Midterm Exam: Statistical Inference

Due: Thu Mar 8 2:15pm

1. Let $x_j \sim f(x \mid \alpha)$ and $y_j \sim g(y \mid \beta)$ be independent random variables and consider the problem about making inference about $\theta = (\alpha, \beta)$. Note that either or both of α, β may be vectors.

   a. Give the likelihood function $L_n(\theta)$ upon observing (x_j, y_j) for $1 \le j \le n$.

   b. Is the Jeffreys prior distribution $\pi_J(\theta)$ just the product $\pi_J(\alpha)\pi_J(\beta)$ of the Jeffreys prior distributions for separate observation of $\mathbf{x}$ and $\mathbf{y}$? Show why.

---

2. For each subject in a clinical trial the time-to-remission T_i is an exponentially distributed random variable whose rate parameter will depend on whether the subject received the Treatment or the Control medication. Thus for treated and control subjects the survival times satisfy, for $t > 0$,

$$\Pr[T_i > t \mid \text{Treated}] = e^{-\lambda_1 t} \qquad \Pr[T_i > t \mid \text{Control}] = e^{-\lambda_0 t}.$$

   Subjects are assigned at random to treated or control groups (neither the subject nor the investigator knows who receives which treatment). For each of the n subjects we have an observed value t_i of T_i (there is no censoring) and we have an indicator $\tau_i = 1$ for Treated group, $\tau_i = 0$ for Control group. Set $\theta = (\lambda_0, \lambda_1) \in \Theta = \mathbb{R}_+^2$ and let $\mathbf{x} = (t_1, \tau_2, ..., t_n, \tau_n)$. There is at least one subject with each treatment.

   a. Is this an exponential family? If so, write $f(\mathbf{x} \mid \theta)$ in standard form $h(\mathbf{x}) \exp\left[\eta(\theta) \cdot T(\mathbf{x}) - B(\theta)\right]$ for suitable $h(\mathbf{x})$, $B(\theta)$, $k \in \mathbb{N}$, and $\eta(\theta)$, $T(\mathbf{x}) \in \mathbb{R}^k$; if not, explain why.

   b. Find the maximum likelihood estimator $\hat{\theta}_n(\mathbf{x})$ for n observations.

   c. For independent gamma $\lambda_j \sim \mathsf{Ga}(\alpha, \beta)$ prior distributions find the posterior mean (Bayes estimator) $\bar{\theta}_n = \mathsf{E}[\theta \mid \mathbf{x}_n]$.

   d. Find the Jeffreys prior $\pi_J(\theta)$ and its posterior mean $T_J(x_n)$.

---

3. For $\alpha > 0$ let $X_j \sim \mathsf{Ga}(4, \theta)$ be independent gamma random variables with density functions $f(x \mid \theta) = \theta^4 \, x^3 \, e^{-\theta x}/6, \quad x > 0$.

   a. For any $\alpha > 0$ and $\theta > 0$ and $p > -\alpha$, let $Y \sim \mathsf{Ga}(\alpha, \theta)$ with density function $f(x \mid \theta) = \theta^\alpha \, x^{\alpha-1} \, e^{-\theta x}/\Gamma(\alpha), \quad x > 0$ and evaluate the expectation $\mathsf{E}[Y^p]$. Note: For all *other* parts of this problem, $\alpha = 4$; this is the only part where general $\alpha > 0$ is considered.

   b. Give the log-likelihood function $\ell(\theta)$ for a sample $\mathbf{x} = \{x_1, ..., x_n\} \in \mathbb{R}_+^n$ of n observations $X_j = x_j$, $X_j \sim \mathsf{Ga}(4, \theta)$.

   c. Find the Fisher information $I(\theta)$ for the one-parameter family $\mathsf{Ga}(4, \theta)$.

   d. Find the MLE $\hat{\theta}_n(\mathbf{x})$, and evaluate its bias

   $$\beta_n(\theta) \equiv \left\{ \mathsf{E}[\hat{\theta}_n \mid \theta] - \theta \right\}.$$

   Is $\hat{\theta}_n$ unbiased, either asymptotically or for all n?

   e. Is $\hat{\theta}_n(x)$ efficient? Answer this by finding the risk function

   $$R(\hat{\theta}_n, \theta) = \mathsf{E}[(\hat{\theta}_n - \theta)^2 \mid \theta]$$

   and the limit $\lim_{n\to\infty} \left[n \, R(\hat{\theta}_n, \theta) \right]$. Is the Cramer-Rao lower bound attained, either asymptotically or for all n?

   f. Find the Jeffreys prior distribution $\pi_J(\theta)$ and the posterior distribution upon observing $n = 3$ data points with sum $\sum X_i = 24$.

   g. Find an *exact* 95% equal-tails **credible** interval $[L, R]$ for θ, using the Jeffreys prior and the data given above. `S-Plus` might be helpful here.

4. Sometimes multiple observations are *not* independent. Let $\lambda_i \sim \mathsf{Ga}(\alpha, \beta)$ be *i.i.d.* and, for each i, let $n_i \in \mathbb{N}$ and let $X_{ij} \sim \mathsf{Po}(\lambda_i z_i)$ for $j \le n_i$. We observe only the $X_{ij} \in \mathbb{N}$ and of course the known constant values of the $n_i \in \mathbb{N}$ and the $z_i > 0$, and we regard $\theta = (\alpha, \beta) \in \mathbb{R}_+^2$ and $\lambda = (\lambda_1, ..., \lambda_n) \in \mathbb{R}_+^n$ as parameters. This is an example of a *hierarchical model*.

   a. Given any prior density $\pi(\theta)$, write down an expression for the joint posterior distribution of θ and λ, given $\mathbf{x} = (\{X_{ij}\})$. (Hint: first write down the joint distribution of *everything* uncertain: α, β, $\{\lambda_j\}$, $\{X_{ij}\}$).

   b. Write down an expression for the marginal posterior distribution of θ, given $\mathbf{x} = (\{X_{ij}\})$. Simplify.

   c. Do you expect the X_{ij} to be mutually independent, or would you expect some or all of them to be positively or negatively correlated (marginally, *not* conditionally on λ or θ)? *why*? Answer in words (no calculations are needed).