PHSX815_Project3:
Estimating Dice Roll Probability

Kurt Hamblin

April 12, 2021

1 Introduction

In the previous projects, we were comparing an observed distribution a given "true" distribution, with the goal of making a quantitative statement relating the two. In this project, rather than comparing two distributions, we study one observed distribution, and attempt to reconstruct the true distributions, and estimate its parameters.

This paper is organized as follow: Sec. 2 provides an overview of the utilized code and experimental design, and Sec. 3 presents the results, analysis, and conclusions.

2 Overview of Code and Experimental Procedure

For this project, we will roll a 6-sided dice, with random side weights, and we will study the distribution of '3' rolls. In doing so, we will attempt to estimate the true probability (or weight) of side '3' on the die, and we will study how well this probability can be estimated from the simulated data via a Neyman Construction.

2.1 Dice Rolling

To simulate dice rolling, we add a method to the Random class, `Random.roll_die(Nsides, weights)`, that returns an integer from 1 to Nsides. We include the ability to roll a biased die with the inclusion of "weights," which allow the user to specify the exact probability of each side of the die being rolled. The algorithm divides the space from 0 to 1 into slices corresponding to each side of the die (either by splitting 0 to 1 evenly, or using the weights as the boundaries for the 'slices'), and then generates a random float from 0 to 1, and returns the side that this random float corresponds to. We decide on the side weights by sampling from a normal distribution using `numpy.random.normal()`, and we will store the weights to a file, in case of later comparison.

2.2 Parameter Estimation

For this project, we choose to simulate 10^4 experiments, with 500 dice rolls per experiment. We will use these simulated rolls to estimate P_3, the true probability of rolling a '3.' We will count the number of times '3' was rolled, and construct a histogram, which we will fit with a normal distribution to obtain a distribution mean and standard deviation.

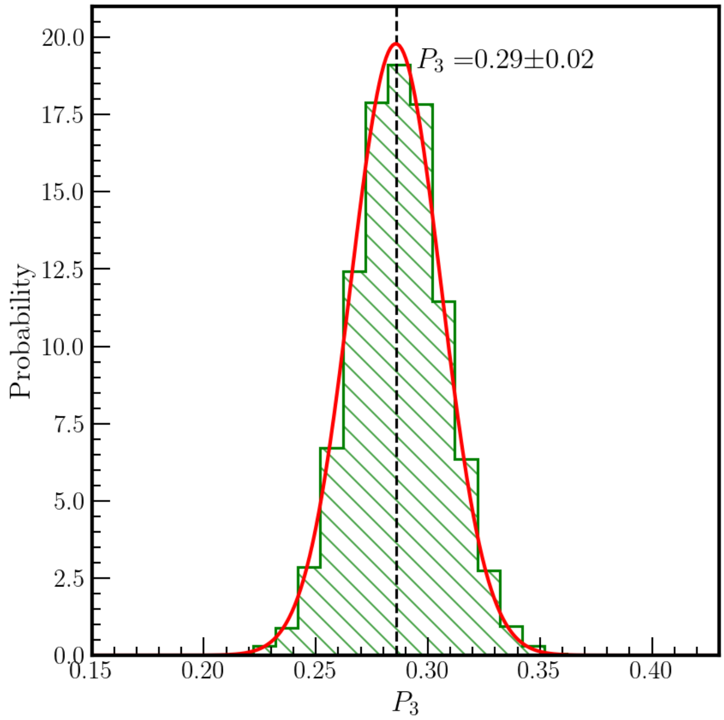

Figure 1: Histogram of observed '3' rolls. A normal distribution is fit and shown in red, with $\mu = 0.29$, and $\sigma = 0.02$.

2.3 Neyman Construction

In order to study how well the true value of P_3 can be estimated from the simulated data, we will make a Neyman Construction. To do so, we will run simulations for 10, 100, and 1000 rolls, iterating over values of $P_{3,true}$ from 0 to 1 in steps of 0.01. Then, we make a 2d-histogram, with the measured probabilities on the y-axis and the true probabilities on the x-axis.

3 Results and Conclusions

We present the histogram results from the parameter estimation in Figure (1), along with the normal distribution fit. From this fit, we find that $P_3 = 0.29 \pm 0.02$, where the error comes from the standard deviation of the fit.

The Neyman Construction discussed in section 2.3 is shown in Figure (2), with 3 separate plots for 10, 100, and 1000 rolls per experiment (with 100 experiments used for each plot). From Figure (2) it is easily seen that, as the number of rolls are increased, the true value of P_3 is better estimated. Even with only 1000 rolls, the Neyman Construction is very tightly spread, and the estimations of P_3 are very effective across all true values of P_3.

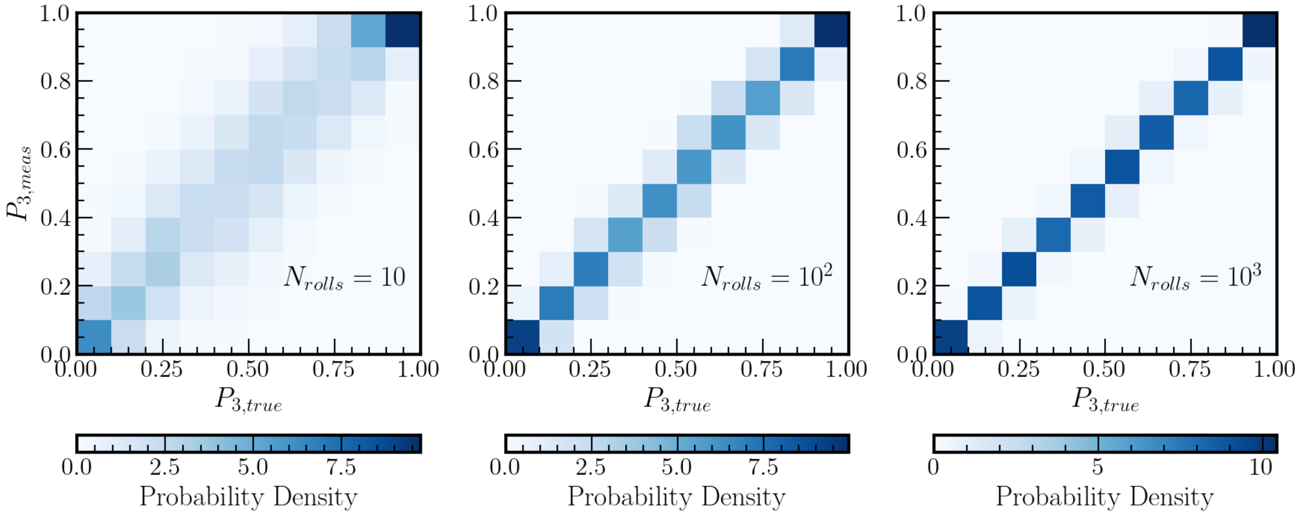

Figure 2: Histogram of observed '3' rolls. A normal distribution is fit and shown in red, with $\mu = 0.29$, and $\sigma = 0.02$.