Problem Set 1 Solutions

JOSEPH AGOR

February 13, 2025

QUESTION ONE

1 Gamma Prior for Poisson Distribution

The Poisson likelihood function for a sample $y_1, ..., y_n$ is:

$$L(\theta|y_1, ..., y_n) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!}$$

$$= \theta^{\sum y_i} e^{-n\theta} \prod_{i=1}^{n} \frac{1}{y_i!}$$

The prior distribution is assumed to be Gamma:

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

The posterior is proportional to:

$$p(\theta|y_1, ..., y_n) \propto L(\theta|y_1, ..., y_n) p(\theta)$$

$$\propto \theta^{\sum y_i} e^{-n\theta} \times \theta^{\alpha-1} e^{-\beta\theta}$$

$$\propto \theta^{(\alpha+\sum y_i)-1} e^{-(\beta+n)\theta}$$

Since this matches the form of a Gamma distribution:

$$\theta|y_1, ..., y_n \sim G(\alpha + \sum y_i, \beta + n)$$

Thus, the Gamma distribution is a conjugate prior for the Poisson likelihood.

2 Maximum Likelihood Estimation (MLE)

The log-likelihood function for the Poisson distribution is:

$$\log L(\theta) = \sum_{i=1}^{n} [y_i \log \theta - \theta - \log(y_i!)]$$

Taking the derivative with respect to θ and setting it to zero:

$$\frac{d}{d\theta} \sum y_i \log \theta - n\theta = \frac{\sum y_i}{\theta} - n = 0$$

Solving for θ:

$$\hat{\theta} = \frac{\sum y_i}{n} = \bar{y}$$

Thus, the MLE for θ is $\bar{y}$.

3 Posterior Mean as a Weighted Average

From the conjugate prior result, the posterior mean is:

$$E[\theta|y] = \frac{\alpha + \sum y_i}{\beta + n}$$

Rewriting it:

$$E[\theta|y] = \frac{\alpha}{\beta + n} + \frac{\sum y_i}{\beta + n}$$

$$= \left(\frac{\beta}{\beta + n}\right) \frac{\alpha}{\beta} + \left(\frac{n}{\beta + n}\right) \bar{y}$$

This expresses the posterior mean as a weighted average of the prior mean $\frac{\alpha}{\beta}$ and the MLE $\bar{y}$.

4 Effect of Increasing Sample Size

As $n \to \infty$, the weight on the prior mean $\frac{\beta}{\beta+n}$ approaches 0, while the weight on the MLE $\frac{n}{\beta+n}$ approaches 1. This means the influence of the prior diminishes as the sample size increases, making the posterior distribution more dependent on the observed data

QUESTION TWO

5 Posterior Distribution for θ_1

The Dirichlet prior is:

$$(\theta_1, ..., \theta_6) \sim \text{Dir}(2, 2, 2, 2, 2, 2)$$

Using the data from the table:

5.1 For $n = 100$

$$\theta_1|\text{data} \sim \text{Beta}(19 + 2, 100 - 19 + 10)$$

$$= \text{Beta}(21, 91)$$

5.2 For $n = 1000$

$$\theta_1|\text{data} \sim \text{Beta}(190 + 2, 1000 - 190 + 10)$$

$$= \text{Beta}(192, 820)$$

6 Comparison of Posterior Distributions

The figure below shows the posterior distributions for θ_1 using both sample sizes.

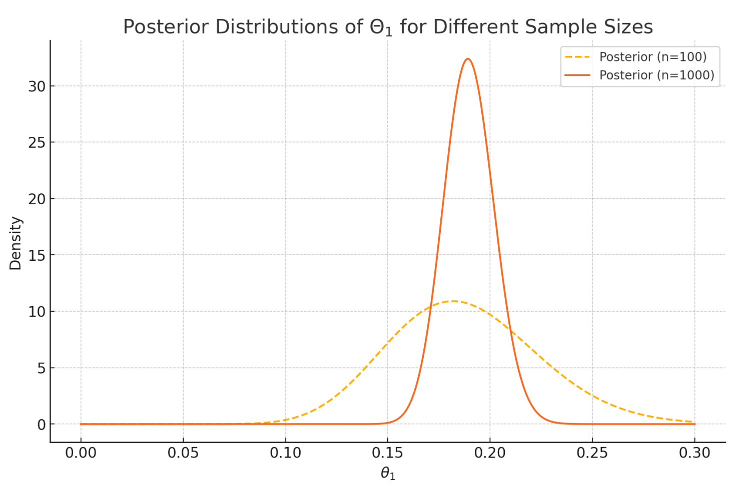

Figure 1: Posterior distributions for different sample sizes

Observations:

- The posterior distribution with $n = 100$ is wider, indicating more uncertainty.
- The posterior with $n = 1000$ is more concentrated around its mean, showing that a larger sample size reduces uncertainty.
- As n increases, the influence of the prior diminishes, and the posterior is more dominated by the data.