

# MA40189 - Solution Sheet One

Simon Shaw, [s.shaw@bath.ac.uk](mailto:s.shaw@bath.ac.uk)  
<https://people.bath.ac.uk/masss/ma40189.html>

2021/22 Semester II

1. For each of the following distributions, write down the probability density function and find a corresponding kernel. Your kernel should be as simple as possible.

(a)  $X | \theta \sim Po(\theta)$ .

$$\begin{aligned} f(x | \theta) = P(X = x | \theta) &= \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, \dots \\ &\propto \frac{\theta^x}{x!}. \end{aligned}$$

(b)  $Y | \theta, \beta \sim Beta(\beta\theta, \beta)$ .

$$\begin{aligned} f(y | \theta, \beta) &= \frac{\Gamma(\beta\theta + \beta)}{\Gamma(\beta\theta)\Gamma(\beta)} y^{\beta\theta-1} (1-y)^{\beta-1}, \quad y \in [0, 1] \\ &\propto y^{\beta\theta-1} (1-y)^{\beta-1}. \end{aligned}$$

(c)  $\theta | \alpha, \beta, x \sim Gamma(\alpha + x + 1, \beta - 3x)$ .

$$\begin{aligned} f(\theta | \alpha, \beta, x) &= \frac{(\beta - 3x)^{\alpha+x+1}}{\Gamma(\alpha + x + 1)} \theta^{\alpha+x} e^{-(\beta-3x)\theta}, \quad \theta > 0 \\ &\propto \theta^{\alpha+x} e^{-(\beta-3x)\theta}. \end{aligned}$$

(d)  $\phi | \mu, \bar{x}, \tau \sim N(\tau\mu + (1-\tau)\bar{x}, \bar{x}^2\tau^{-2})$ .

$$\begin{aligned} f(\phi | \mu, \bar{x}, \tau) &= \frac{\tau}{\sqrt{2\pi\bar{x}}} \exp\left\{-\frac{\tau^2}{2\bar{x}^2} (\phi - \tau\mu - (1-\tau)\bar{x})^2\right\}, \quad -\infty < \phi < \infty \\ &\propto \exp\left\{-\frac{\tau^2}{2\bar{x}^2} (\phi - \tau\mu - (1-\tau)\bar{x})^2\right\} \\ &\propto \exp\left\{-\frac{\tau^2}{2\bar{x}^2} [\phi^2 - 2(\tau\mu + (1-\tau)\bar{x})\phi]\right\}. \end{aligned}$$

2. In each of the following, state the distribution of the corresponding random variable.

(a)  $f(x | \theta) \propto \frac{\theta^x}{x!}$ ,  $x = 0, 1, \dots$ ,  $\theta > 0$ .

From Question 1(a) this is a kernel of a  $Po(\theta)$  random quantity, so  $X | \theta \sim Po(\theta)$ .

(b)  $f(x) \propto e^{-2x}, x > 0$ .

A kernel of the  $Exp(\lambda)$  distribution is  $e^{-\lambda x}$  so that  $X \sim Exp(2)$ . Notice that the  $Exp(\lambda)$  distribution is also the  $Gamma(1, \lambda)$  distribution so that, equivalently,  $X \sim Gamma(1, 2)$ .

(c)  $f(x) \propto 1, 0 \leq x \leq 1$ .

If  $X | a, b \sim U(a, b)$  then a kernel is 1 for  $a \leq x \leq b$ . Here we have  $0 \leq x \leq 1$  so that  $X \sim U(0, 1)$ .

(d)  $f(x | \alpha, \beta) \propto x^{-\alpha} e^{-2\beta/x}, x > 0, \alpha > 1, \beta > 0$ .

A kernel of an inverse gamma distribution with parameters  $a$  and  $b$ , written  $Inv\text{-}gamma(a, b)$ , is  $x^{-(a+1)} e^{-b/x}$  so  $X | \alpha, \beta \sim Inv\text{-}gamma(\alpha - 1, 2\beta)$ .

(e)  $f(x | m) \propto (1 - x)^{(m-1)/2}, 0 \leq x \leq 1, m > -1$ .

A kernel of a  $Beta(\alpha, \beta)$  is  $x^{\alpha-1}(1-x)^{\beta-1}$ . Thus,  $X | m \sim Beta(1, \frac{m+1}{2})$ .

(f)  $f(x | \theta, \phi) \propto \exp\left\{-\frac{\theta}{\phi}x^2 + (\phi + 1)x\right\}, -\infty < x < \infty, \theta > 0, \phi > 0$ .

If  $X | \mu, \sigma^2 \sim N(\mu, \sigma^2)$  then a kernel is  $\exp\left\{-\frac{1}{2\sigma^2}(x^2 - 2\mu x)\right\}$ . Thus,  $X | \theta, \phi \sim N\left(\frac{\phi(\phi+1)}{2\theta}, \frac{\phi}{2\theta}\right)$ .

**3. Find the following sums and integrals by identifying a kernel of a probability density function and using properties of probability density functions.**

(a)  $\sum_{x=0}^{\infty} \frac{\theta^x}{x!}, \theta > 0$ .

From Question 2(a),  $\frac{\theta^x}{x!}$  is a kernel of a  $Po(\theta)$ . Hence,

$$\sum_{x=0}^{\infty} \frac{\theta^x e^{-\theta}}{x!} = 1 \Rightarrow \sum_{x=0}^{\infty} \frac{\theta^x}{x!} = e^{\theta}.$$

(b)  $\sum_{x=0}^{\infty} \frac{x\theta^x}{x!}, \theta > 0$ .

$\frac{x\theta^x}{x!}$  is  $x$  times the kernel of a  $Po(\theta)$ . So, thinking of the expectation of a  $Po(\theta)$ , we have

$$E(X | \theta) = \theta = \sum_{x=0}^{\infty} \frac{x\theta^x e^{-\theta}}{x!} \Rightarrow \sum_{x=0}^{\infty} \frac{x\theta^x}{x!} = \theta e^{\theta}.$$

Alternatively,

$$\begin{aligned} \sum_{x=0}^{\infty} \frac{x\theta^x}{x!} &= \sum_{x=1}^{\infty} \frac{\theta^x}{(x-1)!} = \theta \sum_{x=1}^{\infty} \frac{\theta^{x-1}}{(x-1)!} \\ &= \theta \sum_{y=0}^{\infty} \frac{\theta^y}{y!} = \theta e^{\theta} \end{aligned}$$

where  $y = x - 1$  and the final equation follows from Question 3(a).

(c)  $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$ ,  $\alpha > 0$ ,  $\beta > 0$ .

$x^{\alpha-1}(1-x)^{\beta-1}$  is a kernel of a  $Beta(\alpha, \beta)$  distribution so

$$\begin{aligned} \int_0^1 \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx &= 1 \\ \Rightarrow \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx &= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}. \end{aligned}$$

(d)  $\int_0^1 \frac{1}{2}(\beta + 3)(\beta + 2)(\beta + 1)x(1-x)^\beta dx$ ,  $\beta > -1$ .

$x(1-x)^\beta$  is a kernel of a  $Beta(2, \beta + 1)$  distribution so, from Question 3(c),

$$\int_0^1 x(1-x)^\beta dx = \frac{\Gamma(2)\Gamma(\beta + 1)}{\Gamma(\beta + 3)}.$$

Now, the Gamma function has the property that  $\Gamma(z + 1) = z\Gamma(z)$  for  $z > 0$  so that  $\frac{\Gamma(2)\Gamma(\beta + 1)}{\Gamma(\beta + 3)} = \frac{1}{(\beta + 2)(\beta + 1)}$ . Thus,

$$\begin{aligned} \int_0^1 \frac{1}{2}(\beta + 3)(\beta + 2)(\beta + 1)x(1-x)^\beta dx &= \frac{(\beta + 3)(\beta + 2)(\beta + 1)}{2(\beta + 2)(\beta + 1)} \\ &= \frac{1}{2}(\beta + 3). \end{aligned}$$

(e)  $\int_{-\infty}^{\infty} 4\mu^{(a-1)} \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right)^{\frac{1}{2}} \exp \left\{ - \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right) (\theta - \mu)^2 \right\} d\theta$ ,  $\sigma > 0$ ,  $\tau > 0$ .

$\exp \left\{ - \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right) (\theta - \mu)^2 \right\}$  is a kernel of a  $N \left( \mu, \frac{\sigma\tau}{2(\sigma^2 + \tau^2)} \right)$  distribution. Thus,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \left( \frac{2(\sigma^2 + \tau^2)}{\sigma\tau} \right)^{\frac{1}{2}} \exp \left\{ - \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right) (\theta - \mu)^2 \right\} d\theta &= 1 \\ \Rightarrow \int_{-\infty}^{\infty} \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right)^{\frac{1}{2}} \exp \left\{ - \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right) (\theta - \mu)^2 \right\} d\theta &= \sqrt{\pi} \\ \Rightarrow \int_{-\infty}^{\infty} 4\mu^{(a-1)} \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right)^{\frac{1}{2}} \exp \left\{ - \left( \frac{\sigma^2 + \tau^2}{\sigma\tau} \right) (\theta - \mu)^2 \right\} d\theta &= 4\mu^{(a-1)}\sqrt{\pi}. \end{aligned}$$

4. Suppose that we are interested in  $\theta$ , the probability that a coin will yield a ‘head’ when spun in a specified manner. We judge that the prior distribution is  $\theta \sim Beta(4, 4)$ . The coin is spun ten times (you do not witness the spins) and, rather than being told how many heads were seen, you are only told that the number is less than three.

(a) Find the posterior distribution up to proportionality, and show that the normalising constant  $k$  is given by

$$k = \frac{\Gamma(18)}{1536 \times \Gamma(4)\Gamma(12)}.$$

In this case, for  $X | \theta \sim Bin(n, \theta)$ ,

$$\begin{aligned} f(x | \theta) &= P(X < 3 | \theta) \\ &= \sum_{x=0}^2 \binom{n}{x} \theta^x (1 - \theta)^{n-x}. \end{aligned} \quad (1)$$

For  $\theta \sim Beta(\alpha, \beta)$  we have

$$f(\theta) \propto \theta^{\alpha-1} (1 - \theta)^{\beta-1}. \quad (2)$$

Now as  $f(\theta | x) \propto f(x | \theta) f(\theta)$  then, from (1) and (2),

$$\begin{aligned} f(\theta | x) &\propto \left\{ \binom{n}{x} \theta^x (1 - \theta)^{n-x} \right\} \theta^{\alpha-1} (1 - \theta)^{\beta-1} \\ &= \sum_{x=0}^2 \binom{n}{x} \theta^{\alpha+x-1} (1 - \theta)^{\beta+n-x-1} \end{aligned}$$

so that

$$f(\theta | x) = k \left\{ \sum_{x=0}^2 \binom{n}{x} \theta^{\alpha+x-1} (1 - \theta)^{\beta+n-x-1} \right\}. \quad (3)$$

Now,  $\int_0^1 f(\theta | x) d\theta = 1$  so, from (3),

$$\begin{aligned} k^{-1} &= \int_0^1 \left\{ \sum_{x=0}^2 \binom{n}{x} \theta^{\alpha+x-1} (1 - \theta)^{\beta+n-x-1} \right\} d\theta \\ &= \sum_{x=0}^2 \binom{n}{x} \int_0^1 \theta^{\alpha+x-1} (1 - \theta)^{\beta+n-x-1} d\theta \\ &= \sum_{x=0}^2 \binom{n}{x} B(\alpha + x, \beta + n - x) \end{aligned} \quad (4)$$

where (4) follows since, from Question 3(c),

$$\begin{aligned} \int_0^1 \theta^{\alpha+x-1} (1 - \theta)^{\beta+n-x-1} d\theta &= \frac{\Gamma(\alpha + x) \Gamma(\beta + n - x)}{\Gamma(\alpha + \beta + n)} \\ &= B(\alpha + x, \beta + n - x). \end{aligned}$$

Substituting  $n = 10$ ,  $\alpha = 4$ ,  $\beta = 4$  into (4) we have that

$$\begin{aligned} k^{-1} &= \sum_{x=0}^2 \binom{10}{x} B(4 + x, 14 - x) \\ &= B(4, 14) + 10B(5, 13) + 45B(6, 12) \\ &= \frac{\Gamma(4)\Gamma(14) + 10\Gamma(5)\Gamma(13) + 45\Gamma(6)\Gamma(12)}{\Gamma(18)} \\ &= \frac{\Gamma(4)\Gamma(12)}{\Gamma(18)} \{(13 \times 12) + 10(4 \times 12) + 45(5 \times 4)\} \\ &= \frac{1536 \times \Gamma(4)\Gamma(12)}{\Gamma(18)}. \end{aligned}$$

(b) Show that the posterior mean is  $\frac{39}{128}$ .

From (3) we have that

$$\begin{aligned} E(\theta | X) &= k \sum_{x=0}^2 \binom{n}{x} \int_0^1 \theta \times \theta^{\alpha+x-1} (1-\theta)^{\beta+n-x-1} d\theta \\ &= k \sum_{x=0}^2 \binom{n}{x} B(\alpha+x+1, \beta+n-x). \end{aligned} \quad (5)$$

Substituting  $n = 10$ ,  $\alpha = 4$ ,  $\beta = 4$  into (5) we have that

$$\begin{aligned} E(\theta | X) &= k \sum_{x=0}^2 \binom{n}{x} B(5+x, 14-x) \\ &= k \{B(5, 14) + 10B(6, 13) + 45B(7, 12)\} \\ &= k \left\{ \frac{\Gamma(5)\Gamma(14) + 10\Gamma(6)\Gamma(13) + 45\Gamma(7)\Gamma(12)}{\Gamma(19)} \right\} \\ &= k \frac{\Gamma(5)\Gamma(12)}{\Gamma(19)} \{(13 \times 12) + 10(5 \times 12) + 45(6 \times 5)\} \\ &= k \frac{2106\Gamma(5)\Gamma(12)}{\Gamma(19)} \end{aligned} \quad (6)$$

Substituting for  $k$  in (6) we have

$$\begin{aligned} E(\theta | X) &= \frac{\Gamma(18)}{1536 \times \Gamma(4)\Gamma(12)} \times \frac{2106\Gamma(5)\Gamma(12)}{\Gamma(19)} \\ &= \frac{2106 \times 4}{1536 \times 18} = \frac{39}{128}. \end{aligned}$$

5. Consider the random variables  $X$  and  $Y$  which, for convenience, you may assume are continuous. Recall that the conditional expectation of  $g(X)$  given  $Y$  is defined as

$$E(g(X) | Y) = \int_X g(x)f(x|y) dx$$

for any function  $g(\cdot)$  where  $f(x|y)$  is the conditional distribution of  $X$  given  $Y$ . Prove that

(a)  $E(X) = E(E(X|Y))$ .

Note that  $E(X|Y)$  is a function of  $Y$  so that

$$\begin{aligned} E(E(X|Y)) &= \int_Y E(X|Y)f(y) dy \\ &= \int_Y \left\{ \int_X xf(x|y) dx \right\} f(y) dy \\ &= \int_X x \left\{ \int_Y f(x|y)f(y) dy \right\} dx \\ &= \int_X xf(x) dx = E(X). \end{aligned} \quad (7)$$

(b)  $Var(X) = Var(E(X | Y)) + E(Var(X | Y))$ .

Notice that Question 5(a), by substituting  $g(x)$  for  $x$  in (7), is easily generalised to  $E(g(X)) = E(E(g(X) | Y))$ . Now,

$$\begin{aligned} Var(E(X | Y)) &= E(E^2(X | Y)) - E^2(E(X | Y)) \\ &= E(E^2(X | Y)) - E^2(X). \end{aligned} \tag{8}$$

where (8) follows from Question 5(a). Similarly,

$$\begin{aligned} E(Var(X | Y)) &= E(E(X^2 | Y)) - E(E^2(X | Y)) \\ &= E(X^2) - E(E^2(X | Y)). \end{aligned} \tag{9}$$

where (9) follows from the generalisation of Question 5(a) with  $g(X) = X^2$ . Adding (8) and (9) gives the result.

6. **Consider the random variables  $\theta$  and  $X$ . Consider an estimator  $t(X)$  of  $\theta$  and that we measure the quality of an estimator by its mean square error, the expected square distance between the estimator and  $\theta$ . Show that, given  $X$ , the choice  $t = E(\theta | X)$  minimises this mean square error.**

We seek to minimise  $E((t - \theta)^2 | X)$  where, given  $X = x$ ,  $t = t(X)$  is a constant. Now

$$\begin{aligned} E((t - \theta)^2 | X) &= E(t^2 - 2t\theta + \theta^2 | X) \\ &= t^2 - 2tE(\theta | X) + E(\theta^2 | X) \\ &= t^2 - 2tE(\theta | X) + E^2(\theta | X) + Var(\theta | X) \\ &= \{t - E(\theta | X)\}^2 + Var(\theta | X). \end{aligned}$$

Now  $Var(\theta | X)$  does not depend upon  $t$  so  $E((t - \theta)^2 | X)$  is minimised if we take  $t = E(\theta | X)$ .