STA261 S19: Test 3 Solutions

No aids. 60 minutes. Write all answers directly beneath where the question is asked. Use pages 4 and 5 for rough work.

1. *Basic, 4 marks* Let $X_1, \dots, X_n$ be an IID sample from a parametric family of distributions $\{F_\theta : \theta \in \Theta\}$ with corresponding densities f_θ.

a) (1) Give a mathematical definition of the likelihood ratio for assesing whether $\theta = \theta_0$ against the alternative that $\theta \neq \theta_0$.

Solution. Let $L(\theta \mid X_1, \dots, X_n)$ denote the likelihood function. Then, the likelihood ratio is

$$\lambda(X_1, \dots, X_n) = \frac{L(\theta_0 \mid X_1, \dots, X_n)}{\sup_{\theta \neq \theta_0} L(\theta \mid X_1, \dots, X_n)}. \qquad \text{(1 mark – max is acceptable)}$$

b) (1) Circle True or False: you and I observe two random samples from F_θ, and their sufficient statistics happen to be equal. You and I will make the same decision about whether $\theta = \theta_0$ if we base our decision on the likelihood ratio at the same level.

Solution. True (1 mark – no explanation necessary). Let the first sample be $X_1, \dots, X_n$ and the second be $Y_1, \dots, Y_n$. For a sufficient statistic T, the factorization theorem implies that the likelihood ratio can be written as

$$\begin{aligned}
\lambda(X_1, \dots, X_n) &= \frac{\prod_{i=1}^n h(X_i) f_{\theta_0}(T(X_i))}{\sup_{\theta \neq \theta_0} \prod_{i=1}^n h(X_i) f_\theta(T(X_i))} \\
&= \frac{\prod_{i=1}^n f_{\theta_0}(T(X_i))}{\sup_{\theta \neq \theta_0} \prod_{i=1}^n f_\theta(T(X_i))} && (h(X_i) \text{ does not depend on } \theta) \\
&= \frac{\prod_{i=1}^n f_{\theta_0}(T(Y_i))}{\sup_{\theta \neq \theta_0} \prod_{i=1}^n f_\theta(T(Y_i))} \\
&= \lambda(Y_1, \dots, Y_n).
\end{aligned}$$

Since the two samples have the same likelihood ratio value, they will lead to the same decision.

c) (1) Suppose I go Bayesian on you and put a prior $\pi(\theta)$ on θ. Give an expression for the posterior distribution of $\theta | X$.

Solution.

$$p(\theta \mid X_1, \dots, X_n) = \frac{L(\theta \mid X_1, \dots, X_n)\pi(\theta)}{\int_{\nu \in \Theta} L(\nu \mid X_1, \dots, X_n)\pi(\nu) d\nu}. \qquad \text{(1 mark)}$$

d) (1) Circle True or False: consider the situation in part **b)**. You and I will make the same inferences if we base those inferences off of the posterior distribution of $\theta|X$.

Solution. True (1 mark – no explanation necessary). By the same reasoning as part b),

$$\begin{aligned}
p(\theta \mid X_1, \ldots, X_n) &= \frac{\prod_{i=1}^n h(X_i) f_\theta(T(X_i))\pi(\theta)}{\int_{\nu\in\Theta} \prod_{i=1}^n h(X_i) f_\nu(T(X_i))\pi(\nu)d\nu} \\
&= \frac{\prod_{i=1}^n f_\theta(T(X_i))\pi(\theta)}{\int_{\nu\in\Theta} \prod_{i=1}^n f_\nu(T(X_i))\pi(\nu)d\nu} \\
&= \frac{\prod_{i=1}^n f_\theta(T(Y_i))\pi(\theta)}{\int_{\nu\in\Theta} \prod_{i=1}^n f_\nu(T(Y_i))\pi(\nu)d\nu} \\
&= p(\theta \mid Y_1, \ldots, Y_n).
\end{aligned}$$

Since the two samples have the same posterior distribution, they will lead to the same decision.

2. *Adept, 4 marks.* The gamma function is defined as:

$$\Gamma(n+1) = \int_0^\infty x^n e^{-x} dx \tag{0.1}$$

For $n \in \mathbb{N}$, $\Gamma(n+1) = n! = n \times (n-1) \times \cdots \times 2 \times 1$. Use Laplace approximations to prove Stirling's approximation:

$$n! \approx n^n e^{-n} \sqrt{2\pi n} \tag{0.2}$$

Solution. Since $e^{\log(x)} = x$ and $\log(x^n) = n\log(x)$,

$$n! = \Gamma(n+1) = \int_0^\infty x^n e^{-x} dx = \int_0^\infty e^{n\log(x) - x} dx. \qquad \text{(1 mark)}$$

We wish to write $e^{n\log(x)-x}$ in the form of $e^{nf(x)}$. Making the change of variable $y = x/n$ gives

$$n! = \int_0^\infty e^{n\log(ny) - ny} n dy = n e^{n\log(n)} \int_0^\infty e^{n(\log(y) - y)} dy. \qquad \text{(1 mark)}$$

To approximate this integral, notice $f(y) = \log(y) - y$ achieves its maximum at $y_0 = 1$. Also, $f''(y) = -1/y^2$, so $f(y_0) = f''(y_0) = -1$ (1 mark). So, using the Laplace approximation formula,

$$\int_0^\infty e^{nf(x)} \approx \frac{\sqrt{2\pi} e^{nf(x_0)}}{\sqrt{-nf''(x_0)}},$$

we get

$$n! \approx n e^{n\log(n)} \frac{\sqrt{2\pi} e^{-n}}{\sqrt{n}} = n^n e^{-n} \sqrt{2\pi n}. \qquad \text{(1 mark)}$$

3. *Advanced, 2 marks.* Let $X_1, \dots, X_n$ be an IID random sample from a $\mathrm{Unif}(0, \theta)$ distribution.

(a) Find $\widehat{\theta}$, the maximum likelihood estimator for θ.

Solution. The likelihood function is

$$L(\theta \mid X_1, \dots, X_n) = \prod_{i=1}^{n} f_\theta(X_i) = \prod_{i=1}^{n} \frac{1}{\theta} = \theta^{-n}.$$

The log-likelihood is thus

$$\ell(\theta \mid X_1, \dots, X_n) = \log L(\theta \mid X_1, \dots, X_n) = -n \log(\theta).$$

Differentiating gives the score function

$$S(\theta \mid X_1, \dots, X_n) = \frac{\partial \ell(\theta \mid X_1, \dots, X_n)}{\partial \theta} = -\frac{n}{\theta}. \qquad \text{(0.25 marks)}$$

Setting this to zero does not give us anything reasonable. However, notice that since log is a monotonically increasing function, $\ell(\theta \mid X_1, \dots, X_n)$ is a monotonically decreasing function. Thus, to maximize the log-likelihood, θ should be chosen as small as possible. Further, for every observation, it must hold that $X_i \leq \theta$, since otherwise the value observed is impossible. Thus, the MLE is

$$\widehat{\theta} = \max_{i \in 1, \dots, n} X_i. \qquad \text{(0.75 marks)}$$

(b) Is the asymptotic distribution of $\widehat{\theta}$ Normal? If yes, state why and give the mean and variance. If no, find the cumulative distribution function of $\widehat{\theta}$, and state why it is *not* normal.

Solution. The CDF of $\widehat{\theta}$, defined for $x \in (0, \theta)$, is

$$F(x) = P(\widehat{\theta} \leq x) = P\left(\max_{i \in 1, \dots, n} X_i \leq x\right) = P\left(\forall i : X_i \leq x\right) = \prod_{i=1}^{n} P(X_i \leq x) = \prod_{i=1}^{n} \frac{x}{\theta} = \left(\frac{x}{\theta}\right)^n. \qquad \text{(0.5 marks)}$$

The reason $\widehat{\theta}$ is not asymptotically normal is that the domain of the random variable X_i depends on the parameter θ (0.5 marks).

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED.

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED.