Foundations of Statistical Inference, BS2a, Exercises 2

1. The number of phone calls a man receives in a week follows a Poisson distribution with mean θ. At the start of week 1, the man's opinion about the value of θ corresponds to the gamma distribution

$$\pi(\theta) = \frac{1}{54}\theta^2 e^{-\theta/3},\ \theta > 0.$$

In the 4 weeks following the start of week 1, the man received 3, 7, 6, and 10 phone calls, respectively. Determine the posterior distribution of θ and the predictive distribution for the number of calls that he will receive in week 5.

---

2. In order to measure the intensity, θ, of a source of radiation in a noisy environment a measurement X_1 is taken without the source present and a second, independent measurement X_2 is taken with it present. It is known that X_1 is $N(\mu, 1)$ and X_2 is $N(\mu + \theta, 1)$, where μ is the mean noise level. The prior distribution for μ is $N(\mu_0, 1)$ while the prior for θ is constant.
(a) Write down the joint posterior distribution of μ and θ up to a constant of proportionality.
(b) Hence obtain the posterior marginal distribution of θ.
(c) The usual estimate of θ is $x_2 - x_1$ explain why $\frac{1}{2}(2x_2 - x_1 - \mu_0)$ might be better.

---

3. A function $f(\theta)$ of the parameter is said to be *proper* if its integral is finite

$$\int f(\theta)d\theta < \infty$$

and can thus be normalized to be a probability distribution. Otherwise it is said to be *improper*.

Let $X_1, \dots, X_n$ be a random sample from a normal distribution with unknown mean μ and unknown variance λ^{-1}, with the improper prior

$$\pi_0(\mu, \lambda) = \lambda^{-1},\ \lambda > 0, -\infty < \mu < \infty.$$

(a) Find the joint posterior density $\pi(\mu, \lambda|x)$ of (μ, λ).
(b) Find the marginal posterior density $\pi(\mu|x)$ of μ.

(c) Show that the joint posterior density in (a) is proper.

---

4. Let $E(a, b)$ be the distribution of the shifted exponential with density

$$\frac{1}{b} e^{-(x-a)/b}, \qquad x > a$$

where $a \in \mathbb{R}, b > 0$ are parameters. Let $X_1, \ldots, X_n$ be a random sample from the distribution $E(a, b)$.

1. In this part suppose that a is fixed and known.

   (a) Let T be a statistic from the random sample. Write down the definition of *sufficiency, minimality and completeness* of T for a parameter θ.

   (b) Show that this is a one parameter exponential family. Without calculation, give a sufficient and complete statistic for b.

   (c) Find the minimal variance unbaised estimator (MVUE) $\hat{b}$ of b. (hint: remember that if $G \sim \Gamma(u, v)$ then $E(G) = u/v$) and that a sum of independent exponential variables with same means is Gamma distributed)

   (d) Compute the Fisher information for b and compare to the variance of $\hat{b}$. Does $\hat{b}$ attains the Cramer-Rao lower bound? How could you have inferred this directly from the log-likelihood ℓ? (Hint: if $G \sim \Gamma(u, v)$ then $Var(G) = u/v^2$).

2. Now we suppose that it is b which is known and a that we wish to estimate.

   (a) Can we still say that the distribution $E(a, b)$ with b known is an exponential family? Briefly justify your answer.

   (b) Find the distribution of $X_{(1)} = \min_{i=1,\ldots,n} X_i$?

   (c) Show that $X_{(1)}$ is sufficient and complete.

   (d) Hence deduce the MVUE of a.

3. Suppose now that both a and b are unknown. Define $T_1(X) = X_{(1)}$ and $T_2(X) = \sum_{i=1}^{n} X_i - X_{(1)}$.

(a) Fix $i \in \{1, \ldots, n\}$. Describe the conditional distribution of $X_i - X_{(1)}$ given $X_{(1)}$.

(b) Deduce from c.(i) that $T_2(X)$ has a Gamma distribution and give its parameters.

(c) Show that (T_1, T_2) forms a complete sufficient statistic for (a, b) and thus obtain an MVUE for (a, b) (Hint: you can assume without proof that T_1 and T_2 are independent).

---

5. k different people repeatedly perform a test that results in success or failure. Let X_i be the number of successes the ith person obtains in n_i trials. Assume $X_i \sim B(n_i, p_i)$, where the p_i vary between people but assume that, a priori, the p_i are exchangeable.
(a) Construct a hierarchical model in which the p_i are independent values from a Beta distribution whose parameters have an uninformative (uniform) prior distribution. Write the posterior joint distributions of $(p_1, \ldots, p_k, \alpha, \beta)$ and the marginal posterior distribution of p_1.
(b) Explain how estimates of the p_i could be obtained (but don't compute complicated integrals).

---

6. (a) Let $X \in \{0, 1, 2, ...\}$ be a random sample from a Poisson distribution with unknown mean $\theta \geq 0$.

(a.i) Compute the Jeffreys' prior $\pi_a(\theta)$ for θ.

(a.ii) Write down the posterior for θ given $X = x$, with Jeffreys' prior. Show that the posterior is proper.

(b). Let $Y \in \{0, 1\}$ be a Bernoulli random variable with success probability $e^{-\theta}$, $\theta \geq 0$.

(b.i) Compute the Jeffreys' prior $\pi_b(\theta)$ for θ.

(b.ii) Write down the posterior for θ given $Y = y$, with Jeffreys' prior. Show that the posterior is proper. (*Hint:* $\int_0^\infty \sqrt{\frac{e^{-\theta}}{1 - e^{-\theta}}}\, d\theta = \pi$.)

(c) Show that if $X = 0$ in (a) and $Y = 1$ in (b) the likelihoods and parameter spaces for θ are equal. Hence show that Jeffreys' priors violate the Likelihood Principle.

---

7. (a) Determine the maximum entropy prior for θ, where $\theta \in \Theta = (1, \infty)$ and $\mathbb{E}[\log(\theta)] = \lambda_0$ for some known fixed value of $\lambda_0 > 0$.
(b) Determine the maximum entropy prior for θ, where $\theta \in \Theta = (0, \infty)$, $\mathbb{E}[\log(\theta)] = \lambda_0$ and $\mathbb{E}[\theta] = \theta_0$ for known values λ_0 and θ_0. Derive the distribution explicitly when $\theta_0 = 4$ and $\lambda_0 = 11/6 - \gamma$, where γ is Euler's constant. Hint: Look up the digamma function $\psi(z)$ and note that for integer values $n \geq 2$,

$$\psi(n) = -\gamma + \sum_{k=1}^{n-1} k^{-1}.$$

Optional Exercises

8. Let $X = (X_1, \ldots, X_n)$ be a random sample from a density $f_X(x;\theta)$ belonging to a parametric family $\mathcal{F}$. Let $T = t(X)$ be a function of X and denote the density of T by $f_T(t;\theta)$. Assuming statistical regularity, define $i_X(\theta)$ to be the Fisher information about θ in X. Finally, let $i_{X|t}(\theta)$ denote the Fisher information conditional on $T = t$ and define

$$i_{X|T}(\theta) = \int i_{X|t}(\theta) f_T(t;\theta) dt$$

(a) Show that

$$i_X(\theta) = i_{X|T}(\theta) + i_T(\theta)$$

(b) Show that

$$i_X(\theta) \geq i_T(\theta),$$

with equality for all θ if and only if $T = t(X)$ is sufficient for θ.
Hint: Use the factorization theorem for the density

$$f_X(x;\theta) = f_{X|T}(x \mid t;\theta) f_T(t;\theta).$$

(c) Hence, or otherwise, determine the Fisher information about θ in the first r order statistics

$$X_{(1)} < X_{(2)} < \cdots < X_{(r)}$$

of a sample of size n from the density

$$f(x;\theta) = \theta \exp(-\theta x),\ x > 0$$

---

9. Suppose $T(x)$ is complete sufficient for θ given data x. Show that if a minimal sufficient statistic $S(x)$ for θ exists, then $T(x)$ is also minimal sufficient.

---

10. From 2015.

1. Let $X = (X_1, \ldots, X_n)$ be a sample from a continuous distribution with probability density function $f(x;\theta)$ where θ is an unknown parameter. Let $L(\theta;X)$ be the likelihood function and $\ell(\theta) = \log L(\theta;X)$

   (a) Prove that $\mathbb{E}[\frac{\partial \ell}{\partial \theta}] = 0$.

   (b) If we define $I_\theta = -\mathbb{E}[\frac{\partial^2 \ell}{\partial \theta^2}]$ show that $I_\theta = \mathbb{E}[(\frac{\partial \ell}{\partial \theta})^2] = \mathrm{Var}(\frac{\partial \ell}{\partial \theta})$.

   (c) Show that the variance of an unbiased estimator of θ, denoted $\hat{\theta}(X)$, satisfies the Cramér-Rao inequality

   $$\mathrm{Var}\big[\hat{\theta}(X)\big] \geq I_\theta^{-1},$$

   and provide a brief statement about why regularity conditions are needed for this result to hold.

   (d) Show that there exists an unbiased estimator $\widehat{\theta}$ which attains the Cramér-Rao lower bound (under regularity conditions) if and only if

   $$\frac{\partial \ell}{\partial \theta} = I_\theta(\widehat{\theta} - \theta)$$

   (e) If $\hat{g}(X)$ is an unbiased estimator of the function $g(\theta)$ derive a lower bound for the variance of the estimator.

2. Suppose that

   $$f(x;\theta) = \frac{\theta^3 x(x+1)}{\theta+2} e^{-\theta x}, \quad x \geq 0, \quad \theta > 0.$$

   Find an unbiased estimator of $(\theta^2 - 6)/[\theta(\theta+2)]$ whose variance attains the Cramér-Rao lower bound.