

MAS2317/3317: Problems Class 1

Dr. Lee Fawcett

Semester 2, 2014—2015

Problems Class 1

The work in this problems class will correspond to the (unstarred) questions in Chapter 5 linking in with the course material so far.

This should supplement the assessed (starred) questions for assignment 1 that are due in by no later than **4pm, Tuesday 24th February**.

There is no practical element to this work; there is a practical class next week that will correspond to work in assignment 2.

If you need extra help with any of the material in assignment 1, there is a drop-in session next week.

Question 5

An experiment can result in four possible outcomes x_1 , x_2 , x_3 and x_4 .

The probability of each of these outcomes is affected by a parameter θ which can take one of three values: θ_1 , θ_2 and θ_3 :

$p(x \theta)$	x_1	x_2	x_3	x_4
θ_1	0.1	0.2	0.3	0.4
θ_2	0.3	0.3	0.2	0.2
θ_3	0.5	0.1	0.1	0.3

Question 5

The prior plausibilities attaching to θ_1 , θ_2 and θ_3 are 0.1, 0.4 and 0.5.

Find the posterior probability function for θ after observing x_1 . Repeat the calculations for x_2 , x_3 and x_4 .

Solution to Question 5

The **posterior probabilities** given x_1 are

$$\begin{aligned}Pr(\theta_1|x_1) &= \frac{Pr(x_1|\theta_1)Pr(\theta_1)}{Pr(x_1|\theta_1)Pr(\theta_1) + Pr(x_1|\theta_2)Pr(\theta_2) + Pr(x_1|\theta_3)Pr(\theta_3)} \\&= \frac{0.1 \times 0.1}{0.1 \times 0.1 + 0.3 \times 0.4 + 0.5 \times 0.5} \\&= 0.02632,\end{aligned}$$

$$\begin{aligned}Pr(\theta_2|x_1) &= \frac{Pr(x_1|\theta_2)Pr(\theta_2)}{Pr(x_1|\theta_1)Pr(\theta_1) + Pr(x_1|\theta_2)Pr(\theta_2) + Pr(x_1|\theta_3)Pr(\theta_3)} \\&= \frac{0.3 \times 0.4}{0.1 \times 0.1 + 0.3 \times 0.4 + 0.5 \times 0.5} \\&= 0.31579,\end{aligned}$$

Solution to Question 5

and

$$\begin{aligned}Pr(\theta_3|x_1) &= 1 - Pr(\theta_1|x_1) - Pr(\theta_2|x_1) \\&= 1 - 0.02632 - 0.31579 \\&= 0.65789.\end{aligned}$$

Thus, after observing x_1 , we have the following probability function for θ :

θ	θ_1	θ_2	θ_3
$Pr(\theta = \theta_i x_1)$	0.02632	0.31579	0.65789

Same for x_2 , x_3 and x_4 .

Question 6

Extensive manuscripts of two scribes show that they differ on the relative frequencies with which they use two alternative vowel forms e and œ , as in the following table:

	e	œ
Scribe A	0.3	0.7
Scribe B	0.5	0.5

Question 6

A new manuscript fragment, known to be by one of the scribes, uses the vowel form five times, three times as e and twice as œ .

If an independent assessment of similar fragments favours Scribe A as author with probability 0.7, show that the evidence from the vowel form is sufficient to turn the odds in favour of Scribe B.

Solution to Question 6

Let X = number of e's observed.

Also, let θ_A denote Scribe A and θ_B denote Scribe B.

Then

$$Pr(\theta_A) = 0.7, \quad Pr(\theta_B) = 0.3,$$

$$X|\theta_A \sim \text{Bin}(5, 0.3) \quad \text{and} \quad X|\theta_B \sim \text{Bin}(5, 0.5).$$

Solution to Question 6

Hence

$$Pr(X = 3|\theta_A) = \binom{5}{3} 0.3^3 \times 0.7^2 = 0.1323$$

$$Pr(X = 3|\theta_B) = \binom{5}{3} 0.5^3 \times 0.5^2 = 0.3125.$$

Solution to Question 6

Therefore, using Bayes' Theorem

$$\begin{aligned}Pr(\theta_A|x = 3) &= \frac{Pr(X = 3|\theta_A)Pr(\theta_A)}{Pr(X = 3|\theta_A)Pr(\theta_A) + Pr(X = 3|\theta_B)Pr(\theta_B)} \\&= \frac{0.1323 \times 0.7}{0.1323 \times 0.7 + 0.3125 \times 0.3} \\&= 0.497 \text{ (3 d.p.)}\end{aligned}$$

$$Pr(\theta_B|x = 3) = 1 - Pr(\theta_A|x = 3) = 0.503 \text{ (3 d.p.)}.$$

Thus, the odds have turned from 7 : 3 in favour of Scribe A to 503 : 497 in favour of Scribe B.

Question 7

If $X_1, X_2, \dots, X_n$ are independent random variables with probability function

$$f(x|\theta) = \theta(1 - \theta)^{x-1}, \quad x = 1, 2, \dots,$$

show that $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Question 7

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1-\theta)^{x_i-1} \\&= \theta^n(1-\theta)^{x_1-1}(1-\theta)^{x_2-1} \times \dots \times (1-\theta)^{x_n-1} \\&= \theta^n(1-\theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1-\theta)^{\sum x_i - n} \\&= h(\mathbf{x})g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1-\theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .

Question 8

If $X_1, X_2, \dots, X_n$ are independent $N(5, \sigma^2)$ random variables, show that

$$T = \sum_{i=1}^n (X_i - 5)^2$$

is sufficient for σ .

Solution to Question 8

The joint probability density function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_i - 5)^2}{2\sigma^2}\right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - 5)^2\right\} \\&= 1 \times (2\pi)^{-n/2} \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - 5)^2\right\} \\&= h(\mathbf{x}) g\left(\sum (x_i - 5)^2, \sigma\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \sigma) = (2\pi)^{-n/2} \sigma^{-n} e^{-t/(2\sigma^2)}$.

Therefore, by the Factorisation Theorem, $T = \sum (X_i - 5)^2$ is sufficient for σ .

MAS2317/3317: Problems Class 2

Dr. Lee Fawcett

Semester 2, 2014—2015

Problems Class 2

The work from assignment 1 is being marked as we speak, and will be returned soon (check emails).

The work in this problems class will correspond to the (unstarred) questions in Chapter 5 of the lecture notes associated with the most recent work in class.

This should supplement the assessed (starred) questions for assignment 2 that are due in by no later than **4pm, Tuesday 10th March**.

If you need extra help with any of the material in assignment 2, there is a drop-in session next week.

Question 10*

Similar to Example 2.7 in the lecture notes.

Question 11

Suppose that a random sample $x_1, x_2, \dots, x_n$ is obtained. Derive the posterior distributions for the following models:

- (a) $f(x|\theta) = \theta^{x-1}(1 - \theta)$, $x = 1, 2, \dots$ with a $U(0, 1)$ prior distribution for θ .
- (b) $f(x|\theta) = \frac{e^{-\theta}\theta^x}{x!}$, $x = 0, 1, \dots$ with a $Exp(1)$ prior distribution for θ .
- (c) $f(x|\theta) = \binom{7}{x} \theta^x (1 - \theta)^{7-x}$, $x = 0, 1, \dots, 7$ with a $Beta(3, 2)$ prior distribution for θ .

Solution to Question 11 (a)

The likelihood function is

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

The prior distribution is $\theta \sim U(0, 1)$, with density

$$\pi(\theta) = 1, \quad 0 < \theta < 1.$$

Therefore, using Bayes Theorem, the posterior density is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)f(\mathbf{x}|\theta) \\ &\propto 1 \times \theta^{n(\bar{x}-1)}(1-\theta)^n, \quad 0 < \theta < 1 \\ &\propto \theta^{n(\bar{x}-1)+1-1}(1-\theta)^{n+1-1}, \quad 0 < \theta < 1,\end{aligned}$$

that is, $\theta|\mathbf{x} \sim \text{Be}(n(\bar{x} - 1) + 1, n + 1)$.

Solution to Question 11(b)

The likelihood function is

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

where k is constant w.r.t. θ .

The prior distribution is $\theta \sim \text{Exp}(1)$, with density

$$\pi(\theta) = e^{-\theta}, \quad \theta > 0.$$

Therefore, using Bayes Theorem, the posterior density is

$$\begin{aligned} \pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto e^{-\theta} \times k e^{-n\theta} \theta^{n\bar{x}}, \quad \theta > 0 \\ &\propto \theta^{n\bar{x}+1-1} e^{-(n+1)\theta}, \quad \theta > 0, \end{aligned}$$

that is, $\theta|\mathbf{x} \sim \text{Ga}(n\bar{x} + 1, n + 1)$.

Solution to Question 11(c)

The likelihood function is

$$f(\mathbf{x}|\theta) = \prod_{i=1}^n \binom{7}{x_i} \theta^{x_i} (1 - \theta)^{7-x_i} = k \theta^{n\bar{x}} (1 - \theta)^{n(7-\bar{x})}$$

where k is constant w.r.t. θ .

The prior distribution is $\theta \sim Be(3, 2)$, with density

$$\pi(\theta) = k' \theta^2 (1 - \theta), \quad 0 < \theta < 1$$

where k' is constant w.r.t. θ . Therefore, using Bayes Theorem, the posterior density is

$$\begin{aligned} \pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto k' \theta^2 (1 - \theta) \times k \theta^{n\bar{x}} (1 - \theta)^{n(7-\bar{x})}, \quad 0 < \theta < 1 \\ &\propto \theta^{n\bar{x}+3-1} (1 - \theta)^{n(7-\bar{x})+2-1}, \quad 0 < \theta < 1, \end{aligned}$$

that is, $\theta|\mathbf{x} \sim Be(n\bar{x} + 3, n(7 - \bar{x}) + 2)$.

Question 14*

- Should find it helpful if you have the case study material handy
- Really, no different to any other “Bayes’ Theorem for Distributions” questions we have worked through in class (Examples 2.1–2.7)
- Example 2.6, where I mentioned the **Bayes linear rule**, might also help with part (b)(ii)

Question 17

The data in the table below are measurements of the logarithm of *ornithine carbonyltransferase* (a liver enzyme) in patients suffering from acute viral hepatitis.

2.66	2.38	2.37	2.50	1.96	2.85	2.68	1.76	2.36
2.56	2.09	2.85	2.37	2.40	2.79	1.82	3.00	2.50
2.36	2.48	2.60	2.42	2.51	2.51	2.80	2.50	2.57

Question 17

Assume that the enzyme measurement varies according to a $N(\mu, 0.27^2)$ distribution, and that an expert says his beliefs about the mean level μ are described by a $N(2.7, 0.3^2)$ distribution.

After combining both sources of information, what is the probability that this mean level is less than 2.5?

Sketch the prior and posterior distributions for μ on the same graph. Describe the main changes in beliefs after incorporating the data.

Solution to Question 17 (1/3)

From the data we calculate $\bar{x} = 2.4685$ and $s = 0.297$.

Therefore, the assumed standard deviation $\sigma = 0.27$ is probably okay. Thus, using **Example 2.5** (with $\tau = 1/0.27^2$, $b = 2.7$, $d = 1/0.3^2$, $n = 27$, $\bar{x} = 2.4685$), the posterior distribution is $\mu|\mathbf{x} \sim N(B, 1/D)$, where

$$B = \frac{db + n\tau\bar{x}}{d + n\tau} = \frac{2.7/0.3^2 + 27 \times 2.4685/0.27^2}{1/0.3^2 + 27/0.27^2} = 2.4752$$

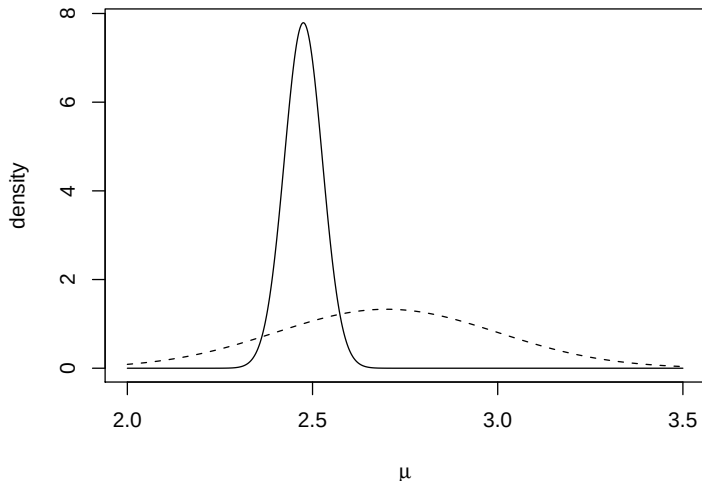
and

$$D = d + n\tau = \frac{1}{0.3^2} + \frac{27}{0.27^2} = \frac{1}{0.0512}.$$

Therefore $Pr(\mu < 2.5|\mathbf{x}) = 0.6859$, using $\text{pnorm}(2.5, 2.4752, 0.0512)$ (or tables!).

Solution to Question 17 (2/3)

Plots of the prior and posterior densities for the mean level μ are given below.



Solution to Question 17 (3/3)

The main change in beliefs about μ is a dramatic reduction in our uncertainty – the standard deviation has decreased from 0.3 to 0.05.

There has also been a reduction in mean value from 2.7 to 2.475.

In summary, the data have led to us being more certain about a lower value of mean enzyme level.

Question 16

The number of defects in a 1200 foot roll of magnetic recording tape has a $Po(\theta)$ distribution.

The prior distribution for θ is $Ga(3, 1)$.

When 5 rolls of this tape are selected at random and inspected, the number of defects found on the rolls are 2, 2, 6, 0 and 3.

Determine the posterior distribution of θ .

Solution to Question 16 (1/2)

Let X_i = number of defects on roll i , $i = 1, 2, 3, 4, 5$.

Then $X_i|\theta \sim Po(\theta)$, independently. Therefore, the likelihood function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= Pr(X_1 = 2, X_2 = 2, X_3 = 6, X_4 = 0, X_5 = 3|\theta) \\&= Pr(X_1 = 2|\theta) \times Pr(X_2 = 2|\theta) \times Pr(X_3 = 6|\theta) \\&\quad \times Pr(X_4 = 0|\theta) \times Pr(X_5 = 3|\theta) \\&= \frac{\theta^2 e^{-\theta}}{2!} \times \frac{\theta^2 e^{-\theta}}{2!} \times \frac{\theta^6 e^{-\theta}}{6!} \times \frac{\theta^0 e^{-\theta}}{0!} \times \frac{\theta^3 e^{-\theta}}{3!} \\&\propto \theta^{13} e^{-5\theta}.\end{aligned}$$

Solution to Question 16(2/2)

The prior distribution has density

$$\pi(\theta) \propto \theta^2 e^{-\theta}, \quad \theta > 0.$$

Therefore, using Bayes Theorem, the posterior density is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \theta^2 e^{-\theta} \times \theta^{13} e^{-5\theta}, \quad \theta > 0 \\ &\propto \theta^{15} e^{-6\theta}, \quad \theta > 0,\end{aligned}$$

that is, $\theta|\mathbf{x} \sim \text{Ga}(16, 6)$.

Question 18

A random sample of size n is to be taken from a $N(\theta, 2^2)$ distribution. The prior for θ is $N(b, 1/d)$.

- (a) If $d = 1$, what is the smallest number of observations that must be included in the sample if the standard deviation of the posterior distribution of θ is to be reduced to 0.1?
- (b) If $n = 100$, show that no matter how large the value of d , the standard deviation of the posterior distribution is less than $1/5$.

Solution to Question 18

From **Example 2.5** (with $\tau = 1/\sigma^2 = 1/4$), the posterior distribution is $\theta|\mathbf{x} \sim N(B, 1/D)$, where

$$B = \frac{db + n\tau\bar{x}}{d + n\tau} = \frac{db + n\bar{x}/4}{d + n/4}$$

and

$$D = d + n\tau = d + \frac{n}{4}.$$

The posterior variance is $1/D$, and so the posterior standard deviation is $1/\sqrt{D}$, i.e.

$$SD(\theta|\mathbf{x}) = \frac{1}{\sqrt{d + n/4}}.$$

Solution to Question 18 (a)

If $d = 1$, then we require

$$\frac{1}{\sqrt{1 + n/4}} \leq 0.1 \quad \text{i.e.}$$

$$\frac{1}{1 + n/4} \leq 0.01 \quad \text{i.e.}$$

$$1 \leq 0.01 \left(1 + \frac{n}{4}\right) \quad \text{i.e.}$$

$$99 \leq \frac{n}{4} \quad \text{i.e.}$$

$$n \geq 396.$$

Solution to Question 18 (b)

If $n = 100$ then

$$SD(\theta|\mathbf{x}) = \frac{1}{\sqrt{d + 100/4}} = \frac{1}{\sqrt{d + 25}} \leq 0.2,$$

since $d \geq 0$.

Question 19

In viewing a section through the pancreas, doctors see what are called “islands”. The numbers of islands seen in 900 patients are summarised below.

Number of islands	0	1	2	3	4	5	6	≥ 7
Frequency	327	340	160	53	16	3	1	0

Question 19

Assuming that the numbers of islands are Poisson distributed with mean θ , obtain the posterior mean and standard deviation for θ from several gamma $Ga(g, h)$ prior distributions (with g and h in the range 1 to 50).

Why is the posterior distribution quite insensitive to variations in the prior distribution?

Solution to Question 19

Let X : number of islands; therefore, $X_i|\theta \sim Po(\theta)$. We are also told that $\theta \sim Ga(g, h)$.

The likelihood is

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

Solution to Question 19 (1/2)

The prior is

$$\begin{aligned}\pi(\theta) &= \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \\ &\propto \theta^{g-1} e^{-h\theta}.\end{aligned}$$

Combining the two via Bayes' Theorem gives:

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \theta^{g-1} e^{-h\theta} \times e^{-n\theta} \theta^{n\bar{x}} \\ &\propto \theta^{[g+n\bar{x}]-1} e^{-[h+n]\theta},\end{aligned}$$

and so $\theta|\mathbf{x} \sim \text{Ga}(g + n\bar{x}, h + n)$.

Solution to Question 19 (2/2)

Now we have $n = 900$ and so

$$\bar{x} = \frac{340 + 2 \times 160 + 3 \times 53 + \dots}{900} = 1.0044.$$

Thus, we have

$$\begin{aligned}\theta|\mathbf{x} &\sim \text{Ga}(g + 900 \times 1.0044, h + 900) \quad \text{i.e.} \\ &\sim \text{Ga}(g + 904, h + 900),\end{aligned}$$

giving

$$E[\theta|\mathbf{x}] = \frac{g + 904}{h + 900} \quad \text{and} \quad SD(\theta|\mathbf{x}) = \frac{\sqrt{g + 904}}{h + 900}.$$

Solution to Question 19

Values for the posterior mean and standard deviation are given in the following tables for a wide range of values for g and h .

		g						
$E(\theta \mathbf{x})$		1	5	10	20	30	40	50
h	1	1.004	1.009	1.014	1.026	1.037	1.048	1.059
	5	1.000	1.004	1.010	1.021	1.032	1.043	1.054
	10	0.995	0.999	1.004	1.015	1.026	1.037	1.048
	20	0.984	0.988	0.993	1.004	1.015	1.015	1.037
	30	0.973	0.977	0.983	0.994	1.043	1.004	1.026
	40	0.963	0.967	0.972	0.983	0.994	0.994	1.015
	50	0.953	0.957	0.962	0.973	0.983	0.983	1.004

Solution to Question 19

$SD(\theta \mathbf{x})$	g						
	1	5	10	20	30	40	50
1	0.0334	0.0335	0.0336	0.0337	0.0339	0.0341	0.0343
5	0.0332	0.0333	0.0334	0.0336	0.0338	0.0339	0.0341
10	0.0331	0.0331	0.0332	0.0334	0.0336	0.0338	0.0339
h 20	0.0327	0.0328	0.0329	0.0330	0.0332	0.0334	0.0336
30	0.0323	0.0324	0.0325	0.0327	0.0329	0.0330	0.0332
40	0.0320	0.0321	0.0322	0.0323	0.0325	0.0327	0.0329
50	0.0317	0.0317	0.0318	0.0320	0.0322	0.0323	0.0325

Solution to Question 19

The posterior mean and standard deviation (and also the posterior distribution) are insensitive to the choice of g and h because the data summaries ($n\bar{x}$ and n) included in their ratios are large in comparison to g and h .

MAS2317/3317: Problems Class 4

Dr. Lee Fawcett

Semester 2, 2014–2015

Problems Class 4

The work from assignment 3 is being marked as we speak, and will be returned soon (check emails).

The work in this problems class will correspond to the (unstarred) questions in Chapter 5 of the lecture notes relating to recent course material.

This should supplement the assessed (starred) questions for assignment 4 that are due in by no later than **4pm, Tuesday 5th May**.

If you need extra help with any of the material in assignment 4, there is a drop-in session after Easter.

There is also an **optional** practical next week.

Questions 34 and 35 correspond to material in the very last part of the notes, and so we will come back to these in the last problems class of the course.

Question 28

Suppose that the random variable Y follows a $Ga(a, b)$ distribution.

Determine the probability density function for the random variable $X = 1/\sqrt{Y}$.

Solution to Question 28

Suppose $Y \sim \text{Ga}(a, b)$ and $X = 1/\sqrt{Y}$. The density of X is

$$\begin{aligned}f_X(x) &= f_Y(x^{-2}) \left| \frac{d}{dx} x^{-2} \right| \\&= \frac{b^a (x^{-2})^{a-1} e^{-bx^{-2}}}{\Gamma(a)} \times 2x^{-3}, \quad x > 0 \\&= \frac{2b^a x^{-2a-1} e^{-b/x^2}}{\Gamma(a)}, \quad x > 0.\end{aligned}$$

Question 29

Using a random sample from a $\text{Bin}(k, \theta)$ (with k known), determine the posterior distribution for θ assuming

- (i) vague prior knowledge;
- (ii) the Jeffreys prior distribution;
- (iii) a very large sample.

Solution to question 29

The **conjugate prior distribution** is a $Beta(g, h)$ distribution. Using this prior distribution, the posterior density is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f(\mathbf{x}|\theta)$$

$$\propto \theta^{g-1}(1-\theta)^{h-1} \times \prod_{i=1}^n \theta^{x_i}(1-\theta)^{k-x_i}, \quad 0 < \theta < 1$$

$$\propto \theta^{g+n\bar{x}-1}(1-\theta)^{h+nk-n\bar{x}-1}, \quad 0 < \theta < 1$$

$$\text{i.e. } \theta|\mathbf{x} \sim Beta(G = g + n\bar{x}, H = h + nk - n\bar{x}).$$

Solution to question 29(i)

We represent **vague prior information** by taking a conjugate prior distribution with large variance.

As the beta distribution restricts values to the range (0,1), there is a finite upper limit to the variance.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g \rightarrow 0$ and $h \rightarrow 0$.

Thus we will take this limit to represent vague prior information. Hence the posterior distribution under vague prior information is

$$\theta | \mathbf{x} \sim \text{Beta}(n\bar{x}, nk - n\bar{x}).$$

The **Jeffreys prior distribution** is

$$\pi(\theta) \propto \sqrt{I(\theta)}.$$

Now

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

Solution to question 29(ii)

Therefore

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

$$\Rightarrow \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta}$$

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

$$\begin{aligned} \Rightarrow I(\theta) &= E_{\mathbf{X}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right] \\ &= \frac{nE_{\mathbf{X}|\theta}(\bar{X})}{\theta^2} + \frac{n[k - E_{\mathbf{X}|\theta}(\bar{X})]}{(1 - \theta)^2}. \end{aligned}$$

Solution to question 29(ii)

Now $E_{X|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

This is a $Beta(1/2, 1/2)$ prior distribution and so the resulting posterior distribution is

$$\theta|\mathbf{x} \sim Beta(1/2 + n\bar{x}, 1/2 + nk - n\bar{x}).$$

The **asymptotic posterior distribution** (as $n \rightarrow \infty$) is

$$\theta|\mathbf{x} \sim N\left(\hat{\theta}, J(\hat{\theta})^{-1}\right),$$

where

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta^2} + \frac{n(k - \bar{x})}{(1 - \theta)^2}.$$

Solution to question 29(iii)

Now

$$\begin{aligned}\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 &\implies \frac{n\bar{x}}{\hat{\theta}} - \frac{n(k - \bar{x})}{1 - \hat{\theta}} = 0 \\ &\implies \hat{\theta} = \frac{\bar{x}}{k} \\ &\implies J(\hat{\theta}) = \frac{nk^3}{\bar{x}(k - \bar{x})} \\ &\implies J(\hat{\theta})^{-1} = \frac{\bar{x}(k - \bar{x})}{nk^3}.\end{aligned}$$

Therefore, for large n , the posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(\frac{\bar{x}}{k}, \frac{\bar{x}(k - \bar{x})}{nk^3}\right) \quad \text{approximately.}$$

Question 4

Use techniques from MAS2302 to obtain 95% frequentist confidence intervals for $\hat{\theta}_C$ and $\hat{\theta}_I$.

Solution to Question 31

From MAS2302, you should (but probably can't!) recall that, for the binomial proportion p ,

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

where n is the number of trials. The variance is estimated using $\hat{p}$, and recall that, for MLEs generally,

$$\text{Var}(\theta) = I(\theta)^{-1},$$

where

$$I(\theta) = E\left[-\frac{\partial^2}{\partial \theta^2} \ell(\theta)\right].$$

Thus, a 95% frequentists CI for p is given by

$$\hat{p} \pm 1.96 \times \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}.$$