Lecture 25 : Reversed Martingales

STAT205 *Lecturer: Jim Pitman* Scribe: *Daisy Huang, Jing Lei* <yanhuang@stat,jinglei@stat>

References: [1], section 4.6.

25.1 Exchangeable σ-field

Let $X_1, X_2, ...$ be a sequence of real-valued random variables. Define $\mathcal{E}_n = \sigma(f(X_1, X_2, ...))$ where $f(X_1, X_2, ...)$ is symmetric with respect to the first n variables; i.e., $f(X_1, X_2, ...) = f(X_{\pi(1)}, X_{\pi(2)}, ...)$ where π is a permutation such that $\pi : \{1, 2, ...n\} \to \{1, 2, ...n\}$ and $\pi(k) = k$ for $k > n$ and f is product Borel measurable. $\mathcal{E}_\infty = \bigcap_n \mathcal{E}_n$ is called the *exchangeable σ-field*.

Remark:

1. X is $\mathcal{E}_n$-measurable iff $X = f(X_1, X_2, ...)$ for some such f.
2. $\mathcal{E}_n \supseteq \mathcal{E}_{n+1}$ because $\mathcal{E}_{n+1}$ requires more symmetries. Also, $\mathcal{E}_n \downarrow \mathcal{E}_\infty$ since

$$\begin{aligned} \mathcal{E}_\infty &\supseteq \text{tail } \sigma\text{-field of } \{X_1, X_2, ...\} && (25.1) \\ &= \bigcap_n \sigma(X_n, X_{n+1}, ...) && (25.2) \end{aligned}$$

   (noting that $\sigma(X_{n+1}, X_{n+2}, ...) \subseteq \mathcal{E}_n$ and shifting the index n).

The "basic" symmetric functions of $X_1, X_2, ...X_n$ are the order statistics

$$\min_{1 \le i \le n} X_i = X_{n,1} \le X_{n,2} \le ... \le X_{n,n} = \max_{1 \le i \le n} X_i.$$

Obviously, each $X_{n,k}$ is a symmetric function of $X_1, X_2, ...X_n$. You can check that $\mathcal{E}_n = \sigma(X_{n,1}, X_{n,2}, ...X_{n,n}, X_{n+1}, X_{n+2}, ...)$.

Example 25.1 *$S_n = X_1 + X_2 + ... + X_n$ is in $\mathcal{E}_n$ but not in the tail σ-field of $\{X_1, X_2, ...\}$.*

25.2 Hewitt-Savage $0-1$ Law

Theorem 25.2 (Hewitt-Savage $0-1$ Law) *If $X_1, X_2, ...$ is i.i.d., then every event in $\mathcal{E}_\infty$ has probability 0 or 1.*

(Compare this with Kolmogorov's $0-1$ Law.)

Recall that if $X_1, X_2, ...$ are i.i.d. and exchangeable (i.e.,

$$(X_1, X_2, ...X_n) \stackrel{d}{=} (X_{\pi(1)}, X_{\pi(2)}, ..., X_{\pi(n)})$$

for all permutations π on n elements) and $\mathbb{E}|X_1| < \infty$, then $(S_n/n, \mathcal{E}_n)_{n\geq 1}$ is a reversed martingale. If $S_n/n = \mathbb{E}(X_1|\mathcal{E}_n)$, then this is obvious because $\mathcal{E}_n \downarrow$.

To see this: exchangeability implies that

$$\mathbb{E}(X_1|\mathcal{E}_n) = \mathbb{E}(X_k|\mathcal{E}_n) \qquad \text{for every } 1 \leq k \leq n.$$

This is because

$$\begin{aligned}
& (X_1, f(X_1, \cdots, X_n, X_{n+1}, X_{n+2}, \cdots)) \\
\stackrel{d}{=} \; & (X_k, f(X_k, \cdots, X_{k-1}, X_1, X_{k+1}, \cdots, X_n, X_{n+1}, X_{n+2}, \cdots)) \\
= \; & (X_k, f(X_1, X_2, ..., X_n, X_{n+1}, X_{n+2} ...))
\end{aligned}$$

Now check the definition of the claim:

$$n\mathbb{E}(X_1|\mathcal{E}_n) = \mathbb{E}(S_n|\mathcal{E}_n) = S_n \text{ since } S_n \subseteq \mathcal{E}_n.$$

Now, we prove the theorem. The plan is to show that for every event $F \in \sigma(X_1, ..., X_n)$, $\mathbb{P}(F|\mathcal{E}_\infty) = \mathbb{P}(F)$. This says that $\sigma(X_1, ..., X_n)$ is independent of $\mathcal{E}_\infty$. To see this, let $n \to \infty$ and learn that $\sigma(X_1, X_2, ...)$ is independent of $\mathcal{E}_\infty$. But this implies that $\mathcal{E}_n$ is independent of $\mathcal{E}_\infty$, which leads to the result of the $0-1$ Law.

Proof: By the Martingale Convergence Theorem,

$$\mathbb{P}(F|\mathcal{E}_\infty) = \lim_{n\to\infty} \mathbb{P}(F|\mathcal{E}_n)$$

Let $F = \{X_1 \in \hat{F}\}$ for some $\hat{F} \subseteq \mathbb{R}$. Then,

$$\begin{aligned}
\mathbb{P}(F \mid \mathcal{E}_n) &= \mathbb{E}(\mathbf{1}_{\hat{F}(X_1)}|\mathcal{E}_n) \\
&= \mathbb{E}(\mathbf{1}_{\hat{F}(X_k)}|\mathcal{E}_n) \\
&= \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{\hat{F}(X_k)}|\mathcal{E}_n\right) \\
&= \frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{\hat{F}(X_k)} \\
&\to \mathbb{E}(\mathbf{1}_{\hat{F}(X_k)}) \\
&= \mathbb{P}(X_1 \in \hat{F}) \\
&= \mathbb{P}(F).
\end{aligned}$$

This proves for the case $F \in \sigma(X_1)$. To deal with the case $F \in \sigma(X_1, X_2)$, we do the same thing. Let $F = \{(X_1, X_2) \in \hat{F}\}$ for some $\hat{F} \subseteq \mathbb{R}^2$ and $\varphi(X_1, X_2) = \mathbf{1}((X_1, X_2) \in \hat{F})$. Then,

$$\begin{aligned}
\mathbb{P}(F \mid \mathcal{E}_n) &= \mathbb{E}(\varphi(X_1, X_2)|\mathcal{E}_n) \\
&= \mathbb{E}(\varphi(X_i, X_j)|\mathcal{E}_n) \text{ for } i \neq j \\
&= \mathbb{E}\left(\frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n}\varphi(X_i, X_j)|\mathcal{E}_n\right) \\
&= \frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n}\varphi(X_i, X_j).
\end{aligned}$$

Consider $\varphi(X_i, X_j) = f(X_i)g(X_j)$, i.e. $\hat{F}$ is rectangular; then

$$\begin{aligned}
\mathbb{P}(F \mid \mathcal{E}_n) &= \frac{1}{n(n-1)}\sum_{1\leq i\neq j\leq n} f(X_i)g(X_j) \\
&= \frac{1}{n(n-1)}\left(\sum_{1\leq i,j\leq n} f(X_i)g(X_j) - \sum_{i=1}^{n} f(X_i)g(X_i)\right) \\
&= \frac{1}{n(n-1)}\left(\sum_{i=1}^{n} f(X_i)\sum_{i=1}^{n} g(X_i) - \sum_{i=1}^{n} f(X_i)g(X_i)\right) \\
&\xrightarrow{a.s.} \mathbb{E}f(X_1)\mathbb{E}g(X_2) \\
&= \mathbb{E}\varphi(X_1, X_2) \\
&= \mathbb{P}(F).
\end{aligned}$$

To finish, for $\hat{F} \in \mathcal{B}(\mathbb{R}^2)$ we just use the $\pi - \lambda$ theorem. Similarly for $\hat{F} \in \mathcal{B}(\mathbb{R}^k)$, $k \geq 3$.

So far, we've shown that

$$\mathbb{P}(F \mid \mathcal{E}_\infty) = \mathbb{P}(F) \text{ for all } F \in \sigma(X_1, \cdots, X_n),$$

i.e. $\mathcal{E}_\infty$ is independent of $\sigma(X_1, \cdots, X_n)$. Similarly as in the proof of Kolmogorov's 0-1 Law, we can learn that $\mathcal{E}_\infty$ is independent of $\sigma(X_1, X_2, \cdots)$ by sending $n \to \infty$. Thus $\mathcal{E}_\infty$ is independent of itself, which completes the proof. ■

25.3 de Finetti's Theorem

Theorem 25.3 *If $X_1, X_2, \cdots,$ are exchangeable, and $F_n(x) := \frac{1}{n}\sum_{1 \le i \le n} \mathbf{1}(X_i \le x)$, then*

$$\lim_{n\to\infty} \sup_x |F_n(x) - F(x)| = 0$$

for some random CDF $(F(x), x \in \mathbb{R})$.

Also, given $\mathcal{E}_\infty$, the $X_1, X_2, \cdots,$ are i.i.d. with common distribution F; i.e.

$$\mathbb{P}(X_1 \le x_1, X_2 \le x_2, \cdots, X_k \le x_k | \mathcal{E}_\infty) = F(x_1)F(x_2)\cdot \cdots \cdot F(x_k). \tag{25.3}$$

Note: Conceptually, it's as if F were first picked at random in some way from the set of probability CDF's on $\mathbb{R}$, and then we sample from F.

Proof Sketch: (This sketch proof is incomplete – see the text book for details.)

1. Look at (25.3) for $k = 1$, and by MGCT and exchangeability we have

$$\begin{aligned} F(x) &= \mathbb{P}(X_1 \le x|\mathcal{E}_\infty) \\ &= \lim_{n\to\infty} \mathbb{P}(X_1 \le x|\mathcal{E}_n) \text{ a.s.} \\ &= \lim_{n\to\infty} F_n(x) \text{ a.s.} \end{aligned}$$

2. Using the fact that $F_n(x)$ is non-decreasing in x for fixed n, we learn that $F(x)$ is non-decreasing in x almost surely.

3. Clean up over rationals: let

$$F^\star(x) := \lim_{q \downarrow x} F(q),$$

where $q \in \mathbb{Q}$, the set of all rational numbers. Then argue that $F^\star(x)$ is a CDF and replace $F(x)$ by $F^\star(x)$.

■

References

[1] Richard Durrett. *Probability: theory and examples, 3rd edition.* Thomson Brooks/Cole, 2005.