SSY316, Hand-in 1

Group: *27*
Student: *Jahanvi B Dinesh*
Student: *Sumukh S Moudghalya*

Exercise 1

We are given two boxes with black and white balls:

- Box 1: 3 black balls, 5 white balls
- Box 2: 2 black balls, 5 white balls
- The prior probabilities for selecting each box are equal: $p(\text{Box 1}) = p(\text{Box 2}) = 0.5$.

We need to calculate the posterior probability that a randomly drawn black ball came from Box 1.

- - Let B_1 be the event "the black ball came from Box 1"
- - Let B_2 be the event "the black ball came from Box 2"
- - Let E be the event "a black ball is drawn."
- - Probability of drawing a black ball from Box 1: $P(E|B_1) = \frac{3}{3+5} = \frac{3}{8}$.
- - Probability of drawing a black ball from Box 2: $P(E|B_2) = \frac{2}{2+5} = \frac{2}{7}$.

We need to Apply Bayes' Theorem to find $P(B_1|E)$:

$$P(B_1|E) = \frac{P(E|B_1) \cdot P(B_1)}{P(E)}$$

Now, Calculate $P(E)$, the total probability of drawing a black ball:

$$P(E) = P(E|B_1) \cdot P(B_1) + P(E|B_2) \cdot P(B_2)$$
$$P(E) = \frac{3}{8} \cdot 0.5 + \frac{2}{7} \cdot 0.5$$
$$P(E) = \frac{3}{16} + \frac{2}{14} = \frac{3}{16} + \frac{1}{7} = \frac{21+16}{112} = \frac{37}{112}$$

Finally, Calculate $P(B_1|E)$:

$$P(B_1|E) = \frac{\frac{3}{8} \cdot 0.5}{\frac{37}{112}}$$

Simplifying, we get:

$$P(B_1|E) = \frac{\frac{3}{16}}{\frac{37}{112}} = \frac{3 \times 112}{16 \times 37} = \frac{336}{592} = \frac{42}{74} \approx 0.568$$

The posterior probability that the black ball came from Box 1 is approximately 0.568.

Exercise 2

The weather follows a pattern:

- - If it rains/snows one day, there is a 60% chance it will rain/snow the next day.
- - If it does not rain/snow one day, there is an 80% chance it will not rain/snow the following day.

Part 1

- Let R be the event "it rained/snowed yesterday."
- Let N be the event "it did not rain/snow yesterday."
- Let D be the event "it does not rain/snow today."

Given probabilities:

- - $P(R) = 0.5$ (prior probability that it rained/snowed yesterday).
- - $P(N) = 0.5$ (prior probability that it did not rain/snow yesterday).
- - $P(D|R) = 1-0.6 = 0.4$ (probability that it does not rain today given it rained yesterday).
- - $P(D|N) = 0.8$ (probability that it does not rain today given it did not rain yesterday).

Now, use Bayes' Theorem to find $P(R|D)$:

$$P(R|D) = \frac{P(D|R) \cdot P(R)}{P(D)}$$

Next, calculate $P(D)$, the total probability that it does not rain today:

$$P(D) = P(D|R) \cdot P(R) + P(D|N) \cdot P(N)$$

$$P(D) = 0.4 \cdot 0.5 + 0.8 \cdot 0.5 = 0.2 + 0.4 = 0.6$$

Finally, calculate $P(R|D)$:

$$P(R|D) = \frac{0.4 \cdot 0.5}{0.6} = \frac{0.2}{0.6} = \frac{1}{3} \approx 0.333$$

The probability that it rained/snowed yesterday, given that it does not rain/snow today, is approximately 0.333.

Part 2

This part involves finding the long-term probability (steady-state probability) of rain/snow on a day, based on the transition probabilities.

- - $P(\text{rain tomorrow} \mid \text{rain today}) = 0.6$.
- - $P(\text{no rain tomorrow} \mid \text{no rain today}) = 0.8$.

Let p be the steady-state probability that it rains/snows on any given day.

In the steady state: The probability of it raining today and tomorrow should equal the probability of it not raining today but raining tomorrow.

So we set up the equation:

$$p = 0.6p + 0.2(1 - p)$$

Now, solve for p:

$$p = 0.6p + 0.2 - 0.2p$$
$$p - 0.4p = 0.2$$
$$0.6p = 0.2$$
$$p = \frac{0.2}{0.6} = \frac{1}{3} \approx 0.333$$

The long-term probability that it will rain/snow on any given day is approximately 0.333.

Part 3

New prior: $P(R) = 0.333$ and $P(N) = 0.667$ (since $P(N) = 1 - P(R)$).

Repeat the Bayes' Theorem calculation with the updated prior:

$$P(R|D) = \frac{P(D|R) \cdot P(R)}{P(D)}$$

where $P(D) = P(D|R) \cdot P(R) + P(D|N) \cdot P(N)$.

Now, calculate $P(D)$:

$$P(D) = 0.4 \cdot 0.333 + 0.8 \cdot 0.667$$
$$P(D) = 0.1332 + 0.5336 = 0.6668$$

Finally, calculate $P(R|D)$:

$$P(R|D) = \frac{0.4 \cdot 0.333}{0.6668} \approx \frac{0.1332}{0.6668} \approx 0.2$$

The probability that it rained/snowed yesterday, given that it does not rain/snow today, with the updated prior is approximately 0.2.

Exercise 3

We need to show that the Beta distribution is correctly normalized. The Beta distribution is defined as:

$$\text{Beta}(\mu; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}$$

where:

- - $\mu \in [0, 1]$,
- - a and b are shape parameters,
- - $\Gamma(x)$ is the Gamma function, defined by $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt$.

We aim to prove that this distribution integrates to 1 over the interval $[0, 1]$:

$$\int_0^1 \text{Beta}(\mu; a, b)\, d\mu = 1$$

Set up the integral:

Substitute the expression for $\text{Beta}(\mu; a, b)$:

$$\int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}\, d\mu$$

Next, factor out the constant term:

Since $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ is a constant with respect to μ, we can pull it out of the integral:

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu$$

Now, recognize the integral as the definition of the Beta function:

The Beta function $B(a, b)$ is defined as:

$$B(a, b) = \int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu$$

Therefore,

$$\int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu = B(a, b)$$

The Beta function $B(a, b)$ is related to the Gamma function by:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Substitute $B(a, b)$ into our original expression:

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot B(a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Finally,

The terms $\Gamma(a+b)$, $\Gamma(a)$, and $\Gamma(b)$ cancel out:

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = 1$$

Conclusion:

$$\int_0^1 \text{Beta}(\mu; a, b)\, d\mu = 1$$

This confirms that the Beta distribution is correctly normalized.

Exercise 4

We need to analyze the probability distribution for the parameter μ of a coin-flipping scenario. Here, $x \in \{0, 1\}$ represents a coin flip, with $x = 1$ indicating heads. We are given the following prior beliefs about the fairness of the coin:

$$p(\mu) = \frac{1}{2}\text{Beta}(\mu; 1, 1) + \frac{1}{2}\delta(\mu - 0.5)$$

This prior represents a mixture model, where:

- - With probability 0.5, we believe μ follows a uniform distribution on $[0, 1]$ (i.e., $\text{Beta}(\mu; 1, 1)$).
- - With probability 0.5, we believe the coin is fair, i.e., $\mu = 0.5$.

We need to update our belief about μ after observing coin flips.

Part 1

For a single observation $x_1 = 1$, the likelihood is given by $p(x_1 = 1|\mu) = \mu$.

The posterior $p(\mu|x_1 = 1)$ is given by:

$$p(\mu|x_1 = 1) = \frac{p(x_1 = 1|\mu)p(\mu)}{p(x_1 = 1)}$$

where $p(x_1 = 1)$ is the marginal probability of observing $x_1 = 1$.

Now, calculate $p(x_1 = 1)$:

$$p(x_1 = 1) = \int_0^1 p(x_1 = 1|\mu)p(\mu)\, d\mu$$

Expanding $p(\mu)$ in terms of the mixture model:

$$p(x_1 = 1) = \frac{1}{2}\int_0^1 \mu\, d\mu + \frac{1}{2} \cdot 0.5$$

- - For the first term: $\int_0^1 \mu\, d\mu = \frac{\mu^2}{2}\Big|_0^1 = \frac{1}{2}$.
- - For the second term: $0.5 \cdot 0.5 = 0.25$.

Therefore,

$$p(x_1 = 1) = \frac{1}{2} \cdot \frac{1}{2} + 0.25 = 0.25 + 0.25 = 0.5$$

Next, compute $p(\mu|x_1 = 1)$:

Substitute back into Bayes' Theorem:

$$p(\mu|x_1 = 1) = \frac{\mu \cdot \left(\frac{1}{2}\text{Beta}(\mu; 1, 1) + \frac{1}{2}\delta(\mu - 0.5)\right)}{0.5}$$

Simplifying:

$$p(\mu|x_1 = 1) = \mu \cdot (\text{Beta}(\mu; 1, 1) + \delta(\mu - 0.5))$$

This posterior shows that observing $x_1 = 1$ slightly shifts our belief, increasing the probability that μ might be closer to 1, but still keeps a significant weight on $\mu = 0.5$. This shows that while fairness hasn't changed after observing $x_1 = 1$ one might conclude that the coin is a little biased to coming up as heads.

Part 2

Posterior $p(\mu|x_1, x_2)$ after Observing $x_1 = 1$ and $x_2 = 1$

Given Information: Prior belief about μ (the probability of heads):

$$p(\mu) = \frac{1}{2}\text{Beta}(\mu; 1, 1) + \frac{1}{2}\delta(\mu - 0.5)$$

This prior represents a mixture model:

- With probability 0.5, we assume μ is uniformly distributed on $[0, 1]$.
- With probability 0.5, we assume the coin is fair ($\mu = 0.5$).

The likelihood of observing two heads in a row ($x_1 = 1$ and $x_2 = 1$), given μ, is:

$$p(x_1 = 1, x_2 = 1|\mu) = \mu^2$$

We need to update our prior based on this new information.

Now, the posterior distribution $p(\mu|x_1 = 1, x_2 = 1)$ is given by:

$$p(\mu|x_1 = 1, x_2 = 1) = \frac{p(x_1 = 1, x_2 = 1|\mu)p(\mu)}{p(x_1 = 1, x_2 = 1)}$$

where $p(x_1 = 1, x_2 = 1)$ is the marginal probability of observing two heads in a row.

Using the prior as a mixture model, we can write $p(x_1 = 1, x_2 = 1)$ as:

$$p(x_1 = 1, x_2 = 1) = \int_0^1 p(x_1 = 1, x_2 = 1|\mu)p(\mu)\, d\mu$$

Expanding $p(\mu)$ with the mixture components:

$$p(x_1 = 1, x_2 = 1) = \frac{1}{2}\int_0^1 \mu^2\, d\mu + \frac{1}{2} \cdot 0.5^2$$

- For the first term: $\int_0^1 \mu^2\, d\mu = \left[\frac{\mu^3}{3}\right]_0^1 = \frac{1}{3}$.
- For the second term: $0.5^2 = 0.25$.

So,

$$p(x_1 = 1, x_2 = 1) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot 0.25 = \frac{1}{6} + \frac{1}{8} = \frac{4 + 3}{24} = \frac{7}{24}$$

Compute $p(\mu|x_1 = 1, x_2 = 1)$:

Substitute the values back into Bayes' theorem:

$$p(\mu|x_1 = 1, x_2 = 1) = \frac{\mu^2 \cdot \left(\frac{1}{2}\,\text{Beta}(\mu; 1, 1) + \frac{1}{2}\,\delta(\mu - 0.5)\right)}{\frac{7}{24}}$$

Simplify by multiplying through by $\frac{24}{7}$:

$$p(\mu|x_1 = 1, x_2 = 1) = \frac{24}{7}\,\mu^2 \cdot \left(\frac{1}{2}\,\text{Beta}(\mu; 1, 1) + \frac{1}{2}\,\delta(\mu - 0.5)\right)$$

Breaking this into the two parts of the mixture model:

$$p(\mu|x_1 = 1, x_2 = 1) = \frac{12}{7}\,\mu^2\,\text{Beta}(\mu; 1, 1) + \frac{12}{7}\,\mu^2\,\delta(\mu - 0.5)$$

Part 3

Probability that the Coin is Fair Given Observations $x_1 = 1$ and $x_2 = 1$

Let fair denote the event that $\mu = 0.5$.

Using Bayes' theorem, the probability that the coin is fair given the observations $x_1 = 1$ and $x_2 = 1$ is:

$$p(\text{fair}|x_1 = 1, x_2 = 1) = \frac{p(x_1 = 1, x_2 = 1|\text{fair})\,p(\text{fair})}{p(x_1 = 1, x_2 = 1)}$$

Since we assume the coin is fair, $\mu = 0.5$, so:

$$p(x_1 = 1, x_2 = 1|\text{fair}) = 0.5^2 = 0.25$$

The prior probability of the coin being fair: $p(\text{fair}) = 0.5$.

We previously found that $p(x_1 = 1, x_2 = 1) = \frac{7}{24}$.

Substitute into Bayes' theorem:

$$p(\text{fair}|x_1 = 1, x_2 = 1) = \frac{0.25 \cdot 0.5}{\frac{7}{24}} = \frac{0.125}{\frac{7}{24}} = \frac{0.125 \times 24}{7} = \frac{3}{7} \approx 0.4286$$

The probability that the coin is fair given the observations $x_1 = 1$ and $x_2 = 1$ is approximately 0.4286.

Exercise 5

We need to model the spread of disease using Bayesian inference, where the daily count of new cases follows a Poisson process, and we use a Gamma prior distribution to express our prior belief about the rate λ of new cases per day.

Part 1: Theoretical Questions

Derive the likelihood of observing a dataset $x = \{x_1, x_2, \dots, x_n\}$, where each x_i follows a Poisson distribution with an unknown rate parameter λ.

Since each x_i is independently Poisson distributed with rate λ, the probability of observing x_i is given by:

$$P(x_i|\lambda) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

The likelihood of observing the entire dataset $x = \{x_1, x_2, \dots, x_n\}$ is the product of individual probabilities:

$$P(x|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

Simplifying, we get:

$$P(x|\lambda) = \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!}$$

Assume a Gamma prior on λ with shape parameter α and rate parameter β. Using Bayes' theorem, we derive the posterior distribution for λ after observing the dataset x.

The Gamma prior on λ is given by:

$$p(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

By Bayes' theorem, the posterior distribution $p(\lambda|x)$ is:

$$p(\lambda|x) \propto P(x|\lambda) p(\lambda)$$

Substitute $P(x|\lambda)$ and $p(\lambda)$:

$$p(\lambda|x) \propto \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} \cdot \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

Simplify:

$$p(\lambda|x) \propto \lambda^{\alpha-1+\sum_{i=1}^{n} x_i} e^{-(\beta+n)\lambda}$$

This is the form of a Gamma distribution with updated shape and rate parameters:

- - Updated shape: $\alpha' = \alpha + \sum_{i=1}^{n} x_i$
- - Updated rate: $\beta' = \beta + n$

To show that the posterior distribution is also a Gamma distribution, and specify the updated shape and rate parameters of the posterior.

From the above expression, we see that $p(\lambda|x)$ has the form of a Gamma distribution with parameters α' and β':

$$p(\lambda|x) = \text{Gamma}(\lambda; \alpha', \beta')$$

where:

- $\alpha' = \alpha + \sum_{i=1}^{n} x_i$
- $\beta' = \beta + n$

Part 2: Programming Simulation

1) Generate a synthetic dataset by simulating counts from a Poisson distribution with a true rate parameter λ_{true}.

Choose a true value for λ_{true} (e.g., $\lambda_{\text{true}} = 3$) and generate a dataset of daily case counts using a Poisson distribution.

2) Assume a Gamma prior on λ with initial parameters $\alpha = 2$ and $\beta = 2$. After observing the synthetic data, update the prior parameters to obtain the posterior distribution.

Using the synthetic data, compute α' and β' as derived above.

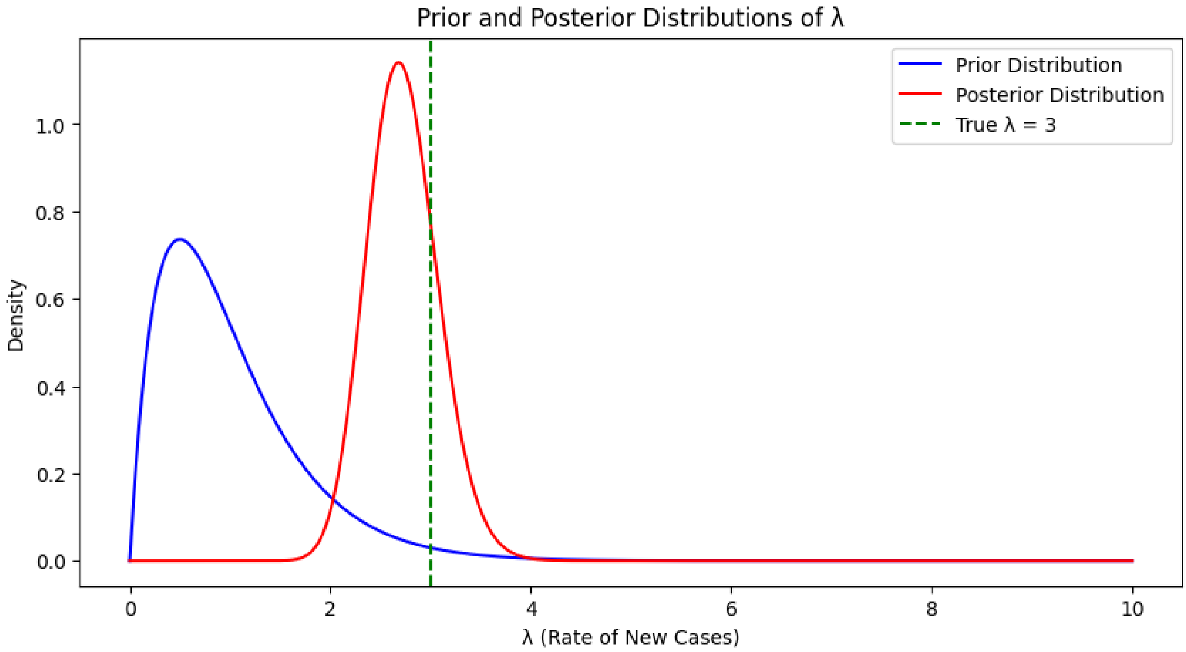
3) Write a Python function that, given data and prior parameters α and β, calculates the updated posterior parameters for λ. [Code is shown under appendix]

Part 3: Visualization

1) Plot the prior and posterior distributions for λ, showing how observing the data has shifted our belief about the rate parameter.

We used Matplotlib to plot the prior and posterior distributions

2) Compare the posterior distribution to the true value of λ_{true} used to generate the data.



Part 4: Analysis

1) Explain the impact of observing more data on the posterior distribution. How does the posterior distribution become more concentrated as the sample size grows?

As more data is observed, the posterior distribution incorporates more evidence about the true rate λ. This leads to a posterior that is more sharply peaked around the true value of λ_{true}, indicating increased confidence in the estimate of λ.

In Bayesian inference, this concentration reflects the reduction of uncertainty as more observations are collected, which makes the posterior distribution narrower and centered closer to the true value. This behavior shows how Bayesian updating refines beliefs with accumulating evidence, especially when the model assumptions are correct.

Appendix

Codes:

1. Programming

```
import numpy as np
import scipy.stats as stats
import matplotlib.pyplot as plt

np.random.seed(0)
lambda_true = 3  #given
n_days = 20
observed_data = np.random.poisson(lambda_true, n_days)

# given
a_prior = 2
b_prior = 2

# Step 3: Update posterior parameters
a_posterior = a_prior + np.sum(observed_data)
b_posterior = b_prior + n_days

def update_posterior(data, a, b):
   return a + np.sum(data), b + len(data)

print("Observed data:", observed_data)
print("Updated posterior shape parameter (alpha):", a_posterior)
print("Updated posterior rate parameter (beta):", b_posterior)
```

2. Visualization:

```
lambda_range = np.linspace(0, 10, 500)
prior = stats.gamma.pdf(lambda_range, a=a_prior, scale=1/b_prior)
posterior = stats.gamma.pdf(lambda_range, a=a_posterior, scale=1/b_posterior)
```

```python
# Plotting
plt.figure(figsize=(10, 5))
plt.plot(lambda_range, prior, label='Prior Distribution', color='blue')
plt.plot(lambda_range, posterior, label='Posterior Distribution', color='red')
plt.axvline(lambda_true, color='green', linestyle='--', label=f'True λ = {lambda_true}')
plt.xlabel('λ (Rate of New Cases)')
plt.ylabel('Density')
plt.title('Prior and Posterior Distributions of λ')
plt.legend()
plt.show()
```