Lecture 18: Martingales Convergence Theorems

Parimal Parag

1 Submartingales, Supermartingales and the Martingale Convergence Theorem

Definition 1.1. A stochastic process $\{Z_n,\ n \geq 1\}$ having $E[|Z_n|] < \infty$ for all n is said to be a submartingale if

$$E[Z_{n+1}|Z_1 \dots Z_n] \geq Z_n \tag{1}$$

and is said to be a supermartingale if

$$E[Z_{n+1}|Z_1 \dots Z_n] \leq Z_n \tag{2}$$

From 1, for a submartingale

$$E[Z_{n+1}] \geq E[Z_n]$$

where the inequality is reversed for a supermartingale.

Theorem 1.2. *If N is a stopping time for $\{Z_n,\ n \geq 1\}$ such that any one of the following sufficient conditions is satisfied:*

1. *$\bar{Z}_n$ is uniformly bounded, or;*
2. *N is bounded, or;*
3. *$E[N] < \infty$, and there is an $M < \infty$ such that*

$$E[|Z_{n+1} - Z_n||Z_1, \dots Z_n] < M,$$

then,

$$E[Z_N] \geq E[Z_1] \text{ for a submartingale}$$
$$E[Z_N] \leq E[Z_1] \text{ for a supermartingale.}$$

Proof. We claim that

$$\bar{Z}_n = \bar{Z}_{n-1} + 1_{N \geq n}(Z_n - Z_{n-1})$$

The above equation can be directly verified by considering the two cases separately viz.

1. $N \geq n$: $\bar{Z}_n = Z_n$.
2. $N < n : \bar{Z}_{n-1} = \bar{Z}_n = Z_N$

$$\begin{aligned} E[\bar{Z}_{n+1}|Z_1 \dots \bar{Z}_n] &= E[\bar{Z}_n + 1_{n\le N}(Z_n - Z_{n-1})|Z_1 \dots \bar{Z}_n] \\ &\overset{(a)}{=} \bar{Z}_n + 1_{n\le N}E[(Z_n - Z_{n-1})|Z_1 \dots \bar{Z}_n] \\ &\ge \bar{Z}_n, \end{aligned}$$

where in (a) we have used the fact that N is a random time. Also, we have $E[\bar{Z}_n] = E[Z_1]$, for all n. Now assume that N is a stopping time. It is immediate that

$$\bar{Z}_n \to Z_N \text{ w.p 1.}$$

But is it true that

$$E[\bar{Z}_n] \to E[Z_N] \text{ as n } \to \infty.$$

which gives that

$$E[Z_N] \ge E[Z_1].$$

□

Before we state and prove martingale convergence theorem, we state some results which will be used in the proof of the theorem.

Lemma 1.3. *If $Z_i,\ i \ge 1$ is a submartingale and N is a stopping time such that $P(N \le n) = 1$ then*

$$E[Z_1] \le E[Z_N] \le E[Z_n].$$

Proof. It follows from Theorem 1.2 that since N is bounded, $E[Z_N] \ge E[Z_1]$. Now,

$$\begin{aligned} E[Z_n|Z_1, \dots, Z_N, N = k] &= E[Z_n|Z_1 \dots Z_k, N = k] \\ &\overset{(a)}{=} E[Z_n|Z_1 \dots Z_k] \\ &= Z_k \\ &= Z_N. \end{aligned}$$

where (a) follows from the fact that N is a stopping time. Result follows by taking expectation on both sides. □

Lemma 1.4. *If $\{Z_n, n \ge 1\}$ is a martingale and f is a convex function, then $\{f(Z_n), n \ge 1\}$ is a submartigale.*

Proof. The result is a direct consequence of Jensen's inequality.

$$E[f(Z_n)|Z_1, \dots Z_n] \quad \ge f(E[Z_{n+1}|Z_1, \dots Z_n]) \quad = f(Z_n).$$

□

Theorem 1.5. ***(Kolmogorov's Inequality for Submartingales)****If $\{Z_n,\ n \ge 1\}$ is a martingale, then*

$$Pr(\max\{Z_1, Z_2 \dots Z_n\} > a) \le \frac{E[Z_n]}{a},\ \textit{for } a > 0.$$

Proof. Let N be the smallest value of i, $i \leq n$, such that $Z_i > a$, and define it to equal n if $Z_i \leq a$ for all $i = 1, \ldots n$. Note that $\max\{Z_1 \ldots Z_n\} > a$ is equivalent to $Z_N > a$. Therefore,

$$\begin{aligned} Pr(\max\{Z_1 \ldots Z_n\} > a) &= Pr(Z_N > a) \\ &\overset{(*)}{\leq} \tfrac{E[Z_N]}{a} \\ &\leq \tfrac{E[Z_n]}{a}, \end{aligned}$$

where the last inequality follows from Lemma 1.3 as $N \leq n$ and $(*)$ follows from Markov's inequality. □

Corollary 1.6. *Let $\{Z_n,\ n \geq 1\}$ be a martingale. Then, for $a > 0$:*

1. $Pr(\max\{|Z_1|, \ldots |Z_n| > a\}) \leq E[|Z_n|]/a$;
2. $Pr(\max\{|Z_1|, \ldots |Z_n| > a\}) \leq E[Z_n^2]/a^2$.

Proof. The proof the above statements follow from Lemma 1.4 and Kolmogorov's inequality for submartingales by considering the convex functions $f(x) = |x|$ and $f(x) = x^2$. □

Theorem 1.7. *If $\{Z_n,\ n \geq 1\}$ is a martingale such that for some $M < \infty$*

$$E[|Z_n|] \leq M, \text{ for all } n$$

then, with probability 1, $\lim_{n\to\infty} Z_n$ exists and is finite.

Proof. Assume $E[Z_n^2] < \infty$ which is stronger than $E[|Z_n|] < \infty$ (as a consequence of Jensen's inequality). Observe that $\{Z_n^2\}$ is a submartingale (from Lemma 1.4). Thus $E[Z_n^2] < \infty$ and is non-decreasing in n. Thus, as $n \to \infty$, $E[Z_n^2]$ converges and let $\mu < \infty$ be given by $\mu = \lim_{n\to\infty} E[Z_n^2]$.

$$Pr(\cup_{k\leq n}\{|Z_{m+k} - Z_m| > \epsilon\}) \tag{3}$$

$$\overset{(a)}{\leq} E[(Z_{m+n} - Z_m)^2]/\epsilon^2 \quad = E[Z_{m+n}^2 - 2Z_mZ_{m+n} + Z_m^2]/\epsilon^2.$$

Note that

$$\begin{aligned} E[Z_{m+n}Z_m] &= E[E[Z_mZ_{m+n}|Z_m]] \\ &= E[Z_mE[Z_{m+n}|Z_m]] \\ &= E[Z_m^2]. \end{aligned}$$

From 3,

$$Pr(\cup_{k\leq n}\{|Z_{m+k} - Z_m| > \epsilon\}) \leq \frac{E[Z_{m+n}^2] - E[Z_m^2]}{\epsilon^2}.$$

Letting $n \to \infty$

$$Pr(\cup_{k\leq 1}\{|Z_{m+k} - Z_m| > \epsilon\}) \leq \frac{\mu - E[Z_m^2]}{\epsilon^2}.$$

Hence,

$$Pr(\cup_{k\leq n}\{|Z_{m+k} - Z_m| > \epsilon\}) \to 0 \text{ as } m \to \infty.$$

Thus with probability 1, $\{Z_n\}$ will be a Cauchy sequence, and thus $\lim_{n\to\infty} Z_n$ will exist and be finite.' □

Corollary 1.8. *If $\{Z_n,\ m \geq 0\}$ is a non-negative martingale, then, with probability 1, $\lim_{n\to\infty} Z_n$ exists and is finite.*

Proof. Since Z_n is non-negative,

$$E[|Z_n|] = E[Z_n] = E[Z_1].$$

$\square$

Theorem 1.9. ***The Strong Law of Large Numbers****Let $X_1, X_2 \ldots$ be a sequence of independent and identically distributed random variables having finite mean μ, and let $S_n = \sum_{i=1}^{n} X_i$. Then*

$$Pr(\lim_{n\to\infty} S_n/n = \mu) = 1.$$

Proof. We will prove the theorem under the assumption that the moment generating function exists. Let $\psi(t) = E[e^{tX}]$. For a given $\epsilon > 0$, let $g(t)$ be defined by

$$g(t) = e^{t(\mu+\epsilon)}/\psi(t).$$

$$g(0) = 1,$$

$$g'(0) = \frac{\psi(0)(\mu+\epsilon)e^{t(\mu+\epsilon)} - \psi'(0)}{e}^{t(\mu+\epsilon)} \psi^2(t) = \epsilon > 0,$$

there exists a value $t_0 > 0$ such that $g(t_0) > 1$. We now show that S_n/n can be as large as $\mu + \epsilon$ only finitely often. For, note that

$$\frac{S_n}{n} \geq \mu + \epsilon \Rightarrow (\frac{e^{t_0 S_n}}{\psi^n(t_0)})^n = (g(t_0))^n \tag{4}$$

But $\frac{e^{t_0 S_n}}{\psi^n(t_0)}$ is the product of independent, non negative random variables with unit mean and hence is a martingale. By Theorem 1.7

$$\lim_{n\to\infty} \frac{e^{t_0 S_n}}{\psi^n(t_0)} \text{ exists and is finite.}$$

Since $g(t_0) > 1$, it follows from 4 that

$$Pr(S_n/n > \mu + \epsilon \text{ for an infinite number of n}) = 0.$$

Similarly, be defining the function $f(t) = e^{t(\mu-\epsilon)}/\psi(t)$ and noting that since $f(0) = 1$, $f'(0) = -\epsilon$, there exists a value $t_0 < 0$ such that $f(t_0) > 1$, we can prove in the same manner that

$$Pr(S_n/n \leq \mu - \epsilon \text{ for an infinite number of n}) = 0.$$

Hence

$$Pr(\mu - \epsilon S_n/n < \mu + epsilon \text{ for all but a finite number of } n) = 1,$$

or, since the above is true for all $\epsilon > 0$,

$$Pr(\lim_{n\to\infty} S_n/n = \mu) = 0.$$

$\square$

Definition 1.10. The sequence of random variables $X_n,\ n \geq 1$, is said to be *uniformly integrable* if for every $\epsilon > 0$, there is a y_ϵ such that

$$\int_{|x|>y_\epsilon} |x| dF_n(x) < \epsilon \ \forall n$$

where F_n is the distribution function of X_n.

Lemma 1.11. *If $X_n,\ n \geq 1$, is uniformly integrable then there exists $M < \infty$ such that $E[|X_n|] < M$ for all n.*

Proof. Let y_1 be as in the definition of uniform integrability. Then

$$\begin{aligned} E[|X_n|] &= \textstyle\int_{|x|\leq y_1} |x| dF_n(x) + \int_{|x|>y_1} |x| dF_n(x) \\ &\leq y_1 + 1 \end{aligned}$$

□

1.1 Generalized Azuma Inequality

Proposition 1.12. *Let $\{Z_n, n \geq 1\}$ be a martingale with mean $Z_0 = 0$, for which*

$$-\alpha \leq Z_n - Z_{n-1} \leq \beta \ \forall \ n \geq 1$$

Then, for any positive values a and b

$$Pr(Z_n \geq a + bn \ for \ some \ n) \leq exp(-8ab/(\alpha+\beta)^2).$$

Proof. Let, for $n \geq 0$

$$W_n = exp\{c(Z_n - a - bn)\}$$

Observe that

$$W_n = W_{n-1}e^{-cb}exp\{c(Z_n - Z_{n-1})\}.$$

Using the fact that knowledge of $W_1, W_2, \ldots W_{n-1}$ is equivalent to that of $Z_1, Z_2 \ldots Z_{n-1}$, we obtain that

$$\begin{aligned} E[W_n | W_1 \ldots W_{n-1}] &= W_{n-1}e^{-cb}E[expc(Z_n - Z_{n-1})|Z_1 \ldots Z_{n-1}] \\ &\stackrel{(a)}{\leq} W_{n-1}e^{-cb}[\beta e^{-c\alpha} + \alpha e^{c\beta}]/(\alpha+\beta) \\ &\stackrel{(b)}{\leq} W_{n-1}e^{-cb}e^{c^2(\alpha+\beta)^2/8} \end{aligned}$$

where (a) follows from Lemma 1.3 (Lecture 17) and (b) from Lemma 1.4 (Lecture 17) with $\theta = \alpha/(\alpha+\beta),\ x = c(\alpha+\beta)$. Hence, fixing the value of c as $c = 8b/(\alpha+\beta)^2$ yields

$$E[W_n | W_1, \ldots W_{n-1}] \leq W_{n-1}, \tag{5}$$

and so $\{W_n,\ ngeq0\}$ is a supermartingale. For a fixed positive integer k, define the bounded stopping time N by

$$N = \min\{n : \text{ either } Z_n \geq a + bn \text{ or } n = k\}.$$

Now,

$$\begin{aligned} Pr(Z_N \geq a + bn) &= P(W_N \geq 1) \\ &\overset{(a)}{\leq} E[W_N] \\ &\overset{(b)}{\leq} E[W_0] \end{aligned}$$

where (a) follows from Markov inequality and (b) follows from supermartingale stopping theorem. But the above inequality is equivalent to

$$Pr(Z_n \geq a + bn \text{ for some } n \leq k) \leq e^{-8ab/(\alpha+\beta)^2}.$$

Letting $k \to \infty$ gives the result. □

Theorem 1.13. ***The generalized Azuma Inequality*** *Let $\{Z_n\ n \geq 1\}$ be a martingale with mean $Z_0 = 0$. If $-\alpha \leq Z_n - Z_{n-1} \leq \beta$ for all $n \geq 1$ then, for any positive constant c and integer m:*

1. *$Pr(Z_n \geq nc$ for some $n \geq m) \leq e^{-2mc^2/(\alpha+\beta)^2}$.*
2. *$Pr(Z_n \leq -nc$ for some $\geq m) \leq e^{-2mc^2/(\alpha+\beta)^2}$.*

Proof. To begin, note that if there is an n such that $n \geq m$ and $Z_n \geq nc$ then, for that n, $Z_n \geq nc \geq mc/2 + nc/2$. Hence,

$$\begin{aligned} Pr(Z_n \geq nc \text{ for some } \geq m) &\leq Pr(Z_n \geq mc/2 + (c/2)n \text{ for some } n) \\ &\leq exp\{-8(mc/2)(c/2)/(\alpha+\beta)^2\} \end{aligned}$$

where the last inequality follows from Proposition 1.12. This proves part (i). Part (ii) follows from part (i) by considering the martingale $\{-Z_n,\ n \geq 0\}$. □