Practice Prelim 1, STAT 611, Spring 2021

In the solutions, we use the notation:

$$\widehat{m}_1 := \sum_{i=1}^{n} X_i, \qquad \widehat{m}_2 := \sum_{i=1}^{n} X_i^2.$$

1. Let $X_1, \ldots, X_n$ be iid random variables from a half normal distribution $HN(\mu, \sigma^2)$. The PDF is $f_X(x) = \sqrt{2/(\pi\sigma^2)} \exp\{-(x-\mu)^2/(2\sigma^2)\}$ for $x > \mu$ and $f_X(x) = 0$ for $x \leq \mu$.

   (a) Find a sufficient statistic for $(\mu, \sigma^2)^T$. Briefly justify your answer.
   $(X_1, \ldots, X_n)^T$ itself is a sufficient statistic by the definition.
   To be more interesting,

$$\begin{aligned} f(X) &= \left(\frac{2}{\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^n (X-\mu)^2}{-2\sigma^2}} I(X_{(1)} > \mu) \\ &= \left(\frac{2}{\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^n X_i^2 + n\mu^2 - 2\mu\sum_{i=1}^n X_i}{-2\sigma^2}} I(X_{(1)} > \mu). \end{aligned}$$

   $(\widehat{m}_2, \widehat{m}_1, X_{(1)})^T$ is a three dimensional sufficient statistic.

   (b) Find a minimum sufficient statistic $(\mu, \sigma^2)^T$. Briefly justify your answer.
   $(\widehat{m}_2, \widehat{m}_1, X_{(1)})^T$ is also a minimum sufficient statistic. Because

$$\begin{aligned} & \left(\frac{2}{\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^n X_i^2 + n\mu^2 - 2\mu\sum_{i=1}^n X_i}{-2\sigma^2}} I(X_{(1)} > \mu) > 0 \\ \iff & \left(\frac{2}{\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^n Y_i^2 + n\mu^2 - 2\mu\sum_{i=1}^n Y_i}{-2\sigma^2}} I(Y_{(1)} > \mu) > 0 \end{aligned}$$

   implies $X_{(1)} = Y_{(1)}$.

$$\left(\frac{2}{\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^n X_i^2 + n\mu^2 - 2\mu\sum_{i=1}^n X_i}{-2\sigma^2}} I(X_{(1)} > \mu)/3 \left(\frac{2}{\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{i=1}^n Y_i^2 + n\mu^2 - 2\mu\sum_{i=1}^n Y_i}{-2\sigma^2}} I(Y_{(1)} > \mu)$$

   free of μ, σ implies $\sum_{i=1}^n X_i = \sum_{i=1}^n Y_i$, $\sum_{i=1}^n X_i^2 = \sum_{i=1}^n Y_i^2$.

   (c) Supposed μ is known, find a complete statistic and the UMVUE for σ^2. Justify your answer. $\sum_{i=1}^n (X_i - \mu)^2$ is a complete minimum sufficient statistic since this is now within the exponential family.

Also,

$$\begin{aligned}
\mathbb{E}_{\sigma^2}(X_i-\mu)^2 &= \int_{\mu}^{\infty}(x-\mu)^2\sqrt{\frac{2}{\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x \\
&= \int_{0}^{\infty} z^2\sqrt{\frac{2}{\pi\sigma^2}}\exp\left(-\frac{z^2}{2\sigma^2}\right)\mathrm{d}z \\
&= \int_{-\infty}^{\infty} z^2\sqrt{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{z^2}{2\sigma^2}\right)\mathrm{d}z \\
&= \sigma^2.
\end{aligned}$$

So the UMVUE for σ^2 is $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$.

2. Let $X_1,\ldots,X_n$ be iid random variables from the Pareto(a,b) distribution. The PDF is $f_X(x) = ba^b x^{-b-1}$ for $x > a$ and $f_X(x) = 0$ for $x \le a$. The parameters a, b satisfy $a > 0, b > 2$.

   (a) Find the moment estimator for a, b. Provide brief derivation.

   $$E(X) = \frac{ab}{b-1}, \qquad E(X^2) = \frac{ba^2}{(b-1)^2(b-2)} + \frac{a^2b^2}{(b-1)^2} = \frac{ba^2}{b-2}$$

   Thus, solve

   $$\widehat{m}_1 = \frac{ab}{b-1}, \qquad \widehat{m}_2 = \frac{ba^2}{b-2}.$$

   We obtain

   $$\widehat{a} = \frac{\widehat{m}_1\sqrt{\widehat{m}_2}}{\sqrt{\widehat{m}_2-\widehat{m}_1^2}+\sqrt{\widehat{m}_2}}, \qquad \widehat{b} = 1+\sqrt{\frac{\widehat{m}_2}{\widehat{m}_2-\widehat{m}_1^2}}.$$

   (b) Find the MLE for a, b. Provide brief derivation.

   $$l = n\log(b) + nb\log(a) - (b+1)\sum_{i=1}^{n}\log(x_i) + \sum_{i=1}^{n}\log I(x_i > a)$$

   Hence, $\widehat{a} = x_{(1)}$. From

   $$l_b' = n/b + n\log(a) - \sum_{i=1}^{n}\log(x_i) = 0,$$

   we obtain $\widehat{b} = n/\{\sum_{i=1}^{n}\log(x_i) - n\log(x_{(1)})\}$. Obviously $l_b'' < 0$. So this is MLE.

3. Let $X_1,\cdots,X_n$ be independent random variables, and $X_i \sim$ Exponential($scale = \lambda$), i.e., $f(x) = \dfrac{1}{\lambda}\exp(-x/\lambda)$, for $x \ge 0$

(a) Find a MLE for λ. The log-likelihood function is

$$\log L(\lambda; \mathbf{x}) = -n \log \lambda - \frac{1}{\lambda} \sum_{i=1}^{n} x_i.$$

Setting the score function to 0 gives:

$$\frac{\partial}{\partial \lambda} \log L(\lambda; \mathbf{x}) = -\frac{n}{\lambda} + \frac{1}{\lambda^2} \sum_{i=1}^{n} x_i = 0. \Rightarrow \hat{\lambda} = \bar{x}.$$

Note that

$$\frac{\partial^2}{\partial \lambda^2} \log L(\lambda; \mathbf{x}) = \frac{n}{\lambda^2} - \frac{2}{\lambda^3} \sum_{i=1}^{n} x_i,$$

which is equal to $-n/(\bar{x})^2 < 0$ at $\lambda = \hat{\lambda}$. This means the log-likelihood function has a local maximum at $\hat{\lambda}$. Since $\hat{\lambda} = \bar{x}$ is the unique stationary point, it is also the global argmax. That is $\hat{\lambda} = \bar{x}$ is the MLE for λ.

(b) Show that the MLE is an unbiased estimator for λ .

Easy to see $\mathbb{E}_\lambda(X_i) = \lambda$ for all i, then the unbiasedness follows.

(c) Calculate the mean squared error of the MLE.

Since the MLE is unbiased, we have

$$MSE_\lambda(\bar{X}) = \text{Var}_\lambda(\bar{X}) = \frac{\lambda^2}{n}.$$

(d) Find the Fisher information $I(\lambda)$ for n observations.

From the log-likelihood function, we have

$$\begin{aligned} I(\lambda) &= -\mathbb{E}_\lambda \left(\frac{\partial^2}{\partial \lambda^2} \log L(\lambda; \mathbf{x}) \right) \\ &= -\frac{n}{\lambda^2} + \frac{2}{\lambda^3} \mathbb{E}_\lambda (\sum_{i=1}^{n} X_i) = \frac{n}{\lambda^2}. \end{aligned}$$

(e) Find a sufficient statistics for λ.

We can write the joint pdf in the form of a one-parameter exponential family:

$$\exp\left(-n \log \lambda - \frac{1}{\lambda} \sum_{i=1}^{n} x_i \right).$$

So the statistic $\bar{X}$ is sufficient.

(f) Find the MLE for $P(X_1 < 1)$.

Note that $P(X_1 < 1) = 1 - e^{-1/\lambda}$. So by the invariance property of MLE, the MLE for $P(X_1 < 1)$ is

$$1 - \exp\left(-\frac{1}{\bar{X}} \right).$$