Bayesian Inference

Writing Assignment 2

Author

Paul Beggs

BeggsPA@Hendrix.edu

Instructor

Prof. Lars Seme, M.S.

Due

OCTOBER 10, 2025

1   Introduction

In probability and statistics, an eponymous theorem developed by Thomas Bayes states

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)},$$

read as "the probability of event A happening given B is equal to the probability of B given A, multiplied by the probability of A, and divided by the probability of B." This theorem is the foundation of Bayesian inference, as it allows us to update our beliefs when presented with new evidence.

Thus, the goal of this paper is to explore exactly that. Given an initial set of beliefs (referred to as a "prior"), how does new information, specifically in the form of successful trials, change those beliefs? This updated belief (called a "posterior") is important because it provides a nuanced method for learning from new evidence, weighing prior assumptions, and updating those beliefs to be as accurate as possible.

2   Probabilistic Distributions

In the following sections, we will examine three different prior distributions and analyze the effect of new information, specifically binomial trials (success or failure), on the resulting posterior distribution. We will use probabilistic distributions that demonstrate the probability density over a given range of values, indicating which value is most probable. We will denote our parameter of interest as p, which represents the probability of success. The "probability density" is the relative likelihood of p occurring.

The three prior distributions to be explored are:

- **Uniform:** Each value of p has an equal probability of occurring.

- **Exponential:** Lower p values have a higher probability of occurring.

- **Bell-curve (Normal or Gaussian):** The probability of p occurring is low at the lower and higher ends of the distribution, but high toward the center.

For each distribution, we will first detail the initial statistics of the prior, such as the mean, median, mode, and variance, along with the middle 50% and 90% credible intervals. Subsequently, we will apply varying amounts of new information from binomial data to the prior to observe how the distribution evolves. We will demonstrate that if we continuously add more trials while maintaining a consistent success rate (specifically 80%), the value $p = 0.80$ will become increasingly likely as the number of trials grows. Furthermore, we will show that the initial state of the prior is critical to the speed at which the posterior changes.

In each of the graphs presented, p ranges from $\{0.01, 0.02, \ldots, 0.99\}$, and the probability density is scaled from 0 to 0.25 to ensure a uniform comparison. This normalization is applied so that all graphs can be viewed with identical axes.

3   Uniform Prior

Initially, we start off with statistics that are self-evident from Figure 1. Our mean and median are both 0.5 because the mean value is just the sum of the lowest p (0.01), and the highest p (0.99) divided by 2, and since each value of p has an equal probability, the halfway point is 0.5. The mode is value that appears the most, but since every p is of equal likelihood, it can be any p. The variance is 0.08. Finally, we also see that the middle 50% and 90% intervals are $[0.25, 0.75]$ and $[0.05, 0.95]$, respectively.
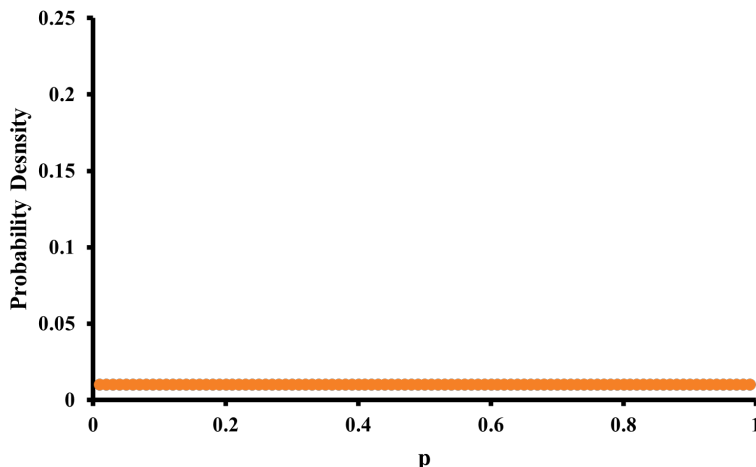

Figure 1: Prior Uniform Distribution

As we start to add more data to the distribution, our statistics start to quickly shift toward 0.80, as is evident in Figure 2. The new mean, median, mode, and variance for 5-trials is 0.71, 0.73, 0.8, and 0.03 and for 50-trials is 0.79, 0.79, 0.8, and 0 respectively. The middle 50% and 90% intervals for 5-trials is $[0.61, 0.83]$ and $[0.41, 0.93]$, and 50-trials is $[0.41, 0.93]$.

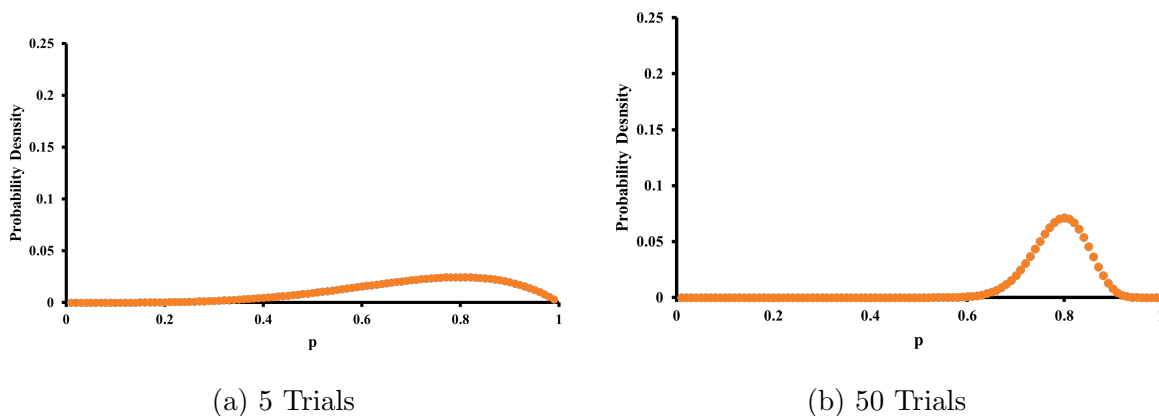(a) 5 Trials   (b) 50 Trials

Figure 2: Posterior Distributions After 5 and 50 Binomial Trials With Uniform Prior

As we get to 500 trials, all statistics (except the variance) are practically equal to 0.80, with the mean, median, mode, and variance being 0.80, 0.79, 0.80, and 0, respectively. In addition, the middle 50% and 90% are also almost entirely centered around .8 with intervals $[0.78, 0.81]$ and $[0.76, 0.82]$. Thus, when we look at Figure 3, it's clear that the new data has a huge impact upon the posterior, and the prior distribution has almost none.
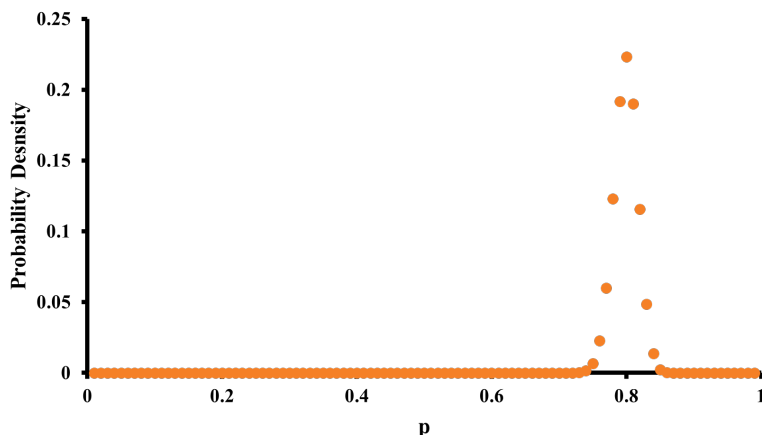

Figure 3: Posterior Distribution After 500 Binomial Trials With Uniform Prior

4   Exponential Prior

The exponential prior distribution looks similar to the uniform distribution, except it has a tail that starts high for low p's and drops toward 0, as can be seen in Figure 4. The initial mean, median, mode, and variance values for the prior distribution are 0.37, 0.32, 0.01, and 0.07. The intervals are $[0.14, 0.57]$ for 50%, and $[0.03, 0.88]$ for 90%. These values make sense as our distribution is skewed toward lower p values.

As we can see in Figure 5, when we add data from the binomial experiments, our posterior will gravitate toward 0.8, but will move slower than the uniform distribution. Thus, for 5-trials, our statistics for the mean, median, mode, and variance are 0.67, 0.68, 0.74, and 0.03, with 50% and 90% intervals $[0.55, 0.80]$ and $[0.36, 0.92]$. We are slowly shifting toward the success rate of 0.8. This becomes more dramatic as we add more trials, such as the 50-trial and 500-trial distributions in Figure 6. The mean, median, mode, and variance for each of these posteriors are 0.78, 0.78, 0.79, and 0, and 0.8, 0.79, 0.8, and 0, respectively. Their 50% and 90% intervals are $[0.74, 0.82]$ and $[0.68, 0.87]$, and $0.78, 0.81$ and $[0.76, 0.82]$, respectively.
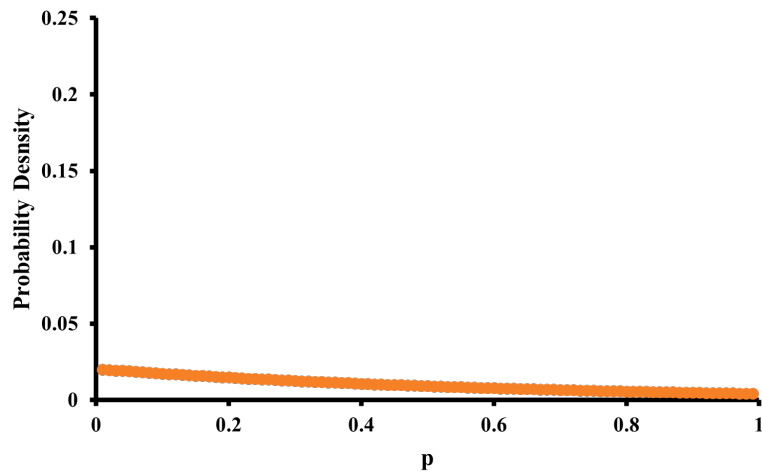
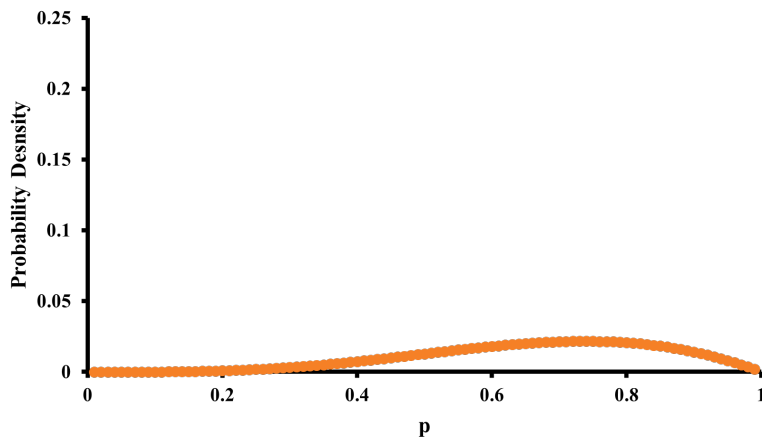Figure 4: Exponential Prior Distribution

Figure 5: Posterior Distribution After 5 Binomial Trials With Exponential Prior
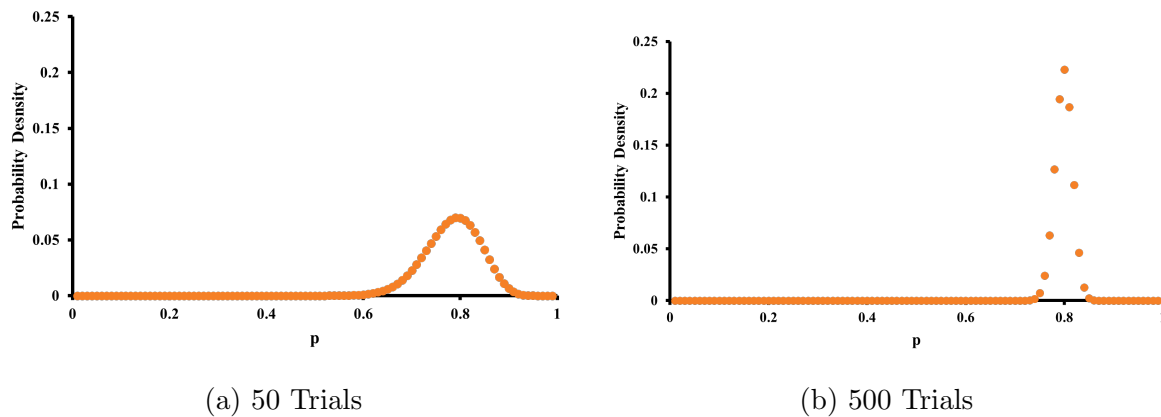

(a) 50 Trials

(b) 500 Trials

Figure 6: Posterior Distributions After 50 and 500 Binomial Trials With Exponential Prior

5 Bell-Shaped Prior

Unlike the previous two distributions, we start off with a moderately high probability of $p = 0.5$. This initial spike has a strong effect upon new posteriors, especially for low trials, as can be seen in Figure 7. For instance, compare the initial statistics of this prior distribution with the 5-trial posterior distribution:

- **Prior:** The mean, median, and mode are all equal to 0.5 with low variance 0.01. The 50% and 90% intervals are $[0.43, 0.56]$ and $[0.33, 0.66]$; and

- **5-Trial Posterior:** The mean, median, and mode all slightly shift away from 0.5 toward 0.55, with the variance equal to 0. The 50% and 90% intervals barely shift too: $[0.48, 0.61]$ and $[0.40, 0.70]$.

This slight change is very important to note. It shows that when we have a strong prior belief, when we add new evidence, we change slightly modify our beliefs, but not by much.
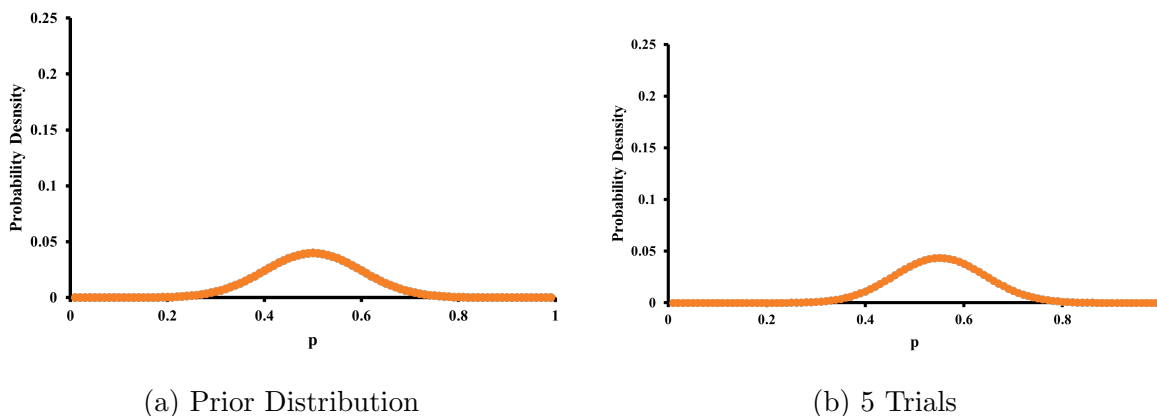

(a) Prior Distribution

(b) 5 Trials

Figure 7: Bell-Shaped Prior and 5-Trial Posterior Distributions

However, when we increase the amount of trials, even though we have a strong prior, the posterior drifts toward 0.8 because of the magnitude of new information, as can be seen in Figure 8. The statistics of these posterior distributions reflect (albeit, to a lesser degree) what we have seen in the previous ones. For the 50-trial distribution, the mean, median, mode, and variance are 0.71, 0.71, 0.71, and 0, with 50% and 90% intervals $[0.67, 0.74]$ and $[0.61, 0.80]$. The 500-trial statistics are even closer to 0.8 with the mean and mode being equal to 0.79, and the median being 0.78 with a variance of 0. The 50% and 90% intervals are $[0.77, 0.80]$ and $[0.75, 0.81]$.

Thus, we can see that with a bell-shaped, high strong initial prior, when we add more information to the distribution, our posterior slowly shifts toward the stronger, new probability. This is in stark contrast to the previous distributions that almost immediately migrated to $p = 0.8$.

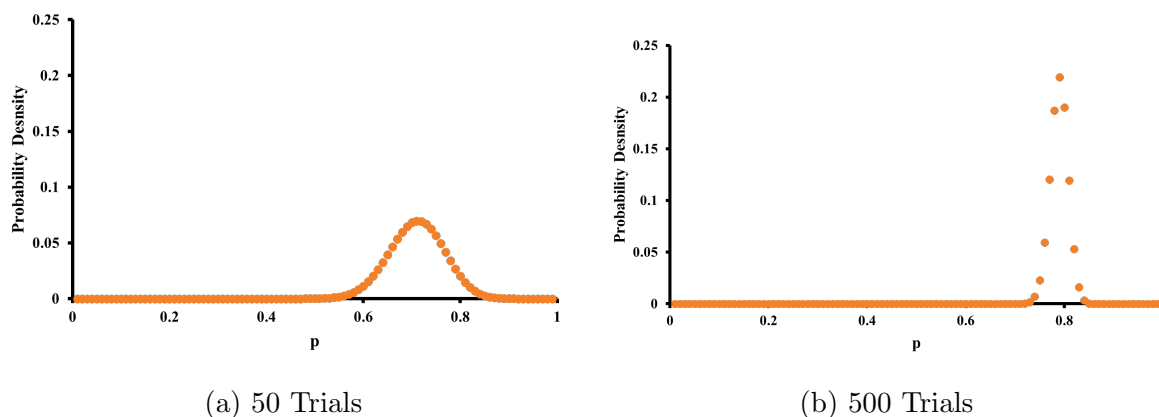(a) 50 Trials   (b) 500 Trials

Figure 8: Posterior Distributions After 50 and 500 Binomial Trials With Bell-Shaped Prior

6   Conclusion

We have reviewed three different distributions, all with different results. We saw that when we start with a uniform prior distribution and add new information, our posterior strongly drifts toward new information with almost no attention being paid to the prior. With the exponential prior, we saw a similar result to the uniform distribution, but with a smaller influence. Finally, we saw that the bell-shaped prior had the strongest opposition to new information because it had an initial high belief centered at $p = 0.5$. Only when we added 500 trials did we truly see a strong shift toward $p = 0.8$.

In our statistics, we saw common trends. For each distribution, the mean, median, and mode were all close, but when we added new information, they all became almost equal to each other. This resulted in an ever shrinking variance, at some points even being approximately equal to 0. Thus, through this exploration, we have found that when the expected outcome for p is approximately equal, the prior has little effect on the posterior; conversely, when there is an initial high value for some p, then the prior has a strong effect on the posterior.