Notes 18 : Optional Sampling Theorem

Math 733-734: Theory of Probability *Lecturer: Sebastien Roch*

References: [Wil91, Chapter 14], [Dur10, Section 5.7].

Recall:

DEF 18.1 (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 18.2 (Necessary and Sufficient Condition for $\mathcal{L}^1$ Convergence) *Let $\{X_n\} \in \mathcal{L}^1$ and $X \in \mathcal{L}^1$. Then $X_n \to X$ in $\mathcal{L}^1$ if and only if the following two conditions hold:*

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

THM 18.3 (Convergence of UI MGs) *Let $\{M_n\}$ be UI MG. Then*

$$M_n \to M_\infty \in \mathcal{F}_\infty = \sigma\left(\cup_n \mathcal{F}_n\right),$$

a.s. and in $\mathcal{L}^1$. Moreover,

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

THM 18.4 (Lévy's upward theorem) *Let $Z \in \mathcal{L}^1$ and define $M_n = \mathbb{E}[Z \,|\, \mathcal{F}_n]$. Then $\{M_n\}$ is a UI MG and*

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

a.s. and in $\mathcal{L}^1$.

1 Optional Sampling Theorem

1.1 Review: Stopping times

Recall:

DEF 18.5 *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 18.6 *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.

LEM 18.7 (Stopping Time Lemma) *Let* $\{M_n\}$ *be a MG and* T *be a stopping time. Then the stopped process* $\{M_{T\wedge n}\}$ *is a MG and in particular*

$$\mathbb{E}[M_{T\wedge n}] = \mathbb{E}[M_0].$$

THM 18.8 *Let* $\{M_n\}$ *be a MG and* T *be a stopping time. Then* $M_T \in \mathcal{L}^1$ *and*

$$\mathbb{E}[M_T] = \mathbb{E}[M_0].$$

if any of the following conditions holds:

1. T *is bounded*
2. $\{M_n\}$ *is bounded and* T *is a.s. finite*
3. $\mathbb{E}[T] < +\infty$ *and* $\{M_n\}$ *has bounded increments*
4. $\{M_n\}$ *is UI. (This one is new. The proof follows from the Optional Sampling Theorem below.)*

Proof: From the previous theorem, we have

$$(*) \quad \mathbb{E}[M_{T\wedge n} - M_0] = 0.$$

1. Take $n = N$ in $(*)$ where $T \leq N$ a.s.
2. Take n to $+\infty$ and use (DOM).

3. Note that
$$|M_{T\wedge n} - M_0| = \left|\sum_{i\leq T\wedge n}(M_i - M_{i-1})\right| \leq KT,$$
where $|M_n - M_{n-1}| \leq K$ a.s. Use (DOM).

■

DEF 18.9 ($\mathcal{F}_T$) *Let T be a stopping time. Denote by $\mathcal{F}_T$ the set of all events F such that $\forall n \in \overline{\mathbb{Z}}_+$*
$$F \cap \{T = n\} \in \mathcal{F}_n.$$

1.2 More on the σ-field $\mathcal{F}_T$

The following two lemmas help clarify the definition of $\mathcal{F}_T$:

LEM 18.10 *$\mathcal{F}_T = \mathcal{F}_n$ 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 $F = F \cap \{T = n\} \in \mathcal{F}_n$. 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 18.11 *If $\{X_n\}$ 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., $X_\infty = \liminf_n X_n$).*

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 18.12 *If S, T are stopping times, then $S \wedge T$ is a stopping time and $\mathcal{F}_{S\wedge T} \subseteq \mathcal{F}_T$.*

Proof: We first show that $S \wedge T$ is a stopping time. Note that
$$\{S \wedge T = k\} = [\{S = k\} \cap \{T \geq k\}] \cup [\{S \geq k\} \cap \{T = k\}] \in \mathcal{F}_k,$$
since all event above are in $\mathcal{F}_k$ by the fact that S and T are themselves stopping times.

For the second claim, let $F \in \mathcal{F}_{S\wedge T}$. Note that
$$F \cap \{T = n\} = \cup_{k\leq n}[(F \cap \{S \wedge T = k\}) \cap \{T = n\}] \in \mathcal{F}_n.$$
Indeed, the expression in parenthesis is in $\mathcal{F}_k \subseteq \mathcal{F}_n$ and $\{T = n\} \in \mathcal{F}_n$. ■

1.3 Optional Sampling Theorem (OST)

We show that the MG property extends to stopping times under UI MGs.

THM 18.13 (Optional Sampling Theorem) *If $\{M_n\}$ 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 $\{M_n\}$ 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 18.14

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

Indeed, we then get the theorem by (TOWER) (and (JENSEN) for the integrability claim).

Proof:(of the lemma) We divide $M_\infty = M_\infty^+ - M_\infty^- \equiv X_\infty - Y_\infty$ into positive and negative parts and write

$$M_n = \mathbb{E}[M_\infty \,|\, \mathcal{F}_n] = \mathbb{E}[X_\infty \,|\, \mathcal{F}_n] - \mathbb{E}[Y_\infty \,|\, \mathcal{F}_n] \equiv X_n - Y_n,$$

by linearity. We show that $\mathbb{E}[X_\infty \,|\, \mathcal{F}_T] = X_T$. The same argument holds for $\{Y_n\}$, which then implies $\mathbb{E}[M_\infty \,|\, \mathcal{F}_T] = X_T - Y_T = M_T$, as claimed.

We have of course that $X_\infty \geq 0$, and hence $X_n = \mathbb{E}[X_\infty \,|\, \mathcal{F}_n] \geq 0 \;\forall n$. Let $F \in \mathcal{F}_T$. Then

$$\mathbb{E}[X_\infty; F \cap \{T = \infty\}] = \mathbb{E}[X_T; F \cap \{T = \infty\}],$$

since $X_n = \mathbb{E}[X_\infty \,|\, \mathcal{F}_n] \to X_\infty$ a.s. by Lévy's Upward Theorem. Therefore it suffices to show

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

In fact, by (MON), it suffices to show

$$\sum_{i \geq 0} \mathbb{E}[X_\infty; F \cap \{T = i\}] = \sum_{i \geq 0} \mathbb{E}[X_T; F \cap \{T = i\}].$$

But note that $F \cap \{T = i\} \in \mathcal{F}_i$ so that

$$\mathbb{E}[X_T; F \cap \{T = i\}] = \mathbb{E}[X_i; F \cap \{T = i\}] = \mathbb{E}[X_\infty; F \cap \{T = i\}],$$

since $X_i = \mathbb{E}[X_\infty \,|\, \mathcal{F}_i]$. That concludes the proof of the stronger claim. ■ ■

2 Wald's identities

Often additional properties of T hold (typically $\mathbb{E}[T] < +\infty$), which can be taken advantage of by considering instead the MG $\{M_{T\wedge n}\}$ in the OST above (noticing of course that $M_{T\wedge S} = M_S$ and $M_{T\wedge T} = M_T$ whenever $S \leq T$ a.s.). In that context, the following is useful.

LEM 18.15 *Suppose $\{M_n\}$ is a MG such that $\mathbb{E}[\,|M_{n+1} - M_n|\,|\,\mathcal{F}_n] \leq B$ a.s. for all n. Suppose T is a stopping time with $\mathbb{E}[T] < +\infty$. Then the stoppped MG $\{M_{T\wedge n}\}$ is UI.*

Proof: Assume WLOG that $M_0 = 0$, to simplify. Observe first that

$$|M_{T\wedge n}| \leq \sum_{m=0}^{+\infty} |M_{m+1} - M_m|\, \mathbb{1}_{\{T>m\}}, \qquad \forall n.$$

Taking expectations on the RHS (which does not depend on n), we get

$$\begin{aligned}
\sum_{m=0}^{+\infty} \mathbb{E}\left[|M_{m+1} - M_m|\, \mathbb{1}_{\{T>m\}}\right] &= \sum_{m=0}^{+\infty} \mathbb{E}\left[|M_{m+1} - M_m|\, \mathbb{1}_{\{T>m\}}\right] \\
&= \sum_{m=0}^{+\infty} \mathbb{E}\left[\mathbb{E}\left[|M_{m+1} - M_m|\, \mathbb{1}_{\{T>m\}}\,|\,\mathcal{F}_m\right]\right] \\
&= \sum_{m=0}^{+\infty} \mathbb{E}\left[\mathbb{E}\left[|M_{m+1} - M_m|\,|\,\mathcal{F}_m\right] \mathbb{1}_{\{T>m\}}\right] \\
&\leq \sum_{m=0}^{+\infty} \mathbb{E}\left[B\mathbb{1}_{\{T>m\}}\right] \\
&\leq B\sum_{m=0}^{+\infty} \mathbb{P}\left[T > m\right] \\
&\leq B\,\mathbb{E}[T] < +\infty,
\end{aligned}$$

where we used that $\{T > m\} \in \mathcal{F}_m$.

■

As an application, we recover Wald's first identity. For $X_1, X_2, \ldots \in \mathbb{R}$, let $S_n = \sum_{i=1}^n X_i$.

THM 18.16 (Wald's first identity) *Let $X_1, X_2, \ldots \in \mathcal{L}^1$ be i.i.d. with $\mu = \mathbb{E}[X_1]$ and let $T \in \mathcal{L}^1$ be a stopping time. Then*

$$\mathbb{E}[S_T] = \mu\mathbb{E}[T].$$

Proof: Recall that $M_n = S_n - n\mu$ is a MG. By LEM 18.15 and the assumption $T \in \mathcal{L}^1$, the MG $\{M_{T\wedge n}\}$ is UI. Indeed

$$\begin{aligned}\mathbb{E}\left[|M_{n+1} - M_n| \,|\, \mathcal{F}_n\right] &= \mathbb{E}\left[|X_{n+1} - \mu| \,|\, \mathcal{F}_n\right] \\ &\leq \mu + \mathbb{E}|X_1| \equiv B < +\infty,\end{aligned}$$

by the triangle inequality and the Role of independence lemma. Apply THM 18.8 to $\{M_{T\wedge n}\}$. ■

We also recall Wald's second identity. We give a MG-based proof (but argue about convergence directly rather than using THM 18.8).

THM 18.17 (Wald's second identity) *Let $X_1, X_2, \ldots \in \mathcal{L}^2$ be i.i.d. with $\mathbb{E}[X_1] = 0$ and $\sigma^2 = \mathrm{Var}[X_1]$ and let $T \in \mathcal{L}^1$ be a stopping time. Then*

$$\mathbb{E}[S_T^2] = \sigma^2 \mathbb{E}[T].$$

Proof: Recall that $M_n = S_n^2 - n\sigma^2$ is a MG. Hence so is $M_{T\wedge n}$ and

$$0 = \mathbb{E}[M_{T\wedge n}] = \mathbb{E}[S_{T\wedge n}^2 - (T\wedge n)\sigma^2] = \mathbb{E}[S_{T\wedge n}^2] - \sigma^2\mathbb{E}[T\wedge n]. \qquad (1)$$

We have that $\mathbb{E}[T\wedge n] \uparrow \mathbb{E}[T]$ as $n \to +\infty$ by (MON).

To argue about the convergence of $\mathbb{E}[S_{T\wedge n}^2]$ we note that, by the assumption $\mathbb{E}[X_1] = 0$, it follows that $\{S_n\}$ is a MG and hence so is $\{S_{T\wedge n}\}$. The latter is bounded in $\mathcal{L}^2$ since, by (1), we have

$$\mathbb{E}[S_{T\wedge n}^2] = \sigma^2\mathbb{E}[T\wedge n] \leq \sigma^2\mathbb{E}[T] < +\infty,$$

for all n. Hence $S_{T\wedge n}$ converges a.s. and in $\mathcal{L}^2$ to S_T (since $T < +\infty$ a.s. by assumption). Convergence in $\mathcal{L}^2$ also implies convergence of the second moment. Indeed, by the triangle inequality,

$$|\|S_{T\wedge n}\|_2 - \|S_T\|_2| \leq \|S_{T\wedge n} - S_T\|_2 \to 0.$$

Hence,

$$0 = \mathbb{E}[S_{T\wedge n}^2] - \sigma^2\mathbb{E}[T\wedge n] \to \mathbb{E}[S_T^2] - \sigma^2\mathbb{E}[T],$$

which concludes the proof. ■

To establish $\mathbb{E}[T] < +\infty$, the following lemma can be used.

LEM 18.18 (Waiting for the inevitable) *Let T be a stopping time. Assume there is $N \in \mathbb{Z}_+$ and $\varepsilon > 0$ such that for every n*

$$\mathbb{P}[T \leq n + N \,|\, \mathcal{F}_n] > \varepsilon \quad a.s.$$

then $\mathbb{E}[T] < +\infty$.

Proof: For any integer $m \geq 1$,

$$\mathbb{P}[T > mN \mid T > (m-1)N] \leq 1 - \varepsilon,$$

by assumption. (Indeed, by definition of $Z = \mathbb{P}[T > n + N \mid \mathcal{F}_n] \leq 1 - \varepsilon$ with $n = (m-1)N$, we have for $F = \{T > (m-1)N\} \in \mathcal{F}_n$

$$\mathbb{P}[T > mN] = \mathbb{E}[\mathbb{1}\{T > n + N\}; F] = \mathbb{E}[Z; F] \leq (1-\varepsilon)\mathbb{P}[F],$$

and apply the definition of the conditional probability.) By the multiplication rule (i.e., the undergraduate rule $\mathbb{P}[A_1 \cap \cdots \cap A_n] = \prod_{i=1}^n \mathbb{P}[A_i \mid A_1 \cap \cdots \cap A_{i-1}]$) and the monotonicity of the events $\{T > mN\}$, we have $\mathbb{P}[T > mN] \leq (1-\varepsilon)^m$. We conclude using $\mathbb{E}[T] = \sum_{k \geq 1} \mathbb{P}[T \geq k]$. ■

3 Application I: Simple RW

DEF 18.19 Simple RW on $\mathbb{Z}$ *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.* $\mathbb{P}[X_1 = 1] = 1/2$.

THM 18.20 *Let* $\{S_n\}$ *as above. Let* $a < 0 < b$*. Define* $T_x = \inf\{n \geq 0 : S_n = x\}$ *and* $T = T_a \wedge T_b$*. Then we have*

1. $$T < +\infty \text{ a.s.}$$

2. $$\mathbb{P}[T_a < T_b] = \frac{b}{b-a}$$

3. $$\mathbb{E}[T] = -ab$$

4. $$T_a < +\infty \text{ a.s. but } \mathbb{E}[T_a] = +\infty$$

Proof:

1) Apply the Waiting for the inevitable lemma with $N = b - a$ and $\varepsilon = (1/2)^{b-a}$ (corresponding to moving right $b - a$ times in a row which takes you to b, no matter where you are within the $\{a, \ldots, b\}$ interval). That shows $\mathbb{E}[T] < +\infty$, from which the claim holds.

2) By Wald's first identity, $\mathbb{E}[S_T] = 0$ or

$$a\,\mathbb{P}[S_T = a] + b\,\mathbb{P}[S_T = b] = 0,$$

that is (taking $b \to \infty$ in the second expression)

$$\mathbb{P}[T_a < T_b] = \frac{b}{b-a} \quad \text{and} \quad \mathbb{P}[T_a < +\infty] \geq \mathbb{P}[T_a < T_b] \to 1.$$

3) Wald's second identity says that $\mathbb{E}[S_T^2] = \mathbb{E}[T]$ (by $\sigma^2 = 1$). Also

$$\mathbb{E}[S_T^2] = \frac{b}{b-a}a^2 + \frac{-a}{b-a}b^2 = -ab,$$

so that $\mathbb{E}[T] = -ab$.

4) Taking $b \to +\infty$ above shows that $\mathbb{E}[T_a] = +\infty$ by monotone convergence. (Note that this case shows that the $\mathcal{L}^1$ condition on the stopping time is necessary in Wald's second identity.)

■

4 Application II: Biased RW

DEF 18.21 Biased simple RW on $\mathbb{Z}$ *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 18.22 *Let* $\{S_n\}$ *as above. Let* $a < 0 < b$*. Define* $T_x = \inf\{n \geq 0 : S_n = x\}$ *and* $T = T_a \wedge T_b$*. Then we have*

1.

$$T < +\infty \text{ a.s.}$$

2.

$$\mathbb{P}[T_a < T_b] = \frac{\phi(0) - \phi(b)}{\phi(a) - \phi(b)}$$

3.

$$\mathbb{P}[T_a < +\infty] = 1/\phi(a) < 1 \text{ and } \mathbb{P}[T_b = +\infty] = 0$$

4.

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

Proof: There are two MGs here:

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

(noting that $|\phi(S_n)| \leq (p/q)^n$ a.s.) and

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

(noting that $|\psi_n(S_n)| \leq (1+p)n$ a.s.)

1) Follows by the same argument as in the unbiased case.

2) Now note that $\{\phi(S_{T\wedge n})\}$ is a bounded MG and, therefore, by THM 18.8, we get

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

or $\mathbb{P}[T_a < T_b] = \frac{\phi(b)-\phi(0)}{\phi(b)-\phi(a)}$ (where we used 1)).

3) By 2), taking $b \to +\infty$, by monotonicity $\mathbb{P}[T_a < +\infty] = \frac{1}{\phi(a)} < 1$ so $T_a = +\infty$ with positive probability. Similarly take $a \to -\infty$.

4) By LEM 18.7 applied to $\{\Psi_n(S_n)\}$,

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

(We cannot use Wald's first identity directly because it is not immediately clear whether T_b is integrable.) By (MON) and the fact that $T_b < +\infty$ a.s. from 3), $\mathbb{E}[T_b \wedge n] \uparrow \mathbb{E}[T_b]$. Finally, $-\inf_n S_n \geq 0$ a.s. and for $x \geq 0$,

$$\mathbb{P}[-\inf_n S_n \geq x] = \mathbb{P}[T_{-x} < +\infty] = \left(\frac{q}{p}\right)^x,$$

so that $\mathbb{E}[-\inf_n S_n] = \sum_{x\geq 1} \mathbb{P}[-\inf_t S_t \geq x] < +\infty$. Hence, we can use (DOM) with $|S_{T_b\wedge n}| \leq \max\{b, -\inf_n S_n\}$ to deduce that

$$\mathbb{E}[T_b] = \frac{\mathbb{E}[S_{T_b}]}{p-q} = \frac{b}{2p-1}.$$

■

References

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

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