M358K: September 15th, 2023.

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 2

Randomized response.

Problem 2.1. How to get an answer without asking the question?
The College of Education wanted to gauge the percentage of cheaters in the student population at Wobegon University. They were inspired by the last line in their motto:

"Where all the women are strong,
All the men are good looking, and
All the children are above average.

Realizing the obvious problems with conducting a survey which outright asks the questions: "Are you or have you ever been a cheater?", they devised the following ruse:

A huge bag was filled with many, many question slips. On exactly 70% of the slips, the question

"Have you ever cheated on an exam?"

was written. On the remaining question slips, the question

"Is the last digit of your social security number even?"

was written. Each subject randomly drew a question slip from the bag, read the question in a clandestine manner, responded to the interviewer with a "yes" or "no", and burnt the question slip at the ritual bonfire provided for this occasion. Of course, the interviewer did not know the actual question asked, just the ultimate response. So, there was no real reason for the subject to lie, and we assume that the subjects responded truthfully.

Assume that precisely 50% of the population has an even last digit of the social security number. Let's first formalize the given information in the language of probability.

Question #1. It turned out that 44% of the subjects answered "yes". Give an estimate of the proportion of *cheaters* in this population. ✓

Question #2. What percentage of "yes" answers would you have obtained had *all* the subjects in the population been *cheaters*?

C = {the subject got the "cheater question"} $\rightarrow$ $\mathbb{P}[C] = 0.70$

D = {the response was "Yes"}

$$\boxed{\mathbb{P}[D \mid C^c] = 0.50}$$

Q#1. $\boxed{\mathbb{P}[D \mid C] = ?}$

The Law of Total Probablity

$$\Rightarrow \mathbb{P}[D] = \mathbb{P}[D \cap C] + \mathbb{P}[D \cap C^c]$$
$$= \boxed{\mathbb{P}[D \mid C]} \cdot \mathbb{P}[C] + \mathbb{P}[D \mid C^c] \cdot \mathbb{P}[C^c]$$
$$\boxed{\mathbb{P}[D \mid C]} = ?$$

$$0.44 = \mathbb{P}[D \mid C] \cdot 0.7 + 0.5 \cdot 0.3$$

$$\mathbb{P}[D \mid C] = \frac{0.44 - 0.15}{0.7} = \dots = 0.4143 \checkmark$$

Q#2. Given $\mathbb{P}[D \mid C] = 1$

$$\mathbb{P}[D] = 1 \cdot 0.7 + 0.5 \cdot 0.3 = 0.85 \checkmark$$

Problem 2.2. (13 points) **Mirrored randomized response**

People text while driving, but they are reluctant to admit having done so. You devise an experiment that would help you estimate the proportion of these reckless individuals in the population.

The experiment is conducted in the following way: The subject spins a spinner which lands either on *Question A*, or on *Question B* with the questions being as follows:

Question A: " I have never, not even once, texted while driving."

Question B: "I have at least once texted while driving."

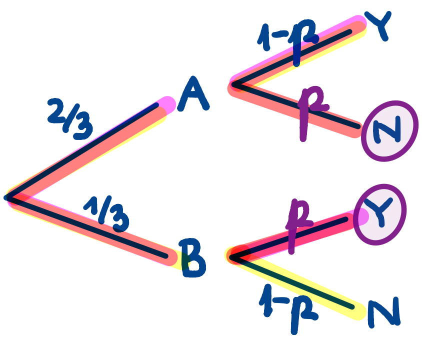
Getting *Question A* is twice as likely as getting *Question B*. The interviewer does not know the actual question asked, just the ultimate response. So, there was no real reason for the subject to lie, and we assume that the subjects responded truthfully.

i. (8 points)

It turned out that 50% of the subjects answered "yes". Give an estimate of the proportion of *terrible texter-drivers* in this population.

$p = ?$

ii. (5 points) What percentage of "yes" answers would you have obtained in an ideal world in which nobody ever endangers the public by simultaneously texting and driving?



$$\frac{2}{3}(1-p) + \frac{1}{3}\cdot p = \frac{1}{2}$$

$$4(1-p) + 2\cdot p = 3$$

$$4 - 4p + 2p = 3$$

$$\boxed{p = \frac{1}{2}}$$

$$\frac{2}{3}\cdot 1 + \frac{1}{3}\cdot 0 = \frac{2}{3} \checkmark$$

Problem 2.3. (13 points) **Forced randomized response**

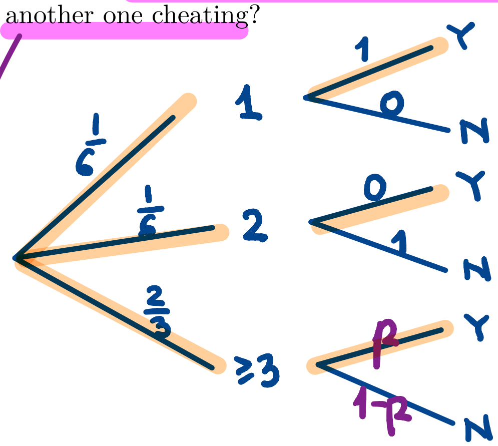
Students sometimes witness another student cheating, but they are reluctant to say so. You devise an experiment that would help you estimate the proportion of these cheaters in the population.

The experiment is conducted in the following way: The subject is instructed to roll an ordinary fair six-sided die without anybody else seeing the outcome. If the die comes up "1", the subject is supposed to say *"Yes"*. If the die comes up "2", the subject is supposed to say *"No"*. If the number of pips on the upturned face of the die is "3" or higher, the subject is supposed to answer the question "Have you ever witnessed a colleague cheating on an exam?"

To reiterate: The interviewer does not know the outcome of the roll of the die, just the ultimate response. So, there was no real reason for the subject to lie, and we assume that the subjects responded truthfully.

i. (8 points) It turned out that 50% of the subjects answered "yes". Give an estimate of the proportion of students who witnessed cheating in this population.

ii. (5 points) What percentage of "yes" answers would you have obtained no student had ever witnessed another one cheating?



$$\frac{1}{2} = \frac{1}{6}\cdot 1 + \frac{2}{3}\cdot p \quad /\cdot 6$$

$$3 = 1 + 4p$$

$$4p = 2$$

$$\boxed{p = \frac{1}{2}}$$

$p = 0 \Rightarrow$ answer: $\frac{1}{6}\cdot 1 = \frac{1}{6}$ ✓

Problem. Four balls are drawn (without replacement) from a box containing 4 black and 5 red balls.

Given that the four drawn balls were not all of the same color, what is the probability that there were exactly two ball of each color among the four?

$\rightarrow$:

$$\begin{cases} E := \{\text{the colors are not all the same}\} \\ F := \{\text{there are exactly 2 black and 2 red}\} \end{cases}$$

$$\mathbb{P}[F \mid E] = \frac{\mathbb{P}[E \cap F]}{\mathbb{P}[E]} = \boxed{\frac{\mathbb{P}[F]}{\mathbb{P}[E]}}$$

$$\mathbb{P}[E] = 1 - \boxed{\mathbb{P}[\text{all } R]} - \boxed{\mathbb{P}[\text{all } B]}$$

$$\mathbb{P}[\text{all } R] = \frac{\binom{5}{4}}{\binom{9}{4}} \qquad \mathbb{P}[\text{all } B] = \frac{1}{\binom{9}{4}}$$

$$\mathbb{P}[E] = 1 - \frac{\binom{5}{4}+1}{\binom{9}{4}} \quad \checkmark$$

$$\mathbb{P}[F] = \frac{\binom{5}{2}\cdot\binom{4}{2}}{\binom{9}{4}} \quad \checkmark$$

$$\mathbb{P}[F \mid E] = \frac{\frac{\binom{5}{2}\binom{4}{2}}{\cancel{\binom{9}{4}}}}{\frac{\binom{9}{4}-\binom{5}{4}-1}{\cancel{\binom{9}{4}}}} = \frac{10\cdot 6}{126-5-1} = \frac{60}{120} = \frac{1}{2} \quad \square$$