

---

**Stat 515: Stochastic Processes I**  
**Final Examination**

---

**Spring 2012**  
**Solutions**

This final exam is worth 30 points. You have 110 minutes. **For full credit, you must explain all of your work!** Naturally, you may use any results that you know; you should not need to prove anything unless you are explicitly asked to do so.

**Problem 1. [14 points]** Suppose that two telephone operators, Andrew and Barbara, work in an office. Each one is always either on or off the phone. Let us assume that each operator's time on and off the phone may be modeled by an off-on process with generator (or rate) matrices as follows:

$$\text{for Andrew: } R_A = \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \quad \text{for Barbara: } R_B = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$$

In each matrix, the first row/column is for "off the phone" and the second is for "on the phone."

(a) [2 points] Suppose that at time zero, Andrew is off the phone. Let  $N(t)$  equal the total number of transitions (from off to on or on to off) that Andrew makes before time  $t$ . Find the mean and variance of  $N(3)$ .

**Solution:** Because each waiting time until the next event in Andrew's case is exponentially distributed with rate 3, the events form a homogeneous Poisson process with rate 3. Therefore,  $N(t) \sim \text{Poisson}(t)$ . We conclude that  $N(3)$  has mean 9 and variance 9.

(b) [2 points] Let  $T_1$  be the time of Andrew's first transition from off to on. Conditional on  $N(3) = 10$ , what is the distribution of  $T_1$ ?

**Solution:** For a homogeneous Poisson process, we have shown that conditional on  $N(t)$ , the event times are uniformly distributed on  $(0, t)$ . Therefore,  $T_1$  has the same distribution as the minimum of a random sample of size 10 from uniform(0, 3). Its cumulative distribution function, conditional on  $N(3) = 10$ , is

$$P(T_1 < x) = 1 - P(T_1 > x) = 1 - P(\text{all 10 uniforms} > x) = 1 - \left(\frac{3-x}{3}\right)^{10} \quad \text{for } 0 < x < 3.$$

To verify this, if you forgot, use the following:

$$P(T_1 > x \mid N(3) = 10) = \frac{P(10 \text{ events in } (x, 3] \text{ and none in } (0, x])}{P(10 \text{ events in } (0, 3])} = \frac{e^{-3x} e^{-3(3-x)} [3(3-x)]^{10} 10!}{10! e^{-9} 9^{10}}.$$

(c) [2 points] Describe how you would find the probability that Barbara is on the phone at time  $t = 1$  assuming that she is off the phone at  $t = 0$ . (You do not have to actually make this calculation, but you should describe how to do so.)

**Solution:** Using a computer, evaluate the matrix exponential  $\exp\{R_B\}$ , whose (1,2) element equals the desired probability.

(d) [2 points] Consider the Markov chain with states 1 through 4, as follows:

$$1 : \text{A off, B off} \quad 2 : \text{A on, B off} \quad 3 : \text{A off, B on} \quad 4 : \text{A on, B on}$$

Assuming that Andrew's process is independent of Barbara's process, write down the rate matrix  $R$  for this four-state process.

**Solution:** The off-diagonal entries are simply taken from  $R_A$  and  $R_B$ . Then the diagonal entries are adjusted so that each row sums to zero:

$$R = \begin{bmatrix} -4 & 3 & 1 & 0 \\ 3 & -4 & 0 & 1 \\ 2 & 0 & -5 & 3 \\ 0 & 2 & 3 & -5 \end{bmatrix}$$

(e) [2 points] Prove that the Markov chain described by the rate matrix  $R$  in part (d) satisfies detailed balance, i.e., that the chain is time-reversible. (You can prove this even if you do not get the correct form of  $R$ .)

**Solution:** Each separate off-on process is time-reversible since it is a birth-death process. Therefore, if multiple processes are run simultaneously and independently, the accumulation of all of these processes must also be time-reversible. However, this line of reasoning is a bit abstract; it may be easier to prove this fact by first answering part (f), then verifying quickly that  $\pi_i r_{ij} = \pi_j r_{ji}$  for all  $i \neq j$ .

(f) [2 points] Find the stationary distribution  $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)^\top$  of the Markov chain described by the rate matrix in part (d).

**Solution:** By standard off-on process results from class, we know that in the long run, Andrew spends 1/2 his time in the off state and Barbara spends 2/3 of her time in the off state. By independence, this means that both are in the off state 1/3 of the time. Similar reasoning yields  $\boldsymbol{\pi} = (1/3, 1/3, 1/6, 1/6)^\top$ .

(g) [2 points] Let  $X_t$ ,  $t = 1, 2, \dots$ , be the  $t$ th state visited by the Markov chain described by the rate matrix in part (d). Then the  $X_t$  describe a discrete-time, discrete-state-space Markov chain in which  $X_t$  is never equal to  $X_{t+1}$ . Derive the probability transition matrix,  $P$ , for the  $\{X_t\}$  Markov chain and then explain whether the  $\{X_t\}$  Markov chain is ergodic.

**Solution:** Each off-diagonal entry  $P_{ij}$  is equal to  $-r_{ij}/r_{ii}$ , which gives

$$P = \begin{bmatrix} 0 & 3/4 & 1/4 & 0 \\ 3/4 & 0 & 0 & 1/4 \\ 2/5 & 0 & 0 & 3/5 \\ 0 & 2/5 & 3/5 & 0 \end{bmatrix}$$

Checking ergodicity means checking aperiodicity and positive recurrence. In this case, the chain is periodic, since the number of people on the phone can only change from odd to even or even to odd. Another way to say this is that the chain oscillates back and forth between sets of states  $\{1, 4\}$  and  $\{2, 3\}$ . Thus, the chain is not ergodic.

**Problem 2. [2 points]** Suppose that  $X \sim \text{gamma}(2.4, 2)$ , so that  $X$  has density function

$$f(x) = 0.15253x^{1.4} \exp\{-x/2\}, \quad x > 0.$$

Explain how to use an i.i.d. sample  $Y_1, \dots, Y_n$  from a standard exponential distribution to construct a 95% confidence interval for  $P(X > 2)$  using importance sampling. The standard exponential density function is

$$f(y) = e^{-y}, \quad y > 0.$$

**Solution:** Letting  $f$  denote the gamma density, we wish to find  $E_f g(X)$  for  $g(x) = I(x > 2)$ . Importance sampling rewrites this as  $E_f g(X) = E_q g(X) f(X)/q(X)$  for some density  $q$ . Here, we let  $q$  be the standard exponential density. Thus, if  $Y_1, \dots, Y_n \sim q$ , we take as the point estimate

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n 0.15253 I(Y_i > 2) Y_i^{1.4} \exp\{Y_i/2\}.$$

Because the  $Y_i$  are i.i.d., we can easily estimate the variance of  $\hat{\mu}$ :

$$\widehat{\text{Var}}(\hat{\mu}) = \frac{\hat{\sigma}^2}{n} = \frac{1}{n(n-1)} \sum_{i=1}^n (0.15253 I(Y_i > 2) Y_i^{1.4} \exp\{Y_i/2\} - \hat{\mu})^2.$$

Then the approximate confidence interval is

$$\hat{\mu} \pm 1.96 \sqrt{\frac{\hat{\sigma}^2}{n}}$$

**Problem 3 [8 points]** Suppose that a joint posterior density for  $(\lambda, \theta)$  is given by

$$p(\lambda, \theta) \propto \lambda^2 \theta \exp\{-\theta\lambda - 3\theta - 2\lambda\}.$$

**(a) [3 points]** Using the fact that a  $\text{gamma}(\alpha, \beta)$  density has the form

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\{-x/\beta\} \quad \text{for } x > 0,$$

explain how to implement Gibbs sampling to construct a Markov chain with stationary distribution  $p(\lambda, \theta)$ .

**Solution:** The full conditionals are

$$\begin{aligned} p(\lambda \mid \theta) &\propto \lambda^2 \exp\left\{-\lambda/\left(\frac{1}{2+\theta}\right)\right\}, \\ p(\theta \mid \lambda) &\propto \theta \exp\left\{-\theta/\left(\frac{1}{3+\lambda}\right)\right\}. \end{aligned}$$

These densities are  $\text{gamma}[3, 1/(2+\theta)]$  and  $\text{gamma}[2, 1/(3+\lambda)]$ , respectively. Therefore, starting with some arbitrarily chosen positive parameter values  $\lambda_0$  and  $\theta_0$ , the Gibbs sampler alternates between choosing  $\lambda_t \sim \text{gamma}[3, 1/(2+\theta_{t-1})]$  and choosing  $\theta_t \sim \text{gamma}[2, 1/(3+\lambda_t)]$  for  $t = 1, 2, \dots$

**(b) [3 points]** Explain one way to implement an “all-at-once” Metropolis-Hastings algorithm to construct a Markov chain with stationary distribution  $p(\lambda, \theta)$ .

**Solution:** We need some kind of bivariate proposal density that will give us a  $(\lambda^*, \theta^*)$  if we are currently at  $(\lambda_t, \theta_t)$ . Perhaps the simplest way to do this is using a symmetric proposal density, such as a bivariate normal centered at  $(\lambda_t, \theta_t)$ , since this will ensure that the M-H acceptance probability only depends on the ratio  $p(\lambda^*, \theta^*)/p(\lambda_t, \theta_t)$ . Things are even easier if we take the bivariate normal proposal to have zero covariance. So the chain operates as follows:

Start with some arbitrarily chosen positive parameter values  $\lambda_0$  and  $\theta_0$  along with a constant  $\tau^2$ . Then, for  $t = 1, 2, \dots$ , let  $\lambda^* \sim N(\lambda_{t-1}, \tau^2)$  and  $\theta^* \sim N(\theta_{t-1}, \tau^2)$ . Take  $(\lambda_t, \theta_t) = (\lambda^*, \theta^*)$  with probability

$$\min \left\{ 1, \frac{I\{\lambda^* > 0\}I\{\theta^* > 0\}(\lambda^*)^2\theta^* \exp\{-\theta^*\lambda^* - 3\theta^* - 2\lambda^*\}}{\lambda_{t-1}^2\theta_{t-1} \exp\{-\theta_{t-1}\lambda_{t-1} - 3\theta_{t-1} - 2\lambda_{t-1}\}} \right\}.$$

otherwise, take  $(\lambda_t, \theta_t) = (\lambda_{t-1}, \theta_{t-1})$ . (We need the indicators in the fraction above because the joint density is zero whenever  $\lambda \leq 0$  or  $\theta \leq 0$ .)

(c) [2 points] If the Markov chain you obtained in part (a) is  $(\lambda_1, \theta_1), (\lambda_2, \theta_2), \dots$ , explain how you could find an approximate 95% confidence interval for  $E_p(\theta)$ .

**Solution:** First, let

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \theta_i.$$

Then, find an estimate of the variance of  $\hat{\mu}$  using a technique such as batch means. Then, the approximate CI would be  $\hat{\mu} \pm 1.96\sqrt{\widehat{\text{Var}}(\hat{\mu})}$ . (To use batch means, we'd take batches of size roughly  $\sqrt{n}$ , find their sample means, then find the sample variance of those sample means. Multiplying the resulting variance by  $1/\sqrt{n}$  gives a reasonable  $\widehat{\text{Var}}(\hat{\mu})$ .)

**Problem 4. [6 points]** We wish to simulate an i.i.d. sample from the distribution with cumulative distribution function  $F(x) = x^{5/3}$ , and therefore density function  $5x^{2/3}/3$ , for  $0 < x < 1$ .

(a) [3 points] Suppose that  $U \sim \text{uniform}(0, 1)$ . Tell how to use the inversion method to obtain a random variable  $X$  with  $X \sim F$  and prove that your  $X$  has the required distribution.

**Solution:** If we let  $X = U^{3/5}$ , then

$$P(X \leq x) = P(U^{3/5} \leq x) = P(U \leq x^{5/3}) = x^{5/3},$$

as desired.

(b) [3 points] Suppose  $U_1$  and  $U_2$  are i.i.d.  $\text{uniform}(0, 1)$  random variables. Derive a function  $h(x)$  such that conditional on  $U_1 < h(U_2)$ ,  $U_2$  has the distribution function  $F$ . (This is implementing a rejection method.) Construct your  $h(x)$  so that it gives the smallest possible probability of rejecting a proposed  $X$ , and calculate this probability.

**Solution:** Since the uniform density, say  $u(x)$ , is a constant 1, here  $h(x)$  should be taken to be  $f(x)/[Ku(x)]$  for some constant  $K$  that ensures the ratio is always less than or equal to 1. Since the maximum value of  $f(x)$ , taken at  $x = 1$ , equals  $5/3$ , we see that the smallest possible value of  $K$  is  $5/3$ . Putting all of this together, we obtain

$$h(x) = x^{2/3}.$$

This gives a rejection probability of  $1 - 1/K$ , which is  $2/5$ .