`PHSX815_Project1`:
Did my friend Bob cheat in a game of chance?

Kenny Couberly

February 2023

1 Introduction

On Saturday afternoons, I like to play dice games with my good friend Bob. However, since Bob and I live far apart, we play these games over zoom, so we have to trust that one another are being honest in our rolls. Last weekend, I noticed that Bob was rolling higher than he normally does. He claimed to roll an average of 5 in 10 rolls throughout the game using a six sided die. Curious, I decided to do an analysis on this likelyhood to determine if I can trust Bob in these games

Sec. 2 presents the hypotheses we are testing to see which best explains Bob's average roll. Sec. 3 provides a description of how the simulated data used to compare the 2 hypotheses was created as well as the results of the simulation. Sec. 4 performs an analysis on the simulated data. Finally, a conclusion between the hypotheses is reached in Sec. 5.

2 Hypotheses to Bob's Average

The first Hypothesis is that Bob is telling the truth and got an average of 5 after rolling a six-sided die 10 times. Since Bob and I also use a ten-sided die in another one of the games we played, the second hypothesis will be that he used the ten-sided die instead. Since anything above 5 would be similarly alarming, I will consider any averages above 5 in the probability distributions of these dice.

3 Code and Experimental Simulation

To create the simulated data, we take a categorical distribution with n divisions to simulate the roll of an n sided die. To simulate Bob's result from the game, we sample this categorical distribution 10 times and record the average. Considering this result to be a trial, we repeat this result a number of times to produce a distribution of averages. The simulated data below is the result of 20000 trials.

The 6 sided trials resulted in an roll average equal or greater than 5 a total of 316 times for a probability of 0.0158. The 10 sided trials had this result 14307 times for a probability of 0.71535. Let us assume that Bob would call averages above 6 as between 5 and 6 to avoid suspicion, as an average above 7 is not possible for a 6 sided die.

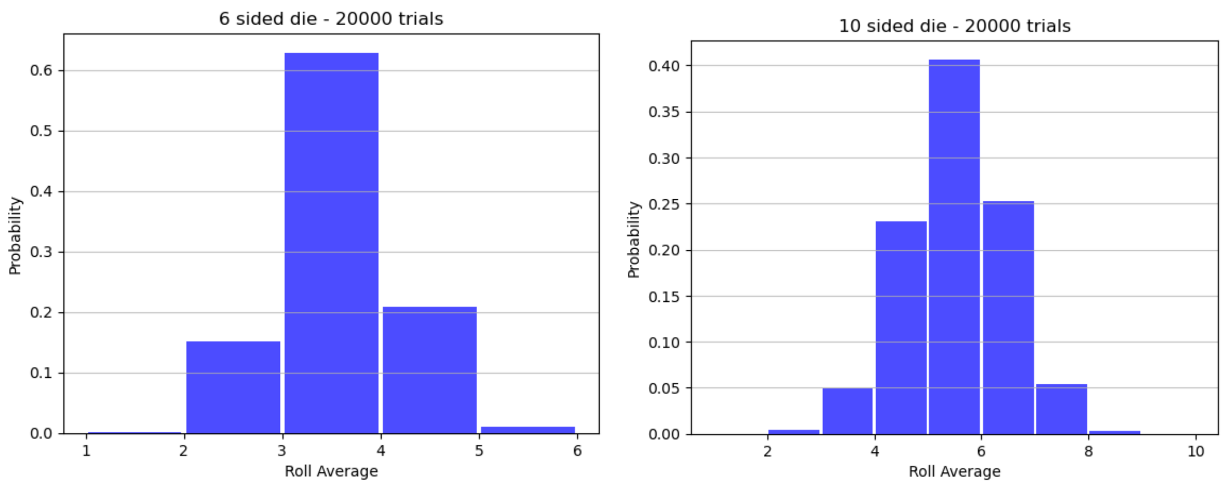

Figure 1: Distributions of the roll averages after 10 rolls for two different hypotheses (Left) Distribution using a 6 sided die. (Right) Distribution using a 10 sided die. Each entry is a simulated experiment where 10 rolls are averaged. 20000 experiments are simulated for each hypothesis. Histograms are normalized to show bins as a portion of the total number of experiments.

4 Analysis

Looking at the previous plots, it is easy to see that through a frequentist inference of statistics the 10 sided die is more likely. However, this interpretation does not take into account any prior experience I have with Bob. Instead, we will use Bayes theorem to compare the hypotheses.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

Hypothesis 1 becomes $(A|B)$ where A is the use of a 6 sided die and B is the average result >= 5. Hypothesis 2 becomes $(A'|B)$ where A' is the use of a 10 sided die and B is the average result >= 5. Bob only has two die so $P(A) + P(A') = 1$. Bob has been my friend for a long time and rarely lies to me, so I estimate the probability that he would lie about the die he used as 0.01. This makes $P(A') = 0.01$ and $P(A) = 0.99$. Since the average result of 5 happened, $P(B) = 1$. Using the result from the simulated data, we get

$$P(A|B) = 0.0158 * 0.99 = 0.0156 \tag{2}$$

and

$$P(A'|B) = 0.71535 * 0.01 = 0.0071535 \tag{3}$$

for hypothesis 1 and 2 respectively. This means that hypothesis 1 is actually about twice as likely when considering the trust I have in Bob.

5 Conclusion

The distributions of roll averages show a much higher probability of getting Bob's result with a 10 sided die than a 6 sided die. However, when I consider the probability that Bob would lie about the die he used, hypothesis 1 remains the more probable explanation. I will continue to monitor what averages Bob gets next time we play. If Bob gets a high average among an even higher number of dice rolls, I may have to update the probability $P(A)$ of how often I consider Bob truthful.