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:

for Andrew:
$$R_A = \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix}$$
 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).

(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 ?

(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.)

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

1: A off, B off 2: A on, B off 3: A off, B on 4: 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.

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

(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).

(g) [2 points] Let X_t , t = 1, 2, ..., be the tth 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.

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, \ldots, 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

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

Problem 3 [8 points] Suppose that a joint posterior density for (λ, θ) 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 gamma(α, β) density has the form

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

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

- (b) [3 points] Explain one way to implement an "all-at-once" Metropolis-Hastings algorithm to construct a Markov chain with stationary distribution $p(\lambda, \theta)$.
- (c) [2 points] If the Markov chain you obtained in part (a) is $(\lambda_1, \theta_1), (\lambda_2, \theta_2), \ldots$, explain how you could find an approximate 95% confidence interval for $E_p(\theta)$.

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.
- (b) [3 points] Suppose U_1 and U_2 are i.i.d. 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.