Part III — Advanced Probability
Definitions

Based on lectures by M. Lis

Notes taken by Dexter Chua

Michaelmas 2017

The aim of the course is to introduce students to advanced topics in modern probability theory. The emphasis is on tools required in the rigorous analysis of stochastic processes, such as Brownian motion, and in applications where probability theory plays an important role.

Review of measure and integration: sigma-algebras, measures and filtrations; integrals and expectation; convergence theorems; product measures, independence and Fubini's theorem.
Conditional expectation: Discrete case, Gaussian case, conditional density functions; existence and uniqueness; basic properties.
Martingales: Martingales and submartingales in discrete time; optional stopping; Doob's inequalities, upcrossings, martingale convergence theorems; applications of martingale techniques.
Stochastic processes in continuous time: Kolmogorov's criterion, regularization of paths; martingales in continuous time.
Weak convergence: Definitions and characterizations; convergence in distribution, tightness, Prokhorov's theorem; characteristic functions, Lévy's continuity theorem.
Sums of independent random variables: Strong laws of large numbers; central limit theorem; Cramér's theory of large deviations.
Brownian motion: Wiener's existence theorem, scaling and symmetry properties; martingales associated with Brownian motion, the strong Markov property, hitting times; properties of sample paths, recurrence and transience; Brownian motion and the Dirichlet problem; Donsker's invariance principle.
Poisson random measures: Construction and properties; integrals.
Lévy processes: Lévy-Khinchin theorem.

Pre-requisites

A basic familiarity with measure theory and the measure-theoretic formulation of probability theory is very helpful. These foundational topics will be reviewed at the beginning of the course, but students unfamiliar with them are expected to consult the literature (for instance, Williams' book) to strengthen their understanding.

Contents

0 Introduction

1 Some measure theory

1.1 Review of measure theory

Definition (σ-algebra)**.** Let E be a set. A subset $\mathcal{E}$ of the power set $\mathcal{P}(E)$ is called a *σ-algebra* (or *σ-field*) if

(i) $\emptyset \in \mathcal{E}$;

(ii) If $A \in \mathcal{E}$, then $A^C = E \setminus A \in \mathcal{E}$;

(iii) If $A_1, A_2, \ldots \in \mathcal{E}$, then $\bigcup_{n=1}^\infty A_n \in \mathcal{E}$.

Definition (Measurable space)**.** A *measurable space* is a set with a σ-algebra.

Definition (Borel σ-algebra)**.** Let E be a topological space with topology $\mathcal{T}$. Then the *Borel σ-algebra* $\mathcal{B}(E)$ on E is the σ-algebra generated by $\mathcal{T}$, i.e. the smallest σ-algebra containing $\mathcal{T}$.

Definition (Measure)**.** A function $\mu : \mathcal{E} \to [0, \infty]$ is a *measure* if

- $\mu(\emptyset) = 0$
- If $A_1, A_2, \ldots \in \mathcal{E}$ are disjoint, then
$$\mu\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i).$$

Definition (Measure space)**.** A *measure space* is a measurable space with a measure.

Definition (Measurable function)**.** Let $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ be measurable spaces. Then $f : E_1 \to E_2$ is said to be *measurable* if $A \in \mathcal{E}_2$ implies $f^{-1}(A) \in \mathcal{E}_1$.

Notation. For $(E, \mathcal{E})$ a measurable space, we write $m\mathcal{E}$ for the set of measurable functions $E \to \mathbb{R}$.

We write $m\mathcal{E}^+$ to be the positive measurable functions, which are allowed to take value ∞.

Definition (Simple function)**.** A function f is *simple* if there exists $\alpha_n \in \mathbb{R}_{\geq 0}$ and $A_n \in \mathcal{E}$ for $1 \leq n \leq k$ such that
$$f = \sum_{n=1}^k \alpha_n \mathbf{1}_{A_n}.$$

Definition (Almost everywhere)**.** We say $f = g$ almost everywhere if
$$\mu(\{x \in E : f(x) \neq g(x)\}) = 0.$$
We say f is a *version* of g.

Definition (Integrable function)**.** We say a function $f \in m\mathcal{E}$ is *integrable* if $\mu(|f|) \leq \infty$. We write $L^1(E)$ (or just L^1) for the space of integrable functions.

We extend μ to L^1 by
$$\mu(f) = \mu(f^+) - \mu(f^-),$$
where $f^\pm = (\pm f) \wedge 0$.

Definition (Product σ-algebra)**.** Let $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ be measure spaces. Then the product σ-algebra$\mathcal{E}_1 \otimes \mathcal{E}_2$ is the smallest σ-algebra on $E_1 \times E_2$ containing all sets of the form $A_1 \times A_2$, where $A_i \in \mathcal{E}_i$.

1.2 Conditional expectation

2 Martingales in discrete time

2.1 Filtrations and martingales

Definition (Filtration)**.** A *filtration* is a sequence of σ-algebras $(\mathcal{F}_n)_{n\geq 0}$ such that $\mathcal{F} \supseteq \mathcal{F}_{n+1} \supseteq \mathcal{F}_n$ for all n. We define $\mathcal{F}_\infty = \sigma(\mathcal{F}_0, \mathcal{F}_1, \ldots) \subseteq \mathcal{F}$.

Definition (Stochastic process in discrete time)**.** A *stochastic process* (in discrete time) is a sequence of random variables $(X_n)_{n\geq 0}$.

Definition (Natural filtration)**.** The *natural filtration* of $(X_n)_{n\geq 0}$ is given by

$$\mathcal{F}_n^X = \sigma(X_1, \ldots, X_n).$$

Definition (Adapted process)**.** We say that $(X_n)_{n\geq 0}$ is *adapted* (to $(\mathcal{F}_n)_{n\geq 0}$) if X_n is $\mathcal{F}_n$-measurable for all $n \geq 0$. Equivalently, if $\mathcal{F}_n^X \subseteq \mathcal{F}_n$.

Definition (Integrable process)**.** A process $(X_n)_{n\geq 0}$ is *integrable* if $X_n \in L^1$ for all $n \geq 0$.

Definition (Martingale)**.** An integrable adapted process $(X_n)_{n\geq 0}$ is a *martingale* if for all $n \geq m$, we have

$$\mathbb{E}(X_n \mid \mathcal{F}_m) = X_m.$$

We say it is a *super-martingale* if

$$\mathbb{E}(X_n \mid \mathcal{F}_m) \leq X_m,$$

and a *sub-martingale* if

$$\mathbb{E}(X_n \mid \mathcal{F}_m) \geq X_m,$$

2.2 Stopping time and optimal stopping

Definition (Stopping time)**.** A *stopping time* is a random variable $T : \Omega \to \mathbb{N}_{\geq 0} \cup \{\infty\}$ such that

$$\{T \leq n\} \in \mathcal{F}_n$$

for all $n \geq 0$.

Definition (X_T)**.** For a stopping time T, we define the random variable X_T by

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

on $\{T < \infty\}$, and 0 otherwise.

Definition (Stopped process)**.** The *stopped process* is defined by

$$(X_n^T)_{n\geq 0} = (X_{T(\omega)\wedge n}(\omega))_{n\geq 0}.$$

Definition ($\mathcal{F}_T$)**.** For a stopping time T, define

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \leq n\} \in \mathcal{F}_n\}.$$

2.3 Martingale convergence theorems

Definition (Upcrossing)**.** Let (x_n) be a sequence and (a, b) an interval. An *upcrossing* of (a, b) by (x_n) is a sequence $j, j+1, \ldots, k$ such that $x_j \leq a$ and $x_k \geq b$. We define

$$\begin{aligned} U_n[a, b, (x_n)] &= \text{number of disjoint upcrossings contained in } \{1, \ldots, n\} \\ U[a, b, (x_n)] &= \lim_{n \to \infty} U_n[a, b, x]. \end{aligned}$$

2.4 Applications of martingales

Definition (Backwards filtration)**.** A *backwards filtration* on a measurable space $(E, \mathcal{E})$ is a sequence of σ-algebras $\hat{\mathcal{F}}_n \subseteq \mathcal{E}$ such that $\hat{F}_{n+1} \subseteq \hat{F}_n$. We define

$$\hat{\mathcal{F}}_\infty = \bigcap_{n \geq 0} \hat{\mathcal{F}}_n.$$

Definition (Transition matrix)**.** A *transition matrix* is a matrix $P = (p_{xy})_{x,y \in E}$ such that each $p_x = (p_{x,y})_{y \in E}$ is a probability measure on E.

Definition (Markov chain)**.** An adapted process (X_n) is called a *Markov chain* if for any n and $A \in \mathcal{F}_n$ such that $\{x_n = x\} \supseteq A$, we have

$$\mathbb{P}(X_{n+1} = y \mid A) = p_{xy}.$$

Definition (Harmonic function)**.** A function $f: E \to \mathbb{R}$ is *harmonic* if $Pf = f$. In other words, for any x, we have

$$\sum_y p_{xy} f(y) = f(x).$$

3 Continuous time stochastic processes

Definition (Continuous time stochastic process). A *continuous time stochastic process* is a family of random variables $(X_t)_{t \geq 0}$ (or $(X_t)_{t \in [a,b]}$).

Definition (Cadlag function). We say a function $X : [0, \infty] \to \mathbb{R}$ is *cadlag* if for all t

$$\lim_{s \to t^+} x_s = x_t, \quad \lim_{s \to t^-} x_s \text{ exists.}$$

Definition (Continuous/Cadlag stochastic process). We say a stochastic process is *continuous* (resp. cadlag) if for any $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$ is continuous (resp. cadlag).

Notation. We write $C([0, \infty), \mathbb{R})$ for the space of all continuous functions $[0, \infty) \to \mathbb{R}$, and $D([0, \infty), \mathbb{R})$ the space of all cadlag functions.

We endow these spaces with a σ-algebra generated by the coordinate functions

$$(x_t)_{t \geq 0} \mapsto x_s.$$

Definition (Finite-dimensional distribution). A *finite dimensional distribution* of $(X_t)_{t \geq 0}$ is a measure on $\mathbb{R}^n$ of the form

$$\mu_{t_1, \ldots, t_n}(A) = \mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in A)$$

for all $A \in \mathcal{B}(\mathbb{R}^n)$, for some $0 \leq t_1 < t_2 < \ldots < t_n$.

Definition (Dyadic numbers). We define

$$D_n = \left\{ s \in [0, 1] : s = \frac{k}{2^n} \text{ for some } k \in \mathbb{Z} \right\}, \quad D = \bigcup_{n \geq 0} D_n.$$

Definition (Continuous time filtration). A *continuous-time filtration* is a family of σ-algebras $(\mathcal{F}_t)_{t \geq 0}$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ if $s \leq t$. Define $\mathcal{F}_\infty = \sigma(\mathcal{F}_t : t \geq 0)$.

Definition (Stopping time). A random variable $t : \Omega \to [0, \infty]$ is a *stopping time* if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.

Definition (Hitting time). Let $A \in \mathcal{B}(\mathbb{R})$. Then the *hitting time* of A is

$$T_A = \inf_{t \geq 0} \{X_t \leq A\}.$$

Definition (Right-continuous filtration). Given a continuous filtration $(\mathcal{F}_t)_{t \geq 0}$, we define

$$\mathcal{F}_t^+ = \bigcap_{s > t} \mathcal{F}_s \supseteq \mathcal{F}_t.$$

We say $(\mathcal{F}_t)_{t \geq 0}$ is *right continuous* if $\mathcal{F}_t = \mathcal{F}_t^+$.

Definition (Usual conditions). Let $\mathcal{N} = \{A \in \mathcal{F}_\infty : \mathbb{P}(A) \in \{0, 1\}\}$. We say that $(\mathcal{F}_t)_{t \geq 0}$ satisfies the *usual conditions* if it is right continuous and $\mathcal{N} \subseteq \mathcal{F}_0$.

Definition (Coninuous time martingale). An adapted process $(X_t)_{t \geq 0}$ is called a *martingale* iff

$$\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s$$

for all $t \geq s$, and similarly for super-martingales and sub-martingales.

Definition (Version). We say a process $(Y_t)_{t \geq 0}$ is a *version* of $(X_t)_{t \geq 0}$ if for all t, $\mathbb{P}(Y_t = X_t) = 1$.

4 Weak convergence of measures

Definition (Law)**.** Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. The *law* of X is the probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$\mu(A) = \mathbb{P}(X^{-1}(A)).$$

Definition (Weak convergence)**.** Let $(\mu_n)_{n\geq 0}$, μ be probability measures on a metric space (M, d) with the Borel measure. We say that $\mu_n \Rightarrow \mu$, or μ_n *converges weakly* to μ if

$$\mu_n(f) \to \mu(f)$$

for all f bounded and continuous.

If $(X_n)_{n\geq 0}$ are random variables, then we say (X_n) converges *in distribution* if μ_{X_n} converges weakly.

Definition (Tight probability measures)**.** A sequence of probability measures $(\mu_n)_{n\geq 0}$ on a metric space (M, e) is *tight* if for all $\varepsilon > 0$, there exists compact $K \subseteq M$ such that

$$\sup_n \mu_n(M \setminus K) \leq \varepsilon.$$

Definition (Characteristic function)**.** Let X be a random variable taking values in $\mathbb{R}^d$. The *characteristic function* of X is the function $\mathbb{R}^d \to \mathbb{C}$ defined by

$$\varphi_X(t) = \mathbb{E}e^{i\langle t, x\rangle} = \int_{\mathbb{R}^d} e^{i\langle t, x\rangle}\;\mathrm{d}\mu_X(x).$$

5 Brownian motion

5.1 Basic properties of Brownian motion

Definition (Brownian motion)**.** A continuous process $(B_t)_{t\geq 0}$ taking values in $\mathbb{R}^d$ is called a *Brownian motion* in $\mathbb{R}^d$ started at $x \in \mathbb{R}^d$ if

(i) $B_0 = x$ almost surely.

(ii) For all $s < t$, the *increment* $B_t - B_s \sim N(0, (t-s)I)$.

(iii) Increments are independent. More precisely, for all $t_1 < t_2 < \cdots < t_k$, the random variables

$$B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}}$$

are independent.

If $B_0 = 0$, then we call it a *standard Brownian motion*.

5.2 Harmonic functions and Brownian motion

Definition (Domain)**.** A *domain* is an open connected set $D \subseteq \mathbb{R}^d$.

Definition (Harmonic function)**.** A function $u : D \to \mathbb{R}$ is called *harmonic* if

$$\Delta f = \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2} = 0.$$

Definition (Maximum principle)**.** Let $u : \bar{D} \to \mathbb{R}$ be continuous and harmonic. Then

(i) If u attains its maximum inside D, then u is constant.

(ii) If D is bounded, then the maximum of u in $\bar{D}$ is attained at ∂D.

Definition (Poincaré cone condition)**.** We say a domain D satisfies the *Poincaré cone condition* if for any $x \in \partial D$, there is an open cone C based at X such that

$$C \cap D \cap B(x, \delta) = \emptyset$$

for some $\delta \geq 0$.

5.3 Transience and recurrence

5.4 Donsker's invariance principle

6 Large deviations