Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE550 Stochastic Processes and Bayesian Inference

Re-exam April 8, 2020, 8:30 - 12:30
Examiner: Petter Mostad, phone 031-772-3579
Allowed aids: All aids are allowed.
For example you may access teaching material on any format and you may use R for computation.
However, you are **not** allowed to communicate with any person other than the examiner and the exam guard.
Total number of points: 30. To pass, at least 12 points are needed.
There is an appendix containing relevant information about some probability distributions.

1. (4 points) Assume the variables $y_1, y_2, \dots, y_k, \dots$ each can have possible values 0 or 1. Assume the parameter p is uniformly distributed on the interval from 0 to 1. Assume that given p, the y_i are independent, with the probablity p of being 1. Assume you have made the observations $y_1 = 1, y_2 = 0, y_3 = 0, y_4 = 1, y_5 = 0$.

   (a) What is the posterior for p given the data?

   (b) Given the information above, what is the probability that $y_6 = 1$?

2. (3 points) A discrete-time Markov chain has states A, B, C, D, and transition matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

   If the chain starts in state C, what is the probability that it is absorbed in state A? (Show the steps in the computation).

3. (4 points) Let $Z_0, Z_1, \dots$ be a branching process where the offspring distribution has probability generating function

$$G(s) = \left(\frac{3+s}{4}\right)^2 e^{(s-1)/3}.$$

   (a) Is the process critical, subcritical, or supercritical? Prove your answer.

   (b) Find the variance of the offspring process.

   (c) Find the extinction probability of the branching process.

4. (6 points) We will consider sequences $S = (x_1, \ldots, x_{50})$ of 50 integers so that $x_1 = 0$, $x_{50} = 0$, and $|x_i - x_{i+1}| \leq 1$ for all $i = 1, \ldots, 49$. In other words, the sequences change with at most 1 at each step. An example is illustrated with

```
0  0  0 -1 -2 -2 -2 -3 -3 -2 -2 -1 -1  0  0  0  1  2  1  2  3  3  2  3  2
2  3  3  4  3  4  4  5  4  4  5  6  5  6  5  5  4  3  3  2  3  2  1  0  0
```

Let $\mathcal{A}$ be the set of all sequences of the type above. For each $S \in \mathcal{A}$ there is a largest value $L(S) = \max_{i=1,\ldots,50} x_i$ that the integers reach; for the sequence above, $L(S) = 6$. We would like to estimate the average of $L(S)$ over all sequences S in the set $\mathcal{A}$.

   (a) Two such sequences are called *neighbours* if they are identical except at a single position in the sequence. Define a graph where the nodes are the possible sequences and an edge connects two sequences if they are neighbours. If you make a random walk on this graph, will the stationary distribution be uniform? Why or why not?

   (b) Assume you have computer code which for any valid sequence S computes $F(S)$, the number of neighbour sequences, and another function $G(S)$ which randomly selects a neighbour sequence uniformly among the neighbours. Write down R code (or pseudo code) which generates a Markov chain of sequences whose stationary distribution is the uniform distribution. Also write down any computation you need to make in order to derive any formulas you use in your code.

   (c) Describe how you can use the code above to estimate[1] the average of $L(S)$ over all sequences S in the set $\mathcal{A}$.

5. (6 points) Anders runs a food truck selling hotdogs and hamburgers. A hotdog meal takes on average 2 minutes to prepare while a hamburger meal on average takes 3 minutes to prepare. 60% of customers choose the hamburger meal. We assume that the preparation times are exponentially distributed. Customers arrive according to a Poisson process with an intensity of 1 person every four minutes. If Anders is already working on an order when a customer arrives, the new customer waits. However, if there is already another customer waiting, the new customer goes away.

   (a) Describe the five states that Ander's food truck can be in. Write down the generator matrix for the corresponding Markov chain.

   (b) Compute the expected proportion of time Anders will be spending tending to customers. You may do this using R. An alternative is to just describe in detail how one can do such a computation.

   (c) Among the customers that do not turn away, what is the average waiting time?

[1] This would not be the most computationally efficient way to get an estimate

6. (4 points) Consider a discrete time Markov chain with the possible states $1, 2, 3, 4$. Assume it has been observed for 30 transitions, and the resulting counts of transitions between states is given in the following table (with the rows listing the state before the transition and the columns listing the state after the transition):

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

   (a) Let P be the unknown transition matrix for the chain. Use as a prior for P the uniform distribution on the set of matrices P where all entries are non-negative and all rows sum to 1. Compute the expectation of the posterior of P given the data above.

   (b) Assume now we get the new information that the Markov chain can only change its value with +1, 0, or -1 in each step. Change the prior distribution to reflect this, so that it is a product of Dirichlet distributions with pseudo counts 1, but restricted to the entries in the rows of P that could be non-zero. Compute the expectation of the posterior of P using this prior and using the data above.

7. (3 points) A discrete-time Markov chain has states 0 and 1 and starts at 1. The probability that it will be 0 after exactly 3 steps is *at least* 0.3, and the probability that it will be 0 after exactly 5 steps will be *at least* 0.5. Find the probability p so that no Markov chain fulfilling the above has a probability below p of being 0 after exactly 8 steps, and give the transition matrix of a Markov chain fulfilling the above and having probability p of being 0 after exactly 8 steps.

Appendix: Some probability distributions

The Bernoulli distribution

If $x \in \{0, 1\}$ has a Bernoulli distribution with parameter $0 \leq p \leq 1$, then the probability mass function is

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

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

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 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 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 $\pi(x \mid \lambda) = \text{Poisson}(x; \lambda)$.