

MAS2903: Introduction to Bayesian Statistics

Solution to 2017 June exam

Question A1

a. The answers are:

- i Exponential
- ii Gamma
- iii (0.00, 3.00)
- iv Lower variance
- v 1.16
- vi 0

(1 mark each for part)

Unseen but should be easy.

b. Which interpretation to use here is a matter of opinion. Marks are awarded for any answer with a sensible justification. Example responses are as follows.

- i Frequency interpretation. We can repeatedly observe whether this event is true many times under similar conditions.
- ii Classical interpretation. Each time we pick a card, any one in the pack is equally likely by symmetry. The probability can be derived from this.

(2 mark each for part: 1 for an interpretation, 1 for a good reason)

Similar to coursework questions.

Question A2

a. Let S_1, S_2, S_3 denote the events “real”, “forged by Fred”, “forged by Carol”. Let R denote the event that the passport is believed to be real. The question supplies us with the information:

$$\begin{aligned}\Pr(S_1) &= \Pr(S_2) = \Pr(S_3) = 1/3 \\ \Pr(R|S_1) &= 0.99 \quad \Pr(R|S_2) = 0.8 \quad \Pr(R|S_3) = 0.4\end{aligned}$$

Using the law of total probability:

$$\begin{aligned}\Pr(R) &= \sum_{i=1}^3 \Pr(R|S_i) \Pr(S_i) \\ &= (0.99 + 0.8 + 0.4) \frac{1}{3} \\ &= 0.73\end{aligned}$$

Bayes theorem is $\Pr(S_i|R) = \Pr(R|S_i) \Pr(S_i) / \Pr(R)$. Hence:

$$\begin{aligned}\Pr(S_1|R) &= 0.99 \times \frac{1}{3} / 0.73 = 0.45 \\ \Pr(S_2|R) &= 0.8 \times \frac{1}{3} / 0.73 = 0.37 \\ \Pr(S_3|R) &= 0.4 \times \frac{1}{3} / 0.73 = 0.18\end{aligned}$$

(1 for mathematical statement + 1 for using Bayes theorem + 1 mark for each correct answer)

Similar to examples in course.

- b. Let RF denote the event that the first expert believes the passport is real and the second that it is false. Using independence:

$$\Pr(RF|S_i) = \Pr(R|S_i)(1 - \Pr(R|S_i))$$

Thus:

$$\Pr(RF|S_1) = 0.0099 \quad \Pr(RF|S_2) = 0.16 \quad \Pr(RF|S_3) = 0.24$$

Now using the law of total probability $\Pr(RF) = 0.137$, and Bayes theorem gives:

$$\Pr(S_1|RF) = 0.0099 \times \frac{1}{3} / 0.137 = 0.02$$

$$\Pr(S_2|RF) = 0.16 \times \frac{1}{3} / 0.137 = 0.39$$

$$\Pr(S_3|RF) = 0.24 \times \frac{1}{3} / 0.137 = 0.59$$

(2 for deriving new conditional probabilities of RF + 1 mark for each correct answer)

Unseen.

Question A3

- a. The likelihood is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{x_i!} \\ &\propto e^{-n\theta} \theta^{n\bar{x}} \\ &= e^{-7\theta} \theta^{12} \end{aligned}$$

(2 marks)

Standard bookwork.

- b. To find the MLE first find the log likelihood:

$$\log f(\mathbf{x}|\theta) = -7\theta + 12 \log \theta + \text{const}$$

Now differentiate:

$$\frac{d}{d\theta} \log f(\mathbf{x}|\theta) = -7 + 12/\theta$$

Setting this equal to zero and solving for θ gives $\theta = 12/7$.

Check this is the global maximum e.g. by noting $f(\mathbf{x}|\theta = 12/7) > 0$ but $f(\mathbf{x}|\theta = 0) = 0$ and $\lim_{\theta \rightarrow \infty} f(\mathbf{x}|\theta) = 0$. (Showing that the 2nd derivative is less than zero for all θ is also another approach here.)

(1 for log likelihood, 1 for finding root, 1 for checking it's global maximum - alternatively the marks can be awarded if students use other methods)

Standard bookwork.

- c. The posterior is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto \theta e^{-b\theta} e^{-7\theta} \theta^{12} \\ &= \theta^{13} e^{-(7+b)\theta}\end{aligned}$$

This is a $Ga(14, b+7)$ distribution.

(1 for Bayes theorem, 1 for some correct working, 1 for correct answer)
Standard bookwork.

- d. From the formula sheet, the posterior mode is $\frac{13}{b+7}$. Solving $13/(b+7) > 12/7$ gives $b < 7/12$. We also need $b > 0$ from the definition of the Gamma distribution.

(1 for using posterior mode formula, 1 for some correct answer)
Unseen but should be easy.

Question A4

- ai. This distribution has the correct mean. Also if $\mu \sim N(a, b^2)$ then $\Pr(a - 2b < \mu < a + 2b) \approx 0.95$. So $\Pr(10 < \mu < 50) \approx 0.95$.

(1 for comment on mean, 1 for comment on other values)
Unseen question on basic material.

- aii. This distribution has the wrong mode.

(1 mark for a sensible comment [e.g. correct mean, variance seems wrong with explanation why], 1 if this shows prior is wrong)
Unseen question on basic material.

- b. "Is it reasonable to expect to see $\mu > 60$ only around 1 time in a thousand?"

(1 mark)
Standard bookwork.

- c. Using the update formulae from the formula sheet, the posterior is $N(B, 1/D)$ where

$$\begin{aligned}D &= 1/100 + 1/10 = 11/100 \\ B &= (1/100 \times 30 + 1/10 \times 8)/D = 10\end{aligned}$$

(1 for using formula, 1 for correct answer)
Standard bookwork.

- d. The posterior is $\mu|x \sim N(10, 100/11)$ truncated to $\mu > 5$. This has density:

$$\pi(\mu|x) \propto \begin{cases} \exp\left[-\frac{11}{200}(\mu - 10)^2\right] & \text{for } \mu > 5 \\ 0 & \text{otherwise} \end{cases}$$

(2 for using truncation of prev posterior, 1 for final density)
Standard coursework question.

Question B5

- a. A prior is conjugate for a specific model if the posterior is in the same family of distributions. This is the case here, as a Gamma prior leads to a Gamma posterior.

(1 for a correct definition, 1 using it to get the correct answer.)

Standard bookwork.

- b. i The credible interval must satisfy:

$$\begin{aligned} \int_a^b \pi(\theta|\mathbf{x})d\theta &= 0.95 \\ \Leftrightarrow \int_a^b \frac{h'^{g'} \theta^{g'-1} \exp(-h'\theta)}{\Gamma(g')} d\theta &= 0.95 \\ \Leftrightarrow \int_a^b \theta^5 e^{-3\theta} d\theta &= 0.95 \frac{\Gamma(6)}{3^6} = \frac{38}{243} \end{aligned}$$

(1 for general integral, 2 for rearranging to correct answer.)

Unseen question applying standard material.

- b. ii A HDI interval must also satisfy (in this case) $\pi(a|\mathbf{x}) = \pi(b|\mathbf{x})$.

(2 marks. Other correct conditions are also acceptable.)

Standard bookwork.

- c. Sequential Bayesian inference can be used here. Use the posterior given $\mathbf{x}$, $Ga(g', h')$ as the prior for the $\mathbf{y}$ data. Using the update formula this produces a new posterior $Ga(g'', h'')$ where $g'' = g' + m = g + n + m$ and $h'' = h' + m\bar{y} = h + n\bar{x} + m\bar{y}$.

(1 mark for applying sequential Bayes, 1 for using update formula, 1 for correct answer. Other correct approaches also acceptable.)

Standard bookwork.

- d. The likelihood is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\ &= \theta^n \exp(-\theta n\bar{x}) \end{aligned}$$

Hence the log-likelihood and its derivatives are:

$$\begin{aligned} \log f(\mathbf{x}|\theta) &= n \log \theta - \theta n\bar{x} \\ \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) &= n/\theta - n\bar{x} \\ \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) &= -n/\theta^2 \end{aligned}$$

Solving $\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0$ gives maximum likelihood estimate $\hat{\theta} = 1/\bar{x}$ (n.b. this is the global maximum as $\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) < 0$ for all θ .)

The observed information is

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = n/\theta^2$$

The general form of the asymptotic posterior is $\theta \sim N(\hat{\theta}, J(\hat{\theta})^{-1})$. So in this particular case it is

$$N\left(\frac{1}{\bar{x}}, \frac{1}{n\bar{x}^2}\right).$$

(1 marks for likelihood, 2 for log likelihood + derivatives, 2 for MLE, 2 for observed information, 2 for general form of asymptotic posterior, 1 for final answer.)

An example from the notes, but remembering all the details may be challenging.

Question B6

a. The likelihood is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \prod_{i=1}^n f(x_i|\theta) \\ &= \prod_{i=1}^n \binom{k}{x_i} \theta^{x_i} \theta^{k-x_i} \\ &= \left[\prod_{i=1}^n \binom{k}{x_i} \right] \theta^{\sum_{i=1}^n x_i} \theta^{\sum_{i=1}^n k-x_i} \\ &= \left[\prod_{i=1}^n \binom{k}{x_i} \right] \theta^{n\bar{x}} \theta^{n(k-\bar{x})} \end{aligned}$$

as required.

(1 for correct approach, 1 for some correct answer)

b. The sufficient statistic is $\bar{X}$ by the factorisation theorem, since we can factorise the likelihood function as $f(\mathbf{x}|\theta) = h(\mathbf{x})g(\bar{x}, \theta)$ where

$$\begin{aligned} h(\mathbf{x}) &= \prod_{i=1}^n \binom{k}{x_i} \\ g(\bar{x}, \theta) &= \theta^{n\bar{x}} \theta^{n(k-\bar{x})} \end{aligned}$$

(1 for using factorisation theorem, 1 for correct answer)

c. This is true because

$$\begin{aligned} E(\bar{X}|\theta) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i|\theta\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i|\theta) && \text{by linearity of expectation} \\ &= \frac{1}{n} nk\theta && \text{using expectation of binomial distribution} \\ &= k\theta && \text{as required} \end{aligned}$$

(1 for correct approach, 1 for details)

d. First we need the log-likelihood:

$$\log f(\mathbf{x}|\theta) = \log c(\mathbf{x}) + n\bar{x} \log \theta + n(k - \bar{x}) \log(1 - \theta)$$

Next we want its second derivative

$$\begin{aligned} \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) &= \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta} \\ \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) &= -\frac{n\bar{x}}{\theta^2} - \frac{n(k - \bar{x})}{(1 - \theta)^2} \end{aligned}$$

Now we need to find Fisher's information $I(\theta)$, which is:

$$\begin{aligned} E \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \middle| \theta \right] &= \frac{nE[\bar{X}|\theta]}{\theta^2} + \frac{n(k - E[\bar{X}|\theta])}{(1 - \theta)^2} \\ &= \frac{nk\theta}{\theta^2} + \frac{nk(1 - \theta)}{(1 - \theta)^2} && \text{using part c} \\ &= nk \left[\frac{1}{\theta} + \frac{1}{1 - \theta} \right] \\ &= \frac{nk}{\theta(1 - \theta)} \end{aligned}$$

Finally, the Jeffreys prior is

$$\begin{aligned} \pi(\theta) &\propto \sqrt{I(\theta)} \\ &= \propto \theta^{-1/2}(1 - \theta)^{-1/2} \end{aligned}$$

as required.

(2 for log-likelihood, 2 for derivatives, 4 for Fisher's information, 2 for Jeffreys prior)

e. No, because it's a $Be(1/2, 1/2)$ distribution. **(1 for a correct answer + a valid reason)**

f. The posterior density is

$$\begin{aligned} \pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto \theta^{-1/2}(1 - \theta)^{-1/2}\theta^{2.5}(1 - \theta)^{12.5} \\ &= \theta^2(1 - \theta)^{12}. \end{aligned}$$

This is a $Be(3, 13)$ distribution.

(1 for Bayes theorem, 1 for correct working, 1 for stating final distribution)

B6 is basically a standard example from the notes, but using an unseen distribution (binomial). This will make part d challenging.

Question B7

a. i The likelihood is:

$$f(X = 6|\theta) = \theta^6 e^{-\theta} / 6! = \begin{cases} 0.122 & \text{if } \theta = 8 \\ 0.050 & \text{if } \theta = 3 \end{cases}$$

Then the probability of busy activity is:

$$\begin{aligned}\Pr(\theta = 8|X = 6) &= \frac{\Pr(X = 6|\theta = 8) \Pr(\theta = 8)}{\Pr(X = 6|\theta = 8) \Pr(\theta = 8) + \Pr(X = 6|\theta = 3) \Pr(\theta = 3)} \\ &= \frac{0.122 \times 0.2}{0.122 \times 0.2 + 0.050 \times 0.8} \\ &= 0.377\end{aligned}$$

Thus the probability of normal activity is $\Pr(\theta = 3|X = 6) = 1 - 0.377 = 0.623$.

(1 for likelihood, 2 for Bayes theorem, 1 for correct answer)

- a. ii Since tropical storms occur at the same rate in August as they do in July, we have

$$\begin{aligned}f(y|x = 6) &= \sum_{\theta=3,8} f(y|\theta)\pi(\theta|x = 6) \\ &= 0.623(3^y e^{-3}/y!) + 0.377(8^y e^{-8}/y!)\end{aligned}$$

(2 for correct predictive formula, 1 for correct answer)

- a. iii Using the values giving rise to highest probabilities, we see that $[0, 8]$ gives close to a 95% coverage.

(2 for correct method, 1 for correct answer. Other intervals with the correct probability are also acceptable.)

- b. The posterior is

$$\pi(\theta|x = 1) \propto \theta(1 - \theta)^{22}\theta(1 - \theta)^4 = \theta^2(1 - \theta)^{26},$$

and so $\theta|x = 1 \sim Be(3, 27)$.

The predictive probability function for Y is, for $Y = 0, 1, \dots$,

$$\begin{aligned}f(y|x = 1) &= \int_{\Theta} f(y|\theta)\pi(\theta|x = 1)d\theta \\ &= \int_0^1 \binom{5}{y} \theta^y (1 - \theta)^{5-y} \frac{\theta^2(1 - \theta)^{26}}{B(3, 27)} d\theta \\ &= \frac{\binom{5}{y}}{B(3, 27)} \int_0^1 \theta^{y+2} (1 - \theta)^{31-y} d\theta \\ &= \binom{5}{y} \frac{B(y+3, 32-y)}{B(3, 27)}\end{aligned}$$

that is, $Y|x = 1 \sim BetaBin(5, 3, 27)$, as required.

(2 for recognising Binomial model, 3 for posterior, 2 for stating predictive dist as an integral, 3 for rearranging this correctly to get correct answer)

All parts of B7 are very similar to examples in the course notes, but will still be challenging