Homework 1 – Solutions

1.1 (a) $p(y) = \frac{1}{2}N(1,\sigma^2) + \frac{1}{2}N(2,\sigma^2)$.
(b) $p(\theta=1|y=1) = e^{-\frac{1}{8}[y-1]^2}/[e^{-\frac{1}{8}[y-1]^2} + e^{-\frac{1}{8}[y-2]^2}] = 1/[1+e^{-\frac{1}{8}}]$.
(c) $p(\theta=1|y=1) = 1/1+e^{-\frac{1}{2\sigma^2}}$ is monotone decreasing in σ^2. And $p(\theta=1|y=1) \to 1.0$ as $\sigma^2 \to 0$ and $p(\theta=1|y=1) \to 0.5$ as $\sigma^2 \to \infty$.

2.1 $p(\theta|Y<3) \propto p(\theta)p(Y<3|\theta)$ where $p(\theta) \propto \theta^3(1-\theta)^3$, and
$p(Y<3|\theta) = \sum_{j=0}^{3} p(Y=j|\theta) = (1-\theta)^{10}+10\theta(1-\theta)^9+90\theta^2(1-\theta)^8 = (1-\theta)^8\left[(1-\theta)^2+10\theta(1-\theta)+45\theta^2\right]$.

2.2 Let X = # additional spins until first head, A = "two tails on first two spins". Then $E(X|A) = E(X|A,C_1)p(C_1|A)+E(X|A,C_2)p(C_2|A)$[1] Now find $p(C_1|A) = p(C_1)p(A|C_1)/[p(C_1)p(A|C_1)+p(C_2)p(A|C_2)] = 0.5\cdot 0.16/(0.5\cdot 0.16+0.5\cdot 0.36) = 0.31$, and $p(C_2|A) = 1-p(C_1|A) = 0.69$.
Also, $E(X|A,C_i) = E(X|C_i)$[2]. Note that given C_j, $X \sim NBin(1,\beta_j = \frac{p_j}{1-p_j})$ (see appendix p483) and $E(X|C_j) = 1/\beta_j = \frac{1-p_j}{p_j}$, i.e. $E(X|C_1) = 4/6$ and $E(X|C_2) = 6/4$. Thus $E(X|A) = 4/6\cdot 0.31 + 6/4\cdot 0.69 = 1.24$.

2.3 $X \sim Bin(n,p) \approx N(\mu,\sigma^2)$ with $n=1000, p=\frac{1}{6}$, $\mu = n\cdot p = 166.7, \sigma = \sqrt{np(1-p)} = 11.8$.

2.7 $Y_i \sim N(\theta, 20^2)$, $\theta \sim N(180, 40^2)$, $\bar{y} = 150$.
(a) $\theta|\bar{y} \sim N(m_1,\tau_1^2)$ with $\frac{1}{\tau_1^2} = \frac{1}{40^2} + \frac{n}{20^2}$ and $m_1 = \tau_1^2(180/40^2 + \bar{y}\frac{n}{20^2})$.
(b) $\tilde{y}|y \sim N(m_1, \tau_1^2 + 20^2)$ (see p44).

2.9a Prior: $p(N) = \frac{1}{100}\left(\frac{99}{100}\right)^{N-1}$. Likelihood: $p(y|N) = \frac{1}{N}$ if $y \le N$ and $p(y|N) = 0$ if $y > N$. And thus posterior: $p(N|y) = C^{-1}a^{N-1}/N$, $N \ge y$, with $a = 0.99$ and $C = \sum_y^\infty a^{N-1}/N = 0.047$ (from Mathematica)[3].

2.10a $p(\theta|y) \propto= C^{-1}\prod_{i=1}^5 \left[1+(y_i-\theta)^2\right]$, $0 \le \theta \le 1$. $C = 0.003390658$ (from Splus). Note that the posterior is monotone decreasing although - at first glance - the data seems to give more evidence for a positive location θ than for a negative location. If we had used a normal likelihood, the posterior would have been centered at $\bar{y} = 0.2$.

2.12ab The conjugate prior for a Poisson likelihood, $Y \sim Poi(\theta)$ is a gamma prior, $\theta \sim Ga(\alpha,\beta)$. Choose α,β to match your prior believes about mean and variance on fatal airline accidents world-wide per year. For example, I believe there might be around 50 give/take 40 fatal airline accidents world wide, i.e. $E(Y) = 50$, $Var(Y) = 40^2$. Note that the marginal distribution of Y is a $NBin(\alpha,\beta)$ (see p49). Matching mean and variance of a neg binom with 50 and 40^2, respectively, I find $\alpha = 1.6, \beta = 0.03$.[4]
(a) Compute $\bar{y} = 23.8$, $n = 10$. Then $p(\theta|y) = Ga(\alpha_1,\beta_1)$, with $\alpha_1 = \alpha + n\bar{y} = 239.6$ and $\beta_1 = \beta + n = 10.03$. The predictive distribution[5] $p(y_{n+1}|y) = \int p(y_{n+1}|\theta)p(\theta|y)d\theta = \int Poi(y_{n+1}|\theta)Ga(\theta|\alpha_1,\beta_1)d\theta$. By an argument analoguous to p49 we find $p(y_{n+1}|y) = NBin(\alpha_1,\beta_1)$. A central 95% interval of the $NBin(239, 10.03)$ is $\{14,\ldots,34\}$ (the problem doesn't explicitely state "central" predictive interval – so any 95% predictive interv is fine).
(b) Let X =passenger miles. In Table 2.2, X = pass deaths/death rate (in 100 mio miles). From p50 we find: $p(\theta|y) = Ga(\alpha_2,\beta_2)$, $\alpha_2 = \alpha + \sum y_i = 239.6$, $\beta_2 = \beta + \sum x_i = 57158.72$. To find the

[1] This is just a variation of $E(X) = E[E(X|Y)]$ (why?).
[2] By conditional indep of A and X given C_j
[3] Alternatively, you can evaluate C analytically by noting that C is the antiderivative of a geometric series etc
[4] Alternatively you could have guessed an a priori mean & var for θ and matched with the moments of a Gamma. It's probably easier though to think about an observable (Y), rather than an unobservable parameter (θ).
[5] The book uses notation $\tilde{y}$ for the future observation y_{n+1}.

posterior predictive $p(y_{n+1}|y)$ use:

$$\begin{aligned} p(y_{n+1}|y) &= \int p(y_{n+1}|\theta)p(\theta|y)d\theta = \int Poi(y_{n+1}|x_{n+1}\theta)Ga(\theta|\alpha_2,\beta_2)d\theta \\ &= \int Poi(y_{n+1}|\eta)Ga(\eta|\alpha_2,\beta_2/x_{n+1}) = NBin(y_{n+1}|\alpha_2,\beta_2/x_{n+1}) = NBin(y_{n+1}|239.6, 7.144) \end{aligned}$$

The third equation comes from a change of variables from θ to $\eta = x_{n+1}\theta$. Note that if $\theta \sim Ga(a,b)$, then $\eta = r\theta \sim Ga(a, b/r)$. The fourth equation is true by the result from p49/bottom (see also p483, appendix). (There's an obivious typo, either $x_{n+1} = 8000 \cdot 100$ mio miles, or death rates are in deats per mio miles). The desired predictive interval is the central 95% interval for a $NBin(\alpha_2, \beta_2/x_{n+1})$. Using, for example Splus, find it to be $\{22, 23, \ldots, 43\}$.

2.15a

$$p(y) = \frac{p(y|\theta)p(\theta)}{p(\theta|y)} = \ldots = \frac{\Gamma(n+1)}{\Gamma(y+1)\Gamma(n-y+1)} \frac{\Gamma(\alpha+y)\Gamma(n+\beta-y)}{\Gamma(\alpha+\beta+n)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

(use $y! = \Gamma(y+1)$ etc.). Alternatively, you could have written $p(y) = \int p(y|\theta)p(\theta|\alpha,\beta)$ and solved the integral by pulling constants out of the integral and recognizing the remaining integrand as a Beta kernel.

2.20 (a) Straightforward.
(b) If $\phi = f(\theta)$ then $p_\phi(\phi) = p_\theta(f^{-1}(\phi))|J|^{-1}$ ("change of variable formula"), where $J = df/d\theta$. See p21 for a general statement. Applying this for $\phi = \theta^{-1}$ gives the inverse gamma p.d.f. for $p_\phi(\phi)$. See p475 for a statement of the pdf's.