MA40189 - Solution Sheet Six

Simon Shaw, s.shaw@bath.ac.uk
https://people.bath.ac.uk/masss/ma40189.html

2021/22 Semester II

1. **Let $X_1, \ldots, X_n$ be exchangeable so that the X_i are conditionally independent given a parameter $\theta = (\mu, \sigma^2)$. Suppose that $X_i \,|\, \theta \sim N(\mu, \sigma^2)$. It is judged that the improper joint prior distribution $f(\mu, \sigma^2) \propto 1/\sigma^2$ is appropriate.**

   (a) **Show that the likelihood $f(x \,|\, \mu, \sigma^2)$, where $x = (x_1, \ldots, x_n)$, can be expressed as**

$$f(x \,|\, \mu, \sigma^2) \;=\; \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\overline{x} - \mu)^2\right]\right\},$$

   where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ are respectively the sample mean and variance. Hence, explain why $\overline{X}$ and S^2 are sufficient for $X = (X_1, \ldots, X_n)$ for learning about θ.

$$\begin{aligned} f(x \,|\, \mu, \sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(x_i - \mu)^2\right\} \\ &= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\} \\ &= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left((x_i - \overline{x}) + (\overline{x} - \mu)\right)^2\right\} \\ &= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\overline{x} - \mu)^2\right]\right\}. \end{aligned}$$

   Recall that a statistic $t(X)$ is said to be sufficient for X for learning about θ if we can write

$$f(x \,|\, \theta) \;=\; g(t, \theta)h(x)$$

   where $g(t, \theta)$ depends upon $t(x)$ and θ and $h(x)$ does not depend upon θ but may depend upon x. In this case we have that

$$f(x \,|\, \mu, \sigma^2) \;=\; g(\overline{x}, s^2, \mu, \sigma^2)\,(2\pi)^{-\frac{n}{2}}$$

   so that $\overline{X}$ and S^2 are sufficient.

   (b) **Find, up to a constant of integration, the posterior distribution of θ given x.**

$$f(\mu, \sigma^2 \,|\, x) \;\propto\; f(x \,|\, \mu, \sigma^2)f(\mu, \sigma^2)$$

$$
\begin{aligned}
&\propto \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right\}\frac{1}{\sigma^2} \\
&\propto \left(\sigma^2\right)^{-\left(\frac{n}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right\}.
\end{aligned}
$$

(c) **Show that $\mu \,|\, \sigma^2, x \sim N(\overline{x}, \sigma^2/n)$. Hence, explain why, in this case, the chosen prior distribution for θ is noninformative.**

$$
\begin{aligned}
f(\mu \,|\, \sigma^2, x) &= \frac{f(\mu, \sigma^2 \,|\, x)}{f(\sigma^2 \,|\, x)} \\
&\propto f(\mu, \sigma^2 \,|\, x) \quad \text{(with respect to } \mu) \\
&\propto \exp\left\{-\frac{n}{2\sigma^2}(\overline{x}-\mu)^2\right\}.
\end{aligned}
$$

We recognise this as a kernel of $N(\overline{x}, \sigma^2/n)$ so $\mu \,|\, \sigma^2, x \sim N(\overline{x}, \sigma^2/n)$. In the classical model for μ when σ^2 is known, the mle is $\overline{x}$ and the standard error is σ^2/n. The distribution is coming only from the data (given σ^2) showing the noninformative prior in this case. Notice that a symmetric $100(1-\alpha)\%$ credible interval for $\mu \,|\, \sigma^2, x$ is

$$
\left(\overline{x} - z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{(1-\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}}\right)
$$

which agrees with the $100(1-\alpha)\%$ confidence interval for μ when σ^2 is known and $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$.

(d) **By integrating $f(\mu, \sigma^2 \,|\, x)$ over σ^2, show that**

$$
f(\mu \,|\, x) \propto \left[1 + \frac{1}{n-1}\left(\frac{\overline{x}-\mu}{s/\sqrt{n}}\right)^2\right]^{-\frac{n}{2}}.
$$

Thus, explain why $\mu \,|\, x \sim t_{n-1}(\overline{x}, s^2/n)$, the non-central t-distribution with $n-1$ degrees of freedom, location parameter $\overline{x}$ and squared scale parameter s^2/n. How does this result relate to the classical problem of making inferences about μ when σ^2 is also unknown?

$$
\begin{aligned}
f(\mu \,|\, x) &= \int_{-\infty}^{\infty} f(\mu, \sigma^2 \,|\, x)\, d\sigma^2 \\
&\propto \int_{0}^{\infty} \left(\sigma^2\right)^{-\left(\frac{n}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right\} d\sigma^2 \qquad (1)
\end{aligned}
$$

as $f(\mu, \sigma^2 \,|\, x) = 0$ for $\sigma^2 < 0$. We recognise the integrand in equation (1) as a kernel of Inv-Gamma $\left(\frac{n}{2}, \frac{1}{2}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right)$ so that

$$
\begin{aligned}
f(\mu \,|\, x) &\propto \Gamma(n/2)\left(\frac{1}{2}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right)^{-\frac{n}{2}} \\
&\propto \left(\frac{1}{2}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right)^{-\frac{n}{2}} \\
&\propto \left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]^{-\frac{n}{2}} \\
&\propto \left[1 + \frac{1}{n-1}\left(\frac{\overline{x}-\mu}{s/\sqrt{n}}\right)^2\right]^{-\frac{n}{2}}.
\end{aligned}
$$

We recognise this as a kernel of $t_{n-1}(\overline{x}, s^2/n)$ so that $\mu \,|\, x \sim t_{n-1}(\overline{x}, s^2/n)$. This gives a further insight into how the prior distribution is noninformative. Inference about μ will mirror the classical approach when μ and σ^2 are unknown and $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ where $t_{n-1} = t_{n-1}(0,1)$. For example, a symmetric $100(1-\alpha)\%$ credible interval for $\mu \,|\, x$ is

$$\left(\overline{x} - t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}, \overline{x} + t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right)$$

which agrees with the $100(1-\alpha)\%$ confidence interval for μ when σ^2 is unknown and $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$.

2. **Let $X_1, \ldots, X_n$ be exchangeable with $X_i \,|\, \theta \sim Bern(\theta)$.**

   (a) **Using the improper prior distribution $f(\theta) \propto \theta^{-1}(1-\theta)^{-1}$ find the posterior distribution of $\theta \,|\, x$ where $x = (x_1, \ldots, x_n)$. Find a normal approximation about the mode to this distribution.**

The likelihood is

$$f(x \,|\, \theta) \;=\; \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} \;=\; \theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}},$$

where $x = (x_1, \ldots, x_n)$. With the given prior the posterior is

$$f(\theta \,|\, x) \;\propto\; f(x \,|\, \theta) f(\theta) \;\propto\; \theta^{n\bar{x}-1}(1-\theta)^{n-n\bar{x}-1}$$

which we recognise as a kernel of a $Beta(n\bar{x}, n-n\bar{x})$ density. Thus, $\theta \,|\, x \sim Beta(n\bar{x}, n-n\bar{x})$. The mode of a $Beta(\alpha, \beta)$ distribution is $\frac{\alpha-1}{\alpha+\beta-2}$ so for $\theta \,|\, x$ the mode is

$$\tilde{\theta} \;=\; \frac{n\bar{x}-1}{n-2}.$$

The observed information is

$$\begin{aligned} I(\theta \,|\, x) &= -\frac{\partial^2}{\partial\theta^2} \log f(\theta \,|\, x) \\ &= -\frac{\partial^2}{\partial\theta^2}\left\{\log \frac{\Gamma(n)}{\Gamma(n\bar{x})\Gamma(n-n\bar{x})} + (n\bar{x}-1)\log\theta + (n-n\bar{x}-1)\log(1-\theta)\right\} \\ &= \frac{n\bar{x}-1}{\theta^2} + \frac{n-n\bar{x}-1}{(1-\theta)^2}. \end{aligned}$$

So, evaluating the observed information at the mode,

$$I(\tilde{\theta} \,|\, x) \;=\; \frac{n\bar{x}-1}{\tilde{\theta}^2} + \frac{n-n\bar{x}-1}{(1-\tilde{\theta})^2}.$$

Noting that $1 - \tilde{\theta} = \frac{n-n\bar{x}-1}{n-2}$ we have that

$$\begin{aligned} I(\tilde{\theta} \,|\, x) &= \frac{(n-2)^2}{n\bar{x}-1} + \frac{(n-2)^2}{n-n\bar{x}-1} \\ &= \frac{(n-2)^3}{(n\bar{x}-1)(n-n\bar{x}-1)}. \end{aligned}$$

So, approximately, $\theta \,|\, x \sim N(\tilde{\theta}, I^{-1}(\tilde{\theta} \,|\, x))$, that is, approximately,

$$\theta \,|\, x \quad \sim \quad N\left(\frac{n\bar{x}-1}{n-2}, \frac{(n\bar{x}-1)(n-n\bar{x}-1)}{(n-2)^3}\right).$$

(b) **Show that the prior distribution $f(\theta)$ is equivalent to a uniform prior on**

$$\beta \quad = \quad \log\left(\frac{\theta}{1-\theta}\right)$$

and find the posterior distribution of $\beta \,|\, x$. Find a normal approximation about the mode to this distribution.

We have $\beta = g(\theta)$. We invert to find $\theta = g^{-1}(\beta)$. We find

$$\theta \quad = \quad \frac{e^{\beta}}{1+e^{\beta}}.$$

The prior $f_{\beta}(\beta)$ for β is given by

$$\begin{aligned} f_{\beta}(\beta) &= \left|\frac{\partial\theta}{\partial\beta}\right| f_{\theta}(\theta) \\ &= \left|\frac{e^{\beta}}{(1+e^{\beta})^2}\right| \left(\frac{e^{\beta}}{1+e^{\beta}}\right)^{-1} \left(1-\frac{e^{\beta}}{1+e^{\beta}}\right)^{-1} \\ &= \frac{e^{\beta}}{(1+e^{\beta})^2} \times \frac{1+e^{\beta}}{e^{\beta}} \times (1+e^{\beta}) \;=\; 1, \end{aligned}$$

which is equivalent to the (improper) uniform on β. The posterior is

$$\begin{aligned} f(\beta \,|\, x) &\propto f(x \,|\, \beta) f(\beta) \\ &\propto \theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}} \\ &= \left(\frac{e^{\beta}}{1+e^{\beta}}\right)^{n\bar{x}} \left(1-\frac{e^{\beta}}{1+e^{\beta}}\right)^{n-n\bar{x}} \;=\; \frac{e^{\beta n\bar{x}}}{(1+e^{\beta})^n}. \end{aligned}$$

Hence, $f(\beta \,|\, x) = \frac{ce^{\beta n\bar{x}}}{(1+e^{\beta})^n}$ where c is the constant of integration. For the normal approximation about the mode, we first need to find the mode of $\beta \,|\, x$. The mode is the maximum of $f(\beta \,|\, x)$ which is, equivalently, the maximum of $\log f(\beta \,|\, x)$. We have

$$\begin{aligned} \log f(\beta \,|\, x) &= \log c + \beta n\bar{x} - n\log(1+e^{\beta}) \\ \Rightarrow \frac{\partial}{\partial\beta}\log f(\beta \,|\, x) &= n\bar{x} - \frac{ne^{\beta}}{1+e^{\beta}}. \end{aligned}$$

The mode $\tilde{\beta}$ satisfies

$$\begin{aligned} n\bar{x} - \frac{ne^{\tilde{\beta}}}{1+e^{\tilde{\beta}}} &= 0 \\ \Rightarrow e^{\tilde{\beta}} &= \frac{n\bar{x}}{n-n\bar{x}} \\ \Rightarrow \tilde{\beta} &= \log\left(\frac{\bar{x}}{1-\bar{x}}\right). \end{aligned}$$

The observed information is

$$\begin{aligned} I(\beta \,|\, x) &= -\frac{\partial^2}{\partial \beta^2} \log f(\beta \,|\, x) \\ &= \frac{n e^{\beta}}{(1+e^{\beta})^2}. \end{aligned}$$

Noting that $1 + e^{\tilde{\beta}} = \frac{1}{1-\bar{x}}$ we have

$$I(\tilde{\beta} \,|\, x) = n \times \frac{\bar{x}}{1-\bar{x}} \times (1-\bar{x})^2 = n\bar{x}(1-\bar{x}).$$

Hence, approximately,

$$\beta \,|\, x \sim N\left(\log \frac{\bar{x}}{1-\bar{x}}, \frac{1}{n\bar{x}(1-\bar{x})}\right).$$

(c) **For which parameterisation does it make more sense to use a normal approximation?**

Whilst we can find a normal approximation about the mode either on the scale of θ or of β, it makes more sense for β. We have $0 < \theta < 1$ and $-\infty < \beta < \infty$ so only β has a sample space which agrees with the normal distribution.

3. **In viewing a section through the pancreas, doctors see what are called "islands". Suppose that X_i denotes the number of islands observed in the ith patient, $i = 1, \ldots, n$, and we judge that $X_1, \ldots, X_n$ are exchangeable with $X_i \,|\, \theta \sim Po(\theta)$. A doctor believes that for healthy patients θ will be on average around 2; he thinks it is unlikely that θ is greater than 3. The number of islands seen in 100 patients are summarised in the following table.**

Number of islands	0	1	2	3	4	5	≥ 6
Frequency	20	30	28	14	7	1	0

(a) **Express the doctor's prior beliefs as a normal distribution for θ. You may interpret the term "unlikely" as meaning "with probability 0.01".**

The doctor thus asserts $\theta \sim N(\mu, \sigma^2)$ with $E(\theta) = 2$ and $P(\theta > 3) = 0.01$. Note that, as θ is continuous, this is equivalent to $P(\theta \geq 3) = 0.01$. We use these two pieces of information to obtain μ and σ^2. Firstly, $E(\theta) = 2 \Rightarrow \mu = 2$. Secondly,

$$P(\theta > 3) = 0.01 \Rightarrow P\left(\frac{\theta - 2}{\sigma} > \frac{1}{\sigma}\right) = 0.01.$$

As $\frac{\theta-2}{\sigma} \sim N(0,1)$ we have that

$$\frac{1}{\sigma} = 2.33 \Rightarrow \sigma^2 = 5.4289^{-1} = 0.1842.$$

Hence, $\theta \sim N(2, 5.4289^{-1})$.

(b) **Find, up to a constant of proportionality, the posterior distribution $\theta \,|\, x$ where $x = (x_1, \ldots, x_{100})$.**

As $X_i \,|\, \theta \sim Po(\theta)$, the likelihood is

$$f(x \,|\, \theta) \;=\; \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} \;=\; \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^{n} x_i!}.$$

As $\theta \sim N(2, 5.4289^{-1})$, the prior is

$$f(\theta) \;=\; \frac{2.33}{\sqrt{2\pi}} \exp\left\{-\frac{5.4289}{2}(\theta - 2)^2\right\}.$$

The posterior is thus

$$\begin{aligned} f(\theta \,|\, x) &\propto f(x \,|\, \theta) f(\theta) \\ &\propto \theta^{n\bar{x}} e^{-n\theta} \times \exp\left\{-\frac{5.4289}{2}(\theta^2 - 4\theta)\right\} \\ &= \theta^{n\bar{x}} \exp\left\{-\frac{5.4289}{2}(\theta^2 - 4\theta) - n\theta\right\} \end{aligned}$$

For the explicit data we have $n = 100$ and

$$\sum_{i=1}^{100} x_i = (0 \times 20) + (1 \times 30) + (2 \times 28) + (3 \times 14) + (4 \times 7) + (5 \times 1) = 161.$$

The posterior is thus

$$f(\theta \,|\, x) \;=\; c\theta^{161} \exp\left\{-\frac{5.4289}{2}(\theta^2 - 4\theta) - 100\theta\right\}$$

where c is the constant of proportionality.

(c) **Find a normal approximation to the posterior about the mode. Thus, estimate the posterior probability that the average number of islands is greater than 2.**

We find the mode of $\theta \,|\, x$ by maximising $f(\theta \,|\, x)$ or, equivalently, $\log f(\theta \,|\, x)$. We have

$$\begin{aligned} \frac{\partial}{\partial\theta} \log f(\theta \,|\, x) &= \frac{\partial}{\partial\theta}\left\{\log c + 161 \log\theta - \frac{5.4289}{2}(\theta^2 - 4\theta) - 100\theta\right\} \\ &= \frac{161}{\theta} - 5.4289\theta + 2(5.4289) - 100. \end{aligned}$$

So, the mode $\tilde{\theta}$ satisfies

$$\begin{aligned} 5.4289\tilde{\theta}^2 + \{100 - 2(5.4289)\}\tilde{\theta} - 161 &= 0 \;\Rightarrow \\ 5.4289\tilde{\theta}^2 + 89.1422\tilde{\theta} - 161 &= 0. \end{aligned}$$

Hence, as $\tilde{\theta} > 0$,

$$\tilde{\theta} \;=\; \frac{-89.1422 + \sqrt{89.1422^2 + 4(5.4289)(161)}}{2(5.4289)} \;=\; 1.6419.$$

The observed information is

$$I(\theta \,|\, x) \;=\; -\frac{\partial^2}{\partial\theta^2}\log f(\theta \,|\, x) \;=\; \frac{161}{\theta^2} + 5.4289$$

so that

$$I(\tilde{\theta} \,|\, x) \;=\; \frac{161}{1.6419^2} + 5.4289 \;=\; 65.1506.$$

So, approximately, $\theta \,|\, x \sim N(1.6419, 65.1506^{-1})$. Thus,

$$P(\theta > 2 \,|\, x) \;=\; P\{Z > \sqrt{65.1506}(2 - 1.6419)\} \;=\; 0.0019.$$

(d) **Why might you prefer to express the doctor's prior beliefs as a normal distribution on some other parameterisation $\phi = g(\theta)$? Suggest an appropriate choice of $g(\cdot)$ in this case. Now express the doctor's beliefs using a normal prior for ϕ; that for healthy patients ϕ will be on average around $g(2)$ and it is "unlikely" that ϕ is greater than $g(3)$. Give an expression for the density of $\phi \,|\, x$ up to a constant of proportionality.**

By definition $\theta > 0$ but both the specified normal prior distribution and the normal approximation for the posterior $\theta \,|\, x$ have a sample space of $(-\infty, \infty)$ so we might want to use some other parametrisation which has the same sample space as the normal distribution. An obvious choice is to use $\phi = g(\theta) = \log\theta$ as if $\theta > 0$ then $-\infty < \log\theta < \infty$. We assert $\phi \sim N(\mu_0, \sigma_0^2)$ with $E(\phi) = \log 2$ and $P(\phi > \log 3) = 0.01$. So, $E(\phi) = \log 2 \Rightarrow \mu_0 = \log 2$ and

$$P(\phi > \log 3) \;=\; 0.01 \quad \Rightarrow P\left(\tfrac{\phi - \log 2}{\sigma_0} > \tfrac{\log 3 - \log 2}{\sigma_0}\right) \;=\; 0.01.$$

As $\frac{\phi - \log 2}{\sigma_0} \sim N(0, 1)$ we have that

$$\frac{1}{\sigma_0} \;=\; \frac{2.33}{\log\frac{3}{2}} \quad\Rightarrow\quad \sigma_0^2 \;=\; 0.0303.$$

As $\phi = \log\theta$ then $\theta = e^\phi$. The likelihood is thus, using (b),

$$f(x \,|\, \phi) \;\propto\; (e^\phi)^{n\bar{x}} e^{-ne^\phi} \;=\; e^{161\phi} e^{-100e^\phi}$$

for the given data. The posterior is thus

$$\begin{aligned} f(\phi \,|\, x) \;&\propto\; e^{161\phi} e^{-100e^\phi} \times \exp\left\{-\frac{1}{0.0606}(\phi^2 - \log 4\phi)\right\} \\ &=\; \exp\left\{-100e^\phi - 16.5017\phi^2 + 183.8761\phi\right\}. \end{aligned}$$

4. **Let $X_1, \ldots, X_{10}$ be the length of time between arrivals at an ATM machine, and assume that the X_is may be viewed as exchangeable with $X_i \,|\, \lambda \sim Exp(\lambda)$ where λ is the rate at which people arrive at the machine in one-minute intervals. Suppose we observe $\sum_{i=1}^{10} x_i = 4$. Suppose that the prior distribution for λ is given by**

$$f(\lambda) \;=\; \begin{cases} c\exp\{-20(\lambda - 0.25)^2\} & \lambda \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

where c is a known constant.

(a) **Find, up to a constant k of proportionality, the posterior distribution $\lambda \,|\, x$ where $x = (x_1, \ldots, x_{10})$. Find also an expression for k which you need not evaluate.**

The likelihood is

$$f(x \,|\, \lambda) \;=\; \prod_{i=1}^{10} \lambda e^{-\lambda x_i} \;=\; \lambda^{10} e^{-4\lambda}$$

with the given data. The posterior is thus

$$f(\lambda \,|\, x) \;\propto\; \lambda^{10} e^{-4\lambda} \times \exp\{-20(\lambda - 0.25)^2\} \;=\; \lambda^{10} \exp\{-20(\lambda^2 - 0.5\lambda) - 4\lambda\}$$

so that, making the fact that $\lambda > 0$ explicit,

$$f(\lambda \,|\, x) \;=\; \begin{cases} k\lambda^{10} \exp\{-20\lambda^2 + 6\lambda\} & \lambda > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where

$$k^{-1} \;=\; \int_0^\infty \lambda^{10} \exp\{-20\lambda^2 + 6\lambda\} \, d\lambda.$$

(b) **Find a normal approximation to this posterior distribution about the mode.**

We find the mode of $\lambda \,|\, x$ by maximising $f(\lambda \,|\, x)$ or, equivalently, $\log f(\lambda \,|\, x)$. We have

$$\begin{aligned} \frac{\partial}{\partial \lambda} \log f(\lambda \,|\, x) &= \frac{\partial}{\partial \lambda} \left\{10 \log \lambda - 20\lambda^2 + 6\lambda\right\} \\ &= \frac{10}{\lambda} - 40\lambda + 6. \end{aligned}$$

So, the mode $\tilde{\lambda}$ satisfies

$$40\tilde{\lambda}^2 - 6\tilde{\lambda} - 10 \;=\; 0 \;\;\Rightarrow\;\; 20\tilde{\lambda}^2 - 3\tilde{\lambda} - 5 \;=\; 0.$$

Hence, as $\tilde{\lambda} > 0$,

$$\tilde{\lambda} \;=\; \frac{3 + \sqrt{9 + 4(20)(5)}}{2(20)} \;=\; 0.5806.$$

The observed information is

$$I(\lambda \,|\, x) \;=\; -\frac{\partial^2}{\partial \lambda^2} \log f(\lambda \,|\, x) \;=\; \frac{10}{\lambda^2} + 40$$

so that

$$I(\tilde{\lambda} \,|\, x) \;=\; \frac{10}{0.5806^2} + 40 \;=\; 69.6651.$$

So, approximately, $\lambda \,|\, x \sim N(0.5806, 69.6651^{-1})$.

(c) **Let Z_i, $i = 1, \ldots, N$ be a sequence of independent and identically distributed standard Normal random quantities. Assuming the normalising constant k is known, explain carefully how the Z_i may be used to obtain estimates of the mean of $\lambda \,|\, x$.**

We shall use importance sampling. If we wish to estimate some $E\{g(\lambda) \,|\, X\}$ with posterior density $f(\lambda \,|\, x)$ and can generate independent samples $\lambda_1, \ldots, \lambda_N$ from some $q(\lambda)$, an approximation of $f(\lambda \,|\, x)$, then

$$\hat{I} = \frac{1}{N} \sum_{i=1}^{N} \frac{g(\lambda_i) f(\lambda_i \,|\, x)}{q(\lambda_i)}$$

is an unbiased estimate of $E\{g(\lambda) \,|\, X\}$.

As $Z_i \sim N(0,1)$ then $\lambda_i = 69.6651^{-\frac{1}{2}} Z_i + 0.5806 \sim N(0.5806, 69.6651^{-1})$ so that we can generate an independent and identically distributed sample from the $N(0.5806, 69.6651^{-1})$ which is an approximation to the posterior of λ. Letting $g(\lambda) = \lambda$,

$$\begin{aligned} f(\lambda \,|\, x) &= \begin{cases} k\lambda^{10} \exp\{-20\lambda^2 + 6\lambda\} & \lambda > 0 \\ 0 & \text{otherwise,} \end{cases} \\ q(\lambda) &= \frac{\sqrt{69.6651}}{\sqrt{2\pi}} \exp\left\{-\frac{69.6651}{2}(\lambda - 0.5806)^2\right\} \end{aligned}$$

then

$$\hat{I} = \frac{1}{N} \sum_{i=1}^{N} \frac{k\lambda_i^{11} \exp\{-20\lambda_i^2 + 6\lambda_i\} \mathbb{I}_{\{\lambda_i > 0\}}}{\frac{\sqrt{69.6651}}{\sqrt{2\pi}} \exp\left\{-\frac{69.6651}{2}(\lambda_i - 0.5806)^2\right\}}$$

is an unbiased estimate of the posterior mean of λ with $\mathbb{I}_{\{\lambda_i > 0\}}$ denoting the indicator function for the event $\{\lambda_i > 0\}$.