Advanced Machine Learning

Assignment Problem Set I

Please solve the problems below and upload a pdf file of the solution by September 27th.

---

1. Suppose that $X_1, \ldots, X_n$ is a random sample from the distribution with pdf

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose also that the value of the parameter θ is unknown ($\theta > 0$) and that the prior distribution of θ is a gamma distribution with parameters α and β ($\alpha > 0$ and $\beta > 0$).

Determine the posterior distribution of θ

2. Suppose k has a geometric distribution with unknown success probability θ

$$\Pr(k \mid \theta) = (1-\theta)^{k-1}\theta, \qquad k = 1, 2, \ldots$$

The geometric distribution is appropriate for modeling the number of independent Bernoilli trials required, each with success probability θ, before observing the first "success."

Let the prior for θ be a beta distribution:

$$p(\theta \mid \alpha, \beta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)} \qquad 0 < \theta < 1$$

where B is the beta function. Show that, given an observation k, the posterior $p(\theta \mid k, \alpha, \beta)$ is a beta distribution with updated parameters (α', β').

3. Suppose that in the last question, we received a sample of n observations $\{k_1, k_2, \ldots, k_n\}$. What is the posterior $p(\theta \mid k_1, k_2, \ldots, k_n, \alpha, \beta)$? What is the posterior mean? The posterior mode?

In light of this and the previous question, can you give an interpretation of the prior parameters α and β? What happens in the limit as $n \to \infty$?

4. Assume $Y_1, Y_2, \ldots, Y_n$ are independent observations which have the normal distribution with mean βx_i and variance σ^2, where the x_is and σ^2 are known constants, and β is an unknown parameter, which has a normal prior distribution with mean β_0 and variance τ^2, where β_0 and τ^2 are known constants.

(a) Derive the posterior distribution of β.

(b) Show that the mean of the posterior distribution is a weighted average of the prior mean β_0, and the maximum likelihood estimator of β.

(c) Find the limit of the posterior distribution as $\tau^2 \to \infty$, and discuss the result.

(d) How would you predict a future observation from the population $N(\beta x_{n+1}, \sigma^2)$, where x_{n+1} is known?