Notes 17 : UI Martingales

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

References: [Wil91, Chapter 13, 14], [Dur10, Section 5.5, 5.6].

1 Uniform Integrability

We give a characterization of $\mathcal{L}^1$ convergence (which has nothing to do per se with MGs). First a simple example.

EX 17.1 ($\mathcal{L}^1$-boundedness is not sufficient) *Let $\{X_n\}$ be a sequence of independent RVs. Let X_n be 0 with probability $1-p_n$ and $f_n > 0$ with probability p_n with $p_n \in [0,1]$. Assume $p_n = 1/n^2$. Then $\sum_n \mathbb{P}[X_n \neq 0] < +\infty$ and, by BC1, $\mathbb{P}[X_n \neq 0 \text{ i.o.}] = 0$ and $X_n \to X_\infty \equiv 0$ a.s. Assume further that $f_n = n^2$. Then $\|X_n - X_\infty\|_1 = \mathbb{E}[X_n] = 1$ for all $n \geq 1$, so the sequence $\{X_n\}$ does not converge in $\mathcal{L}^1$. Observe in particular that $\{X_n\}$ is bounded in $\mathcal{L}^1$, showing that the latter condition is not sufficient for $\mathcal{L}^1$ convergence. On the other hand, if $f_n = n$, we then have $\|X_n - X_\infty\|_1 = \mathbb{E}[X_n] = 1/n \to 0$ and convergence in $\mathcal{L}^1$ holds in that case. In other words, unlike almost sure convergence, convergence in $\mathcal{L}^1$ is sensitive to the size of rare deviations. (For the record, here is an example where one has convergence in $\mathcal{L}^1$ but not a.s. Take $f_n = 1$ for all n above. Then a.s. convergence to 0 occurs iff $\sum_n p_n < +\infty$ by BC1 and BC2. On the other hand, convergence in $\mathcal{L}^1$, which is equivalent to convergence in probability in this case, occurs exactly when $p_n \to 0$.)*

It turns out that what we need is for the following property of integrable variables to hold uniformly over a collection of RVs.

LEM 17.2 *Let $Y \in \mathcal{L}^1$. $\forall \varepsilon > 0$, $\exists K > 0$ s.t.*

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

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

DEF 17.3 (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 17.4 (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*

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

EX 17.5 (UI implies $\mathcal{L}^1$-boundedness) *Let $\mathcal{C}$ be UI and $X \in \mathcal{C}$. Note that*

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

and this bound is the same for any $X \in \mathcal{C}$. So UI implies $\mathcal{L}^1$-boundedness. But the opposite is not true by the construction in EX 17.1 (in that example, when $f(n) = n^2$, for any K we have $\mathbb{E}[|X_n|;|X_n| > K] = 1$ for n large enough).

EX 17.6 ($\mathcal{L}^p$-bounded RVs) *But $\mathcal{L}^p$-boundedness works—for $p > 1$. Let $\mathcal{C}$ be $\mathcal{L}^p$-bounded and $X \in \mathcal{C}$. Then*

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

as $K \to +\infty$, where $A_p = \sup_{X\in\mathcal{C}} \|X\|_p^p < +\infty$ by assumption.

EX 17.7 (Dominated RVs) *Assume $\exists Y \in \mathcal{L}^1$ s.t. $|X| \leq Y$ a.s., $\forall X \in \mathcal{C}$. Then*

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

and apply LEM 17.2 above to establish UI.

2 Proof of main theorem

Proof: We start with the if part. By the bounded convergence theorem (convergence in probability version), convergence in probability implies convergence in $\mathcal{L}^1$ for uniformly bounded variables.

LEM 17.8 (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: 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. ■

It is natural to truncate at K to apply the UI property and extend the claim above to unbounded variables. 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,

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

For the first term, check by case analysis that

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

so that $\phi_K(X_n) \to_P \phi_K(X)$. For K large enough, the 2nd term above is $\leq \varepsilon/3$ by UI and the 3rd term is $\leq \varepsilon/3$ by LEM 17.2 above.

We move on to the proof of the only if part. Suppose $X_n \to X$ in $\mathcal{L}^1$. We know that convergence in $\mathcal{L}^1$ implies convergence in probability by Markov's inequality. So the first claim follows. For the second claim, if $n \geq N$ large enough,

$$\mathbb{E}|X_n - X| \leq \varepsilon. \tag{1}$$

We can choose K large enough so that

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

$\forall n < N$ because $X_n \in \mathcal{L}^1, \forall n$, and N is finite. So we only need to worry about $n \geq N$. To use $\mathcal{L}^1$ convergence, it is natural to write

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

The first term is $\leq \varepsilon$ by (1). The issue with the second term is that we cannot apply LEM 17.2 because the restriction event involves X_n rather than X. In fact, a stronger version of the lemma exists:

LEM 17.9 (Absolute continuity) *Let $X \in \mathcal{L}^1$. $\forall \varepsilon > 0$, $\exists \delta > 0$, s.t. $\mathbb{P}[F] < \delta$ implies*

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

Proof: Argue by contradiction. Suppose there is $\varepsilon > 0$ and F_n s.t. $\mathbb{P}[F_n] \leq 2^{-n}$ and

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

for all n. By BC1,

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

where H is implicitly defined in the equation. By reverse Fatou (applied to $|X|\mathbb{1}_H = \limsup |X|\mathbb{1}_{F_n} \leq |X| \in \mathcal{L}^1$),

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

in contradiction to $\mathbb{P}[H] = 0$. ■

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.

■

Finally, we note that a uniform version of the condition in LEM 17.9 (together with $\mathcal{L}^1$-boundedness) is equivalent to UI.

LEM 17.10 *A collection $\mathcal{C}$ of RVs on $(\Omega, \mathcal{F}, \mathbb{P})$ is UI if and only if:*

1. *$\mathcal{C}$ is bounded in $\mathcal{L}^1$*
2. *$\forall \varepsilon > 0$, $\exists \delta > 0$, s.t. $\mathbb{P}[F] < \delta$ implies*

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

Proof: If $\mathcal{C}$ is UI, then it is bounded in $\mathcal{L}^1$ by EX 17.5. For any $\varepsilon' > 0$, $\varepsilon = \varepsilon'/2$, and $\mathbb{P}[F] < \delta'$,

$$\mathbb{E}[|X|; F] \leq K\mathbb{P}[F] + \mathbb{E}[|X|; \{|X| > K\}] \leq K\delta' + \varepsilon \leq \varepsilon',$$

by taking K large enough (by UI), and then δ' small enough.

On the other hand, if the two conditions above hold, take $F = \{|X| > K\}$ and use Markov's inequality and boundedness in $\mathcal{L}^1$ to choose K large enough that $\mathbb{P}[F] < \delta$ and hence $\mathbb{E}[|X|; F] < \varepsilon$ for all $X \in \mathcal{C}$. ■

3 UI MGs

THM 17.11 (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,

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

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

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

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

by definition of the conditional expectation. We can take a limit by $\mathcal{L}^1$-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 $\mathbb{E}[M_\infty \,|\, \mathcal{F}_n] = M_n$. ■

4 Applications I

THM 17.11 says that any UI MG is a Doob's MG. Conversely:

THM 17.12 (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 $\mathcal{L}^1$.

Proof: $\{M_n\}$ is a MG by (TOWER). We first show it is UI:

LEM 17.13 *Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then*

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

is UI.

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

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

where we used Taking Out What is Known (backwards) on the third line and (TOWER) on the fourth line. By Markov and (JENSEN)

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

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

In particular, we have convergence a.s. and in $\mathcal{L}^1$ 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 the Uniqueness of Extensions lemma, it suffices to prove the equality over all $\mathcal{F}_n$. If $F \in \mathcal{F}_n \subseteq \mathcal{F}_\infty$, then

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

The first equality is by definition of Y; the second equality comes from the fact that $\mathbb{E}[Y \,|\, \mathcal{F}_n] = \mathbb{E}[Z \,|\, \mathcal{F}_n] = M_n$ by (TOWER); the third equality is from our main theorem. ■

A statistical application:

EX 17.14 (Posterior mean consistency) *Let Θ be a RV with a finite mean. Assume we observe the sequence $\{Y_n\}$ with $Y_n = \Theta + Z_n$, where $\{Z_n\}$ is iid with with mean 0. If our goal is to recover Θ from $\{Y_n\}$, a natural strategy is to employ the Strong Law of Large Numbers, which implies*

$$\frac{1}{n}\sum_{i \leq n} Y_i = \Theta + \frac{1}{n}\sum_{i \leq n} Z_i \to \Theta$$

almost surely, showing in particular that $\Theta \in \mathcal{F}_\infty$ if we let $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$. A more "Bayesian" approach to recover Θ is to consider instead the "posterior mean"

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

By Lévy's upward theorem,

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

a.s. and in $\mathcal{L}^1$. Because $\Theta \in \mathcal{F}_\infty$, by Taking Out What is Known we also have

$$M_n \to \Theta,$$

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

We use Lévy's Downward Theorem to prove Lévy's 0-1 Law.

THM 17.15 (Lévy's 0-1 law) *Let $A \in \mathcal{F}_\infty$. Then*

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

Proof: Immediate since $\mathbb{E}[\mathbb{1}_A \,|\, \mathcal{F}_\infty] = \mathbb{1}_A$ by Taking Out What Is Known. ■
Recall that the tail σ-field of a sequence $\{X_n\}$ is

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

COR 17.16 (Kolmogorov's 0-1 law) *Let $X_1, X_2, \ldots$ be iid RVs. 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 the Role of Independence. By Lévy's 0-1 law,

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

■

5 Applications II

Going "backwards in time:"

THM 17.17 (Lévy's downward theorem) *Let $Z \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ 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 $\mathcal{L}^1$.

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

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

Note that

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

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 LEM 17.13, $\{M_{-n}\}$ is UI and hence we have $\mathcal{L}^1$ 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 $\mathcal{L}^1$ convergence. ■

5.1 Law of large numbers

An application:

THM 17.18 (Strong Law; Martingale Proof) *Let* $X_1, X_2, \ldots$ *be iid RVs with* $\mathbb{E}|X_1| < +\infty$. *Let* $S_n = \sum_{i\leq n} X_n$. *Then*

$$n^{-1}S_n \to \mathbb{E}[X_1],$$

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

Proof: Let

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

The key observation is that $\mathbb{E}[X_1 \,|\, \mathcal{G}_{-n}] = n^{-1}S_n$. Indeed 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 and linearity of expectation. By Lévy's Downward Theorem

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

a.s. and in $\mathcal{L}^1$. But the limit must be trivial by Kolmogorov's 0-1 law and we must have $\mathbb{E}[X_1 \,|\, \mathcal{G}_{-\infty}] = \mathbb{E}[X_1]$. ∎

5.2 Hewitt-Savage*

DEF 17.19 *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 X_is that leave $X_{n+1}, X_{n+2}, \ldots$ unchanged. The* exchangeable σ-field *is $\mathcal{E} = \cap_m \mathcal{E}_m$.*

THM 17.20 (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 an application of 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 $\mathcal{L}^2$ 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 Lévy's Downward Theorem implies

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

2. We make ϕ "exchangeable" by averaging over all configurations and taking a limit as $n \to +\infty$. 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. The reason we did this is that now the first k Xs have little influence on this quantity and therefore the limit is independent of them. However, note that

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

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, 2010.

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