Lecture 1 : Overview. Conditional Expectation I.

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Sections 0, 4.8, 9], [Dur10, Section 5.1].

1 Stochastic processes

The course MATH 275B is an introduction to stochastic processes.

DEF 1.1 *A* stochastic process (SP) *is a collection* $\{X_t\}_{t\in\mathcal{T}}$ *of* $(E,\mathcal{E})$*-valued random variables on a triple* $(\Omega,\mathcal{F},\mathbb{P})$*, where* $\mathcal{T}$ *is an arbitrary* index set. *For a fixed* $\omega\in\Omega$, $\{X_t(\omega)\,:\,t\in\mathcal{T}\}$ *is called a* sample path.

EX 1.2 *When* $\mathcal{T}=\mathbb{N}$ *or* $\mathcal{T}=\mathbb{Z}_+$ *we have a* discrete-time SP. *For instance,*

- $X_1, X_2,\ldots$ *iid RVs*
- $\{S_n\}_{n\geq 1}$ *where* $S_n=\sum_{i\leq n} X_i$ *with* X_i *as above*

EX 1.3 *When* $\mathcal{T}=\mathbb{R}_+$*, we have a* continuous-time SP. *For instance,*

- $N_t=\sup\{n\geq 1\,:\,S_n\leq t\}$ *where* S_n *is as above with nonnegative* X_i*s*

In general, $\mathcal{T}$ does not need to represent time.

EX 1.4 *When* $\mathcal{T}$ *is finite, we have a* random vector. *Although seemingly simple, this example encapsulates many non-trivial SPs. For instance,*

- *Let* $V=\{1,\ldots,n\}$ *and* $E=\{e=(u,v)\,:\,u\neq v\in V\}$ *. Consider iid RVs* $X(e)$*,* $e\in E$*, distributed according to Bernoulli(p) for* $0\leq p\leq 1$*. Then* $G_p=(V,E_p)$*, where* $E_p=\{e\in E\,:\,X(e)=1\}$*, is called an* Erdos-Renyi random graph.

2 A Preview of Things to Come

Two main themes:

1. Beyond independence

2. Sample path properties

Here are a few important examples of processes and questions we will answer about them.

2.1 Random walks

DEF 1.5 *A* random walk (RW) *on* $\mathbb{R}^d$ *is an SP of the form:*

$$S_n = \sum_{i \le n} X_i,\ n \ge 1$$

where the X_i*s are iid in* $\mathbb{R}^d$.

EX 1.6 *When* $d = 1$*, recall from MATH 275A that*

- *SLLN:* $n^{-1}S_n \to \mathbb{E}[X_1]$ *a.s. when* $\mathbb{E}|X_1| < +\infty$
- *CLT:*
$$\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\mathrm{Var}[X_1]}} \Rightarrow N(0,1),$$
when $\mathbb{E}[X_1^2] < \infty$.

These are examples of limit theorems. Sample path properties, on the other hand, involve properties of the sequence $S_1(\omega), S_2(\omega), \ldots$*. For instance, let* $A \subset \mathbb{R}^d$

- $\mathbb{P}[S_n \in A$ *for some* $n \ge 1]$*?*
- $\mathbb{P}[S_n \in A \text{ i.o.}]$*?*
- $\mathbb{E}[T_A]$*? where* $T_A = \inf\{n \ge 1 \,:\, S_n \in A\}$

2.2 Branching processes

DEF 1.7 *A* branching process *is an SP of the form:*

- *Let* $X(i,n)$*,* $i \ge 1$*,* $n \ge 1$*, be an array of iid* $\mathbb{Z}_+$*-valued RVs with finite mean* $\mu = \mathbb{E}[X(1,1)] < +\infty$ *and* $\mathbb{P}[X(1,1) = 0] > 0$

- $Z_0 = 1$, *and inductively,*

$$Z_n = \sum_{1 \leq i \leq Z_{n-1}} X(i, n)$$

EX 1.8 *Typical questions about branching processes are:*

- *Extinction:* $\mathbb{P}[Z_n = 0 \text{ for some } n \geq 1]$*?*
- *Exponential growth:* $M_n = \mu^{-n} Z_n \to ?$
- *Limit of expectations: when* $\mu < 1$ *we have* $\mathbb{E}[M_n] = 1$ *for all* n *yet* $\mathbb{E}[M_\infty] = 0$

2.3 Markov chains

The two previous examples are special cases of a large class of SPs.

DEF 1.9 *A* discrete-time countable-space Markov chain(MC) *is an SP of the form:*

- E *countable state space*
- μ *initial distribution, that is,* $\mu_i \geq 0$, $i \in E$, *and* $\sum_{i \in E} \mu_i = 1$
- $\{p_{ij}\}_{i,j \in E}$ *transition matrix, that is,* $p_{ij} \geq 0$, $i, j \in E$, *and* $\sum_{j \in E} p_{ij} = 1$ *for all* $i \in E$
- *Let* $Y(i, n)$, $i \in E$, $n \geq 1$, *be an array of iid RVs distributed according to* $p_{i\cdot}$
- *Define the process recursively by* $Z_0 = 0$, *and,*

$$Z_n = Y(Z_{n-1}, n)$$

3 Review of undergraduate conditional probability

3.1 Conditional probability

For two events A, B, the conditional probability of A given B is defined as

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]},$$

where we assume $\mathbb{P}[B] > 0$.

3.2 Conditional expectation

Let X and Z be RVs taking values $x_1, \ldots, x_m$ and $z_1, \ldots, z_n$ resp. The conditional expectation of X given $Z = z_j$ is given as

$$y_j \equiv \mathbb{E}[X \mid Z = z_j] = \sum_i x_i \mathbb{P}[X = x_i \mid Z = z_j].$$

We assume $\mathbb{P}[Z = z_j] > 0$.

As motivation for the general definition, we make the following observations:

- We can think of the conditional expectation as a RV $Y \equiv \mathbb{E}[X \mid Z]$ defined as follows:
$$Y(\omega) = y_j, \text{ on } G_j \equiv \{\omega \,:\, Z(\omega) = z_j\}.$$
- Then Y is $\mathcal{G}$-measurable where $\mathcal{G} = \sigma(Z)$.
- On sets in $\mathcal{G}$, the expectation of Y agrees with the expectation of X, that is,
$$\begin{aligned} \mathbb{E}[Y; G_j] &= y_j \mathbb{P}[G_j] \\ &= \sum_i x_i \mathbb{P}[X = x_i \mid Z = z_j] \mathbb{P}[Z = z_j] \\ &= \sum_i x_i \mathbb{P}[X = x_i, Z = z_j] \\ &= \mathbb{E}[X; G_j]. \end{aligned}$$
This is also true for all $G \in \mathcal{G}$ by summation.

4 Conditional expectation: definition, existence, uniqueness

4.1 Definition

DEF&THM 1.10 *Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G} \subseteq \mathcal{F}$ a sub σ-field. Then there exists a (a.s.) unique $Y \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ s.t.*

$$\mathbb{E}[Y; G] = \mathbb{E}[X; G], \ \forall G \in \mathcal{G}.$$

Such Y is called a version of $\mathbb{E}[X \mid \mathcal{G}]$.

Further reading

Kolmogorov's extension theorem [Dur10, Section A.3]. Radon-Nikodym theorem [Dur10, Section A.4].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 2 : Conditional Expectation II

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 9], [Dur10, Section 5.1].

1 Conditional expectation: definition, existence, uniqueness

1.1 Definition

DEF&THM 2.1 *Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G} \subseteq \mathcal{F}$ a sub σ-field. Then there exists a (a.s.) unique $Y \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ s.t.*

$$\mathbb{E}[Y; G] = \mathbb{E}[X; G], \; \forall G \in \mathcal{G}.$$

Such Y is called a version of $\mathbb{E}[X \,|\, \mathcal{G}]$.

1.2 Proof of uniqueness

Let Y, Y' be two versions of $\mathbb{E}[X \,|\, G]$ such that w.l.o.g. $\mathbb{P}[Y > Y'] > 0$. By monotonicity, there is $n \geq 1$ with $G = \{Y > Y' + n^{-1}\} \in \mathcal{G}$ such that $\mathbb{P}[G] > 0$. Then, by definition,

$$0 = \mathbb{E}[Y - Y'; G] > n^{-1}\mathbb{P}[G] > 0,$$

which gives a contradiction.

1.3 Proof of existence

There are two main approaches:

1. First approach: Radon-Nikodym theorem. Read [Dur10, Section A.4].
2. Second approach: Hilbert space method.

We begin with a definition.

DEF&THM 2.2 *Let $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G} \subseteq \mathcal{F}$ a sub σ-field. Then there exists a (a.s.) unique $Y \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ s.t.*

$$\Delta \equiv \|X - Y\|_2 = \inf\{\|X - W\|_2 \,:\, W \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})\},$$

and, moreover,

$$\langle Z, X - Y\rangle = 0,\ \forall Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}).$$

Such Y is called an orthogonal projection of X on $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$.

We give a proof for completeness.

Proof: Take (Y_n) s.t. $\|X - Y_n\|_2 \to \Delta$. Remembering that $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is complete we seek to prove that (Y_n) is Cauchy. Using the parallelogram law

$$2\|U\|_2^2 + 2\|V\|_2^2 = \|U - V\|_2^2 + \|U + V\|_2^2,$$

note that

$$\|X - Y_r\|_2^2 + \|X - Y_s\|_2^2 = 2\|X - \frac{1}{2}(Y_r + Y_s)\|_2^2 + 2\|\frac{1}{2}(Y_r - Y_s)\|_2^2.$$

The first term on the LHS is at least Δ^2 so we have what we need. Let Y be the limit of (Y_n).

Note that for any $Z \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and $t \in \mathbb{R}$

$$\|X - Y - tZ\|_2^2 \geq \|X - Y\|_2^2,$$

so that, expanding and rearranging, we have

$$-2t\langle Z, X - Y\rangle + t^2\|Z\|_2^2 \geq 0,$$

which is only possible if the first term is 0.

Uniqueness follows from the parallelogram law again. ∎

We return to the proof of existence of the conditional expectation. We use the standard machinery. The previous theorem implies that conditional expectations exist for indicators and simple functions. Now take $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and write $X = X^+ - X^-$, so we can assume $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})^+$ w.l.o.g. Using the staircase function

$$X^{(r)} = \begin{cases} 0, & \text{if X = 0} \\ (i-1)2^{-r}, & \text{if } (i-1)2^{-r} < X \leq i2^{-r} \leq r \\ r, & \text{if } X > r, \end{cases}$$

we have $0 \leq X^{(r)} \uparrow X$. Let $Y^{(r)} = \mathbb{E}[X^{(r)}\,|\,\mathcal{G}]$. Using an argument similar to the proof of uniqueness, it follows that $U \geq 0$ implies $\mathbb{E}[U\,|\,\mathcal{G}] \geq 0$. Using linearity , we then have $Y^{(r)} \uparrow Y \equiv \limsup Y^{(r)}$ which is measurable in $\mathcal{G}$. By (MON)

$$\mathbb{E}[Y; G] = \mathbb{E}[X; G],\ \forall G \in \mathcal{G}.$$

2 Examples

EX 2.3 *If* $X \in \mathcal{L}^1(\mathcal{G})$*, then* $\mathbb{E}[X \,|\, \mathcal{G}] = X$ *a.s. trivially.*

EX 2.4 *If* $\mathcal{G} = \{\emptyset, \Omega\}$*, then* $\mathbb{E}[X \,|\, \mathcal{G}] = \mathbb{E}[X]$.

EX 2.5 *Let* $A, B \in \mathcal{F}$ *with* $0 < \mathbb{P}[B] < 1$. *If* $\mathcal{G} = \{\emptyset, B, B^c, \Omega\}$ *and* $X = \mathbb{1}_A$*, then*

$$\mathbb{P}[A \,|\, \mathcal{G}] = \begin{cases} \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}, & \text{on } \omega \in B \\ \frac{\mathbb{P}[A \cap B^c]}{\mathbb{P}[B^c]}, & \text{on } \omega \in B^c \end{cases}$$

3 Conditional expectation: properties

We show that conditional expectations behave the way one would expect. Below all Xs are in $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ is a sub σ-field of $\mathcal{F}$.

3.1 Extending properties of standard expectations

LEM 2.6 (cLIN) $\mathbb{E}[a_1 X_1 + a_2 X_2 \,|\, \mathcal{G}] = a_1 \mathbb{E}[X_1 \,|\, \mathcal{G}] + a_2 \mathbb{E}[X_2 \,|\, \mathcal{G}]$ *a.s.*

Proof: Use linearity of expectation and the fact that a linear combination of RVs in $\mathcal{G}$ is also in $\mathcal{G}$. ■

LEM 2.7 (cPOS) *If* $X \geq 0$ *then* $\mathbb{E}[X \,|\, \mathcal{G}] \geq 0$ *a.s.*

Proof: Let $Y = \mathbb{E}[X \,|\, \mathcal{G}]$ and assume $\mathbb{P}[Y < 0] > 0$. There is $n \geq 1$ s.t. $\mathbb{P}[Y < -n^{-1}] > 0$. But that implies, for $G = \{Y < -n^{-1}\}$,

$$\mathbb{E}[X; G] = \mathbb{E}[Y; G] < -n^{-1}\mathbb{P}[G] < 0,$$

a contradiction. ■

LEM 2.8 (cMON) *If* $0 \leq X_n \uparrow X$ *then* $\mathbb{E}[X_n \,|\, \mathcal{G}] \uparrow \mathbb{E}[X \,|\, \mathcal{G}]$ *a.s.*

Proof: Let $Y_n = \mathbb{E}[X_n \,|\, \mathcal{G}]$. By (cLIN) and (cPOS), $0 \leq Y_n \uparrow$. Then letting $Y = \limsup Y_n$, by (MON),

$$\mathbb{E}[X; G] = \mathbb{E}[Y; G],$$

for all $G \in \mathcal{G}$. ■

LEM 2.9 (cFATOU) *If* $X_n \geq 0$ *then* $\mathbb{E}[\liminf X_n \,|\, \mathcal{G}] \leq \liminf \mathbb{E}[X_n \,|\, \mathcal{G}]$ *a.s.*

Proof: Note that, for $n \geq m$,

$$X_n \geq Z_m \equiv \inf_{k \geq m} X_m \uparrow \in \mathcal{G},$$

so that $\inf_{n \geq m} \mathbb{E}[X_n \,|\, \mathcal{G}] \geq \mathbb{E}[Z_m \,|\, \mathcal{G}]$. Applying (cMON)

$$\mathbb{E}[\lim Z_m \,|\, \mathcal{G}] = \lim \mathbb{E}[Z_m \,|\, \mathcal{G}] \leq \lim \inf_{n \geq m} \mathbb{E}[X_n \,|\, \mathcal{G}].$$

■

LEM 2.10 (cDOM) *If* $X_n \leq V \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ *and* $X_n \to X$ *a.s., then*

$$\mathbb{E}[X_n \,|\, \mathcal{G}] \to \mathbb{E}[X \,|\, \mathcal{G}]$$

Proof: Apply (cFATOU) to $W_n = 2V - |X_n - X| \geq 0$

$$\mathbb{E}[2V \,|\, \mathcal{G}] = \mathbb{E}[\liminf W_n] \leq \liminf \mathbb{E}[W_n \,|\, \mathcal{G}] = \mathbb{E}[2V \,|\, \mathcal{G}] - \liminf \mathbb{E}[|X_n - X| \,|\, \mathcal{G}].$$

Use that, by definition, $|\mathbb{E}[X_n - X \,|\, \mathcal{G}]| \leq \mathbb{E}[|X_n - X| \,|\, \mathcal{G}]$. ■

LEM 2.11 (cJENSEN) *If* f *is convex and* $\mathbb{E}[|f(X)|] < +\infty$ *then*

$$f(\mathbb{E}[X \,|\, \mathcal{G}]) \leq \mathbb{E}[f(X) \,|\, \mathcal{G}].$$

Proof: Exercise! ■

3.2 Other properties

LEM 2.12 (Tower) *If* $\mathcal{H} \subseteq \mathcal{G}$ *is a* σ*-field*

$$\mathbb{E}[\mathbb{E}[X \,|\, \mathcal{G}] \,|\, \mathcal{H}] = \mathbb{E}[X \,|\, \mathcal{H}].$$

In particular $\mathbb{E}[\mathbb{E}[X \,|\, \mathcal{G}]] = \mathbb{E}[X]$.

Proof: Let $Y = \mathbb{E}[X \,|\, \mathcal{G}]$ and $Z = \mathbb{E}[X \,|\, \mathcal{H}]$. Then $Z \in \mathcal{H}$ and for $H \in \mathcal{H} \subseteq \mathcal{G}$

$$\mathbb{E}[Z; H] = \mathbb{E}[X; H] = \mathbb{E}[Y; H].$$

■

LEM 2.13 (Taking out what is known) *If* $Z \in \mathcal{G}$ *is bounded then*

$$\mathbb{E}[ZX \,|\, \mathcal{G}] = Z\mathbb{E}[X \,|\, \mathcal{G}].$$

This is also true if $X, Z \geq 0$ *and* $\mathbb{E}[ZX] < +\infty$ *or* $X \in \mathcal{L}^p(\mathcal{F})$ *and* $Z \in \mathcal{L}^q(\mathcal{G})$ *with* $p^{-1} + q^{-1} = 1$ *and* $p > 1$.

Proof: By (LIN), we restrict ourselves to $X \geq 0$. Clear if $Z = \mathbb{1}_{G'}$ is an indicator with $G' \in \mathcal{G}$ since

$$\mathbb{E}[\mathbb{1}_{G'}X; G] = \mathbb{E}[X; G \cap G'] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]; G \cap G'] = \mathbb{E}[\mathbb{1}_{G'}\mathbb{E}[X \mid \mathcal{G}]; G],$$

for all $G \in \mathcal{G}$. Use the standard machine to conclude. ■

LEM 2.14 (Role of independence) *If $\mathcal{H}$ is independent of $\sigma(\sigma(X), \mathcal{G})$, then*

$$\mathbb{E}[X \mid \sigma(\mathcal{G}, \mathcal{H})] = \mathbb{E}[X \mid \mathcal{G}].$$

In particular, if X is independent of $\mathcal{H}$ then $\mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[X]$.

Proof: Let $H \in \mathcal{H}$ and $G \in \mathcal{G}$. Since $Y = \mathbb{E}[X \mid \mathcal{G}] \in \mathcal{G}$, we have

$$\mathbb{E}[X; G \cap H] = \mathbb{E}[X; G]\mathbb{P}[H] = \mathbb{E}[Y; G]\mathbb{P}[H] = \mathbb{E}[Y; G \cap H].$$

We conclude with the following lemma.

LEM 2.15 (Uniqueness of extension) *Let $\mathcal{I}$ be a π-system on a set S, that is, a family of subsets stable under intersection. If μ_1, μ_2 are finite measures on $(S, \sigma(\mathcal{I}))$ with $\mu_1(\Omega) = \mu_2(\Omega)$ that agree on $\mathcal{I}$, then μ_1 and μ_2 agree on $\sigma(\mathcal{I})$.*

Indeed, note that the collection $\mathcal{I}$ of sets $G \cap H$ for $G \in \mathcal{G}, H \in \mathcal{H}$ form a π-system generating $\sigma(\mathcal{G}, \mathcal{H})$. ■

Further reading

Regular conditional probability [Dur10, Section 5.1]. π-λ theorem [Dur10, Section A.1].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 3 : Martingales: definition, examples

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 10], [Dur10, Section 5.2], [KT75, Section 6.1].

1 Definitions

DEF 3.1 *A* filtered space *is a tuple* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$ *where:*

- $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space*
- $\{\mathcal{F}_n\}$ *is a* filtration*, i.e.,*

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_\infty \equiv \sigma(\cup \mathcal{F}_n) \subseteq \mathcal{F}.$$

where each $\mathcal{F}_i$ *is a* σ*-field.*

Intuitively, $\mathcal{F}_i$ is the information up to time i.

EX 3.2 *Let* $X_0, X_1, \ldots$ *be iid RVs. Then a filtration is given by*

$$\mathcal{F}_n = \sigma(X_0, \ldots, X_n),\ \forall n \geq 0.$$

Fix $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$.

DEF 3.3 *A process* $\{W_n\}_{n\geq 0}$ *is* adapted *if* $W_n \in \mathcal{F}_n$ *for all* n.

Intuitively, the value of W_n is known at time n.

EX 3.4 *Continuing. Let* $\{S_n\}_{n\geq 0}$ *where* $S_n = \sum_{i\leq n} X_i$ *is adapted.*

DEF 3.5 *A process* $\{C_n\}_{n\geq 1}$ *is* previsible *if* $C_n \in \mathcal{F}_{n-1}$ *for all* $n \geq 1$.

EX 3.6 *Continuing.* $C_n = \mathbb{1}\{S_{n-1} \leq k\}$.

Our main definition is the following.

DEF 3.7 *A process* $\{M_n\}_{n\geq 0}$ *is a* martingale *(MG) if*

- $\{M_n\}$ *is adapted*
- $\mathbb{E}|M_n| < +\infty$ *for all* n
- $\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] = M_{n-1}$ *for all* $n \geq 1$

A superMG or subMG is similar except that the equality in the last property is replaced with $\leq$ *or* $\geq$ *respectively.*

2 Examples

EX 3.8 (Sums of iid RVs with mean 0**)** *Let*

- $X_0, X_1, \ldots$ *iid RVs integrable and centered with* $X_0 = 0$
- $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$
- $S_n = \sum_{i \leq n} X_i$

Then note that $\mathbb{E}|S_n| < \infty$ *by the triangle inequality and*

$$\begin{aligned}\mathbb{E}[S_n \,|\, \mathcal{F}_{n-1}] &= \mathbb{E}[S_{n-1} + X_n \,|\, \mathcal{F}_{n-1}] \\ &= S_{n-1} + \mathbb{E}[X_n] = S_{n-1}.\end{aligned}$$

EX 3.9 (Variance of a sum) *Same setup with* $\sigma^2 \equiv \mathrm{Var}[X_1] < \infty$. *Define*

$$M_n = S_n^2 - n\sigma^2.$$

Note that

$$\mathbb{E}|M_n| \leq \sum_{i \leq n} \mathrm{Var}[X_i] + n\sigma^2 \leq 2n\sigma^2 < +\infty$$

and

$$\begin{aligned}\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] &= \mathbb{E}[(X_n + S_{n-1})^2 - n\sigma^2 \,|\, \mathcal{F}_{n-1}] \\ &= \mathbb{E}[X_n^2 + 2X_n S_{n-1} + S_{n-1}^2 - n\sigma^2 \,|\, \mathcal{F}_{n-1}] \\ &= \sigma^2 + 0 + S_{n-1}^2 - n\sigma^2 = M_{n-1}.\end{aligned}$$

EX 3.10 (Exponential moment of a sum; Wald's MG) *Same setup with* $\phi(\lambda) = \mathbb{E}[\exp(\lambda X_1)] < +\infty$ *for some* $\lambda \neq 0$. *Define*

$$M_n = \phi(\lambda)^{-n} \exp(\lambda S_n).$$

Note that

$$\mathbb{E}|M_n| \leq \frac{\phi(\lambda)^n}{\phi(\lambda)^n} = 1 < +\infty$$

and

$$\begin{aligned}\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] &= \phi(\lambda)^{-n}\mathbb{E}[\exp(\lambda(X_n + S_{n-1})) \,|\, \mathcal{F}_{n-1}] \\ &= \phi(\lambda)^{-n} \exp(\lambda S_{n-1})\phi(\lambda) = M_{n-1}.\end{aligned}$$

EX 3.11 (Product of iid RVs with mean 1**)** *Same setup with* $X_0 = 1$, $X_i \geq 0$ *and* $\mathbb{E}[X_1] = 1$. *Define*

$$M_n = \prod_{i \leq n} X_i.$$

Note that

$$\mathbb{E}|M_n| = 1$$

and

$$\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] = M_{n-1}\mathbb{E}[X_n \,|\, \mathcal{F}_{n-1}] = M_{n-1}.$$

EX 3.12 (Accumulating data; Doob's MG) *Let* $X \in \mathcal{L}^1(\mathcal{F})$. *Define*

$$M_n = \mathbb{E}[X \,|\, \mathcal{F}_n].$$

Note that

$$\mathbb{E}|M_n| \leq \mathbb{E}|X| < +\infty,$$

and

$$\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] = \mathbb{E}[X \,|\, \mathcal{F}_{n-1}] = M_{n-1},$$

by (TOWER).

EX 3.13 (Eigenvalues of transition matrix) *Recall that a MC on a countable* E *is:*

- $\{\mu_i\}_{i \in E}$, $\{p(i,j)\}_{i,j \in E}$
- $Y(i,n) \sim p(i,\cdot)$ *(indep.)*
- $Z_0 \sim \mu$ *and* $Z_n = Y(Z_{n-1}, n)$.

Suppose $f : E \to \mathbb{R}$ *is s.t.*

$$\sum_j p(i,j) f(j) = \lambda f(i), \ \forall i,$$

with $\mathbb{E}|f(Z_n)| < +\infty$ *for all* n. *Define*

$$M_n = \lambda^{-n} f(Z_n).$$

Note that

$$\mathbb{E}|M_n| < +\infty,$$

and

$$\begin{aligned}\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] &= \lambda^{-n}\mathbb{E}[f(Z_n) \,|\, \mathcal{F}_{n-1}] \\ &= \lambda^{-n}\sum_j p(Z_{n-1}, j) f(j) \\ &= \lambda^{-n}\cdot\lambda\cdot f(Z_{n-1}) = M_{n-1}.\end{aligned}$$

EX 3.14 (Branching Process) *Recall that a branching process is:*

- $X(i,n)$, $i \geq 1$ *and* $n \geq 1$, *iid with mean* m
- $Z_0 = 1$ *and* $Z_n = \sum_{i\leq Z_{n-1}} X(i,n)$

Note that for $f(j) = j$ *we have*

$$\sum_j p(i,j)j = mi,$$

so that $M_n = m^{-n}Z_n$ *is a MG.*

Further reading

Comments on harmonic functions in [Dur10, Seciton 5.2].

Next class

Stopping times and betting systems [Dur10, Section 5.2].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[KT75] Samuel Karlin and Howard M. Taylor. *A first course in stochastic processes*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 4 : Martingales: gambling systems

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 10], [Dur10, Section 5.2].

1 Further definition and example

DEF 4.1 *A process $\{C_n\}_{n\geq 1}$ is* previsible *if $C_n \in \mathcal{F}_{n-1}$ for all $n \geq 1$.*

EX 4.2 *Let $\{X_n\}_{n\geq 0}$ be an integrable adapted process and $\{C_n\}_{n\geq 1}$, a bounded previsible process. Define*

$$M_n = \sum_{i\leq n}(X_i - \mathbb{E}[X_i \,|\, \mathcal{F}_{i-1}])C_i.$$

Then

$$\mathbb{E}|M_n| \leq \sum_{i\leq n} 2\mathbb{E}|X_n|K < +\infty,$$

where $|C_n| < K$ for all $n \geq 1$, and

$$\begin{aligned}\mathbb{E}[M_n - M_{n-1} \,|\, \mathcal{F}_{n-1}] &= \mathbb{E}[(X_n - \mathbb{E}[X_n \,|\, \mathcal{F}_{n-1}])C_n \,|\, \mathcal{F}_{n-1}] \\ &= C_n(\mathbb{E}[X_n \,|\, \mathcal{F}_{n-1}] - \mathbb{E}[X_n \,|\, \mathcal{F}_{n-1}]) = 0.\end{aligned}$$

2 Fair games

Take the previous example with $\{X_n\}_{n\geq 0}$ a MG, that is,

$$M_n = (C \bullet X)_n \equiv \sum_{i\leq n} C_i(X_i - X_{i-1}),$$

where $\{(C \bullet X)_n\}_{n\geq 0}$ is called the *martingale transform* and is a discrete analogue of stochastic integration. If you think of $X_n - X_{n-1}$ as your net winnings per unit stake at time n, then C_n is a gambling strategy and $(C \bullet X)$ is your total winnings up to time n in a *fair game*.

Arguing as in the previous example, we have the following theorem.

THM 4.3 (You can't beat the system) *Let $\{C_n\}$ be a bounded previsible process and $\{X_n\}$ be a MG. Then $\{(C \bullet X)_n\}$ is also a MG. If, moreover, $\{C_n\}$ is nonnegative and $\{X_n\}$ is a superMG, then $\{(C \bullet X)_n\}$ is also a superMG.*

3 Stopping times

DEF 4.4 *A random variable* $T : \Omega \to \overline{\mathbb{Z}}_+ \equiv \{0, 1, \ldots, +\infty\}$ *is called a* stopping time *if*

$$\{T \leq n\} \in \mathcal{F}_n, \ \forall n \in \overline{\mathbb{Z}}_+,$$

or, equivalently,

$$\{T = n\} \in \mathcal{F}_n, \ \forall n \in \overline{\mathbb{Z}}_+.$$

In the gambling context, a stopping time is a time at which you decide to stop playing. That decision should only depend on the history up to time n.

EX 4.5 *Let* $\{A_n\}$ *be an adapted process and* $B \in \mathcal{B}$*. Then*

$$T = \inf\{n \geq 0 \, : \, A_n \in B\},$$

is a stopping time.

4 Stopped supermartingales are supermartingales

DEF 4.6 *Let* $\{X_n\}$ *be an adapted process and* T *be a stopping time. Then*

$$X_n^T(\omega) \equiv X_{T(\omega)\wedge n}(\omega),$$

is called $\{X_n\}$ stopped at T.

THM 4.7 *Let* $\{X_n\}$ *be a superMG and* T *be a stopping time. Then the stopped process* X^T *is a superMG and in particular*

$$\mathbb{E}[X_{T\wedge n}] \leq \mathbb{E}[X_0].$$

The same result holds at equality if $\{X_n\}$ *is a MG.*

Proof: Let

$$C_n^{(T)} = \mathbb{1}\{n \leq T\}.$$

Note that

$$\{C_n^{(T)} = 0\} = \{T \leq n-1\} \in \mathcal{F}_{n-1},$$

so that $C^{(T)}$ is previsible. It is also nonnegative and bounded. Note further that

$$(C^{(T)} \bullet X)_n = X_{T\wedge n} - X_0 = X_n^T - X_0.$$

Apply the previous theorem. ■

5 Optional stopping theorem

When can we say that $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$?

THM 4.8 *Let $\{X_n\}$ be a superMG and T be a stopping time. Then X_T is integrable and*

$$\mathbb{E}[X_T] \leq \mathbb{E}[X_0].$$

if one of the following holds:

1. *T is bounded*
2. *X is bounded and T is a.s. finite*
3. *$\mathbb{E}[T] < +\infty$ and X has bounded increments*
4. *X is nonnegative and T is a.s. finite.*

The first three hold with equality if X is a MG.

Proof: From the previous theorem, we have

$$(*) \quad \mathbb{E}[X_{T\wedge n} - X_0] \leq 0.$$

1. Take $n = N$ in $(*)$ where $T \leq N$ a.s.
2. Take n to $+\infty$ and use (DOM).
3. Note that
$$|X_{T\wedge n} - X_0| \leq |\sum_{i \leq T\wedge n} (X_i - X_{i-1})| \leq KT,$$
where $|X_n - X_{n-1}| \leq K$ a.s. Use (DOM).
4. Use (FATOU).

■

Further reading

No further reading.

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 5 : Martingale convergence theorem

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 10], [Dur10, Section 5.2].

1 A natural gambling strategy

Recall that

$$(C \bullet X)_n = \sum_{i \leq n} C_n (X_n - X_{n-1}),$$

where C_n is predictable and X_n is a superMG, can be interpreted as your net winnings in a game. A natural strategy is to choose $\alpha < \beta$ and apply the following

- REPEAT
  - Wait until X gets below α
  - Play a unit stake until X gets above β and stop playing
- UNTIL TIME N

More formally, let

$$C_1 = \mathbb{1}\{X_0 < \alpha\},$$

and

$$C_n = \mathbb{1}\{C_{n-1} = 1\}\mathbb{1}\{X_{n-1} \leq \beta\} + \mathbb{1}\{C_{n-1} = 0\}\mathbb{1}\{X_{n-1} < \alpha\}.$$

Then $\{C_n\}$ is predictable.

2 Upcrossings

Define the following stopping times. Let $T_0 = -1$,

$$T_{2k-1} = \inf\{n > T_{2k-2} \,:\, X_n < \alpha\},$$

and

$$T_{2k} = \inf\{n > T_{2k-1} \,:\, X_n > \beta\}.$$

The *number of upcrossings of* $[\alpha, \beta]$ *by time* N is

$$U_N[\alpha, \beta] = \sup\{k \,:\, T_{2k} \leq N\}.$$

LEM 5.1 (Doob's Upcrossing Lemma) *Let* X *be a superMG. Then*

$$(\beta - \alpha)\mathbb{E}U_N[\alpha, \beta] \leq \mathbb{E}[(X_N - \alpha)^-].$$

Proof: Let $Y_n = (C \bullet X)_n$. Then Y_n is a superMG and satisfies

$$Y_N \geq (\beta - \alpha)U_N[\alpha, \beta] - (X_N - \alpha)^-,$$

since $(X_N - \alpha)^-$ overestimates the loss during the last interval of play. The result follows from $\mathbb{E}[Y_N] \leq 0$. ■

COR 5.2 *Let* X *be a superMG bounded in* $\mathcal{L}^1$*. Then*

$$U_N[\alpha, \beta] \uparrow U_\infty[\alpha, \beta],$$

$$(\beta - \alpha)\mathbb{E}U_\infty[\alpha, \beta] \leq |\alpha| + \sup_n \mathbb{E}|X_n| < +\infty,$$

so that

$$\mathbb{P}[U_\infty[\alpha, \beta] = \infty] = 0.$$

Proof: Use (MON). ■

3 Convergence theorem

THM 5.3 (Martingale convergence theorem) *Let* X *be a superMG bounded in* $\mathcal{L}^1$*. Then* X_n *converges and is finite a.s. Moreover, let* $X_\infty = \limsup_n X_n$ *then* $X_\infty \in \mathcal{F}_\infty$ *and* $\mathbb{E}|X_\infty| < +\infty$.

Proof: Let $\alpha < \beta \in \mathbb{Q}$ and

$$\Lambda_{\alpha,\beta} = \{\omega \,:\, \liminf X_n < \alpha < \beta < \limsup X_n\}.$$

Note that

$$\begin{aligned} \Lambda &= \{\omega \,:\, X_n \text{ does not converge}\} \\ &= \{\omega \,:\, \liminf X_n < \limsup X_n\} \\ &= \cup_{\alpha<\beta\in\mathbb{Q}} \Lambda_{\alpha,\beta}. \end{aligned}$$

Since

$$\Lambda_{\alpha,\beta} \subseteq \{U_\infty[\alpha, \beta] = \infty\},$$

we have $\mathbb{P}[\Lambda_{\alpha,\beta}] = 0$. By countability, $\mathbb{P}[\Lambda] = 0$. Use (FATOU) on $|X_n|$ to conclude. ■

COR 5.4 *If X is a nonnegative superMG then X_n converges a.s.*

Proof: X is bounded in $\mathcal{L}^1$ since

$$\mathbb{E}|X_n| = \mathbb{E}[X_n] \leq \mathbb{E}[X_0],\ \forall n.$$

■

EX 5.5 (Polya's Urn) *An urn contains 1 red ball and 1 green ball. At each time, we pick one ball and put it back with an extra ball of the same color. Let R_n (resp. G_n) be the number of red balls (resp. green balls) after the nth draw. Let $\mathcal{F}_n = \sigma(R_0, G_0, R_1, G_1, \ldots, R_n, G_n)$. Define M_n to be the fraction of green balls. Then*

$$\begin{aligned} \mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}] &= \frac{R_{n-1}}{G_{n-1}+R_{n-1}} \frac{G_{n-1}}{G_{n-1}+R_{n-1}+1} \\ &\quad + \frac{G_{n-1}}{G_{n-1}+R_{n-1}} \frac{G_{n-1}+1}{G_{n-1}+R_{n-1}+1} \\ &= \frac{G_{n-1}}{G_{n-1}+R_{n-1}} = M_{n-1}. \end{aligned}$$

Since $M_n \geq 0$ and is a MG, we have $M_n \to M_\infty$ a.s. See [Dur10, Section 4.3] for distribution of the limit and a generalization, or decipher,

$$\mathbb{P}[G_n = m+1] = \binom{n}{m} \frac{m!(n-m)!}{(n+1)!} = \frac{1}{n+1},$$

so that

$$\mathbb{P}[M_n \leq x] = \frac{\lfloor x(n+2) - 1 \rfloor}{n+1} \to x.$$

EX 5.6 (Convergence in L^1?) *We give an example that shows that the conditions of the Martingale Convergence Theorem do not guarantee convergence of expectations. Let $\{S_n\}$ be SRW started at 1 and*

$$T = \inf\{n > 0 \,:\, S_n = 0\}.$$

Then $\{S_{T \wedge n}\}$ is a nonnegative MG. It can only converge to 0 . But $\mathbb{E}[X_0] = 1 \neq 0$.

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 6 : Branching Processes	
MATH275B - Winter 2012	*Lecturer: Sebastien Roch*

References: [Wil91, Chapter 0], [Dur10, Section 4.3], [AN72, Section I.1 - I.5].

1 Branching processes

DEF 6.1 *A* branching process *is an SP of the form:*

- *Let $X(i,n)$, $i \geq 1$, $n \geq 1$, be an array of iid $\mathbb{Z}_+$-valued RVs with finite mean $m = \mathbb{E}[X(1,1)] < +\infty$, and inductively,*

$$Z_n = \sum_{1 \leq i \leq Z_{n-1}} X(i,n)$$

To avoid trivialities we assume $\mathbb{P}[X(1,1) = i] < 1$ for all $i \geq 0$.

LEM 6.2 *$M_n = m^{-n} Z_n$ is a nonnegative MG.*

Proof: Note that we have

$$\sum_j j \mathbb{P}[Z_n = j \,|\, Z_{n-1} = i] = mi,$$

so the claim follows from the eigenvector method. Alternatively, use the following lemma (proved in Hwk 1).

LEM 6.3 *If $Y_1 = Y_2$ a.s. on $B \in \mathcal{F}$ then $\mathbb{E}[Y_1 \,|\, \mathcal{F}] = \mathbb{E}[Y_2 \,|\, \mathcal{F}]$ a.s. on B.*

Then, on $\{Z_{n-1} = k\}$

$$\mathbb{E}[Z_n \,|\, \mathcal{F}_{n-1}] = \mathbb{E}[\sum_{1 \leq j \leq k} X(j,n) \,|\, \mathcal{F}_{n-1}] = mk = mZ_{n-1}.$$

This is true for all k. ■

COR 6.4 *$M_n \to M_\infty < +\infty$ a.s. and $\mathbb{E}[M_\infty] \leq 1$.*

2 Extinction

The martingale convergence theorem in itself tells us little about the limit. Here we try to give a more detailed picture of the limiting behavior—starting with extinction.

Let $p_i = \mathbb{P}[X(1,1) = i]$ for all i and for $s \in [0,1]$

$$f(s) = p_0 + p_1 s + p_2 s^2 + \cdots = \sum_{i \geq 0} p_i s^i.$$

Similarly, $f_n(s) = \mathbb{E}[s^{Z_n}]$. Ideally, we would like to compute the generating function of the limit—but this is rarely possible. Instead, we derive some of its properties. In particular, note that

$$\begin{aligned}
\pi &\equiv \mathbb{P}[Z_n = 0 \text{ for some } n \geq 0] \\
&= \lim_{n \to +\infty} \mathbb{P}[Z_n = 0] \\
&= \lim_{n \to +\infty} f_n(0),
\end{aligned}$$

using the fact that 0 is an absorbing state and monotonicity.

Moreover, by the Markov property, f_n as a natural recursive form:

$$\begin{aligned}
f_n(s) &= \mathbb{E}[s^{Z_n}] \\
&= \mathbb{E}[\mathbb{E}[s^{Z_n} \,|\, \mathcal{F}_{n-1}]] \\
&= \mathbb{E}[f(s)^{Z_{n-1}}] \\
&= f_{n-1}(f(s)) = \cdots = f^{(n)}(s).
\end{aligned}$$

So we need to study iterates of f.

We summarize the properties of f next. To make it easier, we assume $p_0 + p_1 < 1$.

LEM 6.5 *The function f on $[0,1]$ satisfies:*

1. $f(0) = p_0$, $f(1) = 1$
2. *f is indefinitely differentiable on $[0,1)$*
3. *f is strictly convex and increasing*
4. $\lim_{s \uparrow 1} f'(s) = m < +\infty$

Proof: 1. is clear by definition. The function f is a power series with radius of convergence $R \geq 1$. This implies 2. In particular,

$$f'(s) = \sum_{i \geq 1} i p_i s^{i-1} \geq 0,$$

and

$$f''(s) = \sum_{i \geq 2} i(i-1) p_i s^{i-2} > 0.$$

because we must have $p_i > 0$ for some $i > 1$ by assumption. This proves 3. Since $m < +\infty$, $f'(1)$ is well defined and f' is continuous on $[0,1]$. ■

COR 6.6 (Fixed points) *We have:*

1. *If $m > 1$ then f has a unique fixed point $\pi_0 \in [0,1)$*
2. *If $m \leq 1$ then $f(t) > t$ for $t \in [0,1)$ (Let $\pi_0 = 1$ in that case.)*

Proof: Since $f'(1) = m > 1$, there is $\delta > 0$ s.t. $f(1-\delta) < 1-\delta$. On the other hand $f(0) \geq 0$ so by continuity of f there must be a fixed point in $[0, 1-\delta)$. Moreover, by strict convexity, if r is a fixed point then $f(s) < s$ for $s \in (r, 1)$, proving uniqueness.

The second part follows by strict convexity and monotonicity. ■

COR 6.7 (Dynamics) *We have:*

1. *If $t \in [0, \pi_0)$, then $f^{(n)}(t) \uparrow \pi_0$*
2. *If $t \in (\pi_0, 1)$ then $f^{(n)}(t) \downarrow \pi_0$*

Proof: We only prove 1. The argument for 2. is similar. By monotonicity, for $t \in [0, \pi_0)$, we have $t < f(t) < f(\pi_0) = \pi_0$. Iterating

$$t < f^{(1)}(t) < \cdots < f^{(n)}(t) < f^{(n)}(\pi_0) = \pi_0.$$

So $f^{(n)}(t) \uparrow L \leq \pi_0$. By continuity of f we can take the limit inside of

$$f^{(n)}(t) = f(f^{(n-1)}(t)),$$

to get $L = f(L)$. So by definition of π_0 we must have $L = \pi_0$. ■

We immediately obtain:

THM 6.8 (Extinction) *The probability of extinction π is given by the smallest fixed point of f in $[0,1]$:*

1. *If $m \leq 1$ then $\pi = 1$.*
2. *If $m > 1$ then $\pi < 1$.*

References

[AN72] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York, 1972.

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 7 : Martingales bounded in L^2

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapters 0, 12], [Dur10, Section 4.4], [AN72, Section I.6].

1 Preliminaries

DEF 7.1 *For $1 \leq p < +\infty$, we say that $X \in \mathcal{L}^p$ if*

$$\|X\|_p = \mathbb{E}[|X^p|]^{1/p} < +\infty.$$

By Jensen's inequality, for $1 \leq p \leq r < +\infty$ we have $\|X\|_p \leq \|X\|_r$ if $X \in \mathcal{L}^r$.

Proof: For $n \geq 0$, let

$$X_n = (|X| \wedge n)^p.$$

Take $c(x) = x^{r/p}$ on $(0, +\infty)$ which is convex. Then

$$(\mathbb{E}[X_n])^{r/p} \leq \mathbb{E}[(X_n)^{r/p}] = \mathbb{E}[(|X| \wedge n)^r] \leq \mathbb{E}[|X|^r].$$

Take $n \to \infty$ and use (MON). ■

DEF 7.2 *We say that X_n converges to X_∞ in $\mathcal{L}^p$ if $\|X_n - X_\infty\|_p \to 0$. By the previous result, convergence on $\mathcal{L}^r$ implies convergence in $\mathcal{L}^p$ for $r \geq p \geq 1$.*

LEM 7.3 *Assume $X_n, X_\infty \in \mathcal{L}^1$. Then*

$$\|X_n - X_\infty\|_1 \to 0,$$

implies

$$\mathbb{E}[X_n] \to \mathbb{E}[X_\infty].$$

Proof: Note that

$$|\mathbb{E}[X_n] - \mathbb{E}[X_\infty]| \leq \mathbb{E}|X_n - X_\infty| \to 0.$$

■

DEF 7.4 *We say that $\{X_n\}_n$ is* bounded in $\mathcal{L}^p$ *if*

$$\sup_n \|X_n\|_p < +\infty.$$

2 L^2 convergence

THM 7.5 *Let M be a MG with $M_n \in \mathcal{L}^2$. Then M is bounded in $\mathcal{L}^2$ if and only if*

$$\sum_{k\geq 1} \mathbb{E}[(M_k - M_{k-1})^2] < +\infty.$$

When this is the case, M_n converges a.s. and in $\mathcal{L}^2$.

Proof:

LEM 7.6 (Orthogonality of increments) *Let $s \leq t \leq u \leq v$. Then,*

$$\langle M_t - M_s, M_v - M_u \rangle = 0.$$

Proof: Use $M_u = \mathbb{E}[M_v \,|\, \mathcal{F}_u]$, $M_t - M_s \in \mathcal{F}_u$ and apply the L^2 characterization of conditional expectations. ■

That implies

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_0^2] + \sum_{1\leq i\leq n} \mathbb{E}[(M_i - M_{i-1})^2],$$

proving the first claim.

By monotonicity of norms, M is bounded in L^2 implies M bounded in L^1 which, in turn, implies M converges a.s. Then using (FATOU) in

$$\mathbb{E}[(M_{n+k} - M_n)^2] = \sum_{n+1\leq i\leq n+k} \mathbb{E}[(M_i - M_{i-1})^2],$$

gives

$$\mathbb{E}[(M_\infty - M_n)^2] \leq \sum_{n+1\leq i} \mathbb{E}[(M_i - M_{i-1})^2].$$

The RHS goes to 0 which proves the second claim.

■

3 Back to branching processes

THM 7.7 *Let Z be a branching process with $Z_0 = 1$, $m = \mathbb{E}[X(1,1)] > 1$ and $\sigma^2 = \mathrm{Var}[X(1,1)] < +\infty$. Then, $M_n = m^{-n} Z_n$ converges in L^2, and in particular, $\mathbb{E}[M_\infty] = 1$.*

Proof: From the orthogonality of increments

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_{n-1}^2] + \mathbb{E}[(M_n - M_{n-1})^2].$$

On $\{Z_{n-1} = k\}$

$$\begin{aligned}
\mathbb{E}[(M_n - M_{n-1})^2 \,|\, \mathcal{F}_{n-1}] &= m^{-2n}\mathbb{E}[(Z_n - mZ_{n-1})^2 \,|\, \mathcal{F}_{n-1}] \\
&= m^{-2n}\mathbb{E}[(\sum_{i=1}^{k} X(i,n) - mk)^2 \,|\, \mathcal{F}_{n-1}] \\
&= m^{-2n}k\sigma^2 \\
&= m^{-2n}Z_{n-1}\sigma^2.
\end{aligned}$$

Hence

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_{n-1}^2] + m^{-n-1}\sigma^2.$$

Since $\mathbb{E}[M_0^2] = 1$,

$$\mathbb{E}[M_n^2] = 1 + \sigma^2 \sum_{i=2}^{n+1} m^{-i},$$

which is uniformly bounded when $m > 1$. So M_n converges in L^2. Finally by (FATOU)

$$\mathbb{E}|M_\infty| \leq \sup \|M_n\|_1 \leq \sup \|M_n\|_2 < +\infty$$

and

$$|\mathbb{E}[M_n] - \mathbb{E}[M_\infty]| \leq \|M_n - M_\infty\|_1 \leq \|M_n - M_\infty\|_2,$$

implies the convergence of expectations. ■

In a homework problem, we will show that under the assumptions of the previous theorem

$$\{M_\infty = 0\} = \{Z_n = 0, \text{ for some } n\},$$

and

$$\mathbb{P}[M_\infty = 0] = \pi,$$

the probability of extinction.

EX 7.8 (Geometric Offspring) *Assume*

$$0 < p < 1,\ q = 1 - p,\ p_i = pq^i,\ \forall i \geq 0,\ m = \frac{q}{p}.$$

Then

$$f(s) = \frac{p}{1 - sq},\ \pi = \min\{\frac{p}{q}, 1\}.$$

- *Case $m \neq 1$.* *If G is a 2×2 matrix, denote*

$$G(s) = \frac{G_{11}s + G_{12}}{G_{21}s + G_{22}}.$$

Then $G(H(s)) = (GH)(s)$. By diagonalization,

$$\begin{pmatrix} 0 & p \\ -q & 1 \end{pmatrix}^n = (q-p)^{-1} \begin{pmatrix} 1 & p \\ 1 & q \end{pmatrix} \begin{pmatrix} p^n & 0 \\ 0 & q^n \end{pmatrix} \begin{pmatrix} q & -p \\ -1 & 1 \end{pmatrix}$$

leading to

$$f_n(s) = \frac{pm^n(1-s) + qs - p}{qm^n(1-s) + qs - p}.$$

In particular, when $m < 1$ we have $\pi = \lim f_n(0) = 1$. On the other hand, if $m > 1$, we have by (DOM) for $\lambda \geq 0$

$$\begin{aligned} \mathbb{E}[\exp(-\lambda M_\infty)] &= \lim_n f_n(\exp(-\lambda/m^n)) \\ &= \frac{p\lambda + q - p}{q\lambda + q - p} \\ &= \pi + (1-\pi)\frac{(1-\pi)}{\lambda + (1-\pi)}. \end{aligned}$$

The first term corresponds to a point mass at 0 and the second term corresponds to an exponential with mean $1/(1-\pi)$.

- *Case $m = 1$.* *By induction*

$$f_n(s) = \frac{n - (n-1)s}{n + 1 - ns},$$

so that

$$\mathbb{P}[Z_n > 0] = 1 - f_n(0) = \frac{1}{n+1},$$

and

$$\mathbb{E}[e^{-\lambda Z_n/n} \mid Z_n > 0] = \frac{f_n(e^{-\lambda/n}) - f_n(0)}{1 - f_n(0)} \to \frac{1}{1+\lambda},$$

which is the Laplace transform of an eponential mean 1. This is consistent with $\mathbb{E}[Z_n] = 1$.

References

[AN72] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York, 1972.

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 8 : MGs in L^p

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 12], [Dur10, Section 4.4].

1 L^p convergence theorem

Recall:

LEM 8.1 (Markov's inequality) *Let $Z \geq 0$ be a RV. Then for $c > 0$*

$$c\mathbb{P}[Z \geq c] \leq \mathbb{E}[Z; Z \geq c] \leq \mathbb{E}[Z].$$

MGs provide a useful generalization.

LEM 8.2 (Doob's submartingale inequality) *Let $Z \geq 0$ a subMG. Then for $c > 0$*

$$c\mathbb{P}[\sup_{1\leq k\leq n} Z_k \geq c] \leq \mathbb{E}[Z_n; \sup_{1\leq k\leq n} Z_k \geq c] \leq \mathbb{E}[Z_n].$$

Proof: Divide $F = \{\sup_{1\leq k\leq n} Z_k \geq c\}$ according to the first time Z crosses c:

$$F = F_0 \cup \cdots \cup F_n,$$

where

$$F_k = \{Z_0 < c\} \cap \cdots \cap \{Z_{k-1} < c\} \cap \{Z_k \geq c\}.$$

Since $F_k \in \mathcal{F}_k$ and $\mathbb{E}[Z_n \,|\, \mathcal{F}_k] \geq Z_k$,

$$c\mathbb{P}[F_k] \leq \mathbb{E}[Z_k; F_k] \leq \mathbb{E}[Z_n; F_k].$$

Sum over k. ■

EX 8.3 (Kolmogorov's inequality) *Let $X_1, \ldots$ be independent RVs with $\mathbb{E}[X_k] = 0$ and $\mathrm{Var}[X_k] < +\infty$. Define $S_n = \sum_{k\leq n} X_k$. Then for $c > 0$*

$$\mathbb{P}[\max_{k\leq n} |S_k| \geq c] \leq c^{-2}\mathrm{Var}[S_n].$$

THM 8.4 (Doob's L^p inequality) *Let $p > 1$ and $p^{-1} + q^{-1} = 1$. Let $Z \geq 0$ a subMG bounded in L^p. Define*

$$Z^* = \sup_{k \geq 0} Z_k.$$

Then

$$\|Z^*\|_p \leq q \sup_k \|Z_k\|_p = q \uparrow \lim_k \|Z_k\|_p.$$

and $Z^* \in L^p$.

Proof: The last equality follows from (JENSEN). Let $Z_n^* = \sup_{k \leq n} Z_k$. By (MON) it suffices to prove:

LEM 8.5

$$\mathbb{E}[(Z_n^*)^p] \leq q^p \mathbb{E}[Z_n^p].$$

Proof: Recall the formula: for $Y \geq 0$ and $p > 0$

$$\mathbb{E}[Y^p] = \int_0^\infty p y^{p-1} \mathbb{P}[Y \geq y] dy.$$

Then for $K > 0$

$$\begin{aligned}
\mathbb{E}[(Z_n^* \wedge K)^p] &= \int_0^\infty p c^{p-1} \mathbb{P}[Z_n^* \wedge K \geq c] dc \\
&\leq \int_0^\infty p c^{p-2} \mathbb{E}[Z_n; Z_n^* \wedge K \geq c] dc \\
&= \mathbb{E}\left[Z_n \left(\frac{p}{p-1} \right) \int_0^\infty (p-1) c^{p-2} \mathbb{P}[Z_n^* \wedge K \geq c] dc \right] \\
&= \mathbb{E}[q Z_n (Z_n^* \wedge K)^{p-1}] \\
&\leq q \mathbb{E}[Z_n^p]^{1/p} \mathbb{E}[(Z_n^* \wedge K)^p]^{1/q}.
\end{aligned}$$

Rearranging and using (MON) gives the result. ■ ■

THM 8.6 (L^p convergence) *Let M be a MG bounded in L^p for $p > 1$. Then $M_n \to M_\infty$ a.s. and in L^p.*

Proof: Note that $|M_n|$ is a subMG bounded in L^p. In particular, it is bounded in L^1 and $M_n \to M_\infty$ a.s. From the previous theorem,

$$|M_n - M_\infty|^p \leq (2 \sup_k |M_k|)^p \in L^1,$$

and by (DOM)

$$\mathbb{E}|M_n - M_\infty|^p \to 0.$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 9 : Martingales in L^2 (continued)

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 12], [Dur10, Section 4.4].

1 Review: Random series

Recall:

THM 9.1 (Three-Series Thm) *Let $\{X_n\}$ be independent. For $K > 0$, let $Y_n = X_n \mathbb{1}\{|X_n| \leq K\}$. Then $\sum_n X_n$ converges a.s. if and only if:*

1. $\sum_n \mathbb{P}[|X_n| > K] < +\infty$
2. $\sum_n \mathbb{E}[Y_n]$ *converges*
3. $\sum_n \mathrm{Var}[Y_n] < +\infty$

We will see a MG generalization of this result.

2 Angle-brackets process

THM 9.2 (Doob decomposition) *Let X be an adapted process in L^1. Then*

- *X has an a.s. unique decomposition*

$$X = X_0 + M + A, \qquad (*)$$

  where M is a MG and A is predictable with $M_0 = A_0 = 0$.

- *X is a subMG if and only if $A_n \uparrow$ a.s.*

Proof: Suppose $(*)$ holds. Observe

$$\mathbb{E}[X_n - X_{n-1} \,|\, \mathcal{F}_{n-1}] = \mathbb{E}[M_n - M_{n-1} \,|\, \mathcal{F}_{n-1}] + \mathbb{E}[A_n - A_{n-1} \,|\, \mathcal{F}_{n-1}] = A_n - A_{n-1},$$

so that

$$A_n = \sum_{k \leq n} \mathbb{E}[X_k - X_{k-1} \,|\, \mathcal{F}_{k-1}].$$

This proves uniqueness—that is, if there is a decomposition such that M is a MG then A has to be of the previous form. Using this equation as definition gives first claim—by the same equation, M will be a MG. Second claim is now obvious. ■

LEM 9.3 *If M is a MG and ϕ is convex with $\mathbb{E}[|\phi(M_n)|] < +\infty$, then $\phi(M_n)$ is a subMG.*

Proof: Using (cJENSEN)

$$\mathbb{E}[\phi(M_n) \,|\, \mathcal{F}_{n-1}] \geq \phi(\mathbb{E}[M_n \,|\, \mathcal{F}_{n-1}]) = \phi(M_{n-1}).$$

■

DEF 9.4 (Angle-brackets process) *Let M be a MG in $\mathcal{L}^2$ with $M_0 = 0$. Then M^2 is a subMG with decomposition*

$$M^2 \equiv N + \langle M \rangle,$$

where $\langle M \rangle_n \uparrow$ a.s. Moreover M is bounded in L^2 if and only if $\mathbb{E}[\langle M \rangle_\infty] < \infty$. Finally note

$$\langle M \rangle_n = \sum_k \mathbb{E}[M_k^2 - M_{k-1}^2 \,|\, \mathcal{F}_{k-1}] = \sum_k \mathbb{E}[(M_k - M_{k-1})^2 \,|\, \mathcal{F}_{k-1}].$$

We finally come to our main theorem.

THM 9.5 *Let M be a MG in L^2. Then*

1. $\lim_n M_n(\omega)$ *exists for a.e. ω s.t. $\langle M \rangle_\infty < \infty$.*
2. *If further $|M_n - M_{n-1}| \leq K$ a.s. $\forall n$ then $\langle M \rangle_\infty(\omega) < +\infty$ for a.e. ω s.t.* $\lim_n M_n(\omega)$ *exists.*

Proof: Proof of 1. Observe that

$$\{\langle M \rangle_\infty < \infty\} = \cup_k \{S(k) = +\infty\},$$

where

$$S(k) = \inf\{n \,:\, \langle M \rangle_{n+1} > k\},$$

defines a stopping time. It suffices to prove:

LEM 9.6 $\langle M^{S(k)} \rangle = \langle M \rangle^{S(k)}$.

Indeed, $\mathbb{E}[\langle M\rangle^{S(k)}] \leq k < +\infty$, hence $\mathbb{E}[\langle M^{S(k)}\rangle] < +\infty$ and the MG $M^{S(k)}$ is bounded in L^2:

$$\lim_n M_n^{S(k)} \text{ exists a.s.}$$

Since $S(k) = +\infty$ for some k we have proved the first claim. It remains to prove the lemma. Note that

$$(M^2 - \langle M\rangle)^{S(k)} = (M^{S(k)})^2 - \langle M\rangle^{S(k)},$$

is a MG. By the uniqueness of Doob's decomposition, it suffices to show that $\langle M\rangle^{S(k)}$ is predictable. Let $B \in \mathcal{B}$. Then

$$\{\langle M\rangle_n^{S(k)} \in B\} = E_1 \cup E_2,$$

where

$$E_1 = \cup_{1\leq r\leq n-1}\{S(k) = r,\ \langle M\rangle_r \in B\} \in \mathcal{F}_{n-1},$$

and

$$E_2 = \{S(k) \leq n-1\}^c \cap \{\langle M\rangle_n \in B\} \in \mathcal{F}_{n-1}.$$

That concludes the proof of the first claim.

Proof of 2. (Sketch.) Proof is similar. Enough to prove that $\sup_n |M_n(\omega)| < +\infty$ implies $\langle M\rangle_\infty < +\infty$ a.s. Observe

$$\{\sup_n |M_n(\omega)| < +\infty\} = \cup_c\{T(c) = +\infty\},$$

where

$$T(c) = \inf\{n\ :\ |M_n| > c\},$$

defines a stopping time. By the above lemma,

$$\mathbb{E}[(M_n^{T(c)})^2 - \langle M\rangle_n^{T(c)}] = 0,$$

so that

$$\mathbb{E}[\langle M\rangle_n^{T(c)}] \leq (c+K)^2.$$

Since $T(c) = +\infty$ for some c , this proves the second claim.

■

3 Applications

THM 9.7 (A strong law for MGs in L^2) *Let M be a MG in $\mathcal{L}^2$ with $M_0 = 0$. Then*

$$\frac{M_n}{\langle M \rangle_n} \to 0, \qquad \text{a.s. on } \{\langle M \rangle_\infty = +\infty\}.$$

Proof: Note that $(1 + \langle M \rangle)^{-1}$ is bounded and predictable so that

$$W_n = ((1 + \langle M \rangle)^{-1} \bullet M)_n = \sum_{k=1}^{n} \frac{M_k - M_{k-1}}{1 + \langle M \rangle_k},$$

is a MG. Note that

$$\begin{aligned}
&\mathbb{E}[(W_n - W_{n-1})^2 \,|\, \mathcal{F}_{n-1}] \\
&\quad = (1 + \langle M \rangle_n)^{-2} \mathbb{E}[(M_n - M_{n-1})^2 \,|\, \mathcal{F}_{n-1}] \\
&\quad = (1 + \langle M \rangle_n)^{-2} (\langle M \rangle_n - \langle M \rangle_{n-1}) \\
&\quad \leq (1 + \langle M \rangle_{n-1})^{-1} (1 + \langle M \rangle_n)^{-1} ((1 + \langle M \rangle_n) - (1 + \langle M \rangle_{n-1})) \\
&\quad = (1 + \langle M \rangle_{n-1})^{-1} - (1 + \langle M \rangle_n)^{-1}.
\end{aligned}$$

In particular, $\langle W \rangle_\infty \leq 1 < +\infty$ so that W_n converges a.s.

LEM 9.8 (Kronecker's Lemma) *If $b_n \uparrow +\infty$ then*

$$\sum_n \frac{x_n}{b_n} \text{ converges} \qquad \Longrightarrow \qquad \frac{\sum_n x_n}{b_n} \to 0.$$

Then on $\{\langle M \rangle_\infty = +\infty\}$, we have $M_n/(1 + \langle M \rangle_n) \to 0$ and the result follows. ∎

THM 9.9 (Levy's extension of Borel-Cantelli) *Suppose $\mathbb{1}_{E_k}$ is adapted. Define*

$$Z_n = \sum_{k=1}^{n} \mathbb{1}_{E_k},$$

and

$$Y_n = \sum_{k=1}^{n} \mathbb{P}[E_k \,|\, \mathcal{F}_{k-1}].$$

Then

1. $Y_\infty < \infty \implies Z_\infty < \infty$
2. $Y_\infty = +\infty \implies Z_n/Y_n \to 1$

Note that the previous theorem implies the classical BC lemmas. For 1, note that $\mathbb{E}[Y_\infty] = \sum_k \mathbb{P}[E_k]$. For 2, note that by independence $\mathbb{P}[E_k \,|\, \mathcal{F}_{k-1}] = \mathbb{P}[E_k]$.
Proof: Z is a subMG, Y is predictable and $M = Z - Y$ is a MG. The proof relies on computing $\langle M \rangle$. Note

$$\begin{aligned}\langle M \rangle_n &= \sum_{k=1}^n \mathbb{E}[(M_k - M_{k-1})^2 \,|\, \mathcal{F}_{k-1}] \\ &= \sum_{k=1}^n \mathbb{E}[(\mathbb{1}_{E_k} - \mathbb{P}[E_k \,|\, \mathcal{F}_{k-1}])^2 \,|\, \mathcal{F}_{k-1}] \\ &= \sum_{k=1}^n \mathbb{E}[\mathbb{1}_{E_k} - \mathbb{P}[E_k \,|\, \mathcal{F}_{k-1}]^2 \,|\, \mathcal{F}_{k-1}] \\ &= \sum_{k=1}^n [\mathbb{P}[E_k \,|\, \mathcal{F}_{k-1}] - \mathbb{P}[E_k \,|\, \mathcal{F}_{k-1}]^2] \\ &\leq Y_n.\end{aligned}$$

We are ready to prove the statements.

1. $\boxed{Y_\infty < +\infty.}$ Then $\langle M \rangle_\infty < +\infty$ and M_n converges. Hence, $Z = M + Y$ also converges.

2. $\boxed{Y_\infty = +\infty.}$ Assume first that $\langle M \rangle_\infty < +\infty$. Then M_n converges and

$$\frac{Z_n}{Y_n} = \frac{M_n + Y_n}{Y_n} \to 1.$$

   On the other hand, if $\langle M \rangle_\infty = +\infty$ the strong law for L^2 MGs gives $M_n / \langle M \rangle_n \to 0$ so that $M_n / Y_n \to 0$ and $Z_n / Y_n \to 1$.

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 10 : Uniform integrability	
MATH275B - Winter 2012	*Lecturer: Sebastien Roch*

References: [Wil91, Chapter 13], [Dur10, Section 4.5].

1 Uniform Integrability

LEM 10.1 *Let* $Y \in L^1$. $\forall \varepsilon > 0$, $\exists K > 0$ *s.t.*

$$\mathbb{E}[|Y|; |Y| > K] < \varepsilon.$$

Proof: Immediate by (MON) to $\mathbb{E}[|Y|; |Y| \leq K]$. ■

DEF 10.2 (Uniform Integrability) *A collection* $\mathcal{C}$ *of RVs on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is uniformly integrable (UI) if:* $\forall \varepsilon > 0$, $\exists K > +\infty$ *s.t.*

$$\mathbb{E}[|X|; |X| > K] < \varepsilon, \qquad \forall X \in \mathcal{C}.$$

THM 10.3 (Necessary and Sufficient Condition for L^1 Convergence) *Let* $\{X_n\} \in L^1$ *and* $X \in L^1$. *Then* $X_n \to X$ *in* L^1 *if and only if:*

- $X_n \to X$ *in prob*
- $\{X_n\}$ *is UI.*

Before giving the proof, we look at a few examples.

EX 10.4 (L^1-bddness is not sufficient) *Let* $\mathcal{C}$ *is UI and* $X \in \mathcal{C}$. *Note that*

$$\mathbb{E}|X| \leq \mathbb{E}[|X|; |X| \geq K] + \mathbb{E}[|X|; |X| < K] \leq \varepsilon + K < +\infty,$$

so UI implies L^1*-bddness. But the opposite is not true by our last example.*

EX 10.5 (L^p-bdd RVs) *Let* $\mathcal{C}$ *be* L^p*-bdd and* $X \in \mathcal{C}$. *Then*

$$\mathbb{E}[|X|; |X| > K] \leq \mathbb{E}[K^{1-p}|X|^p; |X| > K|] \leq K^{1-p}A \to 0,$$

as $K \to +\infty$.

EX 10.6 (Dominated RVs) *Assume* $\exists Y \in L^1$ *s.t.* $|X| \leq Y$ $\forall X \in \mathcal{C}$. *Then*

$$\mathbb{E}[|X|; |X| > K] \leq \mathbb{E}[Y; |X| > K] \leq \mathbb{E}[Y; Y > K],$$

and apply lemma above.

2 Proof of main theorem

Proof: We start with the if part. Fix $\varepsilon > 0$. We want to show that for n large enough:

$$\mathbb{E}|X_n - X| \leq \varepsilon.$$

Let $\phi_K(x) = \mathrm{sgn}(x)[|x| \wedge K]$. Then,

$$\begin{aligned}\mathbb{E}|X_n - X| &\leq \mathbb{E}|\phi_K(X_n) - X_n| + \mathbb{E}|\phi_K(X) - X| + \mathbb{E}|\phi_K(X_n) - \phi_K(X)| \\ &\leq \mathbb{E}[|X_n|; |X_n| > K] + \mathbb{E}[|X|; |X| > K] + \mathbb{E}|\phi_K(X_n) - \phi_K(X)|.\end{aligned}$$

1st term $\leq \varepsilon/3$ by UI and 2nd term $\leq \varepsilon/3$ by lemma above. Check, by case analysis, that

$$|\phi_K(x) - \phi_K(y)| \leq |x - y|,$$

so $\phi_K(X_n) \to_P \phi_K(X)$. By bounded convergence for convergence in probability, the claim is proved.

LEM 10.7 (Bounded convergence theorem (convergence in probability version)) *Let* $X_n \leq K < +\infty$ $\forall n$ *and* $X_n \to_P X$. *Then*

$$\mathbb{E}|X_n - X| \to 0.$$

Proof:(Sketch) By

$$\mathbb{P}[|X| \geq K + m^{-1}] \leq \mathbb{P}[|X_n - X| \geq m^{-1}],$$

it follows that $\mathbb{P}[|X| \leq K] = 1$. Fix $\varepsilon > 0$

$$\begin{aligned}\mathbb{E}|X_n - X| &= \mathbb{E}[|X_n - X|; |X_n - X| > \varepsilon/2] + \mathbb{E}[|X_n - X|; |X_n - X| \leq \varepsilon/2] \\ &\leq 2K\mathbb{P}[|X_n - X| > \varepsilon/2] + \varepsilon/2 < \varepsilon,\end{aligned}$$

for n large enough. ■

Proof of only if part. Suppose $X_n \to X$ in L^1. We know that L^1 implies convergence in probability. So the first claim follows.

For the second claim, if $n \geq N$ (large enough),

$$\mathbb{E}|X_n - X| \leq \varepsilon.$$

We can choose K large enough so that

$$\mathbb{E}[|X_n|; |X_n| > K] < \varepsilon,$$

$\forall n \leq N$. So only need to worry about $n > N$. To use L^1 convergence, natural to write

$$\mathbb{E}[|X_n|; |X_n| > K] \leq \mathbb{E}[|X_n - X|; |X_n| > K] + \mathbb{E}[|X|; |X_n| > K].$$

First term $\leq \varepsilon$. The issue with the second term is that we cannot apply the lemma because the event involves X_n rather than X. In fact, a stronger version exists:

LEM 10.8 (Absolute continuity) *Let $X \in L^1$. $\forall \varepsilon > 0$, $\exists \delta > 0$, s.t. $\mathbb{P}[F] < \delta$ implies*

$$\mathbb{E}[|X|; F] < \varepsilon.$$

Proof: Argue by contradiction. Suppose there is ε and F_n s.t. $\mathbb{P}[F_n] \leq 2^{-n}$ and

$$\mathbb{E}[|X|; F_n] \geq \varepsilon.$$

By BC,

$$\mathbb{P}[H] \equiv \mathbb{P}[F_n \text{ i.o.}] = 0.$$

By (DOM) ,

$$\mathbb{E}[|X|; H] \geq \varepsilon,$$

a contradiction. ■

To conclude note that

$$\mathbb{P}[|X_n| > K] \leq \frac{\mathbb{E}|X_n|}{K} \leq \frac{\sup_{n \geq N} \mathbb{E}|X_n|}{K} \leq \frac{\sup_{n \geq N} \mathbb{E}|X| + \mathbb{E}|X_n - X|}{K} < \delta,$$

uniformly in n for K large enough. We are done.

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 11 : UI MGs

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 14], [Dur10, Section 4.5].

1 UI MGs

THM 11.1 (Convergence of UI MGs) *Let M be UI MG. Then*

$$M_n \to M_\infty,$$

a.s. and in L^1. Moreover,

$$M_n = \mathbb{E}[M_\infty \,|\, \mathcal{F}_n], \qquad \forall n.$$

Proof: UI implies L^1-bddness so we have $M_n \to M_\infty$ a.s. By necessary and sufficient condition, we also have L^1 convergence.

Now note that for all $r \geq n$ and $F \in \mathcal{F}_n$, we know $\mathbb{E}[M_r \,|\, \mathcal{F}_n] = M_n$ or

$$\mathbb{E}[M_r; F] = \mathbb{E}[M_n; F],$$

by definition of CE. We can take a limit by L^1 convergence. More precisely

$$|\mathbb{E}[M_r; F] - \mathbb{E}[M_\infty; F]| \leq \mathbb{E}[|M_r - M_\infty|; F] \leq \mathbb{E}[|M_r - M_\infty|] \to 0,$$

as $r \to \infty$. So plugging above

$$\mathbb{E}[M_\infty; F] = \mathbb{E}[M_n; F],$$

and $\mathbb{E}[M_\infty \,|\, \mathcal{F}_n] = M_n$. ■

2 Applications I

THM 11.2 (Levy's upward thm) *Let $Z \in L^1$ and define $M_n = \mathbb{E}[Z \,|\, \mathcal{F}_n]$. Then M is a UI MG and*

$$M_n \to M_\infty = \mathbb{E}[Z \,|\, \mathcal{F}_\infty],$$

a.s. and in L^1.

Proof: M is a MG by (TOWER). We first show it is UI:

LEM 11.3 *Let* $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. *Then*

$$\{\mathbb{E}[X \,|\, \mathcal{G}] \,:\, \mathcal{G} \textit{ is a sub-}\sigma\textit{-field of } \mathcal{F}\},$$

is UI.

Proof: We use the absolute continuity lemma again. Let $Y = \mathbb{E}[X \,|\, \mathcal{G}] \in \mathcal{G}$. Since $\{|Y| > K\} \in \mathcal{G}$,

$$\begin{aligned} \mathbb{E}[|Y|; |Y| > K] &= \mathbb{E}[|\mathbb{E}[X \,|\, \mathcal{G}]|; |Y| > K] \\ &\leq \mathbb{E}[\mathbb{E}[|X| \,|\, \mathcal{G}]; |Y| > K] \\ &= \mathbb{E}[|X|; |Y| > K]. \end{aligned}$$

By Markov

$$\mathbb{P}[|Y| > K] \leq \frac{\mathbb{E}|Y|}{K} \leq \frac{\mathbb{E}|X|}{K} \leq \delta,$$

for K large enough (uniformly in $\mathcal{G}$). And we are done. ■

In particular, we have convergence a.s. and in L^1 to $M_\infty \in \mathcal{F}_\infty$.

Let $Y = \mathbb{E}[Z \,|\, \mathcal{F}_\infty] \in \mathcal{F}_\infty$. By dividing into negative and positive parts, we assume $Z \geq 0$. We want to show, for $F \in \mathcal{F}_\infty$,

$$\mathbb{E}[Z; F] = \mathbb{E}[M_\infty; F].$$

By Uniqueness Lemma, it suffices to prove equality for all $\mathcal{F}_n$. If $F \in \mathcal{F}_n \subseteq \mathcal{F}_\infty$, then by (TOWER)

$$\mathbb{E}[Z; F] = \mathbb{E}[Y; F] = \mathbb{E}[M_n; F] = \mathbb{E}[M_\infty; F].$$

■

THM 11.4 (Levy's $0-1$ law) *Let* $A \in \mathcal{F}_\infty$. *Then*

$$\mathbb{P}[A \,|\, \mathcal{F}_n] \to \mathbb{1}_A.$$

Proof: Immediate. ■

COR 11.5 (Kolmogorov's $0-1$ law) *Let* $X_1, X_2, \ldots$ *be iid RVs. Recall that the tail* σ*-field is*

$$\mathcal{T} = \cap_n \mathcal{T}_n = \cap_n \sigma(X_{n+1}, X_{n+2}, \ldots).$$

If $A \in \mathcal{T}$ *then* $\mathbb{P}[A] \in \{0, 1\}$.

Proof: Since $A \in \mathcal{T}_n$ is independent of $\mathcal{F}_n$,

$$\mathbb{P}[A \,|\, \mathcal{F}_n] = \mathbb{P}[A],$$

$\forall n$. By Levy's law,

$$\mathbb{P}[A] = \mathbb{1}_A \in \{0, 1\}.$$

■

3 Applications II

THM 11.6 (Levy's Downward Thm) *Let $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\mathcal{G}_{-n}\}_{n\geq 0}$ a collection of σ-fields s.t.*

$$\mathcal{G}_{-\infty} = \cap_k \mathcal{G}_{-k} \subseteq \cdots \subseteq \mathcal{G}_{-n} \subseteq \cdots \subseteq \mathcal{G}_{-1} \subseteq \mathcal{F}.$$

Define

$$M_{-n} = \mathbb{E}[Z \,|\, \mathcal{G}_{-n}].$$

Then

$$M_{-n} \to M_{-\infty} = \mathbb{E}[Z \,|\, \mathcal{G}_{-\infty}]$$

a.s. and in L^1.

Proof: We apply the same argument as in the Martingale Convergence Thm. Let $\alpha < \beta \in \mathbb{Q}$ and

$$\Lambda_{\alpha,\beta} = \{\omega \,:\, \liminf X_{-n} < \alpha < \beta < \limsup X_{-n}\}.$$

Note that

$$\begin{aligned} \Lambda &\equiv \{\omega \,:\, X_n \text{ does not converge}\} \\ &= \{\omega \,:\, \liminf X_{-n} < \limsup X_{-n}\} \\ &= \cup_{\alpha<\beta\in\mathbb{Q}} \Lambda_{\alpha,\beta}. \end{aligned}$$

Let $U_N[\alpha, \beta]$ be the *number of upcrossings of* $[\alpha, \beta]$ *between time* $-N$ *and* -1. Then by the Upcrossing Lemma applied to the MG $M_{-N}, \ldots, M_{-1}$

$$(\beta - \alpha)\mathbb{E}U_N[\alpha, \beta] \leq |\alpha| + \mathbb{E}|M_{-1}| \leq |\alpha| + \mathbb{E}|Z|.$$

By (MON)

$$U_N[\alpha, \beta] \uparrow U_\infty[\alpha, \beta],$$

and

$$(\beta - \alpha)\mathbb{E}U_\infty[\alpha, \beta] \leq |\alpha| + \mathbb{E}|Z| < +\infty,$$

so that

$$\mathbb{P}[U_\infty[\alpha, \beta] = \infty] = 0.$$

Since

$$\Lambda_{\alpha,\beta} \subseteq \{U_\infty[\alpha, \beta] = \infty\},$$

we have $\mathbb{P}[\Lambda_{\alpha,\beta}] = 0$. By countability, $\mathbb{P}[\Lambda] = 0$. Therefore we have convergence a.s.

By lemma in previous class, M is UI and hence we have L^1 convergence as well.

Finally, for all $G \in \mathcal{G}_{-\infty} \subseteq \mathcal{G}_{-n}$,

$$\mathbb{E}[Z; G] = \mathbb{E}[M_{-n}; G].$$

Take the limit $n \to +\infty$ and use L^1 convergence. ■

An application:

THM 11.7 (Strong Law; Martingale Proof) *Let $X_1, X_2, \ldots$ be iid RVs with $\mathbb{E}[X_1] = \mu$ and $\mathbb{E}|X_1| < +\infty$. Let $S_n = \sum_{i \leq n} X_n$. Then*

$$n^{-1}S_n \to \mu,$$

a.s. and in L^1.

Proof: Let

$$\mathcal{G}_{-n} = \sigma(S_n, S_{n+1}, S_{n+2}, \ldots) = \sigma(S_n, X_{n+1}, X_{n+2}, \ldots),$$

and note that, for $1 \leq i \leq n$,

$$\mathbb{E}[X_1 \,|\, \mathcal{G}_{-n}] = \mathbb{E}[X_1 \,|\, S_n] = \mathbb{E}[X_i \,|\, S_n] = \mathbb{E}[n^{-1}S_n \,|\, S_n] = n^{-1}S_n,$$

by symmetry. By Levy's Downward Thm

$$n^{-1}S_n \to \mathbb{E}[X_1 \,|\, \mathcal{G}_{-\infty}],$$

a.s. and in L^1. Note that $\mathcal{G}_{-n} \subseteq \mathcal{E}_n$ and $\mathcal{G}_{-\infty} \subseteq \mathcal{E}$ so that $\mathcal{G}_{-\infty}$ is trivial and we must have $\mathbb{E}[X_1 \,|\, \mathcal{G}_{-\infty}] = \mu$. ■

4 Further material

DEF 11.8 *Let $X_1, X_2, \ldots$ be iid RVs. Let $\mathcal{E}_n$ be the σ-field generated by events invariant under permutations of the Xs that leave $X_{n+1}, X_{n+2}, \ldots$ unchanged. The* exchangeable σ-field *is $\mathcal{E} = \cap_m \mathcal{E}_m$.*

THM 11.9 (Hewitt-Savage 0-1 law) *Let $X_1, X_2, \ldots$ be iid RVs. If $A \in \mathcal{E}$ then $\mathbb{P}[A] \in \{0, 1\}$.*

Proof: The idea of the proof is to show that A is independent of itself. Indeed, we then have

$$0 = \mathbb{P}[A] - \mathbb{P}[A \cap A] = \mathbb{P}[A] - \mathbb{P}[A]\mathbb{P}[A] = \mathbb{P}[A](1 - \mathbb{P}[A]).$$

Since $A \in \mathcal{E}$ and $A \in \mathcal{F}_\infty$, it suffices to show that $\mathcal{E}$ is independent of $\mathcal{F}_n$ for every n (by the π-λ theorem).

WTS: for every bounded ϕ, $B \in \mathcal{E}$,

$$\mathbb{E}[\phi(X_1, \ldots, X_k); B] = \mathbb{E}[\phi(X_1, \ldots, X_k)]\mathbb{E}[B] = \mathbb{E}[\mathbb{E}[\phi(X_1, \ldots, X_k)]; B],$$

or equivalently

$$Y = \mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}] = \mathbb{E}[\phi(X_1, \ldots, X_k)].$$

It suffices to show that Y is independent of $\mathcal{F}_k$. Indeed, by the L^2 characterization of conditional expectation and independence,

$$0 = \mathbb{E}[(\phi(X_1, \ldots, X_k) - Y)Y] = \mathbb{E}[\phi(X_1, \ldots, X_k)]\mathbb{E}[Y] - \mathbb{E}[Y^2] = -\mathrm{Var}[Y],$$

and Y is constant.

1. Since ϕ is bounded, it is integrable and Levy's Downward Thm implies

$$\mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}_n] \to \mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}].$$

2. Define

$$A_n(\phi) = \frac{1}{(n)_k} \sum_{1 \leq i_1 \neq \cdots \neq i_k \leq n} \phi(X_{i_1}, \ldots, X_{i_k}),$$

where $(n)_k = n(n-1)\cdots(n-k+1)$. Note by symmetry

$$A_n(\phi) = \mathbb{E}[A_n(\phi) \,|\, \mathcal{E}_n] = \mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}_n] \to \mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}].$$

3. However, note that

$$\frac{1}{(n)_k} \sum_{1 \in \mathbf{i}} \phi(X_{i_1}, \ldots, X_{i_k}) \le \frac{k(n-1)_{k-1}}{(n)_k} \sup \phi = \frac{k}{n} \sup \phi \to 0,$$

so that the limit of $A_n(\phi)$ is independent of X_1 and

$$\mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}] \in \sigma(X_2, \ldots),$$

and by induction

$$Y = \mathbb{E}[\phi(X_1, \ldots, X_k) \,|\, \mathcal{E}] \in \sigma(X_{k+1}, \ldots).$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 13 : UI MGs: Optional Sampling Thm	
MATH275B - Winter 2012	*Lecturer: Sebastien Roch*

References: [Wil91, Appendix to Chapter 14], [Dur10, Section 4.7].

1 Review: Stopping times

Recall:

DEF 13.1 *A random variable* $T : \Omega \to \overline{\mathbb{Z}}_+ \equiv \{0, 1, \ldots, +\infty\}$ *is called a* stopping time *if*

$$\{T = n\} \in \mathcal{F}_n, \ \forall n \in \overline{\mathbb{Z}}_+.$$

EX 13.2 *Let* $\{A_n\}$ *be an adapted process and* $B \in \mathcal{B}$*. Then*

$$T = \inf\{n \geq 0 \, : \, A_n \in B\},$$

is a stopping time.

THM 13.3 (Optional Stopping Thm) *Let* $\{M_n\}$ *be a MG and* T *be a stopping time. Then* M_T *is integrable and*

$$\mathbb{E}[M_T] = \mathbb{E}[X_0].$$

if one of the following holds:

1. T *is bounded.*
2. M *is bounded and* T *is a.s. finite.*
3. $\mathbb{E}[T] < +\infty$ *and* M *has bounded increments.*
4. M *is UI.*

2 The σ-field $\mathcal{F}_T$

DEF 13.4 ($\mathcal{F}_T$) *Let T be a stopping time. Denote by $\mathcal{F}_T$ the set of all events F such that $\forall n \in \overline{\mathbb{Z}}_+$*

$$F \cap \{T = n\} \in \mathcal{F}_n.$$

The following two lemmas clarify the definition:

LEM 13.5 *$\mathcal{F}_T = \mathcal{F}_n$ if $T \equiv n$, $\mathcal{F}_T = \mathcal{F}_\infty$ if $T \equiv \infty$ and $\mathcal{F}_T \subseteq \mathcal{F}_\infty$ for any T.*

Proof: In the first case, note $F \cap \{T = k\}$ is empty if $k \neq n$ and is F if $k = n$. So if $F \in \mathcal{F}_T$ then $F = F \cap \{T = n\} \in \mathcal{F}_n$ and if $F \in F_n$ then $F = F \cap \{T = n\} \in F_n$. Moreover $\emptyset \in \mathcal{F}_n$ so we have proved both inclusions. This works also for $n = \infty$. For the third claim note

$$F = \cup_{k \in \overline{\mathbb{Z}}_+} F \cap \{T = n\} \in \mathcal{F}_\infty.$$

■

LEM 13.6 *If X is adapted and T is a stopping time then $X_T \in \mathcal{F}_T$ (where we assume that $X_\infty \in \mathcal{F}_\infty$, e.g., $X_\infty = \liminf X_n$).*

Proof: For $B \in \mathcal{B}$

$$\{X_T \in B\} \cap \{T = n\} = \{X_n \in B\} \cap \{T = n\} \in \mathcal{F}_n.$$

■

LEM 13.7 *If S, T are stopping times then $\mathcal{F}_{S \wedge T} \subseteq \mathcal{F}_T$.*

Proof: Let $F \in \mathcal{F}_{S \wedge T}$. Note that

$$F \cap \{T = n\} = \cup_{k \leq n} [(F \cap \{S \wedge T = k\}) \cap \{T = n\}] \in \mathcal{F}_n.$$

■

3 Optional Sampling Theorem (OST)

THM 13.8 (Optional Sampling Theorem) *If M is a UI MG and S, T are stopping times with $S \leq T$ a.s. then $\mathbb{E}|M_T| < +\infty$ and*

$$\mathbb{E}[M_T \,|\, \mathcal{F}_S] = M_S.$$

Proof: Since M is UI, $\exists M_\infty \in \mathcal{L}^1$ s.t. $M_n \to M_\infty$ a.s. and in $\mathcal{L}^1$. We prove a more general claim:

LEM 13.9

$$\mathbb{E}[M_\infty \,|\, \mathcal{F}_T] = M_T.$$

Indeed, we then get the theorem by (TOWER) and (JENSEN).
Proof:(Lemma) Wlog we assume $M_\infty \geq 0$ so that $M_n = \mathbb{E}[M_\infty \,|\, \mathcal{F}_n] \geq 0 \ \forall n$. Let $F \in \mathcal{F}_T$. Then (trivially)

$$\mathbb{E}[M_\infty; F \cap \{T = \infty\}] = \mathbb{E}[M_T; F \cap \{T = \infty\}]$$

so STS

$$\mathbb{E}[M_\infty; F \cap \{T < +\infty\}] = \mathbb{E}[M_T; F \cap \{T < +\infty\}].$$

In fact, by (MON), STS

$$\mathbb{E}[M_\infty; F \cap \{T \leq k\}] = \mathbb{E}[M_T; F \cap \{T \leq k\}] = \mathbb{E}[M_{T\wedge k}; F \cap \{T \leq k\}],$$

$\forall k$. To conclude we make two observations:

1. $\boxed{F \cap \{T \leq k\} \in \mathcal{F}_{T\wedge k}.}$ Indeed if $n \leq k$

   $$F \cap \{T \leq k\} \cap \{T \wedge k = n\} = F \cap \{T = n\} \in \mathcal{F}_n,$$

   and if $n > k$

   $$= \emptyset \in \mathcal{F}_n.$$

2. $\boxed{\mathbb{E}[M_\infty \,|\, \mathcal{F}_{T\wedge k}] = M_{T\wedge k}.}$ Since $\mathbb{E}[M_\infty \,|\, \mathcal{F}_k] = M_k$, STS $\mathbb{E}[M_k \,|\, \mathcal{F}_{T\wedge k}] = M_{T\wedge k}$. But note that if $G \in \mathcal{F}_{T\wedge k}$

   $$\mathbb{E}[M_k; G] = \sum_{l\leq k} \mathbb{E}[M_k; G \cap \{T\wedge k = l\}] = \sum_{l \leq k} \mathbb{E}[M_l; G \cap \{T \wedge k = l\}] = \mathbb{E}[M_{T\wedge k}; G]$$

   since $G \cap \{T \wedge k = l\} \in \mathcal{F}_l$.

■ ■

4 Example: Biased RW

DEF 13.10 *The* asymmetric simple RW *with parameter* $1/2 < p < 1$ *is the process* $\{S_n\}_{n\geq 0}$ *with* $S_0 = 0$ *and* $S_n = \sum_{k\leq n} X_k$ *where the* X_k*s are iid in* $\{-1, +1\}$ *s.t.* $\mathbb{P}[X_1 = 1] = p$*. Let* $q = 1 - p$*. Let* $\phi(x) = (q/p)^x$ *and* $\psi_n(x) = x - (p - q)n$*.*

THM 13.11 *Let $\{S_n\}$ as above. Let $a < 0 < b$. Define $T_x = \inf\{n \geq 0 \,:\, S_n = x\}$. Then*

1. *We have*
$$\mathbb{P}[T_a < T_b] = \frac{\phi(b) - \phi(0)}{\phi(b) - \phi(a)}.$$
In particular, $\mathbb{P}[T_a < +\infty] = 1/\phi(a)$ and $\mathbb{P}[T_b < \infty] = 1$.

2. *We have*
$$\mathbb{E}[T_b] = \frac{b}{2p - 1}.$$

Proof: There are two MGs here:

$$\mathbb{E}[\phi(S_n) \,|\, \mathcal{F}_{n-1}] = p(q/p)^{S_{n-1}+1} + q(q/p)^{S_{n-1}-1} = \phi(S_{n-1}),$$

and

$$\mathbb{E}[\psi_n(S_n) \,|\, \mathcal{F}_{n-1}] = p[S_{n-1}+1-(p-q)(n)]+q[S_{n-1}-1-(p-q)(n)] = \psi_{n-1}(S_{n-1}).$$

Let $N = T_a \wedge T_b$. Now note that $\phi(S_{N\wedge n})$ is a bounded MG and therefore applying the MG property at time n and taking limits as $n \to \infty$ (using (DOM))

$$\phi(0) = \mathbb{E}[\phi(S_N)] = \mathbb{P}[T_a < T_b]\phi(a) + \mathbb{P}[T_a > T_b]\phi(b),$$

where we need to prove that $N < +\infty$ a.s. Indeed, since $(b-a)$ $+1$-steps always take us out of (a, b),

$$\mathbb{P}[T_b > n(b-a)] \leq (1 - q^{b-a})^n,$$

so that
$$\mathbb{E}[T_b] = \sum_{k\geq 0} \mathbb{P}[T_b > k] \leq \sum_n (b-a)(1 - q^{b-a})^n < +\infty.$$

In particular $T_b < +\infty$ a.s. and $N < +\infty$ a.s. Rearranging the formula above gives the first result. (For the second part of the first result, take $b \to +\infty$ and use monotonicity.)

For the third one, note that $T_b \wedge n$ is bounded so that

$$0 = \mathbb{E}[S_{T_b\wedge n} - (p-q)(T_b \wedge n)].$$

By (MON), $\mathbb{E}[T_b \wedge n] \uparrow \mathbb{E}[T_b]$. Finally, using

$$\mathbb{P}[-\inf_n S_n \geq -a] = \mathbb{P}[T_a < +\infty],$$

and the fact that $-\inf_n S_n \geq 0$ shows that $\mathbb{E}[-\inf_n S_n] < +\infty$. Hence, we can use (DOM) with $|S_{T_b\wedge n}| \leq \max\{b, -\inf_n S_n\}$.

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Lecture 13 : Stationary Stochastic Processes

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Var01, Chapter 6], [Dur10, Section 6.1], [Bil95, Chapter 24].

1 Stationary stochastic processes

DEF 13.1 (Stationary stochastic process) *A real-valued process $\{X_n\}_{n\geq 0}$ is stationary if for every k, m*

$$(X_m, \ldots, X_{m+k}) \sim (X_0, \ldots, X_k).$$

EX 13.2 *IID sequences are stationary.*

1.1 Stationary Markov chains

1.1.1 Markov chains

DEF 13.3 (Discrete-time finite-space MC) *Let A be a finite space, μ a distribution on A and $\{p(i,j)\}_{i,j\in A}$ a transition matrix on E. Let $(X_n)_{n\geq 0}$ be a process with distribution*

$$\mathbb{P}[X_0 = x_0, \ldots, X_n = x_n] = \mu(x_0)p(x_0, x_1)\cdots p(x_{n-1}, n_n),$$

for all $n \geq 0$ and $x_0, \ldots, x_n \in A$.

EX 13.4 (RW on a graph) *Let $G = (V, E)$ be a finite, undirected graph. Define*

$$p(i,j) = \frac{\mathbb{1}\{(i,j) \in E\}}{|\{N(i)\}|},$$

where

$$N(i) = \{j \,:\, (i,j) \in E\}.$$

This defines a RW on a graph as the finite MC with the above transition matrix (for each μ, an arbitrary distribution on V). More generally, any finite MC can be seen as a RW on a weighted directed graph.

EX 13.5 (Asymmetric SRW on an interval) *Let $(S_n)_{n\geq 0}$ be an asymmetric SRW with parameter $1/2 < p < 1$. Let $a < 0 < b$, $N = T_a \wedge T_b$. Then $(X_n)_{n\geq 0} = (S_{N\wedge n})_{n\geq 0}$ is a Markov chain.*

1.1.2 Stationarity

DEF 13.6 (Stationary Distribution) *A probability measure π on A is a stationary distribution if*

$$\sum_i \pi(i)p(i,j) = \pi(j),$$

for all $i, j \in A$. In other words, if $X_0 \sim \pi$ then $X_1 \sim \pi$ and in fact $X_n \sim \pi$ for all $n \geq 0$.

EX 13.7 (RW on a graph) *In the RW on a graph example above, define*

$$\pi(i) = \frac{|N(i)|}{2|E|}.$$

Then

$$\sum_{i \in V} \pi(i)p(i,j) = \sum_{i:(i,j)\in E} \frac{|N(i)|}{2|E|}\frac{1}{|N(i)|} = \frac{1}{2|E|}|N(j)| = \pi(j),$$

so that π is a stationary distribution.

EX 13.8 (ASRW on interval) *In the ASRW on $[a, b]$, $\pi = \delta_a$ and $\pi = \delta_b$ as well as all mixtures are stationary.*

EX 13.9 (Stationary Markov chain) *Let X be a MC on A (countable) with transition matrix $\{p_{ij}\}_{i,j\in A}$ and stationary distribution $\pi > 0$. Then X started at π is a stationary stochastic process. Indeed, by definition of π and induction*

$$X_0 \sim X_n,$$

for all $n \geq 0$. Then for all m, k by definition of MCs

$$(X_0, \ldots, X_k) \sim (X_m, \ldots, X_{m+k}).$$

1.2 Abstract setting

EX 13.10 (A canonical example) *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A map $T : \Omega \to \Omega$ is said to be measure-preserving (for $\mathbb{P}$) if for all $A \in \mathcal{F}$,*

$$(\mathbb{P}[\omega : T\omega \in A] =)\mathbb{P}[T^{-1}A] = \mathbb{P}[A].$$

If $X \in \mathcal{F}$ then $X_n(\omega) = X(T^n\omega)$, $n \geq 0$, defines a stationary sequence. Indeed, for all $B \in \mathcal{B}(\mathbb{R}^{k+1})$

$$\begin{aligned}\mathbb{P}[(X_0, \ldots, X_k)(\omega) \in B] &= \mathbb{P}[(X_0, \ldots, X_k)(T^m\omega) \in B] \\ &= \mathbb{P}[(X_m, \ldots, X_{m+k})(\omega) \in B].\end{aligned}$$

Kolmogorov's extension theorem indicates that all real-valued stationary stochastic processes can be realized in the framework of the previous example.

THM 13.11 (Kolmogorov Extension Theorem) *Suppose we are given probability measure μ_n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ s.t.*

$$\mu_{n+1}((a_0, b_0] \times \cdots \times (a_n, b_n] \times \mathbb{R}) = \mu_n((a_0, b_0] \times \cdots \times (a_n, b_n]),$$

for all n and $(n+1)$-dimensional rectangles. Then there exists a unique probability measure $\mathbb{P}$ on $(\mathbb{R}^{\mathbb{Z}_+}, \mathcal{R}^{\mathbb{Z}_+})$ with marginals μ_n.

EX 13.12 (Revisiting stationary processes) *Let $\tilde{X}$ be a stationary process on $\mathbb{R}$. Then by the previous theorem, we can realize $\tilde{X}$ on $\mathbb{R}^{\mathbb{Z}_+}$ as*

$$X_n(\omega) = \omega_n.$$

The corresponding measure-preserving transformation is the shift

$$T\omega = (\omega_1, \ldots).$$

In particular, $X_n(\omega) = X_0(T^n\omega)$.

EX 13.13 *Returning the previous example:*

1. *The only invariant sets are $\emptyset, \Omega$ so that $\mathcal{I}$ is trivial and T is ergodic.*
2. *Both Ω_1 and Ω_2 are invariant so that if $\alpha, \beta \neq 0$ we have that T is not ergodic. Further, note that $\hat{f}$ is measurable with respect to $\mathcal{I} = \{\emptyset, \Omega_1, \Omega_2, \Omega\}$, that is, $\hat{f}$ is invariant.*

Next time, we will prove the ergodic theorem:

THM 13.14 *Let $f \in L^1$. Then there is $\hat{f} \in \mathcal{I}$ s.t.*

$$n^{-1}S_n \to \hat{f},$$

a.s and in L^1. In the ergodic case, $\hat{f} = \mathbb{E}[f]$.

EX 13.15 (IID RVs) *Let $X_n(\omega) = \omega_n$ are iid rvs. If A is invariant then $\{\omega : \omega \in A\} = \{\omega : T\omega \in A\} \in \sigma(X_1, \ldots)$ and by induction*

$$A \in \cap_{n \geq 0}\sigma(X_n, \ldots) = \mathcal{T},$$

where $\mathcal{T}$ is the tail σ-field. Thus $\mathcal{I} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is trivial by Kolmogorov's $0-1$ law, so is $\mathcal{I}$. Therefore T is ergodic and $\mathbb{E}[f \mid \mathcal{I}] = \mathbb{E}[f]$. Applying the ergodic thm to $f = X_0 \in L^1$ we get

$$n^{-1}\sum_{m=0}^{n-1} X_m(\omega) \to \mathbb{E}[X_0],$$

that is, we recover the SLLN.

References

[Bil95] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1995.

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Var01] S. R. S. Varadhan. *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2001.

Lecture 14 : Ergodic Theorem

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Var01, Chapter 6], [Dur10, Section 6.2], [SS05, Section 6.5].

Previous class

In view of the canonical example in the previous lecture, we assume that we have $(\Omega, \mathcal{F}, \mathbb{P})$, $f \in \mathcal{F}$, T a measure-preserving transformation, and we let $X_n(\omega) = f(T^n\omega)$ for all $n \geq 0$.

We are interested in the convergence of empirical averages

$$n^{-1}S_n(\omega) = n^{-1}\sum_{m=0}^{n-1} X_m(\omega) = n^{-1}\sum_{m=0}^{n-1} f(T^m\omega).$$

1 Invariant sets

EX 14.1 *Let $\Omega = \{a, b, c, d, e\}$ and $\mathcal{F} = 2^{\Omega}$. Take $f = \mathbb{1}_A$ for some set $A \in \mathcal{F}$.*

1. *Suppose $T = (a, b, c, d, e)$. For T to be measure-preserving we require $\mathbb{P}[a] = \mathbb{P}[b] = \cdots$ so that $\mathbb{P}[a] = 1/5$ is the only possibility. (It is easy to see that T is indeed measure-preserving because the number of elements of Ω is invariant under T.) In that case, it is immediate that*

$$n^{-1}S_n \to \mathbb{P}[A] = \mathbb{E}[f].$$

2. *Suppose $T = (a, b, c)(d, e)$. Let $\Omega_1 = \{a, b, c\}$, $\mathcal{F}_1 = 2^{\Omega_1}$, $\Omega_2 = \{d, e\}$ and $\mathcal{F}_2 = 2^{\Omega_2}$. For T to be measure-preserving we need $\mathbb{P}[a] = \mathbb{P}[b] = \mathbb{P}[c] = \alpha/3$ and $\mathbb{P}[d] = \mathbb{P}[e] = \beta/2$. (Any choice of α, β with $\alpha + \beta = 1$ works because the number of elements of Ω_1 and Ω_2 is invariant under T.) Take $A = \{a, d\}$. Then $n^{-1}S_n \to 1/3$ with probability α (i.e. if $\omega \in \Omega_1$) and $n^{-1}S_n \to 1/2$ with probability β. Denoting $\hat{f}$ this limit, we note*

$$\mathbb{E}[\hat{f}] = \mathbb{P}[A] = \mathbb{E}[f],$$

but $\hat{f}$ is not constant if $\alpha, \beta \neq 0$. However, it is completely determined by whether $\omega \in \Omega_1$ or $\omega \in \Omega_2$.

DEF 14.2 *A set $A \in \mathcal{F}$ is* invariant *if*

$$(\{\omega \,:\, T\omega \in A\} =)T^{-1}A = A,$$

up to a null set. It is nontrivial *if $0 < \mathbb{P}[A] < 1$. The set of all invariant sets forms a σ-field $\mathcal{I}$. The transformation T is said* ergodic *if $\mathcal{I}$ is trivial, that is, all invariant sets are trivial. A function g is* invariant *if $g(T\omega) = g(\omega)$ a.s. Note that g is invariant iff $g \in \mathcal{I}$. (Exercise 6.1.1 in [Dur10].)*

2 Ergodic Theorem

It will be convenient to think of T as an operator of functions

$$Uf(\omega) = f(T\omega),$$

in which case $U^m f(\omega) = f(T^m\omega)$ and we define

$$A_n f = n^{-1}(I + \cdots + U^{n-1})f.$$

LEM 14.3 *If $g \in L^1$ then*

$$\mathbb{E}[Ug] = \mathbb{E}[g].$$

Moreover if $g, g' \in L^2$ then

$$\|Ug\| = \|g\|,$$

and

$$\langle Ug', Ug\rangle = \langle g', g\rangle.$$

Proof: Start from indicators. ■

THM 14.4 *Let $f \in L^1$. Then there is $\hat{f} \in \mathcal{I}$ s.t.*

$$A_n f \to \hat{f} \equiv \mathbb{E}[f \,|\, \mathcal{I}], \text{ a.s and in } L^1.$$

EX 14.5 (IID RVs) *Let $X_n(\omega) = \omega_n$ are iid rvs. If A is invariant then $\{\omega \,:\, \omega \in A\} = \{\omega \,:\, T\omega \in A\} \in \sigma(X_1, \ldots)$ and by induction*

$$A \in \cap_{n\geq 0}\sigma(X_n, \ldots) = \mathcal{T},$$

where $\mathcal{T}$ is the tail σ-field. Thus $\mathcal{I} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is trivial by Kolmogorov's $0-1$ law, so is $\mathcal{I}$. Therefore T is ergodic and $\mathbb{E}[f \,|\, \mathcal{I}] = \mathbb{E}[f]$. Applying the ergodic thm to $f = X_0 \in L^1$ we get

$$n^{-1}\sum_{m=0}^{n-1} X_m(\omega) \to \mathbb{E}[X_0],$$

that is, we recover the SLLN.

3 L^2 Ergodic Theorem

THM 14.6 *Let $f \in L^2$. Then there is $\hat{f} \in \mathcal{I}$ s.t.*

$$A_n f \to \hat{f} \equiv \mathbb{E}[f \,|\, \mathcal{I}], \textit{ in } L^2.$$

Proof: Let

$$H_0 = \{f \in L^2 \,:\, Uf = f \text{ a.s.}\},$$

and note that $A_n f = f$ for $f \in H_0$. We need the following lemma from basic Hilbert space theory (see [SS05, Lemma 6.5.2]).

LEM 14.7 *The following hold:*

1. $H_0 = \{f \in L^2 \,:\, U^* f = f \textit{ a.s.}\}$.
2. $H_0^\perp = \overline{\textit{Range}(I - U)}$.

Proof: See e.g. [SS05]. ■

For $\varepsilon > 0$, write $f = f_0 + f_1$ where $f_0 \in H_0$ and $\|f_1 - f_1'\|_2 < \varepsilon$ s.t. $f_1' = (I - U)g_1'$. Then

$$A_n f_0 = f_0, \text{ and } A_n f_1' = \frac{1}{n}(I - U^n) g_1',$$

so that

$$\begin{aligned} \|A_n f - f_0\|_2 &= \|n^{-1}(I - U^n)g_1' + A_n(f_1 - f_1')\|_2 \\ &\leq (\|g_1'\|_2 + \|U^n g_1'\|_2) n^{-1} + n^{-1} \sum_{m=0}^{n-1} \|U^m (f_1 - f_1')\|_2 \\ &= 2\|g_1'\|_2 n^{-1} + n^{-1} \sum_{m=0}^{n-1} \|f_1 - f_1'\|_2 \\ &\to \varepsilon. \end{aligned}$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[SS05] Elias M. Stein and Rami Shakarchi. *Real analysis*. Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[Var01] S. R. S. Varadhan. *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2001.

Lecture 15 : Proof of the Ergodic Theorem (cont'd)	
MATH275B - Winter 2012	*Lecturer: Sebastien Roch*

References: [Var01, Chapter 6], [Dur10, Section 6.2], [SS05, Section 6.5].

1 Proof of Ergodic Theorem

We assume we have $(\Omega, \mathcal{F}, \mathbb{P})$, $f \in \mathcal{F}$, T a measure-preserving transformation, and we let $X_n(\omega) = f(T^n\omega)$ for all $n \geq 0$. It will be convenient to think of T as an operator of functions

$$Uf(\omega) = f(T\omega),$$

in which case $U^m f(\omega) = f(T^m\omega)$ and we define

$$A_n f = n^{-1}(I + \cdots + U^{n-1})f.$$

Recall:

LEM 15.1 *If $g, g' \in L^2$ then*

$$\langle Ug', Ug\rangle = \langle g', g\rangle.$$

THM 15.2 *Let $f \in L^1$. Then there is $\hat{f} \in \mathcal{I}$ s.t.*

$$A_n f \to \hat{f} \equiv \mathbb{E}[f \,|\, \mathcal{I}], \text{ a.s and in } L^1.$$

Proof: We first show a.s. convergence to a limit. We proceed as in the L^2 case. Fix ε and let

$$f = F + H = f_0 + (I - U)g_1' + H,$$

where $\|H\|_1 < \varepsilon$ includes both the L^1 and closure error terms. We show that $A_n F$ converges a.s. Note that

$$A_n F(\omega) = f_0(\omega) + n^{-1}(I - U^n)g_1'(\omega) = f_0(\omega) + \frac{g_1'(\omega)}{n} - \frac{g_1'(T^n\omega)}{n}.$$

To deal with the last term, note that

$$\sum_n \frac{g_1'(T^n\omega)^2}{n^2}$$

converges because its norm is bounded by $\|g_1'\|_2^2 \sum_n 1/n^2 < \infty$. To conclude let

$$E_\alpha = \{\lim_N \sup_{m,n\geq N} |A_n f - A_m f| > \alpha\}.$$

Note that

$$\mathbb{P}[E_\alpha] \leq \mathbb{P}[\lim_N \sup_{m,n\geq N} |A_n H - A_m H| > \alpha] \leq \mathbb{P}[2\sup_N |A_N H| > \alpha].$$

To conclude the proof of a.s. convergence, we need the following inequality which is similar to Doob's inequality.

LEM 15.3 (Wiener's Maximal Inequality) *For $f \in L^1$ and $\ell > 0$,*

$$\mathbb{P}\left[\sup_{j\geq 0} |A_j f| \geq \ell\right] \leq \frac{1}{\ell}\mathbb{E}|f|.$$

Proof: The proof is based on the so-called maximal ergodic lemma.

LEM 15.4 (Maximal Ergodic Lemma) *Let*

$$f_n^* = \sup_{1\leq j\leq n} f + \cdots + U^{j-1} f.$$

Then for all $n \geq 0$

$$\mathbb{E}[f; \{f_n^* \geq 0\}] \geq 0.$$

Apply the maximal ergodic lemma to $|f| - \ell$ and take $n \to \infty$. ■

Applying the lemma we have

$$\mathbb{P}[E_\alpha] \leq \mathbb{P}[2\sup_N |A_N H| > \alpha] \leq \frac{2}{\alpha}\mathbb{E}|H| < \frac{2\varepsilon}{\alpha},$$

so that $\mathbb{P}[E_\alpha] = 0$ for all α.

It is clear that the limit satisfies $\hat{f}(\omega) = \hat{f}(T\omega)$. In fact, by the density of L^2 in L^1, writing $f = g_r + h_r$ with $g_r \in L^2$ and $\|h_r\|_1 < 1/r$, we have $\hat{f} = \hat{g}_r + \hat{h}_r$ and for $G \in \mathcal{I}$

$$\mathbb{E}[\hat{f}; G] = \mathbb{E}[\hat{g}_r; G] + \mathbb{E}[\hat{h}_r; G] = \mathbb{E}[g_r; G] + \mathbb{E}[\hat{h}_r; G] \to \mathbb{E}[f; G],$$

where we used the L^2 Ergodic Theorem and

$$\mathbb{E}|\hat{h}_r| \leq \liminf_n \mathbb{E}|A_n h_r| \leq \liminf_n n^{-1} \sum_{m=0}^{n-1} \mathbb{E}|U^m h_r| = \mathbb{E}|h_r| = 1/r,$$

by (FATOU).

A truncation argument gives the L^1 convergence (see [Dur10]). Let

$$f'_M = f\mathbb{1}_{|f|\leq M},$$

and $f''_M = f - f'_M$. By the ergodic theorem and the bounded convergence theorem

$$\mathbb{E}\left|\frac{1}{n}\sum_{m=0}^{n-1} f'_M(T^m\omega) - \mathbb{E}[f'_M \,|\, \mathcal{I}]\right| \to 0.$$

By stationarity and (cJENSEN),

$$\mathbb{E}\left|\frac{1}{n}\sum_{m=0}^{n-1} f''_M(T^m\omega) - \mathbb{E}[f''_M \,|\, \mathcal{I}]\right| \leq 2\mathbb{E}|f''_M| \to 0,$$

as $M \to +\infty$ by (DOM). The result follows.

■

2 Applications

Going back to Markov chains:

DEF 15.5 *Let*

$$T_i = \inf\{n \geq 1 \,:\, X_n = i\},$$

and

$$f_{ij} = \mathbb{P}_i[T_j < +\infty].$$

A chain is irreducible *if $f_{ij} > 0$ for all $i, j \in A$. A state i is recurrent if $f_{ii} = 1$ and is positive recurrent if $\mathbb{E}_i[T_i] < +\infty$.*

LEM 15.6 *If X is irreducible and finite, then every state is positive recurrent.*

THM 15.7 *Let X be an irreducible and positive recurrent MC. Then there exists a unique stationary distribution π. In fact,*

$$\pi(i) = \frac{1}{\mathbb{E}_i[T_i]} > 0.$$

EX 15.8 (MCs) *Let X be a MC on S.*

- *In the ASRW on $[a, b]$ the invariant sets are $\{a\}$ and $\{b\}$ and therefore T is not ergodic if π has positive mass on both of them.*

- *On the other hand, assume X is irreducible and positive recurrent with stationary distribution $\pi > 0$. Let $A \in \mathcal{I}$ and note that $\mathbb{1}_A \circ T^n = \mathbb{1}_A$. Then by the Markov property,*

$$\mathbb{E}[\mathbb{1}_A \,|\, \mathcal{F}_n] = \mathbb{E}[\mathbb{1}_A \circ T^n \,|\, \mathcal{F}_n] = h(X_n),$$

where $h(x) = \mathbb{E}_x[\mathbb{1}_A]$. By Levy's $0-1$ law the LHS $\to \mathbb{1}_A$. By irreducibility and recurrence, any $y \in S$ is visited i.o. and we must have $\mathbb{E}_x[\mathbb{1}_A] \equiv h(x) \equiv 0$ or 1. Therefore $\mathbb{P}[A] \in \{0, 1\}$ and $\mathcal{I}$ is trivial. Then applying the Ergodic Theorem to $f(\omega) = g(X_0(\omega))$ where

$$\sum_y |g(y)|\pi(y) < +\infty,$$

we then have

$$n^{-1} \sum_{m=0}^{n-1} g(X_m(\omega)) \to \sum_y \pi(y) g(y).$$

- *Note finally that the RW on a bipartite graph shows that, even in the irreducible recurrent case, $\mathcal{I}$ may be smaller than $\mathcal{T}$.*

Further reading

See a different proof in [Dur10, Section 6.2].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[SS05] Elias M. Stein and Rami Shakarchi. *Real analysis*. Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[Var01] S. R. S. Varadhan. *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2001.

Lecture 16 : Subadditive Ergodic Theorem

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 6.4].

1 Subadditivity

DEF 16.1 *A sequence* $\{\gamma_n\}_{n\geq 0}$ *is* subadditive *if for all* m, n*:*

$$\gamma_{m+n} \leq \gamma_n + \gamma_m.$$

THM 16.2 (Limit of Subadditive Sequences) *If* γ *is subadditive then*

$$\frac{\gamma_n}{n} \to \inf_m \frac{\gamma_m}{m}.$$

Proof: Clearly

$$\liminf_n \frac{\gamma_n}{n} \geq \inf_m \frac{\gamma_m}{m}.$$

So STS

$$\limsup_n \frac{\gamma_n}{n} \leq \inf_m \frac{\gamma_m}{m}.$$

Fix m and write $n = km + \ell$ with $0 \leq \ell < m$. Applying the subadditivity repeatedly, we have

$$\gamma_n \leq k\gamma_m + \gamma_\ell,$$

so that

$$\frac{\gamma_n}{n} \leq \left(\frac{km}{km+\ell}\right)\frac{\gamma_m}{m} + \frac{\gamma_\ell}{n},$$

and the result follows by taking $n \to +\infty$.

■

EX 16.3 (Longest common subsequence) *Let* $\{X_n\}$ *and* $\{Y_n\}$ *be stationary sequences and let* $L_{m,n}$ *be the longest common subsequence on indices* $m < k \leq n$*. Clearly*

$$L_{0,m} + L_{m,n} \leq L_{0,n},$$

and $\gamma_n = -\mathbb{E}[L_{0,n}]$ *is subadditive.*

2 Statement of Subadditive Ergodic Theorem

THM 16.4 (Subadditive Ergodic Theorem) *Suppose* $\{X_{m,n}\}_{0\leq m<n}$ *satisfy:*

1. $X_{0,m} + X_{m,n} \geq X_{0,n}$.
2. $\{X_{nk,(n+1)k}, n \geq 1\}$ *is a stationary sequence for each* k.
3. *The distribution of* $\{X_{m,m+k}, k \geq 1\}$ *does not depend on* m.
4. $\mathbb{E}X_{0,1}^{+} < \infty$ *and for each* n, $\mathbb{E}X_{0,n} \geq \gamma_0 n$ *where* $\gamma_0 > -\infty$.

Then

- $\lim \mathbb{E}X_{0,n}/n = \inf_m \mathbb{E}X_{0,m}/m \equiv \gamma$.
- $X = \lim X_{0,n}/n$ *exists a.s. and in* L^1 *so* $\mathbb{E}X = \gamma$.
- *If all stationary sequences in* 2. *are ergodic then* $X = \gamma$ *a.s.*

Proof: See [Dur10]. ∎

3 Examples

EX 16.5 (Age-dependent continuous-time branching process) *Start with one individual. Each individual dies independently after time* $T \sim F$ *and at that point produces* $K \sim \{p_k\}_k$ *offsprings (both with finite means). Let* $X_{0,m}$ *be the time of birth of the first individual from generation* m *and* $X_{m,n}$*, the time to the birth of the first descendant of that individual in generation* n*. We check the conditions:*

1. *Clearly*
$$X_{0,m} + X_{m,n} \geq X_{0,n}.$$
2. $\{X_{nk,(n+1)k}\}_n$ *is IID because we are looking at the descendants of a single individual (the first born) over* k *generations which are not overlapping.*
3. *The distribution of* $\{X_{m,m+k}\}_k$ *is independent of* m *for the same reason.*
4. *By nonnegativity and the finite mean of* F*, condition* 4. *is satisfied.*

So we can apply the thm. By the IID remark above in 2. *we get that the limit is trivial. See [Dur10] for a characterization of the limit.*

EX 16.6 (First-passage percolation) *Consider $\mathbb{Z}^d$ as a graph with edges connecting $x, y \in \mathbb{Z}^d$ if $\|x - y\|_1 = 1$. Assign to each edge a nonnegative random variable $\tau(e)$ corresponding to the time it takes to traverse e (in either direction). Define $t(x, y)$ (the passage time) as the minimum time to go from x to y. Let $X_{m,n} = t(mu, nu)$ where $u = (1, 0, \cdots, 0)$. We check the conditions:*

1. *Clearly*
$$X_{0,m} + X_{m,n} \geq X_{0,n}$$
2. *$\{X_{nk,(n+1)k}\}_n$ is stationary by translational symmetry.*
3. *The distribution of $\{X_{m,m+k}\}_k$ is independent of m for the same reason.*
4. *By nonnegativity and the finite mean of τ, condition 4. is satisfied.*

So we can apply the theorem. Enumerating the edges in some order, one can prove (check!) that the limit is tail-measurable and, by the IID assumption, is trivial. See [Dur10] for a characterization of the limit.

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[SS05] Elias M. Stein and Rami Shakarchi. *Real analysis*. Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[Var01] S. R. S. Varadhan. *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2001.

Lecture 17 : Brownian motion: Definition

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 3.9, 8.1], [Lig10, Section 1.2-1.4], [MP10, Section 1.1, Appendix 12].

1 Random vectors

We first develop general tools to study multivariate distributions.

DEF 17.1 (Characteristic function) *The CF of a random vector $X = (X_1, \ldots, X_d)$ is given by, for $t \in \mathbb{R}^d$,*

$$\phi_X(t) = \mathbb{E}\left[\exp\left(i(t_1 X_1 + \cdots + t_d X_d)\right)\right].$$

As in the one-dimensional case, we have an inversion formula:

THM 17.2 (Inversion formula) *Let μ be the probability measure corresponding to the random vector $(X_1, \ldots, X_d)$, that is, for all $B \in \mathcal{B}(\mathbb{R}^d)$,*

$$\mu(B) = \mathbb{P}[(X_1, \ldots, X_d) \in B].$$

If $A = [a_1, b_1] \times \cdots \times [a_d, b_d]$ with $\mu(\partial A) = 0$ then

$$\mu(A) = \lim_{T \to +\infty} (2\pi)^{-d} \int_{[-T,T]^d} \prod_{j=1}^{d} \psi_j(t_j) \phi(t) \mathrm{d}t,$$

where

$$\psi_j(s) = \frac{\exp(-isa_j) - \exp(-isb_j)}{is}.$$

Proof: Follows from the one-dimensional inversion formula. See [Dur10, Theorem 3.9.3]. ■

An important application of the previous formula is:

THM 17.3 *The RVs $X_1, \ldots, X_d$ are independent if and only if*

$$\phi_X(t) = \prod_{j=1}^{d} \phi_{X_j}(t_j),$$

for all $t \in \mathbb{R}^d$ where $X = (X_1, \ldots, X_d)$.

Proof: The "only if" part is obvious. The inversion formula and Fubini's theorem gives the "if" part. ■

DEF 17.4 *A sequence of random vectors X_n converges weakly to X_∞, denoted $X_n \Rightarrow X_\infty$, if*

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X_\infty)],$$

for all bounded continuous functions f. The portmanteau theorem gives equivalent characterizations.

In terms of CFs, we have:

THM 17.5 (Convergence theorem) *Let X_n, $1 \leq n \leq \infty$, be random vectors with CFs ϕ_n. A necessary and sufficient condition for $X_n \Rightarrow X_\infty$ is that*

$$\phi_n(t) \to \phi_\infty(t),$$

for all $t \in \mathbb{R}^d$.

Proof: Follows from the one-dimensional result. See [Dur10, Theorem 3.9.4]. ■

We require one last definition:

DEF 17.6 (Covariance) *Let $X = (X_1, \ldots, X_d)$ be a random vector with mean $\mu = \mathbb{E}[X]$. The* covariance of X *is the $d \times d$ matrix Γ with entries*

$$\Gamma_{ij} = \mathrm{Cov}[X_i, X_j] = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)].$$

2 Multivariate Gaussian distribution

Recall:

DEF 17.7 (Gaussian distribution) *A* standard Gaussian *is a RV Z with CF*

$$\phi_Z(t) = \exp\left(-t^2/2\right),$$

and density

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right).$$

In particular, Z has mean 0 and variance 1. More generally,

$$X = \sigma Z + \mu,$$

is a Gaussian RV with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$.

We will need a multivariate generalization of the standard Gaussian.

DEF 17.8 (Multivariate Gaussian) *A d-dimensional standard Gaussian is a random vector $X = (X_1, \ldots, X_d)$ where the X_is are independent standard Gaussians. In particular, X has mean 0 and covariance matrix I. More generally, a random vector $X = (X_1, \ldots, X_d)$ is Gaussian if there is a vector b, a $d \times r$ matrix A and an r-dimensional standard Gaussian Y such that*

$$X = AY + b.$$

Then X has mean $\mu = b$ and covariance matrix $\Gamma = AA^T$. The CF of X is given by

$$\phi_X(t) = \exp\left(i\sum_{j=1}^d t_j\mu_j - \frac{1}{2}\sum_{j,k=1}^d t_j t_k \Gamma_{jk}\right).$$

From the CF and the theorems above, we get the following:

COR 17.9 (Independence) *Let $X = (X_1, \ldots, X_d)$ be a multivariate Gaussian. Then the X_is are independent if and only if $\Gamma_{ij} = 0$ for all $i \neq j$, that is, if they are* uncorrelated.

COR 17.10 (Convergence) *Let X_n be a sequence of random vectors with means μ_n and covariances Γ_n such that $X_n \to X_\infty$ a.s., $\mu_u \to \mu_\infty$, and $\Gamma_n \to \Gamma_\infty$. Then X_∞ is a multivariate Gaussian with mean μ_∞ and covariance matrix Γ_∞.*

COR 17.11 (Linear combinations) *The random vector $(X_1, \ldots, X_d)$ is multivariate Gaussian if and only if all linear combinations of its components are Gaussian.*

Finally:

THM 17.12 (Multivariate CLT) *Let $X_1, X_2, \ldots$ be IID random vectors with means μ and finite covariance matrix Γ. Let $S_n = \sum_{j=1}^n X_j$, Then*

$$\frac{S_n - n\mu}{\sqrt{n}} \Rightarrow Z,$$

where Z is a multivariate Gaussian with mean 0 and covariance matrix Γ.

Proof: Follows easily from one-dimensional result. See [Dur10, Theorem 3.9.6]. ■

3 Gaussian processes

DEF 17.13 (Gaussian process) *A continuous-time stochastic process* $\{X(t)\}_{t\geq 0}$ *is a* Gaussian process *if for all* $n \geq 1$ *and* $0 \leq t_1 < \cdots < t_n < +\infty$ *the random vector*

$$(X(t_1), \ldots, X(t_n)),$$

is multivariate Gaussian. The mean and covariance functions of X *are* $\mathbb{E}[X(t)]$ *and* $\mathrm{Cov}[X(s), X(t)]$ *respectively.*

4 Definition of Brownian motion

DEF 17.14 (Brownian motion: Definition I) *The continuous-time stochastic process* $X = \{X(t)\}_{t\geq 0}$ *is a* standard Brownian motion *if* X *is a Gaussian process with almost surely continuous paths, that is,*

$$\mathbb{P}[X(t) \text{ is continuous in } t] = 1,$$

such that $X(0) = 0$*,*

$$\mathbb{E}[X(t)] = 0,$$

and

$$\mathrm{Cov}[X(s), X(t)] = s \wedge t.$$

More generally, $B = \sigma X + x$ *is a* Brownian motion started at x.

Further reading

Multivariate CLT in [Dur10, Section 2.9].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 19 : Brownian motion: Construction	
MATH275B - Winter 2012	*Lecturer: Sebastien Roch*

References: [Dur10, Section 8.1], [Lig10, Section 1.5], [MP10, Section 1.1].

1 Definition of Brownian motion

Recall:

DEF 19.1 (Brownian motion: Definition I) *The continuous-time stochastic process $X = \{X(t)\}_{t\geq 0}$ is a* standard Brownian motion *if X is a Gaussian process with almost surely continuous paths, that is,*

$$\mathbb{P}[X(t) \text{ is continuous in } t] = 1,$$

such that $X(0) = 0$,

$$\mathbb{E}[X(t)] = 0,$$

and

$$\mathrm{Cov}[X(s), X(t)] = s \wedge t.$$

More generally, $B = \sigma X + x$ is a Brownian motion started at x.

From the properties of the multivariate Gaussian, we get the following equivalent definition. We begin with a general definition.

DEF 19.2 (Stationary independent increments) *An SP $\{X(t)\}_{t\geq 0}$ has* stationary increments *if the distribution of $X(t) - X(s)$ depends only on $t - s$ for all $0 \leq s \leq t$. It has* independent increments *if the RVs $\{X(t_{j+1} - X(t_j)), 1 \leq j < n\}$ are independent whenever $0 \leq t_1 < t_2 < \cdots < t_n$ and $n \geq 1$.*

DEF 19.3 (Brownian motion: Definition II) *The continuous-time stochastic process $X = \{X(t)\}_{t\geq 0}$ is a* standard Brownian motion *if X has almost surely continuous paths and stationary independent increments such that $X(s+t) - X(s)$ is Gaussian with mean 0 and variance t.*

2 Construction of Brownian motion

Given that standard Brownian motion is defined in terms of finite-dimensional distributions, it is tempting to attempt to construct it by using Kolmogorov's Extension Theorem.

THM 19.4 (Kolmogorov's Extension Theorem: Uncountable Case) *Let*

$$\Omega_0 = \{\omega : [0,\infty) \to \mathbb{R}\},$$

and $\mathcal{F}_0$ be the σ-field generated by the finite-dimensional sets

$$\{\omega : \omega(t_i) \in A_i, 1 \leq i \leq n\},$$

for $A_i \in \mathcal{B}$. There is a unique probability measure ν on $(\Omega_0, \mathcal{F}_0)$ so that

$$\nu(\{\omega : \omega(0) = 0\}) = 1$$

and whenever $0 \leq t_1 < \cdots < t_n$ with $n \geq 1$ we have

$$\nu(\{\omega : \omega(t_i) \in A_i\}) = \mu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n),$$

where the latter is the finite-dimensional distribution of standard Brownian motion.

See [Dur10]. The only problem with this approach is that the event

$$C = \{\omega : \omega(t) \text{ is continuous in } t\},$$

is not in $\mathcal{F}_0$. See Exercise 8.1.1 in [Dur10].

Instead, we proceed as follows. There are several constructions of Brownian motion. We present Lévy's contruction, as described in [MP10]. See [Dur10] and [Lig10] for further constructions.

THM 19.5 (Existence) *Standard Brownian motion $B = \{B(t)\}_{t\geq 0}$ exists.*

Proof: We first construct B on $[0,1]$. The idea is to construct the process on dyadic points and extend it linearly. Let

$$\mathcal{D}_n = \{k2^{-n} : 0 \leq k \leq 2^n\},$$

and

$$\mathcal{D} = \cup_{n=0}^{\infty} \mathcal{D}_n.$$

Note that $\mathcal{D}$ is countable and consider $\{Z_t\}_{t\in\mathcal{D}}$ a collection of independent standard Gaussians. We define $B(d)$ for $d \in \mathcal{D}_n$ by induction. First take $B(0) = 0$

and $B(1) = Z_1$. Note that $B(1) - B(0)$ is Gaussian with variance 1. Then for $d \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}$ we let

$$B(d) = \frac{B(d-2^{-n}) + B(d+2^{-n})}{2} + \frac{Z_d}{2^{(n+1)/2}}.$$

By construction, $B(d)$ is independent of $\{Z_t : t \in \mathcal{D} \backslash \mathcal{D}_n\}$. Moreover, as a linear combination of zero-mean Gaussians, $B(d)$ is a zero-mean Gaussian.

We claim that the differences $B(d) - B(d - 2^{-n})$, for all $d \in \mathcal{D}_n \backslash \{0\}$, are independent Gaussians with variance 2^{-n} .

- We first argue about neighboring increments. Note that, for $d \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}$,

  $$B(d) - B(d-2^{-n}) = \frac{B(d+2^{-n}) - B(d-2^{-n})}{2} + \frac{Z_d}{2 \cdot 2^{(n-1)/2}},$$

  and

  $$B(d+2^{-n}) - B(d) = \frac{B(d+2^{-n}) - B(d-2^{-n})}{2} - \frac{Z_d}{2 \cdot 2^{(n-1)/2}},$$

  are Gaussians and they are independent by the following lemma. By induction the differences above are Gaussians with variance $2^{-(n-1)}$ and independent of Z_d.

  LEM 19.6 *If (X_1, X_2) is a standard Gaussian then so is $\frac{1}{\sqrt{2}}(X_1 + X_2, X_1 - X_2)$.*

- More generally, the two intervals are separated by $d \in \mathcal{D}_j$. Take a minimal such j. Then, by induction, the increments over the intervals $[d - 2^{-j}, d]$ and $[d, d + 2^{-j}]$ are independent. Moreover, the increments over the two intervals of length 2^{-n} of interest (included in the above intervals) are constructed from $B(d) - B(d - 2^{-j})$, respectively $B(d + 2^{-j}) - B(d)$, using a disjoint set of variables $\{Z_t : t \in \mathcal{D}_n\}$. That proves the claim by induction.

We now interpolate linearly between dyadic points. More precisely, let

$$F_0(t) = \begin{cases} Z_1, & t = 1, \\ 0, & t = 0, \\ \text{linearly}, & \text{in between.} \end{cases}$$

and for $n \geq 1$

$$F_n(t) = \begin{cases} 2^{-(n+1)/2} Z_t, & t \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}, \\ 0, & t \in \mathcal{D}_{n-1}, \\ \text{linearly}, & \text{in between.} \end{cases}$$

We then have for $d \in \mathcal{D}_n$

$$B(d) = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{\infty} F_i(d).$$

We want to show that the resulting process is continuous on $[0,1]$. We claim that the series

$$B(t) = \sum_{n=0}^{\infty} F_n(t),$$

is uniformly convergent. From a bound on Gaussian tails we saw last quarter,

$$\mathbb{P}[|Z_d| \geq c\sqrt{n}] \leq \exp\left(-c^2 n/2\right),$$

so that for c large enough

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{P}[\exists d \in \mathcal{D}_n, |Z_d| \geq c\sqrt{n}] &\leq \sum_{n=0}^{\infty} (2^n + 1)\exp\left(-c^2 n/2\right) \\ &< +\infty. \end{aligned}$$

By BC, there is N (random) such that $|Z_d| < c\sqrt{n}$ for all $d \in \mathcal{D}_n$ with $n > N$. In particular, for $n > N$ we have

$$\|F_n\|_\infty < c\sqrt{n} 2^{-(n+1)/2},$$

from which we get the claim.

To show that $B(t)$ has the correct finite-dimensional distributions, note that this is the case for $\mathcal{D}$ by the above argument. Since $\mathcal{D}$ is dense in $[0,1]$ the result holds on $[0,1]$ by taking limits and using the convergence theorem for Gaussians from the previous lecture.

Finally, we extend the process to $[0,+\infty)$ by gluing together independent copies of $B(t)$. ■

Further reading

Other constructions in [Dur10, Section8.1] and [Lig10, Section 1.5].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 19 : Brownian motion: Path properties I

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 8.1], [Lig10, Section 1.5, 1.6], [MP10, Section 1.1, 1.2].

1 Invariance

We begin with some useful invariance properties. The following are immediate.

THM 19.1 (Time translation) *Let $s \geq 0$. If $B(t)$ is a standard Brownian motion, then so is $X(t) = B(t+s) - B(s)$.*

THM 19.2 (Scaling invariance) *Let $a > 0$. If $B(t)$ is a standard Brownian motion, then so is $X(t) = a^{-1}B(a^2 t)$.*

Proof: *Sketch.* We compute the variance of the increments:

$$\begin{aligned}\mathrm{Var}[X(t) - X(s)] &= \mathrm{Var}[a^{-1}(B(a^2 t) - B(a^2 s))] \\ &= a^{-2}(a^2 t - a^2 s) \\ &= t - s.\end{aligned}$$

■

THM 19.3 (Time inversion) *If $B(t)$ is a standard Brownian motion, then so is*

$$X(t) = \begin{cases} 0, & t = 0, \\ tB(t^{-1}), & t > 0. \end{cases}$$

Proof: *Sketch.* We compute the covariance function for $s < t$:

$$\begin{aligned}\mathrm{Cov}[X(s), X(t)] &= \mathrm{Cov}[sB(s^{-1}), tB(t^{-1})] \\ &= st\left(s^{-1} \wedge t^{-1}\right) \\ &= s.\end{aligned}$$

It remains to check continuity at 0. Note that

$$\left\{\lim_{t\downarrow 0} B(t) = 0\right\} = \bigcap_{m\geq 1} \bigcup_{n\geq 1} \left\{|B(t)| \leq 1/m,\ \forall t \in \mathbb{Q}\cap(0,1/n)\right\},$$

and

$$\left\{\lim_{t\downarrow 0} X(t) = 0\right\} = \bigcap_{m\geq 1} \bigcup_{n\geq 1} \left\{|X(t)| \leq 1/m,\ \forall t \in \mathbb{Q}\cap(0,1/n)\right\}.$$

The RHSs have the same probability because the distributions on all finite-dimensional sets —and therefore on the rationals—are the same. The LHS of the first one has probability 1. ∎

Typical applications of these are:

COR 19.4 *For $a < 0 < b$, let*

$$T(a,b) = \inf\left\{t \geq 0\ :\ B(t) \in \{a,b\}\right\}.$$

Then

$$\mathbb{E}[T(a,b)] = a^2\mathbb{E}[T(1,b/a)].$$

In particular, $\mathbb{E}[T(-b,b)]$ is a constant multiple of b^2.

Proof: Let $X(t) = a^{-1}B(a^2t)$. Then,

$$\begin{aligned}\mathbb{E}[T(a,b)] &= a^2\mathbb{E}[\inf\{t\geq 0,:\ X(t) \in \{1,b/a\}\}]\\ &= a^2\mathbb{E}[T(1,b/a)].\end{aligned}$$

∎

COR 19.5 *Almost surely,*

$$t^{-1}B(t) \to 0.$$

Proof: Let $X(t)$ be the time inversion of $B(t)$. Then

$$\lim_{t\to\infty}\frac{B(t)}{t} = \lim_{t\to\infty} X(1/t) = X(0) = 0.$$

∎

2 Modulus of continuity

By construction, $B(t)$ is continuous a.s. In fact, we can prove more.

DEF 19.6 (Hölder continuity) *A function f is said* locally α-Hölder continuous at x *if there exists $\varepsilon > 0$ and $c > 0$ such that*

$$|f(x) - f(y)| \leq c|x - y|^\alpha,$$

for all y with $|y - x| < \varepsilon$. We refer to α as the Hölder exponent and to c as the Hölder constant.

THM 19.7 (Holder continuity) *If $\alpha < 1/2$, then almost surely Brownian motion is everywhere locally α-Hölder continuous.*

Proof:

LEM 19.8 *There exists a constant $C > 0$ such that, almost surely, for every sufficiently small $h > 0$ and all $0 \leq t \leq 1 - h$,*

$$|B(t + h) - B(t)| \leq C\sqrt{h \log(1/h)}.$$

Proof: Recall our construction of Brownian motion on $[0, 1]$. Let

$$\mathcal{D}_n = \{k2^{-n} \,:\, 0 \leq k \leq 2^n\},$$

and

$$\mathcal{D} = \cup_{n=0}^{\infty} \mathcal{D}_n.$$

Note that $\mathcal{D}$ is countable and consider $\{Z_t\}_{t \in \mathcal{D}}$ a collection of independent standard Gaussians. Let

$$F_0(t) = \begin{cases} Z_1, & t = 1, \\ 0, & t = 0, \\ \text{linearly}, & \text{in between.} \end{cases}$$

and for $n \geq 1$

$$F_n(t) = \begin{cases} 2^{-(n+1)/2} Z_t, & t \in \mathcal{D}_n \backslash \mathcal{D}_{n-1}, \\ 0, & t \in \mathcal{D}_{n-1}, \\ \text{linearly}, & \text{in between.} \end{cases}$$

Finally

$$B(t) = \sum_{n=0}^{\infty} F_n(t).$$

Each F_n is piecewise linear and its derivative exists almost everywhere. By construction, we have

$$\|F_n'\|_\infty \leq \frac{\|F_n\|_\infty}{2^{-n}}.$$

Recall that there is N (random) such that $|Z_d| < c\sqrt{n}$ for all $d \in \mathcal{D}_n$ with $n > N$. In particular, for $n > N$ we have

$$\|F_n\|_\infty < c\sqrt{n}2^{-(n+1)/2}.$$

Using the mean-value theorem, assuming $l > N$,

$$\begin{aligned}
|B(t+h) - B(t)| &\leq \sum_{n=0}^{\infty} |F_n(t+h) - F_n(t)| \\
&\leq \sum_{n=0}^{l} h\|F_n'\|_\infty + \sum_{n=l+1}^{\infty} 2\|F_n\|_\infty, \\
&\leq h\sum_{n=0}^{N} \|F_n'\|_\infty + ch\sum_{n=N}^{l} \sqrt{n}2^{n/2} + 2c\sum_{n=l+1}^{\infty} \sqrt{n}2^{-n/2}.
\end{aligned}$$

Take h small enough that the first term is smaller than $\sqrt{h\log(1/h)}$ and l defined by $2^{-l} < h \leq 2^{-l+1}$ exceeds N. Then approximating the second and third terms by their largest element gives the result. ■

We go back to the proof of the theorem. For each k, we can find an $h(k)$ small enough so that the result applies to the standard BMs

$$\{B(k+t) - B(k) \,:\, t \in [0,1]\},$$

and

$$\{B(k+1-t) - B(k+1) \,:\, t \in [0,1]\}.$$

Since there are countably many intervals $[k, k+1)$, such $h(k)$'s exist almost surely on all intervals simultaneously. Then note that for any $\alpha < 1/2$, if $t \in [k, k+1)$ and $h < h(k)$ small enough,

$$|B(t+h) - B(t)| \leq C\sqrt{h\log(1/h)} \leq Ch^\alpha (= Ch^{1/2}(1/h)^{(1/2-\alpha)}).$$

This concludes the proof. ■

In fact:

THM 19.9 (Lévy's modulus of continuity) *Almost surely,*

$$\limsup_{h\downarrow 0} \sup_{0\leq t\leq 1-h} \frac{|B(t+h) - B(t)|}{\sqrt{2h\log(1/h)}} = 1.$$

For the proof, see [MP10].

This result is tight. See [MP10, Remark 1.21].

3 Non-Monotonicity

THM 19.10 *Almost surely, for all $0 < a < b < +\infty$, standard BM is not monotone on the interval $[a, b]$.*

Proof: It suffices to look at intervals with rational endpoints because any general non-degenerate interval of monotonicity must contain one of those. Since there are countably many rational intervals, it suffices to prove that any particular one has probability 0 of being monotone. Let $[a, b]$ be such an interval. Note that for any finite sub-division

$$a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b,$$

the probability that each increment satisfies

$$B(a_i) - B(a_{i-1}) \geq 0, \qquad \forall i = 1, \ldots, n,$$

or the same with negative, is at most

$$2\left(\frac{1}{2}\right)^n \to 0,$$

as $n \to \infty$ by symmetry of Gaussians. ■

More generally, we can prove the following. For a proof see [Lig10].

THM 19.11 *Almost surely, BM satisfies:*

1. *The set of times at which local maxima occur is dense.*
2. *Every local maximum is strict.*
3. *The set of local maxima is countable.*

Proof: Part (3). We use part (2). If t is a strict local maximum, it must be in the set

$$\bigcup_{n=1}^{+\infty} \left\{t \,:\, B(t,\omega) > B(s,\omega),\ \forall s,\ |s-t| < n^{-1}\right\}.$$

But for each n, the set must be countable because two such t's must be separated by n^{-1}. So the union is countable. ■

Further reading

Other constructions in [Dur10, Section8.1] and [Lig10, Section 1.5]. Proof of modulus of continuity [MP10, Theorem 1.14].

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 20 : Path properties II

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 8.1], [Lig10, Section 1.6], [MP10, Section 1.3].

1 Previous class

THM 20.1 *If $\alpha < 1/2$, then almost surely Brownian motion is everywhere locally α-Hölder continuous.*

Recall:

THM 20.2 (Scaling invariance) *Let $a > 0$. If $B(t)$ is a standard Brownian motion, then so is $X(t) = a^{-1}B(a^2 t)$.*

THM 20.3 (Time inversion) *If $B(t)$ is a standard Brownian motion, then so is*

$$X(t) = \begin{cases} 0, & t = 0, \\ tB(t^{-1}), & t > 0. \end{cases}$$

LEM 20.4 (LLN) *Almost surely, $t^{-1}B(t) \to 0$ as $t \to +\infty$.*

2 Non-differentiability

So $B(t)$ grows slower than t. But the following lemma shows that its limsup grows faster than $\sqrt{t}$.

LEM 20.5 *Almost surely*

$$\limsup_{n\to+\infty} \frac{B(n)}{\sqrt{n}} = +\infty.$$

Proof: By (FATOU),

$$\mathbb{P}[B(n) > c\sqrt{n} \text{ i.o.}] \geq \limsup_{n\to+\infty} \mathbb{P}[B(n) > c\sqrt{n}] = \limsup_{n\to+\infty} \mathbb{P}[B(1) > c] > 0,$$

by the scaling property. Thinking of $B(n)$ as the sum of $X_n = B(n) - B(n-1)$, the event on the LHS is exchangeable and the Hewitt-Savage 0-1 law implies that it has probability 1. ■

DEF 20.6 (Upper and lower derivatives) *For a function f, we define the* upper *and* lower right derivatives *as*

$$D^* f(t) = \limsup_{h \downarrow 0} \frac{f(t+h) - f(t)}{h},$$

and

$$D_* f(t) = \liminf_{h \downarrow 0} \frac{f(t+h) - f(t)}{h}.$$

We begin with an easy first result.

THM 20.7 *Fix $t \geq 0$. Then almost surely Brownian motion is not differentiable at t. Moreover, $D^* B(t) = +\infty$ and $D_* B(t) = -\infty$.*

Proof: Consider the time inversion X. Then

$$D^* X(0) \geq \limsup_{n \to +\infty} \frac{X(n^{-1}) - X(0)}{n^{-1}} = \limsup_{n \to +\infty} B(n) = +\infty,$$

by the lemma above. This proves the result at 0. Then note that $X(s) = B(t+s) - B(s)$ is a standard Brownian motion and differentiability of X at 0 is equivalent to differentiability of B at t. ■

In fact, we can prove something much stronger.

THM 20.8 *Almost surely, BM is nowhere differentiable. Furthermore, almost surely, for all t*

$$D^* B(t) = +\infty,$$

or

$$D_* B(t) = -\infty,$$

or both.

Proof: Suppose there is t_0 such that the latter does not hold. By boundedness of BM over $[0, 1]$, we have

$$\sup_{h \in [0,1]} \frac{|B(t_0 + h) - B(t_0)|}{h} \leq M,$$

for some $M < +\infty$. Assume t_0 is in $[(k-1)2^{-n}, k2^{-n}]$ for some k, n. Then for all $1 \leq j \leq 2^n - k$, in particular, for $j = 1, 2, 3$,

$$\begin{aligned} &|B((k+j)2^{-n}) - B((k+j-1)2^{-n})| \\ &\qquad \leq |B((k+j)2^{-n}) - B(t_0)| + |B(t_0) - B((k+j-1)2^{-n})| \\ &\qquad \leq M(2j+1)2^{-n}, \end{aligned}$$

by our assumption. Define the events

$$\Omega_{n,k} = \{|B((k+j)2^{-n}) - B((k+j-1)2^{-n})| \le M(2j+1)2^{-n},\ j = 1,2,3\}.$$

It suffices to show that $\cup_{k=1}^{2^n-3}\Omega_{n,k}$ cannot happen for infinitely many n. Indeed,

$$\mathbb{P}\left[\exists t_0 \in [0,1],\ \sup_{h\in[0,1]} \frac{|B(t_0+h) - B(t_0)|}{h} \le M\right] \le \mathbb{P}\left[\bigcup_{k=1}^{2^n-3} \Omega_{n,k} \text{ for infinitely many } n\right]$$

But by the independence of increments

$$\begin{aligned}
\mathbb{P}[\Omega_{n,k}] &= \prod_{j=1}^{3} \mathbb{P}[|B((k+j)2^{-n}) - B((k+j-1)2^{-n})| \le M(2j+1)2^{-n}] \\
&\le \mathbb{P}\left[|B(2^{-n})| \le \frac{7M}{2^n}\right]^3 \\
&= \mathbb{P}\left[\left|\frac{1}{\sqrt{2^{-n}}}B\left(\left[\sqrt{2^{-n}}\right]^2\right)\right| \le \frac{7M}{\sqrt{2^{-n}}\cdot 2^n}\right]^3 \\
&= \mathbb{P}\left[|B(1)| \le \frac{7M}{\sqrt{2^n}}\right]^3 \\
&\le \left(\frac{7M}{\sqrt{2^n}}\right)^3,
\end{aligned}$$

because the density of a standard Gaussian is bounded by $1/2$. Hence

$$\mathbb{P}\left[\bigcup_{k=1}^{2^n-3}\Omega_{n,k}\right] \le 2^n\left(\frac{7M}{\sqrt{2^n}}\right)^3 = (7M)^3 2^{-n/2},$$

which is summable. The result follows from BC. ■

3 Quadratic variation

Recall:

DEF 20.9 (Bounded variation) *A function $f : [0,t] \to \mathbb{R}$ is of* bounded variation *if there is $M < +\infty$ such that*

$$\sum_{j=1}^{k} |f(t_j) - f(t_{j-1})| \le M,$$

for all $k \geq 1$ *and all partitions* $0 = t_0 < t_1 < \cdots < t_k = t$*. Otherwise, we say that it is of* unbounded variation.

THM 20.10 (Quadratic variation) *Suppose the sequence of partitions*

$$0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{k(n)}^{(n)} = t,$$

is nested*, that is, at each step one or more partition points are added, and the* mesh

$$\Delta(n) = \sup_{1 \leq j \leq k(n)} \{t_j^{(n)} - t_{j-1}^{(n)}\},$$

converges to 0*. Then, almost surely,*

$$\lim_{n \to +\infty} \sum_{j=1}^{k(n)} (B(t_j^{(n)}) - B(t_{j-1}^{(n)}))^2 = t.$$

Proof: By considering subsequences, it suffices to consider the case where one point is added at each step. Let

$$X_{-n} = \sum_{j=1}^{k(n)} (B(t_j^{(n)}) - B(t_{j-1}^{(n)}))^2.$$

Let

$$\mathcal{G}_{-n} = \sigma(X_{-n}, X_{-n-1}, \ldots)$$

and

$$\mathcal{G}_{-\infty} = \bigcap_{k=1}^{\infty} \mathcal{G}_{-k}.$$

CLAIM 20.11 *We claim that* $\{X_{-n}\}$ *is a reversed MG.*

Proof: We want to show that

$$\mathbb{E}[X_{-n+1} \,|\, \mathcal{G}_{-n}] = X_{-n}.$$

In particular, this will imply by induction

$$X_{-n} = \mathbb{E}[X_{-1} \,|\, \mathcal{G}_{-n}].$$

Assume that, at step n, the new point s is added between the old points $t_1 < t_2$. Write

$$X_{-n+1} = (B(t_2) - B(t_1))^2 + W,$$

and

$$X_{-n} = (B(s) - B(t_1))^2 + (B(t_2) - B(s))^2 + W,$$

where W is independent of the other terms. We claim that

$$\mathbb{E}[(B(t_2) - B(t_1))^2 \,|\, (B(s) - B(t_1))^2 + (B(t_2) - B(s))^2] \\ = (B(s) - B(t_1))^2 + (B(t_2) - B(s))^2,$$

which follows from the following lemma.

LEM 20.12 *Let $X, Z \in \mathcal{L}^2$ be independent and assume Z is symmetric. Then*

$$\mathbb{E}[(X + Z)^2 \,|\, X^2 + Z^2] = X^2 + Z^2.$$

Proof: By symmetry of Z,

$$\begin{aligned} \mathbb{E}[(X + Z)^2 \,|\, X^2 + Z^2] &= \mathbb{E}[(X - Z)^2 \,|\, X^2 + (-Z)^2] \\ &= \mathbb{E}[(X - Z)^2 \,|\, X^2 + Z^2]. \end{aligned}$$

Taking the difference we get

$$\mathbb{E}[XZ \,|\, X^2 + Z^2] = 0.$$

■

The fact that X_{-n} is a reversed MG follows from the argument above. (Exercise.) ■

We return to the proof of the theorem. By Lévy's Downward Theorem,

$$X_{-n} \to \mathbb{E}[X_{-1} \,|\, \mathcal{G}_{-\infty}],$$

almost surely. Note that $\mathbb{E}[X_{-1}] = \mathbb{E}[X_{-n}] = t$. Moreover, by (FATOU), the variance of the limit

$$\begin{aligned} \mathbb{E}[(\mathbb{E}[X_{-1} \,|\, \mathcal{G}_{-\infty}] - t)^2] &\leq \liminf_n \mathbb{E}[(X_{-n} - t)^2] \\ &\leq \liminf_n \mathrm{Var}\left[\sum_{j=1}^{k(n)} (B(t_j^{(n)}) - B(t_{j-1}^{(n)}))^2\right] \\ &= \liminf_n 3 \sum_{j=1}^{k(n)} (t_j^{(n)} - t_{j-1}^{(n)})^2 \\ &\leq 3t \liminf_n \Delta(n) \\ &= 0. \end{aligned}$$

So finally

$$\mathbb{E}[X_{-1} \,|\, \mathcal{G}_{-\infty}] = t.$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 21 : Markov property

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 8.2], [Lig10, Section 1.7], [MP10, Section 2.1].

1 Filtrations

Recall:

DEF 21.1 (Filtration) *A* filtration *is a family* $\{\mathcal{F}(t) : t \geq 0\}$ *of sub-*σ*-fields such that* $\mathcal{F}(s) \subseteq \mathcal{F}(t)$ *for all* $s \leq t$.

We will consider two natural filtrations for BM.

DEF 21.2 *Let* $\{B(t)\}$ *be a BM. Then we denote*

$$\mathcal{F}^0(t) = \sigma(B(s) : 0 \leq s \leq t).$$

Moreover, we let

$$\mathcal{F}^+(t) = \bigcap_{s>t} \mathcal{F}^0(s).$$

Clearly $\mathcal{F}^0(t) \subseteq \mathcal{F}^+(t)$. *The latter has the advantage of being* right-continuous, *that is,*

$$\bigcap_{\varepsilon>0} \mathcal{F}^+(t+\varepsilon) = \mathcal{F}^+(t).$$

DEF 21.3 (Germ field) *The* germ σ-field *is* $\mathcal{F}^+(0)$.

EX 21.4 *Let* $B(t)$ *be a standard BM and define*

$$T = \inf\{t > 0 : B(t) > 0\}.$$

Then $\{T = 0\} \in \mathcal{F}^+(0)$ *since*

$$\{T = 0\} = \bigcap_{n\geq 1} \{\exists 0 < \varepsilon < n^{-1},\ B(\varepsilon) > 0\}.$$

2 Markov property

The basic Markov property for BM is the following.

THM 21.5 (Markov property I) *Suppose that $\{B(t)\}$ is a BM started at x. Let $s \geq 0$. Then the process $\{B(s+t) - B(s)\}_{t\geq 0}$ is a BM started at 0 and is independent of the process $\{B(t) : 0 \leq s \leq t\}$, that is, the σ-fields*

$$\sigma(B(s+t) - B(s) : t \geq 0),$$

and

$$\sigma(B(t) : 0 \leq t \leq s),$$

are independent.

Proof: We have already proved that $\{B(s+t) - B(s)\}_{t\geq 0}$ is a BM started at 0.

Further, recall:

LEM 21.6 (Independence and π-systems) *Suppose that $\mathcal{G}$ and $\mathcal{H}$ are sub-σ-algebras and that $\mathcal{I}$ and $\mathcal{J}$ are π-systems (i.e., families of subsets stable under finite intersections) such that*

$$\sigma(\mathcal{I}) = \mathcal{G}, \quad \sigma(\mathcal{J}) = \mathcal{H}.$$

Then $\mathcal{G}$ and $\mathcal{H}$ are independent if and only if $\mathcal{I}$ and $\mathcal{J}$ are, i.e.,

$$\mathbb{P}[I \cap J] = \mathbb{P}[I]\mathbb{P}[J], \quad \forall I \in \mathcal{I}, J \in \mathcal{J}.$$

Note that sets of the form

$$\{\omega : B(t_j) \in A_j,\ 0 \leq t_j \leq t,\ j = 1, \ldots, n\},$$

for $A_j \in \mathcal{B}$ are a π-system generating $\mathcal{F}^0(t)$. Similarly for $\sigma(B(s+t) - B(s) : t \geq 0)$. Therefore the independence statement immediately follows from the independence of increments. ■

In fact, we can prove a stronger statement:

THM 21.7 (Markov property II) *Suppose that $\{B(t)\}$ is a BM started at x. Let $s \geq 0$. Then the process $\{B(s+t) - B(s)\}_{t\geq 0}$ is a BM started at 0 and is independent of $\mathcal{F}^+(s)$.*

Proof: By continuity,

$$B(t+s) - B(s) = \lim_n B(s_n + t) - B(s_n),$$

for a strictly decreasing sequence $\{s_n\}_n$ converging to s. But note that for any $0 \leq t_1 < \cdots < t_j$

$$(B(t_1 + s_n) - B(s_n), \ldots, B(t_j + s_n) - B(s_n)),$$

is independent of $\mathcal{F}^+(s) \subseteq \mathcal{F}^0(s_n)$ and so is the limit. ■

3 Applications

As a first application, we get the following.

THM 21.8 (Blumenthal's 0-1 law) *For any x, the germ σ-field $\mathcal{F}^+(0)$ of a BM started at x is trivial.*

Proof: Let

$$A \in \mathcal{F}^+(0) \subseteq \sigma(B(t)\,:\,t \geq 0) = \sigma(B(t) - x\,:\,t \geq 0).$$

By the previous theorem, the two σ-fields above are independent and therefore A is independent of itself, that is,

$$\mathbb{P}[A] = \mathbb{P}[A \cap A] = \mathbb{P}[A]^2,$$

or $\mathbb{P}[A] \in \{0, 1\}$. ■

We come back to our example.

EX 21.9 *Let $B(t)$ be a standard BM and define*

$$T = \inf\{t > 0\,:\,B(t) > 0\}.$$

Then $\{T = 0\} \in \mathcal{F}^+(0)$ since

$$\{T = 0\} = \bigcap_{n \geq 1} \{\exists 0 < \varepsilon < n^{-1},\ B(\varepsilon) > 0\}.$$

Hence,

$$\mathbb{P}[T = 0] \in \{0, 1\}.$$

We show that it is 1 by showing that it is positive. Note that

$$\mathbb{P}[T \leq t] \geq \mathbb{P}[B(t) > 0] = \frac{1}{2},$$

for $t > 0$, by symmetry of the Gaussian. It also follows by continuity that

$$\inf\{t > 0\,:\,B(t) = 0\} = 0,$$

almost surely.

An immediate application of Blumenthal's 0-1 law (by time inversion) is:

THM 21.10 (0-1 law for tail events) *Let $B(t)$ be a BM. Then the tail of B, that is,*

$$\mathcal{T} = \bigcap_{t \geq 0} \mathcal{G}(t) = \bigcap_{t \geq 0} \sigma(B(s)\,:\,s \geq t),$$

is trivial.

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 22 : Strong Markov property

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 8.3], [Lig10, Section 1.8], [MP10, Section 2.2].

1 Stopping times

We first generalize stopping times to continuous time.

DEF 22.1 (Stopping time) *A RV T with values in $[0, +\infty]$ is a stopping time with respect to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ if for all $t \geq 0$,*

$$\{T \leq t\} \in \mathcal{F}(t).$$

THM 22.2 *If the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ is right-continuous in the previous definition, then an equivalent definition is obtained by using a strict inequality.*

EX 22.3 *Let G be an open set. Then*

$$T = \inf\{t \geq 0 \,:\, B(t) \in G\},$$

is a stopping time with respect to $\{\mathcal{F}^+(t)\}$. Indeed, note

$$\{T < t\} = \bigcup_{s<t,\ s\in\mathbb{Q}} \{B(s) \in G\} \in \mathcal{F}^+(t),$$

by continuity of paths and the fact that G is open.

To define the strong Markov property, we will need the following.

DEF 22.4 *Let T be a stopping time with respect to $\{\mathcal{F}^+(t)\}_{t\geq 0}$. Then we let*

$$\mathcal{F}^+(T) = \{A \,:\, A \cap \{T \leq t\} \in \mathcal{F}^+(t),\ \forall t \geq 0\}.$$

The following lemma will be useful in extending properties about discrete-time stopping times to continuous time.

LEM 22.5 *The following hold:*

1. *If T_n is a sequence of stopping times with respect to $\{\mathcal{F}(t)\}$ such that $T_n \uparrow T$, then so is T.*
2. *Let T be a stopping time with respect to $\{\mathcal{F}(t)\}$. Then the following are also stopping times:*

$$T_n = (m+1)2^{-n} \text{ if } m2^{-n} \leq T < (m+1)2^{-n}.$$

EX 22.6 *Let F be a closed set. Then*

$$T = \inf\{t \geq 0 \,:\, B(t) \in F\},$$

is a stopping time with respect to $\{\mathcal{F}^+(t)\}$. See [Lig10] for a proof.

2 Strong Markov property

THM 22.7 (Strong Markov property) *Let $\{B(t)\}_{t\geq 0}$ be a BM and T, an almost surely finite stopping time. Then the process*

$$\{B(T+t) - B(T) \,:\, t \geq 0\},$$

is a BM started at 0 independent of $\mathcal{F}^+(T)$.

Proof: Let T_n be a discretization of T as above. Let

$$B_k(t) = B(t + k2^{-n}) - B(k2^{-n}),$$

and

$$B_*(t) = B(t + T_n) - B(T_n).$$

Suppose $E \in \mathcal{F}^+(T_n)$. Then for every "finite-dimensional" event A we have, by the Markov property and time translation invariance,

$$\begin{aligned}
\mathbb{P}[\{B_* \in A\} \cap E] &= \sum_{k=1}^{+\infty} \mathbb{P}[\{B_k \in A\} \cap E \cap \{T_n = k2^{-n}\}] \\
&= \sum_{k=1}^{+\infty} \mathbb{P}[B_k \in A]\mathbb{P}[E \cap \{T_n = k2^{-n}\}] \\
&= \mathbb{P}[B \in A] \sum_{k=1}^{+\infty} \mathbb{P}[E \cap \{T_n = k2^{-n}\}] \\
&= \mathbb{P}[B \in A]\mathbb{P}[E].
\end{aligned}$$

That is, B_* is independent of $\mathcal{F}^+(T_n)$. Since $\mathcal{F}^+(T) \subseteq \mathcal{F}^+(T_n)$, B_* is also independent of $\mathcal{F}^+(T)$. Moreover, $T_n \downarrow T$ so that by continuity $\{B(t+T) - B(T)\}_{t\geq 0}$ is itself independent of $\mathcal{F}^+(T)$. The same argument shows that the increments have the correct distribution. ∎

3 Applications

We discuss one application.

THM 22.8 (Reflection principle) *Let* $\{B(t)\}_{t\geq 0}$ *be a standard BM and* T*, a stopping time. Then the process*

$$B^*(t) = B(t)\mathbb{1}\{t \leq T\} + (2B(T) - B(t))\mathbb{1}\{t > T\},$$

called BM reflected at T*, is also a standard BM.*

Proof: Follows immediately from the strong Markov property and symmetry. ■

A remarkable consequence is the following.

THM 22.9 *Let* $\{B(t)\}$ *be a standard BM and let*

$$M(t) = \max_{0\leq s\leq t} B(s).$$

Then, if $a > 0$*,*

$$\mathbb{P}[M(t) \geq a] = 2\mathbb{P}[B(t) \geq a] = \mathbb{P}[|B(t)| \geq a].$$

Proof: Let

$$T = \inf\{t \geq 0 \,:\, B(t) = a\}.$$

Then we have the disjoint union

$$\begin{aligned} \{M(t) \geq a\} &= \{B(t) \geq a\} \cup \{B(t) < a, M(t) \geq a\} \\ &= \{B(t) \geq a\} \cup \{B^*(t) > a\}. \end{aligned}$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 23 : Martingale property	
MATH275B - Winter 2012	*Lecturer: Sebastien Roch*

References: [Dur10, Section 8.5], [Lig10, Section 1.9], [MP10, Section 2.4].

1 Martingales

We first generalize MGs to continuous time.

DEF 23.1 (Continuous-time martingale) *A real-valued SP* $\{X(t)\}_{t\geq 0}$ *is a* martingale *with respect to a filtration* $\{\mathcal{F}(t)\}$ *if it is* adapted, *that is,* $X(t) \in \mathcal{F}(t)$ *for all* $t \geq 0$, *if* $E|X(t)| < +\infty$ *for all* $t \geq 0$, *and if*

$$\mathbb{E}[X(t)\,|\,\mathcal{F}(s)] = X(s),$$

almost surely, for all $0 \leq s \leq t$.

EX 23.2 *Let* $\{B(t)\}$ *be a standard BM. Then*

$$\begin{aligned}\mathbb{E}[B(t)\,|\,\mathcal{F}^+(s)] &= \mathbb{E}[B(t) - B(s)\,|\,\mathcal{F}^+(s)] + B(s)\\ &= \mathbb{E}[B(t) - B(s)] + B(s)\\ &= B(s),\end{aligned}$$

by the Markov property. Hence BM is a MG.

2 Optional stopping theorem

THM 23.3 (Optional stopping theorem) *Suppose* $\{X(t)\}_{t\geq 0}$ *is a continuous MG, and* $0 \leq S \leq T$ *are stopping times. If the process* $\{X(T \wedge t)\}_{t\geq 0}$ *is dominated by an integrable RV* X, *then*

$$\mathbb{E}[X(T)\,|\,\mathcal{F}(S)] = X(S),$$

almost surely.

Proof: Fix N and consider the discrete-time MG

$$X_n = X(T \wedge n2^{-N})$$

and the stopping times

$$S'_N = \lfloor 2^N S \rfloor + 1$$

and

$$T'_N = \lfloor 2^N T \rfloor + 1$$

with respect to the filtration

$$\mathcal{G}_n = \mathcal{F}(n2^{-N}).$$

The discrete-time optional stopping theorem gives

$$\mathbb{E}[X_{T'_N} \,|\, \mathcal{G}_{S'_N}] = X_{S'_N},$$

which is equivalent to

$$\mathbb{E}[X(T \wedge 2^{-N} T'_N) \,|\, \mathcal{F}(2^{-N} S'_N)] = \mathbb{E}[X(T) \,|\, \mathcal{F}(2^{-N} S'_N)] = X(T \wedge 2^{-N} S'_N).$$

For $A \in \mathcal{F}(S) \subseteq \mathcal{F}(2^{-N} S'_N)$, by the definition of the conditional expectation and the dominated convergence theorem,

$$\begin{aligned}
\mathbb{E}[X(T); A] &= \lim_N \mathbb{E}[\mathbb{E}[X(T) \,|\, \mathcal{F}(2^{-N} S'_N)]; A] \\
&= \mathbb{E}[\lim_N X(T \wedge 2^{-N} S'_N); A] \\
&= \mathbb{E}[X(S); A],
\end{aligned}$$

where we used continuity. ■

3 Applications

A typical application is Wald's lemma.

THM 23.4 (Wald's lemma for BM) *Let $\{B(t)\}$ be a standard BM and T a stopping time with respect to $\{\mathcal{F}^+(t)\}$ such that either:*

1. $\mathbb{E}[T] < +\infty$*, or*
2. $\{B(t \wedge T)\}$ *is dominated by an integrable RV.*

Then $\mathbb{E}[B(T)] = 0$*.*

Proof: The result under the second condition follows immediately from the optional stopping theorem with $S = 0$. We show that the first condition implies the second one.

Assume $\mathbb{E}[T] < +\infty$. Define

$$M_k = \max_{0 \leq t \leq 1} |B(t+k) - B(k)|,$$

and

$$M = \sum_{k=1}^{\lceil T \rceil} M_k,$$

and note that $|B(t \wedge T)| \leq M$.

Then

$$\begin{aligned} E[M] &= \sum_k \mathbb{E}[\mathbb{1}\{T > k-1\} M_k] \\ &= \sum_k \mathbb{P}[T > k-1] \mathbb{E}[M_k] \\ &= \mathbb{E}[M_0] \mathbb{E}[T+1] < +\infty \end{aligned}$$

by our result on the maximum from the previous lecture. ∎

We state without proof:

THM 23.5 (Wald's second lemma) *Let $\{B(t)\}$ be a standard BM and T a stopping time with respect to $\{\mathcal{F}^+(t)\}$ such that $\mathbb{E}[T] < +\infty$. Then*

$$\mathbb{E}[B(T)^2] = E[T].$$

Proof: The proof is based on the fact that $B(t)^2 - t$ is a MG. Consider

$$T_n = \inf\{t \geq 0 \,:\, |B(t)| = n\},$$

and take an appropriate limit. See [MP10] for details. ∎

An immediate application of Wald's lemma gives:

THM 23.6 *Let $\{B(t)\}$ be a standard BM. For $a < 0 < b$ let*

$$T = \inf\{t \geq 0 \,:\, B(t) \in \{a, b\}\}.$$

Then

$$\mathbb{P}[B(T) = a] = \frac{b}{|a| + b}.$$

Moreover,

$$\mathbb{E}[T] = |a| b.$$

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

Lecture 24 : Skorokhod embedding

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Dur10, Section 8.6, 8.8], [Lig10, Section 1.10], [MP10, Section 5.1, 5.3].

1 Previous class

Recall:

THM 24.1 (Wald's lemma for BM) *Let $\{B(t)\}$ be a standard BM and T a stopping time with respect to $\{\mathcal{F}^+(t)\}$ such that $\mathbb{E}[T] < +\infty$. Then*

$$\mathbb{E}[B(T)] = 0.$$

THM 24.2 (Wald's second lemma) *Let $\{B(t)\}$ be a standard BM and T a stopping time with respect to $\{\mathcal{F}^+(t)\}$ such that $\mathbb{E}[T] < +\infty$. Then*

$$\mathbb{E}[B(T)^2] = E[T].$$

THM 24.3 *Let $\{B(t)\}$ be a standard BM. For $a < 0 < b$ let*

$$T = \inf\{t \geq 0 \,:\, B(t) \in \{a, b\}\}.$$

Then

$$\mathbb{P}[B(T) = a] = \frac{b}{|a| + b}.$$

Moreover,

$$\mathbb{E}[T] = |a|b.$$

2 Skorokhod embedding

THM 24.4 (Skorokhod embedding) *Suppose $\{B(t)\}_t$ is a standard BM and that X is a RV with $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] < +\infty$. Then there exists a stopping time T with respect to $\{\mathcal{F}^+(t)\}_t$ such that $B(T)$ has the law of X and $\mathbb{E}[T] = \mathbb{E}[X^2]$.*

The proof uses a binary splitting MG:

DEF 24.5 *A* $\{X_n\}_n$ *is* binary splitting *if, whenever the event*

$$A(x_0, \ldots, x_n) = \{X_0 = x_0, \ldots, X_n = x_n\},$$

for some $x_0, \ldots, x_n$, *has positive probability, then the RV* X_{n+1} *conditioned on* $A(x_0, \ldots, x_n)$ *is supported on at most two values.*

LEM 24.6 *Let* X *be a RV with* $\mathbb{E}[X^2] < +\infty$. *Then there is a binary splitting MG* $\{X_n\}_n$ *such that* $X_n \to X$ *almost surely and in* $\mathcal{L}^2$.

Proof:(of Lemma) The MG is defined recursively. Let

$$\mathcal{G}_0 = \{\emptyset, \Omega\},$$

and

$$X_0 = \mathbb{E}[X].$$

For $n > 0$, we let

$$\xi_n = \begin{cases} 1, & \text{if } X \geq X_n \\ -1, & \text{if } X < X_n, \end{cases}$$

and

$$\mathcal{G}_n = \sigma(\xi_0, \ldots, \xi_{n-1}),$$

and

$$X_n = \mathbb{E}[X \mid \mathcal{G}_n].$$

Then $\{X_n\}_n$ is a binary splitting MG. It remains to prove the convergence claim.

By (cJENSEN)

$$\mathbb{E}[X_n^2] \leq \mathbb{E}[X^2],$$

so $\{X_n\}_n$ is bounded in $\mathcal{L}^2$ and we have by Lévy's upward theorem

$$X_n \to X_\infty = \mathbb{E}[X \mid \mathcal{G}_\infty],$$

almost surely and in $\mathcal{L}^2$, where

$$\mathcal{G}_\infty = \sigma\left(\bigcup_i \mathcal{G}_i\right).$$

We need to show that $X = X_\infty$.

CLAIM 24.7 *Almost surely,*

$$\lim_n \xi_n(X - X_{n+1}) = |X - X_\infty|.$$

We first finish the proof of the lemma. Note that

$$\mathbb{E}[\xi_n(X - X_{n+1})] = \mathbb{E}[\xi_n \mathbb{E}[X - X_{n+1} \,|\, \mathcal{G}_{n+1}]] = 0.$$

Since $\{\xi_n(X - X_{n+1})\}_n$ is bounded in $\mathcal{L}^2$, the expectations converge and

$$\mathbb{E}|X - X_\infty| = 0.$$

Finally we prove the claim. If $X = X_\infty$, both sides are 0. If $X < X_\infty$, then for n large enough, $X < X_n$ and $\xi_n = -1$ by construction and the result holds. Similarly for the other case. ■

Proof:(of Theorem) Take a binary splitting MG as in the previous lemma. Since X_n conditioned on $A(x_0, \ldots, x_{n-1})$ is supported on two values, we can use the stopping time from last time and we get a sequence of stopping times

$$T_0 \leq T_1 \leq \cdots \leq T_n \leq \cdots \uparrow T$$

for some T such that

$$B(T_n) \sim X_n,$$

and

$$\mathbb{E}[T_n] = \mathbb{E}[B(T_n)^2].$$

By (MON) and $\mathcal{L}^2$ convergence

$$\mathbb{E}[T] = \lim_n \mathbb{E}[T_n] = \lim_n \mathbb{E}[X_n^2] = \mathbb{E}[X].$$

By continuity of paths,

$$B(T_n) \to B(T), \qquad \text{a.s.}$$

and

$$B(T) \sim X.$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[Lig10] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.