Statistical Theory - Problem Sheet 1

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. Show that the following families of distributions are exponential families.

(i) $X \sim \text{Beta}(a,b)$, $a,b>0$, with $f_{a,b}(x) = x^{a-1}(1-x)^{b-1}/B(a,b)$ for $x \in [0,1]$.

(ii) $X \sim \text{Geometric}(p)$, $p \in (0,1)$, with $P(X=k) = p(1-p)^{k-1}$ for $k = 1,2,\ldots$.

(iii) $X \sim \text{Gamma}(\alpha,\beta)$, $\alpha,\beta>0$, with $f_{\alpha,\beta}(x) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}e^{-\beta x}$ for $x>0$.

2. Let $X_1,\ldots,X_n \sim^{iid} U[0,\theta]$ with $\theta > 0$. Estimate θ using both the method of moments and maximum likelihood. Calculate the means and variances of the two estimators. Which one should you prefer in terms of mean-squared error?

3. Let $X_1,\ldots,X_n$ be i.i.d. random variables from a parametric statistical model. For the following cases, find sufficient statistics for the parameters with the same dimension as the number of model parameters.

(a) $X_i \sim^{iid} \text{Beta}(a,b)$, $a,b>0$.

(b) $X_i \sim^{iid} \text{Beta}(a,a)$, $a>0$.

(c) $X_i \sim^{iid} f_{\mu,\sigma}$, where $f_{\mu,\sigma}(x) = \frac{1}{\sigma}e^{-(x-\mu)/\sigma}$ for $\mu < x < \infty$ and $\sigma > 0$.

(d) $X_i \sim^{iid} \Gamma(\alpha,\beta)$, $\alpha,\beta>0$.

4. (i) Let $X_1,\ldots,X_n$ be independent random variables with densities

$$f_{X_i,\theta}(x) = \begin{cases} e^{i\theta - x}, & x \geq i\theta, \\ 0, & x < i\theta. \end{cases}$$

Show that $T = \min_i(X_i/i)$ is a sufficient statistic for θ.

(ii) Now suppose

$$f_{X_i,\theta}(x) = \begin{cases} \frac{1}{2i\theta}, & -i(\theta-1) < x < i(\theta+1), \\ 0, & \text{otherwise}. \end{cases}$$

Find a two-dimensional sufficient statistic for θ.

5. Let $X_1,\ldots,X_n \sim^{iid} \text{Bernoulli}(\theta)$, $0<\theta<1$, with n even. Show that $T(x_1,\ldots,x_n) = \sum_{i=1}^{n/2} x_i$ is *not* a sufficient statistic for θ.

6. Let $X_1,\ldots,X_n$ be independent Poisson random variables with $X_i \sim \text{Poisson}(i\theta)$ for some $\theta>0$. Find a minimal sufficient statistic T for θ and compute its sampling distribution. Find the maximum likelihood estimator of θ and show that it is unbiased.

why this is required?

7. For $n \geq 3$, let $X_1,\ldots,X_n \sim^{iid} \text{Exp}(\theta)$.

(a) Find the method of moments estimator of θ.

(b) Find a minimal sufficient statistic T and compute its sampling distribution. Show that the maximum likelihood estimator is biased for fixed n, but asymptotically unbiased. Show that its variance tends to zero as $n \to \infty$.

(c) Find an injective function $h : (0, \infty) \to \mathbb{R}$ such that, writing $\phi = h(\theta)$, the maximum likelihood estimator $\hat{\phi}$ of the new parameter ϕ is unbiased.

8. Let $X_1, \dots, X_n \sim^{iid} \text{Bernoulli}(p)$, $0 \leq p \leq 1$.

(i) Show that a sufficient statistic for $\theta = (1-p)^2$ is $T(X) = \sum_{i=1}^n X_i$ and that the MLE for θ is $(1 - \frac{1}{n}T)^2$.

Hint: use the chain rule $\frac{df}{d\theta} = \frac{df}{dp}\frac{dp}{d\theta}$

(ii) Show that $\tilde{\theta} = \mathbb{1}\{X_1 + X_2 = 0\}$ is an unbiased estimator of θ.

(iii) Find a function of T which is an unbiased estimator for θ.

9*. Let $X_1, \dots, X_n \sim^{iid} U[\theta, 2\theta]$ for some $\theta > 0$. Show that $\tilde{\theta} = \frac{2}{3}X_1$ is an unbiased estimator for θ. Find a two-dimensional minimal sufficient statistic T. Using the Rao-Blackwell theorem or otherwise, find an unbiased estimator $\hat{\theta}$ of θ, which is a function of T, and satisfies $\text{Var}_\theta(\hat{\theta}) < \text{Var}_\theta(\tilde{\theta})$ for all $\theta > 0$.

10*. Let $X \sim f_\theta$ for some pmf/pdf f_θ, $\theta \in \Theta$. Show that if $\hat{\theta}_{ML}$ is both the *unique* maximum likelihood estimator for θ and also a sufficient statistic for θ, then it is minimal sufficient for θ.

11*. For $n \geq 3$, let $u_1, \dots, u_n \sim^{iid} N(0,1)$ and set $X_1 = u_1$ and $X_i = \theta X_{i-1} + (1-\theta^2)^{1/2} u_i$ for $i = 2, \dots, n$ and some $-1 < \theta < 1$. Find a sufficient statistic for θ that takes values in a subset of $\mathbb{R}^3$.