SSY316 Assignment 1

Edoardo Mangia
edoardom@student.chalmers.se

Nuree Kim
nuree@student.chalmers.se

November 2024

Contents

1 Exercise 1

Consider two boxes with white and black balls. Box 1 contains three black and five white balls, and Box 2 contains two black and five white balls. First, a box is chosen at random with a prior probability $p(\text{box} = 1) = p(\text{box} = 2) = 0.5$. Then, a ball is picked at random from the chosen box. This ball turns out to be black. What is the posterior probability that this black ball came from Box 1?

Solution

To find $P(\text{Box 1} \mid \text{Black})$, we apply Bayes' Theorem, which is stated as:

$$P(\text{Box 1} \mid \text{Black}) = \frac{P(\text{Black} \mid \text{Box 1}) \cdot P(\text{Box 1})}{P(\text{Black})}$$

Where:

$$P(\text{Black}) = P(\text{Black} \mid \text{Box 1}) \cdot P(\text{Box 1}) + P(\text{Black} \mid \text{Box 2}) \cdot P(\text{Box 2})$$

The probability of drawing a black ball for each box is

$$P(\text{Black} \mid \text{Box 1}) = \frac{3}{8}$$
$$P(\text{Black} \mid \text{Box 2}) = \frac{2}{7}$$

Therefore, the probability of drawing a black ball is

$$P(\text{Black}) = \left(\frac{3}{8} \times 0.5\right) + \left(\frac{2}{7} \times 0.5\right) = \frac{37}{112}$$

Now, applying Bayes' Theorem

$$P(\text{Box 1} \mid \text{Black}) = \frac{P(\text{Black} \mid \text{Box 1}) \cdot P(\text{Box 1})}{P(\text{Black})} = \frac{\left(\frac{3}{8}\right) \times 0.5}{\frac{37}{112}} = \frac{21}{37}$$

Answer

The posterior probability that the black ball came from Box 1 is approximately 56.8%.

2 Exercise 2

The weather in Gothenburg can be summarized as follows: if it rains or snows on one day, there is a 60% chance it will also rain or snow the following day; if it does not rain or snow on one day, there is an 80% chance it will not rain or snow the following day either.

(i) Assuming that the prior probability it rained or snowed yesterday is 50%, what is the probability that it was raining or snowing yesterday given that it does not rain or snow today?

(ii) If the weather follows the same pattern as above, day after day, what is the probability that it will rain or snow on any given day (based on an effectively infinite number of days of observing the weather)?

(iii) Use the result from part (ii) above as a new prior probability of rain/snow yesterday and recompute the probability that it was raining or snowing yesterday given that it does not rain or snow today.

Solution

Here is the description of the notation and probabilities used in the formulation:

- **States:**
  - R: It rains or snows.

- N: It does not rain or snow.

- **Transition Probabilities:**

$$P(R_{\text{today}} \mid R_{\text{yesterday}}) = 0.6$$

$$P(N_{\text{today}} \mid R_{\text{yesterday}}) = 0.4$$

$$P(N_{\text{today}} \mid N_{\text{yesterday}}) = 0.8$$

$$P(R_{\text{today}} \mid N_{\text{yesterday}}) = 0.2$$

Part (i)

Find the probability that it was raining or snowing yesterday given that it does not rain or snow today, $P(R_{\text{yesterday}} \mid N_{\text{today}})$.

The prior probabilities we were given are

$$P(R_{\text{yesterday}}) = 0.5$$

$$P(N_{\text{yesterday}}) = 0.5$$

Using Bayes' Theorem, we will have

$$P(R_{\text{yesterday}} \mid N_{\text{today}}) = \frac{P(N_{\text{today}} \mid R_{\text{yesterday}}) \cdot P(R_{\text{yesterday}})}{P(N_{\text{today}})}$$

Where:

$$P(N_{\text{today}}) = P(N_{\text{today}} \mid R_{\text{yesterday}}) \cdot P(R_{\text{yesterday}}) + P(N_{\text{today}} \mid N_{\text{yesterday}}) \cdot P(N_{\text{yesterday}})$$

The probability that it won't rain today is

$$P(N_{\text{today}}) = (0.4 \times 0.5) + (0.8 \times 0.5) = 0.6$$

And applying Bayes' Theorem

$$P(R_{\text{yesterday}} \mid N_{\text{today}}) = \frac{0.4 \times 0.5}{0.6} = \frac{0.2}{0.6} \approx 0.3333$$

Answer to Part (i)

The probability that it was raining or snowing yesterday given that it does not rain or snow today is approximately 33.33%.

Part (ii)

Find the probability that it will rain or snow on any given day in the long run (i.e., the stationary probability), assuming the weather follows the same pattern day after day indefinitely.

For starters, let

$$\pi_R = \text{Long-term probability that it rains or snows on any given day.}$$

$$\pi_N = \text{Long-term probability that it does not rain or snow on any given day.}$$

Since there are only two states:

$$\pi_N = 1 - \pi_R$$

At stationarity, the probabilities remain unchanged after a transition. Therefore

$$\pi_R = P(R_{\text{today}} \mid R_{\text{yesterday}}) \cdot \pi_R + P(R_{\text{today}} \mid N_{\text{yesterday}}) \cdot \pi_N$$

Substituting $\pi_N = 1 - \pi_R$

$$\pi_R = 0.6\pi_R + 0.2(1 - \pi_R)$$

Solving for π_R

$$\begin{aligned}
\pi_R &= 0.6\pi_R + 0.2 - 0.2\pi_R \\
\pi_R - 0.6\pi_R + 0.2\pi_R &= 0.2 \\
(1 - 0.6 + 0.2)\pi_R &= 0.2 \\
0.6\pi_R &= 0.2 \\
\pi_R &= \frac{0.2}{0.6} = \frac{1}{3} \approx 0.333
\end{aligned}$$

Answer to Part (ii)

The stationary probability that it will rain or snow on any given day is approximately 33.33%.

Part (iii)

Use the result from part (ii) above as a new prior probability of rain/snow yesterday and recompute the probability that it was raining/snowing yesterday given that it does not rain or snow today, $P(R_{\text{yesterday}} \mid N_{\text{today}})$.

Previously, in part (i), we used a prior probability of $P(R_{\text{yesterday}}) = 0.5$. Now, we use the stationary probability $\pi_R = \frac{1}{3}$ as the new prior.

$$P(R_{\text{yesterday}}) = \frac{1}{3}$$

$$P(N_{\text{yesterday}}) = 1 - \frac{1}{3} = \frac{2}{3}$$

Applying Bayes' Theorem

$$P(R_{\text{yesterday}} \mid N_{\text{today}}) = \frac{P(N_{\text{today}} \mid R_{\text{yesterday}}) \cdot P(R_{\text{yesterday}})}{P(N_{\text{today}})}$$

Where

$$P(N_{\text{today}}) = P(N_{\text{today}} \mid R_{\text{yesterday}}) \cdot P(R_{\text{yesterday}}) + P(N_{\text{today}} \mid N_{\text{yesterday}}) \cdot P(N_{\text{yesterday}})$$

The probability that it won't rain today is

$$P(N_{\text{today}}) = (0.4 \times \frac{1}{3}) + (0.8 \times \frac{2}{3}) = \frac{0.4}{3} + \frac{1.6}{3} = \frac{2}{3}$$

Applying Bayes' Theorem

$$P(R_{\text{yesterday}} \mid N_{\text{today}}) = \frac{\frac{0.4}{3}}{\frac{2}{3}} = \frac{0.4}{2} = 0.2$$

Answer to Part (iii)

When using the stationary probability as the prior, the probability that it was raining or snowing yesterday given that it does not rain or snow today is 20%.

Summary of Answers

(i) $P(R_{\text{yesterday}} \mid N_{\text{today}}) = 33.33\%$.

(ii) The long-term (stationary) probability of rain or snow on any day is 33.33%.

(iii) Using the stationary probability as the new prior, $P(R_{\text{yesterday}} \mid N_{\text{today}}) = 20\%$.

3 Exercise 3

Prove that the Beta distribution

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

where $\mu \in [0, 1]$ and

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1}\, dx$$

is correctly normalized, i.e.,

$$\int_0^1 \text{Beta}(\mu; a, b)\, d\mu = 1 \qquad \Longleftrightarrow \qquad \int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Solution

Let's prove the statement above both ways:

$\Rightarrow$

Given the definition of $\mathrm{Beta}(\mu; a, b)$:

$$\begin{aligned}\int_0^1 \mathrm{Beta}(\mu; a, b)\, d\mu &= \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}\, d\mu = 1 \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu = 1\end{aligned}$$

Therefore:

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

$\square$

$\Leftarrow$

The reverse implication is trivial:

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

$$\Rightarrow \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}\, d\mu = \int_0^1 \mathrm{Beta}(\mu; a, b)\, d\mu = 1$$

$\square$

The Beta distribution is thus correctly normalized.

4 Exercise 4

Consider a variable $x \in \{0, 1\}$ and $p(x = 1) = \mu$ representing the flipping of a coin. With 50% probability, we believe the coin is fair, i.e., $\mu = 0.5$, and with 50% probability, we believe it is unfair, i.e., $\mu \neq 0.5$. We encode this prior belief with the following prior:

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

(i) Assume that we get one observation $x_1 = 1$. What is the posterior $p(\mu \mid x_1)$? In particular, how does the belief in the fairness of the coin change under this observation?

(ii) Assume that we get one additional observation $x_2 = 1$. What is the posterior $p(\mu \mid x_1, x_2)$? In particular, how does the belief in the fairness of the coin change under this observation?

(iii) Compute the probability of the coin being fair by defining an event *fair* with the prior probability $p(\text{fair}) = 0.5$. Compute $p(\text{fair} \mid x_1, x_2)$ using Bayes' theorem based on the observations $x_1 = 1, x_2 = 1$.

Solution

Part (i)

We are given one observation $x_1 = 1$

The Likelihoods in the case we assume the coin toss is unfair and fair are:

$$p(x_1 = 1 \mid \text{Unfair}) = \int_0^1 \mu \cdot \text{Beta}(\mu; 1, 1)\, d\mu = \int_0^1 \mu \cdot 1\, d\mu = \frac{1}{2}$$

$$p(x_1 = 1 \mid \text{Fair}) = \frac{1}{2}$$

The Total Evidence is:

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

The Posterior Probabilities are:

$$p(\text{Unfair} \mid x_1 = 1) = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2}} = \frac{1}{2}$$

$$p(\text{Fair} \mid x_1 = 1) = \frac{0.5 \cdot \frac{1}{2}}{\frac{1}{2}} = \frac{1}{2}$$

The Posterior Distribution is:

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

where the $\text{Beta}(2, 1)$ distribution comes from updating the Beta prior with one success.

Answer to Part (i)

The posterior $p(\mu \mid x_1)$ is still:

50% belief that the coin is fair ($\mu = 0.5$).

50% belief that the coin is unfair, with μ following a $\text{Beta}(2, 1)$ distribution.

The belief in fairness remains unchanged at 50%.

Part (ii)

We are now given two observations $x_1 = 1$ and $x_2 = 1$.

After two successes, the posterior under the unfair hypothesis becomes $\text{Beta}(3, 1)$.

The Likelihoods in the case we assume the coin toss is unfair and fair are:

$$p(x_2 = 1 \mid \text{Unfair}, x_1 = 1) = \mathbb{E}[\mu \mid \text{Beta}(2, 1)] = \frac{2}{2+1} = \frac{2}{3}$$

$$p(x_2 = 1 \mid \text{Fair}, x_1 = 1) = 0.5$$

The Total Evidence is:

$$p(x_2 = 1 \mid x_1 = 1) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot 0.5 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12} \approx 0.583$$

The Posterior Probabilities are:

$$p(\text{Unfair} \mid x_1, x_2) = \frac{\frac{2}{3} \cdot \frac{1}{2}}{\frac{7}{12}} = \frac{4}{7} \approx 0.571$$

$$p(\text{Fair} \mid x_1, x_2) = \frac{0.5 \cdot \frac{1}{2}}{\frac{7}{12}} = \frac{3}{7} \approx 0.428$$

The Posterior Distribution is:

$$p(\mu \mid x_1, x_2) = \frac{4}{7} \cdot \text{Beta}(\mu; 3, 1) + \frac{3}{7} \cdot \delta(\mu - 0.5)$$

Answer to Part (ii)

The posterior $p(\mu \mid x_1 = 1, x_2 = 1)$ is:

57.14% belief that the coin is unfair (μ follows $\text{Beta}(3, 1)$).

42.86% belief that the coin is fair ($\mu = 0.5$).

The belief in fairness decreases from 50% to approximately 42.86%.

Part (iii): Probability of the Coin Being Fair $p(\textbf{Fair} \mid x_1, x_2)$

With the two observations $x_1 = 1$ and $x_2 = 1$, from Part (ii), we have already calculated:

$$p(\text{Fair} \mid x_1, x_2) = \frac{3}{7}$$

Answer to Part (iii)

As seen in Part (ii), the probability that the coin is fair after observing two heads is $\frac{3}{7}$, which is approximately 42.86%.

Now, using Bayes' Theorem to explicit this result:

$$p(\text{Fair} \mid x_1 = 1, x_2 = 1) = \frac{p(x_1 = 1, x_2 = 1 \mid \text{Fair}) \cdot p(\text{Fair})}{p(x_1 = 1, x_2 = 1)}$$

Calculating each term:

$$p(x_1 = 1, x_2 = 1 \mid \text{Fair}) = 0.5 \times 0.5 = 0.25$$

$$p(x_1 = 1, x_2 = 1 \mid \text{Unfair}) = \int_0^1 \mu^2 \cdot p(\mu \mid \text{Unfair})\, d\mu = \int_0^1 \mu^2\, d\mu = \left[\frac{\mu^3}{3}\right]_0^1 = \frac{1}{3}$$

The above results from that if the coin is unfair, then $\mu \sim \text{Beta}(1, 1)$, a uniform distribution on $[0, 1]$.

The marginal likelihood $p(x_1 = 1, x_2 = 1)$ is given by:

$$p(x_1 = 1, x_2 = 1) = p(x_1 = 1, x_2 = 1 \mid \text{Fair}) \cdot p(\text{Fair}) + p(x_1 = 1, x_2 = 1 \mid \text{Unfair}) \cdot p(\text{Unfair})$$

$$p(x_1 = 1, x_2 = 1) = 0.25 \cdot 0.5 + \frac{1}{3} \cdot 0.5 = 0.125 + \frac{1}{6} = \frac{7}{24} \approx 0.2917$$

Applying Bayes' Theorem:

$$p(\text{Fair} \mid x_1 = 1, x_2 = 1) = \frac{p(x_1 = 1, x_2 = 1 \mid \text{Fair}) \cdot p(\text{Fair})}{p(x_1 = 1, x_2 = 1)} = \frac{0.25 \times 0.5}{0.2917} \approx 0.4286$$

Summary of Answers

(i) After observing one head, the posterior distribution is a 50-50 mixture of $\text{Beta}(2, 1)$ and a point mass at $\mu = 0.5$. Thus, the belief that the coin is fair ($\mu = 0.5$) remains at 50%.

(ii) After observing a second head, the posterior becomes a mixture of $\text{Beta}(3, 1)$ with weight $\frac{4}{7}$ and a point mass at $\mu = 0.5$ with weight $\frac{3}{7}$. Therefore, the belief that the coin is fair decreases to 42.86%.

(iii) The probability that the coin is fair after two heads is 42.86%.

5 Exercise 5

Theoretical Questions

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

Given a dataset $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, where each x_i is an independent observation from a Poisson distribution with rate λ, the likelihood function $L(\lambda|\mathbf{x})$ represents the probability of observing the data given λ.

Since the observations are independent, the joint likelihood is the product of the individual probabilities:

$$
\begin{aligned}
L(\lambda|\mathbf{x}) &= \prod_{i=1}^{n} P(X_i = x_i|\lambda) \\
&= \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} \\
&= e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \frac{1}{x_i!}
\end{aligned}
$$

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

The Probability Density Function (PDF) of a Gamma-distributed random variable λ with shape parameter α and rate parameter β is:

$$
f(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}, \quad \lambda > 0
$$

where $\Gamma(\alpha)$ is the Gamma function, which generalizes the factorial function.

Bayes' Theorem allows us to update our prior beliefs in light of new evidence. It is given by:

$$
P(\lambda|\mathbf{x}) = \frac{P(\mathbf{x}|\lambda)P(\lambda)}{P(\mathbf{x})}
$$

where:

- $P(\lambda|\mathbf{x})$ is the posterior distribution of λ given the data $\mathbf{x}$.
- $P(\mathbf{x}|\lambda)$ is the likelihood of the data given λ.
- $P(\lambda)$ is the prior distribution of λ.

- $P(\mathbf{x})$ is the marginal likelihood or evidence, which acts as a normalizing constant.

Using the expressions for the likelihood and the prior derived before:

$$\begin{aligned} P(\lambda|\mathbf{x}) &\propto P(\mathbf{x}|\lambda)P(\lambda) \\ &= \left(e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}\prod \frac{1}{x_i!}\right)\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\right) \\ &\propto \lambda^{\sum x_i}e^{-n\lambda}\lambda^{\alpha-1}e^{-\beta\lambda} \\ &= \lambda^{\alpha+\sum x_i-1}e^{-(\beta+n)\lambda} \end{aligned}$$

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

The unnormalized posterior distribution derived above:

$$P(\lambda|\mathbf{x}) \propto \lambda^{\alpha+\sum x_i-1}e^{-(\beta+n)\lambda}$$

matches the form of a Gamma distribution's PDF:

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

where α' and β' are the updated shape and rate parameters, respectively.

By comparing the exponents and coefficients, we can identify the updated (posterior) parameters:

$$\text{Shape Parameter:} \quad \alpha' = \alpha + \sum_{i=1}^{n} x_i$$

$$\text{Rate Parameter:} \quad \beta' = \beta + n$$

Thus, the posterior distribution is:

$$\lambda|\mathbf{x} \sim \text{Gamma}(\alpha', \beta') = \text{Gamma}\left(\alpha + \sum_{i=1}^{n} x_i, \beta + n\right)$$

To confirm that the posterior is a Gamma distribution, we can derive the normalizing constant. The Gamma distribution's normalization ensures that the integral of the PDF over all possible values of λ equals 1.

Given the unnormalized posterior:

$$P(\lambda|\mathbf{x}) = \lambda^{\alpha'-1}e^{-\beta'\lambda}$$

the normalizing constant is:

$$\text{Normalization Constant} = \frac{(\beta')^{\alpha'}}{\Gamma(\alpha')}$$

Thus, the normalized posterior PDF is:

$$P(\lambda|\mathbf{x}) = \frac{(\beta')^{\alpha'}}{\Gamma(\alpha')}\lambda^{\alpha'-1}e^{-\beta'\lambda}$$

which is the PDF of a Gamma distribution with parameters α' and β'.

Programming Simulation

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

```
import numpy as np

# Set seed for reproducibility
np.random.seed(42)

# True rate parameter
lambda_true = 5

# Number of observations
n = 50

# Generate synthetic data from a Poisson distribution
synthetic_data = np.random.poisson(lam=lambda_true, size=n)

print("Synthetic Data:", synthetic_data)
print("Total Counts:", synthetic_data.sum())
```

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.

```
# Prior parameters
alpha_prior = 2
beta_prior = 2

print(f"Prior Gamma Parameters: alpha = {alpha_prior}, beta =
    {beta_prior}")

# Sum of observed counts
sum_x = synthetic_data.sum()

# Update posterior parameters
alpha_post = alpha_prior + sum_x
```

```python
beta_post = beta_prior + n

print(f"Posterior Gamma Parameters: alpha = {alpha_post}, beta =
    {beta_post}")
```

Write a Python function that, given data and prior parameters α and β, calculates the updated posterior parameters for λ.

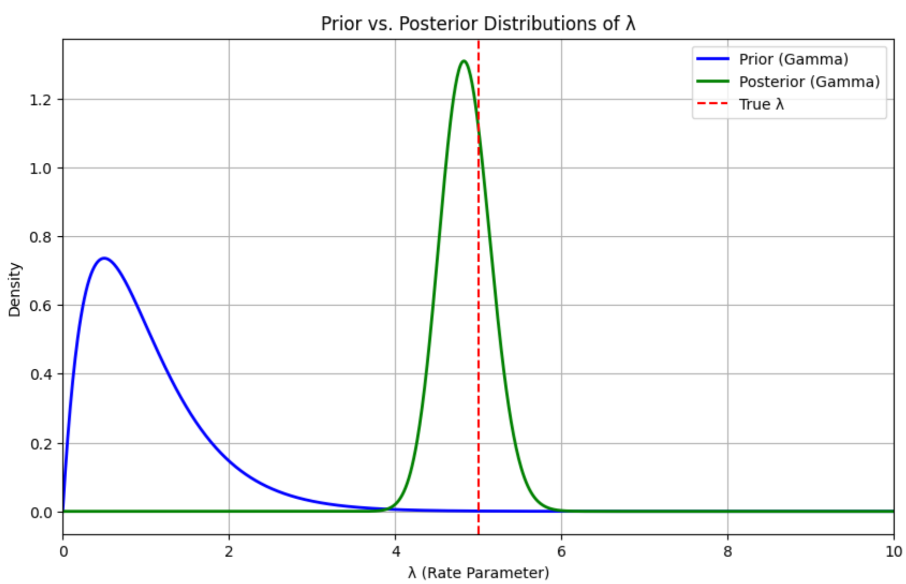
```python
def update_posterior_gamma_poisson(data, alpha_prior, beta_prior):

# Ensure data is a NumPy array for efficient computation
data = np.array(data)

# Number of observations
n = len(data)

# Sum of observed counts
sum_x = data.sum()

# Update posterior parameters
alpha_post = alpha_prior + sum_x
beta_post = beta_prior + n

return alpha_post, beta_post
```

Visualization

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

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import gamma

# Define the prior parameters
alpha_prior = 2
beta_prior = 2

# Define the posterior parameters from previous simulation
alpha_post = 252
beta_post = 52

# True lambda used to generate data
lambda_true = 5

# Generate a range of lambda values
lambda_values = np.linspace(0, 10, 1000)

# Compute the prior and posterior PDFs
```

```python
prior_pdf = gamma.pdf(lambda_values, a=alpha_prior, scale=1/beta_prior)
posterior_pdf = gamma.pdf(lambda_values, a=alpha_post, scale=1/beta_post)

# Plotting
plt.figure(figsize=(10, 6))
plt.plot(lambda_values, prior_pdf, label='Prior (Gamma)', color='blue', lw=2)
plt.plot(lambda_values, posterior_pdf, label='Posterior (Gamma)', color='green', lw=2)
plt.axvline(lambda_true, color='red', linestyle='--', label='True ')
plt.title('Prior vs. Posterior Distributions of  ')
plt.xlabel('  (Rate Parameter)')
plt.ylabel('Density')
plt.xlim(0, 10)
plt.legend()
plt.grid(True)
plt.show()
```

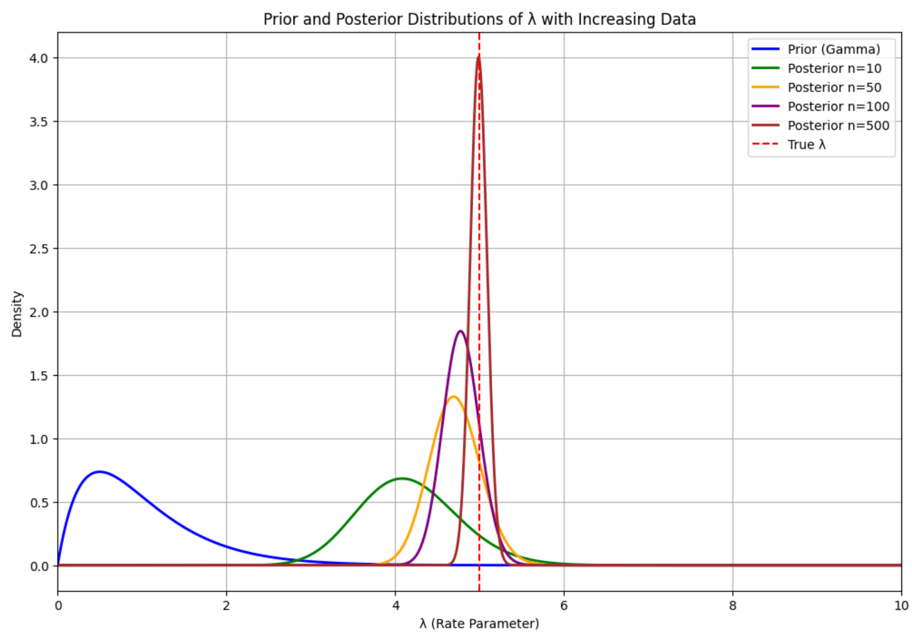


Compare the posterior distribution to the true value of λ_{true} used to generate the data. As you increase the number of observations, observe how the posterior distribution becomes more concentrated around λ_{true}.

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import gamma
```

```python

def update_posterior_gamma_poisson(data, alpha_prior, beta_prior):
    """
    Updates the posterior Gamma distribution parameters for
        given Poisson-distributed data.
    """
    data = np.array(data)
    n = len(data)
    sum_x = data.sum()
    alpha_post = alpha_prior + sum_x
    beta_post = beta_prior + n
    return alpha_post, beta_post

# Set seed for reproducibility
np.random.seed(42)

# True rate parameter
lambda_true = 5

# Prior parameters
alpha_prior = 2
beta_prior = 2

# Define sample sizes to simulate
sample_sizes = [10, 50, 100, 500]

# Generate a range of lambda values for plotting
lambda_values = np.linspace(0, 10, 1000)

# Compute the prior PDF
prior_pdf = gamma.pdf(lambda_values, a=alpha_prior,
    scale=1/beta_prior)

# Initialize the plot
plt.figure(figsize=(12, 8))
plt.plot(lambda_values, prior_pdf, label='Prior (Gamma)',
    color='blue', lw=2)

# Colors for different sample sizes
colors = ['green', 'orange', 'purple', 'brown']

# Iterate over different sample sizes
for idx, n in enumerate(sample_sizes):
    # Generate synthetic data
    synthetic_data = np.random.poisson(lam=lambda_true, size=n)

    # Update posterior parameters
    alpha_post, beta_post =
        update_posterior_gamma_poisson(synthetic_data,
        alpha_prior, beta_prior)

    # Compute posterior PDF
    posterior_pdf = gamma.pdf(lambda_values, a=alpha_post,
        scale=1/beta_post)

    # Plot posterior
```

```
    plt.plot(lambda_values, posterior_pdf, label=f'Posterior
        n={n}', color=colors[idx], lw=2)

# Plot the true lambda
plt.axvline(lambda_true, color='red', linestyle='--', label='True
      ')

# Final plot adjustments
plt.title('Prior and Posterior Distributions of     with Increasing
    Data')
plt.xlabel('    (Rate Parameter)')
plt.ylabel('Density')
plt.xlim(0, 10)
plt.legend()
plt.grid(True)
plt.show()
```



Analysis

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

This effect can be understood in terms of the parameters of the Gamma distribution, which governs both the prior and posterior distributions.

When we assume a Gamma prior for λ and observe data points from a Poisson distribution with rate parameter λ, the posterior distribution of λ also follows a Gamma distribution. The posterior parameters are updated based on the prior parameters and the observed data, as seen above.

Posterior Shape and Scale Parameters

The posterior shape parameter, α_{post}, increases with the sum of observed data points. As more data points are observed, α_{post} grows, which increases the concentration of the posterior distribution. Similarly, the posterior rate parameter, β_{post}, also adjusts based on the sum of the observed data.

As we accumulate more observations, this sum grows proportionally to the true λ, increasing β_{post} and thus sharpening the posterior distribution around the mean $\frac{\alpha_{\text{post}}}{\beta_{\text{post}}}$.

Increased Concentration

As the sample size increases, the posterior distribution becomes more concentrated around the true value of λ. This happens because each new observation provides additional evidence about λ, allowing the Bayesian model to "learn" the true parameter more precisely.

Mathematically, as n (the number of observations) grows, both α_{post} and β_{post} increase, resulting in a posterior distribution with a higher peak and less spread. The variance of a Gamma distribution $\text{Var}(\lambda) = \frac{\alpha_{\text{post}}}{\beta_{\text{post}}^2}$ decreases as β_{post} increases, leading to a tighter distribution around the mean.

Bayesian Convergence

With enough data, the posterior distribution will converge to a very narrow distribution centered around the true λ. This reflects the certainty gained by observing a large number of samples, as the model increasingly aligns with the true parameter value. The posterior becomes increasingly dominated by the observed data rather than the prior as the sample size grows.