B. We cannot reject the null hypothesis.

(g) Interpret what your decision to part (f) means.

- A. We do not have evidence to suggest that the control group has a lower death rate.
- B. We do not have evidence that suggests that the treatment group has a lower death rate.
- C. We have evidence that suggests that the control group has a lower death rate.
- D. We have evidence that suggests that the death rates for the control and treatment groups are the same.
- E. We have evidence that suggests that the treatment group has a lower death rate.
- F. none of the above.

(h) Is this study a controlled experiment?

- A. No
- B. Yes
- C. Cannot tell from the given information

(i) Can we conclude a cause and effect relationship based on this study?

- A. No
- B. Yes

Answer(s) submitted:

-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-

(incorrect)

---

18. **(1 point)** `carroll_problib/statistics/simple_hypothesis_testing/pvalue_and_samplesize.pg`
Consider the hypothesis test:

$$H_0 : p = 0.5$$
$$H_A : p < 0.5$$

where p is the proportion of hats coming off of a factory line where the logo has been glued incorrect. Past data indicates that 50% of the hats had glue defects in the past, and the quality control group in the factory is trying to fix this issue. They think with some new machines they have decreased the proportion of defective glueing jobs.

Now that the new machines are in place, the quality control group wants to understanding the effect of sample size on their decision making.

They simulate two sampling distributions for this test:
* one with sample size $N_1 = 36$, and
* one with sample size $N_2 = 144$ (4 times as many as in sample 1).

In each simulation they find a simulated sample proportion, $\hat{p}_{sim}$. They run many thousands of random trials on a computer to get a reasonably accurate picture of the sampling distribution of $\hat{p}$. Recall that a sampling distribution from a simulation shows the results of all of the randomly simulated trials so we can decide if our data would have possibly occurred just by random chance.

The quality control group is most interested in how these two sampling distributions compare to each other and how the different samples sizes affect their statistical decision making.
(a) Fill in the blanks and choose the correct descriptor from the dropdown menu.

	Scenario 1	Scenario 2
Sample Size:	$N_1 = 36$ hats	$N_2 = 144$ hats
Expected center of the Sampling Distribution of $\hat{p}$:	(incorrect)	___
Expected standard deviation of the sampling distribution of $\hat{p}$:	0.08333	• ? • equal to 0.08333 • less than 0.08333 • more than 0.08333

(b) Under this null hypothesis, finding a sample where $\hat{p} = 0.375 = 37.5\%$ of the hats have glue defects is:

- A. more likely from samples of size 144 than from samples of size 36
- B. less likely from samples of size 144 than from samples of size 36
- C. equally likely no matter the sample size

(c) Which of the following is most likely to be a correct statement regarding finding $\hat{p} = 0.375 = 37.5\%$ defective glue jobs in a sample?

- A. The p-value associated with $N_1 = 36$ is $p = 0.001350$, and
  the p-value associated with $N_2 = 144$ is $p = 0.066799$.
- B. The p-value associated with $N_1 = 36$ is $p = 0.066799$, and
  the p-value associated with $N_2 = 144$ is $p = 0.001350$.
- C. The p-value associated with $N_1 = 36$ is $p = 0.001350$, and
  the p-value associated with $N_2 = 144$ is $p = 0.001350$.
- D. The p-value associated with $N_1 = 36$ is $p = 0.066799$, and
  the p-value associated with $N_2 = 144$ is $p = 0.066799$.

Use the decision level of $\alpha = 0.01$ for the following two questions.
(d) Based on your correct answer to part (c), we would
- ?
- fail to reject the null hypothesis
- reject the null hypothesis

when we use a sample size of $N_1 = 36$ to test the hypothesis with $\hat{p} = 0.375$.
(e) Based on your correct answer to part (c), we would
- ?
- reject the null hypothesis
- fail to reject the null hypothesis

when we use a sample size of $N_2 = 144$ to test the hypothesis with $\hat{p} = 0.375$.

Answer(s) submitted:
-
-
-
-
-

(incorrect)

---

19. **(1 point)** carroll_problib/statistics/simple_hypothesis_testing/pvalue_and_samplesize2.pg

(a) According to a recent census a city has a 25% Hispanic population. Three samples are drawn from this population:
* A jury of 12 people was drawn and only 1 of them was Hispanic. What proportion of the jury is this? ____
* A jury panel of 48 was drawn and 4 of them were Hispanic. What proportion of the jury panel is this? ____
* A large committee of 240 people is drawn and only 20 were Hispanic. What proportion of the committee was Hispanic? ____

(b) Which of the following samples is least likely to result from a random sampling of members of the county mentioned in part (a)?

- A. A jury of 12 people from the county is selected with only 1 Hispanic person
- B. A jury panel of 48 people from the county is selected with only 4 Hispanic people
- C. A committee of 240 people from the county is selected with only 20 Hispanic people.

(c) A school district is known to have a 20% black population. Which of the following samples is most likely to result from a random sampling of the students with disciplinary records assuming that the district is disciplining races equally?

- A. A random sample of 1200 disciplinary files finds that 30% of them are of black students.
- B. A random sample of 500 disciplinary files finds that 30% of them are of black students.
- C. A random sample of 100 disciplinary files finds that 30% of them are of black students.

Answer(s) submitted:
-
-
-
-
-

(incorrect)

---

20. **(1 point)** carroll_problib/statistics/InterStats/NormalDist01.pg

What percent of a standard normal distribution $N(\mu = 0, \sigma = 1)$ is found in each region.
Be sure to draw, label, and shade an appropriate normal distribution curve.

In Excel you should be using the "norm.s.dist" command to answer these questions.

(a) $P(z < -1.32) =$ _____
(b) $P(z > 1.37) =$ _____
(c) $P(-0.55 < z < 1.46) =$ _____
(a) $P(-2.09 < z < 2.09) =$ _____

Answer(s) submitted:

-
-
-
-

(incorrect)

---

21. **(1 point)** `carroll_problib/statistics/InterStats/NormalDist02.pg`

A college senior who took the Graduate Record Examination (GRU) scored 622 on the verbal reasoning section and 650 on the quantitative reasoning section. The mean for the verbal reasoning section was 454 with a standard deviation of 118, and the mean score for the quantitative reasoning was 598 with standard deviation 156. Suppose that both distributions are nearly normal.

(a) What is the z-score for the verbal reasoning section? _____
(b) What is the z-score for the quantitative reasoning section? _____
(c) Which section did she do better on relative to the other people that took the exam?

- A. Quantitative Reasoning
- B. The scores were the same
- C. Verbal Reasoning

(d) What is the percentile score for the verbal reasoning? _____%
(e) What is the percentile score for the quantitative reasoning? _____%

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

22. **(1 point)** `carroll_problib/statistics/InterStats/NormalDist03.pg`

In triathlons, it is common for racers to be placed into age and gender groups. Friends Leo and Mary both completed the Helena Triathlon, where Leo competed in the Men, Ages 30-34 group while Mary complete in the Women, Age 25 - 29 group. Leo completed the race in 1 hour, 24 minutes, and 22 seconds (a total of 5062 seconds), while Mary completed the race in 1 hour, 32 minutes, and 25 seconds (a total of 5545 seconds). Obviously Leo finished faster, but they are curious about how they did within their respective groups.

Here is some information on the performance of their groups.
The finishing times of the Men, Ages 30 - 34 group has a mean of 4342 seconds with a standard deviation of 595 seconds.
The finishing times of the Women, Ages 25 - 29 group has a mean of 5356 seconds with a standard deviation of 769 seconds.
The distributions of finishing times for both groups are approximately normally distributed.

(a) What is the z-score for Leo's time? _____
(b) What is the z-score for Mary's time? _____
(c) Did Leo or Mary rank better in their respective groups? (remember that faster times are better)

- A. Mary did better in her group.
- B. Leo did better in his group.
- C. They ranked the same.

For the next two questions remember that a person that went faster in the triathlon has a smaller time.
(d) What percent of the triathletes did Leo finish faster than in his group? _____%
(e) What percent of the triathletes did Mary finish faster than in her group? _____%

Answer(s) submitted:

-
-
-
-
-

(incorrect)

---

23. **(1 point)** `carroll_problib/statistics/InterStats/NormalDist04.pg`

The Graduate Record Examination is built on the following two distributions: the verbal part has distribution N($\mu = 467$, $\sigma = 119$) and the quantitative reasoning part has distribution N($\mu = 583, \sigma = 152$). Use this information to compute the following.

(a) Compute the score of a student who scored in the 80th percentile on the quantitative reasoning section _____
(b) Compute the score of a student who scored worse than 70% of the test takers in the verbal reasoning section? _____

Answer(s) submitted:

-
-

(incorrect)

---

24. **(1 point)** `carroll_problib/statistics/InterStats/NormalDist05.pg`

In two categories for a triathlon, the times were distributed with the following normal distributions:

The times for the Men, age 30-34 were distributed according to N(4389,598).
The times for the Women, age 25 - 29 were distributed according to N(5338,799).
Use this information to compute each of the following:
(a) The cutoff time for the fastest 5% of athletes in the men's group. (In other words, those who took the shortest 5% of time to finish): _____
(b) The cutoff time for the slowest 10% of athletes in the women's group: _____

Answer(s) submitted:

-
-

(incorrect)

---

25. **(1 point)** `carroll_problib/statistics/InterStats/NormalDist06.pg`

The Capital Asset Pricing Model is a financial model that assumes returns on a portfolio are normally distributed. Suppose a portfolio has an average annual return of 13.8

(a) In what percent of years does this portfolio lose money? __________

(b) What is the cutoff for the highest 25__________

Answer(s) submitted:

-
-

(incorrect)

---

26. **(1 point)** `carroll_problib/statistics/InterStats/NormalDist07.pg`

Heights of 10 year olds, regardless of gender, closely follow a normal distribution with mean 55 inches and standard deviation 6 inches.

(a) What is the probability that a randomly chosen 10 year old is shorter than 59 inches? __________
(b) What is the probability that a randomly chosen 10 year old is between 46 and 53 inches? __________
(c) If the tallest 10__________
(d) The height requirement for a roller coaster is 55 inches. What percent of 10 year olds cannot go on this ride? __________

Answer(s) submitted:

-
-
-
-

(incorrect)

---

27. **(1 point)** `carroll_problib/statistics/prob1.pg`

The probability of event A is 1/2. The probability of event B is 1/3. Determine each probability below. If there is not enough information to calculate the requested probability, enter "info" (without quotes).

What is the chance that A and B both happen? ____________________

If A and B are independent, what is the chance that they both happen? ____________________

If A and B are mutually exclusive, what is the chance that they both happen? ____________________

What is the chance that at least one of A or B happens? ____________________

Answer(s) submitted:

-
-
-
-

(incorrect)

---

28. **(1 point)** `carroll_problib/probability/hw2_2.pg`

A deck of six cards consists of three black cards numbered 1, 2, 3, and three red cards numbered 1, 2, 3. First Jenn draws a card at random and keeps it. Then Paul draws a card at random from the remaining cards. Let A be the event that Paul's card has a higher number than Jenn's. Let B be the event that Jenn's card has a higher number than Paul's.

Are A and B mutually exclusive events?

- A. No, because these events are not independent.
- B. Yes, because it is not possible that Paul's number is higher than Jenn's and Jenn's number is higher than Paul's.
- C. Yes, because Jenn and Paul cannot draw the same card.
- D. No, because Jenn and Paul cannot draw the same card.
- E. No, because Jenn and Paul might draw cards that have equal numbers.

Are A and B complements of one another?

- A. No, because these events are not independent.
- B. Yes, because it is not possible that Paul's number is higher than Jenn's and Jenn's number is higher than Paul's.
- C. No, because Jenn and Paul might draw cards that have equal numbers.
- D. No, because Jenn and Paul cannot draw the same card.
- E. Yes, because Jenn and Paul cannot draw the same card.

Answer(s) submitted:

-
-

(incorrect)

---

29. **(1 point)** `carroll_problib/statistics/InterStats/Probability_Dice.pg`

You have two dice: A red 4 sided die and a blue 20 sided die.
(a) You roll both dice. What is the probability of getting a 4 on both dice? _____
(b) What is the probability of rolling the red die 4 times and getting a 1 every time? _____
(c) You roll both dice. What is the probability of getting an odd number on the red die and a 3 on the blue die? _____
(d) You roll both dice. What is the probability of getting an even number on the red die or a 4 on the blue die? _____

Answer(s) submitted:

-
-
-
-

(incorrect)

---

30. **(1 point)** `carroll_problib/statistics/InterStats/Probability_FemaleMale.pg`

A college has a population that is 46% female and 54% male.
(a) If 4 students are picked at random for a student wellness committee, what's the chance that all 4 are female? _____
(b) If 6 students are picked at random for a student wellness committee, what's the chance that all 6 are male? _____
(c) If 6 students are picked at random, what's the chance that none are female? _____
(d) If 3 students are picked at random, what's the chance that none are male? _____

Answer(s) submitted:

-
-
-
-

(incorrect)

---

31. **(1 point)** `carroll_problib/statistics/InterStats/CondProbability_Table.pg`

Are students more likely to use marijuana when their parents used drugs? The following table shows a sample of 450 cases with two variables, "student" and parent", and summarizes the categorical breakdown of the sample. The student either "uses" or "not" where a student is labeled as "uses" if he/she recently used marijuana. The "parents" variables takes the value "used" if at least one of the parents used drugs, including alcohol.

(a) Presume that from the data set we know the following probability:
$P(student = uses|parents = used) = 0.548523206751055$
Use this probability to help fill in the entire table (round to the nearest number of people.

	Parents used	Parents not	Total
Students uses	_____	_____	210
Students not	_____	_____	240
Total	237	213	450

(b) Now complete the following probabilities:
* $P(student = not|parent = not) =$ _____
* $P(student = not|parent = used) =$ _____
* $P(parents = used|student = uses) =$ _____
* $P(parents = used) =$ _____
* $P(students = uses) =$ _____

Answer(s) submitted:

-
-
-
-
-
-
-
-
-

(incorrect)

---

32. **(1 point)** `carroll_problib/statistics/InterStats/CondProbability_SmallPox.pg`

In 1721 there was an outbreak of smallpox in Boston. As modern evidence based medicine as well as vaccinations were emerging at the time doctors gathered a very large data set about inoculations. Doctors at the time believed that inoculations, which involves exposing a person to the disease in a controlled form, could reduce the likelihood of death.

Each case represents one person with two variables: "inoculated" and "result". The variable "inoculated" takes two levels: "yes" and "no", indicating whether the person was inoculated or not. The variable "result" has outcomes "lived" or "died". These data are given in the tables below. In the top table we display the counts of the individual people, and in the bottom table we display the proportion of the people in the study. For the purposes of this problem each table only shows a portion of the information. Start by filling in the remainder of the tables.

	Inoculated	Not Inoculated	Total
Lived	_____	_____	5681
Died	_____	829	_____
Total	_____	_____	6519

	Inoculated	Not Inoculated	Total
Lived	0.034208	_____	_____
Died	_____	_____	_____
Total	0.035588	_____	1

Now determine the probability that an inoculated person died from smallpox.
probability = _____

Answer(s) submitted:

-

•
•
•
•
•
•
•
•
•
•
•
•

(incorrect)

---

33. **(1 point)** `carroll_problib/statistics/InterStats/SensitiveQuestion.pg`

Determining the prevalence of illicit substance in society can be very difficult. Simply looking at police records doesn't give the whole picture, and any surveys designed to determine usage are subject to all sorts of biases. In particular, if responding "yes" to a particular question carries with it a stigma or implication of guilt, then many respondents will lie. To alleviate this issue, we can use a survey design which is based on conditional probability.

Suppose we want to estimate the proportion of the Helena population who have used marijuana within the last month. We select 4896 participants and ask them each to first privately draw a card from a standard 52-card deck.
* If they get a "club" they are asked to respond truthfully to the question "Have you used marijuana within the last month?"
* If they get any other suit they are asked to respond to the question "Does your phone number end in an odd digit?"

The study is blinded in the sense that the researcher doesn't know which question the participant is answering. Hence, participants are guaranteed anonymity and are more likely to be truthful in their responses. The researchers only get a "yes" or "no" answer from each respondent and don't know the actual results of the card draw.

(a) What is the probability of getting a club? ____
(b) What is the probability of not getting a club? ____
(c) What is the expected probability of a person's phone number ending in an odd digit? ____
(d) What is the expected probability of a person's phone number not ending in an odd digit? ____
(e) We don't know the probability that a Helena resident has used marijuana in the last month (that is what we want!), but from this study we know the probability of answering the marijuana question and the percent of participants that answered "yes". Assume that 4896 people are surveyed and 2008 of them say "yes". Use this information to fill in the entire table.

	drew a club	did not draw a club	Total
response = yes	____	____	2008
response = no	____	____	____
Total	____	____	4896

(f) Based on the table, what is a point estimate for the percent of people in the Helena community that have used marijuana in the last month?
Estimate = ____

Answer(s) submitted:

•
•
•
•
•
•
•
•
•
•
•
•

(incorrect)

---

34. **(1 point)** `carroll_problib/statistics/InterStats/SensitivitySpecificity.pg`

In evidence-based medicine, the decision making process uses conditional probability similar to what we have studied in our Statistics class. The language that doctors, nurses, and therapists often use is different than our mathematical language. Instead, you might hear them discuss "sensitivity", "specificity", "accuracy", "positive predictive value" (PPV), and "negative predictive value" (NPV) for a given test.
Here are the definitions:

Sensitivity $= \frac{\text{number of true positives}}{\text{number of true positives}+\text{number of false negatives}}$

The sensitivity is the probability of a positive test given that the patient has the disease.

Specificity $= \frac{\text{number of true negatives}}{\text{number of true negatives}+\text{number of false positives}}$

The specificity is the probability of a negative test given that the patient is well.

Accuracy $= \frac{\text{number of true positives}+\text{number of true negatives}}{\text{total population}}$

The accuracy is the probability of getting a correct test.

PPV $= \frac{\text{number of true positives}}{\text{number of positive outcomes}}$

The PPV is the probability of having a true positive given that the test is positive.

NPV $= \frac{\text{number of true negatives}}{\text{number of negative outcomes}}$

The NPV is the probability of having a true negative given that the test is negative.

In the 2008 study *Physical examination tests of the shoulder: a systemati* the authors discuss tests for **shoulder impingement**. They state that sensitivity and specificity for the "Neer test" is 79% and 53% respectively. (The "Neer test" is named for the inividual that invented the test.)

Let's presume, for the sake of this problem, that 2500 people with shoulder pain were tested with the Neer test. Let's also presume that 1600 of the 2500 people actually have shoulder impingement (even though there is no way to know that ahead of time).

Use the above information to fill in the table below. For the sake of this problem it is ok to have decimal values in the table entries.

	Positive Test	Negative Test	Total
Has Shoulder Impingement	____	____	1600
Does Not Have Shoulder Impingement	____	____	____
Total	____	____	2500

(b) What is the Accuracy of the Neer test? ____
(c) What is the Positive Predictive Value of the Neer test? ____
(d) What is the Negative Predictive Value of the Neer test? ____

Answer(s) submitted:

-
-
-
-
-
-
-
-
-
-

(incorrect)

---

35. **(1 point)** `carroll_problib/statistics/InterStats/SensitivitySpecificity2.pg`

In evidence-based medicine, the decision making process uses conditional probability similar to what we have studied in our Statistics class. The language that doctors, nurses, and therapists often use is different than our mathematical language. Instead, you might hear them discuss "sensitivity", "specificity", "accuracy", "positive predictive value" (PPV), and "negative predictive value" (NPV) for a given test.

Here are the definitions:

Sensitivity $= \frac{\text{number of true positives}}{\text{number of true positives}+\text{number of false negatives}}$

The sensitivity is the probability of a positive test given that the patient has the disease.

Specificity $= \frac{\text{number of true negatives}}{\text{number of true negatives}+\text{number of false positives}}$

The specificity is the probability of a negative test given that the patient is well.

Accuracy $= \frac{\text{number of true positives}+\text{number of true negatives}}{\text{total population}}$

The accuracy is the probability of getting a correct test.

PPV $= \frac{\text{number of true positives}}{\text{number of positive outcomes}}$

The PPV is the probability of having a true positive given that the test is positive.

NPV $= \frac{\text{number of true negatives}}{\text{number of negative outcomes}}$

The NPV is the probability of having a true negative given that the test is negative.

In the 2003 study *A Review of the Special Tests Associated with Shoulder Examination. Part I: The Rotator Cuff Tests.* **American Journal o**

the authors discuss tests for rotator cuff injury in the shoulder. They note that for the Yoccum test (which tests for injury to the **supraspinatus muscle**) the PPV was 88% and the NPV was 94%. This was taken from a population of people that the therapist had already indicated had should impingement.

Let's presume that 2700 people with shoulder impingement are tested with the Yoccum test. Let's also presume that 1600 of the 2700 people have a positive Yoccum test.

(a) Use the above information to fill in the table below. For the sake of this problem it is ok to have decimal values in the table entries.

	Positive Test	Negative Test	T
Has Shoulder Impingement	____	____	_
Does Not Have Shoulder Impingement	____	____	_
Total	1600	____	2

(b) What is the Accuracy of the Yoccum test? ____
(c) What is the Sensitivity of the Yoccum test? ____
(d) What is the Specificity of the Yoccum test? ____

Answer(s) submitted:

-
-
-
-
-
-
-
-
-
-

(incorrect)

---

36. **(1 point)** `carroll_problib/statistics/InterStats/Probability_Cond_Innoculation.pg`

In a large population the probability of someone being stricken with the euphoric mental state "StatsOnMyMind-itis" is 0.104. The probability of a person with StatsOnMyMind-itis of getting into a stats-related conversation is 0.897. The probability of a person that doesn't have StatsOnMyMind-itis of getting into a stats-related conversation is 0.817.

(a) What is the probability of having StatsOnMyMind-itis and ending up in a stats-related conversation? ____
(b) What is the probability of having StatsOnMyMind-itis and not ending up in a stats-related conversation? ____
(c) What is the probability of not having StatsOnMyMind-itis and ending up in a stats-related conversation? ____
(d) What is the probability of not having StatsOnMyMind-itis and not ending up in a stats-related conversation? ____
(e) What is the probability of having StatsOnMyMind-itis given that you just ended up in a stats-related conversation? ____
(f) What is the probability of not having StatsOnMyMind-itis given that you just ended up in a stats-related conversation? ____
(g) What is the probability of having StatsOnMyMind-itis given that you have not ended up in a stats-related conversation? ____
(h) What is the probability of not having StatsOnMyMind-itis given that you have not ended up in a stats-related conversation? ____

Answer(s) submitted:

-

-
-
-
-
-
-
-

(incorrect)

---

37. **(1 point)** `carroll_problib/statistics/InterStats/Probability_tent.pg`

A tent manufacturer knows from past data that there is an 6.3% chance of having a defect in the zipper of a tent and a 6.1% chance of having a defect in the seam sealing treatment on the tent. The machines that do these parts of the tent manufacturing are independent of each other, but every tent gets both a zipper and a seam sealing treatment.

(a) What is the probability that the tent will have a defective zipper AND a defective seam sealing treatment? ____

(b) What is the probability that the tent will have neither a defective zipper nor a defective seam sealing treatment? ____

(c) What is the probability of having at least one of the two possible defects?

Hint: It might be easier to use the complement rule on this one: $P(notA) = 1 - P(A)$

____

(d) What is the probability of getting exactly 1 of the defects?

Hint: If you have two events, A and B, and you want exactly one then calculate the probability of getting A and not B OR the probability of getting B and not A.

____

(e) What is the probability of getting a zipper defect but not a seam sealing defect? ____

(f) What is the probability of getting a seam sealing defect but not a zipper defect? ____

Answer(s) submitted:

-
-
-
-
-
-

(incorrect)

---

38. **(1 point)** `carroll_problib/statistics/lane/Chapter09/problem14.pg`

In a city, 70% of the people prefer Candidate A. Suppose 30 people from this city were sampled.

(a) What is the mean of the sampling distribution of p? ____

(b) What is the standard error of p? ____

(c) What is the probability that 80% or more of this sample will prefer Candidate A? ____

(d) What is the probability that 45% or more of this sample will prefer some other candidate? ____

(Relevant section: **Sampling Distribution of p**)

Answer(s) submitted:

-
-
-
-

(incorrect)