MATH/STAT 3460, Intermediate Statistical Theory
Winter 2014
Toby Kenney
Midterm Examination
Model Solutions

Basic Questions

1. *The weight (in kg) of a certain species of rabbit is believed to follow a Normal distribution with mean 5 and variance $7e^{-2\sqrt{a}}$. Eight of these rabbits are collected, and their weights are measured as 4.9, 6.8, 3.6, 8.1, 2.3, 3.1, 6.4, and 4.2. What is the maximum likelihood estimate for a?*

   Let $\sigma = \sqrt{7}e^{-\sqrt{a}}$ be the standard deviation of the normal distribution. Then we have

$$\begin{aligned} l(\sigma) &= -8\log(\sigma) - \tfrac{(4.9-5)^2+(6.8-5)^2+(3.6-5)^2+(8.1-5)^2+(2.3-5)^2+(3.1-5)^2+(6.4-5)^2+(4.2-5)^2}{2\sigma^2} \\ &= -8\log(\sigma) - \tfrac{28.32}{2\sigma^2} \end{aligned}$$

   The score function is therefore, $S(\sigma) = \frac{28.32}{\sigma^3} - \frac{8}{\sigma}$. Setting this equal to zero gives $\sigma^2 = \frac{28.32}{8} = 3.54$. The MLE for a is therefore given by $7e^{-2\sqrt{a}} = 3.54$, or $a = \left(\frac{\log(0.50571)}{2}\right)^2 = 0.116207159$.

2. *In a trial for a new drug, the probability of a response to dose d is assumed to be $1 - \frac{1}{1+e^{\alpha+\beta d}}$ for some α and β. The data from a study of the drug are given in the following table:*

dose	*0*	*1*	*2*
number	*26*	*23*	*21*
number of responses	*3*	*12*	*21*

   (a) Show that $\alpha = -2.40315$ and $\beta = 2.78828$ is the maximum likelihood estimate for α and β

   The likelihood function is

$$L(\alpha,\beta) = \left(1-\tfrac{1}{1+e^{\alpha}}\right)^3 \left(\tfrac{1}{1+e^{\alpha}}\right)^{23} \left(1-\tfrac{1}{1+e^{\alpha+\beta}}\right)^{12} \left(\tfrac{1}{1+e^{\alpha+\beta}}\right)^{11} \left(1-\tfrac{1}{1+e^{\alpha+2\beta}}\right)^{21}$$

   So the score function is given by:

$$\frac{\partial l(\alpha,\beta)}{\partial\alpha} = 3\frac{1}{(1+e^{\alpha})} - 23\frac{e^{\alpha}}{(1+e^{\alpha})} + 12\frac{1}{(1+e^{\alpha+\beta})} - 11\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})} + 21\frac{1}{(1+e^{\alpha+2\beta})}$$

$$= 26\frac{1}{1+e^{\alpha}} - 23 + 23\frac{1}{1+e^{\alpha+\beta}} - 11 + 21\frac{1}{1+e^{\alpha+2\beta}}$$

and

$$\frac{\partial l(\alpha,\beta)}{\partial\beta} = -12\frac{1}{(1+e^{\alpha+\beta})^2} + 11\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})^2} - 42\frac{1}{(1+e^{\alpha+2\beta})^2}$$

$$= 11 - 23\frac{1}{1+e^{\alpha+\beta}} - 42\frac{1}{(1+e^{\alpha+2\beta})}$$

Substituting in the values $\alpha = -2.40315$ and $\beta = 2.78828$ gives $\frac{\partial l(\alpha,\beta)}{\partial\alpha} = -0.000025585$ and $\frac{\partial l(\alpha,\beta)}{\partial\beta} = -0.000019771$, so this is the maximum likelihood estimate.

(b) Use a normal approximation to calculate a 10% likelihood region for (α, β).

The second derivatives are:

$$\frac{\partial^2 l(\alpha,\beta)}{\partial\alpha^2} = -26\frac{e^{\alpha}}{(1+e^{\alpha})^2} - 23\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})^2} - 21\frac{e^{\alpha+2\beta}}{(1+e^{\alpha+2\beta})^2}$$

$$\frac{\partial^2 l(\alpha,\beta)}{\partial\alpha\partial\beta} = -23\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})^2} - 42\frac{e^{\alpha+2\beta}}{(1+e^{\alpha+2\beta})^2}$$

$$\frac{\partial^2 l(\alpha,\beta)}{\partial\beta^2} = -23\frac{e^{\alpha+\beta}}{(1+e^{\alpha+\beta})^2} - 84\frac{e^{\alpha+2\beta}}{(1+e^{\alpha+2\beta})^2}$$

Evaluating these at $\alpha = -2.40315$ and $\beta = 2.78828$ gives the observed information matrix

$$\begin{pmatrix} 8.3292241141 & 7.1616549515 \\ 7.1616549515 & 8.7813618861 \end{pmatrix}$$

The relative log-likelihood function is

$$r(\alpha,\beta) = -\frac{8.329}{2}(\alpha+2.40315)^2 - 7.162(\alpha+2.40315)(\beta-2.78828) - \frac{8.781}{2}(\beta-2.78828)^2$$

The 10% likelihood region is therefore

$$\frac{8.329}{2}(\alpha+2.40315)^2 + 7.162(\alpha+2.40315)(\beta-2.78828) + \frac{8.781}{2}(\beta-2.78828)^2 < \log(10)$$

3. *The remaining lifetime (in years) of a patient undergoing a certain kind of treatment is exponentially distributed with parameter λ. In a study which follows 10 patients for a period of 3 years, seven of the patients have lifetimes: 0.3, 0.8, 0.9, 1.4, 1.8, 2.5, and 2.9, while the remaining three patients survive to the end of the three-year period. What is the maximum likelihood estimate for λ?*

   The log-likelihood function is $7\log(\lambda) - (0.3+0.8+0.9+1.4+1.8+2.5+2.9+3\times 3)\lambda$, so the score function is $S(\lambda) = \frac{7}{\lambda} - 19.6$. Setting this to zero gives $\lambda = \frac{7}{19.6}$.

4. *We observe two samples from a Poisson distribution with parameter λ. If the true value of λ is 0.7, what is the probability that this value lies within a 10% likelihood interval?*

   Let the samples be X_1 and X_2. The likelihood is $e^{-2\lambda}\frac{\lambda^{X_1+X_2}}{X_1!X_2!}$, and the maximum likelihood estimate for λ is $\frac{X_1+X_2}{2}$. The relative likelihood of 0.7 is therefore $R(0.7) = \frac{e^{-1.4}0.7^{X_1+X_2}}{e^{-(X_1+X_2)}\left(\frac{X_1+X_2}{2}\right)^{X_1+X_2}}$. Evaluating this for different values of X_1+X_2 gives:

X_1+X_2	$R(0.7)$
0	0.246596964
4	0.202040219
5	0.062986908

   So 0.7 is in the 10% likelihood interval whenever $X_1+X_2 < 5$. X_1+X_2 has a Poisson distribution with mean 1.4, so the probability of this is $e^{-1.4}\left(1+1.4+\frac{1.4^2}{2}+\frac{1.4^3}{3!}+\frac{1.4^4}{4!}\right) = .9857467038$.

5. *The probability of a particular genetic condition is $p = \theta^2$. Let N be a sample from a binomial distribution with parameters 100 and θ. The MLE for p is $\left(\frac{N}{100}\right)^2$. What is the bias of this estimate?*

   Since $N \sim B(100,\theta)$, we have that $\mathbb{E}(N) = 100\theta$, and $\mathrm{Var}(N) = 100\theta(1-\theta)$, so $\mathbb{E}(N^2) = 10000\theta^2 + 100\theta(1-\theta)$, and $\mathbb{E}\left(\left(\frac{N}{100}\right)^2\right) = \theta^2 + \frac{\theta(1-\theta)}{100}$, so the bias of this estimator is $\frac{\theta(1-\theta)}{100}$.