Statistical Theory - Problem Sheet 3

Spring 2023. Email corrections to Kolyan Ray: kolyan.ray@ic.ac.uk

Instructions: Please attempt the non-starred questions. If you have time, attempt the starred questions (they are not necessarily more difficult).

1. Let $\Theta \subseteq \mathbb{R}$ have nonempty interior and let S_n be a sequence of random real-valued continuous functions defined on Θ such that, as $n \to \infty$, $S_n(\theta) \to^P S(\theta)\ \forall \theta \in \Theta$, where $S : \Theta \to \mathbb{R}$ is non-random. Suppose for some θ_0 in the interior of Θ and every $\varepsilon > 0$ small enough we have $S(\theta_0 \pm \varepsilon) < 0 < S(\theta_0 \mp \varepsilon)$, and that S_n has *exactly one* zero $\hat{\theta}_n$ for every $n \in \mathbb{N}$. Deduce that $\hat{\theta}_n \to^P \theta_0$ as $n \to \infty$.

2. Consider i.i.d. random variables $X_1, \dots, X_n$ arising from the statistical model

$$\left\{ f_\theta(x) = \theta x^{\theta-1} \exp\{-x^\theta\},\ x > 0,\ \theta \in (0, \infty) \right\}$$

of *Weibull distributions*. Show that the MLE exists with probability one and is consistent.

Hint: For proving consistency, use the previous question. You may also interchange differentiation $d/d\theta$ and dx-integration without justification in your argument.

3. Consider the parameter $\phi = EX^4$ equal to the fourth moment of a $N(0, \theta)$ distribution. Find the MLE $\hat{\phi}$ of ϕ and derive the asymptotic distribution of $\sqrt{n}(\hat{\phi} - \phi)$ as $n \to \infty$.

4. Consider the maximum likelihood estimator $\hat{\theta}$ from $X_1, \dots, X_n$ i.i.d. $N(\theta, 1)$, where $\theta \in \Theta = [0, \infty)$. Show that $\sqrt{n}(\hat{\theta} - \theta)$ is asymptotically normal whenever $\theta > 0$. What happens when $\theta = 0$? Comment on your findings in light of the general asymptotic theory for maximum likelihood estimators.

5. Let $X_1, \dots, X_n \sim^{iid} f_\theta$, with density $f_\theta(x) = e^{-(x-\theta)}$ for $x > \theta$.

(i) Find the maximum likelihood estimator $\hat{\theta}$ of θ. By finding the density function of $\hat{\theta}$, show that $\hat{\theta}$ is a consistent but biased estimator of θ with $E_\theta \hat{\theta} = \theta + 1/n$.

(ii) Find the variance of the unbiased estimator $\tilde{\theta} = \hat{\theta} - 1/n$. Is it appropriate to compare it with the Cramèr-Rao inequality?

(iii) Find the asymptotic distribution of $n(\tilde{\theta} - \theta)$ as $n \to \infty$.

6. Suppose $X_1, \dots, X_n \sim^{iid}$ Poisson(θ) , $\theta > 0$ and consider prior $\theta \sim \Gamma(\alpha, \beta)$. Show that the posterior distribution of $\theta | X_1, \dots, X_n$ is also a Gamma distribution and find its parameters.

7. (i) For $\sigma^2 > 0$ a fixed constant, consider the Bayesian model $X_1, \dots, X_n \sim^{iid} N(\theta, \sigma^2)$ with prior distribution $\theta \sim N(\mu, v^2)$, $\mu \in \mathbb{R}$, $v^2 > 0$. Show that the posterior distribution of θ given the observations is

$$\theta | X_1, \dots, X_n \sim N\left(\frac{\frac{n\bar{X}}{\sigma^2} + \frac{\mu}{v^2}}{\frac{n}{\sigma^2} + \frac{1}{v^2}}, \frac{1}{\frac{n}{\sigma^2} + \frac{1}{v^2}} \right), \quad \text{where } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

(ii)* The *total variation* distance between two probability distributions with densities f and g is defined as $\|f - g\|_{TV} = \int |f(x) - g(x)| dx$. Compute the total variation distance between a $N(\theta_1, \tau^2)$ and $N(\theta_2, \tau^2)$ distribution in terms of the standard normal cumulative distribution function $\Phi(x) = P(N(0, 1) \leq x)$.

(iii)* Returning to (i), suppose that $\sigma^2 = 1$ and $\mu = 0$ for simplicity, so that $X_1, \ldots, X_n \sim^{iid} N(\theta, 1)$ and $\theta \sim N(0, v^2)$. Compute the posterior distribution of the rescaled quantity $\sqrt{n}(\theta - \hat{\theta}_{ML})|X_1, \ldots, X_n$, and evaluate the total variation distance between this distribution and a $N(0, \frac{n}{n+1/v^2})$ distribution.

Under the frequentist assumption that $X_1, \ldots, X_n \sim^{iid} N(\theta_0, 1)$ for a true $\theta_0 \in \mathbb{R}$, how does the rescaled posterior compare to the limiting distribution of $\sqrt{n}(\hat{\theta}_{ML} - \theta_0)$?

Hint: recall the MLE and other useful facts from Q1(b) and Q10(b) on Problem Sheet 2.

8. (i) Let $X \sim \text{Poisson}(\theta)$ with $\theta \geq 0$. Find the Jeffreys prior for θ. Write down the corresponding posterior distribution of $\theta|X$ and show that it is proper.

(ii) Let $Y \sim \text{Bernoulli}(e^{-\theta})$ with $\theta \geq 0$. Find the Jeffreys priors for θ. Write down the corresponding posterior distribution of $\theta|Y$ and show that it is proper.

Hint: $\int_0^\infty \left(\frac{e^{-\theta}}{1-e^{-\theta}}\right)^{1/2} d\theta = \pi$.

(iii) The *likelihood principle* states that if two likelihoods are proportional to one another (as functions of θ), then one should make the same inference based on either one. Show that if $X = 0$ in (i) and $Y = 1$ in (ii), the likelihoods and parameter spaces are equal. Hence show that Jeffreys priors do not satisfy the likelihood principle.

9. Consider $X_1, \ldots, X_n|(\mu, \sigma^2) \sim^{iid} N(\mu, \sigma^2)$ with *improper prior* density $\pi(\mu, \sigma)$ proportional to σ^{-2} (constant in μ). Argue that the resulting 'posterior distribution' has a density proportional to

$$\sigma^{-(n+2)} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2\right\},$$

and that thus the distribution of $\mu|(\sigma^2, X_1, \ldots, X_n)$ is $N(\bar{X}_n, \sigma^2/n)$, where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.