MA20226 - Question Sheet Two

Simon Shaw, s.shaw@bath.ac.uk
http://people.bath.ac.uk/masss/ma20226.html

2011/12 Semester I

Set: Tutorials of Week 4, commencing Monday 24th October 2011.
Due in: Tutorials of Week 5, commencing Monday 31st October 2011. You should write both your name and tutorial group on your work. If you are unable to make your practical then you should ensure that you hand the work to me personally *before* the practical is held; my office is 4W4.10.
Task: Attempt questions 1-3; questions 4-6 are additional questions and may be used for tutorial discussion.

1. The independent observations $X_1, \ldots, X_n$ are assumed to come from a $N(\mu, \sigma^2)$ distribution. In the case when μ and σ^2 are unknown, we showed, in Lecture 4, that the maximum likelihood estimators were $\overline{X}$ for μ, which is unbiased, and $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ for σ^2, which is biased.
   (a) In the case where σ^2 is known and μ is unknown, find the maximum likelihood estimator of μ. Is the estimator biased?
   (b) In the case where μ is known and σ^2 is unknown, find the maximum likelihood estimator of σ^2. Is the estimator biased?

2. A random quantity X has a uniform distribution on the interval $(0, \theta)$ so that its pdf is given by

$$f(x \,|\, \theta) = \begin{cases} \frac{1}{\theta} & 0 < x \leq \theta, \\ 0 & \text{otherwise.} \end{cases}$$

   A random sample $X_1, \ldots, X_n$ is drawn from this distribution in order to learn about the value of θ.
   (a) Show that the joint pdf of the X_i is given by

$$f(x_1, \ldots, x_n \,|\, \theta) = \begin{cases} \frac{1}{\theta^n} & 0 < m \leq \theta, \\ 0 & \text{otherwise,} \end{cases}$$

   where $m = \max\{x_1, \ldots, x_n\}$.
   (b) Sketch the likelihood function, $L(\theta)$, of θ. Hence, explain why the maximum likelihood estimate of θ, $\hat{\theta} = m$.
   (c) Derive the exact sampling distribution of $M = \max\{X_1, \ldots, X_n\}$. [**Hint:** To calculate the sampling distribution of M, first calculate its cdf, $P(M \leq m \,|\, \theta)$, by noting that $M \leq m$ if and only if all of the X_i are less than or equal to m. Then, differentiate to give the pdf, $f(m \,|\, \theta)$. Recall that we gave the sampling distribution in Lecture 4 where we went on to show that $E(M \,|\, \theta) = (1 - \frac{1}{n+1})\theta$ so that M "under-estimates" θ. On Practical Sheet 1 you calculate $Var(M \,|\, \theta)$.]

3. If $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ random quantities, then an unbiased estimator of σ^2 is

$$S^2 \quad = \quad \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2.$$

This estimator has variance $2\sigma^4/(n-1)$.

(a) Write down the estimator's MSE.

(b) Consider a second estimator constructed as αS^2 where α is any positive constant.

i. Find the expectation and variance of αS^2, and hence its MSE.

ii. Find the value of α which minimises this mean square error.

iii. What is the efficiency of S^2 relative to the smallest MSE αS^2?

4. Let $x_1, \ldots, x_n$ be a random sample from a geometric distribution

$$P(X = x \,|\, p) \quad = \quad (1-p)p^{x-1}, \; x = 1, 2, \ldots$$

where $p \in (0, 1)$ so that X is the number of trials until the first failure in a sequence of independent trials with success probability p.

(a) Explain why the sample average $\overline{x}$ and sample size n are sufficient for calculating the likelihood function.

(b) Find the maximum likelihood estimator of p.

(c) Show that $P(X > x \,|\, p) = p^x$.

5. As part of a quality control procedure for a certain mass production process, batches containing very large numbers of components from the production are inspected for defectives. We will assume the process is in equilibrium so that each component is independent and is either acceptable, with probability p, or defective, with probability $q = 1 - p$.

The inspection procedure is as follows. During each shift n batches are selected from the production and for each such batch components are inspected until a defective one if found, and the number of inspected components is recorded. At the end of the shift, there may be some inspected batches which have not yet yielded a defective component; and for such batches the number of inspected components is recorded.

Suppose that at the end of one such inspection shift, a defective component was detected in each of r of the batches, the recorded numbers of inspected batches being $x_1, \ldots, x_r$. Inspection of the remaining $s = n - r$ batches was incomplete, the recorded numbers of inspected components being $c_1, \ldots, c_s$.

(a) By considering question 4, argue that the likelihood for $q = 1 - p$ based on these data is $L(q) = q^r (1-q)^{x+c-r}$, where $x = \sum_{i=1}^{r} x_i$ and $c = \sum_{i=1}^{s} c_i$.

(b) Show that the maximum likelihood estimate of q is $\hat{q} = 1/a$, where $a = (x + c)/r$. Interpret a.

6. The independent observations $x_1, \ldots, x_{10}$ are assumed to come from a $N(\mu, \sigma^2)$ distribution, and $x_{11}, \ldots, x_{15}$ from a $N(2\mu, \sigma^2/2)$ distribution, where σ^2 is assumed known.

(a) Write down the joint probability density function of $X_1, \ldots, X_{10}$, and the joint probability density function of $X_{11}, \ldots, X_{15}$. Hence write down the likelihood function for μ based on all 15 observations.

(b) Find the maximum likelihood estimator of μ and determine its bias and its mean square error.