58/60.

Stat 515 Final, Penn State Statistics
May 4, 2015.

NAME Wanjun Lin

1. (a) A certain town never has two sunny days in a row. Each day is classified as either being sunny, cloudy (but dry), or rainy. If it is sunny one day, then it is equally likely to be either cloudy or rainy the next day. If it is rainy one day, then there is a 50% chance it will be the same the next day and there is a 25% chance each that it will be cloudy or sunny the next day. If it is cloudy one day, then there is a 50% chance that it will be the same the next day and there is a 25% chance it will be rainy or sunny the next day. In the long run, what proportion of days are sunny and what proportion of days are cloudy? [3pts]

   (b) Consider a Markov chain with transition matrix with $p_1, p_2 \in (0,1)$:

$$P = \begin{bmatrix} 0 & p_1 & p_1 \\ p_2 & p_1 & p_2 \\ p_2 & p_2 & p_1 \end{bmatrix}$$

   Does this Markov chain have a limiting distribution (you do not have to find this limiting distribution)? If so, carefully show how the conditions for this theoretical result are satisfied. If not, state any conditions that are violated. [4pts]

1. (a) $\Omega = \{1 = \text{sunny},\ 2 = \text{cloudy},\ 3 = \text{rainy}\}$

t.p.m $P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$, Let $\underset{\sim}{\pi} = (\pi_1, \pi_2, \pi_3)$.

$\underset{\sim}{\pi} P = \underset{\sim}{\pi} \Rightarrow \pi_1 = \frac{1}{5},\ \pi_2 = \frac{2}{5},\ \pi_3 = \frac{2}{5}$

thus proportion of sunny days is $\frac{1}{5}$, proportion of cloudy days is $\frac{2}{5}$.

(b)
- Any two different state are communicable, this M.C. is irreducible.
- $P_{11}^2 > 0$, $P_{11}^3 > 0$, $P_{22} > 0$, $P_{33} > 0 \Rightarrow d(1) = d(2) = d(3) = 1$, thus MC is aperiodic.
- Since we have finite states, all states are recurrent and thus positive recurrent.

$\Rightarrow$ This MC has a limiting distribution.

10

2. Suppose n lightbulbs are placed in a room and switched on at time 0. Assume the lifetime of each bulb is independently distributed according to Exponential(θ) with $\theta > 0$, i.e., they have expected lifetimes of θ. Suppose you walk into the room at some time $\tau > 0$ and the only information you have is the number of bulbs still working, W.

(a) Suppose the time you walk into the room is random, with $\tau \sim U(a, b)$, $b > a > 0$. What is the expected value of W ? [3pts]

(b) Now assume τ is fixed (but the rest of the description of the experiment stays the same as above.) What is the expected *total* lifetime of all the bulbs that are still working at time τ? [3pts]

(a) $$E(W) = E(E(W|\tau)) = E(ne^{-\frac{\tau}{\theta}}) = \int_a^b \frac{1}{b-a} ne^{-\frac{\tau}{\theta}} d\tau = \frac{n\theta}{b-a}\left(e^{-\frac{a}{\theta}} - e^{-\frac{b}{\theta}}\right).$$

(b) For fixed τ, the expected # of bulbs still working at time τ is $ne^{-\frac{\tau}{\theta}}$

By memorylessness,

The expected total time of all the bulbs still working at time τ

is $n\theta e^{-\frac{\tau}{\theta}}$. $(+\tau)$

-1

-1

Wanjun Liu

3. Consider a (time-homogeneous) Poisson process $\{N(t),\, t > 0\}$ with $\lambda = 5$.

   (a) What is the probability of seeing 10 events in the time interval (0,2) *and* 5 events in the time interval (1,2) ? [3pts]

   (b) Suppose you know 10 events occurred in the interval (0,2). What is the distribution of the occurrence of the very first event? [3pts]

   (c) Show that this process satisfies the Markov assumption for continuous-time discrete state space Markov processes. Hint: You can use the connection between Poisson processes and exponential random variables (and properties of exponential random variables) to show this. [3pts]

(a) $P(\text{10 events in } (0,2)) = e^{-10}\frac{10^{10}}{10!}$

$P(\text{5 events in } (1,2)) = e^{-5}\frac{5^5}{5!}$

(b) Given 10 events, the joint distribution of $(S_1, \cdots S_{10})$ is the order statistic of $u(0,2)$.

The CDF of S_1 is $(\frac{x}{2})^{10} I(0<x<2)$.

The PDF of S_i is $5(\frac{x}{2})^9 I(0<x<2)$

Thus the density of the first event is: $5(\frac{x}{2})^9 I(0<x<2)$.

(c) $P(N(t) > n \mid N(u), u\in[0,s])$

$= P(N(t) - N(s) + N(s) > n \mid n(u), u\in[0,s])$

not clear

$= P(N(t) - N(s) > n - N(s) \mid n(u), u\in[0,s])$

$= P(N(t) - N(s) > n - N(s) \mid n(s))$ (by independent increasement)

-1

$= P(N(t) > n \mid n(s))$

∴ Markov assumption is satisfied.

Also the holding time for each state is Exp(5) because the interarrival time between two events is $\sim$ Exp(5).

(-1)

4. Helicopters land at a small hangar (a place for storing aircraft) at the Poisson rate of 20 per day. However, they will only land if there are either 0 or 1 helicopters (including the one being worked on) at the hangar. That is, they will not land if there are 2 or more helicopters at the hangar. The hangar can only work on one helicopter at a time. Assume that the amount of time required to service a helicopter is exponentially distributed with a mean of 2 hours.

   (a) What is the generator matrix (following class notation this is denoted by G) for this continuous-time Markov chain? [3pts]

   12

   (b) Does this process satisfy the detailed balance condition ? Justify your answer. [3pts]

   (c) Every hour without any helicopters results in a loss of \$1,000 for the hangar, while every hour with at least one helicopter results in a profit of \$5,000. In the long run, how much profit can the hangar expect to make per hour? (Note: you do not have to simplify your final answer.) [3pts]

(a) $\Omega = \{0, 1, 2\}$.

Suppose T_0, T_1, T_2 are the holding time for each state.

$T_0 \sim Exp(20)$, $T_1 \sim Exp(32)$, $T_2 \sim Exp(12)$.

$$G = \begin{bmatrix} -20 & 20 & 0 \\ 12 & -32 & 20 \\ 0 & 12 & -12 \end{bmatrix} =: (\gamma_{ij})_{3\times 3}$$

(b) Let π_0, π_1, π_2 satisfy.

$$\begin{cases} 20\pi_0 = 12\pi_1 \\ 20\pi_1 = 12\pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases} \Rightarrow \begin{cases} \pi_0 = \frac{9}{49} \\ \pi_1 = \frac{15}{49} \\ \pi_2 = \frac{25}{49} \end{cases}$$

Thus this process satisfies detailed balance condition.

(c) $\left(\frac{15}{49} + \frac{25}{49}\right) \cdot 5000 - \frac{9}{49} \cdot 1000$

$= \frac{191000}{49}$

Expect to make profit of $\$\frac{191000}{49}$ per hour.

10

Wanjun Liu

5. Let $\{X(t), t \geq 0\}$ be standard Brownian motion. That is, $X(0) = 0$, every increment $X(s+t) - X(s)$ is $N(0, t)$, and for every set of n disjoint time intervals, the increments are independent random variates.

   (a) Show that $\{X(t), t \geq 0\}$ is an example of a continuous-time continuous-state-space Markov process. [4pts]

   (b) Show that the stochastic process $\{Y(t), t \geq 0\}$ with $Y(t) = X^2(t) - t$ is a martingale. [4pts]

   (c) Define $\{Z(t), t \geq 0\}$ as $Z(t) = \exp(\lambda X(t) - \lambda^2 t/2)$, λ is a constant. $\{Z(t), t \geq 0\}$ is known to be a martingale (you do not have to prove this). Let T be the first time that $X(t)$ reaches 2-4t, that is, $T = \min\{t : X(t) = 2 - 4t\}$. What is $E(T)$? Fully justify your answer. [4pts]

(a). $P(X(t) > x \mid X(u), u \in [0, s])$

$= P(X(t) - X(s) + X(s) > x \mid X(u), u \in [0,s])$

$= P(X(t) - X(s) > x - X(s) \mid X(u), u \in [0,s])$

$= P(X(t) - X(s) > x - X(s) \mid X(s))$ (by independent increasement)

$= P(X(t) > x \mid X(s))$

Also time and space state are continuous

$\Rightarrow$ This is a continuous-time and continuous state space M.C.

(b) $E(Y(t) \mid X(u), u \in [0,s])$

$= E(X^2(t) - t \mid X(u). u \in [0,s])$

$= -t + E([X(t) - X(s) + X(s)]^2 \mid X(s))$

$= -t + E\left((X(t) - X(s))^2 + X^2(s) + 2X(s)(X(t) - X(s)) \mid X(s)\right)$

$= -t + (t - s) + X^2(s)$

$= X^2(s) - s = Y(s)$

$E|Y(t)|$

$\leq E|X^2(t)| + t$

$= 2t < \infty$

$\Rightarrow \{Y(t), t \geq 0\}$ is a mg.

(c). $E(Z_0) = \exp(0) = 1$

$X(t) = 2 - 4t$

$\Leftrightarrow Z(t) = \exp(\lambda(2 - 4t) - \lambda^2 t/2)$

$E(Z_T) = E(\exp(\lambda X(T) - \lambda^2 t/2))$

$= 1$

10

6. Consider a regression of a variable Y on X where the regression model is as follows, $Y_i \sim EMG(\beta_0 + \beta_1 X, \sigma, \lambda)$, where the exponentially modified Gaussian random variable, $\text{EMG}(\mu, \sigma, \lambda)$, has pdf $f(x; \mu, \sigma, \lambda) = \frac{\lambda}{2} \exp(\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x))\text{erfc}\left(\frac{\mu+\lambda\sigma^2-x}{\sqrt{2}\sigma}\right)$, and erfc is the complementary error function defined as

$$\text{erfc}(x) = \frac{2}{\pi} \int_x^{\infty} e^{-t^2} dt.$$

Assume that σ is known. Let the independent priors for $\beta_0, \beta_1, \lambda$ be $p(\beta_0), p(\beta_1), p(\lambda)$ respectively.

(a) Provide pseudocode for a Metropolis-Hastings algorithm to construct a Markov chain with stationary distribution $\pi(\beta_0, \beta_1, \lambda \mid \mathbf{Y}, \mathbf{X})$ for a data set of size n, $(X_1, Y_1), \ldots, (X_n, Y_n)$. $\mathbf{Y}, \mathbf{X}$ are $(Y_1, \ldots, Y_n)$, $(X_1, \ldots, X_n)$ respectively. You should provide enough detail so anyone reading it should be able to write code based on your description. You do not have to provide starting values or specific tuning parameters since those will depend on the particulars of the data. You should, however, list at the beginning of the algorithm any/all tuning parameters for the algorithm that you will have to adjust in order to make it work well. [5pts]

(b) Suppose your Markov chain is $(\beta_0^{(1)}, \beta_1^{(1)}, \lambda^{(1)}), \ldots (\beta_0^{(n)}, \beta_1^{(n)}, \lambda^{(n)})$. Provide estimators for (i) $E_\pi(\lambda)$, and (ii) $E_\pi\left(\frac{1}{\beta_1+\lambda}\right)$. [2pts]

(c) State the theoretical result that justifies the use of the above estimator of $E_\pi(\lambda)$. State sufficient conditions for the theorem to hold (you do not have to prove that these conditions hold). [3pts]

(d) You are worried about the influence of the priors on the posterior so you would like to see how the posterior changes if the prior is modified to $p^*(\beta_0), p^*(\beta_1), p^*(\lambda)$, independent of each other. Describe in detail how you would use the Markov chain above (do not construct a new Metropolis-Hastings algorithm) to approximate $E_{\pi^*}(\lambda)$ where π^* is the new posterior pdf. Briefly explain when your approach is likely to work well and when it will not. [4pts]

(e) Now suppose that σ is also assumed to be unknown, and has prior $p(\sigma)$. Describe all the ways in which your MCMC algorithm from part (a) will change when you now have to approximate expectations with respect to $\pi(\beta_0, \beta_1, \lambda, \sigma \mid \mathbf{Y}, \mathbf{X})$. [3pts]

Wanjun Liu

6(a) ~~full~~ Let full conditional dist'n for $\beta_0, \beta_1, \lambda$ be $h_0(\beta_0)$, $h_1(\beta_1)$, $h_2(\lambda)$

$$h_0(\beta_0) \propto \prod_{i=1}^{n} f(y_i; \beta_0+\beta_1 x_i, \sigma, \lambda) \cdot p(\beta_0)$$

$$h_1(\beta_1) \propto \prod_{i=1}^{n} f(y_i; \beta_0+\beta_1 x_i, \sigma, \lambda) \cdot p(\beta_1)$$

$$h_2(\lambda) \propto \prod_{i=1}^{n} f(y_i; \beta_0+\beta_1 x_i, \sigma, \lambda) \cdot p(\lambda)$$

Let g_0, g_1, g_2 be the proper proposals for $\beta_0, \beta_1, \lambda$ with tuning parameters τ_0, τ_1, τ_2

Initilization: sample size N, starting values $\beta_0^{(1)}, \beta_1^{(1)}, \lambda^{(1)}, \tau_1, \tau_2, \tau_3$.

For $i=2:N$

draw $y_0 \sim g_0$

$$P_0 = \min\left(1, \frac{h_0(y_0)}{h_0(\beta_0^{(i-1)})} \cdot \frac{g_0(y_0, \beta_0^{(i-1)})}{g_0(\beta_0^{(i-1)}, y_0)}\right) \quad (\text{with } \beta_1^{(i-1)}, \lambda^{(i-1)}).$$

accept $\beta_0^{(i)} \leftarrow y_0$ with prob p_0, otherwise $\beta_0^{(i)} \leftarrow \beta_0^{(i-1)}$.

draw $y_1 \sim g_1$

$$P_1 = \min\left(1, \frac{h_1(y_1)}{h_1(\beta_1^{(i-1)})} \cdot \frac{g_1(y_1, \beta_1^{(i-1)})}{g_1(\beta_1^{(i-1)}, y_1)}\right) \quad (\text{with } \beta_0^{(i)}, \lambda^{(i-1)})$$

accept $\beta_1^{(i)} \leftarrow y_1$ with prob p_1, otherwise $\beta_1^{(i)} \leftarrow \beta_1^{(i-1)}$

draw $y_2 \sim g_2$

$$P_2 = \min\left(1, \frac{h_2(y_2)}{h_2(\lambda^{(i-1)})} \cdot \frac{g_2(y_2, \lambda^{(i-1)})}{g_1(\lambda^{(i-1)}, y_2)}\right) \quad (\text{with } \beta_0^{(i)}, \beta_1^{(i)})$$

accept ~~β~~ $\lambda^{(i)} \leftarrow y_2$ with prob p_2, otherwise $\lambda^{(i)} \leftarrow \lambda^{(i-1)}$.

End for

Return $(\beta_0^{(1)} \cdots \beta_0^{(N)}), (\beta_1^{(1)}, \cdots, \beta_1^{(N)}), (\lambda^{(1)} \cdots, \lambda^{(N)})$.

(b). $E_\pi(\lambda)$: $\sum_{i=1}^{n} \lambda^{(i)} / n$

$E_\pi\left(\frac{1}{\beta_1+\lambda}\right)$: $\sum_{i=1}^{n} \frac{1}{\beta_1^{(i)}+\lambda^{(i)}} / n$

(c) Theoretical result ① $\pi(y) = \int_\Omega K(x,y)\, \pi(x)\, dx$.

② $$\frac{\sum_{i=1}^{n} g(X^{(i)})}{n} \rightarrow E_\pi\, g(X) \text{ a.s.}$$

sufficient conditions: Harris-positive, irreducible and aperiodic

10

(d) Using importance sampling.

$$E_{\pi^*}(\lambda) \approx \sum_{i=1}^{n} \frac{\pi^*(\lambda^{(i)})}{\pi(\lambda^{(i)})} \Big/ n.$$

If π^* and π match well, then it is likely to work well.

If π^* and π do not match well, it will not.

(e). 1 Propose a new proposal g_3 with tuning parameter τ_3 for σ.

2. Starting value $\sigma^{(i)}$ and full conditional density $h_3(\sigma)$

3. After updating λ (in part (a)).

draw $y_3 \sim g_3$

$$p_3 = \min\left(1, \frac{h_3(y_3)}{h_3(\sigma^{(i-1)})} \cdot \frac{g_3(y_3, \sigma^{(i-1)})}{g_3(\sigma^{(i-1)}, y_3)}\right) \text{ with } (\beta_0^{(i)}, \beta_1^{(i)}, \lambda^{(i)})$$

Accept $\sigma^{(i)} \leftarrow y_3$ with prob p_3, otherwise $\sigma^{(i)} \leftarrow \sigma^{(i-1)}$.

Ben Sheng

57/60

Stat 515 Final, Penn State Statistics
May 4, 2015.

NAME

1. (a) A certain town never has two sunny days in a row. Each day is classified as either being sunny, cloudy (but dry), or rainy. If it is sunny one day, then it is equally likely to be either cloudy or rainy the next day. If it is rainy one day, then there is a 50% chance it will be the same the next day and there is a 25% chance each that it will be cloudy or sunny the next day. If it is cloudy one day, then there is a 50% chance that it will be the same the next day and there is a 25% chance it will be rainy or sunny the next day. In the long run, what proportion of days are sunny and what proportion of days are cloudy? [3pts]

   (b) Consider a Markov chain with transition matrix with $p_1, p_2 \in (0,1)$:

$$P = \begin{bmatrix} 0 & p_1 & p_1 \\ p_2 & p_1 & p_2 \\ p_2 & p_2 & p_1 \end{bmatrix}$$

   Does this Markov chain have a limiting distribution (you do not have to find this limiting distribution)? If so, carefully show how the conditions for this theoretical result are satisfied. If not, state any conditions that are violated. [4pts]

(a)
$$P = \begin{array}{c|ccc} & S & C & R \\ S & 0 & .5 & .5 \\ C & .25 & .5 & .25 \\ R & .25 & .25 & .5 \end{array}$$

Solve $\pi P = \pi$

$\pi_S = .25\pi_C + .25\pi_R$

$\pi_C = .5\pi_S + .5\pi_C + .25\pi_R$

$\Rightarrow .5\pi_C = .5\pi_S + .25\pi_R$

$\frac{1}{2.5} = \frac{10}{25} = \frac{2}{5}$

$\pi_S = .25\pi_R + .25\pi_R = .5\pi_R$

$\Rightarrow \pi_C = \pi_S + .5\pi_R = .25\pi_C + .25\pi_R + .5\pi_R = .25\pi_C + .75\pi_R$

$1 = \pi_S + \pi_C + \pi_R = .5\pi_R + \pi_R + \pi_R = 2.5\pi_R$

$\Rightarrow .75\pi_C = .75\pi_R \Rightarrow \pi_C = \pi_R$

$\Rightarrow \pi_R = \frac{2}{5}, \pi_C = \frac{2}{5}, \pi_S = \frac{1}{5}$. $\boxed{\frac{1}{5} \text{ are sunny and } \frac{2}{5} \text{ are cloudy}}$

(b) Positive probability of going from one state to another, so irreducible.
There are self-loops and periodicity is a class property, so aperiodic.
Finite and irreducible, so positive recurrent.

By Ergodic Thm, Markov Chain has a limiting distribution.

10

Ben Sheng

2. Suppose n lightbulbs are placed in a room and switched on at time 0. Assume the lifetime of each bulb is independently distributed according to Exponential(θ) with $\theta > 0$, i.e., they have expected lifetimes of θ. Suppose you walk into the room at some time $\tau > 0$ and the only information you have is the number of bulbs still working, W.

   (a) Suppose the time you walk into the room is random, with $\tau \sim U(a, b)$, $b > a > 0$. What is the expected value of W ? [3pts]

   (b) Now assume τ is fixed (but the rest of the description of the experiment stays the same as above.) What is the expected *total* lifetime of all the bulbs that are still working at time τ? [3pts]

use this parametrization!

(a) $\mathbb{E}[W] = \mathbb{E}[\mathbb{E}[W|\tau]]$

$$W|\tau = \sum_{i=1}^{n} I(L_i > \tau), \quad \text{where } L_i \sim Exp(\theta)$$

$$\mathbb{E}[W|\tau] = \sum_{i=1}^{n} \mathbb{P}(L_i > \tau) = \sum_{i=1}^{n} e^{-\theta\tau} = ne^{-\theta\tau}$$

$$\mathbb{E}[W] = \mathbb{E}[\mathbb{E}[W|\tau]] = \int_a^b ne^{-\theta x} \cdot \frac{1}{b-a}\, dx = \frac{n}{b-a}\int_a^b e^{-\theta x}\, dx$$

$$= \frac{n}{b-a}\left[\frac{e^{-\theta x}}{-\theta}\right]_a^b = \frac{n}{b-a}\left[\frac{e^{-\theta x}}{\theta}\right]_b^a = \boxed{\frac{n}{\theta(b-a)}\left[e^{-\theta a} - e^{-\theta b}\right]}$$

(b) By memoryless property, each bulb that is still working has expected remaining lifetime of $\frac{1}{\theta}$ at time τ.

Thus, the expected total lifetime is

$$(\tau + \tfrac{1}{\theta})\mathbb{E}[W] = \boxed{\frac{n}{\theta^2(b-a)}\left[e^{-\theta a} - e^{-\theta b}\right]}$$

-1.

-1

Ben Sheng

3. Consider a (time-homogeneous) Poisson process $\{N(t),\, t > 0\}$ with $\lambda = 5$.

   (a) What is the probability of seeing 10 events in the time interval (0,2) *and* 5 events in the time interval (1,2) ? [3pts]

   (b) Suppose you know 10 events occurred in the interval (0,2). What is the distribution of the occurrence of the very first event? [3pts]

   (c) Show that this process satisfies the Markov assumption for continuous-time discrete state space Markov processes. Hint: You can use the connection between Poisson processes and exponential random variables (and properties of exponential random variables) to show this. [3pts]

(a) $\mathbb{P}(10 \text{ events in } (0,2), 5 \text{ events in } (1,2)) = \mathbb{P}(5 \text{ events in } (0,1), 5 \text{ events in } (1,2))$

$= \mathbb{P}(5 \text{ events in } (0,1)) \cdot \mathbb{P}(5 \text{ events in } (1,2))$, by $\perp\!\!\!\perp$ increments

$= (\mathbb{P}(5 \text{ events in } (0,1)))^2$, by stationary increments

$= \left(\frac{e^{-\lambda t}(\lambda t)^n}{n!}\right)^2 = \boxed{\left(\frac{e^{-5}(5)^5}{5!}\right)^2}$ ✓

(b) The minimum of 10 order statistics with distribution Unif$(0,2)$.

(c) $\mathbb{P}(N(s+t) = i+1 \mid N(s) = i, N(u), u \in (0,s))$

$= \mathbb{P}(0 < T_{i+1} < t \mid N(s) = i, N(u), u \in (0,s))$,

where $T_i \sim Exp(\lambda)$ is the ith interarrival time

$= \mathbb{P}(0 < T_{i+1} < t \mid N(s) = i)$, by the memoryless property } skipped step -0.5

$= \mathbb{P}(N(s+t) = i+1 \mid N(s) = i)$

The larger steps follow from Chapman-Kolmogorov and conditioning.

-0.5

Ben Sheng

4. Helicopters land at a small hangar (a place for storing aircraft) at the Poisson rate of 20 per day. However, they will only land if there are either 0 or 1 helicopters (including the one being worked on) at the hangar. That is, they will not land if there are 2 or more helicopters at the hangar. The hangar can only work on one helicopter at a time. Assume that the amount of time required to service a helicopter is exponentially distributed with a mean of 2 hours.

   (a) What is the generator matrix (following class notation this is denoted by G) for this continuous-time Markov chain? [3pts]

   (b) Does this process satisfy the detailed balance condition ? Justify your answer. [3pts]

   (c) Every hour without any helicopters results in a loss of \$1,000 for the hangar, while every hour with at least one helicopter results in a profit of \$5,000. In the long run, how much profit can the hangar expect to make per hour? (Note: you do not have to simplify your final answer.) [3pts]

(a)

$$P = \begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 0 \\ 1 & 3/8 & 0 & 5/8 \\ 2 & 0 & 1 & 0 \end{array}$$

service $\sim Exp(\lambda = \frac{1}{2})$

arrival $\sim Exp(\lambda = \frac{20}{24} = \frac{5}{6})$

$\frac{1}{2} + \frac{5}{6} = \frac{3}{6} + \frac{5}{6} = \frac{8}{6} = \frac{4}{3}$ is rate of min of service and arrival

$\frac{1}{2} / \frac{4}{3} = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$

$\frac{5}{6} / \frac{4}{3} = \frac{5}{6} \cdot \frac{3}{4} = \frac{5}{8}$

$$G = \begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & -5/6 & 5/6 & 0 \\ 1 & \frac{1}{2} & -\frac{4}{3} & \frac{5}{6} \\ 2 & 0 & \frac{1}{2} & -\frac{1}{2} \end{array}$$

(b) $\pi G = 0 \Rightarrow$

$\pi_0 = -\frac{5}{6}\pi_0 + \frac{1}{2}\pi_1 \Rightarrow \frac{11}{6}\pi_0 = \frac{1}{2}\pi_1 \Rightarrow \pi_0 = \frac{6}{11} \cdot \frac{1}{2}\pi_1 = \frac{3}{11}\pi_1$

$\pi_2 = \frac{5}{6}\pi_1 + -\frac{1}{2}\pi_2 \Rightarrow \frac{3}{2}\pi_2 = \frac{5}{6}\pi_1 \Rightarrow \pi_2 = \frac{\cancel{2}}{3} \cdot \frac{5}{\cancel{6}_3}\pi_1 = \frac{5}{9}\pi_1$

$1 = \pi_0 + \pi_1 + \pi_2 = \frac{3}{11}\pi_1 + \pi_1 + \frac{5}{9}\pi_1 = \pi_1\left(\frac{27}{99} + \frac{99}{99} + \frac{55}{99}\right) = \frac{181}{99}\pi_1$

$\pi_1 = \frac{99}{181}, \quad \pi_0 = \frac{27}{181}, \quad \pi_2 = \frac{55}{181}$

Check $\pi_i \gamma_{ij} = \pi_j \gamma_{ji}$: $\quad \pi_0 \gamma_{01} = \frac{9\cancel{27}}{181} \cdot \frac{5}{2\cancel{6}} = \frac{45}{362}$, but $\pi_1 \gamma_{10} = \frac{99}{181} \cdot \frac{1}{2} = \frac{99}{362}$

$\boxed{\text{Detailed Balance is } \cancel{\text{not}} \text{ satisfied}}$

-1.5

(c) $(\pi_1 + \pi_2)\,5000 - \pi_0\,1000 = 1000\left[5 \cdot \frac{154}{181} - \frac{27}{181}\right] = 1000\left[\frac{770}{181} - \frac{27}{181}\right]$

$= 1000\left[\frac{743}{181}\right] = \boxed{\frac{743000}{181}}$

-1.5

Ben Sheng

5. Let $\{X(t), t \geq 0\}$ be standard Brownian motion. That is, $X(0) = 0$, every increment $X(s+t) - X(s)$ is $\mathrm{N}(0, t)$, and for every set of n disjoint time intervals, the increments are independent random variates.

   (a) Show that $\{X(t), t \geq 0\}$ is an example of a continuous-time continuous-state-space Markov process. [4pts]

   (b) Show that the stochastic process $\{Y(t), t \geq 0\}$ with $Y(t) = X^2(t) - t$ is a martingale. [4pts]

   (c) Define $\{Z(t), t \geq 0\}$ as $Z(t) = \exp(\lambda X(t) - \lambda^2 t/2)$, λ is a constant. $\{Z(t), t \geq 0\}$ is known to be a martingale (you do not have to prove this). Let T be the first time that $X(t)$ reaches 2-4t, that is, $T = \min\{t : X(t) = 2 - 4t\}$. What is $E(T)$? Fully justify your answer. [4pts]

(b) $\mathbb{E}[Y(t) \mid X(u) : u \in [0,s]] = \mathbb{E}[X^2(t) \mid X(u) : u \in [0,s]] - t$

$= \mathbb{E}[X^2(t) \mid X(s)] - t$, by the Markov Property

$= \mathbb{E}[((X(t) - X(s)) + X(s))^2 \mid X(s)] - t$

$= \mathbb{E}[(X(t) - X(s))^2 \mid X(s)] + 2\,\mathbb{E}[(X(t) - X(s))X(s) \mid X(s)]$

$+ \mathbb{E}[X^2(s) \mid X(s)] - t$

$= \mathbb{E}[(X(t) - X(s))^2] + 2\,\mathbb{E}[X(t) - X(s)]X(s) + X^2(s) - t$, by ⫫ increments

$= t - s + 0 + X^2(s) - t = X^2(s) - s = Y(s)$ ✓

(a) Need $\mathbb{P}(X(s+t) \leq y \mid X(s) = x, X(u) = x_u \text{ for } u \in (0,s)) = \mathbb{P}(X(s+t) \leq y \mid X(s) = x)$

$\mathbb{P}(X(s+t) \leq y \mid X(s) = x, X(u) = x_u \text{ for } u \in (0,s))$

$= \mathbb{P}(X(s+t) - X(s) \leq y - x \mid X(s) = x, X(u) = x_u \text{ for } u \in (0,s))$

$= \mathbb{P}(X(s+t) - X(s) \leq y - x \mid X(s) = x)$, by ⫫ increments

$= \mathbb{P}(X(s+t) \leq y \mid X(s) = x)$

(c) $\mathbb{E}[Z(0)] = 1$. By optimal Sampling thm, since T is bounded, $\mathbb{E}[Z(T)] = \mathbb{E}[Z(0)] = 1$.

$\mathbb{E}[Z(T)] = \exp(-\lambda^2 t/2)\,\mathbb{E}[\lambda X(T)]$

0

10

6. Consider a regression of a variable Y on X where the regression model is as follows, $Y_i \sim EMG(\beta_0 + \beta_1 X, \sigma, \lambda)$, where the exponentially modified Gaussian random variable, $\text{EMG}(\mu, \sigma, \lambda)$, has pdf $f(x; \mu, \sigma, \lambda) = \frac{\lambda}{2}\exp(\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x))\text{erfc}\left(\frac{\mu+\lambda\sigma^2-x}{\sqrt{2}\sigma}\right)$, and erfc is the complementary error function defined as

$$\text{erfc}(x) = \frac{2}{\pi}\int_x^\infty e^{-t^2}\,dt.$$

Assume that σ is known. Let the independent priors for $\beta_0, \beta_1, \lambda$ be $p(\beta_0), p(\beta_1), p(\lambda)$ respectively.

(a) Provide pseudocode for a Metropolis-Hastings algorithm to construct a Markov chain with stationary distribution $\pi(\beta_0, \beta_1, \lambda \mid \mathbf{Y}, \mathbf{X})$ for a data set of size n, $(X_1, Y_1), \ldots, (X_n, Y_n)$. $\mathbf{Y}, \mathbf{X}$ are $(Y_1, \ldots, Y_n)$, $(X_1, \ldots, X_n)$ respectively. You should provide enough detail so anyone reading it should be able to write code based on your description. You do not have to provide starting values or specific tuning parameters since those will depend on the particulars of the data. You should, however, list at the beginning of the algorithm any/all tuning parameters for the algorithm that you will have to adjust in order to make it work well. [5pts]

(b) Suppose your Markov chain is $(\beta_0^{(1)}, \beta_1^{(1)}, \lambda^{(1)}), \ldots (\beta_0^{(n)}, \beta_1^{(n)}, \lambda^{(n)})$. Provide estimators for (i) $E_\pi(\lambda)$, and (ii) $E_\pi\left(\frac{1}{\beta_1+\lambda}\right)$. [2pts]

(c) State the theoretical result that justifies the use of the above estimator of $E_\pi(\lambda)$. State sufficient conditions for the theorem to hold (you do not have to prove that these conditions hold). [3pts]

(d) You are worried about the influence of the priors on the posterior so you would like to see how the posterior changes if the prior is modified to $p^*(\beta_0), p^*(\beta_1), p^*(\lambda)$, independent of each other. Describe in detail how you would use the Markov chain above (do not construct a new Metropolis-Hastings algorithm) to approximate $E_{\pi^*}(\lambda)$ where π^* is the new posterior pdf. Briefly explain when your approach is likely to work well and when it will not. [4pts]

(e) Now suppose that σ is also assumed to be unknown, and has prior $p(\sigma)$. Describe all the ways in which your MCMC algorithm from part (a) will change when you now have to approximate expectations with respect to $\pi(\beta_0, \beta_1, \lambda, \sigma \mid \mathbf{Y}, \mathbf{X})$. [3pts]

Ben Sheng

$$\pi(\beta_0,\beta_1,\lambda \mid Y, X) = \prod_{i=1}^{n} f(Y_i;\ \beta_0+\beta_1 X_i, \sigma, \lambda)\cdot p(\beta_0), p(\beta_1), p(\lambda).$$

(a) 1. Choose initial state $(\beta_0^{(1)}, \beta_1^{(1)}, \lambda^{(1)})$.

Choose tuning parameters $\tau_0^2, \tau_1^2, \tau_\lambda^2, \rho_{01}, \rho_{0\lambda}, \rho_{1\lambda}$.

2. Propose candidate $(\beta_0^*, \beta_1^*, \log\lambda^*)$ from the multivariate Normal w/ mean $(\beta_0^{(i)}, \beta_1^{(i)}, \log\lambda^{(i)})$ and

covariance matrix $\Sigma = \begin{bmatrix} \tau_0^2 & \rho_{01}\tau_0\tau_1 & \rho_{0\lambda}\tau_0\tau_\lambda \\ \rho_{01}\tau_0\tau_1 & \tau_1^2 & \rho_{1\lambda}\tau_1\tau_\lambda \\ \rho_{0\lambda}\tau_0\tau_\lambda & \rho_{1\lambda}\tau_1\tau_\lambda & \tau_\lambda^2 \end{bmatrix}$

3. W/ probability $\alpha = \min\left\{1, \dfrac{\pi(\beta_0^*, \beta_1^*, \lambda^* \mid Y, X)}{\pi(\beta_0^{(i)}, \beta_1^{(i)}, \lambda^{(i)} \mid Y, X)}\right\}$,

set $(\beta_0^{(i+1)}, \beta_1^{(i+1)}, \lambda^{(i+1)}) = (\beta_0^*, \beta_1^*, \lambda^*)$.

Otherwise, set $(\beta_0^{(i+1)}, \beta_1^{(i+1)}, \lambda^{(i+1)}) = (\beta_0^{(i)}, \beta_1^{(i)}, \lambda^{(i)})$.

4. Repeat steps 2,3 for desired # of iterations.

(b) (i) $\frac{1}{n}\sum_{i=1}^{n} \lambda^{(i)}$ (ii) $\frac{1}{n}\sum_{i=1}^{n} \frac{1}{\beta_1^{(i)}+\lambda^{(i)}}$

(c) Strong Law of Large Numbers

(i) The proposal $q(x,y) > 0$ over the set of sample values

(ii) For each set of values, $\mathbb{P}(\pi(x)q(x,y) \le \pi(y)q(y,x) <)$.

(iii) Detailed balance satisfied.

We need $\mathbb{E}_f[|g(x)|] < \infty$.

(d) We use importance sampling. We compute $\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} \frac{\lambda^{(i)} f(\beta_0^{(i)}, \beta_1^{(i)}, \lambda^{(i)})}{q(\beta_0^{(i)}, \beta_1^{(i)}, \lambda^{(i)})}$

We use a new density f.

This will work well when the proposal approximates the density well.

This will not work well when we have numerical underflow or overflow issues.

(e) We could do all-at-once M-H w/ 4 variables.

Or, we could do variable-at-a-time or block sampling.

We would need to choose new starting values and tuning parameters.

10