Mathematical Statistics I
Assignment 7 - Solutions

1. a. For one observation (X, Y) we have

$$I(\theta)=-\mathrm{E}\left(\frac{\partial^2}{\partial \theta^2} \log f(X, Y \mid \theta)\right)=-\mathrm{E}\left(-\frac{2 Y}{\theta^3}\right)=\frac{2 \mathrm{E} Y}{\theta^3} .$$

But, $Y \sim$ exponential (θ), and $\mathrm{E} Y=\theta$. Hence, $I(\theta)=2 / \theta^2$ for a sample of size one, and $I(\theta)=2 n / \theta^2$ for a sample of size n.
b. (i) The cdf of T is

$$P(T \leq t)=P\left(\frac{\sum_i Y_i}{\sum_i X_i} \leq t^2\right)=P\left(\frac{2 \sum_i Y_i / \theta}{2 \sum_i X_i \theta} \leq t^2 / \theta^2\right)=P\left(F_{2 n, 2 n} \leq t^2 / \theta^2\right)$$

where $F_{2 n, 2 n}$ is an F random variable with $2 n$ degrees of freedom in the numerator and denominator. This follows since $2 Y_i / \theta$ and $2 X_i \theta$ are all independent exponential(1), or χ_2^2. Differentiating (in t) and simplifying gives the density of T as

$$f_T(t)=\frac{\Gamma(2 n)}{\Gamma(n)^2} \frac{2}{t}\left(\frac{t^2}{t^2+\theta^2}\right)^n\left(\frac{\theta^2}{t^2+\theta^2}\right)^n,$$

and the second derivative (in θ) of the log density is

$$2 n \frac{t^4+2 t^2 \theta^2-\theta^4}{\theta^2\left(t^2+\theta^2\right)^2}=\frac{2 n}{\theta^2}\left(1-\frac{2}{\left(t^2 / \theta^2+1\right)^2}\right),$$

and the information in T is

$$\frac{2 n}{\theta^2}\left[1-2 \mathrm{E}\left(\frac{1}{T^2 / \theta^2+1}\right)^2\right]=\frac{2 n}{\theta^2}\left[1-2 \mathrm{E}\left(\frac{1}{F_{2 n, 2 n}^2+1}\right)^2\right] .$$

The expected value is

$$\mathrm{E}\left(\frac{1}{F_{2 n, 2 n}^2+1}\right)^2=\frac{\Gamma(2 n)}{\Gamma(n)^2} \int_0^{\infty} \frac{1}{(1+w)^2} \frac{w^{n-1}}{(1+w)^{2 n}}=\frac{\Gamma(2 n)}{\Gamma(n)^2} \frac{\Gamma(n) \Gamma(n+2)}{\Gamma(2 n+2)}=\frac{n+1}{2(2 n+1)} .$$

Substituting this above gives the information in T as

$$\frac{2 n}{\theta^2}\left[1-2 \frac{n+1}{2(2 n+1)}\right]=I(\theta) \frac{n}{2 n+1},$$

(ii) Let $W=\sum_i X_i$ and $V=\sum_i Y_i$. In each pair, X_i and Y_i are independent, so W and V are independent. $X_i \sim$ exponential$(1 / \theta)$; hence, $W \sim$ gamma$(n, 1 / \theta) . Y_i \sim$ exponential(θ); hence, $V \sim$ gamma(n, θ). Use this joint distribution of (W, V) to derive the joint pdf of (T, U) as

$$f(t, u \mid \theta)=\frac{2}{[\Gamma(n)]^2 t} u^{2 n-1} \exp \left(-\frac{u \theta}{t}-\frac{u t}{\theta}\right), \quad u>0, \quad t>0 .$$

Now, the information in (T, U) is

$$-\mathrm{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(T, U \mid \theta)\right) = -\mathrm{E}\left(-\frac{2UT}{\theta^3}\right) = \mathrm{E}\left(\frac{2V}{\theta^3}\right) = \frac{2n\theta}{\theta^3} = \frac{2n}{\theta^2}.$$

(iii) The pdf of the sample is $f(\mathbf{x}, \mathbf{y}) = \exp\left[-\theta\left(\sum_i x_i\right) - \left(\sum_i y_i\right)/\theta\right]$. Hence, (W, V) defined as in part (ii) is sufficient. (T, U) is a one-to-one function of (W, V), hence (T, U) is also sufficient. But, $\mathrm{E}U^2 = \mathrm{E}WV = (n/\theta)(n\theta) = n^2$ does not depend on θ. So $\mathrm{E}\left(U^2 - n^2\right) = 0$ for all θ, and (T, U) is not complete.

2. Let $T(x) = \left(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i^3, \sum_{i=3}^n x_i^4\right)$ for any $x = (x_1, \ldots, x_n)$ and let $\eta(\theta) = \sigma^{-4}\left(-4\mu^3, 6\mu^2, -4\mu, 1\right)$. The joint density of $(X_1, \ldots, X_n)$ is

$$f_\theta(x) = \exp\left\{[\eta(\theta)]^\tau T(x) - n\mu^4/\sigma^4 - n\xi(\theta)\right\},$$

which belongs to an exponential family. For any two sample points $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$,

$$\begin{aligned}\frac{f_\theta(x)}{f_\theta(y)} = \exp\Bigg\{-\frac{1}{\sigma^4}\Bigg[&\left(\sum_{i=1}^n x_i^4 - \sum_{i=1}^n y_i^4\right) - 4\mu\left(\sum_{i=1}^n x_i^3 - \sum_{i=1}^n y_i^3\right)\\ &+6\mu^2\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2\right) - 4\mu^3\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\Bigg]\Bigg\}\end{aligned}$$

which is free of parameter (μ, σ) if and only if $T(x) = T(y)$. Hence, $T(X)$ is minimal sufficient for θ.

3. We're going to use the following theorem. Let $\mathcal{U}$ be the set of all unbiased estimators of 0 with finite variances and T be an unbiased estimator of ϑ with $E\left(T^2\right) < \infty$. A necessary and sufficient condition for $T(X)$ to be a UMVUE of ϑ is that $E[T(X)U(X)] = 0$ for any $U \in \mathcal{U}$ and any $P \in \mathcal{P}$.
Let U be an unbiased estimator of 0. Since T is a UMVUE of $\theta, E(TU) = 0$ for any P, which means that TU is an unbiased estimator of 0 . Then $E\left(T^2U\right) = E[T(TU)] = 0$ if $E\left(T^4\right) < \infty$. By the mentioned theorem, T^2 is a UMVUE of $E\left(T^2\right)$. By repeating the argument, we can show that T^k is a UMVUE of $E\left(T^k\right)$.

4. For all $\epsilon > 0$, by Chebyshev's inequality,

$$P\left(|W_n - \mu| \geq \epsilon\right) \leq \frac{b}{n^p\epsilon^2} \to 0,$$

as $n \to \infty$. Hence by definition W_n converges in probability to μ i.e. W_n is a consistent estimator of μ.

5. (i) Let $\ell(\theta)$ be the likelihood function. Note that

$$s(\theta) = \frac{\partial\log\ell(\theta)}{\partial\theta} = \sum_{i=1}^n \frac{f_1(X_i) - f_2(X_i)}{f_2(X_i) + \theta\left[f_1(X_i) - f_2(X_i)\right]},$$

which has derivative

$$s'(\theta) = -\sum_{i=1}^n \frac{\left[f_1(X_i) - f_2(X_i)\right]^2}{\left\{f_2(X_i) + \theta\left[f_1(X_i) - f_2(X_i)\right]\right\}^2} < 0.$$

Therefore, $s(\theta) = 0$ has at most one solution. The necessary and sufficient condition that $s(\theta) = 0$ has a solution (which is unique if it exists) is that $\lim_{\theta \to 0} s(\theta) > 0$ and $\lim_{\theta \to 1} s(\theta) < 0$, which is equivalent to

$$\sum_{i=1}^{n} \frac{f_1(X_i)}{f_2(X_i)} > n \quad \text{and} \quad \sum_{i=1}^{n} \frac{f_2(X_i)}{f_1(X_i)} > n.$$

The solution, if it exists, is the MLE since $s'(\theta) < 0$.
(ii) If $\sum_{i=1}^{n} \frac{f_2(X_i)}{f_1(X_i)} \leq n$, then $s(\theta) \geq 0$ and $\ell(\theta)$ is nondecreasing and, thus, the MLE of θ is 1. Similarly, if $\sum_{i=1}^{n} \frac{f_1(X_i)}{f_2(X_i)} \leq n$, then the MLE of θ is 0.