Bayesian Problems

Suryasis Jana

December 29, 2025

Problem 1:
Let N be the number of robins (a bird species) in the forest. Suppose we have conducted an expedition in the forest and observed Y number of robins. Also, suppose that the detection probability of each robin in an expedition is 0.2. Then, given N, it is wise to take the data distribution of Y as

$$Y|N \sim \text{Binomial}(N, 0.2).$$

Further, we have the prior information that there are at most 19 robins in the forest. Then we can consider the prior distribution as

$$N \sim \text{Uniform}(\{0, 1, 2, \ldots, 19\}).$$

Now, think about the following queries:

a) Given that we do not observe any robin, what is the probability that no robin is there in the forest? (0.2023)

b) Intuitively, how would this change if the prior is considered as a uniform distribution over $\{0, 1, 2, \ldots, 99\}$? (0.2000) Why does the probability decrease? Because we have changed our prior belief that there is a higher number of robins in the forest, and $P[N = 0]$ is smaller in this case.

c) Intuitively, how would this change if the detection probability increased from 0.2 to 0.9? (0.9) Why does the probability increase? Because, with better detection probability, we are more confident about the data.

Problem 2:
Suppose you are interested in balloon shooting in a village fair. Say, you shoot $n = 100$ times and hit $Y = 60$ balloons. Let the parameter $\theta \in [0, 1]$ denotes the probability of success. Then, given the value of θ, a reasonable model for the number of hits would be

$$Y|\theta \sim \text{Binomial}(n, \theta).$$

If you have no prior belief about your shooting capability, you can take a non-informative prior as $\pi_1 : \theta \sim \text{Uniform}(0, 1)$, treating all the values in [0,1] as equally likely.

a) Plot the prior density and the posterior density in the same plot.

b) Let us take some other priors as well, given by $\pi_2 : \theta \sim \text{Beta}(0.5, 0.5)$, $\pi_3 : \theta \sim \text{Beta}(2, 2)$, $\pi_4 : \theta \sim \text{Beta}(20, 1)$. For each of the three cases, plot the prior density and the posterior density in the same plot. Out of the four priors, which one do you think is the most informative?

c) Finally, plot all the prior densities together in the same graph and all the posterior densities together in another graph. Comment on the plots.

d) For each of the 4 cases, report the mean and SD for the prior and the posterior distribution. Observe how your prior belief is getting updated due to the use of the data, $Y = 60$. Further, find a 95% posterior credible interval of θ for each case.

e) Suppose we are interested in testing $H_0 : \theta \geq 0.5$ against $H_1 : \theta < 0.5$. For each of the four cases, find the posterior probability of the null and alternative hypotheses.

Problem 3: (Monte Carlo sampling from posterior distribution)
Let $Y_1, Y_2, \ldots, Y_n$ be the sample data from the conditional data distribution $Y_i|\theta \sim \text{Binomial}(20, \theta)$, independently, $\forall i$. The prior distribution of θ is $\text{Beta}(a = 3,\ b = 2)$.

- Considering the true value of θ as 0.4, simulate the sample of size $n = 50$.
- Using this sample data, generate $R = 1000$ sample values from the posterior distribution. These posterior samples can be used to approximate the posterior summaries.
- Use these posterior samples to approximate the posterior mean, posterior SD, and a 95% credible interval of the odds $\gamma = \frac{\theta}{1-\theta}$.
- Repeat the above steps again, considering the prior distribution of θ as $\text{Beta}(a = 3,\ b = 5)$. Comment on the findings.