Review Handout 1 – long version

Math/Stat 491: Introduction to Stochastic Processes

Wellner; 10/25/2013

Part 1: Terminology and Definitions

1. sigma field:
   a collection $\mathcal{F}$ of subsets of the sample space Ω satisfying: (a) $\Omega \in \mathcal{F}$; (b) if $A \in \mathcal{F}$, then $A^c = \Omega \setminus A \in \mathcal{F}$; (c) if $A_1, A_2, \ldots \in \mathcal{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

2. probability measure:
   a function $P : \mathcal{F} \to [0,1]$ for a sigma-field $\mathcal{F}$ of subsets of Ω satisfying: (a) $P(A) \ge 0$ for all $A \in \mathcal{F}$ and (b) $P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ for any collection of sets $\{A_n\}$ with $A_i \cap A_j = \emptyset$ for $i \ne j$.

3. probability space:
   a triple $(\Omega, \mathcal{F}, P)$ where Ω is some non-empty set (the sample space), $\mathcal{F}$ is sigma - field of subsets of Ω, and P is a probability measure defined on $\mathcal{F}$.

4. random variable:
   a real valued function defined on Ω that is $\mathcal{F} - \mathcal{B}$ measurable: that is, $X^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{B}$, the Borel sigma-field on $\mathbb{R}$.

5. independent events $A_1, \ldots, A_n$; independent random variables $X_1, \ldots, X_n$.
   events $A_1, \ldots, A_n$ are *independent* if $P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$ for any choice of $1 \le i_1 < i_2 < \cdots < i_k \le n$.
   random variables $X_1, \ldots, X_n$ are independent if $E\{g_1(X_1) \cdots g_n(X_n)\} = E\{g_1(X_1)\} \cdots E\{g_n(X_n)\}$ for all Borel functions g_i, $i = 1, \ldots, n$.

6. stochastic process (sample path of a stochastic process):
   a stochastic process $\{X_t : \ t \in T\}$ for some index set T is a collection of random variables $X_t : \Omega \to \mathbb{R}$ for each $t \in T$.
   the sample paths of $\{X_t : \ t \in T\}$ are the functions $t \mapsto X_t(\omega)$ for $\omega \in \Omega$ fixed.

7. Borel function:
A Borel function g is a $\mathcal{B}-\mathcal{B}$-measurable function g from $\mathbb{R}$ to $\mathbb{R}$; that is, $g^{-1}(B) \in \mathcal{B}$ for every $B \in \mathcal{B}$.

8. characteristic function of a random variable; joint characteristic function of random variables $X_1, \ldots, X_n$:
The characteristic function of a random variable X is defined by $\phi_X(t) = E\{e^{itX}\}$.
The characteristic function of $X_1, \ldots, X_n$ is defined by
$\phi_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = E\{\exp(i \sum_{j=1}^{n} t_j X_j)\}$.

9. conditional expectation (in terms of fundamental identity):
The expectation $E(Y|X)$ is a measurable function of X satisfying

$$E\{g(X)E(Y|X)\} = E\{g(X)Y\}$$

for every bounded Borel function g.

10. martingale, sub-martingale, super-martingale:
A process $\{X_n : n \geq 0\}$ is a martingale if $E|X_n| < \infty$ and $E(X_{n+1}|X_n) = X_n$ for all $n \geq 0$. It is a sub-martingale if $E(X_{n+1}|X_n) \geq X_n$ for all $n \geq 0$; and it is a super-martingale if $E(X_{n+1}|X_n) \leq X_n$ for all $n \geq 0$.

11. stopping time: T is a stopping time relative to σ-fields $\{\mathcal{F}\}_n$ if $[T \leq n] \in \mathcal{F}_n$ for each n.

12. predictable process:
A process A_n is predictable with respect to $\sigma-$ fields $\mathcal{F}_n$ if A_n is $\mathcal{F}_{n-1}$ measurable for each n.

13. random walk: If $Y_1, Y_2, \ldots$ are independent then $S_n = \sum_{j=1}^{n} Y_j$, $n \geq 1$ is a *random walk*.

Part 2: Results and theorems

1. Properties of conditional expectation:
   (a) $E(Y|X)$ is unique up to a set with probability 0;
   (b) $E(E(Y|X)) = E(Y)$;
   (c) $E(Y|X) = E(Y)$ if X and Y are independent;
   (d) $E\{\phi(X)Y|X\} = \phi(X)E(Y|X)$;
   (e) $E\{Y + Z|X\} = E(Y|X) + E(Z|X)$;
   (f) $E\{E(Y|X,Z)|X\} = E(Y|X)$;
   (g) If $Y \geq 0$ then $E(Y|X) \geq 0$.

2. Variance decomposition:
   (a) $Var(Y) = E\{Var(Y|X)\} + Var\{E(Y|X)\}$; (b) proof; (c) geometric interpretation in terms of orthogonal components

3. martingales connected with sums of independent random variables:
   $S_n = \sum_{i=1}^n X_i$ with X_i independent with $E(X_i) = 0$;
   $S_n^2 - n\sigma^2$ if the X_i's are i.i.d. with $E(X_i) = 0$ and $Var(X_i) = \sigma^2$;
   $\exp(\lambda S_n)/\phi(\lambda)^n$ if X_i are i.i.d. with moment generating function $\phi(\lambda) = Ee^{\lambda X_1}$.

4. martingales connected with products:
   (products of independent mean 1 random variables; exponential/mgf martingales; likelihood ratio martingales)

5. Doob's optional sampling theorem:
   If $\{X_n : n \geq 0\}$ is a martingale and T is a stopping time, then $\{X_0, X_T\}$ is a martingale (and hence $E(X_T) = E(X_0)$) if any of the following hold:
   (i) $T \leq N < \infty$ with probability 1 for some fixed integer N.
   (ii) $\{X_n\}$ is bounded (so $|X_n| \leq K < \infty$ for all n) and $P(T < \infty) = 1$.
   (iii) $E(T) < \infty$ and $|X_n - X_{n-1}| \leq K$ for all n with probability 1.

6. Expected number of visits to origin for (simple) random walk. For a simple random walk in $\mathbb{R}^1$, $E\{V_0(1, 1/2)\} = \infty$ where $V_0(1,p) = \sum_{n=0}^{\infty} 1_{[S_{2n}=0]}$.
   For a simple random walk in $\mathbb{R}^2$, $E\{V_0(2, 1/2)\} = \infty$ where $V_0(2,p) = \sum_{n=0}^{\infty} 1_{[\underline{S}_{2n}=\underline{0}]}$.
   For a simple random walk in $\mathbb{R}^d$, $d \geq 3$, $E\{V_0(d, 1/2)\} < \infty$ where $V_0(d,p) = \sum_{n=0}^{\infty} 1_{[\underline{S}_{2n}=\underline{0}]}$.

7. Doob's decomposition of a submartingale: If $\{X_n : n \geq 0\}$ is a submartingale, then $X_n = (X_n - A_n) + A_n = Y_n + A_n$ where $\{Y_n : n \geq 0\}$ is a martingale and A_n is a predictable and increasing process: A_n is measurable with respect to $\mathcal{F}_{n-1}$ for each n and $A_0 \leq A_1 \leq \cdots \leq A_n$.

8. Jensen's inequality for $Eg(X)$ with g convex: $E\{g(X)\} \geq g(E(X))$ if g is convex.

9. Jensen's inequality for conditional expectation $E\{g(Y)|X\}$ with g convex: $E\{g(Y)|X\} \geq g(E(Y|X))$.

Part 3: Facts and calculations to know:

1. Binomial and Bernoulli process facts: see 394 Handouts # x and # y.

2. Poisson process facts: see 394 Handouts # 3 and #4.

3. the distribution of $X \sim N(\mu, \sigma^2)$ distribution:
   the density of X is given by
   $$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

4. bivariate Gaussian distribution of $(X, Y) \sim N_2((\mu, \nu), \Sigma)$ distribution with
   $$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}.$$
   Marginal and conditional distributions of bivariate Gaussian distribution:
   The joint density is given by
   $$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma\tau(1-\rho^2)^{1/2}} \exp\left(-\frac{1}{2}\left(\frac{(x-\mu)^2}{\sigma^2} - \frac{2\rho(x-\mu)(y-\nu)}{\sigma\tau} + \frac{(y-\nu)^2}{\tau^2}\right)\right).$$
   When $\mu = \nu = 0$, $(Y|X = x) \sim N(\rho x, 1 - \rho^2)$ and $X \sim N(0, 1)$.

5. Stirling's formula: $n! \sim \sqrt{2\pi n}(n/e)^n$

6. Newton's binomial formula and Newton's series: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$;
   $$(1+t)^a = 1 + \binom{a}{1}t + \binom{a}{2}t^2 + \binom{a}{3}t^3 + \cdots = \sum_{m=0}^{\infty} \binom{a}{m} t^m.$$