Backwards Martingales and Exchangeability

김지수 (Jisu KIM)

확률론 2 (Probability Theory 2), 2025 2nd semester (fall)

Some parts of this lecture note are from the lecture notes from Prof. Alessandro Rianldo's "Advanced Probability Overview".

Backwards Martingales

A backwards martingale (some authors call them *reversed*) is a martingale indexed by the negative integers, i.e. X_n, $n \leq 0$, adapted to an increasing sequence of σ-fields $\mathcal{F}_n$ with

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n, \quad n \leq -1.$$

Because the σ-fields decrease as $n \downarrow -\infty$, the convergence theory for backwards martingales is particularly simple.

Theorem ([2, Theorem 4.7.1])**.** *$X_{-\infty} = \lim_{n \to -\infty} X_n$ exists a.s. and in L^1.*

Proof. Let U_n be the number of upcrossings of $[a, b]$ by $X_{-n}, \ldots, X_0$. The upcrossing inequality [2, Theorem 4.2.10] implies

$$(b-a)\mathbb{E}U_n \leq \mathbb{E}(X_0 - a)^+.$$

Letting $n \to \infty$ and using the monotone convergence theorem, we have $\mathbb{E}U_\infty < \infty$, so by the remark after [2, Theorem 4.2.11], the limit $X_{-\infty}$ exists a.s.. The martingale property implies $X_n = \mathbb{E}[X_0 \mid \mathcal{F}_n]$, so [2, Theorem 4.6.1] implies X_n is uniformly integrable and [2, Theorem 4.6.3] tells us that the convergence occurs in L^1. □

The next result identifies the limit in [2, Theorem 4.7.1].

Theorem ([2, Theorem 4.7.2])**.** *If $X_{-\infty} = \lim_{n \to -\infty} X_n$ and $\mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_n$, then*

$$X_{-\infty} = \mathbb{E}[X_0 \mid \mathcal{F}_{-\infty}].$$

Proof. Clearly $X_{-\infty} \in \mathcal{F}_{-\infty}$. Since $X_n = \mathbb{E}[X_0 \mid \mathcal{F}_n]$, for $A \in \mathcal{F}_{-\infty} \subset \mathcal{F}_n$,

$$\int_A X_n \, dP = \int_A X_0 \, dP.$$

[2, Theorem 4.7.1] and [2, Lemma 4.6.5] imply $\mathbb{E}[X_n; A] \to \mathbb{E}[X_{-\infty}; A]$, so

$$\int_A X_{-\infty} \, dP = \int_A X_0 \, dP,$$

for all $A \in \mathcal{F}_{-\infty}$, proving the claim. □

The next result is [2, Theorem 4.6.8] backwards

Theorem ([2, Theorem 4.7.3])**.** *If $\mathcal{F}_n \downarrow \mathcal{F}_{-\infty}$ as $n \downarrow -\infty$ (i.e. $\mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_n$), then*

$$\mathbb{E}[Y \mid \mathcal{F}_n] \to \mathbb{E}[Y \mid \mathcal{F}_{-\infty}] \qquad \textit{a.s. and in } L^1.$$

Proof. $X_n = \mathbb{E}[Y \mid \mathcal{F}_n]$ is a backwards martingale, so [2, Theorem 4.7.1]-[2, Theorem 4.7.2] imply that as $n \downarrow -\infty$, $X_n \to X_{-\infty}$ a.s. and in L^1, where

$$X_{-\infty} = \mathbb{E}[X_0 \mid \mathcal{F}_{-\infty}] = \mathbb{E}\left[\mathbb{E}[Y \mid \mathcal{F}_0] \mid \mathcal{F}_{-\infty}\right] = \mathbb{E}[Y \mid \mathcal{F}_{-\infty}].$$

□

Exchangeability

Even though the convergence theory for backwards martingales is easy, there are some nice applications, in particular for exchangeable random variables.

Definition. A sequence of random quantities $\{X_n\}_{n=1}^\infty$ is *exchangeable* if, for every n and all distinct $j_1, \dots, j_n$, $(X_{j_1}, \dots, X_{j_n})$ and $(X_1, \dots, X_n)$ have the same joint distribution.

Remark ([2, Example 4.7.8])*.* This definition is equivalent to that, for every n and any permutation π of $\{1, \dots, n\}$, $(X_1, \dots, X_n)$ and $(X_{\pi(1)}, \dots, X_{\pi(n)})$ have the same joint distribution.

Example (Conditionally iid random quantities)**.** Let $\{X_n\}_{n=1}^\infty$ be conditionally iid given a σ–field $\mathcal{F}$. Then $\{X_n\}_{n=1}^\infty$ is an exchangeable sequence. The result follows easily from the fact that

$$\mu_{X_{j_1},\dots,X_{j_n}|\mathcal{F}} = \mu_{X_1,\dots,X_n|\mathcal{F}} \quad \text{a.s..}$$

To analyze an exchangeable sequence of random variables, we return to the special space utilized for Hewitt-Savage 0-1 law in Section 2,5. That is, we suppose

$$\Omega = \{(\omega_1, \omega_2, \dots) : \omega_i \in S\}, \qquad \mathcal{F} = \mathcal{S} \times \mathcal{S} \times \cdots, \qquad X_n(\omega) = \omega_n.$$

Let $\mathcal{E}_n$ be the σ-field generated by events that are invariant under permutations leaving $n+1, n+2, \dots$ fixed, and let $\mathcal{E} = \bigcap_n \mathcal{E}_n$ be the exchangeable σ-field.

We recall Hewitt-Savage 0-1 law, which is a generalization of Kolmogorov's 0-1 law. We provide martingale proof of this at Appendix.

Theorem ([2, Theorem 2.5.4, Example 4.7.6])**.** *Hewitt–Savage 0–1 law. If $X_1, X_2, \dots$ are i.i.d. and $A \in \mathcal{E}$ then $P(A) \in \{0, 1\}$.*

Theorem. *Strong Law of Large Numbers. Let $\{X_n\}_{n=1}^\infty$ be an exchangeable sequence of random variables with $\mathbb{E}|X_i| < \infty$. Let $S_n = \xi_1 + \cdots + \xi_n$. Then $\lim_{n\to\infty} S_n/n$ exists a.s. and has mean equal to $\mathbb{E}[X_1]$. If the X_j's are independent, then the limit equals $\mathbb{E}[X_1]$ a.s.*

Proof. Let $Y_{-n} = S_n/n$, and let $\{\mathcal{F}_{-n}\}_{n\geq 0}$ be defined as

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

To compute $\mathbb{E}[Y_{-n} \mid \mathcal{F}_{-n-1}]$, note that if $j, k \leq n+1$, symmetry (specifically, HW#1 Problem 7) implies $\mathbb{E}[X_j \mid$

$\mathcal{F}_{-n-1}] = \mathbb{E}[X_k \mid \mathcal{F}_{-n-1}]$, so

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_{-n-1}] = \frac{1}{n+1}\sum_{k=1}^{n+1}\mathbb{E}[X_k \mid \mathcal{F}_{-n-1}] = \frac{1}{n+1}\mathbb{E}[S_{n+1} \mid \mathcal{F}_{-n-1}] = \frac{S_{n+1}}{n+1}.$$

Since $Y_{-n} = (S_{n+1} - X_{n+1})/n$, it follows that

$$\mathbb{E}[Y_{-n} \mid \mathcal{F}_{-n-1}] = \mathbb{E}[S_{n+1}/n \mid \mathcal{F}_{-n-1}] - \mathbb{E}[X_{n+1}/n \mid \mathcal{F}_{-n-1}] = \frac{S_{n+1}}{n+1} = Y_{-n-1}.$$

Thus Y_{-n} is a backwards martingale adapted to $\{\mathcal{F}_{-n}\}_{n\geq 0}$, so by [2, Theorem 4.7.1]-[2, Theorem 4.7.2],

$$\lim_{n\to\infty} Y_{-n} = \lim_{n\to\infty} S_n/n = \mathbb{E}[Y_{-1} \mid \mathcal{F}_{-\infty}] = \mathbb{E}[X_1 \mid \mathcal{F}_{-\infty}].$$

In particular,

$$\mathbb{E}\left[\lim_{n\to\infty} Y_{-n}\right] = \mathbb{E}\left[\mathbb{E}[X_1 \mid \mathcal{F}_{-\infty}]\right] = \mathbb{E}[X_1].$$

Suppose further that X_j's are independent. Since $\mathcal{F}_{-n} \subset \mathcal{E}_n$, $\mathcal{F}_{-\infty} \subset \mathcal{E}$. The Hewitt–Savage 0–1 law [2, Theorem 2.5.4, Example 4.7.6] says $\mathcal{E}$ is trivial, so

$$\lim_{n\to\infty} S_n/n = \mathbb{E}[X_{-1}] \qquad \text{a.s.}$$

□

Example. Let $\{X_n\}_{n=1}^{\infty}$ be Bernoulli random variables such that

$$P(X_1 = x_1, \ldots, X_n = x_n) = \frac{1}{(n+1)\binom{n}{y}},$$

where $y = \sum_{j=1}^{n} x_j$. One can show that this specifies consistent joint distributions. One can also check that the X_n's are not independent:

$$P(X_1 = 1) = \tfrac{1}{2}, \qquad P(X_1 = 1, X_2 = 1) = \tfrac{1}{3} \neq \left(\tfrac{1}{2}\right)^2.$$

From Strong law of large number, we know that $Y_{-n} := S_n/n$ converges a.s., hence it converges in distribution. We can compute the exact distribution:

$$P(Y_{-n} = k/n) = \frac{1}{n+1}, \quad k = 0, 1, \ldots, n.$$

Hence Y_{-n} converges in distribution to Uniform$[0,1]$, which must be the distribution of the limit. The limit is not a.s. constant.

De Finetti's theorem says that a sequence of random quantities is exchangeable if and only if it is conditionally iid given exchangeable σ-field. That is, the Conditionally iid example is essentially the only example of exchangeable sequences. We provide the proof at Appendix.

Theorem ([2, Theorem 4.7.9])**.** *de Finetti's Theorem.* $X_1, X_2, \ldots$ *are exchangeable if and only if, conditional on* $\mathcal{E}$, $X_1, X_2, \ldots$ *are independent and identically distributed.*

When the X_i take values in a nice space, there is a regular conditional distribution for $(X_1, X_2, \ldots)$ given $\mathcal{E}$, and the sequence can be represented as a mixture of i.i.d. sequences. [4] call the sequence *presentable* in this case. For general measurable spaces the result may fail; see [1] and [3] for counterexamples.

References

[1] Lester E. Dubins and David A. Freedman. Exchangeable processes need not be mixtures of independent, identically distributed random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 1979.

[2] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition.

[3] David A. Freedman. A mixture of independent identically distributed random variables need not admit a regular conditional probability given the exchangeable σ-field. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 1980.

[4] Edwin Hewitt and Leonard J. Savage. Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.*, 80:470–501, 1955.

Appendix

The key to the martingale proof for Hewitt-Savage 0-1 law and de Finetti's Theorem is:

Lemma ([2, Lemma 4.7.7])**.** *Suppose $X_1, X_2, \dots$ are i.i.d. and let*

$$A_n(\varphi) = \frac{1}{(n)_k} \sum_i \varphi(X_{i_1}, \dots, X_{i_k}),$$

where the sum is over all sequences of distinct integers $1 \le i_1, \dots, i_k \le n$, and

$$(n)_k = n(n-1)\cdots(n-k+1)$$

is the number of such sequences. If φ is bounded, $A_n(\varphi) \to \mathbb{E}\varphi(X_1, \dots, X_k)$ a.s..

Proof. $A_n(\varphi) \in \mathcal{E}_n$, so

$$A_n(\varphi) = \mathbb{E}(A_n(\varphi) \mid \mathcal{E}_n) = \frac{1}{(n)_k} \sum_i \mathbb{E}(\varphi(X_{i_1}, \dots, X_{i_k}) \mid \mathcal{E}_n) = \mathbb{E}(\varphi(X_1, \dots, X_k) \mid \mathcal{E}_n),$$

since all the terms in the sum are the same. [2, Theorem 4.7.3] with $\mathcal{F}_{-m} = \mathcal{E}_m$ for $m \ge 1$ implies that

$$\mathbb{E}(\varphi(X_1, \dots, X_k) \mid \mathcal{E}_n) \to \mathbb{E}(\varphi(X_1, \dots, X_k) \mid \mathcal{E}).$$

We want to show that the limit is $\mathbb{E}(\varphi(X_1, \dots, X_k))$. The first step is to observe that there are $k(n-1)^{k-1}$ terms in $A_n(\varphi)$ involving X_1, and φ is bounded, so if we let $1 \in i$ denote the sum over sequences that contain 1,

$$\frac{1}{(n)_k} \sum_{1 \in i} \varphi(X_{i_1}, \dots, X_{i_k}) \le \frac{k(n-1)^{k-1}}{(n)_k} \sup \varphi \to 0.$$

This shows that $\mathbb{E}(\varphi(X_1, \dots, X_k) \mid \mathcal{E}) \in \sigma(X_2, X_3, \dots)$. Repeating the argument for $2, 3, \dots, k$ shows

$$\mathbb{E}(\varphi(X_1, \dots, X_k) \mid \mathcal{E}) \in \sigma(X_{k+1}, X_{k+2}, \dots).$$

Intuitively, if the conditional expectation of a r.v. is independent of the r.v., then

$$\mathbb{E}(\varphi(X_1, \ldots, X_k) \mid \mathcal{E}) = \mathbb{E}(\varphi(X_1, \ldots, X_k)). \tag{1}$$

To show this, we prove:

$$\text{If } \mathbb{E}X^2 < \infty \text{ and } \mathbb{E}(X \mid \mathcal{G}) \in \mathcal{F} \text{ with } X \text{ independent of } \mathcal{F}, \text{ then } \mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}X. \tag{2}$$

To prove (2), let $Y = \mathbb{E}(X \mid \mathcal{G})$ and note that [2, Theorem 4.1.11] implies $\mathbb{E}Y^2 \leq \mathbb{E}X^2 < \infty$. By independence, $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y = (\mathbb{E}Y)^2$ since $\mathbb{E}Y = \mathbb{E}X$. From the geometric interpretation of conditional expectation [2, Theorem 4.1.15], $\mathbb{E}((X-Y)Y) = 0$, so $\mathbb{E}Y^2 = \mathbb{E}XY = (\mathbb{E}Y)^2$ and hence $\text{var}(Y) = 0$. □

Now we use this Lemma to show Hewitt-Savage 0-1 law.

Proof. Statement (1) holds for all bounded φ, so $\mathcal{E}$ is independent of $\mathcal{G}_k = \sigma(X_1, \ldots, X_k)$. Since this holds for all k, and $\bigcup_k \mathcal{G}_k$ is a π-system containing Ω, [2, Theorem 2.1.6] implies $\mathcal{E}$ is independent of $\sigma(\bigcup_k \mathcal{G}_k) \supset \mathcal{E}$, and we get the usual 0–1 law conclusion: if $A \in \mathcal{E}$, it is independent of itself, and hence

$$P(A) = P(A \cap A) = P(A)P(A),$$

i.e. $P(A) \in \{0, 1\}$. □

We also use this Lemma to show de Finitti's Theorem.

Proof. Repeating the first calculation in [2, Lemma 4.7.7] and using the notation introduced there shows that for any exchangeable sequence:

$$A_n(\varphi) = \mathbb{E}(A_n(\varphi) \mid \mathcal{E}_n) = \frac{1}{(n)_k} \sum_i \mathbb{E}(\varphi(X_{i_1}, \ldots, X_{i_k}) \mid \mathcal{E}_n) = \mathbb{E}(\varphi(X_1, \ldots, X_k) \mid \mathcal{E}_n),$$

since all terms are the same. Again, [2, Theorem 4.7.3] implies that

$$A_n(\varphi) \to \mathbb{E}(\varphi(X_1, \ldots, X_k) \mid \mathcal{E}). \tag{3}$$

This time, however, $\mathcal{E}$ may be nontrivial, so we cannot hope that the limit is $\mathbb{E}(\varphi(X_1, \ldots, X_k))$.

Let f and g be bounded functions on $\mathbb{R}^{k-1}$ and $\mathbb{R}$, respectively. If we let $I_{n,k}$ be the set of all sequences of distinct integers $1 \leq i_1, \ldots, i_k \leq n$, then

$$\begin{aligned}
(n)_{k-1} A_n(f) n A_n(g) &= \sum_{i \in I_{n,k-1}} f(X_{i_1}, \ldots, X_{i_{k-1}}) \sum_m g(X_m) \\
&= \sum_{i \in I_{n,k}} f(X_{i_1}, \ldots, X_{i_{k-1}}) g(X_{i_k}) + \sum_{i \in I_{n,k-1}} \sum_{j=1}^{k-1} f(X_{i_1}, \ldots, X_{i_{k-1}}) g(X_{i_j}).
\end{aligned}$$

If we let $\varphi(x_1, \ldots, x_k) = f(x_1, \ldots, x_{k-1}) g(x_k)$, note that

$$\frac{(n)_{k-1} n}{(n)_k} = \frac{n}{n-k+1}, \qquad \frac{(n)_{k-1}}{(n)_k} = \frac{1}{n-k+1},$$

then rearrange to get

$$A_n(\varphi) = \frac{n}{n-k+1} A_n(f) A_n(g) - \frac{1}{n-k+1} \sum_{j=1}^{k-1} A_n(\varphi_j),$$

where $\varphi_j(x_1, \ldots, x_{k-1}) = f(x_1, \ldots, x_{k-1})g(x_j)$. Applying (3) to φ, f, g, and all φ_j gives

$$\mathbb{E}(f(X_1, \ldots, X_{k-1})g(X_k) \mid \mathcal{E}) = \mathbb{E}(f(X_1, \ldots, X_{k-1}) \mid \mathcal{E})\mathbb{E}(g(X_k) \mid \mathcal{E}).$$

It follows by induction that

$$\mathbb{E}\left(\prod_{j=1}^{k} f_j(X_j) \middle| \mathcal{E}\right) = \prod_{j=1}^{k} \mathbb{E}(f_j(X_j) \mid \mathcal{E}).$$

$\square$