DEPARTMENT OF MATHEMATICAL SCIENCES
Outline solution to Examination Questions

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS
Academic Year 2019/20	Examiner SIMON SHAW
Semester II	Question No. 1 — Page 1 of 4

Part		Mark
(a)	Let Π denote the class of prior distributions. Conjugacy means that, given the likelihood $f(\theta \mid x)$, the posterior distribution will also be in the class Π for all x.	2
(b)	In order for a conjugate family to exist, the likelihood $\prod_{i=1}^{n} f(x_i \mid \theta)$ must involve only a finite number of different functions of $x = (x_1, \ldots, x_n)$ for n arbitrarily large. Thus the likelihood must contain a finite number of sufficient statistics which implies (given regularity conditions) that the likelihood is a member of a regular exponential family.	3

Total

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS
Academic Year 2019/20	Examiner SIMON SHAW
Semester II	Question No. 1 — Page 2 of 4

Part | Mark

(c) $$f(x \mid \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi v}} \exp\left\{-\frac{1}{2v}(x_i-\mu)^2\right\}$$
$$= (2\pi)^{-\frac{n}{2}} v^{-\frac{n}{2}} \exp\left\{-\frac{1}{2v}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

Now, $$\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}((x_i-\bar{x})+(\bar{x}-\mu))^2$$
$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2 = (n-1)s^2 + n(\bar{x}-\mu)^2$$

Thus, $$f(x \mid \theta) = (2\pi)^{-\frac{n}{2}} v^{-\frac{n}{2}} \exp\left\{-\frac{1}{2v}[(n-1)s^2+n(\bar{x}-\mu)^2]\right\}$$
$$\propto v^{-\frac{n}{2}} \exp\left\{-\frac{1}{2v}[(n-1)s^2+n(\bar{x}-\mu)^2]\right\}$$
$$= g(\theta, \bar{x}, s^2).$$

Hence, as $f(\theta \mid x) \propto f(x \mid \theta) f(\theta) \propto g(\theta, \bar{x}, s^2) f(\theta)$, $\bar{X}$ and S^2 are sufficient for X for learning about θ. 3

(d)(i) $$f(\mu, v \mid x) \propto v^{-\frac{n}{2}} \exp\left\{-\frac{1}{2v}[(n-1)s^2+n(\bar{x}-\mu)^2]\right\} v^{-1}$$
$$= v^{-(\frac{n}{2}+1)} \exp\left\{-\frac{1}{2v}[(n-1)s^2+n(\bar{x}-\mu)^2]\right\}$$

Total

DEPARTMENT OF MATHEMATICAL SCIENCES
Outline solution to Examination Questions

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 1	Page 3 of 4

Part | Mark

Thus, $f(\nu \mid x) \propto \int_{-\infty}^{\infty} \nu^{-(\frac{n}{2}+1)} \exp\left\{-\frac{1}{2\nu}[(n-1)s^2 + n(\bar{x}-\mu)^2]\right\} d\mu$

$$= \nu^{-(\frac{n}{2}+1)} \exp\left\{-\frac{1}{2\nu}(n-1)s^2\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\nu} n(\bar{x}-\mu)^2\right\} d\mu$$

We recognise the integrand as a kernel of $N(\bar{x}, \nu/n)$ so that

$$f(\nu \mid x) \propto \nu^{-(\frac{n}{2}+1)} \exp\left\{-\frac{1}{2\nu}(n-1)s^2\right\} \nu^{\frac{1}{2}}$$

$$= \nu^{-(\frac{n}{2}-\frac{1}{2}+1)} \exp\left\{-\frac{1}{2\nu}(n-1)s^2\right\}$$

Only marks for the whole of the part given. 4

We recognise this as a kernel of Inv-gamma$\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$ and thus this is the distribution of $\nu \mid x$.

(ii) Note that if $Z \sim$ Inv-gamma(α, β) then $Z^{-1} \sim$ Gamma(α, β).

Hence, $\nu^{-1} \mid x \sim$ Gamma$\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right)$ and, from the hint,

$(n-1)s^2\nu^{-1} \mid x \sim$ Gamma$\left(\frac{n-1}{2}, \frac{1}{2}\right)$ which is the chi-squared distribution with $(n-1)$ degrees of freedom. Hence, $Y \mid x \sim \chi^2_{n-1}$.

(total marks for (a)(i) and (ii)) 8

This corresponds to the pivot $\frac{(n-1)S^2}{\nu}$ used to construct confidence intervals in the classical model. Thus, classical and Bayesian inference will coincide showing that the prior is noninformative.

Only marks for whole part given. 4

Total

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 1	Page 4 of 4

Part | Mark

(e) $f(\mu,\nu) = f(\mu|\nu)f(\nu)$

$$\propto \nu^{-\frac{1}{2}}\exp\left\{-\frac{\lambda}{2\nu}(\mu-\mu_0)^2\right\}\nu^{-(\alpha+1)}\exp\left\{-\frac{\beta}{\nu}\right\}$$

$$= \nu^{-(\alpha+\frac{3}{2})}\exp\left\{-\frac{1}{2\nu}\left[\lambda(\mu-\mu_0)^2+2\beta\right]\right\}$$

$$f(\mu,\nu|x) \propto \nu^{-(\alpha+\frac{n}{2}+\frac{3}{2})}\exp\left\{-\frac{1}{2\nu}\left[\lambda(\mu-\mu_0)^2+n(\mu-\bar{x})^2+2\beta+(n-1)s^2\right]\right\}$$

Now, from the hint, $\lambda(\mu-\mu_0)^2+n(\mu-\bar{x})^2 = (\lambda+n)\left(\mu-\frac{\lambda\mu_0+n\bar{x}}{\lambda+n}\right)^2 + \frac{n\lambda}{\lambda+n}(\mu_0-\bar{x})^2$

so that

$$f(\mu,\nu|x) \propto \nu^{-(\alpha+\frac{n}{2}+\frac{3}{2})}\exp\left\{-\frac{1}{2\nu}\left[(\lambda+n)\left(\mu-\frac{\lambda\mu_0+n\bar{x}}{\lambda+n}\right)^2 + 2\left(\beta + \frac{(n-1)s^2}{2} + \frac{n\lambda}{2(\lambda+n)}(\mu_0-\bar{x})^2\right)\right]\right\}$$

which is of the same form as the prior giving conjugacy.

[In terms of the hyperparameters, $\alpha \mapsto \alpha+\frac{n}{2}$, $\lambda \mapsto \lambda+n$, $\mu_0 \mapsto \frac{\lambda\mu_0+n\bar{x}}{\lambda+n}$, $\beta \mapsto \beta + \frac{(n-1)s^2}{2} + \frac{n\lambda}{2(\lambda+n)}(\mu_0-\bar{x})^2$.] 4

Total 20

DEPARTMENT OF MATHEMATICAL SCIENCES
Outline solution to Examination Questions

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 2	Page 1 of 4

Part | Mark

(a) $$f(x \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta) = \prod_{i=1}^{n} \theta \exp(-x_i \theta), \quad \theta > 0$$
$$= \theta^n \exp\left\{-\left(\sum_{i=1}^{n} x_i\right)\theta\right\}$$

$$\log f(x \mid \theta) = n \log \theta - \left(\sum_{i=1}^{n} x_i\right)\theta$$

$$\frac{\partial}{\partial \theta} \log f(x \mid \theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i, \quad \frac{\partial^2}{\partial \theta^2} \log f(x \mid \theta) = -\frac{n}{\theta^2}$$

$$I(\theta) = -E\left(-\frac{n}{\theta^2} \,\middle|\, \theta\right) = \frac{n}{\theta^2}.$$

The Jeffreys prior is $f(\theta) \propto \sqrt{I(\theta)}$
$\propto \theta^{-1}$ (the improper Gamma$(0,0)$)

$$f(\theta \mid x) \propto f(x \mid \theta) f(\theta)$$
$$\propto \theta^n \exp\left\{-\left(\sum_{i=1}^{n} x_i\right)\theta\right\} \theta^{-1}$$
$$= \theta^{n-1} \exp\left\{-\left(\sum_{i=1}^{n} x_i\right)\theta\right\}$$

We recognise this as a kernel of Gamma$\left(n, \sum_{i=1}^{n} x_i\right)$ so that $\theta \mid x \sim \text{Gamma}(\alpha_n, \beta_n)$ where $\alpha_n = n$ and $\beta_n = \sum_{i=1}^{n} x_i$.

Total

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 2	Page 2 of 4

Part | Mark

$E(\theta|X) = \dfrac{\alpha_n}{\beta_n} = \dfrac{n}{\sum_{i=1}^{n} x_i}$ which is the maximum likelihood estimate of θ. The posterior mean is thus based solely on the data and the prior has no influence; it is noninformative. 6

(b) As Z is exchangeable then $Z|\theta \sim Exp(\theta)$ and $(X \perp\!\!\!\perp Z)|\theta$. Thus,

$$\begin{aligned} E(Z|X) &= E(E(Z|\theta)|X) \quad (\text{using } (X \perp\!\!\!\perp Z)|\theta) \\ &= E(\theta^{-1}|X) \quad (\text{as } Z|\theta \sim Exp(\theta)) \\ &= \frac{\beta_n}{\alpha_n - 1} = \frac{\sum_{i=1}^{n} x_i}{n-1} \quad (\text{as } \theta|x \sim Gamma(\alpha_n, \beta_n) \text{ so } \theta^{-1}|x \sim \text{Inv-gamma}(\alpha_n, \beta_n)) \end{aligned}$$

Similarly,

$$\begin{aligned} Var(Z|X) &= E(Var(Z|\theta)|X) + Var(E(Z|\theta)|X) \\ &= E(\theta^{-2}|X) + Var(\theta^{-1}|X) \\ &= \frac{\beta_n^2}{(\alpha_n-1)(\alpha_n-2)} + \frac{\beta_n^2}{(\alpha_n-1)^2(\alpha_n-2)} \\ &= \frac{\alpha_n \beta_n^2}{(\alpha_n-1)^2(\alpha_n-2)} = \frac{n\left(\sum_{i=1}^{n} x_i\right)^2}{(n-1)^2(n-2)} \end{aligned}$$

(both values well defined for $n > 2$). 3

Total	

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 2	Page 3 of 4

Part | Mark

~~(iv)~~ (c)

$$f(\theta) = \frac{\beta^{\alpha}}{4\Gamma(\alpha)}\theta^{\alpha-1}\exp(-\beta\theta) + \frac{3\beta^{\alpha+1}}{4\Gamma(\alpha+1)}\theta^{\alpha}\exp(-\beta\theta)$$

$$\propto \theta^{\alpha-1}\exp(-\beta\theta) + \frac{3\beta}{\alpha}\theta^{\alpha}\exp(-\beta\theta)$$

$$\propto \alpha\,\theta^{\alpha-1}\exp(-\beta\theta) + 3\beta\,\theta^{\alpha}\exp(-\beta\theta).$$

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

$$\propto \alpha\,\theta^{(\alpha+n)-1}\exp\left(-\left(\beta+\sum_{i=1}^{n}x_i\right)\theta\right) + 3\beta\,\theta^{(\alpha+1+n)-1}\exp\left(-\left(\beta+\sum_{i=1}^{n}x_i\right)\theta\right)$$

We recognise kernels of $Gamma(\alpha+n, \beta+\sum_{i=1}^{n}x_i)$ and $Gamma(\alpha+1+n, \beta+\sum_{i=1}^{n}x_i)$ in this sum. Letting $f_1(\theta)$ be the pdf of $Gamma(\alpha+n, \beta+\sum_{i=1}^{n}x_i)$ and $f_2(\theta)$ the pdf of $Gamma(\alpha+1+n, \beta+\sum_{i=1}^{n}x_i)$ we have:

$$f(\theta\mid x) \propto \alpha\frac{\Gamma(\alpha+n)}{(\beta+\sum_{i=1}^{n}x_i)^{\alpha+n}}f_1(\theta) + \frac{3\beta\,\Gamma(\alpha+n+1)}{(\beta+\sum_{i=1}^{n}x_i)^{\alpha+n+1}}f_2(\theta)$$

$$\propto \alpha f_1(\theta) + \frac{3\beta(\alpha+n)}{(\beta+\sum_{i=1}^{n}x_i)}f_2(\theta)$$

$$\propto \alpha\left(\beta+\sum_{i=1}^{n}x_i\right)f_1(\theta) + 3\beta(\alpha+n)f_2(\theta)$$

Total

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 2	Page 4 of 4

Part | Mark

Thus, as $\int_\theta f_1(\theta)\,d\theta = 1 = \int_\theta f_2(\theta)\,d\theta$ we have:

$$f(\theta \mid x) = c f_1(\theta) + (1-c) f_2(\theta) \text{ where } c = \frac{\alpha(\beta + \sum_{i=1}^n x_i)}{\alpha(\beta+\sum_{i=1}^n x_i) + 3\beta(\alpha+n)}$$

5

(d)(i) If $X_1, \ldots, X_n$ are finitely exchangeable then the labelling of the outcomes in the joint distribution is uninformative and only the outcomes matter and not the random variable of which they were a realisation of.

i.e. $f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$ for all permutations π on $\{1, \ldots, n\}$.

(ii) $$f(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^n f(x_i \mid \theta)$$

$$= \prod_{i=1}^n f(x_i \mid \theta_i)$$

Now, $$f(x_1, \ldots, x_n) = \int_\theta f(x \mid \theta) f(\theta)\,d\theta$$

$$= \int_\theta \left\{\prod_{i=1}^n f(x_i \mid \theta_i)\right\}\left\{\prod_{i=1}^n \pi(\theta_i)\right\} d\theta$$

$$= \prod_{i=1}^n \int_{\theta_i} f(x_i \mid \theta_i)\pi(\theta_i)\,d\theta_i = \prod_{i=1}^n f(x_i)$$

as each marginal X_i is identically distributed. Thus, the ordering of the x_i's is irrelevant (they're iid) and thus they are finitely exchangeable.

6

Total 20

DEPARTMENT OF MATHEMATICAL SCIENCES
Outline solution to Examination Questions

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT.

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 3	Page 1 of ~~4~~ 5

Part | Mark

(a) As $X_i \mid \theta \sim Po(s_i\theta_i)$ then $\mathbb{P}(X_i = x_i \mid \theta) = \frac{1}{x_i!}(s_i\theta_i)^{x_i}\exp(-s_i\theta_i)$

$$\propto \theta_i^{x_i}\exp(-s_i\theta_i)$$

Hence, $f(x\mid\theta) = \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n \mid \theta) \propto \prod_{i=1}^{n} \theta_i^{x_i}\exp(-s_i\theta_i)$

$$= \left(\prod_{i=1}^{n}\theta_i^{x_i}\right)\exp\left(-\sum_{i=1}^{n}s_i\theta_i\right)$$

$\theta_i \mid \phi \sim Exp(\phi)$ so that $f(\theta\mid\phi) = \prod_{i=1}^{n}\phi\exp(-\phi\theta_i) = \phi^n\exp\left(-\phi\sum_{i=1}^{n}\theta_i\right)$

$\phi \sim Gamma(\alpha, \beta)$ so that $f(\phi) \propto \phi^{\alpha-1}\exp(-\beta\phi)$

Then, $f(\theta, \phi \mid x) \propto f(x\mid\theta)f(\theta\mid\phi)f(\phi)$

$$\propto \left(\prod_{i=1}^{n}\theta_i^{x_i}\right)\phi^{\alpha+n-1}\exp\left\{-\phi\left(\beta + \sum_{i=1}^{n}\theta_i\right) - \sum_{i=1}^{n}s_i\theta_i\right\}.$$

2

(b) Let $\theta_{-i} = (\theta_1, \ldots, \theta_n) / \theta_i$.

$$f(\theta_i \mid \theta_{-i}, \phi, x) = \frac{f(\theta, \phi\mid x)}{f(\theta_{-i}, \phi \mid x)} \propto f(\theta, \phi \mid x)$$

as a function of θ_i

Total

Part | Mark

Thus, $f(\theta_i \mid \theta_{-i}, \phi, x) \propto \theta_i^{x_i} \exp\{-(\phi + s_i)\theta_i\}$
We recognise this as a kernel of Gamma$(x_i + 1, \phi + s_i)$ so

$$\theta_i \mid \theta_{-i}, \phi, x \sim \text{Gamma}(x_i + 1, \phi + s_i).$$

Similarly, $f(\phi \mid \theta, x) \propto f(\theta, \phi \mid x)$ (wrt ϕ)

$$\propto \phi^{\alpha + n - 1} \exp\left\{-\phi\left(\beta + \sum_{i=1}^{n} \theta_i\right)\right\}$$

We recognise this as a kernel of Gamma$\left(\alpha + n, \beta + \sum_{i=1}^{n} \theta_i\right)$ so

$$\phi \mid \theta, x \sim \text{Gamma}\left(\alpha + n, \beta + \sum_{i=1}^{n} \theta_i\right).$$

The Gibbs sampler algorithm is:

1. Choose a starting value $(\theta^{(0)}, \phi^{(0)})$ for which $f(\theta^{(0)}, \phi^{(0)} \mid x) > 0$
2. At iteration t generate new values $(\theta^{(t)}, \phi^{(t)})$ as follows
   - draw $\theta_1^{(t)}$ from Gamma$(x_1 + 1, \phi^{(t-1)} + 1)$

     $\vdots$
   - draw $\theta_n^{(t)}$ from Gamma$(x_n + 1, \phi^{(t-1)} + 1)$
   - draw $\phi^{(t)}$ from Gamma$\left(\alpha + n, \beta + \sum_{i=1}^{n} \theta_i^{(t)}\right)$
3. Repeat step 2.

Total

Part | Mark

The algorithm will produce a Markov chain with stationary distribution $f(\theta, \phi | x)$. After a sufficiently long time to allow for convergence, values $\{\theta^{(t)}, \phi^{(t)}\}$ for $t > b$ may be viewed as a sample from $f(\theta, \phi | x)$ (b denotes the burn-in). 7

(c) We wish to sample from $\theta_i \,|\, \phi, \theta_{-i}, x \sim Gamma(x_i + 1, \phi + s_i)$ using the Metropolis-Hastings algorithm. At time t of the Gibbs sampler we sample from $Gamma(x_i + 1, \phi^{(t-1)} + s_i)$ and thus our target distribution is $\pi(\theta_i) \propto \theta_i^{x_i} \exp(-(\phi^{(t-1)} + s_i)\theta_i)$

The M-H algorithm is thus:

1) Choose an arbitrary starting point $\theta_i^{(0)}$ for which $\pi(\theta_i) > 0$.

2) At time s,

a) Sample proposal θ_i^* from $q(\theta_i \,|\, \theta_i^{(s-1)})$, the proposal distribution.

b) Calculate the acceptance probability

$$\alpha(\theta_i^{(s-1)}, \theta_i^*) = \min\left(1, \frac{\pi(\theta_i^*)\, q(\theta_i^{(s-1)} \,|\, \theta_i^*)}{\pi(\theta_i^{(s-1)})\, q(\theta_i^* \,|\, \theta_i^{(s-1)})}\right)$$

$$= \min\left(1, \frac{\theta_i^{*\,x_i} \exp(-(\phi^{(t-1)} + s_i)\theta_i^*)\, q(\theta_i^{(s-1)} \,|\, \theta_i^*)}{\theta_i^{(s-1)\,x_i} \exp(-(\phi^{(t-1)} + s_i)\theta_i^{(s-1)})\, q(\theta_i^* \,|\, \theta_i^{(s-1)})}\right)$$

Total

Part		Mark

c) Generate $U \sim U(0,1)$

d) If $U \leq \alpha(\theta_i^{(s-1)}, \theta_i^*)$ accept the proposal, $\theta_i^{(s)} = \theta_i^*$.
Otherwise reject the proposal and set $\theta_i^{(s)} = \theta_i^{(s-1)}$.

3) Repeat step 2).

The chain is run until convergence. After this observations are from $\pi(\theta_i)$. Thus, if b denotes the length of the burn-in, $\theta_i^{(b+1)}$ is a sample from $\pi(\theta_i)$ and so, in the Gibbs sampler, set $\theta_i^{(t)} = \theta_i^{(b+1)}$. 4

(d) $$P(\theta > a \mid x) = \int_a^\infty f(\theta \mid x)\, d\theta = \int_{-\infty}^{\infty} \mathbb{I}_{\{\theta > a\}} f(\theta \mid x)\, d\theta$$

$$= \int_{-\infty}^{\infty} \frac{\mathbb{I}_{\{\theta > a\}} f(\theta \mid x)}{q(\theta)} q(\theta)\, d\theta = E_q\left(\frac{\mathbb{I}_{\{\theta > a\}} f(\theta \mid x)}{q(\theta)}\right)$$

Draw a random sample $\theta_1, \ldots, \theta_N$ from $q(\theta)$ and estimate $P(\theta > a \mid x)$ by $$\hat{I} = \frac{1}{N} \sum_{i=1}^{N} \frac{\mathbb{I}_{\{\theta_i > a\}} f(\theta_i \mid x)}{q(\theta_i)}.$$ 3

(e) $$Var(\hat{I} \mid \theta \sim q(\theta)) = \frac{1}{N} Var\left(\frac{g(\theta) f(\theta \mid X)}{q(\theta)} \,\Big|\, \theta \sim q(\theta)\right)$$

as θ_i iid from $q(\theta)$. Thus,

$$Var(\hat{I} \mid \theta \sim q(\theta)) = \frac{1}{N}\left\{E\left(\frac{g^2(\theta) f^2(\theta \mid X)}{q^2(\theta)} \,\Big|\, \theta \sim q(\theta)\right) - E^2\left(\frac{g(\theta) f(\theta \mid X)}{q(\theta)} \,\Big|\, \theta \sim q(\theta)\right)\right\}$$

	Total	

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 3	Page 5 of 5

Part		Mark
	$= \frac{1}{N}\left\{E\left(\frac{g^2(\theta)f(\theta\|X)}{q(\theta)}\,\middle\|\,X\right) - E^2(g(\theta)\|X)\right\}$ by changing the measure	
	Thus, choosing $q(\theta)$ to minimise $E\left(\frac{g^2(\theta)f(\theta\|X)}{q(\theta)}\,\middle\|\,X\right)$ will minimise $Var(\hat{I}\|\theta \sim q(\theta))$ and any $q(\theta)$ doing so is optimal.	4

Total	20

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 4	Page 1 of 4

Part | Mark

(a) The risk of decision d is

$$\rho(\pi, d) = E\{g(\theta)(\theta-d)^2\} \quad (\text{wrt } \pi(\theta))$$
$$= E(g(\theta)\theta^2) - 2dE(g(\theta)\theta) + d^2E(g(\theta))$$

We choose d to minimise this:

$$\frac{\partial}{\partial d}\rho(\pi, d) = -2E(g(\theta)\theta) + 2dE(g(\theta))$$

$$\text{so} \quad d^* = \frac{E(g(\theta)\theta)}{E(g(\theta))}$$

which is a minimum for $g(\theta) > 0$. The corresponding Bayes risk is

$$\rho^*(\pi) = \rho(\pi, d^*) = E(g(\theta)\theta^2) - 2\frac{E^2(g(\theta)\theta)}{E(g(\theta))} + \frac{E^2(g(\theta)\theta)}{E(g(\theta))}$$
$$= E(g(\theta)\theta^2) - \frac{E^2(g(\theta)\theta)}{E(g(\theta))}$$

When $g(\theta) = 1$, $d^* = E(\theta)$ (the mean) and $\rho^*(\pi) = Var(\theta)$ (the variance). 4

In the change to the open book exam, this question was re-phrased to ask for the solution rather than to show that the solution was of a given form. As such, the model solution is unchanged.

Total

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS	
Academic Year 2019/20	Examiner SIMON SHAW	
Semester II	Question No. 4	Page 2 of 4

Part | Mark

(b) We have $g(\theta) = \theta^{-3}$. For the immediate decision, the Bayes rule is $d^* = \dfrac{E(\theta^{-2})}{E(\theta^{-3})}$

Now as $\theta \sim Gamma(\alpha, \beta)$,

$$E(\theta^{-k}) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{(\alpha-k)-1} e^{-\beta\theta} \, d\theta$$

$$= \frac{\beta^k \Gamma(\alpha-k)}{\Gamma(\alpha)} \int_0^\infty \frac{\beta^{\alpha-k}}{\Gamma(\alpha-k)} \theta^{(\alpha-k)-1} e^{-\beta\theta} \, d\theta$$

$$= \frac{\beta^k \Gamma(\alpha-k)}{\Gamma(\alpha)} \text{ provided that } \alpha > k.$$

Hence, $E(\theta^{-3}) = \dfrac{\beta^3 \Gamma(\alpha-3)}{\Gamma(\alpha)}$, $E(\theta^{-2}) = \dfrac{\beta^2 \Gamma(\alpha-2)}{\Gamma(\alpha)}$

so that $\lambda^* = \dfrac{\beta^2 \Gamma(\alpha-2)}{\Gamma(\alpha)} \times \dfrac{\Gamma(\alpha)}{\beta^3 \Gamma(\alpha-3)} = \dfrac{\alpha-3}{\beta}$.

The Bayes risk is $\rho^*(f(\theta)) = E(\theta^{-1}) - \dfrac{E^2(\theta^{-2})}{E(\theta^{-3})}$

$$= \frac{\beta\Gamma(\alpha-1)}{\Gamma(\alpha)} - \frac{\beta^2 \Gamma(\alpha-2)}{\Gamma(\alpha)} \left(\frac{\alpha-3}{\beta}\right)$$

$$= \frac{\beta}{\alpha-1} - \frac{\beta(\alpha-3)}{(\alpha-1)(\alpha-2)} = \frac{\beta}{(\alpha-1)(\alpha-2)}$$

5

Total

UNIVERSITY OF BATH

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS.		
Academic Year 2019/20	Examiner SIMON SHAW		
Semester II	Question No. 4	Page 3	of 4

Part | Mark

(c) $f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \propto \theta^n \exp\left(-\theta \sum_{i=1}^{n} x_i^{\lambda}\right)$

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

Thus, $f(\theta|x) \propto \theta^{(\alpha+n)-1} \exp\left(-\theta\left(\beta + \sum_{i=1}^{n} x_i^{\lambda}\right)\right)$ so that $\theta | x \sim \text{Gamma}\left(\alpha+n, \beta + \sum_{i=1}^{n} x_i^{\lambda}\right)$.

We may exploit conjugacy to find $d^*(x) = \dfrac{\alpha+n-3}{\beta + \sum_{i=1}^{n} x_i^{\lambda}}$ with risk $\dfrac{\beta + \sum_{i=1}^{n} x_i^{\lambda}}{(\alpha+n-1)(\alpha+n-2)}$. 3

(d) With $\lambda = 1$, the risk of the sampling procedure is $E\left(\dfrac{\beta + \sum_{i=1}^{n} X_i}{(\alpha+n-1)(\alpha+n-2)}\right)$

Now, $E(X_i) = E(E(X_i|\theta)) = E(\theta^{-1}) = \dfrac{\beta}{\alpha-1}$ so that

$$E\left(\frac{\beta + \sum_{i=1}^{n} X_i}{(\alpha+n-1)(\alpha+n-2)}\right) = \frac{\beta + \frac{n\beta}{\alpha-1}}{(\alpha+n-1)(\alpha+n-2)} = \frac{\beta}{(\alpha-1)(\alpha+n-2)}$$

The total risk is $R_n = nc + \dfrac{\beta}{(\alpha-1)(\alpha+n-2)}$

Hence, $\dfrac{dR_n}{dn} = c - \dfrac{\beta}{(\alpha-1)(\alpha+n-2)^2}$

Total

ADAPTED FOR OPEN BOOK ASSESSMENT

Unit Code MA40189	Unit Title TOPICS IN BAYESIAN STATISTICS		
Academic Year 2019/20	Examiner SIMON SHAW		
Semester II	Question No. 4	Page 4	of 4

Part | Mark

Solving to 0 gives $(\alpha+n-2)^2 = \dfrac{\beta}{c(\alpha-1)}$

$$\Rightarrow n = \sqrt{\frac{\beta}{c(\alpha-1)}} - (\alpha-2).$$

(Sample if $n>0$ otherwise don't sample). 4.

(c) $E(L_1(\theta,d)) = (\theta_2-\theta_1) + \displaystyle\int_{-\infty}^{\theta_1} \frac{2}{\alpha}(\theta_1-\theta)\pi(\theta)\,d\theta + \int_{\theta_2}^{\infty}\frac{2}{\alpha}(\theta-\theta_1)\pi(\theta)\,d\theta$

$$= (\theta_2-\theta_1) + \frac{2}{\alpha}\theta_1 P(\theta\le\theta_1) - \frac{2}{\alpha}\int_{-\infty}^{\theta_1}\theta\pi(\theta)\,d\theta$$

$$+ \frac{2}{\alpha}\left[E(\theta) - \int_{-\infty}^{\theta_2}\theta\pi(\theta)\,d\theta\right] - \frac{2}{\alpha}\theta_2 P(\theta\ge\theta_2)$$

$$\frac{\partial}{\partial\theta_1}E(L_1(\theta,d)) = -1 + \frac{2}{\alpha}P(\theta\le\theta_1) + \frac{2}{\alpha}\theta_1\pi(\theta_1) - \frac{2}{\alpha}\theta_1\pi(\theta_1)$$

Hence, $\dfrac{\partial}{\partial\theta_1}E(L_1(\theta,d)) = 0$ when $P(\theta\le\theta_1) = \dfrac{\alpha}{2}$.

$$\frac{\partial}{\partial\theta_2}E(L_1(\theta,d)) = 1 - \frac{2}{\alpha}\theta_2\pi(\theta_2) - \frac{2}{\alpha}P(\theta\ge\theta_2) + \frac{2}{\alpha}\theta_2\pi(\theta_2)$$

Hence, $\dfrac{\partial}{\partial\theta_2}E(L_1(\theta,d)) = 0$ when $1 - \dfrac{2}{\alpha}P(\theta\ge\theta_2) = 0$

i.e. $P(\theta\ge\theta_2) = \dfrac{\alpha}{2}$

Hence, the Bayes rule is $d^* = (\theta_1^*, \theta_2^*)$ when $P(\theta\le\theta_1^*) = \dfrac{\alpha}{2} = P(\theta\ge\theta_2^*)$ 4

Total 20