STAT 713 sp 2023 Exam 1

1. Let $X_1, \dots, X_n \stackrel{\text{ind}}{\sim} f(x;\beta) = \beta^2 x e^{-\beta x}\mathbf{1}(x > 0)$ for some $\beta > 0$.

   (a) Find the maximum likelihood estimator for β.

   The likelihood function is given by

   $$\mathcal{L}(\beta;\mathbf{X}) = \beta^{2n}(\textstyle\prod_{i=1}^n X_i)e^{-\beta n \bar{X}_n}$$

   and the log-likelihood is given by

   $$\ell(\beta;\mathbf{X}) = 2n\log\beta + \textstyle\sum_{i=1}^n \log X_i - \beta n \bar{X}_n.$$

   Now we have

   $$\frac{\partial}{\partial\beta}\ell(\beta;\mathbf{X}) = \frac{2n}{\beta} - n\bar{X}_n = 0 \iff \beta = \frac{2}{\bar{X}_n},$$

   so the MLE for β is $\hat{\beta}_n = 2/\bar{X}_n$.

   (b) Find the maximum likelihood estimator for $\tau = \tau(\beta) = 1/\beta$.

   The MLE for τ is $\tau(\hat{\beta}) = 1/\hat{\beta}_n = \bar{X}_n/2$.

   (c) Check whether $T(X_1,\dots,X_n) = (\prod_{i=1}^n X_i, \sum_{i=1}^n X_i)$ is a minimal sufficient statistic.

   This is a sufficient, but not a minimal sufficient statistic. We can see this by writing, for two samples $\mathbf{x}$ and $\mathbf{y}$ the ratio of joint densities

   $$\frac{\beta^{2n}(\prod_{i=1}^n x_i)e^{-\beta\sum_{i=1}^n x_i}}{\beta^{2n}(\prod_{i=1}^n y_i)e^{-\beta\sum_{i=1}^n y_i}} = \frac{(\prod_{i=1}^n x_i)e^{-\beta\sum_{i=1}^n x_i}}{(\prod_{i=1}^n y_i)e^{-\beta\sum_{i=1}^n y_i}}.$$

   we see that this constant in β if and only if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. The values $\prod_{i=1}^n x_i$ and $\prod_{i=1}^n y_i$ may be different and yet the ratio of densities still be constant, so the statistic $(\prod_{i=1}^n X_i, \sum_{i=1}^n X_i)$ is *not* a minimal sufficient statistic.

   (d) Find the value of $\text{Cov}(\hat{\beta}_n, S_n\hat{\beta}_n)$, where $\hat{\beta}_n$ is the MLE for β and S_n^2 is the sample variance.

   The quantity $S_n\hat{\beta}_n$ is an ancillary statistic, while $\hat{\beta}_n$ is a complete sufficient statistic; therefore $S_n\hat{\beta}_n$ and $\hat{\beta}_n$ are independent and $\text{Cov}(\hat{\beta}_n, S_n\hat{\beta}_n) = 0$.

The ancillarity of $S_n\hat{\beta}_n$ comes from the fact that the function

$$S_n\hat{\beta}_n = \frac{2\sqrt{(n-1)^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2}}{\bar{X}_n}$$

is scale-invariant (returns the same value for $X_1, \ldots, X_n$ as for $cX_1, \ldots, cX_n$ for any $c > 0$) and that we can write $f(x;\beta) = 1/\beta^{-1} f_X(x/\beta^{-1})$ for any $\beta > 0$ and $x \in \mathbb{R}$, where $f_Z(z) = ze^{-z}\mathbf{1}(z > 0)$. That is, $f(x;\beta)$ belongs to a scale family.

The completeness of $\hat{\beta}_n$ comes from the fact that $f(x;\beta)$ belongs to a full exponential family with $w_1(\beta) = -\beta$ and $t_1(x) = x$, giving that $\sum_{i=1}^n X_i$ is a complete minimal sufficient statistic. Since $\hat{\beta}_n$ is a one-to-one function of $\sum_{i=1}^n X_i$, $\hat{\beta}_n$ is also a complete sufficient statistic.

2. A randomly selected spectator of a USC basketball game will shoot free throws until making two baskets. Suppose the ability $\theta \in (0,1)$ of a randomly selected spectator has the pdf $\pi(\theta) = 6\theta(1-\theta)\mathbf{1}(0 < \theta < 1)$ and, given the spectator's ability θ, the number of shots Y required by the spectator to make two baskets has pmf $p(y|\theta) = (y-1)\theta^2(1-\theta)^{y-2}\mathbf{1}(y \in \{2,3,\dots\})$.

(a) Give the posterior distribution of θ given Y.

We have

$$\pi(\theta|Y) \propto (y-1)\theta^2(1-\theta)^{y-2}6\theta(1-\theta) \propto \theta^3(1-\theta)^{y-1},$$

so $p|Y \sim \text{Beta}(4, y)$.

(b) Give a Bayesian estimate of the ability θ of a spectator who makes the 2nd basket on the 5th shot.

The posterior mean is $\mathbb{E}[\theta|Y] = 4/(Y+4)$. If the spectator makes the 2nd basket on the 5th shot, we would estimate his or her ability with $\hat{\theta}_{\text{Bayes}} = 4/9$.

(c) Give the value of the maximum likelihood estimator of θ for the same spectator (treat θ as fixed).

Using on the distribution of $Y|\theta$ and considering θ as fixed, the likelihood function for θ is given by

$$\mathcal{L}(\theta; Y) = (Y-1)\theta^2(1-\theta)^{Y-2}$$

and the log-likelihood by

$$\ell(\theta; Y) = \log(Y-1) + 2\log\theta + (Y-2)\log(1-\theta).$$

Setting

$$\frac{\partial}{\partial\theta}\ell(\theta; Y) = \frac{2}{\theta} - \frac{Y-2}{1-\theta} = 0$$

and solving for θ gives

$$\hat{\theta} = \frac{2}{Y}.$$

For a spectator who makes the 2nd basket on the 5th shot, we would estimate his or her ability as $\hat{\theta} = 2/5$.

3. Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} f(x;\rho) = (1-\sqrt{\rho})^{-1}\mathbf{1}(\sqrt{\rho} \le x \le 1)$.

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

We have $m_1 = (1+\sqrt{\rho})/2$, giving $\rho = (2m_1 - 1)^2$. So the method of moments estimator for ρ is $\bar{\rho} = (2\hat{m}_1 - 1)^2$, where $\hat{m}_1 = \bar{X}_n$.

(b) Find the bias of the method of moments estimator.

We have

$$\operatorname{Bias}\bar{\rho} = \mathbb{E}(2\hat{m}_1 - 1)^2 - \rho = 4\mathbb{E}\hat{m}_1^2 - 4\mathbb{E}\hat{m}_1 + 1 - \rho = 4[\operatorname{Var}\hat{m}_1 + (\mathbb{E}m_1)^2] - 4\mathbb{E}\hat{m}_1 + 1 - \rho.$$

Substituting $\mathbb{E}\hat{m}_1 = (1+\sqrt{\rho})/2$ and $\operatorname{Var}\hat{m}_1 = (1-\sqrt{\rho})^2/(12n)$ gives

$$\operatorname{Bias}\bar{\rho} = \frac{(1-\sqrt{\rho})^2}{3n}.$$

(c) Find the maximum likelihood estimator of ρ.

The likelihood function for ρ is given by

$$\mathcal{L}(\rho;\mathbf{X}) = (1-\sqrt{\rho})^{-n}\mathbf{1}(X_{(1)} \ge \sqrt{\rho}).$$

We see that this is maximized at $\rho = X_{(1)}^2$, so the MLE for ρ is $\hat{\rho} = X_{(1)}^2$.

(d) Which estimator uses all the information in the sample about the parameter? Justify your answer.

The MLE is a function of the sufficient statistic $X_{(1)}$, so it uses all the information in the sample about the parameter ρ. The sufficiency of $X_{(1)}$ comes from the factorization theorem. The method of moments estimator, which is a function of $\bar{X}_n$, is not based on a sufficient statistic. If we keep the value of $\bar{X}_n$ and erase the values of $X_1, \dots, X_n$, we will lose information about ρ; not so with $X_{(1)}$. If we keep the value of $X_{(1)}$ and throw away the values $X_1, \dots, X_n$, we will not lose any information about ρ.