Lent Term 2000 Richard Weber

Statistics Examples Sheet 1

This examples sheet covers the material of the first 5 lectures and is appropriate for your first supervision. There will be two further examples sheets and a sheet of supplementary questions. A copy of this sheet can be found from the WWW site: `http://www.statslab.cam.ac.uk/~rrw1/stats/`

1. (Lecture 1, unbiased estimation) Suppose X_1, X_2 are independent samples from $B(1,p)$. Let $T = X_1 + X_2$. In cases (a)–(c) show that $\hat\theta$ is an unbiased estimator of θ. Prove the statement made in case (d).

(a) $\theta = 1997 - p$, $\hat\theta = 1997 - \frac{1}{2}T$.

(b) $\theta = (1-p)^2$, $\hat\theta = 1$ if $T = 0$ and $\hat\theta = 0$ otherwise.

(c) $\theta = (1-3p)^2$, $\hat\theta = (-2)^T$.

(d) $\theta = \left(1 - \frac{1}{2}p\right)^{-1}$, there is no unbiased estimator of θ.

Hint: Note that $T \sim B(2,p)$ and $\mathbb{E}\hat\theta(T) = (1-p)^2\hat\theta(0) + 2p(1-p)\hat\theta(1) + p^2\hat\theta(2)$.

You should note from this example that an unbiased estimator can be silly (as in case (c) where $\hat\theta = -2$ when $T = 1$ even though we know $\theta > 0$), or may not even exist (as in case (d)).

2. (Lecture 2, MLE) In a genetics experiment, a sample of n individuals was found to include a, b, c of the three possible genotypes GG, Gg, gg respectively. The population frequency of a gene of type G is $\theta/(\theta+1)$, where θ is unknown, and it is assumed that the individuals are unrelated and that two genes in a single individual are independent. Show that the likelihood of θ is proportional to

$$\theta^{2a+b} / (1+\theta)^{2a+2b+2c}$$

and that the maximum likelihood estimate of θ is $(2a+b)/(b+2c)$.

3. (Lecture 2, MLE and sufficiency) Suppose $X_1, \dots, X_n$ is a random sample from a gamma (α, λ)distribution with density function

$$f(x \mid \alpha, \lambda) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0.$$

Let $\theta = (\alpha, \lambda)$. What is meant by saying that $T(X)$ is sufficient for θ? Find a sufficient statistic for θ. How might you find MLEs for α and λ?

Hint. In this example the sufficient statistic is a vector with two components.

4. (Lecture 2, MLE and sufficiency) In each of cases (a)–(c) write down the likelihood of θ and show that the stated $T(X)$ is a sufficient statistic for θ.

In each case also find a MLE of θ and show that it is a function of $T(X)$. Find the distribution of $T(X)$ and determine whether or not the MLE is an unbiased estimator of θ. If it is not, verify that it is asymptotically unbiased, and find some other estimator which is unbiased.

(a) $X_1, \dots, X_n$ are independent Poisson random variables, with X_i having mean $i\theta$, where $\theta > 0$. $T(X) = \sum_{i=1}^n X_i$.

(b) $X_1, \dots, X_n$ are independent normal random variables, with $X_i \sim N(\theta, \sigma_i^2)$ and σ_i^2, $i = 1, \dots, n$, known. $T(X) = \sum_{i=1}^n X_i/\sigma_i^2$.

(c) $X_1, \dots, X_n$ are $n > 2$ independent and exponentially distributed random variables, with parameter θ, i.e., with density $f(x \mid \theta) = \theta e^{-\theta x}$, $x > 0$. $T(X) = \sum_{i=1}^n X_i$.

Hint: In case (a), $T(X) \sim P\left(\frac{1}{2}n(n+1)\theta\right)$. In case (b), $T(X) \sim N\left(\theta \sum_i \sigma_i^{-2}, \sum_i \sigma_i^{-2}\right)$. In case (c), $T(X) \sim \text{gamma}(n, \theta)$. Do you understand why?

5. (Lecture 3, Rao-Blackwell theorem) Suppose $X_1, \dots, X_n$ are independent random variables with distribution $B(1,p)$.

(a) Show that a sufficient statistic for $\theta = (1-p)^2$ is $T(X) = \sum_{i=1}^n X_i$ and that the MLE for θ is $\left(1 - \frac{1}{n}T\right)^2$.

Hint: use the chain rule, $df/d\theta = (df/dp)(dp/d\theta)$.

(b) The MLE is a biased estimator for θ. Find a function of T which is an unbiased estimator for θ.

Hint: $\theta = \mathbb{P}(X_1 + X_2 = 0)$. Recall example 1 (b) above.

6. (Lecture 3, Rao-Blackwell theorem) Suppose $X_1, \dots, X_n$ are independent random variables uniformly distributed over $(\theta, 2\theta)$. Show that a sufficient statistic for θ is $T(X) = (\min_i X_i, \max_i X_i)$ and that an unbiased estimator of θ is $\hat\theta = \frac{2}{3}X_1$. Find an unbiased estimator of θ which is a function of $T(X)$ and whose mean square error is no more than that of $\hat\theta$.

Note that this is another example in which the sufficient statistic turns out to be a vector, despite the fact that the parameter θ is only a scalar.

7. (Lecture 4, confidence intervals) A random variable is uniformly distributed over $(0, \theta)$. Show that the maximum of a random sample of n values of this variable is sufficient for θ and that this is also the MLE for θ. Show also that a $100\gamma\%$ confidence interval for θ is $(y_n, y_n/(1-\gamma)^{1/n})$, y_n being the maximum of the sample.

8. (Lecture 4, confidence intervals) Suppose that $X_1 \sim N(\theta_1, 1)$ and $X_2 \sim N(\theta_2, 1)$ independently, where θ_1 and θ_2 are unknown. For this model, $(\theta_1 - X_1)^2 + (\theta_2 - X_2)^2$ has the distribution $\mathcal{E}(\frac{1}{2})$, i.e., the exponential distribution with mean 2. (A fact you may recall from Probability IA, and which we will prove again later.)

Show that both the square S and circle C in $\mathbb{R}^2$, given by

$$S = \{(\theta_1, \theta_2) : |\theta_1 - X_1| \le 2.236;\ |\theta_2 - X_2| \le 2.236\}$$
$$C = \left\{(\theta_1, \theta_2) : (\theta_1 - X_1)^2 + (\theta_2 - X_2)^2 \le 5.991\right\}$$

are 95% confidence regions for (θ_1, θ_2), in the sense that $\mathbb{P}(S \text{ contains } (\theta_1, \theta_2)) = 0.95$ and $\mathbb{P}(C \text{ contains } (\theta_1, \theta_2)) = 0.95$. Hint: $\Phi(2.236) = (1 + \sqrt{.95})/2$, where Φ is the cdf of $N(0, 1)$.

Which of S and C would you prefer, and why?

9. (Lecture 5, Bayes estimation) Each word that baby Hamlet speaks is chosen independently and with equal probability from a set of k words. Suppose your prior belief is that k is equally likely to be either 5, 6, 7 or 8. You hear him say 'to not be or be to'. Show that the posterior probability mass function of k is proportional to $q(k) := (k-1)(k-2)(k-3)/k^5$, $k = 5, 6, 7, 8$, and is 0 otherwise.

Given that $q(k)$ has values 0.00768, 0.00772, 0.00714, 0.00641 for $k = 5, 6, 7, 8$ respectively, find a point estimate of k under the loss function

$$L(k, \hat{k}) = \begin{cases} 0 & \text{if } \hat{k} = k, \\ 1 & \text{if } \hat{k} \ne k. \end{cases}$$

How does this particular choice of prior distribution and loss function relate to maximum likelihood estimation?

10. (Lecture 5, Bayes estimation) Suppose that the number of defects on a roll of magnetic recording tape has a Poisson distribution for which the mean λ is known to be either 1 or 1.5. Suppose the prior mass function for λ is

$$p_\lambda(1) = 0.4, \quad p_\lambda(1.5) = 0.6\ .$$

A collection of 5 rolls of tape are found to have $x = (3, 1, 4, 6, 2)$ defects respectively. Show that the posterior distribution for λ is

$$f_\lambda(1 \mid x) = 0.012, \quad f_\lambda(1.5 \mid x) = 0.988\,.$$

You will have to use your calculator for this one.

11. (Lecture 5, Bayes estimation) Suppose $X_1, \dots, X_n$ are IID from a distribution uniform on $(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, and that the prior for θ is uniform on (10,20). Calculate the posterior distribution for θ, given $x = X_1, \dots, X_n$ and show that the point estimate for θ under both quadratic and absolute error loss functions is

$$\hat{\theta} = \tfrac{1}{2}\left[\max_i(x_i - \tfrac{1}{2}) \vee 10 + \min_i(x_i + \tfrac{1}{2}) \wedge 20\right].$$

The notation here is $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$.

12. (Lecture 5, Bayes estimation) Suppose $X_1, \dots, X_n$ form a random sample from the following pdf:

$$f(x \mid \theta) = \begin{cases} \theta x^{\theta - 1} & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

and that the prior for θ is gamma(α, β), $\alpha > 0, \beta > 0$, with density

$$p(\theta) = \frac{\beta^\alpha \theta^{\alpha-1} e^{-\beta\theta}}{\Gamma(\alpha)}, \quad \theta > 0\,.$$

Show that the posterior distribution of θ is gamma $(\alpha + n, \beta - \sum_i \log x_i)$ and hence that a point estimate for θ under quadratic loss function is

$$\frac{\alpha + n}{\beta - \sum_{i=1}^n \log x_i}.$$

Hint: You may want to refer to the notes for Lecture 1 to remind yourself of some basic facts about the gamma distribution.

13. (Lecture 5, Bayes estimation) Suppose that X is distributed as a binomial random variable $B(n, \theta)$. Suppose the prior distribution for θ is the uniform distribution on $[0, 1]$ and the loss function is

$$L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2/\theta(1 - \theta)\,.$$

Show that, based on the single observation x, the point estimate for θ is $\hat{\theta} = x/n$.

Hint: You may want to refer to the notes for Lecture 1 to remind yourself of some basic facts about the beta distribution. Recall

$$\int_0^1 x^{a-1}(1 - x)^{b-1}\, dx = B(a, b) := \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}$$

and that $\Gamma(a) = (a - 1)!$ when a is an integer.

May 11, 2000