BST 676 — Spring 2010 — Dr. Charnigo

Written Assignment 2 Solutions

1a. We have $m_1(\theta) = \theta/2$. Solving $m_1(\hat{\theta}) = n^{-1}\sum_{i=1}^n y_i$ yields the method of moments estimate $\hat{\theta}_{MM} = 2n^{-1}\sum_{i=1}^n y_i$, which is obviously different from the maximum likelihood estimate $\max\{y_1, y_2, \ldots, y_n\}$.

1b. The likelihood function is

$$L(\zeta; \mathbf{y}) = \prod_{i=1}^n (2\pi\zeta_2)^{-1/2} \exp[-(y_i - \zeta_1)^2/(2\zeta_2)] = (2\pi\zeta_2)^{-n/2} \exp[-\sum_{i=1}^n (y_i - \zeta_1)^2/(2\zeta_2)],$$

and so the log likelihood function is

$$l(\zeta; \mathbf{y}) = (-n/2)\log(2\pi) + (-n/2)\log(\zeta_2) - \sum_{i=1}^n (y_i - \zeta_1)^2/(2\zeta_2).$$

We have

$$\frac{\partial l(\zeta; \mathbf{y})}{\partial \zeta_1} = \sum_{i=1}^n (y_i - \zeta_1)/\zeta_2 \tag{1}$$

and

$$\frac{\partial l(\zeta; \mathbf{y})}{\partial \zeta_2} = (-n/2)/\zeta_2 + \sum_{i=1}^n (y_i - \zeta_1)^2/(2\zeta_2^2). \tag{2}$$

Setting (1) to zero yields $\zeta_1 = \bar{y}$, and then setting (2) to zero yields $\zeta_2 = n^{-1}\sum_{i=1}^n (y_i - \bar{y})^2$. Thus, the maximum likelihood estimate is $\hat{\theta}_{ML} = (\bar{y}, n^{-1}\sum_{i=1}^n (y_i - \bar{y})^2)^T$. This differs from the method of moments estimate, which had a divisor of $(n-1)$ in its second component.

2a. We have $m_1(\theta) = \int_\theta^\infty 3\theta^3/x^3\, dx = (3/2)\theta$. Solving $m_1(\hat{\theta}) = n^{-1}\sum_{i=1}^n x_i$ yields $\hat{\theta}_{MM} = (2/3)n^{-1}\sum_{i=1}^n x_i$.

2b. The likelihood function is

$$L(\zeta; \mathbf{x}) = \prod_{i=1}^n 1_{\{x_i \geq \zeta\}} 3\zeta^3/x_i^4 = 1_{\{\min\{x_1, x_2, \ldots, x_n\} \geq \zeta\}} \zeta^{3n} 3^n / \prod_{i=1}^n x_i^4.$$

Since ζ^{3n} is a strictly increasing function on $(0, \infty)$, the likelihood is maximized at the largest ζ for which the indicator does not vanish. Thus, $\hat{\theta}_{ML} = \min\{x_1, x_2, \ldots, x_n\}$.

3a. The likelihood function is

$$L(\zeta; \mathbf{x}) = \prod_{i=1}^n \zeta^{x_i}(1-\zeta)^{1-x_i} = \zeta^{\sum_{i=1}^n x_i}(1-\zeta)^{n - \sum_{i=1}^n x_i}. \tag{3}$$

If $\sum_{i=1}^n x_i = 0$, then (3) simplifies to $(1-\zeta)^n$, whose maximum on $[0,1]$ occurs at $\zeta = 0 = n^{-1}\sum_{i=1}^n x_i$.
If $\sum_{i=1}^n x_i = n$, then (3) simplifies to ζ^n, whose maximum on $[0,1]$ occurs at $\zeta = 1 = n^{-1}\sum_{i=1}^n x_i$.
Otherwise, (3) equals zero at $\zeta = 0$ and at $\zeta = 1$, so that the maximum occurs in the interior of $[0,1]$. The log likelihood, defined on the interior of $[0,1]$, is

$$l(\zeta; \mathbf{x}) = \sum_{i=1}^n x_i \log\zeta + (n - \sum_{i=1}^n x_i)\log(1-\zeta).$$

Setting

$$\frac{\partial l(\zeta;\mathbf{x})}{\partial \zeta} = \sum_{i=1}^n x_i/\zeta - (n - \sum_{i=1}^n x_i)/(1-\zeta)$$

to zero yields $\zeta = n^{-1}\sum_{i=1}^n x_i$.

Thus, in any event, the maximum likelihood estimate is $\hat{\theta}_{ML} = n^{-1}\sum_{i=1}^n x_i$.

3b. We have

$$p(\theta;\mathbf{x}) \propto p(\theta)L(\theta;\mathbf{x}) = \theta^{a+\sum_{i=1}^n x_i}(1-\theta)^{b+n-\sum_{i=1}^n x_i}.$$

Using calculations virtually identical to those in item 3a, we find that the posterior mode is $\hat{\theta}_{PMo} = (a+\sum_{i=1}^n x_i)/(a+b+n)$.

Recognizing that the posterior distribution is beta with parameters $a+\sum_{i=1}^n x_i+1$ and $b+n-\sum_{i=1}^n x_i+1$, we can quote the known result that the beta distribution with parameters α and β has mean $\alpha/(\alpha+\beta)$ to determine that the posterior mean is $\hat{\theta}_{PMe} = (a+\sum_{i=1}^n x_i+1)/(a+b+n+2)$.

Remark. If we do not remember the known result about the mean of the beta distribution with parameters α and β, we can derive it by the kernel method:

$$\begin{aligned}
& \int_0^1 \theta \, \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}\theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta \\
= \;& \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}\frac{\Gamma[\alpha+1]\Gamma[\beta]}{\Gamma[\alpha+\beta+1]}\int_0^1 \frac{\Gamma[\alpha+\beta+1]}{\Gamma[\alpha+1]\Gamma[\beta]}\theta^{\alpha+1-1}(1-\theta)^{\beta-1}\, d\theta \\
= \;& \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}\frac{\Gamma[\alpha+1]\Gamma[\beta]}{\Gamma[\alpha+\beta+1]} \\
= \;& \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]}\frac{\alpha\Gamma[\alpha]}{(\alpha+\beta)\Gamma[\alpha+\beta]} \\
= \;& \alpha/(\alpha+\beta).
\end{aligned}$$

4a. We can choose a and ν so that $\tau_1 Y_1 + \tau_2 Y_2$ and aY have the same mean and variance. Recalling that the mean of a chi-square random variable equals the degrees of freedom and that the variance equals twice the degrees of freedom, we obtain the equations

$$E[\tau_1 Y_1 + \tau_2 Y_2] = \tau_1 + \tau_2 = a\nu = E[aY]$$

and

$$Var[\tau_1 Y_1 + \tau_2 Y_2] = 2\tau_1^2 + 2\tau_2^2 = 2a^2\nu = Var[aY].$$

Dividing the second equation by twice the first yields $a = (\tau_1^2+\tau_2^2)/(\tau_1+\tau_2)$, and dividing twice the square of the first equation by the second yields $\nu = (\tau_1+\tau_2)^2/(\tau_1^2+\tau_2^2)$.

4b. The given values of τ_1 and τ_2 yield $a = 0.126$ and $\nu = 1.96$. Since $\chi^2_{1.96,0.95} = 5.914$, we have $a\chi^2_{\nu,0.95} = 0.745$. The following R code conducts the requested simulation study.

```
Y1<-rchisq(10000,1)
Y2<-rchisq(10000,1)
tau1<-1/(4*sqrt(pi))
tau2<-3/(16*sqrt(pi))
sum(tau1*Y1+tau2*Y2>(tau1^2+tau2^2)/(tau1+tau2)*qchisq(0.95,(tau1+tau2)^2/(tau1^2+tau2^2)))/10000
```

A different result will be obtained each time the code is run since new random chi-square variables will be generated. However, the result will always be fairly close to 0.05.

This suggests that, if we are comfortable treating nD_n as if its null-hypothesis distribution were that of $1/(4\sqrt{\pi})\ \chi^2_1 + 3/(16\sqrt{\pi})\ \chi^2_1$, then we may as well take the further step and treat nD_n as if its null-hypothesis distribution were that of $0.126\ \chi^2_{1.96}$.

What the simulation study does not clarify is what n must be for us to feel comfortable treating nD_n as if its null-hypothesis distribution were that of $1/(4\sqrt{\pi})\ \chi^2_1 + 3/(16\sqrt{\pi})\ \chi^2_1$. To get an idea about that, we would need to conduct another simulation study in which we generated 10000 realizations of nD_n and then examined a histogram of the 10000 realizations, for each of various n (small, medium, and large).