XXXX 2017
xx.xx – x.xx

EXAMINATION FOR THE DEGREES OF M.Sci., M.Sc. and M.Res.

Bayesian Statistics (Level M) Solutions

“Hand calculators with simple basic functions (log, exp, square root, etc.) may be used in examinations. No calculator which can store or display text or graphics may be used, and any student found using such will be reported to the Clerk of Senate”.

NOTE: Candidates should attempt ALL questions.

CONTINUED OVERLEAF/

1. Consider the Poisson model with parameter θ:

$$p(y|\theta) = \frac{\theta^y}{y!}\exp(-\theta), \qquad \theta > 0, y = 0, 1, \ldots.$$

(a) Write down the likelihood function appropriate to n i.i.d. observations $y_1, \ldots, y_n$. **[2 MARKS]**

Solution

$$\begin{aligned} L(\theta) &= p(y_1, \ldots, y_n|\theta) = \prod_{i=1}^{n} p(y_i|\theta) = \prod_{i=1}^{n} \frac{\theta^{y_i}e^{-\theta}}{y_i!} \\ &= \frac{\theta^{\sum y_i}}{\prod y_i!}e^{-n\theta} \end{aligned}$$

(b) Explain the Bayesian concept of a conjugate prior. Show that the Ga(α, β) distribution is conjugate for i.i.d. Poisson data $y_1, \ldots, y_n$ using the likelihood derived in the previous question. **[2,3 MARKS]**

Solution

Suppose the prior distribution comes from a particular family of distributions, e.g. the gamma distributions. If the posterior distribution (obtained via Bayes' theorem) is from the same family of distributions (with parameters modified by the data), then the prior is conjugate to the likelihood. *(2 marks)*

$$p(\theta) \propto \theta^{\alpha-1}e^{-\beta\theta}$$

$$\begin{aligned} p(\theta|y) &\propto p(\theta)p(y|\theta) \\ &\propto \theta^{\alpha-1}e^{-\beta\theta}\theta^{\sum y_i}e^{-n\theta} \\ &= \theta^{\alpha+\sum y_i - 1}e^{-(\beta+n)\theta} \end{aligned}$$

$$\Rightarrow \theta|y \sim Ga\left(\alpha + \sum_{i=1}^{n} y_i, \beta + n\right)$$

i.e. the posterior is also gamma, therefore this prior is conjugate. *(3 marks)*

(c) Derive the Jeffreys' prior $p(\theta)$ for the parameter θ in this Poisson model. **[7 MARKS]**

Solution

Jeffreys' prior $p(\theta) = \sqrt{J(\theta)}$, where $J(\theta)$ is the Fisher information.

CONTINUED OVERLEAF/

$$J(\theta) = -\mathbb{E}\left[\left.\frac{\partial^2 \ln L(\theta)}{\partial \theta^2}\right| \theta\right]$$

$$\ln L(\theta) = \text{const.} + \left(\sum y_i\right) \ln \theta - n\theta = \text{const.} + n\bar{y} \ln \theta - n\theta$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n\bar{y}}{\theta} - n$$

$$\frac{\partial^2 \ln L(\theta)}{\partial \theta^2} = -\frac{n\bar{y}}{\theta^2}$$

$$\mathbb{E}\left[\left.\frac{\partial^2 \ln L(\theta)}{\partial \theta^2}\right| \theta\right] = -\frac{n}{\theta^2}\mathbb{E}\left[\bar{y}\right] = -\frac{n}{\theta^2} \cdot \theta = -\frac{n}{\theta}$$

$$J(\theta) = \frac{n}{\theta}$$

$$p(\theta) = \sqrt{\frac{n}{\theta}} \propto \theta^{-\frac{1}{2}}$$

(d) Is the Jeffreys' prior for θ a proper prior distribution? Explain. **[2 MARKS]**

Solution

$$\int_0^\infty \theta^{-\frac{1}{2}} d\theta = \left[2\theta^{\frac{1}{2}}\right]_0^\infty = \infty$$

i.e. the normalisation constant does not exist. Therefore the prior is improper.

(e) Is the posterior distribution that results from using the Jeffreys' prior in this problem a proper distribution? Explain. **[4 MARKS]**

Solution

The Jeffreys' prior is equivalent to $Ga\left(\frac{1}{2}, 0\right)$, as

$$Ga\left(\frac{1}{2}, 0\right) \propto \theta^{\frac{1}{2}-1} e^{-0 \cdot \theta} = \theta^{-\frac{1}{2}}$$

and therefore the posterior is

$$Ga\left(\frac{1}{2} + n\bar{y}, n\right)$$

and as long as $n \geq 1$, it is a proper (normalised) probability distribution. So, the posterior is proper.

CONTINUED OVERLEAF/

2. Consider the hierarchical model

$$\begin{aligned} y_{ij}|\theta,\beta &\sim \text{Exp}(\theta_j), \qquad i=1,\ldots,n_j;\ j=1,\ldots,J, \text{ independently;} \\ \theta_j|\beta &\sim \text{Ga}(\alpha,\beta), \qquad j=1,\ldots,J, \text{ independently;} \\ \beta &\sim \text{Exp}(\psi), \end{aligned}$$

where $\theta = (\theta_1,\ldots,\theta_J)$ and α and ψ are two fixed positive real numbers.

(a) Show that the joint probability density function of all the random quantities in the model is

$$p(\beta,\theta,y) = \psi\frac{\beta^{\alpha J}}{[\Gamma(\alpha)]^J}\left[\prod_{j=1}^J \theta_j^{\alpha+n_j-1}\right]\exp\left\{-\left[\psi\beta+\beta\sum_{j=1}^J\theta_j+\sum_{j=1}^J n_j\bar{y}_j\theta_j\right]\right\},$$

where $y = (y_{11},\ldots,y_{1n_1},\ldots,y_{J1},\ldots,y_{Jn_J})$ and $\bar{y}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{ij}$. **[4 MARKS]**

Solution

$$\begin{aligned} p(\beta,\theta,y) &= p(y|\theta,\beta)p(\theta|\beta)p(\beta) \\ &= \left[\prod_{j=1}^J\prod_{i=1}^{n_j} p(y_{ij}|\theta_j)\right]\left[\prod_{j=1}^J p(\theta_j|\beta)\right]p(\beta) \\ &= \left[\prod_{j=1}^J\prod_{i=1}^{n_j}\theta_j e^{-\theta_j y_{ij}}\right]\left[\prod_{j=1}^J\frac{\beta^\alpha}{\Gamma(\alpha)}\theta_j^{\alpha-1}e^{-\beta\theta_j}\right]\psi e^{-\psi\beta} \\ &= \left[\prod_{j=1}^J\prod_{i=1}^{n_j}\theta_j e^{-\theta_j y_{ij}}\right]\frac{\beta^{\alpha J}}{[\Gamma(\alpha)]^J}\left[\prod_{j=1}^J\theta_j^{\alpha-1}\right]e^{-\beta\sum\theta_j}\psi e^{-\psi\beta} \\ &= \psi\frac{\beta^{\alpha J}}{[\Gamma(\alpha)]^J}\left[\prod_{j=1}^J\theta_j^{\alpha+n_j-1}\right]\exp\left\{-\left[\sum_{j=1}^J n_j\theta_j\bar{y}_j+\beta\sum_{j=1}^J\theta_j+\psi\beta\right]\right\} \end{aligned}$$

as required.

(b) Find the full conditional distributions of β and θ. **[5 MARKS]**

Solution

Dropping irrelevant parts of the joint distribution, we obtain:

$$p(\beta|\theta,y)\propto p(\beta,\theta,y)\propto\beta^{\alpha J}\exp\left\{-\beta\left[\sum_{j=1}^J\theta_j+\psi\right]\right\}$$

So,

$$\beta|\theta,y\sim Ga\left(\alpha J+1,\sum_{j=1}^J\theta_j+\psi\right)$$

CONTINUED OVERLEAF/

$$
\begin{aligned}
p(\theta|\beta, y) &\propto p(\beta,\theta,y) \propto \left[\prod_{j=1}^{J} \theta_j^{\alpha+n_j-1}\right] exp\left\{-\left[\sum_{j=1}^{J} n_j\theta_j\bar{y}_j + \beta\sum_{j=1}^{J}\theta_j\right]\right\} \\
&= p(\beta,\theta,y) \propto \left[\prod_{j=1}^{J} \theta_j^{\alpha+n_j-1}\right]\left[\prod_{j=1}^{J} e^{-[n_j\bar{y}_j+\beta]\theta_j}\right]
\end{aligned}
$$

So,

$$\theta_j|\beta, y, \theta_{-j} \sim Ga\left(\alpha + n_j, \beta + n_j\bar{y}_j\right),$$

where $\theta_{-j} = (\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_J)$.

(c) Explain how the full conditional distributions could be used to implement a Gibbs sampler to draw from $p(\beta, \theta|y)$. **[3 MARKS]**

Solution

Initialise $\beta = \beta^{(0)}$ and $\theta = \theta^{(0)}$

for $k = 1, \ldots, K$

Draw $\beta^{(k)}$ from $Ga\left(\alpha J + 1, \sum_{j=1}^{J} \theta_j^{(k-1)} + \psi\right)$,

Draw $\theta_j^{(k)}$ from $Ga\left(\alpha + n_j, \beta^{(k)} + n_j\bar{y}_j\right)$ *(note $\beta^{(k)}$)*

Discard burn-in and optionally thin-out.

(d) In an Empirical Bayes approach, one would drop from the model the higher-level prior on β (third line) and instead estimate β from the data. Explain how you would do that. [Hint: recall that the expected value of a Ga(α, β) random variable is α/β.] **[4 MARKS]**

Solution

Estimate rates θ from sample means:

$$\hat{\theta}_j = \frac{1}{y_{\cdot j}} = \frac{1}{\frac{1}{n_j}\sum_{i=1}^{n_j} y_{ij}}$$

Since $\mathbb{E}\left[Ga(\alpha, \beta)\right] = \alpha/\beta$, estimate β as

$$\hat{\beta} = \frac{\alpha}{\text{average of } \hat{\theta}_j} = \frac{\alpha}{\frac{1}{J}\sum_{j=1}^{J}\frac{1}{\bar{y}_{\cdot j}}}$$

(e) Suppose that a sample $(\theta^{(t)}, \beta^{(t)})$, $t = 1, \ldots, T$, from the joint posterior distribution of θ and β is available. Explain how you can use it to compute an estimate of the posterior predictive distribution of $\tilde{y}_j$, a future observation from the jth group of observations. **[4 MARKS]**

CONTINUED OVERLEAF/

Solution

For each sample from the joint posterior distribution of θ and β, take a draw of

$$\tilde{y}_j^{(t)} \sim Exp\left(\theta_j^{(t)}\right)$$

we can use `rexp` in R.

The set $\left\{\tilde{y}_j^{(t)} : t = 1, \ldots, T\right\}$ are samples from the posterior predictive distribution of $\tilde{y}_j$.

3. The effectiveness of a proposed gene therapy for a genetic condition that affects the liver was explored in mice (prior to potential application in humans). In the ith of six replicate experiments, n_i ($i = 1, \ldots, 6$) mice with the liver condition were administered with the gene therapy and after a certain period of time the number y_i of mice with liver function improved by a certain amount was determined, with the following results:

Experiment, i	1	2	3	4	5	6
Sample size, n_i	91	88	102	96	110	113
Number improved, y_i	24	26	7	25	18	18

To explore the effect of the treatment, a Bayesian hierarchical model was fitted in `WinBUGS` with the following model code:

```
model {
  for (i in 1:6) {
    y[i] ~ dbin(theta[i], n[i])
    theta[i] ~ dbeta(alpha, beta)
  }
  alpha ~ dexp(1)
  beta ~ dexp(1)
}
```

where `y[i]` corresponds to y_i and `n[i]` corresponds to n_i.

(a) Convert the `WinBUGS` model specification into standard statistical notation, making it clear in your answer which random quantities are independent. **[4 MARKS]**

CONTINUED OVERLEAF/

Solution

$$\begin{aligned} y_i|\theta,\alpha,\beta &\sim Bin(n_i,\theta_i) \quad i=1,\ldots,6 \text{ independently}\\ \theta_i|\alpha,\beta &\sim Be(\alpha,\beta) \qquad i=1,\ldots,6 \text{ independently}\\ \alpha &\sim exp\,(1)\\ \beta &\sim exp\,(1) \end{aligned}$$

(b) Describe the model in words and state one advantage of explaining the data with a model which is hierarchical. **[3,1 MARKS]**

Solution

The θ_is represent the probability of improvement of an individual rat in replicate experiment i. The number of improvements in replicate experiment i is binomially distributed with the stated sample sizes (as number of trials n_i) and the θ_i as the probabilities of success, independently across replicate experiments.

The θ_is can vary between replicate experiments to model variability of other uncontrolled effects. The θ_is are modelled as coming independently from a beta distribution (chosen for the reasons of conjugacy) with unknown hyperparameters α and β, which themselves are given independent exponential priors.

One advantage of the hierarchical structure is 'borrowing strengths' (or 'sharing information'): we allow information from other replicates to influence estimation of θ_i for any particular replicate (via their effects on α and β).

(c) Some of the `WinBUGS` output is shown below:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha	1.086	0.4834	0.002433	0.3689	1.012	2.227	20001	100000
beta	3.084	1.518	0.007458	0.8676	2.839	6.698	20001	100000
theta[1]	0.2633	0.04501	1.532E-4	0.1802	0.2616	0.3562	20001	100000
theta[2]	0.294	0.04726	1.508E-4	0.206	0.2926	0.3908	20001	100000
theta[3]	0.07596	0.02578	8.242E-5	0.03364	0.07325	0.1336	20001	100000
theta[4]	0.2603	0.04365	1.525E-4	0.1792	0.2587	0.3499	20001	100000
theta[5]	0.1672	0.03489	1.093E-4	0.1049	0.1652	0.2409	20001	100000
theta[6]	0.163	0.034	1.103E-4	0.1021	0.1611	0.2346	20001	100000

(Below, the symbols θ_i, α and β will be used to represent the variables called `theta[i]`, `alpha` and `beta`, respectively, in `WinBUGS`.)

For each of the posterior summaries, (i. to iv.), listed below, state whether each can be determined using the `WinBUGS` output above. If your answer is "yes", explain why and compute the estimate. If your answer is "no", explain why not and explain how it could be computed in `WinBUGS`.

i. The posterior mean of $\theta_6 - \theta_5$. **[2 MARKS]**

CONTINUED OVERLEAF/

Solution

Yes, since $\mathbb{E}[\theta_6 - \theta_5|y] = \mathbb{E}[\theta_6|y] - \mathbb{E}[\theta_5|y] \approx 0.163 - 0.1672 = -0.0042$

ii. The posterior variance of $\theta_6 - \theta_5$. **[2 MARKS]**

Solution

No, since $\mathrm{Var}[\theta_6 - \theta_5|y] = \mathrm{Var}[\theta_6|y] + \mathrm{Var}[\theta_5|y] - 2\mathrm{Cov}[\theta_6, \theta_5|y]$ depends on posterior covariance of θ_6 and θ_5, and we don't have estimates of those.
We can get posterior correlation ρ of θ_6 and θ_5 from `WinBUGS`. For example,

```
correlation
theta[5] theta[6] 0.00215912
```

Then the covariance of θ_6 and θ_5 is $\mathrm{Cov}[\theta_6, \theta_5|y] = \rho \cdot sd[\theta_6|y]\, sd[\theta_5|y]$. And so $\mathrm{Var}[\theta_6 - \theta_5|y] = sd[\theta_6|y]^2 + sd[\theta_5|y]^2 - 2\rho \cdot sd[\theta_6|y]\, sd[\theta_5|y]$.

iii. The posterior mean of α/β. **[2 MARKS]**

Solution

No, since $\mathbb{E}[\alpha/\beta|y]$ cannot be expressed through available summary statistics for α and β.
We can add additional expression to the model specification:

```
pmab <- alpha/beta
```

monitor this quantity, and find posterior mean of `pmab` as out answer.

iv. The central 95% posterior interval for the log odds, $\log(\theta/[1-\theta])$, in experiment 2.

[2 MARKS]

Solution

Yes, since log odds is a 1-1 transformation of the probabilities of success, so we can apply this transformation to the quantiles of θ_2:
Central 95% posterior interval of $\log\left(\frac{\theta_2}{1-\theta_2}\right)$ can be estimated by

$$\left(\log\left(\frac{0.206}{1-0.206}\right), \log\left(\frac{0.3908}{1-0.3908}\right)\right)$$
$$= (-1.3492, -0.4440)$$

(d) Explain how you would modify the model to use independent inverse gamma priors on α and β. Inverse gamma distribution is not supported by `WinBUGS` directly, but it supports the Gamma distribution using function `dgamma(a,b)`. **[4 MARKS]**

Solution

The model should be

CONTINUED OVERLEAF/

```
model {
  for (i in 1:6) {
    y[i] ~ dbin(theta[i], n[i])
    theta[i] ~ dbeta(alpha, beta)
  }
  inv.alpha ~ dgamma(a,b)
  inv.beta ~ dgamma(c,d)
  alpha <- 1/inv.alpha
  beta <- 1/inv.beta
}
```

where `a, b, c,` and `d` can be set externally in the data file or the model description file.

---

4. (a) Imagine you are running a radio manufacturing company. You receive a shipment of transistors from a supplier to be used in the assembly of the final product. Checking the performance of each transistor in the shipment is too expensive, so you establish a sampling plan to either accept or reject the shipment. A random sample of 100 transistors from the whole shipment will be tested. Based upon the number y of defective transistors in that random sample, you will take one of the two decisions: a_1 = *accept the shipment,* or a_2 = *reject the shipment.*

As the test sample is relatively small in comparison to the shipment size, we are safe to assume:

$$y \sim Bi(100, \theta)$$

where θ is the proportion of the defective transistors in the shipment.

From your previous experience with this supplier, you know that θ is relatively small, and assume *a priori*:

$$\theta \sim Be(0.05, 1)$$

You select the following loss function:

$$L(\theta, a) = \begin{cases} 10\ \theta, & a = a_1 \\ 1, & a = a_2 \end{cases}$$

as if the shipment is rejected (a_2 decided) the loss is a constant cost due to inconvenience, delay, and testing of a replacement shipment; while if the shipment is accepted (a_1 decided) the loss is proportional to θ, since θ will also reflect the proportion of defective radios produced. The factor 10 indicates the relative costs involved in two types of errors.

You found on Wikipedia that

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = B(a, b), \quad \text{for } a \in \mathbb{R}, b \in \mathbb{R}.$$

CONTINUED OVERLEAF/

and you found an excerpt of the statistical table for B(a,b):

$B(a,b)$	b				
a	1	98	99	100	101
0.05	20.0000	15.4850	15.4771	15.4693	15.4616
1.05	0.9524	0.0079	0.0078	0.0077	0.0077
2.05	0.4878	8.3708×10^{-5}	8.1993×10^{-5}	8.0330×10^{-5}	7.8716×10^{-5}
3.05	0.3279	1.7152×10^{-6}	1.6634×10^{-6}	1.6136×10^{-6}	1.5659×10^{-6}
4.05	0.2469136	5.1769×10^{-8}	4.9714×10^{-8}	4.7761×10^{-8}	4.5902×10^{-8}

You receive a shipment that has $y = 2$ defective transistors in a random sample of 100.

i. What is your posterior for θ given this latest test result with $y = 2$? **[2 MARKS]**

Solution

The posterior for θ can be found using conjugate prior update:

$$p(\theta|y) = Be(0.05 + 2, 1 + 100 - 2) = Be(2.05, 99)$$

ii. Find the Bayes expected loss $p(\pi, a)$ for the decision on accepting this shipment. **[5 MARKS]**

Solution

$$\begin{aligned}
p(\pi, a) &= \int_0^\infty L(\theta, a) p(\theta|y) d\theta \\
&= \begin{cases} \int_0^\infty \frac{10}{B(2.05, 99)} \theta^{2.05} (1-\theta)^{98} d\theta, & a = a_1 \\ \int_0^\infty \frac{1}{B(2.05, 99)} \theta^{1.05} (1-\theta)^{98} d\theta, & a = a_2 \end{cases} \\
&= \begin{cases} \frac{10B(3.05, 99)}{B(2.05, 99)}, & a = a_1 \\ \frac{B(2.05, 99)}{B(2.05, 99)}, & a = a_2 \end{cases} \\
&\approx \begin{cases} 0.2029, & a = a_1 \\ 1 & a = a_2 \end{cases}
\end{aligned}$$

iii. Should you accept or reject this shipment? **[1 MARK]**

Solution

As $p(\pi, a)$ takes its minimum at a_1, that should be the preferred decision as a_1 is the Bayes action for this problem.

CONTINUED OVERLEAF/

Therefore, you should accept the shipment.

(b) What is the Bayes action for the absolute error loss $L(\theta, a) = |\theta - a|$? **[1 MARK]**

Solution

The Bayes action for the absolute error loss is the posterior median of θ.

(c) What is the Bayes action for the squared error loss $L(\theta, a) = (\theta - a)^2$? **[1 MARK]**

Solution

The Bayes action for the squared error loss is the posterior mean of θ.

(d) It is common to report results of an experiment rounded to certain precision. Consider studying average waiting time for a Route 6 bus at a bus stop next to the Botanic Gardens in Glasgow during commute hours. It is natural to consider an exponential model for this problem:

$$y_i \overset{i.i.d}{\sim} exp(\theta)$$

where y_i are the observed waiting times, and θ is the rate of the exponential distribution, $\mathbb{E}[y] = 1/\theta$.
Assume non-informative Gamma prior on θ:

$$\theta \sim Ga(1, 0)$$

I recorded three waiting times between buses as 5, 6, 10 minutes, rounded to a minute.

i. Infer the posterior distribution of θ considering observed waiting times as exact (not rounded). What is the posterior mean waiting time? **[2 MARKS]**

Solution

Assuming that the data are measured exactly, the posterior is

$$\theta|y \sim Ga(4, 21)$$

Therefore

$$\mathbb{E}[\theta|y] = \frac{4}{21} \approx 0.1905$$

And consequently, the posterior mean waiting time is $21/4 = 5.25$ minutes.

CONTINUED OVERLEAF/

ii. Infer the posterior distribution of θ taking rounding into account. What is the posterior mean waiting time in this case?

[5,3 MARKS]

Solution

$$\begin{aligned}
p(\theta|y) &= p(\theta|4.5 \leq y_1 < 5.5, 5.5 \leq y_2 < 6.5, 9.5 \leq y_3 < 10.5) \\
&\propto p(\theta)p(4.5 \leq y_1 < 5.5|\theta)p(5.5 \leq y_2 < 6.5|\theta)p(9.5 \leq y_3 < 10.5|\theta) \\
&= p(\theta)\int_{4.5}^{5.5} p(y|\theta)dy \int_{5.5}^{6.5} p(y|\theta)dy \int_{9.5}^{10.5} p(y|\theta)dy \\
&= \int_{4.5}^{5.5} \theta e^{-\theta y}dy \int_{5.5}^{6.5} \theta e^{-\theta y}dy \int_{9.5}^{10.5} \theta e^{-\theta y}dy \\
&= \left[-e^{-\theta y}\right]_{y=4.5}^{5.5} \left[-e^{-\theta y}\right]_{y=5.5}^{6.5} \left[-e^{-\theta y}\right]_{y=9.5}^{10.5} \\
&= \left(e^{-4.5\theta} - e^{-5.5\theta}\right)\left(e^{-5.5\theta} - e^{-6.5\theta}\right)\left(e^{-9.5\theta} - e^{-10.5\theta}\right) \\
&= \left(e^{-4.5\theta}e^{-5.5\theta} - e^{-4.5\theta}e^{-6.5\theta} - e^{-5.5\theta}e^{-5.5\theta} + e^{-5.5\theta}e^{-6.5\theta}\right) \\
&\quad \left(e^{-9.5\theta} - e^{-10.5\theta}\right) \\
&= \left(e^{-10\theta} - 2e^{-11\theta} + e^{-12\theta}\right)\left(e^{-9.5\theta} - e^{-10.5\theta}\right) \\
&\propto \left(e^{-19.5\theta} - 3e^{-20.5\theta} + 3e^{-21.5\theta} - e^{-22.5\theta}\right) \quad \textit{(3 marks)}
\end{aligned}$$

Finding normalising constant:

$$Z\int_0^\infty \left(e^{-19.5\theta} - 3e^{-20.5\theta} + 3e^{-21.5\theta} - e^{-22.5\theta}\right) d\theta = 1$$

$$\begin{aligned}
I &= \int_0^\infty \left(e^{-19.5\theta} - 3e^{-20.5\theta} + 3e^{-21.5\theta} - e^{-22.5\theta}\right) d\theta \\
&= \int_0^\infty e^{-19.5\theta}d\theta - 3\int_0^\infty e^{-20.5\theta}d\theta + 3\int_0^\infty e^{-21.5\theta}d\theta - \int_0^\infty e^{-22.5\theta}d\theta \\
&= \frac{1}{19.5} - \frac{3}{20.5} + \frac{3}{21.5} - \frac{1}{22.5} = \frac{2}{39} - \frac{6}{41} + \frac{6}{43} - \frac{2}{45} = \frac{32}{1031355}
\end{aligned}$$

$$Z = \frac{1031355}{32}$$

$$p(\theta|y) = \frac{1031355}{32}\left(e^{-19.5\theta} - 3e^{-20.5\theta} + 3e^{-21.5\theta} - e^{-22.5\theta}\right) \quad \textit{(2 marks)}$$

$$\begin{aligned}
\mathbb{E}\left[\theta|y\right] &= \frac{1031355}{32}\int_0^\infty \theta\left(e^{-19.5\theta} - 3e^{-20.5\theta} + 3e^{-21.5\theta} - e^{-22.5\theta}\right) d\theta \\
&= \frac{1031355}{32}\left(\frac{1}{19.5^2} - \frac{3}{20.5^2} + \frac{3}{21.5^2} - \frac{1}{22.5^2}\right) \\
&= \frac{1031355}{32} \cdot \frac{6304256}{1063693136025} = \frac{197008}{1031355} \approx 0.1910 \quad \textit{(2 marks)}
\end{aligned}$$

CONTINUED OVERLEAF/

And therefore, the posterior mean waiting time is $1031355/197008 \approx 5.2351$ minutes. *(1 mark)*

Total: 80

END OF QUESTION PAPER.