MAST20005/MAST90058: Week 3 Problems

1. (a) Let $X_1, \dots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$ where $-\infty < \mu < \infty$ and $\sigma^2 > 0$. Assume that σ^2 is known (i.e. it is a fixed, known value). Show the maximum likelihood estimator of μ is $\hat{\mu} = \bar{X}$.

   (b) A random sample $X_1, \dots, X_n$ of size n is taken from a Poisson distribution with mean $\lambda > 0$.

   i. Show the maximum likelihood estimator of λ is $\hat{\lambda} = \bar{X}$.

   ii. Suppose with $n = 40$ we observe 5 zeros, 7 ones, 12 twos, 9 threes, 5 fours, 1 five, and 1 six. What is the maximum likelihood estimate of λ?

   (c) Let $X_1, \dots, X_n$ be random samples from the following probability density functions. In each case find the maximum likelihood estimator $\hat{\theta}$.

   i. $f(x \mid \theta) = \frac{1}{\theta^2} x \exp(-x/\theta), \quad 0 < x < \infty, \quad 0 < \theta < \infty$

   ii. $f(x \mid \theta) = \frac{1}{2\theta^3} x^2 \exp(-x/\theta), \quad 0 < x < \infty, \quad 0 < \theta < \infty$

   iii. $f(x \mid \theta) = \frac{1}{2} \exp(-|x - \theta|), \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$

   Hint: The last part involves minimizing $\sum_{i=1}^{n} |x_i - \theta|$, which is tricky. Try $n = 5$ and the sample $\{6.1, \text{-}1.1, 3.2, 0.7, 1.7\}$. Then deduce the MLE in general.

2. Consider a random sample of n observations on X having the following pmf:

x	0	1	2
$p(x)$	$1-\theta$	$3\theta/4$	$\theta/4$

   (a) Find two unbiased estimators for θ: one based on $\bar{X}$, and one based on $Z = \mathrm{freq}(0) = \sum_{i=1}^{n} I(X_i = 0)$ (i.e. the frequency of $x = 0$).

   (b) Compare the two estimators above in terms of their variance.

3. Let $f(x \mid \theta) = \theta x^{\theta - 1}$, $0 < x < 1$, $0 < \theta < \infty$ and let $X_1, \dots, X_n$ denote a random sample from this distribution. Note that,

$$\int_0^1 x\theta\, x^{\theta-1} dx = \frac{\theta}{\theta + 1}.$$

   (a) Sketch the pdf of X for $\theta = 1/2$ and $\theta = 2$.

   (b) Show that $\hat{\theta} = -n / \left(\sum_{i=1}^{n} \ln X_i\right)$ is the maximum likelihood estimator of θ.

   (c) For each of the following three sets of observations from this distribution, compute the maximum likelihood estimates and the method of moments estimates.

X	Y	Z
0.0256	0.9960	0.4698
0.3051	0.3125	0.3675
0.0278	0.4374	0.5991
0.8971	0.7464	0.9513
0.0739	0.8278	0.6049
0.3191	0.9518	0.9917
0.7379	0.9924	0.1551
0.3671	0.7112	0.0710
0.9763	0.2228	0.2110
0.0102	0.8609	0.2154

   ($\sum_{i=1}^{n} \ln(x_i) = -18.2063$, $\sum_{i=1}^{n} \ln(y_i) = -4.5246$, $\sum_{i=1}^{n} \ln(z_i) = -10.42968$, $\sum_{i=1}^{n} x_i = 3.7401$, $\sum_{i=1}^{n} y_i = 7.0592$, $\sum_{i=1}^{n} z_i = 4.6368$)

4. Let $X_1, \ldots, X_n$ be a random sample from the exponential distribution whose pdf is $f(x \mid \theta) = (1/\theta)\exp(-x/\theta)$, $0 < x < \infty$, $0 < \theta < \infty$.

   (a) Show that $\bar{X}$ is an unbiased estimator of θ.

   (b) Show that the variance of $\bar{X}$ is θ^2/n.

   (c) Calculate an estimate of θ if a random sample gave the following values:

   ```
   3.5  8.1  0.9  4.4  0.5
   ```

5. Let $X_1, \ldots, X_n$ be a random sample from a distribution having finite variance σ^2. Show that

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

   is an unbiased estimator of σ^2.

   Hint: Use a result from question 4(a)(i) from week 2 to derive an alternative expression for S^2 and then compute $E(S^2)$.

6. Let $X_1, \ldots, X_n$ be iid observations from $\mathrm{N}(0, \theta^2)$. Consider the estimators S^2 and $\hat{\theta}^2 = n^{-1}\sum_{i=1}^{n} X_i^2$. Show that $\hat{\theta}^2$ is unbiased and $\mathrm{var}(\hat{\theta}^2) < \mathrm{var}(S^2)$ for any $n > 1$.

7. Let $X_1, \ldots, X_n$ be iid observations from $X \sim \mathrm{N}(\mu, \sigma^2)$. Since X has a symmetric pdf, we might expect that both the sample mean $\bar{X}$ and the sample median $\hat{\pi}_{0.5}$ will be good estimators of the population mean μ.

   (a) Find the variance of $\bar{X}$.

   (b) In general, the sample median will approximately follow a normal distribution, $\hat{\pi}_{0.5} \sim \mathrm{N}\left(\pi_{0.5}, \pi/2 \times \sigma^2/n\right)$, where $\pi_{0.5}$ is the true median (we will learn more about this later in the semester). How does the variance of the sample median compare with that of the sample mean?
   (n.b. $\pi/2$ is just the usual mathematical constant π divided by 2, it is not the same as the median $\pi_{0.5}$.)

   (c) Are the estimators biased?

   (d) Which estimator do you expect to be more accurate?