Petter Mostad
Applied Mathematics and Statistics
Chalmers

Suggested solutions for
MVE550 Stochastic Processes and Bayesian Inference
Re-exam April 9 2021

1. (a) We try with the Beta family: Assume $\theta \sim \text{Beta}(\alpha,\beta)$. Then

$$\begin{aligned}\pi(\theta \mid x) &\propto_\theta \pi(x \mid \theta)\pi(\theta) \\ &\propto_\theta \theta(1-\theta)^x\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \theta^\alpha(1-\theta)^{\beta+x-1}\end{aligned}$$

so $\theta \mid x \sim \text{Beta}(\alpha+1,\beta+x)$ and the Beta family is a conjugate family of priors.

(b) One way to compute is the following

$$\pi(x) = \int_0^1 \theta(1-\theta)^x\, d\theta = \frac{\Gamma(2)\Gamma(x+1)}{\Gamma(3+x)} = \frac{1}{(x+1)(x+2)}$$

where we have used the formula for the density of a $\text{Beta}(2, x+1)$ distribution to compute the integral. Another way is to compute

$$\begin{aligned}\pi(x) &= \frac{\pi(x \mid \theta)\pi(\theta)}{\pi(\theta \mid x)} = \frac{\theta(1-\theta)^x \cdot \text{Beta}(\theta; 1, 1)}{\text{Beta}(\theta; 2, 1+x)} \\ &= \frac{\theta(1-\theta)^x}{\frac{\Gamma(3+x)}{\Gamma(2)\Gamma(1+x)}\theta(1-\theta)^x} = \frac{\Gamma(2)\Gamma(1+x)}{\Gamma(3+x)} = \frac{1}{(x+1)(x+2)}.\end{aligned}$$

(c) Let p be the vector of length $n-1$ containing the prior, so that $p_i = \Pr[\theta = i/n]$. If we have observations $x_1, \dots, x_k$, define vectors $v_1, \dots, v_k$ by

$$v_{ji} = \frac{i}{n}\left(1 - \frac{i}{n}\right)^{x_j}$$

for $j = 1, \dots, k$, $i = 1, \dots, n-1$. Then compute the vector v with

$$v_i = p_i \cdot v_{1i} \cdot \dots \cdot v_{ki}$$

and normalize it so that it sums to 1: This probability vector is then the posterior for θ.

2. (a) The transition matrix, after changing d into an absorbing state, becomes

$$P' = \begin{bmatrix} 0 & 0.3 & 0.7 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0.1 & 0.4 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} Q & R \\ 0 & 1 \end{bmatrix}$$

where

$$Q = \begin{bmatrix} 0 & 0.3 & 0.7 \\ 0 & 0 & 0.2 \\ 0.1 & 0.4 & 0 \end{bmatrix}.$$

The fundamental matrix is then $F = (I - Q)^{-1}$ and the expected number of steps before hitting d is the sum of the entries in the third row of this matrix. In R we can write

```
Q <- matrix(c(0, 0, 0.1, 0.3, 0, 0.4, 0.7, 0.2, 0), 3, 3)
F <- solve(diag(3)-Q)
print(sum(F[3,]))
```

which yields the numeric result 1.812796. It is also possible to use what Dobrow calls “first step analysis” to obtain the same result.

(b) We need a Markov chain which records not only the current state, but also how far we might have come in constructing the sequence *abc*. We can use the transition graph in Figure 1.

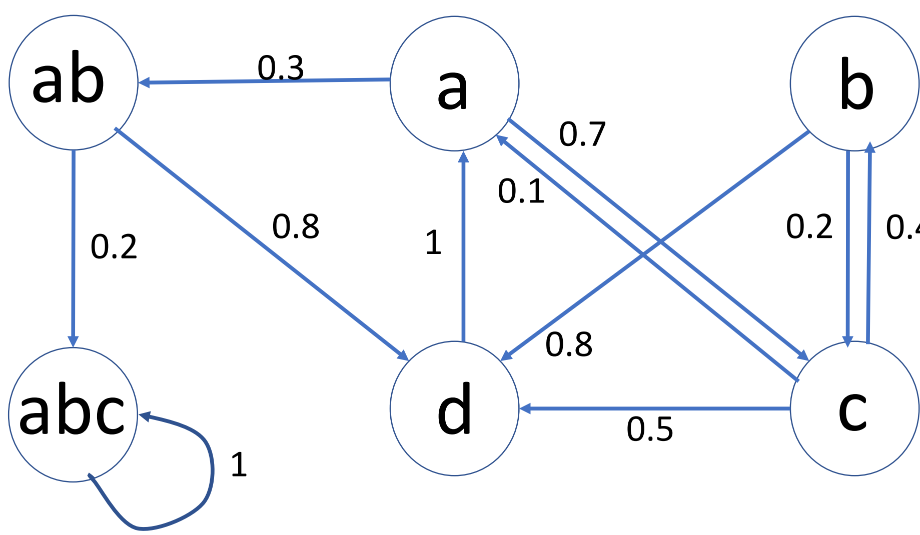

Figure 1: The graph for question 2

3. (a) We get

$$\begin{aligned} G(s) &= \mathrm{E}\left[s^X\right] = a_0 + (1-a_0)\sum_{k=2}^{\infty}\left(\frac{1}{2}\right)^{k-1} s^k \\ &= a_0 + (1-a_0)s\sum_{k=2}^{\infty}\left(\frac{s}{2}\right)^{k-1} \\ &= a_0 + (1-a_0)s\frac{s/2}{1-s/2} = a_0 + (1-a_0)\frac{s^2}{2-s} \end{aligned}$$

(b) To find the expectation we may differentiate $G(s)$:

$$G'(s) = (1-a_0)\frac{(2-s)2s+s^2}{(2-s)^2} = (1-a_0)\frac{4s-s^2}{(2-s)^2}$$

Thus $\mathrm{E}[X] = G'(1) = 3(1-a_0)$. The Branching process is supercritical if and only if $\mathrm{E}[X] > 1$, i.e., if

$$3(1-a_0) > 1$$

which gives $a_0 < \frac{2}{3}$.

(c) The extinction probability is the smallest positive root of the equation $G(s) = s$, i.e., of

$$a_0 + (1-a_0)\frac{s^2}{2-s} = s,$$

which yields the 2nd degree equation

$$s^2 - \frac{2+a_0}{2-a_0}s + \frac{2a_0}{2-a_0} = 0.$$

We know that $G(1) = 1$, so 1 is a root of this equation. Using that, we get the factorization

$$(s-1)\left(s - \frac{2a_0}{2-a_0}\right) = 0$$

and the smallest positive root, when $a_0 < \frac{2}{3}$, is $\frac{2a_0}{2-a_0}$. In summary: When $0 < a_0 < \frac{2}{3}$, the excinction probability is $\frac{2a_0}{2-a_0}$, while when $\frac{2}{3} \leq a_0 < 1$, the extinction probability is 1.

(d) There are 4 observations of the offspring distribution in Figure 2. In one of those there is no offspring, while in the other 3 there are 2 or more offspring. The likelihood for the first observation is a_0, while the likelihoods for the other three observations are proportional to $1-a_0$ as a function of a_0. With a prior that is uniform on $(0,1)$ we get that the posterior is proportional to

$$a_0^1(1-a_0)^3$$

Comparing with the Beta density, we see that

$$p_0 \mid \text{data} \sim \mathrm{Beta}(2,4).$$

4. (a) Let X_A and X_B be the number of customers of type A and B, respectively, during the first two hours. We get $X_A \sim \text{Poisson}(2 \cdot 3) = \text{Poisson}(6)$ and $X_B \sim \text{Poisson}(2 \cdot 2) = \text{Poisson}(4)$. The answer to the question becomes

$$\begin{aligned} & \Pr[X_A \geq 3] \Pr[X_B = 2] \\ = & (1 - \Pr[X_A = 0] - \Pr[X_A = 1] - \Pr[X_A = 2]) \Pr[X_B = 2] \\ = & (1 - e^{-6}(1 + 6 + 6^2/2)) e^{-4} 4^2/2 = 0.1374451 \end{aligned}$$

This can also be computed in R with

```
(1-ppois(2, 6))*dpois(2, 4)
```

(b) Given that a fixed number of customers arrive, the arrival time of a randomly selected customer among these will be uniformly distributed. Thus the probability is $\frac{3/4}{2} = 0.375$.

5. (a) We get

$$Q = \begin{bmatrix} -2.5 & 2 & 0.5 \\ 0.3 & -0.4 & 0.1 \\ 1.5 & 0 & -1.5 \end{bmatrix}$$

for the generator matrix. To find the limiting distribution $v = (v_1, v_2, v_3)$ we need to solve the equations $vQ = 0$ and $v_1 + v_2 + v_3 = 1$. If we let Q' be the matrix Q with the last column replaced by 1's, we get that we need to solve the equation

$$vQ' = (0, 0, 1)$$

Possible R code is

```
Q <- matrix(c(-2.5, 0.3, 1.5, 2, -0.4, 0, 1, 1, 1), 3, 3)
print(c(0, 0, 1)%*%solve(Q))
```

yielding the numerical answer

$$(0.15, 0.75, 0.1)$$

Thus the answer to the original question is 0.75.

(b) We first find the transition matrix for the embedded chain:

$$\tilde{P} = \begin{bmatrix} 0 & 0.8 & 0.2 \\ 0.75 & 0 & 0.25 \\ 1 & 0 & 0 \end{bmatrix}.$$

In order to find the limiting distribution $w = (w_1, w_2, w_3)$ for the discrete-time Markov chain, we need to solve the equations $w_1 + w_2 + w_3 = 1$ and $w\tilde{P} = w$, or equivalently $w(\tilde{P} - I) = 0$. With similar computations as in (a), we get

```
Q <- matrix(c(-1, 0.75, 1, 0.8, -1, 0, 1, 1, 1), 3, 3)
print(c(0, 0, 1)%*%solve(Q))
```

yielding the numerical answer

$$(0.4545455, 0.3636364, 0.1818182)$$

Thus the answer to the original question is 0.3636364. Note that the result can also be found directly from (a) using the relationship between the limiting distributions of a continuous-time Markov chain and its embedded chain:

$$\psi_2 = \frac{\pi_2 q_2}{\pi_1 q_1 + \pi_2 q_2 + \pi_3 q_3} = \frac{0.75 \cdot 0.4}{0.15 \cdot 2.5 + 0.75 \cdot 0.4 + 0.1 \cdot 1.5} = \frac{4}{11} = 0.3636364.$$

6. (a) We get

$$\begin{aligned} & aB_t + bB_{2t} + cB_{3c} \\ = \quad & aB_t + b(B_{2t} - B_t) + bB_t + c(B_{3t} - B_{2t}) + c(B_{2t} - B_t) + cB_t \\ = \quad & (a+b+c)B_t + (b+c)(B_{2t} - B_t) + c(B_{3t} - B_{2t}) \end{aligned}$$

This is a sum of three independent normally distributed variables, and it has a normal distribution. We see directly that the expectation is zero, and for the variance we get

$$\begin{aligned} & \text{Var}\left[(a+b+c)B_t + (b+c)(B_{2t} - B_t) + c(B_{3t} - B_{2t})\right] \\ = \quad & (a+b+c)^2 t + (b+c)^2 t + c^2 t \\ = \quad & \left((a+b+c)^2 + (b+c)^2 + c^2\right) t \end{aligned}$$

So

$$aB_t + bB_{2t} + cB_{3c} \sim \text{Normal}\left(0, \left((a+b+c)^2 + (b+c)^2 + c^2\right) t\right)$$

(b) One may prove this directly from the definition: One must then prove each of the 5 defining properties of Brownian motion mentioned in Dobrow. Alternative one may first argue that $-B_t$ is a Gaussian process: As Brownian motion is a Gaussian process, any linear combination of variables from the process has a multivariate normal distribution, so this is also true for any linear combination of variables from the process $-B_t$, so $-B_t$ is a Gaussian process. It also satisfies $-B_0 = 0$, $\text{E}[-B_t] = 0$, and $\text{Cov}[-B_s, -B_t] = \text{Cov}[B_s, B_t] = \min\{s, t\}$. Finally, $t \mapsto -B_t$ is clearly a continuous map. By a theorem in Dobrow, $-B_t$ is Brownian motion.

(c) If there is exactly one such t, that implies that there is at least one such t, which implies that $T_{1.4} < 1$, where $T_{1.4}$ is the first hitting time for 1.4. As the first hitting time is a stopping time, we get that $B_{T_{1.4}+t} - B_{T_{1.4}}$ is brownian motion. We know that the probability that this process has a zero in the interval $(0, \epsilon)$ is 1, for any ϵ. Thus we can find another t, with $t < 1$, where the original Brownian motion will be equal to 1.4. In fact, with probability 1, there will be infinitely many t with $t < 1$ where $B_t = 1.4$, as long as we assume there is at least one such t. But this means that the probability that there is exactly one such t is zero.