

**SMF SOLUTIONS 8. 13.12.2013**

Q1. (i) With  $X$  the number of successes, and as the prior in  $p$  is uniform,

$$P(X = x|p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad P(X = x) = \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} dp :$$

$$\begin{aligned} P(a < p < b|X = x) &= \int_a^b \binom{n}{x} p^x (1-p)^{n-x} dp / \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} dp \\ &= \int_a^b \binom{n}{x} p^x (1-p)^{n-x} dp / B(x+1, n-x+1). \end{aligned}$$

So the posterior is  $B(x+1, n-x+1)$ .

(ii) If the prior is now  $B(\alpha, \beta)$ , as in (i)

$$\begin{aligned} P(a < p < b|X = x) &\propto \int_a^b \binom{n}{x} p^x (1-p)^{n-x} \cdot p^{\alpha-1} (1-p)^{\beta-1} dp \\ &= \int_a^b \binom{n}{x} p^{x+\alpha-1} (1-p)^{n-x+\beta-1} dp. \end{aligned}$$

So the posterior is  $B(x+\alpha, n-x+\beta)$  (observe that  $U(0, 1) = B(1, 1)$ , so (i) is the case  $\alpha = \beta = 1$ ).

Q2. (i) For the Bernoulli distribution  $B(p)$ ,  $f(x; p) = p^x (1-p)^{1-x}$ ,

$$\ell = x \log p + (1-x) \log(1-p), \quad \ell' = \frac{x}{p} - \frac{1-x}{1-p}, \quad \ell'' = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2},$$

$$I(p) = -E[\ell''] = \frac{(1-p)}{(1-p)^2} + \frac{p}{p^2} = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}.$$

(ii) So the Jeffreys prior is  $\pi(p) \propto \sqrt{I(p)} = 1/\sqrt{p(1-p)}$ . This is the Beta distribution  $B(\frac{1}{2}, \frac{1}{2})$ , and  $B(\frac{1}{2}, \frac{1}{2}) = \Gamma(\frac{1}{2}) \cdot \Gamma(\frac{1}{2}) / \Gamma(1) = \pi$ , as  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ . So the Jeffreys prior is the *arc-sine law*:

$$\pi(x) = \frac{1}{\pi \sqrt{x(1-x)}} \quad (x \in [0, 1]).$$

Q3. Recall  $\Gamma(z+1) = z\Gamma(z)$ .  $B(\alpha, \beta)$  has mean

$$\begin{aligned} E[X] &= \int_0^1 x \cdot x^{\alpha-1} (1-x)^{\beta-1} dx / B(\alpha, \beta) = \int_0^1 x^\alpha (1-x)^{\beta-1} dx / B(\alpha, \beta) \\ &= B(\alpha+1, \beta) / B(\alpha, \beta) = \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta)} / \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \alpha / (\alpha + \beta), \end{aligned}$$

Q4. So the posterior mean in Q1(ii) is  $(x + \alpha) / (n + \alpha + \beta)$ . As the amount of data increases,  $n \rightarrow \infty$ , and by SLLN  $x/n \rightarrow p$  a.s., where  $p$  is the true parameter value. With no data,  $x = n = 0$ , and the mean is the prior mean  $\alpha / (\alpha + \beta)$ . The value above is a compromise between these two.

Q5.

$$\begin{aligned} (f_\alpha * f_\beta)(x) &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x y^{\alpha-1} e^{-y} \cdot (x-y)^{\beta-1} e^{-(x-y)} dy \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \cdot e^{-x} \int_0^x y^{\alpha-1} (x-y)^{\beta-1} dy. \end{aligned}$$

In the integral, *I* say, substitute  $y = xu$ . Then  $I = x^{\alpha+\beta-1} \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du = x^{\alpha+\beta-1} B(\alpha, \beta)$ . Combining, the RHS has the form of  $f_{\alpha+\beta}(x)$  (to within constants!):

$$(f_\alpha * f_\beta)(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot B(\alpha, \beta) f_{\alpha+\beta}(x).$$

As both sides are densities, both integrate to 1. So the constant on the RHS is 1, which gives Euler's integral for the Beta function.

*Note.* It is remarkable that this purely probabilistic argument (convolutions of Gamma densities) yields a purely analytic result (Euler's integral for the Beta function).

Q6. (i) The likelihood is  $L = \prod_1^n \theta^{-1} I(x_i \in (0, \theta)) = \theta^{-n} I(\theta > \max)$ ,  $\max := \max(x_1, \dots, x_n)$ . To maximise this, one minimises  $\theta$ , subject to the constraint  $\theta > \max$ . So the MLE is  $\hat{\theta} = \max$ .

(ii) By Fisher-Neyman, for each  $n$   $\max(x_1, \dots, x_n)$  is a sufficient statistic.

(iii) Posterior is proportional to prior times likelihood, so

$$f(\theta | x_1, \dots, x_n) \propto \lambda e^{-\lambda\theta} \cdot \theta^{-n} \quad (\theta > \max(x_1, \dots, x_n)). \quad \text{NHB}$$