Equitable Equations: *Bias and variability*

Problem 1

In each of the following cases, identify both the parameter and statistic. Do you think the difference between the two is more likely attributable to bias or variability, or might it be impossible to tell? Briefly explain your choice.

(a) The average SAT math score in 2022 was 521. In a simple random sample of 25 test-takers from that year, the average score was 595.

(b) On a certain day, 150 reported gas prices at their local gas stations to a tracking website. The mean price was \$5.53/gallon, while the national average was 4.12.

(c) In ten flips of a particular coin, seven were heads

Problem 2

In an opinion poll, researchers asked 200 random customers at Chicago-area McDonalds if they identified as republicans. Of those surveyed, 31% answered "yes."

(a) Identify at least two possible sources of bias in this study.

(b) Suppose the actual percentage of republicans in the Chicago-area was in fact 31%. Does this mean that the study was not biased? Briefly explain.

Problem 3

Why is the following a poor question to ask in order to learn how much people exercise? Suggest better wording.

How much do you exercise most weeks?

Problem 4

Give an example of a statistic with high variability but low bias. Briefly explain your answer.

Problem 1

A) The parameter is the overall SAT score of 521 while the statistic is the average of 595 from the random sample. The difference is likely due to variability because the sample was random.

B) The parameter is the national average price of \$4.12 while the statistic is the average of \$5.52 from the 150 local stations. The difference is due to variability in price from place to place (ex: Wisconsin vs California) and bias due to the fact that local gas stations are more likely to have a higher price than commercial ones.

C)Parameter is the chance of any coin flip landing on heads while the sample is the 10 coin flips. The difference is due to random variability.

Problem 2

A) There may be bias in the demographic of the study, that is people who live in the Chicago area and go to MacDonalds. There is also non-response bias as not everyone who was asked probably actually responded.

B) It does not mean there was no bias it just means the study got lucky. With high variability it is possible to be bias and still get the right answer.

Problem 3

The first problem with this question is that a possible answer to this question is “a lot” with bo further information on how much exercise that person actually gets. People are also likely to lie for social desirability. A better way would be to ask how many hours on average per week or look at heart rate differences before and after exercise.

Problem 4

The average grade of one hour of a class on a scantron test is an example of a statistic with high variability and low bias. The parameter is the average of all the classes. The scantron will not be bias with the test grade (unless it is somehow hacked or broken) and the averages from each class are variable.