Petter Mostad
Applied Mathematics and Statistics
Chalmers

Suggested solutions for
MVE550 Stochastic Processes and Bayesian Inference
Exam January 8 2022

1. (a) We get

$$\begin{aligned}
\mathrm{E}[Z_n] &= \mathrm{E}\left[\sum_{i=1}^{Z_{n-1}} X_i\right] = \mathrm{E}\left[\mathrm{E}\left[\sum_{i=1}^{k} X_i \mid Z_{n-1}=k\right]\right] = \mathrm{E}\left[\sum_{i=1}^{k} \mathrm{E}[X_i] \mid Z_{n-1}=k\right] \\
&= \mathrm{E}[k\mu \mid Z_{n-1}=k] = \mathrm{E}[\mu Z_{n-1}] = \mu\,\mathrm{E}[Z_{n-1}].
\end{aligned}$$

Application of recursion and the fact that $\mathrm{E}[Z_0] = \mathrm{E}[1] = 1$ gives

$$\mathrm{E}[Z_n] = \mu^n.$$

(b) We get

$$\begin{aligned}
\mathrm{Var}[Z_n] &= \mathrm{Var}\left[\sum_{i=1}^{Z_{n-1}} X_i\right] \\
&= \mathrm{Var}\left[\mathrm{E}\left[\sum_{i=1}^{k} X_i \mid Z_{i-1}=k\right]\right] + \mathrm{E}\left[\mathrm{Var}\left[\sum_{i=1}^{k} X_i \mid Z_{n-1}=k\right]\right] \\
&= \mathrm{Var}[\mu Z_{n-1}] + \mathrm{E}\left[\sigma^2 Z_{n-1}\right] = \mu^2\,\mathrm{Var}[Z_{n-1}] + \mu^{n-1}\sigma^2.
\end{aligned}$$

(c) We get directly $\mathrm{Var}[Z_0] = 0$ and $\mathrm{Var}[Z_1] = \mathrm{Var}[X_1] = \sigma^2$ Using the result from (b) repeatedly we get

$$\begin{aligned}
\mathrm{Var}[Z_2] &= \mu^2\sigma^2 + \mu\sigma^2 \\
\mathrm{Var}[Z_3] &= \mu^4\sigma^2 + \mu^3\sigma^2 + \mu^2\sigma^2 \\
\mathrm{Var}[Z_4] &= \mu^6\sigma^2 + \mu^5\sigma^2 + \mu^4\sigma^2 + \mu^3\sigma^2
\end{aligned}$$

We hypothesize that for $n \geq 1$

$$\mathrm{Var}[Z_n] = \sigma^2 \sum_{i=n-1}^{2n-2} \mu^i$$

and prove the formula by induction: First, it is true for $n = 1$, and secondly, assuming it is true for $n-1$ we get

$$\begin{aligned}
\mathrm{Var}\left[Z_n\right] &= \mu^2\,\mathrm{Var}\left[Z_{n-1}\right] + \mu^{n-1}\sigma^2 \\
&= \mu^2\sigma^2 \sum_{i=n-2}^{2(n-1)-2} \mu^i + \mu^{n-1}\sigma^2 \\
&= \sigma^2 \left(\sum_{i=n}^{2n-2} \mu^i + \mu^2 \right) \\
&= \sigma^2 \sum_{i=n-1}^{2n-2} \mu^i
\end{aligned}$$

so the proof is complete. Note that we can write the result as

$$\mathrm{Var}\left[Z_n\right] = \sigma^2 \sum_{i=n-1}^{2n-2} \mu^i = \sigma^2\mu^{n-1} \sum_{i=0}^{n-1} \mu^i = \sigma^2\mu^{n-1}\frac{\mu^n - 1}{\mu - 1}.$$

2. (a) There have been a total of $23 + 44 = 67$ customers during the $5 \cdot 8 = 40$ hours of observation. Thus the likelihood function is Poisson$(67; 40\lambda)$. We get

$$\begin{aligned}
\pi(\lambda \mid \text{data}) &\propto_\lambda \pi(\text{data} \mid \lambda)\pi(\lambda) \\
&\propto_\lambda e^{-40\lambda}\frac{(40\lambda)^{67}}{67!} \cdot \frac{1}{\lambda} \\
&\propto_\lambda e^{-40\lambda}\lambda^{67-1}
\end{aligned}$$

so that

$$\pi(\lambda \mid \text{data}) = \mathrm{Gamma}(\lambda; 67, 40).$$

The probability p asked for can be expressed as an integal as

$$p = \int_{1.2}^{\infty} \frac{40^{67}}{67!}\lambda^{67-1}\exp(-40\lambda)\,d\lambda$$

or in R

```
1 - pgamma(1.2, 67, 40)
```

(b) We get for the posterior

$$\begin{aligned}
\pi(\lambda, \mu \mid \text{data}) &\propto_{\lambda,\mu} \pi(\text{data} \mid \lambda, \mu)\pi(\lambda, \mu) \\
&\propto_{\lambda,\mu} \mathrm{Poisson}(23; 20\lambda) \cdot \mathrm{Poisson}(44, 20\lambda\mu) \cdot \frac{1}{\lambda} \cdot \frac{1}{\mu} \\
&\propto_{\lambda,\mu} e^{-20\lambda}(20\lambda)^{23}e^{-20\lambda\mu}(20\lambda\mu)^{44}\frac{1}{\lambda\mu} \\
&= \exp(-20\lambda(1+\mu))\lambda^{67-1}\mu^{44-1}
\end{aligned}$$

(c) Alex might use a Metropolis Hastings algorithm to obtain such an posterior. The algorithm would start with reasonable values for the parameters (for example the value 1) and use a proposal density $q(\theta^* \mid \theta)$ in each iteration. Generally, the algorithm would iterate between making a proposed new density according to $q(\theta^* \mid \theta)$ and accepting it with probability

$$a = \min\left(1, \frac{\pi(\theta^* \mid \text{data})q(\theta \mid \theta^*)}{\pi(\theta \mid \text{data})q(\theta^* \mid \theta)}\right).$$

If θ^* is not accepted, the old value θ would be repeated.

For the posterior $\pi(\theta \mid \text{data})$ we get (writing $\mu_0 = 1$, assuming the counts of the different hours are $c_1, c_2, \dots, c_8$, respectively, and using the priors $\mu_i \propto_{\mu_i} 1/\mu_i$)

$$\begin{aligned}
\pi(\theta \mid \text{data}) &\propto_\theta \prod_{i=1}^{8} \text{Poisson}(c_i; 5\lambda\mu_{i-1}) \cdot \frac{1}{\lambda\mu_1 \dots \mu_7} \\
&\propto_\theta \prod_{i=1}^{8} e^{-5\lambda\mu_{i-1}}(5\lambda\mu_{i-1})^{c_i} \cdot \frac{1}{\lambda\mu_1 \dots, \mu_7} \\
&\propto_\theta \lambda^{67-1}\mu_1^{c_2-1}\mu_2^{c_3-1} \dots \mu_7^{c_8-1} \exp(-5\lambda(1 + \mu_1 + \dots + \mu_7))
\end{aligned}$$

An alternative would be to use Gibbs sampling, in which case one would cycle through simulating from the conditional distribution of each of the parameters given fixed values for the others. From the expression of the posterior above we see that these conditional distributions would all be Gamma distributions.

3. (a) This will be a Markov chain, as the position at each time step only depends on the position at the previous time step. However, this Markov chain is not ergodic: In fact it is periodic, of period 2, as the knight will alternate between black and white squares. Because of the periodicity, there is also no limiting distribution.

(b) The Markov chain may be viewed as a random walk on a graph: The graph would consist of all the 64 squares in the board game, and each square has a degree 8 because of the extended way we allow the knight to move. Given the comment in the question about getting from any square to any other, the Markov chain is irreducible. Thus there is a unique stationary distribution. Because all the 64 states have degree 8, the stationary distribution is uniform. The long-term proportion of steps spent at the starting square is 1/64, and the expected return time to the starting square becomes 64.

(c) It is still possible to look at this as an irreducible random walk on a graph, but now the states do not all have degree 8. To do computations, one needs to find the degree of each of the 64 states. If d denotes the sum of all the degrees, we know that the long-term proportion of steps spent at the start square is $\frac{2}{d}$, as the start square has degree 2. Thus the expected number of steps to return to the start square becomes $d/2$.

In fact, there are 4 squares with degree 2, 3 with degree 3, 20 with degree 4, 16 with degree 6, and 16 with degree 8. This means that $d = 336$ and that the expected number of steps to return to the start square is 168.

4. (a) We get

$$Q = \begin{bmatrix} -2 & 2 & 0 & 0 & 0 & 0 \\ 1 & -6 & 3 & 2 & 0 & 0 \\ 0 & 4 & -4 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4 & 1 & 2 \\ 0 & 0 & 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 4 & 0 & -4 \end{bmatrix}.$$

(b) As the rate graph is a tree, the process is time reversible.

(c) It is clear that the process is irreducible and ergodic. To find the limiting distribution we find the v with positive values summing to 1 such that $vQ = 0$. We can do this by replacing the first column of Q with ones, producing Q', and then requiring that vQ' should be the vector $(1, 0, 0, 0, 0, 0)$. Finally, to find the long term proportion of time that the process will spend in C, we take the third element of v. In matrix terms we need to compute

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & -6 & 3 & 2 & 0 & 0 \\ 1 & 4 & -4 & 0 & 0 & 0 \\ 1 & 1 & 0 & -4 & 1 & 2 \\ 1 & 0 & 0 & 3 & -3 & 0 \\ 1 & 0 & 0 & 4 & 0 & -4 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

(d) To find this expected time we make states A and F absorbing states. Removing rows and columns connected to these states we are left with a matrix

$$Q_0 = \begin{bmatrix} -6 & 3 & 2 & 0 \\ 4 & -4 & 0 & 0 \\ 1 & 0 & -4 & 1 \\ 0 & 0 & 3 & -3 \end{bmatrix}.$$

The matrix F of expected times spent in each state before absorbtion is given by $F = -Q^{-1}$. The answer is given by the sum of the second line (corresponding to state C) of this matrix, so we must compute

$$-\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -6 & 3 & 2 & 0 \\ 4 & -4 & 0 & 0 \\ 1 & 0 & -4 & 1 \\ 0 & 0 & 3 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

(e) To answer this question, we first find the transition matrix of the embedded discrete-time Markov chain, which becomes

$$\tilde{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1/6 & 0 & 1/2 & 1/3 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 & 1/4 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The fundamental matrix when making A into an absorbing state will be $F = (I - \tilde{P}_{-A})^{-1}$, where $\tilde{P}_{-A}$ is $\tilde{P}$ with the row and columnd representing A removed. Our desired answer is the sum of the second row of this matrix, i.e, the answer is

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1/2 & -1/3 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -1/4 & 0 & 1 & -1/4 & -1/2 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

5. (a) A Brownian bridge is Brownian motion on the interval $[0, 1]$ conditional on $B_1 = 1$.

(b) One may simulate Brownian motion as usual and then for each t subtract tB_1 from the simulated values.

(c) Using that we can write $X_t = B_t - tB_1$ we get

$$\begin{aligned} \mathrm{Cov}[X_s, X_r] &= \mathrm{Cov}[B_s - xB_1, B_r - rB_1] \\ &= \mathrm{Cov}[B_s, B_r] - s\,\mathrm{Cov}[B_1, B_r] - r\,\mathrm{Cov}[B_s, B_1] + sr\,\mathrm{Cov}[B_1, B_2] \\ &= s - sr - rs + sr\,\mathrm{Var}[B_1] \\ &= s - sr \end{aligned}$$