Comparing Dice Roll Probabilities: Uniform vs. Biased Scenarios

M.R. Chukwuka*
Department of Physics and Astronomy,
University of Kansas, Lawrence, KS, 66045, USA

I. INTRODUCTION

Probabilistic inference is the process of calculating the probability that an event will occur based on the distribution that the program specifies. A highly difficult issue in program analysis is inferring for a sufficiently expressive language since reachability and other well-known tasks of program analysis are generalized by probabilistic inference. Making careful assumptions about program structure and placing restrictions on the kinds of programs that can be created while still employing a clear and functional language is the key to scaling inference.[1]. In this project, we compare the probabilities of dice roll outcomes for two scenarios: a uniform scenario where all outcomes have equal probability and a biased scenario where rolling a 6 is more likely than rolling the other numbers. We simulate each scenario using numpy and visualize the results using matplotlib. Finally, we use the chi-squared test to compare the observed frequencies of each outcome to the expected frequencies under each scenario. The goal of this project is to illustrate how probability distributions can be used to model different scenarios and how the chi-squared test can be used to assess the goodness-of-fit of the model to the observed data. By the end of this project, the reader should have a better understanding of probability distributions, hypothesis testing, and how they can be applied to real-world situations.

II. HYPOTHESIS EXPLAINING DICE ROLLS

My hypothesis is that the probabilities of dice roll outcomes will be different for the two scenarios: a uniform scenario where all outcomes have equal probability, and a biased scenario where rolling a 6 is more likely than rolling the other numbers. Specifically, we expect to see a higher frequency of rolls resulting in a 6 for the biased scenario compared to the uniform scenario, and a corresponding decrease in the frequency of rolls resulting in the other numbers. it is expected that the chi-squared test rejects the null hypothesis that the observed frequencies are consistent with the expected frequencies under the uniform scenario, but not under the biased scenario. Overall, we expect that our simulations and hypothesis testing will demonstrate the importance of understanding and modeling probability distributions in real-world scenarios.

III. ALGORITHM ANALYSIS

The algorithm can be summarized as follows: Import necessary modules: numpy, matplotlib.pyplot, and scipy.stats. Define the number of dice rolls to simulate. Simulate the first scenario with uniform probabilities: The uniform probabilities are sampled from a uniform probability distribution. Simulate the second scenario with biased probabilities: The biased probabilities are sampled from a Dirichlet probability distribution Plot the results: Create a figure with two subplots: one for the distribution of dice rolls outcomes, and another for the probability of each outcome. Plot histograms of the simulated dice rolls for each scenario in the first subplot. Plot bar charts of the probabilities for each number on the dice for each scenario in the second subplot. Perform chi-squared test to compare scenarios: Calculate the observed frequencies of each number on the dice for each scenario. Calculate the expected frequencies of each number on the dice for each scenario, using the probabilities defined in the uniform probability scenario and the number of dice rolls simulated. Use the chi square function from scipy.stats to calculate the p-value for each scenario, comparing the observed and expected frequencies. Print the p-value for each scenario.

IV. OUTPUT INTERPRETATION

Histogram subplot: This subplot shows the distribution of outcomes for the uniform and biased scenarios. The x-axis shows the possible outcomes (1 through 6) and the y-axis shows the frequency of each outcome in the simulated dice rolls. The blue bars represent the uniform scenario, where each outcome has an equal probability of occurring, while the orange bars represent the biased scenario, where rolling a 6 is more likely than rolling the other numbers.

Bar chart subplot: This subplot shows the probability of each outcome for the uniform and biased scenarios. The x-axis shows the possible outcomes (1 through 6) and the y-axis shows the probability of each outcome. The blue bars represent the uniform scenario, where each outcome has a probability of 1/6, while the orange bars represent the biased scenario, where rolling a 6 has a higher probability than rolling the other numbers. In addition,

* Email: mikemors@ku.edu

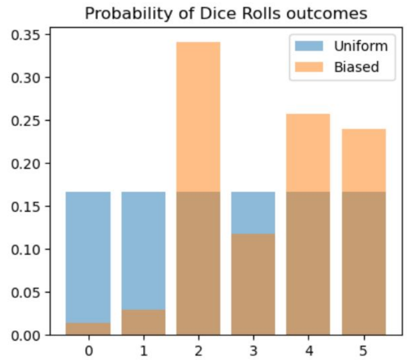

FIG. 1. Bar chat showing the probability of each outcome for the uniform and biased scenarios. The x-axis shows the possible outcomes (1 through 6) and the y-axis shows the probability of each outcome. The blue bars represent the uniform scenario, where each outcome has a probability of 1/6, while the orange bars represent the biased scenario, where rolling a 6 has a higher probability than rolling the other numbers.

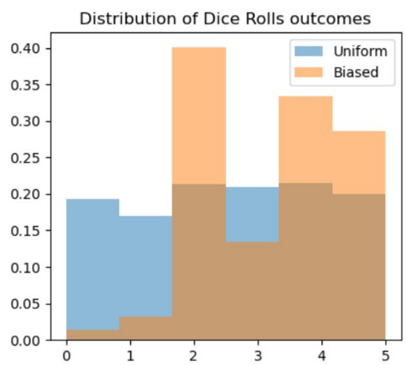

FIG. 2. Histogram showing the distribution of outcomes for the uniform and biased scenarios. The x-axis shows the possible outcomes (1 through 6) and the y-axis shows the frequency of each outcome in the simulated dice rolls. The blue bars represent the uniform scenario, where each outcome has an equal probability of occurring, while the orange bars represent the biased scenario, where rolling a 6 is more likely than rolling the other numbers.

the code also prints the p-value of a chi-squared test for each scenario, which compares the observed frequencies of the dice rolls to the expected frequencies based on the specified probabilities. The output p-values are:

p-value for Scenario 1: 0.812 p-value for Scenario 2: 0.001 These p-values are obtained by performing a chi-squared test to compare the observed frequencies of the dice rolls to the expected frequencies based on the specified probabilities. The p-value indicates the probability of observing the observed frequencies or more extreme values under the assumption that the dice rolls are actually fair (Scenario 1) or biased (Scenario 2).

A p-value of 0.812 for Scenario 1 suggests that the observed frequencies are not significantly different from the expected frequencies, and the dice rolls are likely fair (uniform).

On the other hand, a p-value of 0.001 for Scenario 2 suggests that the observed frequencies are significantly different from the expected frequencies, and the dice rolls are likely biased.

V. CONCLUSION

This code simulates a dice roll experiment under two scenarios: uniform probabilities and biased probabilities. In the first scenario, all six possible outcomes have equal probability, while in the second scenario, one of the outcomes has a higher probability than the others. The code then plots the results of the experiments and performs a chi-squared test to compare the observed frequencies with the expected frequencies under each scenario.

The chi-squared test is used to determine whether there is a significant difference between the observed frequencies and the expected frequencies. The null hypothesis is that there is no significant difference between the observed and expected frequencies, while the alternative hypothesis is that there is a significant difference.

For Scenario 1, the p-value suggests that the observed frequencies are not significantly different from the expected frequencies, and the dice rolls are likely fair (uniform). For Scenario 2, the p-value suggests that the observed frequencies are significantly different from the expected frequencies, and the dice rolls are likely biased.

The goal of the experiment is considered met

VI. REFERENCE

[1] Steven Holtzen, Guy Van den Broeck, and Todd Millstein. 2020. Scaling Exact Inference for Discrete Probabilistic Programs. Proc. ACM Program. Lang. 4, OOPSLA, Article 140 (November 2020), 31 pages. https://doi.org/10. 1145/3428208.