Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE550 Stochastic Processes and Bayesian Inference

Exam April 16, 2025, 8:30 - 12:30
Examiner: Petter Mostad will be available by phone 0707163235,
and will visit the exam at 9:30 and 11:30.
Allowed aids: Chalmers-approved calculator
Total number of points: 30. At least 12 points are needed to pass.
See appendix for some information about some probability distributions.
All answers need to be explicitly computed or explicitly argued for.

1. (4 points) Consider a discrete-time Markov chain with state space $\{1, 2, 3\}$. Assume you have observed the counts of transitions from states given in the rows to states given in the columns listed in the following table:

	1	2	3
1	4	3	0
2	3	4	2
3	0	3	3

   Assume we use a prior for the transition matrix where all possible transition matrices have equal prior probability.

   (a) If the process currently is in state 1, what is the posterior probability that a process goes to state 2 in the next step?

   (b) If the process currently is in state 1, what is the posterior probability that the following happens: It stays for one step in state 1 and then moves to state 2 in the following step?

   (c) Assume instead of the prior given above you use a prior where the process cannot jump from 1 to 3 or from 3 to 1, but where all transition matrices respecting this have equal probability. Answer again questions (a) and (b) under this assumption.

2. (4 points) Consider the discrete-time Markov chain illustrated in Figure 1.

   (a) For each state, describe whether it is transient[1] or not, and give its period.

   (b) Is the Markov chain ergodic? Why or why not?

[1] Unfortunately, this was written "transitory" in the original exam

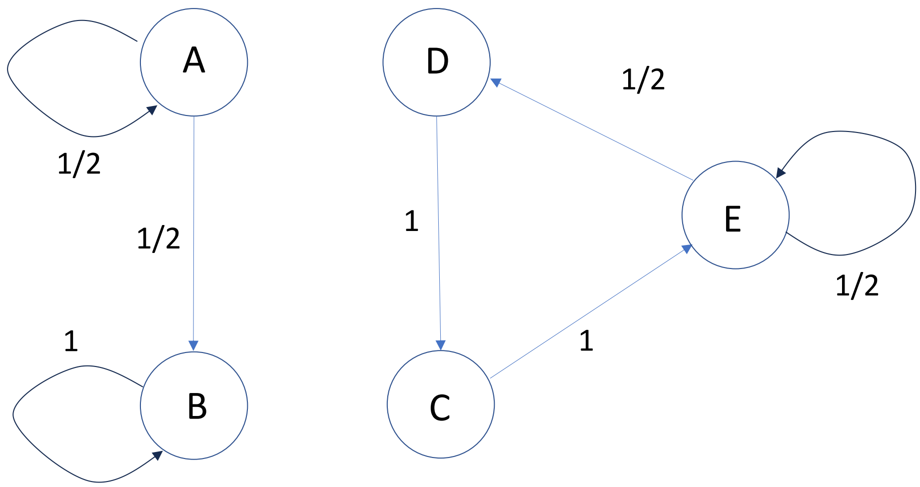

Figure 1: Illustration for question 2

(c) Does the chain have a limiting distribution? If so, present such a limiting distribution.

(d) Does the chain have a stationary distribution? If so, present such a stationary distribution.

3. (6 points) Consider a branching process. Let X be a random variable with the offspring distribution of the branching process. Let G_X be the probability generating function for X.

   (a) Compute (i.e., write down and prove) a recursive formula for the expectation of the branching process generation size Z_n for any $n > 0$.

   (b) Compute (i.e., write down and prove) a recursive formula for the variance of the branching process generation size Z_n for any $n > 0$.

   (c) Let e_n denote the probability that the branching process is extinct at step n. Compute (i.e., write down and prove) a recursive formula for e_n for any $n > 0$. (Hint: Consider what happens in the first step of the branching process).

   (d) Prove that G_X is convex on the interval $[0, 1]$.

4. (4 points) Assume $f(x, y, z)$ is a positive real function proportional to a density, where x, y, z are numbers that can take on any real values. Assume $q(x^*, y^*, z^* \mid x, y, z)$ is a positive density for any real values x^*, y^*, z^*, given any fixed real values x, y, z.

   (a) Write down the details of a Metropolis Hastings algorithm that constructs a Markov chain whose limiting distribution is the density proportional to f and that uses q as a proposal density.

(b) Describe how you can use a realization $(x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_n, y_n, z_n)$ from the chain above to find the approximate probability that $x > y$ under the distribution whose density is porportional to f.

5. (5 points) Male customers arrive at a shop according to a Poisson process with an average of 5 arriving each hour. Independently, female customers also arrive according to a Poisson process with an average of 2 per hour.

   (a) What is the probability that the first customer that arrives will be male?

   (b) What is the probability that the three first customers that arrive will be female?

   (c) What is the probability distribution of the arrival time of the fifth customer?

   (d) What is the probability that there are no customers the first 30 minutes?

   (e) If exactly 5 customers arrive during the first hour, what is the probability that exactly one customer arrives during the first 20 minutes, and that this customer is male?

6. (2 points) Compute the derivative with respect to t of e^{tA}, where A is a matrix, using the definition of matrix exponentials.

7. (3 points) Assume $P(t)$ is the transition function of a continuous-time discrete state space Markov process, and that the derivative of $P(t)$ at $t = 0$ is the matrix

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

   Does the process have a limiting distribution? If so, compute it.

8. (2 points) Let B_t denote Brownian motion. Find the probability distribution of $B_2 + 5B_3 - 3B_6$.

Appendix: Some probability distributions

The Beta distribution

If $x \in [0, 1]$ has a Beta distribution with parameters with $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}.$$

We write $x \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Beta}(x; \alpha, \beta)$.

The Beta-Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Beta-Binomial distribution, with n a positive integer and parameters $\alpha > 0$ and $\beta > 0$, then the probability mass function is

$$\pi(x \mid n, \alpha, \beta) = \binom{n}{x} \frac{\Gamma(x + \alpha)\Gamma(n - x + \beta)\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n + \alpha + \beta)}.$$

We write $x \mid n, \alpha, \beta \sim \text{Beta-Binomial}(n, \alpha, \beta)$ and $\pi(x \mid n, \alpha, \beta) = \text{Beta-Binomial}(x; n, \alpha, \beta)$.

The Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Binomial distribution, with n a positive integer and $0 \leq p \leq 1$, then the probability mass function is

$$\pi(x \mid n, p) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

We write $x \mid n, p \sim \text{Binomial}(n, p)$ and $\pi(x \mid n, p) = \text{Binomial}(x; n, p)$.

The Dirichlet distribution

If $x = (x_1, x_2, \ldots, x_n)$ has a Dirichlet distribution, with $x_i \geq 0$ and $\sum_{i=1}^n x_i = 1$ and with parameters $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_1 > 0, \ldots, \alpha_n > 0$, then the density function is

$$\pi(x \mid \alpha) = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_n)} p_1^{\alpha_1-1} p_2^{\alpha_2-1} \cdots p_n^{\alpha_n-1}.$$

We write $x \mid \alpha \sim \text{Dirichlet}(\alpha)$ and $\pi(x \mid \alpha) = \text{Dirichlet}(x; \alpha)$.

The Exponential distribution

If $x \geq 0$ has an Exponential distribution with parameter $\lambda > 0$, then the density is

$$\pi(x \mid \lambda) = \lambda \exp(-\lambda x)$$

We write $x \mid \lambda \sim \text{Exponential}(\lambda)$ and $\pi(x \mid \lambda) = \text{Exponential}(x; \lambda)$. The expectation is $1/\lambda$ and the variance is $1/\lambda^2$.

The Gamma distribution

If $x > 0$ has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x).$$

We write $x \mid \alpha, \beta \sim \text{Gamma}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Gamma}(x; \alpha, \beta)$. The $\text{Gamma}(\alpha, \beta)$ distribution has expectation $\frac{\alpha}{\beta}$ and variance $\frac{\alpha}{\beta^2}$.

The Inverse Gamma distribution

If $x > 0$ has an Inverse Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp(-\beta/x).$$

We write $x \mid \alpha, \beta \sim \text{Inverse-Gamma}(\alpha, \beta)$ and $\pi(x \mid \alpha, \beta) = \text{Inverse-Gamma}(x; \alpha, \beta)$. If $x \sim \text{Gamma}(\alpha, \beta)$ then $1/x \sim \text{Inverse-Gamma}(\alpha, \beta)$. The $\text{Inverse-Gamma}(\alpha, \beta)$ distribution has expectation $\frac{\beta}{\alpha-1}$ and variance $\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$.

The Geometric distribution

If $x \in \{1, 2, 3, \ldots\}$ has a Geometric distribution with parameter $p \in (0, 1)$, the probability mass function is

$$\pi(x \mid p) = p(1-p)^{x-1}$$

We write $x \mid p \sim \text{Geometric}(p)$ and $\pi(x \mid p) = \text{Geometric}(x; p)$. The expectation is $1/p$ and the variance $(1-p)/p^2$.

The Negative Binomial distribution

A stochastic variable x taking on as possible values any nonnegative integer has a Negative Binomial distribution if its probability mass function is given by

$$\pi(x \mid r, p) = \binom{x+r-1}{x} \cdot (1-p)^x p^r = \frac{\Gamma(x+r)}{\Gamma(x+1)\Gamma(r)} (1-p)^x p^r$$

where $r > 0$ and $p \in (0, 1)$ are parameters. We write $x \sim \text{Negative-Binomial}(r, p)$ and $\pi(x \mid r, p) = \text{Negative-Binomial}(x; r, p)$.

The Normal distribution

If the real x has a Normal distribution with parameters μ and σ^2, its density is given by

$$\pi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

We write $x \mid \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$ and $\pi(x \mid \mu, \sigma^2) = \text{Normal}(x; \mu, \sigma^2)$.

The Poisson distribution

If $x \in \{0, 1, 2, \ldots\}$ has Poisson distribution with parameter $\lambda > 0$ then the probability mass function is

$$e^{-\lambda} \frac{\lambda^x}{x!}.$$

We write $x \mid \lambda \sim \text{Poisson}(\lambda)$ and $\pi(x \mid \lambda) = \text{Poisson}(x; \lambda)$. The Poisson distribution has expectation λ and variance λ.

Petter Mostad
Applied Mathematics and Statistics
Chalmers

Suggested solutions for
MVE550 Stochastic Processes and Bayesian Inference
Exam April 16, 2025

1. (a) If we let P_1, P_2 and P_3 denote the three rows of the transition matrix, the prior indicated corresponds to using an independent Dirichlet(1, 1, 1) prior for each P_i. Because of the conjugacy between the Dirichlet and the Multinomial, the posterior for P_1 becomes Dirichlet(5, 4, 1), and the probability to move form state 1 to state 2 becomes $4/(5 + 4 + 1) = 0.4$.

   (b) Similar to the above calculation, the probability that it stays in the first state will be $5/10 = 0.5$. However, when computing the probability that it then jumps to state 2, one needs to use counts updated with the count that it has stayed once in the first state. In other words, one should then use the posterior Dirichlet(6, 4, 1) for predictions. So the answer becomes
$$0.5 \cdot \frac{4}{11} = \frac{2}{11}.$$

   (c) The changed prior implies using Dirichlet(1, 1, 0) for the prior for P_1. The answer to (a) thus becomes $4/9$, and the answer to (b) becomes
$$\frac{5}{9} \cdot \frac{4}{10} = \frac{2}{9}.$$

2. (a) State A is transient, while all the other states are recurrent (i.e., not transient). State A has period ∞, while all the other states have period 1.

   (b) To be ergodic, the chain needs to be irreducible. But the given chain has two communication classes, and is thus not irreducible. So it is not ergodic.

   (c) A limiting distribution needs to be independent of the starting point. However, as we have two communication classes, the state will always depend on which class we start in, so there is no limiting distribution.

   (d) Even if there is no limiting distribution, there may be stationary distributions. The simplest example of a stationary distribution is $(0, 1, 0, 0, 0)$, i.e., a distribution where one is in state B with probability 1. Considering the Figure in the question, this will then also be the case for the following step, and this is a stationary distribution. But there are also a number of other stationsry distributions.

3. (a) Writing X_i for independent copies of X, we can write
$$\mathrm{E}[Z_n] = \mathrm{E}\left[\sum_{i=1}^{Z_{n-1}} X_i\right] = \mathrm{E}\left[\mathrm{E}\left[\sum_{i=1}^{Z_{n-1}} X_i \mid Z_{n-1}\right]\right] = \mathrm{E}[Z_{n-1} \cdot \mathrm{E}[X_i]] = \mathrm{E}[Z_{n-1}]\,\mathrm{E}[X].$$

(b) We can write

$$\begin{aligned}
\operatorname{Var}[Z_n] &= \operatorname{Var}\left[\sum_{i=1}^{Z_{n-1}} X_i\right] \\
&= \operatorname{E}\left[\operatorname{Var}\left[\sum_{i=1}^{Z_{n-1}} X_i \mid Z_{n-1}\right]\right] + \operatorname{Var}\left[\operatorname{E}\left[\sum_{i=1}^{Z_{n-1}} X_i \mid Z_{n-1}\right]\right] \\
&= \operatorname{E}[Z_{n-1} \operatorname{Var}[X]] + \operatorname{Var}[Z_{n-1} \operatorname{E}[X]] \\
&= \operatorname{E}[Z_{n-1}] \operatorname{Var}[X] + \operatorname{E}[X]^2 \operatorname{Var}[Z_{n-1}]
\end{aligned}$$

(c) We can write

$$e_n = \operatorname{E}[Z_n = 0] = \operatorname{E}[\operatorname{E}[Z_n = 0 \mid Z_1]] = \operatorname{E}\left[e_{n-1}^{Z_1}\right] = \operatorname{E}\left[e_{n-1}^{X}\right] = G_X(e_{n-1})$$

(d) As

$$G_X(s) = \operatorname{E}\left[s^X\right] = \sum_{k=0}^{\infty} s^k \Pr(X = k)$$

we have

$$G_X'(s) = \sum_{k=1}^{\infty} k s^{k-1} \Pr(X = k)$$

and

$$G_X''(s) = \sum_{k=2}^{\infty} k(k-1) s^{k-2} \Pr(X = k)$$

As the sum is nonnegative, we get that $G_X''(s) \geq 0$ for $s \in [0, 1]$, and the probability generating function is convex.

Alternatively, one can simply observe that $G_X(s)$ can be written as a sum of convex functions: All the terms in the infinite sum are convex for non-negative s.

4. (a) The algorithm starts with choosing a starting value (x_0, y_0, z_0). Then, for $i = 1, 2, \ldots, n$,

   i. Simulate (x^*, y^*, z^*) according to the density $q(x^*, y^*, z^* \mid x_{i-1}, y_{i-1}, z_{i-1})$.

   ii. Compute

$$\rho = \min\left(1, \frac{f(x^*, y^*, z^*) q(x_{i-1}, y_{i-1}, z_{i-1} \mid x^*, y^*, z^*)}{f(x_{i-1}, y_{i-1}, z_{i-1}) q(x^*, y^*, z^* \mid x_{i-1}, y_{i-1}, z_{i-1})}\right).$$

   iii. With probability ρ set $(x_i, y_i, z_i) = (x^*, y^*, z^*)$, otherwise set $(x_i, y_i, z_i) = (x_{i-1}, y_{i-1}, z_{i-1})$.

(b) One can use that, for the density proportional to f,

$$\operatorname{E}[x > y] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} I(x_i > y_i).$$

where I is the indicator function, and thus compute

$$\operatorname{E}[x > y] \approx \frac{1}{n} \sum_{i=1}^{n} I(x_i > y_i).$$

5. (a) It is $\frac{5}{5+2} = \frac{5}{7} = 0.7143$.

   (b) The probability that the first customer is female is $\frac{2}{7}$. That the first three customers are female corresponds to this happening (independently) three times. So the probability is $\left(\frac{2}{7}\right)^3 = \frac{8}{343} = 0.0233$.

   (c) It is Gamma(5, 7).

   (d) The number of customers during the first 30 minutes is Poisson distributed with parameter $7/2$, so the probability that there are no customers in this period is $\exp(-7/2) = 0.030$.

   (e) As exactly 5 customers arrive during the first hour, we know that their arrival times are independent and uniform over this hour. Thus the probability that there is exactly one customer during the first 20 minutes can be computed as

   $$\text{Binomial}(1; 5, 1/3).$$

   Each customer has a probability 5/7 of beming male, so the answer is

   $$\text{Binomial}(1; 5, 1/3) \cdot \frac{5}{7} = 5 \cdot (1/3)^1 (2/3)^4 \cdot \frac{5}{7} = \frac{400}{1701} = 0.2351558$$

6. We may compute

$$\begin{aligned} \frac{d}{dt}e^{tA} &= \frac{d}{dt}\sum_{k=0}^{\infty}\frac{1}{k!}(tA)^k \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\frac{d}{dt}t^k A^k \\ &= \sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}A^k \\ &= A\sum_{k=1}^{\infty}\frac{1}{(k-1)!}(tA)^{k-1} \\ &= Ae^{tA} \end{aligned}$$

7. The derivative of $P(t)$ when $t = 0$ is Q, the infinitesimal generator matrix. As all the elements of Q are non-zero, the Markov chain is irreducible, and thus ergodic, and it has a unique limiting distribution. To find this limiting distribution, we need to find the vector v such that $vQ = 0$ and such that the elements of v sum to 1. The equation $vQ = 0$ resolves to the three equations

$$\begin{aligned} v_1 &= \frac{v_2 + v_3}{2} \\ v_2 &= \frac{v_3 + v_1}{2} \\ v_3 &= \frac{v_1 + v_2}{2} \end{aligned}$$

One may solve these equations simultaneously with the equation $v_1 + v_2 + v_3 = 1$, or one may simply observe that, by symmetry, all values of the v vector must be equal, leading to the answer

$$v = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

as the limiting distribution.

8. We may transform the expression as follows:

$$\begin{aligned} & B_2 + 5B_3 - 3B_6 \\ = \quad & B_2 + 5(B_3 - B_2) + 5B_2 - 3(B_6 - B_3) - 3(B_3 - B_2) - 3B_2 \\ = \quad & 3B_2 + 2(B_3 - B_2) - 3(B_6 - B_3). \end{aligned}$$

This is a sum of three independent normally distributed variables, and is thus a normally distributed variable. Its expectation is the sum of the expectations of the terms, which is zero. The variance can be computed as

$$\begin{aligned} & \text{Var}\,[3B_2 + 2(B_3 - B_2) - 3(B_6 - B_3)] \\ = \quad & \text{Var}\,[3B_2] + \text{Var}\,[2(B_3 - B_2)] + \text{Var}\,[3(B_3 - B_3)] \\ = \quad & 9\,\text{Var}\,[B_2] + 4\,\text{Var}\,[B_1] + 9\,\text{Var}\,[B_3] \\ = \quad & 9 \cdot 2 + 4 \cdot 1 + 9 \cdot 3 \\ = \quad & 49 \end{aligned}$$

Thus the answer is Normal$(0, 49)$.