Discussion Problems - Week 2

Instructor: Dr. Paul Parker
TA: Antonio Aguirre

University of California, Santa Cruz
Winter 2025

Problem 1

Suppose that the prior distribution of some parameter θ is a Beta distribution for which the mean is $\frac{1}{3}$ and the variance is $\frac{1}{45}$. Determine the prior pdf of θ.

Solution

We are given that the prior distribution of a parameter θ follows a Beta distribution with:

- **Mean:** $\mathbb{E}[\theta] = \frac{1}{3}$
- **Variance:** $\text{Var}(\theta) = \frac{1}{45}$

We aim to determine the prior probability density function (pdf) of θ.

Step 1: Beta Distribution Properties

A Beta distribution $\theta \sim \text{Beta}(\alpha, \beta)$ has the pdf:

$$f(\theta \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad 0 \le \theta \le 1 \tag{1}$$

where $\alpha, \beta > 0$ are shape parameters. The mean and variance of $\text{Beta}(\alpha, \beta)$ are:

$$\mathbb{E}[\theta] = \frac{\alpha}{\alpha+\beta} \tag{2}$$

$$\text{Var}(\theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{3}$$

Step 2: Solve for α and β

From the mean equation:

$$\frac{\alpha}{\alpha+\beta} = \frac{1}{3} \Rightarrow \beta = 2\alpha \tag{4}$$

Substituting into the variance equation:

$$\frac{\alpha(2\alpha)}{(3\alpha)^2(3\alpha+1)} = \frac{1}{45} \tag{5}$$

Multiplying both sides by $9\alpha^2(3\alpha+1)$:

$$2\alpha^2 = \frac{9\alpha^2(3\alpha+1)}{45} \tag{6}$$

Simplifying:

$$81\alpha^2 = 27\alpha^3 \Rightarrow \alpha = 3, \quad \beta = 6 \tag{7}$$

Step 3: Write the Prior PDF

Since $\theta \sim \text{Beta}(3, 6)$, the pdf is:

$$f(\theta) = \frac{\Gamma(9)}{\Gamma(3)\Gamma(6)}\theta^2(1-\theta)^5 \tag{8}$$

Using $\Gamma(n) = (n-1)!$:

$$\frac{8!}{2!5!} = 168 \tag{9}$$

Thus, the final prior pdf is:

Prior PDF of θ

$$f(\theta) = 168\theta^2(1-\theta)^5, \quad 0 \leq \theta \leq 1 \tag{10}$$

Final Answer

Final Answer

$$f(\theta) = 168\theta^2(1-\theta)^5, \quad 0 \leq \theta \leq 1 \tag{11}$$

This is the required prior pdf for θ.

Problem 2

Suppose that the time a student spends studying each week follows an exponential distribution with rate parameter λ. After randomly sampling 3 students, you found that they studied 2, 2.5, and 3 hours respectively. Write out the likelihood function for λ, and then draw a sketch of this function.

Solution

We are given that the time a student spends studying per week follows an exponential distribution with rate parameter λ. That is,

$$X_i \sim \text{Exp}(\lambda), \quad i = 1, 2, 3. \tag{12}$$

The probability density function (pdf) of an exponential distribution is:

$$f(x \mid \lambda) = \lambda e^{-\lambda x}, \quad x > 0. \tag{13}$$

Given that we observe three independent study times $x_1 = 2$, $x_2 = 2.5$, and $x_3 = 3$, the likelihood function is given by:

$$L(\lambda) = \prod_{i=1}^{3} f(x_i \mid \lambda) = \lambda^3 e^{-\lambda(x_1+x_2+x_3)}. \tag{14}$$

Substituting the observed values:

$$L(\lambda) = \lambda^3 e^{-\lambda(2+2.5+3)} = \lambda^3 e^{-7.5\lambda}, \quad \lambda > 0. \tag{15}$$

Final Answer

Likelihood Function

$$L(\lambda) = \lambda^3 e^{-7.5\lambda}, \quad \lambda > 0. \tag{16}$$

This is the required likelihood function for λ. We now proceed to sketch this function.

Likelihood Function Plot

To visualize the likelihood function, we include the following plot:

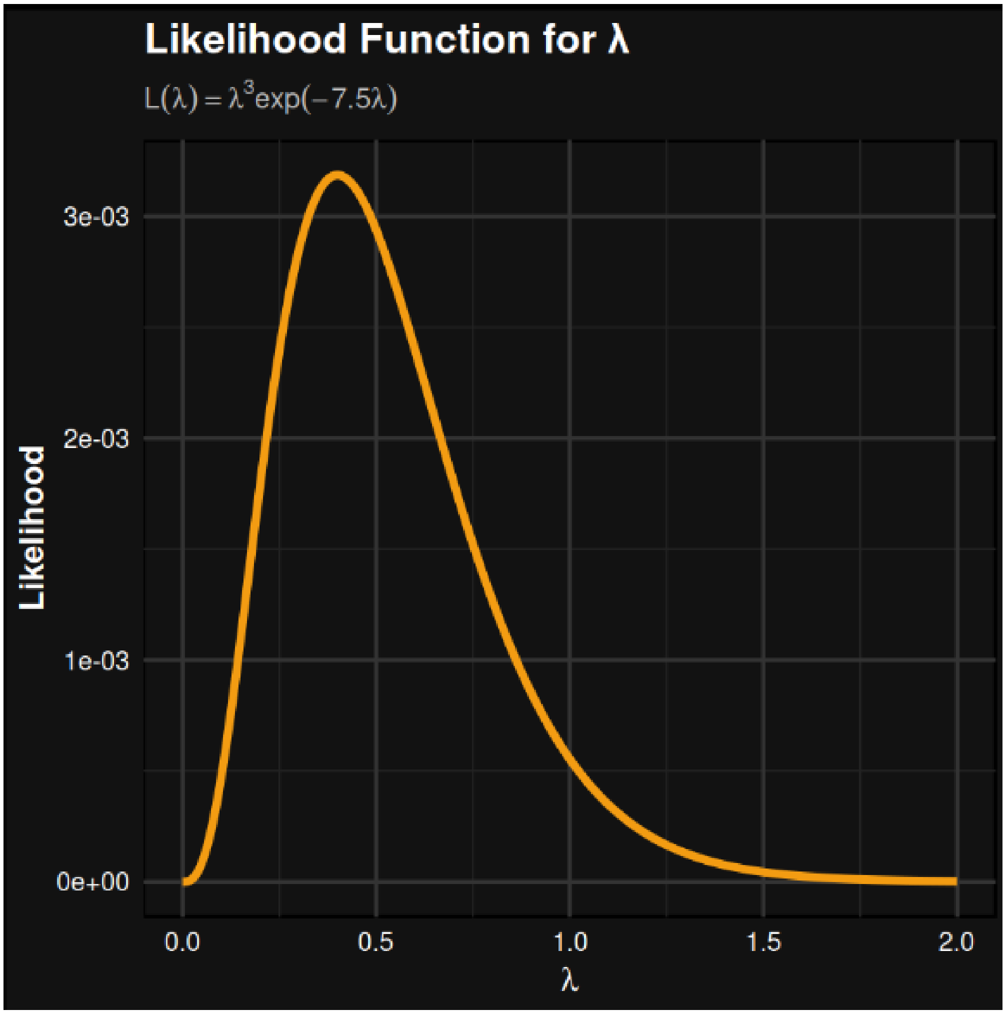

Figure 1: Plot of the likelihood function $L(\lambda) = \lambda^3 e^{-7.5\lambda}$.

.

Problem 3

Suppose that the time a student spends studying each week follows an exponential distribution with rate parameter λ. Your friend suggests a uniform distribution over the interval $[0, 10]$ for λ. Is this a reasonable prior distribution for this problem? Why or why not?

Solution

To evaluate whether a uniform prior on $[0, 10]$ is reasonable for the rate parameter λ, we need to consider:

- **The Nature of λ in an Exponential Distribution:**

  The rate parameter λ in an exponential distribution must be strictly positive ($\lambda > 0$), ensuring that the density function remains well-defined. The choice of a uniform prior on $[0, 10]$ satisfies this constraint.

- **Informative vs. Non-informative Priors:**
  - A uniform prior assumes that all values of λ in the range $[0, 10]$ are equally likely a priori.
  - This is a weakly informative prior because it restricts λ to a finite interval rather than allowing it to take values over a wider, possibly more realistic domain (e.g., an unbounded prior like a Gamma or Jeffreys prior).

- **Potential Issues with a Bounded Prior:**
  - **Artificially Constraining the Parameter:** The prior imposes a strict upper bound of $\lambda = 10$, meaning that any value greater than 10 is considered impossible a priori, even if the data suggest otherwise. This could introduce bias and distort inference.
  - **Edge Effects in Bayesian Updating:** If posterior mass accumulates near the boundary (e.g., near 10), then the inference may be overly influenced by this arbitrary choice of bound, leading to misleading results.
  - **Lack of Heavy-Tailed Behavior:** Many real-world scenarios involve small but nonzero probabilities of very high λ values. A uniform prior does not allow for such behavior, unlike an exponential or gamma prior, which can model such possibilities.

- **Missing Information and Contextual Factors:**
  - The choice of prior should depend on domain knowledge. If past studies suggest a plausible range for λ, a more structured prior (such as a gamma prior) could be preferable.

- If λ represents the inverse of an expected study duration, practical constraints (such as observed study times being rarely less than a certain amount) could justify an alternative prior form.
- A proper Bayesian approach should consider prior predictive checks—how well the prior aligns with plausible observed data.

Final Answer

Conclusion

A uniform prior on $[0, 10]$ is weakly informative but may not be ideal due to its artificial constraint and potential boundary issues. A more flexible alternative, such as a gamma prior, could provide better modeling, especially if prior knowledge or empirical data suggests a broader or more structured distribution for λ. The appropriateness of the uniform prior ultimately depends on additional context, including past empirical knowledge and the range of plausible values for λ in real-world scenarios.

Problem 4

Suppose that the proportion p of defective items in a large manufactured lot is known to be either 0.1 or 0.2. Come up with a reasonable prior distribution for this problem, and then write out $\pi(p)$.

Solution

To specify a reasonable prior distribution for p, we consider three types of priors: a vague prior, a weakly informative prior, and an informative prior.

- **Vague Prior:** A vague prior represents minimal prior knowledge about p. Since p is known to take values in $\{0.1, 0.2\}$, a reasonable vague prior is a discrete uniform prior:

$$\pi(p) = \begin{cases} 0.5, & p = 0.1 \\ 0.5, & p = 0.2 \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

> This prior assumes that before observing any data, we believe both values of p are equally likely.

- **Weakly Informative Prior:** A weakly informative prior incorporates slight preference for one of the values based on limited prior knowledge. For instance, if past data suggest that defects are slightly more common at $p = 0.1$, we can assign a higher probability to it:

$$\pi(p) = \begin{cases} 0.7, & p = 0.1 \\ 0.3, & p = 0.2 \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

> This prior reflects a mild belief that the lower defect rate is more common but still allows for uncertainty.

- **Informative Prior:** An informative prior strongly reflects prior knowledge or expert opinion. Suppose historical defect rates from a similar manufacturing process indicate that $p = 0.1$ occurs 90% of the time, while $p = 0.2$ occurs only 10%. Then, the prior is:

$$\pi(p) = \begin{cases} 0.9, & p = 0.1 \\ 0.1, & p = 0.2 \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

> This prior heavily favors $p = 0.1$ and would have a strong influence on the posterior unless overwhelming data suggest otherwise.

- **Beta Prior:** If we were to use a Beta prior, a natural choice might be $p \sim \text{Beta}(\alpha, \beta)$. For example, setting $\alpha = 2$ and $\beta = 8$ gives a prior with mean $\mathbb{E}[p] = 0.2$, suggesting a belief that defects are generally low but not nonexistent:

$$\pi(p) \propto p^{\alpha-1}(1-p)^{\beta-1}, \quad 0 \leq p \leq 1. \tag{20}$$

> However, using a Beta prior in this case is not ideal because it does not incorporate the critical information that p can only take on the values 0.1 or 0.2. Instead, it allows p to take any value in $[0, 1]$, which does not reflect the problem's structure. This could lead to incorrect inference if posterior updates assign probability mass to values of p that are not actually possible.

When to Use Each Prior

- **Vague Prior:** Use when there is no past knowledge or when objectivity is required in an analysis.
- **Weakly Informative Prior:** Use when some evidence suggests one value is more likely but prior knowledge is weak.
- **Informative Prior:** Use when strong empirical or expert knowledge supports a particular distribution of p.
- **Beta Prior:** Generally not recommended here since it does not respect the known discrete nature of p.

Final Answer

Conclusion

Depending on the level of prior knowledge available, a vague, weakly informative, or informative prior can be selected. A Beta prior, while commonly used in Bayesian modeling, is not appropriate in this case because it does not incorporate the fact that p is known to take only two specific values.

Problem 5

Suppose that the proportion p of defective items in a large manufactured lot is unknown. The following prior distribution is used:

$$\pi(p) = \begin{cases} 2(1-p), & 0 \leq p \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

What is the prior mean and variance for p?

Solution

To compute the prior mean and variance of p, we use the definitions of expectation and variance.

Step 1: Compute the Prior Mean

The mean of p is given by:

$$\mathbb{E}[p] = \int_0^1 p\pi(p)dp. \tag{22}$$

Substituting $\pi(p) = 2(1-p)$:

$$\mathbb{E}[p] = \int_0^1 p \cdot 2(1-p)dp. \tag{23}$$

Expanding the integral:

$$\mathbb{E}[p] = 2\int_0^1 (p - p^2)dp. \tag{24}$$

Evaluating each term:

$$\int_0^1 pdp = \frac{1}{2}, \quad \int_0^1 p^2 dp = \frac{1}{3}. \tag{25}$$

Thus,

$$\mathbb{E}[p] = 2\left(\frac{1}{2} - \frac{1}{3}\right) = 2 \times \frac{1}{6} = \frac{1}{3}. \tag{26}$$

Step 2: Compute the Prior Variance

The variance of p is given by:

$$\text{Var}(p) = \mathbb{E}[p^2] - (\mathbb{E}[p])^2. \tag{27}$$

First, compute $\mathbb{E}[p^2]$:

$$\mathbb{E}[p^2] = \int_0^1 p^2\pi(p)dp. \tag{28}$$

Substituting $\pi(p) = 2(1-p)$:

$$\mathbb{E}[p^2] = 2\int_0^1 p^2(1-p)dp. \tag{29}$$

Expanding the integral:

$$\mathbb{E}[p^2] = 2\left(\int_0^1 p^2 dp - \int_0^1 p^3 dp\right). \tag{30}$$

Using standard integrals:

$$\int_0^1 p^2 dp = \frac{1}{3}, \quad \int_0^1 p^3 dp = \frac{1}{4}. \tag{31}$$

Thus,

$$\mathbb{E}[p^2] = 2\left(\frac{1}{3} - \frac{1}{4}\right) = 2 \times \frac{1}{12} = \frac{1}{6}. \tag{32}$$

Now, compute the variance:

$$\text{Var}(p) = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{6} - \frac{1}{9}. \tag{33}$$

Finding a common denominator:

$$\text{Var}(p) = \frac{3}{18} - \frac{2}{18} = \frac{1}{18}. \tag{34}$$

Final Answer

Prior Mean and Variance

$$\mathbb{E}[p] = \frac{1}{3}, \quad \text{Var}(p) = \frac{1}{18}. \tag{35}$$

This completes the computation of the prior mean and variance for p.