STAT 753 EXAM - SPRING 2020

Name: **Score:**

INSTRUCTIONS: Due on **Tuesday March 24** at the beginning of lecture. Please show all work and justify your claims. You must not collaborate with others. Submit a PDF file (preferably using R Markdown, but other digital formats are acceptable), and include this sheet as the coversheet. The exam is worth 100 points total.

1. (25 pts) At all times an urn contains N balls - some green balls and some white balls. At each stage a coin having probability $p \in (0, 1)$ of landing heads is tossed. If it lands heads up, then a ball is chosen at random from the urn and is replaced by a green ball; if tails appears, then a ball is chosen at random from the urn and is replaced by a white ball. Let X_n denote the number of green balls in the urn after the n^{th} stage.

   (a) Explain why $\{X_n, n \geq 0\}$ is a Markov chain.

   (b) What are the states of this Markov chain?

   (c) Draw a diagram of this Markov chain for the case $N = 4$.

   (d) State the transition probability matrix P. (Hint: make sure the rows sum to 1.)

   (e) What are the classes of states? What are their periods? Are they transient or recurrent?

   (f) Explain why a unique stationary distribution exists for this Markov chain, and compute the stationary distribution for the case $N = 4$ and $p = 1/3$ (Use R).

   (g) Now suppose that $p = 1$. What is the transition probability matrix and what are the state classes in this case?

   (h) **BONUS:** Let W be the time until there are only green balls in the urn, given that $X_0 = 1$. Compute $E[W]$ for the case $p = 1$.

2. (10 pts) Give an example of a discrete-time Markov chain (DTMC) that does not have a unique stationary distribution.

   (a) Clearly define the DTMC in terms of the state space S, time index T, and the transition matrix P.

   (b) Discuss what prevents the DTMC from having a unique stationary distribution.

3. (25 pts) Write a function in R to simulate a branching process. Start with 1 individual at generation 0 and suppose the offspring distribution is given by the binomial distribution with parameters $n = 6$ and $p = 0.2$. Let X_n denote the number of individuals in the n^{th} generation for $n = 0, 1, 2, \ldots$.

   (a) Discuss whether this population will eventually die out or persist (in theory). Plot a sample path of the process for 40 generations. Include R code and the figure.

   (b) Run the function 100 times, each time stopping at generation 40. Make a vector of population sizes at generation 40 and plot the sizes (y-axis) as a function of the 100 samples (x-axis). Compute the mean of the 100 population sizes at generation 40.

   (c) What proportion of your 100 branching process simulations died out? Compute the mean of the 100 population sizes for those simulations that did not die out.

4. (20 pts) Write a function in R to simulate a hypoexponential random variable Y which has PDF
$$f_Y(y) = C_{1,2} f_{X_1}(y) + C_{2,2} f_{X_2}(y)$$
where $X_1 \sim \text{exponential}(\lambda_1)$, $X_2 \sim \text{exponential}(\lambda_2)$, and $C_{i,2} = \dfrac{\lambda_j}{\lambda_j - \lambda_i}$ for $j \neq i \in \{1, 2\}$. Generate a histogram of 10,000 samples for the case $\lambda_1 = 1$ and $\lambda_2 = 5$. Include R code and the histogram. Briefly discuss how this histogram differs from the two single exponential histograms (corresponding to X_1 and X_2).

5. (20 pts) Modify the Metropolis-Hastings MCMC R code given in class to sample from the hypoexponential distribution described in Problem 4 using a standard normal proposal distribution. Include R code and a figure consisting of two plots: time series and histogram for a sample size of $n = 50,000$. Discuss your observations (e.g. Does the Markov chain seem to converge to the target stationary distribution? Compare the shape of the histogram to the one from Problem 4).

Note about figures: Please include appropriate labels (x-axis, y-axis, title) and use breaks=50 to show sufficient detail.