Midterm Exam: Statistical Inference

Due: Thu Mar 9 2:15pm

1. For $\theta > 0$ set

$$f(x \mid \theta) = \theta\, x^{\theta-1}, \qquad 0 < x < 1 \tag{1}$$

and answer the following questions.

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

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

c. For a gamma $\mathsf{Ga}(\alpha, \lambda)$ prior distribution find the posterior mean (Bayes estimator) $\bar{\theta}_n = \mathsf{E}[\theta \mid x_n]$.

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

e. Find a consistent Method of Moments estimator $T_M(x_n)$ based on $\bar{x}_n$. Show that it is consistent in probability, i.e., show that $T_M(x_n) \to \theta$ as $n \to \infty$.

f. In a paragraph or so at most, which of these estimators would you recommend, and why?

---

2. Let $x_j \sim \mathsf{No}(\theta, 1)$ be independent Normal random variables with unit variance for $1 \le j \le n$.

a. For a normal $\mathsf{No}(\xi, \tau^2)$ prior distribution find the posterior mean (Bayes estimate) $\bar{\theta}_n = \mathsf{E}[\theta \mid x]$ based on $x = \{x_1, ..., x_n\}$.

b. Find and plot the risk functions $R(\hat{\theta}_n, \theta)$ and $R(\bar{\theta}_n, \theta)$ for $n = 5$ with $\xi = \tau^2 = 1$. For which θ is $R(\bar{\theta}_n, \theta) < R(\hat{\theta}_n, \theta)$?

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

d. Which estimator would you recommend, T_J or $\bar{\theta}_n$? Why? If the question does not give enough information for you to give a recommendation, explain what else you would need to know.

---

3. Let T be the natural statistic in a one-dimensional exponential family

$$f(x \mid \eta) = h(x)\, e^{\eta\, T(x) - c(\eta)}$$

with natural parameter η. Show (in detail) that T is *minimal* sufficient.

---

4. For $\theta > 0$ let $X_j \sim \mathsf{Un}(0, \theta)$ be independent uniforms.

   a. Find the maximum likelihood estimator, $\hat{\theta}_n(x)$, for a sample $x = \{x_1, ..., x_n\} \in \mathbb{R}_+^n$ (**Warning**: this is not a "regular" problem in the sense that $f_n(x \mid \theta)$ is not differentiable).

   b. Is $\hat{\theta}_n(x)$ unbiased or not? Find $\beta_n(\hat{\theta}, \theta) = \big\{\mathsf{E}[\hat{\theta}_n \mid \theta] - \theta\big\}$.

   c. Is $\hat{\theta}_n(x)$ efficient? Find the risk function $R(\hat{\theta}_n, \theta) = \mathsf{E}[(\hat{\theta}_n - \theta)^2 \mid \theta]$ and the limit $\lim_{n\to\infty} \big[n\, R(\hat{\theta}_n, \theta)\big]$; discuss the implications.

   d. Find an *exact* 95% equal-tails **confidence** interval $[L_x, R_x]$ for θ, based on $\hat{\theta}_n$— for example, we must have $\mathsf{P}[\theta < L_x \mid \theta] = 0.025$. How long is your interval, on average?

   e. A common choice for "noninformative" prior distribution for this model would be $\pi(\theta) = \theta^{-1}$, $\theta > 0$. Using this prior, find an exact 95% equal-tails **credible** interval for θ— for example, we must have $\mathsf{P}[\theta < L_x \mid x] = 0.025$.

   f. Find the Bayesian posterior mean $\bar{\theta}_n(x) = \mathsf{E}[\theta \mid x]$ for this prior, and compute its risk function $R(\bar{\theta}_n, \theta) = \mathsf{E}[(\bar{\theta}_n - \theta)^2 \mid \theta]$.

   g. Show that $T_n(x) = 2\bar{x}_n$ is a consistent unbiased estimator of θ, and compute its risk function $R(T_n, \theta)$.

   h. Find an approximate 95% confidence interval for θ based on $T_n(x)$ (Hint: what is the asymptotic *distribution* of T_n?). Is it better or worse than the interval you found in (d.) above? Why?

   i. Which is the best estimator for θ: $\hat{\theta}_n$, $\bar{\theta}_n$, or $T_n(x)$? Why?
