MA20033 - Question Sheet Two

Simon Shaw
s.c.shaw@maths.bath.ac.uk

2004/05 Semester I

Attempt questions 1-3; questions 4-6 are for tutorial discussion. If you wish to attempt these, I'm happy to mark them. Hand in by 5.00pm Friday 29 October 2004 either to me, or the envelope on my door, 1W4.8.

1. In Lecture 3, we showed that if $X_1, \ldots, X_n$ are independently and identically distributed as $Po(\lambda)$, and we observe $X_1 = x_1, \ldots, X_n = x_n$, then the maximum likelihood estimate of λ is $\hat{\lambda} = \overline{x}$.

   (a) Write down the probability mass function of X_i under the reparameterisation $\nu = 1/\lambda$ (i.e. change all the λ to $1/\nu$).

   (b) Under this reparameterisation, what is the likelihood function, $L(\nu)$, given the data $X_1 = x_1, \ldots, X_n = x_n$?

   (c) Thus, find the maximum likelihood estimate, $\hat{\nu}$, of ν based on this data. [You should find that $\hat{\nu} = 1/\hat{\lambda}$. This is an example of the invariance property of maximum likelihood estimates: if $\hat{\theta}$ is the maximum likelihood estimate of θ and $u(\theta)$ a function of θ, then $u(\hat{\theta})$ is the maximum likelihood estimate of $u(\theta)$.]

2. The independent observations $x_1, \ldots, x_n$ are assumed to come from a $N(\mu, \sigma^2)$ distribution.

   (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?

   (c) In the case when μ and σ^2 are unknown, we showed, in Lecture 3, that the respective maximum likelihood estimators were $\overline{X}$ and $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Are these estimators biased?

3. 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, $f(m|\theta)$.]

(d) Thus, show that $E(M|\theta) = (1 - \frac{1}{n+1})\theta$. Is it surprising that the maximum likelihood estimate "under-estimates" θ? Provide an unbiased estimator of θ.

4. The independent observations $x_1, \ldots, x_n$ are assumed to come from a $U(a, b)$ distribution. Find the maximum likelihood estimates of a and b.

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

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

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 $P(X > x|p) = p^x$.

6. 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 5, 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.