Probability Qualifying Review Exam
Friday May 24, 2013
9a.m. to 1p.m.

Instructions:

- You are allowed only pen (or pencil) and paper. Calculators, cell phones, and any other electronic device should not be brought into the exam room.
- You may answer up to seven of the eight questions. Each question carries equal credit. If you submit solutions to all eight questions then only the first seven will be graded.
- Questions 1-4 cover material from Stat 620. Questions 5-8 cover material from Stat 621.

1. This question studies a random variable with a random probability mass function, constructed as follows. Let $Z_1, \dots, Z_N$ be independent, identically distributed, positive random variables. Let

$$W_n = \frac{Z_n}{\sum_{m=1}^{N} Z_m} \qquad \text{for } n = 1, \dots N.$$

Let $x_1, \dots, x_N$ be a sequence of real numbers. Let X be a random variable defined such that $\mathbb{P}(X = x_n \,|\, W_1, \dots, W_N) = W_n$.

(a) Find $\mathbb{E}[W_n]$ and $\mathbb{E}[W_n \,|\, W_m]$ for $m \neq n$. Hint: you can appeal to symmetry to argue that the answer should not depend on n or m.

(b) Let $\mathrm{Var}(W_n) = \alpha$. Find $\mathrm{Cov}(W_m, W_n)$ in terms of α, for $m \neq n$.

(c) Find an expression for $\mathbb{E}[\mathrm{Var}(X \,|\, W_1, \dots, W_N)]$ in terms of $x_1, \dots, x_N$ and α. Check whether your expression behaves appropriately in two situations; (i) when $\alpha = 0$, and (ii) when $x_1, x_2, \dots, x_N$ is replaced by $x_1 + c, x_2 + c, \dots, x_N + c$ for some constant c.

2. Let $\{X(t), 0 \leq t \leq T\}$ be a Brownian motion with initial distribution $X(0) \sim \mathrm{Normal}(\nu, \tau^2)$. In other words, $\{X(t)\}$ has continuous sample paths and stationary independent increments having distribution $X(t+s) - X(t) \sim \mathrm{Normal}(0, \sigma^2 s)$ for $s > 0$. Let $Y(t) = X(T-t)$ for $0 \leq t \leq T$.

(a) Explain why $\{Y(t), 0 \leq t \leq T\}$ is a diffusion process.

(b) Obtain the infinitesimal parameters of $\{Y(t)\}$.

(c) In the special case with $\nu = 0$, $\tau = 0$ and $T = 1$, comment on how the behavior of $\{Y(t), 0 \leq t \leq T\}$ relates to a well-known diffusion process.

3. We wish to show that not every discrete time Markov chain can be embedded in a continuous time Markov chain. To formalize this, let $\{X_n, n = 0, 1, 2, \dots\}$ be a discrete time Markov chain with states $\{1, 2, \dots, K\}$ having transition probabilites P_{ij}, and let $\{X(t), t \geq 0\}$ be a continuous time Markov chain with states $\{1, 2, \dots, K\}$ having transition rates q_{ij}. We say that $\{X_n, n = 0, 1, 2, \dots\}$ is embedded in $\{X(t), t \geq 0\}$ if $\{X(n), n = 0, 1, 2, \dots\}$ has the same distribution as $\{X_n, n = 0, 1, 2, \dots\}$.

Specifically, you are asked to consider the case $K = 2$ with $\{X_n\}$ having a symmetric transition probability matrix,

$$P = [P_{ij}] = \begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{pmatrix} \qquad \text{for } 0 < \alpha < 1.$$

For what values of α does an embedding exist?

4. Consider a partially observed discrete stochastic process $\{(X_n, Y_n), n = 1, 2, \dots\}$, where $\{Y_n, n = 1, 2, \dots\}$ is observed and $\{X_n, n = 1, 2, \dots\}$ is unobserved. Suppose that, marginally, $\{X_n\}$ is a Markov chain taking values in a finite set $\mathbb{X}$. Suppose $\{Y_n\}$ takes values in a finite set $\mathbb{Y}$. To model Y_n as a noisy observation of X_n, we suppose $\{Y_n\}$ has a conditional independence property that

$$\mathbb{P}(Y_n = j_n \,|\, X_1 = i_1, \dots, X_n = i_n, Y_1 = j_1, \dots, Y_{n-1} = j_{n-1}) = \mathbb{P}(Y_n = j_n \,|\, X_n = i_n).$$

We are interested in computing the following quantities, concerning the conditional distribution of $\{X_n\}$ given observed values $Y_1 = j_1, Y_2 = j_2, \dots$

$$\begin{aligned} C(i,n) &= \mathbb{P}(X_n = i \,|\, Y_1 = j_1, Y_2 = j_2, \dots, Y_n = j_n) \\ P(i,n) &= \mathbb{P}(X_n = i \,|\, Y_1 = j_1, Y_2 = j_2, \dots, Y_{n-1} = j_{n-1}) \\ S(i,n) &= \mathbb{P}(X_n = i \,|\, Y_1 = j_1, Y_2 = j_2, \dots, Y_N = j_N) \quad \text{for } N > n \end{aligned}$$

Here, $C(i,n)$ is current state estimation given observations up to time n; $P(i,n)$ is prediction of the state at time n using observations up to time $n-1$; $S(i,n)$ is so-called smoothed estimation of the state at time n using observations up to time N.

(a) Prove the following recursions:

$$\begin{aligned} C(i,n) &= \frac{P(i,n)\,\mathbb{P}(Y_n = j_n \,|\, X_n = i)}{\sum_{k\in\mathbb{X}} P(k,n)\,\mathbb{P}(Y_n = j_n \,|\, X_n = k)} \\ P(i,n+1) &= \textstyle\sum_{k\in\mathbb{X}} C(k,n)\,\mathbb{P}(X_n = i \,|\, X_{n-1} = k) \end{aligned}$$

(b) Prove the relationship

$$S(i,n) = \frac{P(i,n)R(i,n)}{\sum_{k\in\mathbb{X}} P(k,n)R(k,n)},$$

where

$$R(i,n) = \mathbb{P}(Y_n = j_n, Y_{n+1} = j_{n+1}, \dots Y_N = j_N \,|\, X_n = i).$$

5. Let $X_n,\ n = 1, 2, \cdots$ be independent, identically distributed random variables.

(a) If $X_1 \in L^p,\ p \geq 1$, then show that

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{L^p} \mathbb{E}(X_1), \qquad \text{as } n \to \infty.$$

(b) If X_n/n converges almost surely to zero, as $n \to \infty$, then show that $\mathbb{E}|X_1| < \infty$.

6. Let $X_n,\ n = 1, 2, \cdots$ be independent random variables such that $X_n \sim \text{Gamma}(\alpha_n, 1),\ \alpha_n > 0$. That is, the probability density of X_n is

$$f_{X_n}(x) = \begin{cases} x^{\alpha_n - 1} e^{-x}/\Gamma(\alpha_n) & , \text{ if } x > 0 \\ 0 & , \text{ otherwise,} \end{cases}$$

where $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1} e^{-x} dx,\ \alpha > 0$ is the Euler gamma function.

(a) If $\sum_{n=1}^\infty \alpha_n < \infty$, then show that the series $\sum_{n=1}^\infty X_n$ converges almost surely and in the L^2 sense.

(b) If $\sum_{n=1}^\infty \alpha_n = \infty$, then show that

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}} \Longrightarrow \text{Normal}(0,1), \qquad \text{as } n \to \infty,$$

where $S_n = \sum_{k=1}^n X_k$.

Hint: Although it would in principle be possible to use a Lindeberg-Feller argument, you are not advised to follow that approach.

7. Let $X_0 = 1$ and define the random variables $X_n,\ n \geq 1$ recursively as follows. Introduce the σ-algebra $\mathcal{F}_n := \sigma(X_0, \cdots, X_n),\ n \geq 0$. Conditionally on $\mathcal{F}_n$, the random variable X_{n+1} is uniformly distributed in the interval $(-S_n, S_n)$, where $S_n = X_0 + \cdots + X_n,\ n \geq 0$.

   (a) Show that $\{S_n,\ n \geq 1\}$ is a martingale and compute $\mathbb{E}(S_n)$ and $\mathrm{Var}(S_n)$.

   (b) Show that $S_n > 0$ almost surely, and that for some positive sequences a_n and b_n, the random variables $S_n^{a_n}/b_n$ converge in distribution to a non-trivial limit, as $n \to \infty$. Identify explicitly the sequences $\{a_n\}$ and $\{b_n\}$, and the limit distribution.

   Hint: Express S_n in terms of independent random variables.

8. Suppose that $Y_k,\ k = 1, 2, \cdots$ are independent and identically distributed random variables such that $\mathbb{E}(Y_1) = 0$ and $\mathrm{Var}(Y_1) = 1$. Let $\{X_k,\ k = 1, 2, \cdots\}$ be a sequence of positive random variables, independent of $\{Y_k, k = 1, 2, \cdots\}$, such that

$$\frac{\max_{k=1,\cdots,n} X_k^2}{\sum_{j=1}^n X_j^2} \xrightarrow{a.s.} 0, \quad \text{as } n \to \infty.$$

   (a) Let

$$Z_n := \frac{\sum_{k=1}^n X_k Y_k}{\sqrt{\sum_{j=1}^n X_j^2}}$$

   and calculate $\mathbb{E}(Z_n)$ and $\mathrm{Var}(Z_n)$.

   (b) Let $\mathcal{F}_X := \sigma(X_k,\ k \geq 1)$ and show that for all bounded and continuous functions f, we have

$$\mathbb{E}(f(Z_n)|\mathcal{F}_X) \longrightarrow \mathbb{E}f(Z), \quad \text{almost surely,}$$

   as $n \to \infty$, where Z is a standard Normal random variable.

   (c) Using part (b), argue that

$$Z_n \Longrightarrow \mathrm{Normal}(0, 1), \quad \text{as } n \to \infty.$$