University of Texas at Austin

Problem Set # 3

Experiments. Causation.

Problem 3.1. Consider a study performed by a medical center to determine which of two heart surgeries is most effective: angioplasty (running plastic tubes through the arteries) or bypass (rerouting arteries). The purpose of either procedure is to prolong the life of the patient. The study records the survival time of each patient (measured from the time of the surgery). Identify the response and explanatory variables in this study!

Explanatory: type of surgery

Response: survival time

Problem 3.2. A statistics instructor wants to know which route will get her to school the fastest. Each day from October 2 to November 15, when she gets to the turn point she checks the odometer on her car. If it shows an even number, she takes the freeway; if it shows an odd number, she takes the in-town route. She records the total time each day. What is the explanatory variable in this study? What is the response variable in this study? What kind of a study is this (an *observations study* or an *experiment* or ...)?

Explanatory: Route < in-town / freeway

Response: Commute Time

Problem 3.3. A researcher is studying the effects of a new drug on reducing high blood pressure. He recruits 250 men to test the new active drug against a current standard. At the end of six weeks, the decrease in systolic blood pressure will be evaluated. He believes the drug will be more effective for black men than for white men. What kind of experimental design should he use to properly test his belief?

Blocking.

Problem 3.4. Two amateur gardeners are interested in comparing the yields of two varieties of tomatoes. They each have small backyard gardens. Each gardener is going to plant three plants of each variety in his garden. The first gardener will select six small areas in his garden for planting, then choose three of these at random for the three plants of the first variety and then use the remaining three for the second variety. The second gardener will follow the same procedure with his own randomization in his garden. At the end of the growing season they will compare the yields of the two varieties. What do the gardens represent in this example?

Blocks.

Problem 3.5. (5 points) To examine the relationship between two variables, the variables must be measured from the same ...

a.: ... units.
b.: ... cases.
c.: ... values.
d.: ... parameters.
e.: None of the above is correct.

Problem 3.6. (5 points)
Midsomer University "flips" the classroom for a group of 250 students. Of these students, 150 receive an *A* in the end of the term. From this information you conclude ...

a.: ... "flipping" the classroom is an extremely effective way of teaching.
b.: ... nothing, because the sample size is too small.
c.: ... "flipping" the classroom should become priority on the public education front because it evidently works.
d.: ... nothing, because there is no control group.
e.: None of the above.

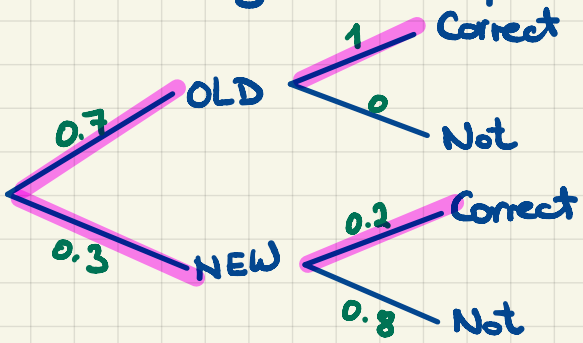
Problem 3.7. (5 points) A medical researcher thinks that adding calcium to the diet will help reduce blood pressure. She believes that the effect is different for men and women. 20 men and 20 women are willing to participate in the study. The researcher chooses 10 of the men and 10 of the women at random. These chosen 20 men and women take a calcium pill every day. The other 20 men and women take a placebo. This is a ...

a.: stratified random sample design.
b.: simple random sample design.
c.: randomized block experimental design.
d.: completely randomized experimental design.
e.: None of the above is correct.

Problem. A student prepares to take a multiple-choice final exam. There is a practice problem set out of which 70% of the problems will be chosen verbatim. The student diligently works through these and decides to guess @ random for the rest out of the 5 offered choices.

(i) What is the probability that the student responds correctly to a randomly chosen question in the exam?

→:



$$\mathbb{P}[\text{Correct}] = 0.7 \cdot 1 + 0.3 \cdot 0.2 = 0.76$$

(ii) Given that the student answered correctly to a question, what's the probability the question had been on the practice test?

→:

$$\mathbb{P}[\text{OLD} \mid \text{Correct}] = \frac{\mathbb{P}[\text{OLD} \cap \text{Correct}]}{\mathbb{P}[\text{Correct}]} = \frac{0.7}{0.76} = \frac{70}{76} = \frac{35}{38}$$

(iii) Let the final have 20 questions. What is the expected value of the number of questions answered correctly?

→: $20 \cdot \mathbb{P}[\text{Correct}] = 20 \cdot 0.76 = 13.2$

(iv) What is the standard deviation of the number of questions answered correctly?

→: N ... # of questions answered correctly

$$SD[N] = \sqrt{Var[N]}$$

$Var[N] = ?$

$$N = 14 + N'$$

$$N' \sim \text{Binomial}(n=6, p=\tfrac{1}{5})$$

$$Var[N] = Var[14 + N'] = Var[N'] = 6 \cdot (\tfrac{1}{5})(\tfrac{4}{5}) = \frac{24}{25}$$

$$SD[N] = \sqrt{\tfrac{24}{25}} \quad \square$$