Week 10 HW: Stats

Srikanth Pai, Asst. Prof, MSE.

Exercises

1. The formal definition of sufficient statistic is given in Definition 5.6.1 and the criterion that we have used in class is stated as Theorem 5.6.1. Using the theorem solve the following problems: L&M 5.6.1, 5.6.2,5.6.5, 5.6.6, 5.6.11.

2. The following exercises are practice for Bayesian estimation: L&M Examples 5.8.2, 5.8.1, 5.8.3, 5.8.6.

3. For each of the following models, begin with the naive estimator $\hat{\theta}_1 = X_1$:

   (a) $X_1, \dots, X_n \sim \text{Bernoulli}(p)$.

   (b) $X_1, \dots, X_n \sim \text{Poisson}(\lambda)$.

   (c) $X_1, \dots, X_n \sim \text{Uniform}(0, \theta)$.

   In each case:

   (i) Identify a sufficient statistic for the parameter.

   (ii) Compute the Rao–Blackwellised version $\hat{\theta}_2 = \mathbb{E}[\hat{\theta}_1 \mid T]$.

   (iii) Compare the bias and variance of $\hat{\theta}_1$ and $\hat{\theta}_2$.

4. Let $X_1, \dots, X_n \sim N(\mu, 1)$ i.i.d.

   (a) Show that $\hat{\mu}_1 = X_1 - X_2$ is an unbiased estimator of μ.

   (b) Show that $\bar{X}$ is sufficient for μ.

   (c) Compute the Rao–Blackwellised estimator of $\hat{\mu}_1$ with respect to $\bar{X}$, and compare its variance to $\hat{\mu}_1$.

(d) Prove that the Rao–Blackwellised version always has variance less than or equal to the original estimator.

5. Suppose $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$ i.i.d. and prior $\lambda \sim \text{Gamma}(\alpha, \beta)$. [Larsen and Marx example 5.8.3, 5.8.6 and Case Study 5.8.1 that estimates the annual number of Hurricanes.]

   (a) Derive the posterior distribution of λ.

   (b) Find the posterior mean and mode.

   (c) Compare the posterior mean with the maximum likelihood estimator of λ.

6. Suppose $X_1, \ldots, X_n \sim \text{Uniform}(0, \theta)$ i.i.d.

   (a) Write down the likelihood function of θ given the sample.

   (b) Let the prior distribution for θ be Pareto(α, θ_0), that is,

$$\pi(\theta) = \frac{\alpha\, \theta_0^\alpha}{\theta^{\alpha+1}}, \qquad \theta \geq \theta_0,\ \alpha > 0.$$

   Show that the posterior distribution is also Pareto, with updated parameters.

   (c) Interpret the role of $\max(X_1, \ldots, X_n)$ in this posterior distribution.

7. (German tank problem.) During the Second World War, the Allied forces needed to estimate the number of German tanks in production. Serial numbers on captured tanks provided indirect evidence. Suppose the serial numbers are modeled as

$$X_1, X_2, \ldots, X_n \sim \text{Uniform}\{1, 2, \ldots, \theta\},$$

   where θ is the maximum serial number and represents total production. Now I urge you to read the wiki article on German Tank problem.