University of Glasgow

XXXX 2018
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. Let y be a number of draws required until the first success in a series of independent Bernoulli trials. Consider the geometric model with parameter θ:

$$p(y|\theta) = (1-\theta)^{y-1}\theta, \qquad 0 < \theta < 1, \quad y = 1, 2, \dots$$

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

Solution

$$\begin{aligned} L(\theta) &= p(y_1, \dots, y_n|\theta) = \prod_{i=1}^{n} p(y_i|\theta) = \prod_{i=1}^{n}(1-\theta)^{y_i-1}\theta \\ &= (1-\theta)^{\sum_{i=1}^{n}(y_i-1)}\theta^n = (1-\theta)^{\sum_{i=1}^{n} y_i - n}\theta^n \end{aligned}$$

(b) Explain the Bayesian concept of a conjugate prior. Show that the Be(α, β) distribution is conjugate for i.i.d. geometric data $y_1, \dots, y_n$ 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 beta 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)*

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

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

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

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

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

Solution

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

CONTINUED OVERLEAF/

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

$$\ln L(\theta) = \left(\sum y_i - n\right)\ln(1-\theta) + n\ln\theta = (n\bar{y} - n)\ln(1-\theta) + n\ln\theta$$

$$\frac{\partial \ln L(\theta)}{\partial\theta} = -\frac{n\bar{y}-n}{(1-\theta)} + \frac{n}{\theta} = \frac{n - \theta n - \theta n\bar{y} + \theta n}{(1-\theta)\theta} = \frac{n - \theta n\bar{y}}{(1-\theta)\theta}$$

$$\frac{\partial^2 \ln L(\theta)}{\partial\theta^2} = \frac{n - n\bar{y}}{(\theta-1)^2} - \frac{n}{\theta^2} = \frac{n}{(\theta-1)^2} - \frac{n\bar{y}}{(\theta-1)^2} - \frac{n}{\theta^2}$$

$$\begin{aligned}\mathbb{E}\left[\frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\middle|\theta\right] &= \frac{n}{(\theta-1)^2} - \frac{n}{(\theta-1)^2}\mathbb{E}\left[\bar{y}\right] - \frac{n}{\theta^2} = \frac{n\theta^2 - n\theta - n(\theta-1)^2}{(\theta-1)^2\theta^2} \\ &= \frac{n(\theta^2 - \theta - \theta^2 + 2\theta - 1)}{(\theta-1)^2\theta^2} = \frac{n(\theta-1)}{(\theta-1)^2\theta^2} = \frac{n}{(\theta-1)\theta^2}\end{aligned}$$

$$J(\theta) = \frac{n}{(1-\theta)\theta^2}$$

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

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

Solution

$$\begin{aligned}\int_0^1 \theta^{-1}(1-\theta)^{-1/2}d\theta &= \left[\log\left(1 - \sqrt{1-\theta}\right) - \log\left(\sqrt{1-\theta}+1\right)\right]_{\theta=0}^1 = \\ &= \ln(1) - \ln(1) - \ln(0) + \ln(2) = 0 - 0 + \infty + \log 2 = \infty\end{aligned}$$

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 $Be\,(0, 1/2)$, as

$$Be\left(0, \frac{1}{2}\right) \propto \theta^{-1}(1-\theta)^{1/2-1} = \theta^{-1}(1-\theta)^{-1/2}$$

CONTINUED OVERLEAF/

and therefore the posterior is

$$Be\left(n, 1/2 + \sum_{i=1}^{n} y_i - n\right)$$

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

2. Consider the hierarchical model

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

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_J)$ is a vector of unknown rates, and α and ψ are fixed positive real numbers.

(a) Derive the joint probability density function of all the random quantities in the model $p(\beta, \boldsymbol{\lambda}, y)$. **[4 MARKS]**

Solution

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

(b) Find the full conditional distributions of β and $\boldsymbol{\lambda}$. **[5 MARKS]**

CONTINUED OVERLEAF/

Solution

Dropping irrelevant parts of the joint distribution, we obtain:

$$p(\beta|\boldsymbol{\lambda}, y) \propto p(\beta, \boldsymbol{\lambda}, y) \propto \beta^{\alpha J} e^{-\psi\beta} \exp\left\{-\beta \sum_{j=1}^{J} \lambda_j\right\} \propto \beta^{\alpha J} \exp\left\{-\left(\psi + \sum_{j=1}^{J} \lambda_j\right)\beta\right\}$$

So,

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

$$p(\lambda_j|\beta, y) \propto p(\beta, \boldsymbol{\lambda}, y) \propto \left[\lambda_j^{n_j \bar{y}_j + \alpha - 1}\right] e^{-(n_j+\beta)\lambda_j}$$

So,

$$\lambda_j|\beta, y, \boldsymbol{\lambda}_{-j} \sim Ga\left(\alpha + n_j \bar{y}_j, \beta + n_j\right),$$

where $\boldsymbol{\lambda}_{-j} = (\lambda_1, \ldots, \lambda_{j-1}, \lambda_{j+1}, \ldots, \lambda_J)$.

(c) Explain how the full conditional distributions could be used to implement a Gibbs sampler to draw from $p(\beta, \boldsymbol{\lambda}|y)$. **[4 MARKS]**

Solution

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

for $k = 1, \ldots, K$

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

Draw $\lambda_j^{(k)}$ from $Ga\left(\alpha + n_j \bar{y}_j, \beta^{(k)} + n_j\right)$ *(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 λ from sample means:

$$\hat{\lambda}_j = \bar{y}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}$$

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

CONTINUED OVERLEAF/

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

(e) Suppose that a sample $(\boldsymbol{\lambda}^{(t)}, \beta^{(t)})$, $t = 1, \ldots, T$, from the joint posterior distribution of $\boldsymbol{\lambda}$ 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. **[3 MARKS]**

Solution

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

$$\tilde{y}_j^{(t)} \sim Poi\left(\lambda_j^{(t)}\right)$$

we can use `rpois` 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. (a) Suppose you perform ten independent Bernoulli trials and observe eight successes. Let θ denote the probability of success on each trial.

i. Compute the MAP estimate and posterior expectation of θ using $Be(0.5, 0.5)$ prior distribution. Note that the mode of the Beta distribution is $(\alpha-1)/(\alpha+\beta-1)$ when $\alpha, \beta > 1$. **[2 MARKS]**

Solution

The model is:

$$\begin{aligned} y &\sim Bin(\theta, n) \\ \theta &\sim Be(0.5, 0.5) \end{aligned}$$

We know that the prior is conjugate, and the posterior must be:

$$\theta | y \sim Be(0.5 + y, 0.5 + n - y) = Be(0.5 + 8, 0.5 + 10 - 8) = Be(8.5, 2.5)$$

The MAP estimate of θ is the mode of the posterior distribution, which is

$$\frac{\alpha_n - 1}{\alpha_n + \beta_n - 2} = \frac{7.5}{9} = \frac{5}{6} \approx 0.8333$$

And the posterior expectation of θ is the mean of the posterior:

$$\mathbb{E}\left[\theta | y\right] = \frac{\alpha_n}{\alpha_n + \beta_n} = \frac{8.5}{11} = \frac{17}{22} \approx 0.7727$$

CONTINUED OVERLEAF/

ii. Consider that we are using a mixture of Beta priors for this example:

$$p(\theta) = \frac{1}{2} Be\,(6,4) + \frac{1}{2} Be\,(4,6)$$

Derive the posterior using the same data $y = 8, n = 10$. Make sure that your posterior is properly normalised. **[9 MARKS]**

Solution

$$\begin{aligned}
p(\theta|y) &\propto p(\theta)p(y|\theta) \\
&\propto \theta^8(1-\theta)^2 \left[\frac{1}{2} Be\,(6,4) + \frac{1}{2} Be\,(4,6)\right] \\
&\propto \theta^8(1-\theta)^2 \left[\frac{\Gamma(10)}{2\Gamma(6)\Gamma(4)}\theta^5(1-\theta)^3 + \frac{\Gamma(10)}{2\Gamma(6)\Gamma(4)}\theta^3(1-\theta)^5\right] \\
&\propto \theta^{13}(1-\theta)^5 + \theta^{11}(1-\theta)^7 = B(14,6)Be(14,6) + B(12,8)Be(12,8)
\end{aligned}$$

To normalise it with some constant Z we need to solve:

$$\int_0^1 ZB(14,6)Be(\theta;14,6) + ZB(12,8)Be(\theta;12,8)d\theta = 1$$

therefore

$$\begin{aligned}
Z(B(14,6) + B(12,8)) &= 1 \\
Z &= \frac{1}{B(14,6) + B(12,8)}
\end{aligned}$$

and finally

$$\begin{aligned}
p(\theta|y) &= \frac{B(14,6)}{B(14,6)+B(12,8)}Be(14,6) + \frac{B(12,8)}{B(14,6)+B(12,8)}Be(12,8) \\
&= \frac{13!5!}{13!5!+11!7!}Be(14,6) + \frac{11!7!}{13!5!+11!7!}Be(12,8) \\
&= \frac{12\cdot 13}{12\cdot 13 + 6\cdot 7}Be(14,6) + \frac{6\cdot 7}{12\cdot 13+6\cdot 7}Be(12,8) \\
&= \frac{26}{33}Be(14,6) + \frac{7}{33}Be(12,8) \\
&\approx 0.787879Be(14,6) + 0.212121Be(12,8)
\end{aligned}$$

iii. Using the $Be(0.5, 0.5)$ prior, and after observing the data, what probability would you attach to the event that the eleventh trial will result in a success? **[4 MARKS]**

Solution

CONTINUED OVERLEAF/

The posterior predictive distribution for $\tilde{y}$ given the results of the previous question is

$$\begin{aligned}
Pr(\tilde{y} = \text{success}|y) &= \int_0^1 Pr(\tilde{y} = \text{success}|\theta) p(\theta|y) d\theta \\
&= \int_0^1 \theta \frac{1}{B(8.5, 2.5)} \theta^{7.5} (1-\theta)^{1.5} d\theta \\
&= \mathbb{E}\left[\theta|y\right] = \frac{17}{22} \approx 0.7727
\end{aligned}$$

$$\begin{aligned}
Pr(\tilde{y} = \text{failure}|y) &= 1 - Pr(\tilde{y} = \text{success}|y) \\
&= 1 - \frac{17}{22} = \frac{5}{22} \approx 0.2273
\end{aligned}$$

---

iv. You may believe that Beta priors are not flexible enough and prefer to specify a prior on θ by means of a normal prior on the logits:

$$\phi = logit(\theta) = \log \frac{\theta}{1-\theta} \qquad \phi \sim \mathcal{N}(\mu, \sigma^2)$$

You want the induced prior on θ to satisfy the following conditions:

$$Pr\left[\theta > 0.7\right] = 1/2 \qquad \text{and} \qquad Pr\left[\theta < 0.9\right] = 0.975$$

Determine the values of μ and σ^2 required. **[5 MARKS]**

Solution

$$\phi = logit(\theta) = \log \frac{\theta}{1-\theta} \qquad \Rightarrow \qquad \theta = \frac{e^{\phi}}{1+e^{\phi}}$$

$$\begin{aligned}
\theta > 0.7 \;&\Rightarrow\; \phi > \log \frac{0.7}{0.3} \approx 0.847 \\
\theta < 0.9 \;&\Rightarrow\; \phi < \log \frac{0.9}{0.1} \approx 2.197
\end{aligned}$$

We defined a normal prior on ϕ:

$$\phi \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

Now we need to find μ and σ^2 to match the required constraints.

$$\begin{aligned}
Pr\left[\theta > 0.7\right] = 1/2 \;&\Rightarrow\; Pr\left[\phi > 0.847\right] = 1/2 \\
&\Rightarrow\; 1 - \Phi\left(\frac{0.847 - \mu}{\sigma}\right) = 0.5 \\
&\Rightarrow\; \frac{0.847 - \mu}{\sigma} = \Phi^{-1}(0.5) = 0 \\
&\Rightarrow\; \mu = 0.847
\end{aligned}$$

CONTINUED OVERLEAF/

$$
\begin{aligned}
Pr\left[\theta < 0.9\right] = 0.975 \;\; &\Rightarrow \;\; Pr\left[\phi < 2.197\right] = 0.975 \\
&\Rightarrow \;\; \Phi\left(\frac{2.197 - 0.847}{\sigma}\right) = 0.975 \\
&\Rightarrow \;\; \frac{1.35}{\sigma} = \Phi^{-1}(0.975) = 1.96 \\
&\Rightarrow \;\; \sigma = \frac{1.35}{1.96} = 0.689 \\
&\Rightarrow \;\; \sigma^2 = 0.474721
\end{aligned}
$$

4. (a) Consider the following model:

$$
\begin{aligned}
y &\sim Bin(n, \theta) \\
\theta &\sim Be(\alpha, \beta)
\end{aligned}
$$

Using the following loss function:

$$
L(\theta, a) = \frac{(\theta - a)^2}{\theta(1-\theta)}
$$

i. Derive the expression for the Bayes action. **[8 MARKS]**

Solution

Expected loss:

$$
\begin{aligned}
\rho(\pi, a) &= \int_0^1 \frac{(\theta - a)^2}{\theta(1-\theta)} \times \frac{1}{B(\alpha + y, \beta + n - y)} \times \theta^{\alpha+y-1}(1-\theta)^{\beta+n-y-1} d\theta \\
&= \mathbb{E}\left[\left.\frac{(\theta - a)^2}{\theta(1-\theta)}\right| y\right] \\
&= \mathbb{E}\left[\left.\frac{\theta^2}{\theta(1-\theta)}\right| y\right] - 2a\mathbb{E}\left[\left.\frac{\theta}{\theta(1-\theta)}\right| y\right] + a^2\mathbb{E}\left[\left.\frac{1}{\theta(1-\theta)}\right| y\right]
\end{aligned}
$$

When expected loss is minimised, its derivative is 0, therefore to find the

CONTINUED OVERLEAF/

Bayes action, we need to solve the equation:

$$
\begin{aligned}
\frac{\partial \rho}{\partial a} =& 0 \\
\frac{\partial \rho}{\partial a} =& -2\mathbb{E}\left[\frac{1}{1-\theta}\middle| y\right] + 2a\mathbb{E}\left[\frac{1}{\theta(1-\theta)}\middle| y\right] \\
=& -2\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+y)\Gamma(\beta+n-y)}\int_0^1 \theta^{\alpha+y-1}(1-\theta)^{\beta+n-y-2}d\theta \\
&+ 2a\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+y)\Gamma(\beta+n-y)}\int_0^1 \theta^{\alpha+y-2}(1-\theta)^{\beta+n-y-2}d\theta \\
=& -2\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+y)\Gamma(\beta+n-y)} \times \frac{\Gamma(\alpha+y)\Gamma(\beta+n-y-1)}{\Gamma(\alpha+\beta+n-1)} \\
&+ 2a\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+y)\Gamma(\beta+n-y)} \times \frac{\Gamma(\alpha+y-1)\Gamma(\beta+n-y-1)}{\Gamma(\alpha+\beta+n-2)} \\
=& -2\frac{\alpha+\beta+n-1}{\beta+b-y-1} + 2a\frac{(\alpha+\beta+n-2)(\alpha+\beta+n-1)}{(\alpha+y-1)(\beta+n-y-1)} = 0
\end{aligned}
$$

$$
\begin{aligned}
a =& \frac{2(\alpha+\beta+n-1)}{\beta+n-y-1} \times \frac{(\alpha+y-1)(\beta+n-y-1)}{2(\alpha+\beta+n-2)(\alpha+\beta+n-1)} \\
a =& \frac{\alpha+y-1}{\alpha+\beta+n-2}
\end{aligned}
$$

This is our resulting Bayes action.

ii. Evaluate this Bayes action when $\alpha = \beta = 1$, and comment on your result. **[2 MARKS]**

Solution

When $\alpha = \beta = 1$, the Bayes action is

$$\frac{y}{n}$$

This demonstrates that when using loss function of the form

$$L(\theta, a) = \frac{(\theta - a)^2}{\theta(1-\theta)}$$

under uniform prior, the Bayes action is the maximum likelihood estimator of θ.

(b) Six batteries are put on test for 200 hours. Four fail after 96, 130, 160, and 180 hours. The other two batteries are still working after 200 hours. Assume the lifetime of these batteries has an exponential distribution with mean θ^{-1}. Assume that the battery lifetimes are independent of each other.

CONTINUED OVERLEAF/

i. Derive the likelihood for this problem. **[3 MARKS]**

Solution

The overall likelihood is the product of individual likelihoods:

$$y_i \sim exp(\theta), \quad i = 1, \dots, 6$$

$$\begin{aligned}
p(y|\theta) =& p(96|\theta)p(130|\theta)p(160|\theta)p(180|\theta) \cdot \\
& \cdot p(y > 200|\theta)p(y > 200|\theta) \\
=& \theta e^{-96\theta} \cdot \theta e^{-130\theta} \cdot \theta e^{-160\theta} \cdot \theta e^{-180\theta} \cdot \int_{200}^{\infty} \theta e^{-y\theta} dy \cdot \int_{200}^{\infty} \theta e^{-y\theta} dy \\
=& \theta^4 e^{-566\theta} \cdot \left[-e^{-\theta y}\right]_{y=200}^{\infty} \cdot \left[-e^{-\theta y}\right]_{y=200}^{\infty} \\
=& \theta^4 e^{-566\theta} \cdot \left[0 + e^{-200\theta}\right] \cdot \left[0 + e^{-200\theta}\right] = \theta^4 e^{-966\theta}
\end{aligned}$$

ii. Assuming the gamma prior

$$p(\theta) = \frac{\beta^\alpha \theta^{\alpha-1} \exp(-\beta\theta)}{\Gamma(\alpha)}$$

with parameters $\alpha = 10, \beta = 1800$. Find the posterior mean and variance. **[3 MARKS]**

Solution

$$p(\theta|y) \propto p(\theta)p(y|\theta) \propto \theta^9 e^{-1800\theta} \cdot \theta^4 e^{-966\theta} = \theta^{13} e^{-2766\theta} \propto Gam(\theta; 14, 2766)$$

Mean of the posterior is $\alpha/\beta = 14/2766 = 7/1383 \approx 0.00506$.
The variance of the posterior is $\alpha/\beta^2 = 14/7650756 = 7/3825378 \approx 1.830 \times 10^{-6}$.

iii. Find the density of the predictive distribution of another independent battery lifetime and find the probability it is more than 200 hours. **[4 MARKS]**

Solution

$$\begin{aligned}
p(\tilde{y}|y) &= \int_0^{\infty} p(\tilde{y}|\theta)p(\theta|y)d\theta = \int_0^{\infty} \frac{2766^{14}}{\Gamma(14)} \theta e^{-\tilde{y}\theta} \theta^{13} e^{-2766\theta} d\theta \\
&= \frac{2766^{14}}{\Gamma(14)} \int_0^{\infty} \theta^{14} e^{-\theta(2766+\tilde{y})} d\theta = \frac{2766^{14}}{\Gamma(14)} \times \frac{\Gamma(15)}{(2766 + \tilde{y})^{15}} \\
&= \frac{14 \cdot 2766^{14}}{(2766 + \tilde{y})^{15}}
\end{aligned}$$

which is Lomax distribution with $\lambda = 2766$, and $\alpha = 14$ (but the students don't have to know the name of this distribution).

CONTINUED OVERLEAF/

$$
\begin{aligned}
p(\tilde{y} > 200|y) &= \int_{200}^{\infty} \frac{14 \cdot 2766^{14}}{(2766+\tilde{y})^{15}} d\tilde{y} = 14 \cdot 2766^{14} \cdot \left[-\frac{1}{14(2766+\tilde{y})^{14})} \right]_{\tilde{y}=200}^{\infty} \\
&= 14 \cdot 2766^{14} \cdot \left[0 + \frac{1}{14 \cdot 2966^{14}} \right] = \frac{2766^{14}}{2966^{14}} \approx 0.3763
\end{aligned}
$$

Total: 80

END OF QUESTION PAPER.