M347 TMA03 R2698663

Louis Botterill

04-March-2018

Question 1

1.a

$$f(x_i|\theta) = \theta^2 x_i e^{-\theta x_i} \text{ on } x_i > 0$$

if we take logs we can equate this to a standard natural exponential family NEF.

$$log(f(x_i|\theta)) = log(\theta^2) + log(x_i) + -\theta x_i$$

The standard form of the NEF is

$$log(f(x|\theta)) = xa(\theta) - b(\theta) + c(x)$$

So we have

$$\begin{aligned} b(\theta) &= -log(\theta^2) \\ c(x_i) &= log(x_i) \\ a(\theta) &= -\theta \end{aligned}$$

Now we can make use of the standard result for the NEF,

$$\begin{aligned} L(\theta) &\propto e^{n\bar{x}a(\theta) - nb(\theta)} \\ &\propto e^{-n\bar{x}\theta - n(-log(\theta^2))} \\ &\propto e^{-n\bar{x}\theta + nlog(\theta^2)} \\ &\propto e^{-n\bar{x}\theta} e^{nlog(\theta^2)} \\ &\propto e^{-n\bar{x}\theta} e^{log(\theta^{2n})} \end{aligned}$$

Therefore we obtain

$$L(\theta) \propto \theta^{2n} e^{-\theta n\bar{x}}$$

as required.

1.b

The posterior density is given by:

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)}$$

With a $gamma(a, b)$ prior we have

$$f(\theta) = gamma(a, b) \propto \theta^{a-1} e^{-b\theta}$$

ignoring constant terms (the normalisation constant $f(x)$), we then obtain

Posterior $\propto$ Likelyhood x Prior

$$f(\theta|x) \propto f(x|\theta)f(\theta)$$

$$f(x|\theta) = L(\theta)$$

$$\theta^{2n}e^{-\theta n\bar{x}} \times \theta^{a-1}e^{-b\theta}$$

Grouping terms

$$\theta^{2n}\theta^{a-1}e^{-\theta n\bar{x}} \times e^{-b\theta}$$

$$\theta^{2n+a-1}e^{-\theta(n\bar{x}+b)}$$

Therefore, the posterior can be written as

$$f(\theta|x) \propto \theta^{a+2n-1}e^{-(b+n\bar{x})\theta}$$

as required.

1.c

The posterior corresponds to the $Gamma(a + 2n, b + n\bar{x})$ distribution.

1.d

Given:

$E[\theta] = \frac{a}{b}$

MLE $\hat{\theta}$ of θ under $gamma(2, \theta) = \hat{\theta} = \frac{2}{\bar{x}}$

Posterior mean is:

$$E[\theta|x] = \frac{a + 2n}{b + n\bar{x}}$$

Expanding the posterior mean we have

$$E[\theta|x] = \frac{a}{b + n\bar{x}} + \frac{2n}{b + n\bar{x}}$$

To write this in the form

$$E[\theta|x] = t\hat{\theta} + (1 - t)E[\theta] \text{ over } 0 < t < 1$$

Start with

$$E[\theta|x] = t\frac{2}{\bar{x}} + (1 - t)\frac{a}{b}$$

We can equate the terms, the 2nd term can be written

$$t\frac{2}{\bar{x}} = \frac{n\bar{x}}{b + n\bar{x}} \times \frac{2}{\bar{x}} = \frac{2n}{b + n\bar{x}}$$

So t is therefore

$$t = \frac{n\bar{x}}{b + n\bar{x}}$$

Similarly then for (1 - t) we have

$$1 - t = \frac{b + n\bar{x} - n\bar{x}}{b + n\bar{x}} = \frac{b}{b + n\bar{x}}$$

and indeed this equates to the first term

$$(1 - t)\frac{a}{b} = \frac{b}{b + n\bar{x}}\frac{a}{b} = \frac{a}{b + n\bar{x}}$$

So we've shown we can write the posterior in the form

$$E[\theta|x] = t\frac{2}{\bar{x}} + (1 - t)\frac{a}{b}$$

Therefore $E(\theta|x)$ can be written as

$$E[\theta|x] = t\hat{\theta} + (1 - t)E[\theta] \text{ over } 0 < t < 1$$

as required.

Question 2

2.a

$$E[L(d,\theta)|x] = 1 + F(d-\alpha|x) - F(d+\alpha|x)$$

$$F(\theta|x) = 1 - F(2\theta_0 - \theta|x)$$

$$E[L(d,\theta)|x] = 1 + F(d-\alpha|x) - 1 - F(2\theta_0 - (d+\alpha)|x)$$

$$E[L(d,\theta)|x] = F(d-\alpha|x) - F(2\theta_0 - (d+\alpha)|x)$$

2.b

$$E[L(d,\theta)|x] = F(d-\alpha|x) - F(2\theta_0 - (d+\alpha)|x)$$

recall that

$$F(d|x) = \int_{\infty}^{d} f(\theta|x)d\theta$$

so that

$$\frac{d}{dd}F(d|x) = \frac{d}{dd}\int_{\infty}^{d} f(\theta|x)d\theta = f(\theta|x)$$

$$\frac{d}{dd}E[L(d,\theta)|x] = f(d-\alpha|x) - f(2\theta_0 - d - \alpha|x)$$

when $d = \theta_0$

$$\frac{d}{dd}E[L(d,\theta)|x] = f(d-\alpha|x) - f(2\theta_0 - \theta_0 - \alpha|x)$$

$$\frac{d}{dd}E[L(d,\theta)|x] = f(d-\alpha|x) - f(\theta_0 - \alpha|x)$$

$$\frac{d}{dd}E[L(d,\theta)|x] = f(d-\alpha|x) - f(d-\alpha|x)$$

$$\frac{d}{dd}E[L(d,\theta)|x] = 0$$

2.c

$$\frac{d^2}{dd}E[L(d,\theta)|x] = \frac{d}{dd}f(d-\alpha|x) - f(2\theta_0 - d - \alpha|x)$$

Using the chain rule for differentiation of nested functions, which is

$$k(x) = g(f(x)) \text{ then } k'(x) = g'(f(x)) \times f'(x)$$

$$\frac{d}{dd}f(d-\alpha|x) = 1 \times f'(d-\alpha|x)$$

$$\frac{d}{dd} - f(2\theta_0 - d - \alpha|x) = -1 \times -f'(2\theta_0 - d - \alpha|x) = f'(2\theta_0 - d - \alpha|x)$$

So we have

$$\frac{d^2}{dd}E[L(d,\theta)|x] = f'(d-\alpha|x) + f'(2\theta_0 - d - \alpha|x)$$

at $d = \theta_0$ when then get

$$\frac{d^2}{dd}E[L(d,\theta)|x] = f'(\theta_0-\alpha|x) + f'(2\theta_0 - \theta_0 - \alpha|x)$$

$$\frac{d^2}{dd}E[L(d,\theta)|x] = 2f'(\theta_0-\alpha|x)$$

At $\theta = \theta_0 - \alpha$ we know $\theta < \theta_0$ always (since $\alpha > 0$) so, using the differentiation properties of the posterior density $f(\theta|x)$ as given,

$$f'(\theta|x) > 0 \text{ for } \theta < \theta_0$$

Therefore

$$2f'(\theta_0 - \alpha|x) > 0$$

and so

$$\frac{d^2}{dd}E[L(d,\theta)|x] > 0$$

So we have shown that $\frac{d^2}{dd}E[L(d,\theta)|x]$ is positive at $d = \theta_0$, as required (for any $\alpha > 0$).

2.d

In this question we have shown that the 1^{st} derivative of the expected loss function is 0 (therefore a stationary point at $d = \theta_0$) and that the 2^{nd} derivative of the expected loss about the mode of the posterior density $d = \theta_0$ is positive, showing that $d = \theta_0$ is a minimiser of the expected loss.
Furthermore we are told that θ_0 is the mode of the posterior. So the minimiser of the expected loss $E[L(d,\theta)|x]$ is at the mode of the posterior at θ_0.
Also since the posterior is unimodal, this is the only minimiser, it is therefore the global minimiser of the expected loss function. We have made this proof under the basis $\alpha > 0$ and so α is also non-zero. We don't know the behaviour of derivative of the posterior density at θ_0 so we can not extend the proof to $alpha = 0$.

So the optimal estimator of θ for the symmetric posterior under 0-1 loss function is the posterior mode, for positive α.

Question 3

3.a

We have

$$\mu_2^* \sim N(\mu_1', 2)$$
$$\mu_1' = 0$$

Therefore the Metropolis-Hastings algorithm will use the following distribution to generate candidate value μ^* for μ_2

$$\mu_2^* \sim N(0, 2)$$

3.b

The general acceptance probability for Metropolis-Hastings sampling is

$$\alpha(\mu^*|\mu_t) = min\Big(\frac{q(\mu_t|\mu^*)f(\mu^*)}{q(\mu^*|\mu_t)f(\mu_t)}, 1\Big)$$

This can be simplified under a symmetric proposal density where

$$q(\mu_{t+1}|\mu_t) = q(\mu_t|\mu_{t+1})$$

The proposal density is $N \sim (\mu = \mu_t, \sigma^2 = 2)$

Where the Normal distribution is defined as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \text{ on } \mathbb{R}$$

So ignoring linear constants of proportionality we have

$$q(\mu_{t+1}|\mu_t) \propto e^{-\frac{1}{2x2}(x_{t+1}-x_t)^2}$$

$$q(\mu_t|\mu_{t+1}) \propto e^{-\frac{1}{2x2}(x_t-x_{t+1})^2}$$

since $(x_{t+1} - x_t)^2 = (x_t - x_{t+1})^2$

these two expressions are the same and therefore cancel out, leaving

$$\alpha(\mu^*|\mu_t) = min\Big(\frac{f(\mu^*)}{f(\mu_t)}, 1\Big)$$

So when X was observed to be 1 we have for the simulated value of $\mu_{t=1}'$ the candidate acceptable probability of

$$\alpha(\mu^*|\mu_1') = min\Big(\frac{f(\mu^*|x=1)}{f(\mu_1'|x=1)}, 1\Big)$$

as required.

3.c

Given

$$\mu_1' = 0$$

$$\mu^* = 1.2$$

$$f(\mu|x=1) \propto (1+e^{-(\mu+2)})^{-1}e^{-\mu^2/2}$$

So we have

$$f(\mu^*|x=1) \propto (1+e^{-(1.2+2)})^{-1}e^{-1.2^2/2} = (1+e^{-3.2})^{-1}e^{-0.72}$$

$$f(\mu^*|x=1) \propto 0.9608 \times 0.4868 = 0.4677$$

$$f(\mu_1'|x=1) \propto (1+e^{-(0+2)})^{-1}e^{0^2/2} = (1+e^{-2})^{-1} = 0.8808$$

Using

$$\alpha(\mu^*|\mu_1') = min\Big(\frac{f(\mu^*|x=1)}{f(\mu_1'|x=1)}, 1\Big)$$

We get

$$\alpha(\mu^*|\mu_1') = min\Big(\frac{0.4677}{0.8808}, 1\Big)$$

$$\alpha(\mu^*|\mu_1') = 0.531 \text{ (to 3 d.p.)}$$

3.d

Since $\alpha(\mu^*|\mu_t') = 0.531 < u \sim U(0,1) = 0.81$ this candidate value is accepted.

Therefore the value of μ_2' is updated to 1.2

Therefore the candidate μ^{**} for μ_3 will be drawn from $\sim N(1.2, 2)$ which is the same normal distribution with mean updated to be centred around the latest accepted value of μ.