Quantitative Analysis of Physical data: Assignment 2

Madhurima Nath and Zachary Stottler
Department of Physics, Virginia Tech, Blacksburg, VA, USA 24061
(Dated: February 16, 2017)

When a coin is tossed, there are two possibly outcomes. It can be a head or a tail, which are both equally likely for a fair coin toss. The posterior probability is calculated for the three different values of prior probabilities for the two sets of data obtained from the experiment of tossing a coin. Further, the statistical quantities like median and the interquartile distance are calculated. A measure defined as the figure of merit is introduced to quantify the assessment of the fairness of a coin.

I. INTRODUCTION

The toss of a fair coin has two possible outcomes, heads (H) or tails (T), both of which are equally likely to occur. For a single coin toss the probabilities are given by

$$p(H) = r \qquad \text{and} \qquad p(T) = 1 - r; \tag{1}$$

with r being a real number in the interval $[0, 1]$. We call a coin fair if the probability for heads is the same as for tails, that is $p(H) = p(T)$, which is the case for $r = 1/2$. When this is repeated N times, it is an example of a Bernoulli trial (or binomial trial), which is defined as a random experiment with exactly two possible outcomes, "success" and "failure", with the same probability of success every time the experiment is conducted. Assuming that a successful outcome is a head when the coin is tossed N times, the probability for n_H times heads is given by the binomial distribution

$$p(n_H|rC) = \binom{N}{n_H} r^{n_H}(1-r)^{N-n_H} \tag{2}$$

where C stands for the information describing the problem, and n_T times tails with $n_T = N - n_H$.

In Bayesian statistics, a prior probability distribution of a random event is the probability distribution that expresses the knowledge about the event that is taken into account before an experiment is performed. It is a probability that is assigned before any relevant information is considered. Similarly, a posterior probability is defined as the conditional probability when the relevant information or background is taken into account. If $p(r|C)$ is called prior probability for r and $p(r|n_HC)$ is called posterior probability for r, using the product rule

$$p(n_H|C)p(r|n_HC) = p(r|C)p(n_H|rC) \tag{3}$$

which can be solved for $p(r|n_HC)$

$$p(r|n_HC) = \frac{p(r|C)p(n_H|rC)}{p(n_H|C)} \tag{4}$$

where the quantity $p(n_H|C)$ can be obtained by requiring $p(r|n_HC)$ to satisfy the normalization condition which translates to

$$p(n_H|C) = \int_0^1 dr p(r|C)p(n_H|rC) \tag{5}$$

N	n_H	n_T
10	6	4
26	14	12
50	27	23
75	40	35
100	51	49

TABLE I. Data obtained from coin toss.

These concepts are applied to analyze the our dataset, which was obtained by tossing a coin 100 times, cf. Table I. Three different prior probabilities are utilized in the analysis of this sample. The determined posterior probabilities are then used to assess the fairness of our coin.

Two statistical tools are employed to evaluate the fairness of the coin used to accumulate our dataset, so it will be prudent to define them here. The median is the "middle" value that bifurcates the data into the lower half and the upper half of the sample. It is a measure of the central tendency that requires the arrangement of the data in a monotonic fashion. The interquartile distance is a measure of variability, based on dividing a data set into quartiles. It is the difference between the 25^{th} and the 75^{th} percentiles, denoted as Q_1 and Q_2 respectively, and quantifies the statistical dispersion.

II. EVALUATION OF THE PRIORS

Three distinct prior probabilities, or priors, are employed to characterize the posterior probability of the coin toss. The following three prior probabilities are used in this analysis:

(1) Flat Prior: $p(r|C) = 1$

(2) Gaussian Prior: $p(r|C) = \eta exp\left[-\left(\frac{r-1/2}{\sigma}\right)^2\right]$, where $\sigma = 1/2$ and η is the normalization constant.

(3) δ-Distribution Prior: $p(r|C) = \delta(r - 1/4)$

In Bayesian statistics, a prior is used to express information about an event before any evidence is amassed.

As such, a prior can express both knowledge and uncertainty. For example, prior (1) is what as known as an objective prior. It assigns equal (and nonzero) likelihood to each input probability, thus maintaining the possibility of any input. Contrast this with prior (3), which allows for only one possible input value, namely $P(H) = 1/2$. The $\delta-$distribution expresses absolute certainty in the input and thus dominates the form of the resultant posterior probability density (PDF). Lying between these two extremes, prior (2) assigns nonzero probability to each of the possible inputs and includes information about which inputs are expected to have a higher likelihood. As the gaussian prior is an informative prior, it wields a stronger influence on the final form of the posterior probability density than the flat prior. For few trials, this influence can be distinctly seen, cf. FIG. 1 (top). However, unlike prior (2) this influence diminishes as the number of trials is increased, cf. FIG. 1 (bottom).

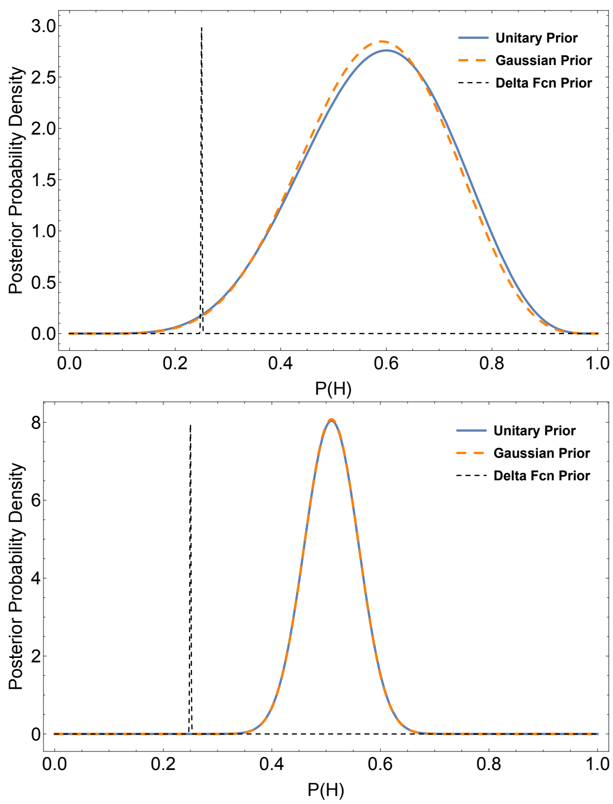

FIG. 1. The posterior probability densities resulting from data runs after 10 (top) and 100 (bottom) tosses. In order to visualize the $\delta-$function prior, a gaussian of very narrow width has been plotted.

III. ANALYSIS FOR ASSESSING FAIRNESS

As we seek to objectively assess the fairness of our coin, we will utilize the flat prior, $p(r|C) = 1$, in our determination of posterior PDFs. Posterior PDFs will be determined from the data after 10, 26, 50, and 100 tosses.

	Median	Q1	Q2	Interquartile Range
Data: 10 trials	0.59	0.49	0.68	0.19
Fair Coin: 10 trials	0.5	0.40	0.60	0.20
Data: 26 trials	0.56	0.49	0.62	0.13
Fair Coin: 26 trials	0.5	0.43	0.56	0.13
Data: 50 trials	0.54	0.49	0.58	0.09
Fair Coin: 50 trials	0.5	0.45	0.55	0.10
Data: 100 trials	0.51	0.48	0.54	0.06
Fair Coin: 100 trials	0.5	0.47	0.53	0.06

TABLE II. The value for the quartile measurements.

The individual data driven posterior PDFs will be compared to simulated posterior PDFs that are generated for an fair (ideal) coin. We define a fair coin to be one that always yields equal numbers of heads and tails. (Note: 26 was chosen to ease the simulation of an ideal coin, ie., a posterior PDF describing 13 heads and 13 tails).

For each subset of the data, the overlap of the distributions will be assessed by examining the fractional overlap of the interquartile ranges of each distribution. For any given subset of the data, the simulated and derived posterior PDFs are of the same form, ie., a normalized binomial distribution. Therefore, we need not concern ourselves with comparisons that rely on the exact overlap of the signal shapes, as this is irrelevant. Rather, we need only concern ourselves with the relative positions of the medians of the PDFs. To quantify the relative overlap of the PDFs, we define a figure of merit:

$$\frac{X_{overlap}}{R_{fair}} \tag{6}$$

where R_{fair} is the interquartile range of the projected posterior PDF of the fair coin and $X_{overlap}$ is the length of the overlap between the interquartile ranges of the data and the simulated ideal coin.

The table II shows the median and the interquartile distances are calculated for the data and the fair coin. As expected the medians for the fair trials are 0.5. The medians for the data approaches this value as we increase the number of trials. This suggests that we need to increase the sample size of our experiment to reduce the errors and obtain a result that is close to that of the fair coin. The decrease in the interquartile range suggests that the widths of the of the posterior PDFs for both the data and the fair trails are decreasing signifying that the outcomes of the trails are clustered closer to the median and hence these are more reliable.

To assess the fairness of the coin, we inspect the fractional overlap of the interquartile ranges of each distribution for the data and the fair coin. The FIG. 2 shows the posterior PDFs for 10, 26, 50, and 100 tosses of the coin and those for the fair coin. It can be seen that as the number of trials increase, the outcomes of our coin approach those of the fair coin. This suggests that large number of samples are required to conclude whether the coin is fair. The figure of merit defined by 6 quantifies

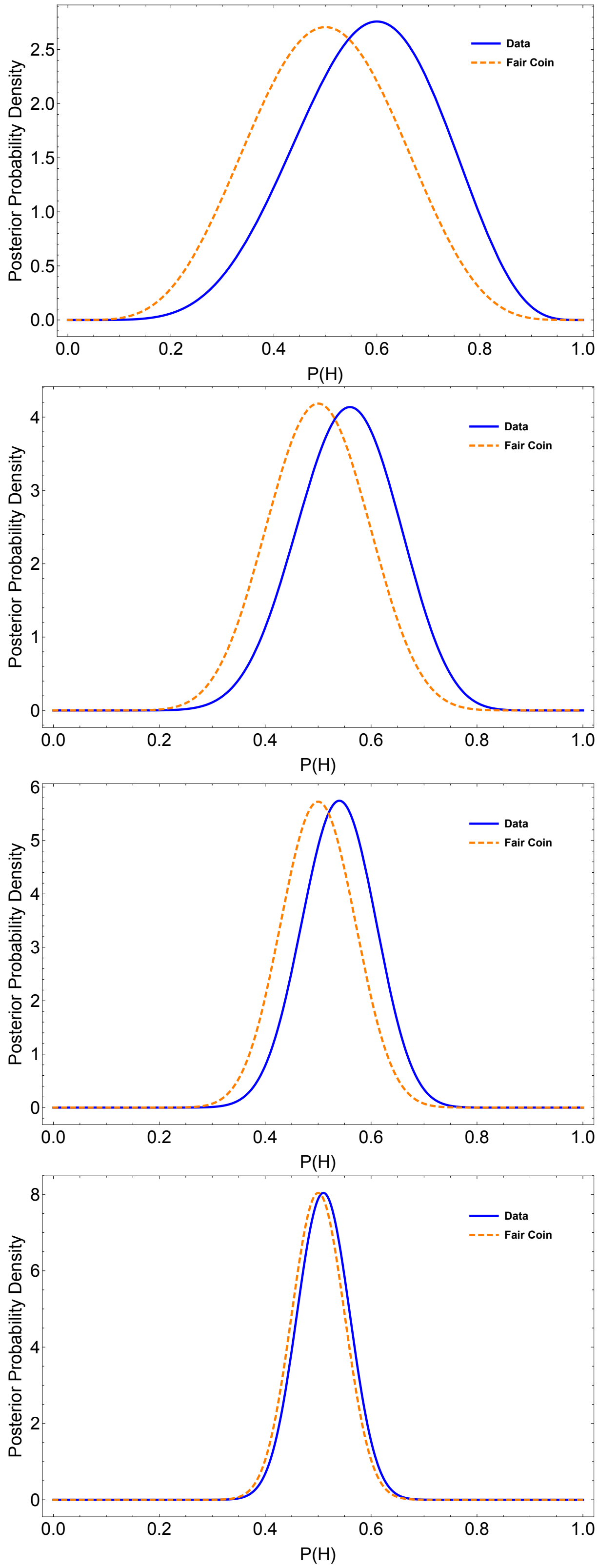

FIG. 2. Posterior probability density functions generated for (top to bottom) 10, 26, 50, and 100 tosses. The data is shown as the blue solid line, and the simulated fair coin is shown as the dashed orange line.

the percentage pf this overlap for the trails. The corresponding values of the figure of merit is tabulated in the following table.

Trials	$\frac{X_{overlap}}{R_{fair}}$
10	0.556
26	0.566
50	0.585
100	0.852

It can be concluded that the overlap is large, $\sim 85\%$ for large number of trials, indicating that the outcomes from the tossing of the coin is tending to the results obtained from a fair one.

IV. CONCLUSIONS

The experiment was to toss a coin 100 times and to analyze the results for different prior probabilities. We observe that the gaussian prior wields a stronger influence on the final form of the posterior probability density than the flat prior, since it is an informative prior. We also define a measure called the figure of merit that quantifies the amount of the overlap of the interquartile distances and suggests that to be certain about the fairness of a coin large number of trials are necessary.

HONOR CODE

I have neither given nor received unauthorized assistance on this assignment.