Chapter 4 FRAPPY!

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

In a recent study, 166 adults from the St. Louis area were recruited and randomly assigned to receive one of two treatments for a sinus infection. Half of the subjects received an antibiotic (amoxicillin) and the other half received a placebo.

(a) Describe how the researchers could have assigned treatments to subjects if they wanted to use a completely randomized design.

(b) All the subjects in the experiment had moderate, severe, or very severe symptoms at the beginning of the study. Describe one statistical benefit and one statistical drawback for using subjects with moderate, severe, or very severe symptoms instead of just using subjects with very severe symptoms.

Chapter 4 FRAPPY!
Scoring Guidelines

Intent of the question The primary goals of this question are to assess a student's ability to: (1) describe how to assign treatments in a completely randomized design; (2) identify a benefit and drawback to using subjects with varied initial symptoms; (3) explain the concept of statistical significance in the context of a randomized experiment; (4) explain how blocking can be incorporated in an experiment.

Start here →

Model Solution

(a) Label 83 note cards with "A" and 83 note cards with "B." Shuffle the cards well and hand one card out to each subject at random. Subjects with "A" cards will receive the antibiotic and subjects with "B" cards will receive the placebo.

(b) Benefit: We can make inferences about subjects with moderate, severe, or very severe symptoms and not just those with very severe symptoms.
Drawback: Because subjects with very severe symptoms will likely have different test scores than subjects with moderate symptoms, there will be more variability in test scores if subjects with a range of symptoms are included. This will make it more difficult to find convincing evidence that the antibiotic is more effective than a placebo.

(c) If the difference is not statistically significant, then the difference wasn't large enough to rule out random chance as a plausible explanation. That is, the observed difference could be due to the random assignment and not to the effects of the treatments.

(d) To incorporate blocking, form blocks based on the initial conditions of the patients. That is, put all the patients with very severe symptoms into one block and so on. Then within each block, randomly assign the subjects to treatments as in part (a). Blocking by initial severity will help us account for the additional variability in test scores caused by the differences in severity.

HOW TO SCORE YOUR PAPER

Scoring
Parts (a)–(d) are scored essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows

Essentially correct (E) if the response describes a method of random assignment that is described in sufficient detail and results in a completely randomized design.

Partially correct (P) if the response describes a method of random assignment that results in a completely randomized design, but does not contain sufficient detail.

Incorrect (I) otherwise.

Notes:

- A response that only says "use a random number generator" or "use a random digit table" to assign the treatments is not sufficiently detailed.
- If a response uses random digits, to have sufficient detail the explanation of method must address what to do with repeated numbers.
- If a response uses a random digit table, the subjects must be identified with labels of the same length (e.g., 001 not 1).
- A response that incorporates blocking is scored Incorrect (I).
- A response that uses coin flipping (or equivalent) and stops assigning subjects to a treatment when the number of subjects reaches 83 must put the subjects in random order initially to result in a completely randomized design. Otherwise, score this type of response Incorrect (I).

Part (b) is scored as follows

Essentially correct (E) if the response identifies the benefit of larger scope of inference and drawback of increased variability in test scores.

Partially correct (P) if the response identifies either the benefit or the drawback, but not both.

Incorrect (I) otherwise.

Note:

- Saying only that there will be "more variability" without specifying that there will be more variability in the response variable (test scores) is not sufficient for the drawback component.

Part (c) is scored as follows

Essentially correct (E) if the response indicates that the difference in average test scores could be due to chance and explains that the chance is due to the random assignment of treatments.

Partially correct (P) if the response only states that the difference in average test scores could be due to chance.

Incorrect (I) otherwise.

Part (d) is scored as follows

Essentially correct (E) if the response describes forming blocks by grouping subjects with similar characteristics (e.g., severity of symptoms) and indicates that the subjects will be randomly assigned to treatments within each block.

Partially correct (P) if the response describes forming blocks by grouping subjects with similar characteristics but does not indicate that the subjects will be randomly assigned to treatments within each block.

Incorrect (I) otherwise.

Notes:

- A response that uses a variable other than severity of symptoms must include a justification for the choice of blocking variable that addresses variability in test scores. If the justification is not included but the response is otherwise correct, score the response partially correct (P).
- In part (d), the random assignment within blocks only needs to be mentioned, not described.
- A response that pairs subjects by severity of symptoms and randomly assigns the members of the pair to the two treatments is essentially correct (E).
- In this context, a response that assigns both treatments to each subject, in random order, should be scored partially correct (P).

Each essentially correct (E) section counts as 1 point. Each partially correct (P) section counts as ½ point. If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the overall strength of the response and communication, particularly in parts (a) and (d).

Overall Score

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

Two sample Students with Their Scoring

Chapter 4 FRAPPY! Sample #1

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

In a recent study, 166 adults from the St. Louis area were recruited and randomly assigned to receive one of two treatments for a sinus infection. Half of the subjects received an antibiotic (amoxicillin) and the other half received a placebo.

(a) Describe how the researchers could have assigned treatments to subjects if they wanted to use a completely randomized design.

To assign treatments randomly, you could write A and B on 166 note cards (83 with A and 83 with B), put the cards in a hat and draw at random and give one card to each adult recruited.

(b) All the subjects in the experiment had moderate, severe, or very severe symptoms at the beginning of the study. Describe one statistical benefit and one statistical drawback for using subjects with moderate, severe, or very severe symptoms instead of just using subjects with very severe symptoms.

Benefit: Since the subjects had all kinds of levels of sinus infections, the results and conclusions drawn could be applied generally to all people with sinus infections. Therefore, conclusions can be drawn about the antibiotic's overall effectiveness.

Drawback: Because the levels of sinus infections that the subjects had varies, results are going to have more variability. Thus, precise conclusions cannot be drawn about the level of effectiveness of the antibiotic.

Sample #1

(c) At different stages during the next month, all subjects took the sino-nasal outcome test. After 10 days, the difference in average test scores was *not* statistically significant. In this context, explain what it means for the difference to be not statistically significant.

The results could be due to random chance.

(d) One possible way that researchers could have improved the study is to use a randomized block design. Explain how the researchers could have incorporated blocking in their design.

The researchers could have grouped the subjects according to the severity of their sinus infections (moderate, severe, very severe). They could then randomly assign half of the people in each group to the antibiotic treatment and the other halves to the placebo.

Chapter 4 FRAPPY! Sample #2

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

In a recent study, 166 adults from the St. Louis area were recruited and randomly assigned to receive one of two treatments for a sinus infection. Half of the subjects received an antibiotic (amoxicillin) and the other half received a placebo.

(a) Describe how the researchers could have assigned treatments to subjects if they wanted to use a completely randomized design.

Give each person a number from 1 to 166. Then, use a random number generator to divide the subjects into 2 groups of 83. Flip a coin. If it is heads, the first group gets the antibiotic. If it is tails, the second group gets the antibiotic.

(b) All the subjects in the experiment had moderate, severe, or very severe symptoms at the beginning of the study. Describe one statistical benefit and one statistical drawback for using subjects with moderate, severe, or very severe symptoms instead of just using subjects with very severe symptoms.

One benefit of using people with moderate, severe, or very severe symptoms is that we have a much larger population than if we only used people with very severe symptoms. However, a drawback is that we won't be able to tell if the results were due to the antibiotic or to the symptoms.

Sample #2

(c) At different stages during the next month, all subjects took the sino-nasal outcome test. After 10 days, the difference in average test scores was *not* statistically significant. In this context, explain what it means for the difference to be not statistically significant.

The average test scores for the antibiotic and placebo groups weren't very different.

(d) One possible way that researchers could have improved the study is to use a randomized block design. Explain how the researchers could have incorporated blocking in their design.

The researchers could have blocked their design by separating the subjects with moderate, severe, + very severe symptoms and given half of each block groups the placebo and the other half the antibiotic and then compared the results of the three groups.

Chapter 4 FRAPPY!
Student Samples Commentary

Sample #1

In part (a), the response describes a completely randomized design, but doesn't include sufficient detail about the random assignment process. The response doesn't indicate that the note cards need to be shuffled and doesn't identify which treatments correspond to A and B. Part (a) was scored partially correct (P). In part (b), the response identifies the larger scope of inference as the benefit. However, the response states that the results can be applied to "all people," instead of just people with moderate, severe, or very severe symptoms. This was judged to be a minor error with no loss of credit. The response also provides the correct drawback, stating that there will be more variation in the response variable ("results"). Part (b) was scored essentially correct (E). In part (c), the response correctly states that the results could be due to random chance, but does not address the random assignment to treatments or answer in context. Part (c) was scored partially correct (P). In part (d), the response correctly forms blocks based on severity of symptoms and randomly assigns treatments within each block. Part (d) was scored essentially correct (E). With two parts essentially correct and two parts partially correct, the entire answer judged substantial and earned a score of 3.

Sample #2

In part (a), the response describes a completely randomized design, but doesn't include sufficient detail about the random assignment process. The response doesn't describe how to use the random number generator. Also, flipping a coin to determine which group receives which treatment is unnecessary, as the groups were formed at random. Part (a) was scored partially correct (P). In part (b), the response identifies a larger population as the benefit. However, because the response doesn't address the ability to make an inference about a larger population, no credit is earned for the benefit. The response identifies confounding between symptoms and treatments as the drawback, but these variables will not be confounded in a completely randomized design. Because neither the benefit nor the drawback are correct, part (b) was scored incorrect (I). In part (c), the response does not address the need to consider the role of chance when determining statistical significance. Part (c) was scored incorrect (I). In part (d), the response correctly describes forming blocks based on severity of symptoms but does not indicate that the treatments should be randomly assigned within each block. Part (d) was scored partially correct (P). With two parts partially correct and two parts incorrect, the entire answer was judged minimal and earned a score of 1.