Makeup Stat 515 Final, Penn State Statistics
June 1, 2015.

NAME

1. (a) Consider three urns, one colored red, one white, and one blue. The red urn contains 1 red and 4 blue balls; the white urn contains 3 white balls, 2 red balls, and 2 blue balls; the blue urn contains 4 white balls, 3 red balls, and 2 blue balls. At the initial stage, a ball is randomly selected from the red urn and then returned to that urn. At every subsequent stage, a ball is randomly selected from the urn whose color is the same as that of the ball previously selected and is then returned to that urn. Write the state space of this process and the associate initial distribution and transition probability matrix. Explain why this process is a Markov chain. Let X_0 be the initial state of the Markov chain. What is $E(X_2)$ if red, white and blue are coded as 1,2,3 respectively? [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]

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]

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 (0,4) *and* 5 events in the time interval (1,4) ? [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) A doctor has scheduled two appointments, one at 1pm and the other at 1:30pm. Amount of time that appointments last are independent exponential random variables with mean 30 minutes. Assuming that both patients are on time, find the expected amount of time that the 1:30 appointment spends at the doctor's office. [3pts]

4. Helicopters land at a small hangar (a place for storing aircraft) at the Poisson rate of 50 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]

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) Let $\{X(t), t > 0\}$ be standard Brownian motion. Prove that the process $\{M(t), t > 0\}$ where $M(t) = \exp\left(\lambda X(t) - \frac{1}{2}\lambda^2 t\right)$, is a martingale.

   (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]

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]