Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MVE550 Stochastic Processes and Bayesian Inference

Re-exam August 22, 2022, 8:30 - 12:30
Examiner: Petter Mostad, phone 031-772-3579, visits 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

1. (6 points) Assume you have observed the values $x_1 = 2.3, x_2 = 5.1, x_3 = 7.9$ and you believe they are sampled from a Normal$(\mu, 1/\tau)$ distribution. You have some information about the parameters μ and τ; assume first that you know that $\mu = 4$ but are uncertain about τ.

   (a) Write down the likelihood function for τ. Find a probability density function of τ that is proportional to the likelihood.

   (b) You would like to compute the probability that your next observed value will be larger than 8. Describe in detail the steps in the Bayesian way of making such a computation. Describe and make the assumptions you need to make. You do not need to compute the actual probability, just describe how to compute it in mathematical detail and/or with R code.

   (c) Now assume that instead of knowing that $\mu = 4$, your prior information is a uniform distribution on the interval $[2, 6]$ for μ, and a uniform distribution on the interval $[0.1, 10]$ for τ. There are now several ways of (approximately) computing the posterior probability that $\mu > 4$; describe one of them.

2. (6 points)

   (a) Let $Z_0, Z_1, Z_2, \ldots$ be a branching process, so that $Z_n = \sum_{i=1}^{Z_{n-1}} X_i$, where the X_i are independent copies of a random variable X. If

$$\begin{aligned} \Pr[X = 0] &= 0.1 \\ \Pr[X = 1] &= 0.5 \\ \Pr[X = 3] &= 0.4 \end{aligned}$$

   find the probability of extinction.

   (b) Assume instead that the offspring processes are different in even and odd generations, so that

$$\begin{aligned} Z_n &= \sum_{i=1}^{Z_{n-1}} X_i \qquad n \text{ even} \\ Z_n &= \sum_{i=1}^{Z_{n-1}} Y_i \qquad n \text{ odd} \end{aligned}$$

where the Y_i are independent copies of a random variable Y. Define

$$W = \sum_{i=1}^{Y} X_i.$$

Find and prove a relationship between the probability generating functions $G_X(s), G_Y(s), G_W(s)$ of X, Y, W, respectively.

(c) Assume $G_Y(s) = \exp(s-1)$. Describe the steps in a numerical way to compute the extinction probability for the process in part (b).

3. (6 points)

(a) For a counting process $\{N_t\}_{t\geq 0}$ to be a Poisson process with parameter λ, we must have $N_0 = 0$ and $N_t \sim \text{Poisson}(t\lambda)$ for all $t \geq 0$. Precisely describe two additional properties so that if N_t has these properties it must be a Poisson process.

(b) Using the definition above, write down a proof that if $\{N_t\}_{t\geq 0}$ is a Poisson process and T is the arrival time of the first event, then T has an exponential distribution.

(c) Assume that customers arrive at a carnival stand as a Poisson process with parameter λ. Each customer has a probability 0.01 of winning a grand price, a probability 0.1 of winning a smaller price, and a probability 0.89 of not winning. Find and simplify the formula for the following:

Given that no grand prices are won during the first hour of operation, and the first grand price is won before the end of the second hour, find the probability that the second grand price is also won before the end of the second hour.

4. (6 points) A continuous-time Markov chain has states 1,2,3,4,5. The expected holding times for these states are 3,2,1,1,2, respectively. The *embedded chain* has transition matrix

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

Let $v = (3/7, 2/7, 1/14, 1/14, 1/7)$.

(a) Prove that v is a stationary distribution for the continuous-time Markov chain.

(b) Prove that v is *not* a limiting distribution for the continuous-time Markov chain.

(c) Find $\lim_{t\to\infty} \Pr[N_t = 3 \mid N_0 = 4]$.

5. (6 points) Each round in a game works as follows: You first pay 1 kroner. Then with a probability 1/10 you win 10 kroner and with a probability 9/10 you win nothing. Let X_i denote your total winnings or losses after i rounds.

   (a) Compute the expectation and variance of X_i.

   (b) Write $Y_i = aX_i$ for some $a > 0$, and find the a such that the variance of Y_i is i for all i.

   (c) What does the Donsker invariance principle say about the behaviour of Y_i when i is large?

   (d) Use the above to find an (approximate) value m so that with 95% probability the total winnings will never go above m during 10000 played rounds. (You may use that a variable with a standard normal distribution is in the interval $[-1.96, 1.96]$ with 95% probability).

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 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.

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 and GU

Suggested solutions to
MVE550 Stochastic Processes and Bayesian Inference

Re-exam August 22, 2022, 8:30 - 12:30

1. (a) We get

$$\begin{aligned}
\pi(\text{data} \mid \tau) &= \prod_{i=1}^{3} \text{Normal}(x_i; 4, 1/\tau) \\
&= \left(\frac{1}{\sqrt{2\pi/\tau}}\right)^3 \exp\left(-\frac{\tau}{2}\left((x_1-4)^2 + (x_2-4)^2 + (x_3-4)^2\right)\right) \\
&\propto_\tau \tau^{3/2} \exp\left(-\frac{\tau}{2}\left(1.7^2 + 1.1^2 + 3.9^2\right)\right) = \tau^{3/2}\exp(-9.655\tau).
\end{aligned}$$

This means that the likelihood $\pi(\text{data} \mid \tau)$ is proportional to the probability density $\text{Gamma}(\tau; 5/2, 9.655)$.

(b) To compute this probability you need to find the posterior predictive probability. To find this, you first need to find a posterior for τ, and this means you need to assume some prior. A choice corresponding with the computation in (a) is to choose a flat prior on the positive real values as a prior: With such a prior, the posterior becomes $\text{Gamma}(\tau; 5/2, 9.655)$.

The posterior predictive then becomes

$$\pi(x \mid \text{data}) = \int_0^\infty \text{Normal}(x; 4, 1/\tau)\,\text{Gamma}(\tau; 5/2, 9.655)\,d\tau$$

and the required probability can be computed as

$$\int_8^\infty \pi(x \mid \text{data})\,dx.$$

(One may provide more detail in several ways: One is to write

$$\begin{aligned}
&\int_8^\infty \int_0^\infty \text{Normal}(x; 4, 1/\tau)\,\text{Gamma}(\tau; 5/2, 9.655)\,d\tau\,dx \\
= &\int_0^\infty \left[\int_8^\infty \text{Normal}(x; 4, 1/\tau)\,d\tau\right] \text{Gamma}(\tau; 5/2, 9.655)\,dx
\end{aligned}$$

and note that this can be computed in R as a numerical integral of

```
(1-pnorm(8,4,1/sqrt(tau)))*dgamma(tau, 5/2, 9.655).
```

Another is to compute

$$\begin{aligned}
\pi(x \mid \text{data}) &= \int_0^\infty \frac{1}{\sqrt{2\pi/\tau}} \exp\left(-\frac{\tau}{2}(x-4)^2\right) \frac{9.655^{5/2}}{\Gamma(5/2)} \tau^{3/2} \exp(-9.655\tau)\, d\tau \\
&= \frac{9.655^{5/2}}{\sqrt{2\pi}\Gamma(5/2)} \int_0^\infty \tau^{3-1} \exp\left(-\tau(9.655 + (x-4)^2/2)\right) d\tau \\
&= \frac{9.655^{5/2}}{\sqrt{2\pi}\Gamma(5/2)} \cdot \frac{\Gamma(3)}{(9.655 + (x-4)^2/2)^3}.
\end{aligned}$$

A third is to express this integral as a non-centered t-distribution.)

(c) A simple procedure is to use gridding: Make a uniform 2D grid for μ in the interval $[2, 6]$ and τ in the interval $[0.1, 10]$, for example with 100 grid points in each direction, for a total of 10000 grid points. Then compute the likelihood function

$$\pi(\text{data} \mid \mu, \tau) = \prod_{i=1}^{3} \text{Normal}(x_i; \mu, 1/\tau)$$

in each grid point, and normalize so that the values sum to 1. Then compute the sum at the grid points where $\mu > 4$.

2. (a) We get for the probability generating function

$$G_X(s) = 0.1 + 0.5s + 0.4s^2$$

and then

$$\begin{aligned}
G_X(s) - s &= 0.1 \cdot (1 + 5s + 4s^2) - s \\
&= 0.1 \cdot (1 - 5s + 4s^2) \\
&= 0.1 \cdot (s-1)(4s^2 + 4s - 1) \\
&= 0.1 \cdot (s-1) \cdot 4 \cdot (s + 1/2 + \sqrt{2}/2)(s + 1/2 - \sqrt{2}/2)
\end{aligned}$$

We see from this that the smallest positive root of $G_X(s) = s$, and thus the probability of extinction, is $\sqrt{2}/2 - 1/2$.

(b) We get

$$\begin{aligned}
G_W(s) = \mathrm{E}(s^W) = \mathrm{E}(\mathrm{E}(s^W \mid Y)) = \mathrm{E}\left(\mathrm{E}\left(s^{\sum_{i=1}^Y X_i} \mid Y\right)\right) \\
= \mathrm{E}\left(\mathrm{E}\left(\prod_{i=1}^{Y} s^{X_i} \mid Y\right)\right) = \mathrm{E}\left(\prod_{i=1}^{Y} \mathrm{E}(s^{X_i})\right) = \mathrm{E}\left(G_X(s)^Y\right) = G_Y(G_X(s))
\end{aligned}$$

(c) By considering two consecutive generations as one generation, we see that the branching process can be viewed as a standard branching process with offspring process given by W. We also have

$$G_W(s) = G_Y(G_X(s)) = G_Y(0.1 + 0.5s + 0.4s^3) = \exp(0.5s + 0.4s^3 - 0.9).$$

To find the smallest positive root of $G_W(s)$ we can apply for example the R function `uniroot` to

$$f(s) = \exp(0.5s + 0.4s^3 - 0.9) - s$$

on the interval $[0, 1]$.

3. (a)
   - Stationary increments: For all $s, t > 0$ $N_{t+s} - N_s$ has the same distribution as N_t.
   - Independent increments: For $0 \leq q < r \leq s < t$, $N_t - N_s$ and $N_r - N_q$ are independent.

   (b) For any $t > 0$ we have that

   $$\Pr[T > t] = \Pr[N_t = 0] = e^{-t\lambda},$$

   using the probability mass function for the Poisson. Thus

   $$\Pr[T \leq t] = 1 - e^{-t\lambda}$$

   and taking derivative we get for the probability density for T

   $$\pi(T) = \lambda e^{-t\lambda}.$$

   Comparing with the density for the exponential distribution, we get $T \sim \text{Exponential}(\lambda)$.

   (c) As the winning of grand prices is a Poisson process and such processes have stationary increments, we can ignore the first hour and start the Poisson process at the start of the second hour. The required probability is the probability of two or more grand prices during this hour divided by the probability of one or more grand prices during this hour. If $(N_t)_{t\geq 0}$ is the Poisson process for grand prices, this can be computed as

   $$\frac{\Pr[N_1 \geq 2]}{\Pr[N_1 \geq 1]} = \frac{1 - \Pr[N_1 = 0] - \Pr[N_1 = 1]}{1 - \Pr[N_1 = 0]} = \frac{1 - e^{-0.01\lambda}(1 + 0.01\lambda)}{1 - e^{-0.01\lambda}},$$

4. (a) From the expected holding times we get that $(q_1, q_2, \ldots, q_5) = (1/3, 1/2, 1, 1, 1/2)$. Using $\tilde{P}$ we can now compute the generator matrix as

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

We then get

$$vQ = \left(-\frac{1}{3}\frac{3}{7}+\frac{1}{2}\frac{2}{7}, \frac{1}{3}\frac{3}{7}-\frac{1}{2}\frac{2}{7}, -\frac{1}{14}+\frac{1}{2}\frac{1}{14}+\frac{1}{4}\frac{1}{7}, \frac{1}{2}\frac{1}{14}-\frac{1}{14}+\frac{1}{4}\frac{1}{7}, \frac{1}{2}\frac{1}{14}+\frac{1}{2}\frac{1}{14}-\frac{1}{2}\frac{1}{7}\right) = 0$$

proving that v is a stationary distribution.

(b) v cannot be a limiting distribution, as this Markov chain has no limiting distribution. The reason is that it is reducible, it has the two closed communication classes $\{1, 2\}$ and $\{3, 4, 5\}$. The most direct proof that the chain does not have a limiting distribution is to observe that the state when $t \to \infty$ depends on the starting state: It cannot move out of the communication class it starts in.

(c) When $N_0 = 3$, we know that the chain starts in the second communication class. Restricting the Markov chain to this class, it has generator matrix

$$Q' = \begin{bmatrix} -1 & 1/2 & 1/2 \\ 1/ & -1 & 1/2 \\ 1/4 & 1/4 & -1/2 \end{bmatrix}.$$

We have seen above that $(1/14, 1/14, 1/7)Q' = 0$. Normalizing so that this vector is a probability vector, we get that $v' = 14/4 \cdot (1/14, 1/14, 1/7) = (1/4, 1/4, 1/2)$ is the unique limiting distribution for the restricted chain. Thus

$$\lim_{t\to\infty} \Pr[N_t = 3 \mid N_0 = 4] = 1/4.$$

5. (a) Let Z denote the outcome of a single round. Then

$$\mathrm{E}(X_i) = i\,\mathrm{E}(Z) = i\left(9 \cdot \frac{1}{10} - 1 \cdot \frac{9}{10}\right) = 0$$

and

$$\mathrm{Var}(X_i) = i\,\mathrm{Var}(Z) = i(\mathrm{E}(Z^2) - \mathrm{E}(Z)^2) = i\,\mathrm{E}(Z^2) = i\left(9^2 \cdot \frac{1}{10} + 1 \cdot \frac{9}{10}\right) = 9i.$$

(b) We get $\mathrm{Var}(Y_i) = a^2\,\mathrm{Var}(X_i) = a^2 9i$, so setting $a = 1/3$ will lead to $\mathrm{Var}(Y_i) = i$.

(c) When i is large, Y_i behaves approximately like Brownian motion B_t with $t = i$.

(d) The maximum value of a Brownian motion on the interval $[0, 10000]$ can be written M_{10000} where we know from theory that M_{10000} has the same distribution as $|B_{10000}|$. But B_{10000} is normally distributed with expectation 0 and variance 10000, i.e., standard deviation $\sqrt{10000} = 100$. Using the hint, we know that B_{10000} is in the interval $[-196, 196]$ with 95% probability, so that

$$\Pr[M_{10000} < 196] = \Pr[|B_{10000}| < 196] = 0.95.$$

Thus, approximately, the maximum value of Y_i is below 196 with 95% probability during 10000 played rounds, so that the maximum value of X_i is below $3 \cdot 196 = 588$ with 95% probability during 10000 played rounds.