Ma 551 — Advanced Probability

Take-Home Final — Due Mon. Dec. 20, 2007

Prof. Sawyer — Washington University

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

In the following, i.r.v. means "independent random variables', i.i.d. means "independent and identically distributed random variables", and d.f. means "(probability) distribution function".

Problems and parts of problems may not be equally weighted.

Warning: There should be no collaboration on the Take-Home Final.

1. Let $\phi(\theta) = E\left(e^{i\theta X}\right)$ for a random variable X.

(i) Show that $\phi(\theta)$ is real for all θ if and only if X is symmetrically distributed (that is, if and only if $P(X \le y) = P(X \ge -y)$ for all y).

(ii) Let $Z = X - Y$ where X and Y are i.i.d. Prove that $\phi_Z(\theta) = |\phi(\theta)|^2$ for all θ where $\phi_Z(\theta) = E\left(e^{i\theta Z}\right)$.

2. (i) Let $\{ F_n(y) : n = 1, 2, \ldots \}$ be a sequence of probability distribution functions. Define what it means for $\{ F_n(y) \}$ to be tight.

(ii) Set $F_n(y) = P(X_n \le y)$ for random variables X_n. Assume that $\sup_n E\big(\lambda(X_n)\big) \le C < \infty$ where (a) $\lambda(y) = \lambda(-y) \ge 0$ and (b) $\lim_{y\to\infty} \lambda(y) = \infty$. Prove that $\{ F_n(y) \}$ is tight.

3. Let $\{ Z_n, Y_n \}$ be random variables such that

(i) $\lim_{n\to\infty} P(Z_n \le y) = H(y)$ for a.e. y for some d.f. $H(y)$

(ii) the family $\{ F_n(y) \}$ is tight for $F_n(y) = P(Y_n \le y)$.

Assume $h_n \to 0$ for numbers h_n. Prove that

$$\lim_{n\to\infty} P(Z_n + h_n Y_n \le y) = H(y) \text{ for a.e. } y \tag{1}$$

4. Let $N, X_1, X_2, \ldots$ be independent random variables such that $X_1, X_2, \ldots$ are identically distributed with $E(X_1^2) < \infty$ and $N \ge 0$ is integer-valued with

$$P(N = n) = e^{-\mu} \frac{\mu^n}{n!}, \qquad n = 0, 1, 2, \ldots$$

Let $Y = \sum_{i=1}^{N} X_i$, so that Y is the sum of a random number of random variables. Prove that

(i) $E(Y) = \mu E(X)$ and $\mathrm{Var}(Y) = \mu \,\mathrm{Var}(X) + \mu E(X)^2 = \mu E(X^2)$
(ii) For all θ

$$E(e^{i\theta Y}) = \exp\left(\mu \int \left(e^{i\theta y} - 1\right) F_X(dy)\right) \tag{2}$$

where $F_X(y) = P(X_1 \le y)$. (*Remarks*: N is said to have a *Poisson* distribution with mean μ and Y a *compound Poisson* distribution with *steps* X_i. Note the similarly between (2) and the limiting distribution of one of the example Central Limit Theorems proven in class.)

Remark. A function $f(z)$ on an open set $\mathcal{O} \subseteq R^2$ $(z = x + iy)$ is called *analytic* if $f'(z) = \lim_{h\to 0}(f(z+h) - f(z))/h$ exists for complex h for all complex $z \in \mathcal{O}$ and is continuous in z. It then follows from Cauchy's Theorem that (i) $f^{(n)}(z) = (d/dz)^n f(z)$ exists in the same sense for all $n \ge 0, z \in \mathcal{O}$ and (ii) for any $z_0 \in \mathcal{O}$, $f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$ converges uniformly in any closed disk $\{\, z : |z - z_0| \le r \,\} \subseteq \mathcal{O}$ for $a_n = f^{(n)}(z_0)/n!$.

A useful property of analytic functions is that if $f(z), g(z)$ are analytic on the same connected open set $\mathcal{O}$ and $f(z_n) = g(z_n)$ on a sequence of distinct points $z_n \in \mathcal{O}$ such that $z_n \to z_0 \in \mathcal{O}$, then $f(z) = g(z)$ identically for all $z \in \mathcal{O}$. This follows from the facts that (i) the coefficients a_n is the power series $\sum_{k=0}^{\infty} a_n (z - z_0)^n$ for $f(z)$ and $g(z)$ can be computed from $f(z_n) = g(z_n)$ and (ii) once we conclude $f(z) = g(z)$ for $|z - z_0| \le r$, we can pick other values $z_1 \in \mathcal{O}$ with $|z_1 - z_0| < r$ and extend the argument.

5. Let X, Y be random variables with the same *moment-generating function*

$$g(s) = E\left(e^{sX}\right) = E\left(e^{sY}\right) < \infty \qquad \text{for real } s,\ -h \le s \le h$$

Prove that $P(X \le y) = P(Y \le y)$ for a.e. y; i.e., that X and Y have the same probability distribution.
(*Hints*: Write $g(z) = E\left(e^{(s+i\theta)X}\right)$ for $z = s + i\theta$. Note that (i) $|g(z)| \le E(e^{sX})$, so that $g(z)$ is uniformly bounded in the vertical strip $\mathcal{S} = \{\, z = s + i\theta : |s| < h \,\}$, (ii) $(d/dz)e^{cz} = ce^{cz}$ for complex c, z and $|(e^a - e^b)/(b - a)| \le e^M$ for complex a, b with $|a|, |b| \le M$, so that $g_X(z) = E(e^{zX})$ and $g_Y(z) = E(e^{zY})$ are analyztic in $\mathcal{S}$, and (iii) the real line $\mathcal{R} = \{\, s + i0 : s \in R \,\}$ crosses $\mathcal{S}$. Can you use this to get information about the characteristic functions of X and Y?)

6. The *Laplace transform* of $f \in L^1(I)$ for $I = [0, \infty)$ is

$$Lf(s) = \int_0^\infty e^{-sx} f(x)\, dx \qquad (s > 0) \tag{3}$$

Prove

(i) If $f_n(x) = (-x)^n f(x) \in L^1(I)$ for all $n \geq 0$, then $Lf(s)$ is infinitely differentiable for real $s > 0$ and $(d/ds)^n Lf(s) = Lf_n(s)$.

(ii) If $f, g \in L^1(I)$ and $Lf(s) = Lg(s)$ for all real $s > 0$, then $f(x) = g(x)$ a.e.

(*Hint*: Define $Lf(z)$ and $Lg(z)$ by (3) for $z = s + i\theta$ for $s > 0$. Are $Lf(z)$ and $Lg(z)$ analytic on $\mathcal{H} = \{\, s + i\theta : s > 0 \,\} \subseteq R^2$? If so, what can you conclude?)

Remark. (*Normal families arguments.*) Another useful result from complex variables that also follows from Cauchy's theorem is the following. If $f_n(z)$ are analytic functions that are uniformly bounded on an open set $\mathcal{O}$, then, for any subsequence n_k, there exists a further subsequence n_{k_ℓ} and an analytic function $g(z)$ on $\mathcal{O}$ such that $\lim_{\ell\to\infty} f_{n_{k_\ell}}(z) = g(z)$ uniformly on all compact subsets $K \subset \mathcal{O}$ (u.c.c.). Such a collection of analytic functions on an open set $\mathcal{O}$ is called a *normal family*.

One can then show by standard compactness arguments that if there is only one possible limit function $g(z)$ — for example, if $g(z)$ is uniquely determined on a subset of $\mathcal{O}$ with an accumulation point within $\mathcal{O}$ — then the entire sequence $f_n(z) \to g(z)$ u.c.c.

7. (*The Method of Moments*) Let $Z_1, Z_2, \dots, Z_k, \dots$ and Z be random variables such that

(i) $\lim_{k\to\infty} E\big(Z_k^n\big) = E\big(Z^n\big)$ for all integers $n = 1, 2, 3, \dots$ and

(ii) $\sup_k E\big(e^{sZ_k}\big) \leq C < \infty$ for $|s| \leq h$ for constants $h > 0$ and C.

Use the characteristic functions $\phi_n(\theta) = E\big(e^{i\theta Z_n}\big)$ to prove that

$$\lim_{k\to\infty} P(Z_k \leq y) \;=\; P(Z \leq y) \text{ for a.e. } y$$

at all points of continuity of $F(y) = P(Z \leq y)$.
(*Hints*: First show that $E(e^{sZ_n}) \to E(e^{sZ})$ for real s with $|s| < h$.)