Statistics 352 Winter 2008 Midterm # 1 — Solutions

1. (a) Using Bayes' rule we find

$$P\{\text{Box } i \mid \text{black ball}\} = \frac{P\{\text{black ball} \mid \text{Box } i\}P\{\text{Box } i\}}{P\{\text{black ball}\}}$$

where

$$P\{\text{black ball}\} = \sum_{i=1}^{3} P\{\text{black ball} \mid \text{Box } i\}P\{\text{Box } i\} = \frac{1}{2}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{1}{3} + \frac{1}{4}\cdot\frac{1}{3} = \frac{13}{36}.$$

Therefore,

$$P\{\text{Box 1} \mid \text{black ball}\} = \frac{P\{\text{black ball} \mid \text{Box 1}\}P\{\text{Box 1}\}}{P\{\text{black ball}\}} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{13}{36}} = \frac{6}{13},$$

$$P\{\text{Box 2} \mid \text{black ball}\} = \frac{P\{\text{black ball} \mid \text{Box 2}\}P\{\text{Box 2}\}}{P\{\text{black ball}\}} = \frac{\frac{1}{3}\cdot\frac{1}{3}}{\frac{13}{36}} = \frac{4}{13},$$

and

$$P\{\text{Box 3} \mid \text{black ball}\} = \frac{P\{\text{black ball} \mid \text{Box 3}\}P\{\text{Box 3}\}}{P\{\text{black ball}\}} = \frac{\frac{1}{4}\cdot\frac{1}{3}}{\frac{13}{36}} = \frac{3}{13}$$

which implies that Barry should pick Box 1.

1. (b) Since all three boxes are equally likely, if Barry does not see the ball, then he has the same chance of being correct no matter which box he picks. Thus, Barry should pick Box 2 because 2 is my favourite number. (Of course, Box 1 and Box 3 are also correct answers.)

2. (a) She must choose a and b to satisfy $ab = 8$ and $ab^2 = 4$. Dividing these two expressions implies $b = 1/2$ and so $a = 16$.

2. (b) The posterior density satisfies

$$f(\theta|y = 11) \propto f(y = 11|\theta)g(\theta) \propto \theta^{11}e^{-\theta}\cdot\theta^{15}e^{-2\theta} = \theta^{26}e^{-3\theta}.$$

Thus, the posterior distribution for θ given $y = 11$ is $\Gamma(27, 1/3)$.

2. (c) We can use the posterior distribution from **(b)** as the new prior for θ. Thus, the new posterior density satisfies

$$f(\theta|y_1 = 11, y_2 = 10) \propto f(y_2 = 10|\theta)f(\theta|y_1 = 11) \propto \theta^{10}e^{-\theta}\cdot\theta^{26}e^{-3\theta} = \theta^{36}e^{-4\theta}$$

and so the new posterior distribution for θ given $y = 10$ is $\Gamma(37, 1/4)$.

3. Since the likelihood function

$$f(y|\theta) = \binom{n}{y}\theta^y(1-\theta)^{n-y}, \quad 0 < \theta < 1,$$

satisfies

$$\log f(y|\theta) = \log\binom{n}{y} + y\log\theta + (n-y)\log(1-\theta),$$

we find

$$\frac{\partial}{\partial\theta}\log f(y|\theta) = \frac{y}{\theta} - \frac{n-y}{1-\theta} \quad \text{and} \quad \frac{\partial^2}{\partial\theta^2}\log f(y|\theta) = -\frac{y}{\theta^2} - \frac{n-y}{(1-\theta)^2}.$$

Therefore, the Fisher information is

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(Y|\theta)\right) = \frac{\mathbb{E}(Y)}{\theta^2} + \frac{n-\mathbb{E}(Y)}{(1-\theta)^2} = \frac{n\theta}{\theta^2} + \frac{n-n\theta}{(1-\theta)^2} = \frac{n}{\theta(1-\theta)}.$$

Thus, we take our Jeffreys prior to satisfy

$$g(\theta) \propto \sqrt{I(\theta)} \propto \theta^{-1/2}(1-\theta)^{-1/2}$$

for $0 < \theta < 1$. In other words, the Jeffreys prior distribution for θ is $\beta(1/2, 1/2)$.

4. The posterior density satisfies

$$f(\theta|y=1) \propto f(y=1|\theta)g(\theta) = \frac{1}{\theta}e^{-1/\theta}\cdot\frac{1}{\theta^2} = \frac{1}{\theta^3}e^{-1/\theta}$$

provided that $\theta > 1$. Therefore, the posterior density is

$$f(\theta|y=1) = \frac{\frac{1}{\theta^3}e^{-1/\theta}}{\int_1^\infty \frac{1}{\theta^3}e^{-1/\theta}\,\mathrm{d}\theta}, \quad \theta > 1.$$

Since

$$\int_1^\infty \frac{1}{\theta^3}e^{-1/\theta}\,\mathrm{d}\theta = \int_0^1 xe^{-x}\,\mathrm{d}x = -e^{-1} + \int_0^1 e^{-u}\,\mathrm{d}u = 1 - 2e^{-1}$$

we conclude that

$$f(\theta|y=1) = \frac{\frac{1}{\theta^3}e^{-1/\theta}}{1-2e^{-1}} = \frac{e^{1-1/\theta}}{(e-2)\theta^3}, \quad \theta > 1.$$

5. The posterior density satisfies

$$f(\theta|y) \propto f(y|\theta)g(\theta) = \frac{1}{\theta}\cdot ab^a\theta^{-a-1} \propto \theta^{-a-2} = \theta^{-(a+1)-1}$$

provided that $\theta > b$ AND $\theta > y$; that is, provided $\theta > b'$ where $b' = \max\{y, b\}$ Thus, the posterior density is

$$f(\theta|y) = (a+1)(b')^{a+1}\theta^{-(a+1)-1}, \quad \theta > b',$$

which we recognize as a $\mathrm{Pa}(a', b')$ distribution with $a' = a+1$ and $b' = \max\{b, y\}$.

6. (a) We compute

$$f(y=\{3,0,2,1\}|\theta=1) = \frac{e^{-1}}{3!}\cdot\frac{e^{-1}}{0!}\cdot\frac{e^{-1}}{2!}\cdot\frac{e^{-1}}{1!} = \frac{e^{-4}}{12}$$

and

$$f(y=\{3,0,2,1\}|\theta=2) = \frac{e^{-2}2^3}{3!}\cdot\frac{e^{-2}2^0}{0!}\cdot\frac{e^{-2}2^2}{2!}\cdot\frac{e^{-2}2^1}{1!} = \frac{64e^{-8}}{12} = \frac{16e^{-8}}{3}$$

so that

$$\begin{aligned} f(y=\{3,0,2,1\}) &= f(y=\{3,0,2,1\}|\theta=1)g(\theta=1) + f(y=\{3,0,2,1\}|\theta=2)g(\theta=2) \\ &= \frac{e^{-4}}{12}\cdot\frac{1}{2} + \frac{16e^{-8}}{3}\cdot\frac{1}{2} \\ &= \frac{e^{-4}}{24} + \frac{8e^{-8}}{3}. \end{aligned}$$

Therefore,

$$f(\theta=1|y=\{3,0,2,1\}) = \frac{f(y=\{3,0,2,1\}|\theta=1)g(\theta=1)}{f(y=\{3,0,2,1\})} = \frac{\frac{e^{-4}}{24}}{\frac{e^{-4}}{24}+\frac{8e^{-8}}{3}} = \frac{1}{1+64e^{-4}}.$$

6. (b) From **(a)** we have

$$f(\theta=1|y=\{3,0,2,1\}) = \frac{1}{1+64e^{-4}} \quad (\approx 0.4603626)$$

so that

$$f(\theta=2|y=\{3,0,2,1\}) = \frac{64e^{-4}}{1+64e^{-4}} \quad (\approx 0.5396374).$$

Therefore,

$$\begin{aligned} \mathbb{E}(\theta|y=\{3,0,2,1\}) &= 1\cdot f(\theta=1|y=\{3,0,2,1\}) + 2\cdot f(\theta=2|y=\{3,0,2,1\}) \\ &= \frac{1}{1+64e^{-4}} + 2\cdot\frac{64e^{-4}}{1+64e^{-4}} \\ &= \frac{1+128e^{-4}}{1+64e^{-4}} \quad (\approx 1.539637). \end{aligned}$$

7. We see that the probability of a "Yes" answer is

$$P\{\text{Yes}\} = \frac{1}{2}+\frac{\theta}{2}.$$

Therefore, the binomial model

$$Y \sim \text{Bin}\left(n, \frac{1+\theta}{2}\right)$$

is appropriate for this problem and so the likelihood function is

$$f(y|\theta) = \binom{n}{y}\left(\frac{1+\theta}{2}\right)^y\left(\frac{1-\theta}{2}\right)^{n-y}, \quad 0<\theta<1.$$

(continued)

The corresponding posterior density for θ given y satisfies

$$f(\theta|y) \propto (1+\theta)^y(1-\theta)^{n-y} \cdot 1 = (1+\theta)^y(1-\theta)^{n-y}, \quad 0 < \theta < 1.$$

Therefore, an expression for the posterior mean is

$$\mathbb{E}(\theta|y) = \int_0^1 \theta \cdot f(\theta|y)\,\mathrm{d}\theta = \frac{\int_0^1 \theta(1+\theta)^y(1-\theta)^{n-y}\,\mathrm{d}\theta}{\int_0^1 (1+\theta)^y(1-\theta)^{n-y}\,\mathrm{d}\theta}.$$

8. The posterior density satisfies

$$f(\alpha,\beta|y) \propto f(y|\alpha,\beta)g(\alpha,\beta) \propto \frac{1}{\beta-\alpha} \cdot (\beta-\alpha)^{-4} = (\beta-\alpha)^{-5}$$

provided that $\beta > 2$, $\alpha < 1$. Therefore, the posterior density is

$$f(\alpha,\beta|y) = \frac{(\beta-\alpha)^{-5}}{\int_2^\infty \int_{-\infty}^1 (\beta-\alpha)^{-5}\,\mathrm{d}\alpha\,\mathrm{d}\beta}, \quad \beta > 2, \alpha < 1.$$

Since

$$\int_{-\infty}^1 (\beta-\alpha)^{-5}\,\mathrm{d}\alpha = \frac{1}{4}(\beta-\alpha)^{-4}\Big|_{\alpha=-\infty}^{\alpha=1} = \frac{1}{4}(\beta-1)^{-4}$$

we find

$$\int_2^\infty \int_{-\infty}^1 (\beta-\alpha)^{-5}\,\mathrm{d}\alpha\,\mathrm{d}\beta = \frac{1}{4}\int_2^\infty (\beta-1)^{-4}\,\mathrm{d}\beta = -\frac{1}{12}(\beta-1)^{-3}\Big|_2^\infty = \frac{1}{12},$$

and we conclude that

$$f(\alpha,\beta|y) = 12(\beta-\alpha)^{-5}, \quad \beta > 2, \alpha < 1.$$