Foundations of Statistical Inference, BS2a, Exercises 1

1. Let $X_1, \dots, X_n$ be a random sample from a distribution with density

$$f(x;\theta) = \frac{1}{2}\theta^3 x^2 e^{-\theta x},\ x > 0.$$

(a) Rewrite the density in standard exponential form, giving $A(\theta)$, $B(X)$, $C(X)$, $D(\theta)$ explicitly.
(b) Find a minimal sufficient statistic for θ. Find the expected value of the statistic. Is it complete?
(c) Find the maximum likelihood estimator for θ. Is it unbiassed for θ?
(d) Show that $\theta^* = (2/n)\sum_{i=1}^n X_i^{-1}$ is an unbiased estimator of θ and find its variance.
(e) Find the Cramér-Rao lower bound for the variance of unbiased estimators of θ and show that the variance of θ^* is strictly larger than this bound.
(f) What does Lehmann-Scheffé Theorem tells us?

2. Let $X_1, \dots, X_n$ be independent Poisson random variables with means $\mathbb{E}(X_i) = \lambda m_i$, $i = 1, \dots, n$ where $\lambda > 0$ is unknown and $m_1, \dots, m_n$ are known constants.
(a) Show that the model defines a canonical exponential family with canonical parameter $\theta = \log\lambda$.
(b) What is the canonical sufficient statistic? Find its mean and variance.
(c) Find the MLE $\widehat{\theta}$ of θ.
(d) Find the Fisher information for θ. What statements can you make about the variance of $\widehat{\theta}$.

3. Let $X_1, \dots, X_n$ be a random sample from the density

$$f(x;\theta) = e^{-(x-\theta)},\ x > \theta$$

(a) Show that the MLE $\widehat{\theta}$ of θ is the minimum of $X_1, \dots, X_n$.
(b) By finding the density function of $\widehat{\theta}$ show that $\widehat{\theta}$ is a consistent, but biassed estimator of θ with $\mathbb{E}[\widehat{\theta}] = \theta + 1/n$. Suggest an unbiassed and consistent estimator and find its variance.
(c) Compare the sampling properties of the MLE with those of the method of moments estimator. Is it appropriate to compare the variances of these estimators with that suggested by the Cramér-Rao inequality?

4. Let $X_1, \dots, X_n$ be a sample from $N(\mu, \sigma^2)$.

(a) Show that the MLE of σ^2 is

$$\widehat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

(b) Show that $\widehat{\sigma}^2$ has a smaller mean square error than

$$(n-1)^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

(c) For which value of a is the MSE of

$$(n+a)^{-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

the smallest. Hint: For (b) and (c) you will need to find $\mathrm{Var}(\chi^2_{n-1})$ which is a special case of the variance of a gamma distribution.

5. Let $X_1, \ldots, X_n$ be a random sample from a truncated Poisson distribution with distribution

$$f(x;\lambda) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \cdot \frac{\lambda^x}{x!},\ x = 1, 2, \ldots.$$

For $i = 1, \ldots, n$ a random variable Z_i is defined by

$$Z_i = X_i \text{ if } X_i \geq 2 \text{ or } Z_i = 0 \text{ if } X_i = 1$$

Show that $\bar{Z}$ is an unbiased estimator of λ with efficiency

$$\frac{1-e^{-\lambda}}{1-\left(\frac{\lambda e^{-\lambda}}{1-e^{-\lambda}}\right)^2}.$$

6. (a) (optional bookwork) Let X be a discrete random variable with pmf $f(x;\theta)$ with parameter $\theta \in \Theta$ and sample space $X \in \chi$. Let $T(x)$ be a function of x. Suppose $f(x;\theta)/f(y;\theta)$ is not a function of θ if and only if $T(x) = T(y)$. Show that $T(x)$ is minimal sufficient for θ.

(b) Let $N = N(0, S]$ be the number of events in a Poisson arrival process of rate λ acting over time s in the interval $0 < s \leq S$. Suppose we observe arrivals in the process at times $X_1, X_2, \ldots, X_N$, and wish to use these data to estimate λ. Show that N is minimal sufficient for λ (assume the result in (a) holds for any sufficiently regular family of probability distributions).

7. The random variable X has a discrete distribution such that $P(X = r) = \theta^{-1}$ for $r = 1, 2, \ldots, \theta$, where θ is an unknown positive integer. Show that Y, the maximum of a sample of n independent observations of X, is a complete sufficient statistic for θ, and hence verify that

$$\frac{Y^{n+1} - (Y-1)^{n+1}}{Y^n - (Y-1)^n}$$

is a minimum-variance unbiased estimator for θ.

8. Consider a binomial experiment with probability of success p in which m fixed trials are conducted, resulting in R successes; a further set of trials is then conducted until s (fixed) further successes have occurred. The number of trials necessary in the second set is a random variable N. By considering the function

$$U(R, N) = \frac{R}{m} - \frac{s-1}{N-1}$$

show that (R, N) are jointly sufficient for p, but not complete.

9. (GJJ 2.19) The random variables $X_1, X_2, ..., X_n$ are iid with density $f(x;\theta) = \theta x^{\theta-1}$ for $0 < x < 1$ and $\theta > 0$ unknown.

(i) Find a sufficient statistic T for θ.

(ii) Given that $-\log(X_1)$ is unbiased for θ^{-1}, find another unbiased estimator with smaller variance.Give a simple expression of this estimator involving T.