Ma 551 — Advanced Probability

Problem Set #2 — Due October 30, 2007

Prof. Sawyer — Washington University

Text references are to Kai Lai Chung, *A Course in Probability Theory*, 3rd edition, Academic Press, 2001.

1. Let $X_1, X_2, \ldots$ be i.d.i.r.v. (identically-distributed independent random variables) such that $E(X_1^2) < \infty$ and $E(X_1) = 0$. Prove that

$$\lim_{n\to\infty} X_n/\sqrt{n} \;=\; 0 \quad \text{a.s.}$$

where "a.s." stands for "almost surely". (*Hint*: Use the Borel-Cantelli Lemma somehow.)

2. Let $X(\omega)$, $Y(\omega)$ be random variables on a probability space $(\Omega, \mathcal{F}, P)$ satisfying

$$P(X \le \lambda, Y \le \mu) \;=\; \int_{-\infty}^{\lambda}\int_{-\infty}^{\mu} f(x,y)dxdy$$

for all $\lambda, \mu \in R$ for some measurable function $f(x,y) \ge 0$. Prove that

$$E\big(\phi(Y) \mid \mathcal{B}(X)\big)(\omega) \;=\; \frac{\int_{-\infty}^{\infty} \phi(z) f\big(X(\omega), z\big)dz}{\int_{-\infty}^{\infty} f\big(X(\omega), z\big)dz} \quad \text{a.s.} \tag{1}$$

for any bounded Borel function $\phi(y) \ge 0$. (*Hint*: Find $f_X(x)$ such that $P(X \le \lambda) = \int_{-\infty}^{\lambda} f_X(x)dx$ and use the definition of conditional expectation. Use the fact that a random variable $Z(\omega)$ is $\mathcal{B}(X)$-measurable if and only if $Z(\omega) = g\big(X(\omega)\big)$ for some Borel function $g(y)$.)

In the following problems, let $p_{ij} = p(i,j)$ be an $N \times N$ matrix with

$$\sum_{k=1}^{N} p_{ik} = 1, \quad \text{all } i, \quad \text{and } p_{ij} > 0 \quad \text{all } i,j \tag{2}$$

3. (i) Choose $\nu_i > 0$ with $\sum_{k=1}^N \nu_k = 1$ and let $p(i,j)$ be as above. Prove that there exists a probability space $(\Omega, \mathcal{F}, P)$ and random variables $X_0(\omega), X_1(\omega), \ldots$ on $(\Omega, \mathcal{F})$ such that

$$P(X_0 = i_0, X_1 = i_1, \ldots X_k = i_k) \;=\; \nu_{i_0} p(i_0,i_1)p(i_1,i_2)\ldots p(i_{k-1},i_k) \tag{3}$$

for all k and integers i_a such that $1 \le i_a \le N$ for $0 \le a \le N$.

(ii) Where did your proof of (i) use the fact that $\sum_{i=1}^N \nu_i = 1$? Could you get a stronger result by dropping the condition $\sum_{i=1}^N \nu_i = 1$?

4. (i) Show that the random variables in the last problem satisfy the conditional probability relation

$$P\big(X_{n+1} = j \mid \mathcal{B}(X_1, X_2, \ldots, X_n)\big)(\omega) \;=\; p(X_n(\omega), j) \quad \text{a.s.} \tag{4}$$

A set of random variables $\{ X_n \}$ that satisfies a set of relations like (4), where the right-hand side of (4) depends only on $X_n(\omega)$, is called a *Markov process*, after the Russian mathematician Andrei Markov (1856-1922).

(ii) Show more generally that

$$P(X_{n+m} = j \mid \mathcal{B}(X_1, X_2, \ldots, X_n))(\omega) \;=\; p^m(X_n(\omega), j) \quad \text{a.s.} \tag{5}$$

for all $m \ge 1$, where $p^m(i,j)$ is the m^{th} matrix power of $p(i,j)$ (i.e.,

$$\sum_{j=1}^N p(i,j)p^{m-1}(j,k) = p^m(i,k)$$

etc. *Hint*: Write (3) for $X_0, X_1, \ldots, X_n, X_{n+1}, \ldots, X_{n+m}$ and sum the indices between n and $n+m$.)

Definitions: For $p(i,j)$ as in (2), one can then show that there exists a unique vector $\nu \in R^N$ such that

$$\sum_{i=1}^N \nu_i p_{ij} = \nu_j \quad \text{for } 1 \le j \le N, \quad \nu_j > 0, \quad \sum_{k=1}^N \nu_k = 1 \tag{6}$$

Any vector ν satisfying the first equation in (6) is called a *stationary vector* for $p(i,j)$. One can also show that there exist constants C, λ such that $0 < \lambda < 1$ and

$$|p^m(i,j) - \nu_j| \;\le\; C\lambda^m, \quad \text{for all } i, j, m \ge 1 \tag{7}$$

where ν is the unique stationary vector in (6). In general, a sequence of random variables $X_1, X_2, \ldots$ is called *strongly mixing* if there exist constants $\rho_m > 0$ satisfying $\sum_{m=0}^\infty \rho_m < \infty$ and

$$\big|P(A \cap B) - P(A)P(B)\big| \;\le\; \rho_m P(A)P(B) \tag{8}$$

whenever $A \in \mathcal{B}(X_0, X_1, \ldots, X_n)$ and $B \in \mathcal{B}(X_{n+m}, X_{n+m+1}, \ldots)$ for $n, m \ge 1$. Without loss of generality in (8), we can assume that $\Omega = R^\infty$, $\mathcal{F} = \mathcal{B}(R^\infty)$, and $X_i(\omega) = \omega_{i+1}$ for $\omega = (\omega_1, \omega_2, \ldots, \omega_n, \ldots)$. In that case, the unilateral shift transformation $\theta(\omega) = (\omega_2, \omega_3, \ldots)$ on R^∞ is also called strongly mixing. (Recall that then $X_n(\omega) = X_0\big(\theta^n(\omega)\big)$ for $n \ge 1$.)

5. (i) For $p(i,j)$ as above and ν as in (6), prove that the random variables $X_0, X_1, \ldots$ are *strictly stationary*. That is,

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) \;=\; P(X_m = i_0, X_{m+1} = i_1, \ldots, X_{m+n} = i_n) \tag{9}$$

for all integers $m, n \geq 0$ and $1 \leq i_a \leq N$.

(ii) Use part (i) to show that

$$\lim_{n\to\infty} \frac{f\big(X_1(\omega)\big) + f\big(X_2(\omega)\big) + \cdots + f\big(X_n(\omega)\big)}{n} \;=\; Y(\omega) \quad \text{converges a.s.}$$

for any function $f(j)$ on $S = \{\, 1, 2, \ldots, N \,\}$.

6. Show that the random variables $\{\, X_i \,\}$ in Problem 5 are strongly mixing; i.e., satisfy (8). (*Hint*: Use (7).) Use this or otherwise to prove that the shift operator $\theta(\omega)$ is ergodic. Finally, prove that

$$\lim_{n\to\infty} \frac{f\big(X_1(\omega)\big) + f\big(X_2(\omega)\big) + \cdots + f\big(X_n(\omega)\big)}{n} \;=\; \sum_{i=1}^{N} \nu_i f(i) \quad \text{a.s.} \tag{10}$$

for any function $f(x)$ on $S = \{\, 1, 2, \ldots, N \,\}$. (*Hints*: Recall that $X_1, X_2, \ldots$ on a probability space $(\Omega, \mathcal{F}, P)$ is *ergodic* if $A \in \mathcal{F}$, $\theta^{-1}(A) = A$ for the unilateral shift $\theta(\omega)$ implies $P(A) = 0$ or $P(A) = 1$. Consider that, while the Kolmogorov Zero-One Law does not apply to X_i since the X_i are in general not independent, its proof might generalize. Connect the conclusion to ergodicity.)