Lecture-27: Martingale: convergence and concentration

1 Martingale convergence theorem

Before we state and prove martingale convergence theorem, we state some results which will be used in the proof of the theorem. Consider a discrete time filtration $\mathcal{F}_\bullet \triangleq (\mathcal{F}_n : n \in \mathbb{N})$.

Lemma 1.1. *Consider a submartingale $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ and a stopping time $\tau : \Omega \to \mathbb{N}$ both adapted to a filtration $\mathcal{F}_\bullet$. If there exists some $N \in \mathbb{N}$ such that $P\{\tau \leqslant N\} = 1$, then $\mathbb{E}X_1 \leqslant \mathbb{E}X_\tau \leqslant \mathbb{E}X_N$.*

Proof. Recall that for any random time τ, the stopped process X^τ is submartingale. Hence $\mathbb{E}X_\tau \geqslant \mathbb{E}X_1$. Since τ is a stopping time, we see that for the event $\{\tau = k\}$ for any $k \leqslant N$

$$\mathbb{E}[X_N \mathbb{1}_{\{\tau=k\}} | \mathcal{F}_k] \geqslant X_k \mathbb{1}_{\{\tau=k\}} = X_\tau \mathbb{1}_{\{\tau=k\}}.$$

Since $\tau \leqslant N$ almost surely, we have the following almost sure inequality $X_N = \sum_{k=1}^N X_\tau \mathbb{1}_{\{\tau=k\}}$. Result follows by taking expectation on both sides, using linearity of expectation, and applying tower property of conditional expectation. That is, $\mathbb{E}X_N = \mathbb{E}\sum_{k=1}^N X_N \mathbb{1}_{\{\tau=k\}} \geqslant \mathbb{E}\sum_{k=1}^N X_\tau \mathbb{1}_{\{\tau=k\}} = \mathbb{E}X_\tau$. □

Definition 1.2. Consider a discrete random process $X : \Omega \to \mathbb{R}^{\mathbb{Z}^+}$ adapted to the filtration $\mathcal{F}_\bullet = (\mathcal{F}_n \subseteq \mathcal{F} : n \in \mathbb{Z}^+)$. Let $N_0 \triangleq 0$. For the two thresholds $a < b$, we define the stopping times corresponding to kth downcrossing and upcrossing times as

$$N_{2k-1} \triangleq \inf\{m > N_{2k-2} : X_m \leqslant a\}, \qquad N_{2k} \triangleq \inf\{m > N_{2k-1} : X_m \geqslant b\}.$$

We next define the indicator to the event that the process is in kth upcrossing transition from a to b at time m, as $H_m^k \triangleq \mathbb{1}_{\{N_{2k-1} < m \leqslant N_{2k}\}}$, and the indicator to the event that the process is in some upcrossing transition from a to b at time m, as

$$H_m \triangleq \sum_{k \in \mathbb{N}} H_m^k = \sum_{k \in \mathbb{N}} \mathbb{1}_{\{N_{2k-1} < m \leqslant N_{2k}\}}.$$

The number of upcrossings completed in time n is defined by

$$U_n \triangleq \sup\{k \in \mathbb{N} : N_{2k} \leqslant n\} = \sum_{k \in \mathbb{N}} \mathbb{1}_{\{N_{2k} \leqslant n\}}.$$

Remark 1. For each $k \in \mathbb{N}$, the kth downcrossing time N_{2k-1} and the kth upcrossing time N_{2k}, are integer stopping times. Hence, we have $H_m^k = \{N_{2k-1} < m \leqslant N_{2k}\} = \{N_{2k-1} \leqslant m-1\} \cap \{N_{2k} \leqslant m-1\}^c \in \mathcal{F}_{m-1}$. It follows that $\sigma(H_m^k) \subseteq \mathcal{F}_{m-1}$ for each $k \in \mathbb{N}$. Further, since $H_m = \sum_{k \in \mathbb{N}} H_m^k$ and addition is a Borel measurable function, it follows that $\sigma(H_m) \subseteq \mathcal{F}_{m-1}$. Hence, the event that the process X is in an upcrossing transition at time m is predictable. Further, we notice that the indicator $H_m \in \{0,1\}$ is nonnegative and bounded. Since $N_0 = 0$, it follows that the first downcrossing time $N_1 \geqslant 1$ and $H_1 = 0$.

Lemma 1.3 (Upcrossing inequality). *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a submartingale adapted to a filtration $\mathcal{F}_\bullet$. Then, we have $(b-a)\mathbb{E}U_n \leqslant \mathbb{E}(X_n - a)^+$.*

Proof. Define a random sequence $Y : \Omega \to \mathbb{R}^{\mathbb{N}}$ for each $n \in \mathbb{N}$, as $Y_n \triangleq a + (X_n - a)^+ = X_n \vee a$. Since X is a submartingale adapted to $\mathcal{F}_\bullet$ and the map $x \mapsto f(x) = x \vee a$ is convex and nondecreasing, it follows that Y is also a submartingale adapted to $\mathcal{F}_\bullet$. By definition of $(H \cdot Y)_n$, exchanging order of summation, the fact that $\mathbb{1}_{\{m \leqslant n\}} H_m^k = \mathbb{1}_{\{N_{2k-1} < m \leqslant N_{2k} \wedge n\}} \geqslant H_m^k \mathbb{1}_{\{N_{2k} \leqslant n\}}$, and observing that each upcrossing has a gain lower bounded by $b - a$, we obtain

$$(H \cdot Y)_n = \sum_{m \in \mathbb{N}} \mathbb{1}_{\{m \leqslant n\}} \sum_{k \in \mathbb{N}} H_m^k (Y_m - Y_{m-1}) \geqslant \sum_{k \in \mathbb{N}} \mathbb{1}_{\{N_{2k} \leqslant n\}} \sum_{m=N_{2k-1}+1}^{N_{2k}} (Y_m - Y_{m-1}) = \sum_{k=1}^{U_n} (Y_{N_{2k}} - Y_{N_{2k-1}}) \geqslant (b-a)U_n.$$

Let $K_m \triangleq 1 - H_m$ for each $m \in \mathbb{N}$. From the telescopic sum, the definition of H, K, and $(H \cdot K)$, we obtain

$$Y_n - Y_0 = \sum_{i=1}^{n} (Y_i - Y_{i-1}) = \sum_{i=1}^{n} (H_i + K_i)(Y_i - Y_{i-1}) = (H \cdot Y)_n + (K \cdot Y)_n.$$

Since H is predictable, then so is K with respect to $\mathcal{F}_\bullet$. Further, both $H : \Omega \to \{0,1\}^{\mathbb{N}}$ and $K : \Omega \to \{0,1\}^{\mathbb{N}}$ are nonnegative and bounded sequences. It follows that since Y is a submartingale, so are sequences $((H \cdot Y)_n : n \in \mathbb{Z}^+)$ and $((K \cdot Y)_n : n \in \mathbb{Z}^+)$. Therefore, we can write

$$\mathbb{E}[(K \cdot Y)_n] \geqslant \mathbb{E}[(K \cdot Y)_1] = \mathbb{E}[K_1(Y_1 - Y_0)] = \mathbb{E}[Y_1 - Y_0] \geqslant -\mathbb{E}(X_0 - a)^+.$$

Therefore, it follows that

$$\mathbb{E}(Y_n - Y_0) = \mathbb{E}(H \cdot Y)_n + \mathbb{E}(K \cdot Y)_n \geqslant \mathbb{E}(H \cdot Y)_n - \mathbb{E}(X_0 - a)^+ \geqslant (b - a)\mathbb{E}U_n - \mathbb{E}(X_0 - a)^+.$$

The result follows from the fact that $\mathbb{E}Y_n - \mathbb{E}Y_0 = \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_0 - a)^+$. $\square$

Theorem 1.4 (Martingale convergence theorem). *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a submartingale adapted to filtration $\mathcal{F}_\bullet$. If $\sup_{n \in \mathbb{N}} \mathbb{E}X_n^+ < \infty$ then (a) $\lim_{n \in \mathbb{N}} X_n = X_\infty$ a.s. and (b) $\mathbb{E}|X_\infty| < \infty$. That is, X converges almost surely and in mean.*

Proof. We will show this in two parts.

(a) To show that $\lim_{n \in \mathbb{N}} X_n$ exists almost surely, we show that $\limsup_{n \in \mathbb{N}} X_n = \liminf_{n \in \mathbb{N}} X_n$ almost surely. From the density of rationals in reals, it suffices to show that

$$P\left(\bigcup_{a,b \in \mathbb{Q}} \{\liminf X_n < a < b < \limsup X_n\} \right) = 0.$$

To this end, it suffices to show that $P\{X_n \leqslant a \text{ and } X_n \geqslant b \text{ infinitely often}\} = 0$ for any $a, b \in \mathbb{Q}$. We fix $a, b \in \mathbb{Q}$, and observe that U_n is the number of upcrossings until time n, then it suffices to show that $\limsup_{n \in \mathbb{N}} U_n$ is finite almost surely. To this end, we observe that $(X - a)^+ \leqslant X^+ + |a|$, and hence it follows from the upcrossing inequality and the theorem hypothesis that

$$\sup_{n \in \mathbb{N}} \mathbb{E}U_n \leqslant \sup_{n \in \mathbb{N}} \frac{\mathbb{E}X_n^+ + |a|}{b - a} < \infty.$$

That is, even though the number of upcrossings U_n increases with n, the mean $\mathbb{E}U_n$ is uniformly bounded above for each $n \in \mathbb{N}$. Hence, $\lim_{n \in \mathbb{N}} \mathbb{E}U_n$ exists and is finite. From monotone convergence theorem, we have $\mathbb{E} \lim_{n \in \mathbb{N}} U_n = \lim_{n \in \mathbb{N}} \mathbb{E}U_n$. We define $U \triangleq \lim_{n \in \mathbb{N}} U_n$ and since $\mathbb{E}U \leqslant \sup_n \mathbb{E}U_n < \infty$, it follows that $U < \infty$ almost surely.

(b) Since $X_\infty \triangleq \lim_{n \in \mathbb{N}} X_n$ exists almost surely, to show the convergence in mean, it suffices to show that $\lim_{n \in \mathbb{N}} \mathbb{E}X_n = \mathbb{E} \lim_{n \in \mathbb{N}} X_n$. It suffices to show that $\mathbb{E}|X_\infty| < \infty$, then the result would follow from the application of dominated convergence theorem to exchange the limit and expectation. To this end, we observe $\mathbb{E}X_\infty^+ \leqslant \liminf_{n \in \mathbb{N}} \mathbb{E}X_n^+ < \infty$ from Fatou's Lemma, that $X_\infty < \infty$ almost surely. Further, we have $\mathbb{E}X_n^- = \mathbb{E}X_n^+ - \mathbb{E}X_n \leqslant \mathbb{E}X_n^+ - \mathbb{E}X_0$ from the submartingale property of X. From this propety and application of Fatou's lemma to the sequence $(X_n^- : n \in \mathbb{N})$, we get

$$\mathbb{E}X_\infty^- \leqslant \liminf_{n \in \mathbb{N}} \mathbb{E}X_n^- \leqslant \sup_{n \in \mathbb{N}} \mathbb{E}X_n^+ - \mathbb{E}X_0 < \infty.$$

This implies $X_\infty > -\infty$ almost surely, completing the proof.

$\square$

2 Martingale concentration inequalities

Consider a discrete time filtration $\mathcal{F}_\bullet \triangleq (\mathcal{F}_n \subseteq \mathcal{F} : n \in \mathbb{N})$ defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a random sequence and $\tau : \Omega \to \mathbb{N}$ a stopping time, both adapted to the filtration $\mathcal{F}_\bullet$.

Remark 2. Recall that for a submartingale X and a stopping time τ bounded above by n, both adapted to the same filtration $\mathcal{F}_\bullet$, we have $\mathbb{E}X_1 \leqslant \mathbb{E}X_\tau \leqslant \mathbb{E}X_n$.

Theorem 2.1 (Kolmogorov's inequality for submartingales). *For a non-negative submartingale $X : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ and $a > 0$, we have $P\left\{\max_{i \in [n]} X_i > a\right\} \leqslant \frac{\mathbb{E}[X_n]}{a}$.*

Proof. We define a random time $\tau_a \triangleq \inf\{i \in \mathbb{N} : X_i > a\}$ and stopping time $\tau \triangleq \tau_a \wedge n$. It follows that,

$$\left\{\max_{i\in[n]} X_i > a\right\} = \cup_{i\in[n]} \{X_i > a\} = \{X_\tau > a\}.$$

Using this fact and Markov inequality, we get $P\left\{\max_{i\in[n]} X_i > a\right\} = P\{X_\tau > a\} \leqslant \frac{\mathbb{E}[X_\tau]}{a}$. Since $\tau \leqslant n$ is a bounded stopping time, the result follows from Remark 2. $\square$

Corollary 2.2. *For a martingale X adapted to $\mathcal{F}_\bullet$ and positive constant a,*

$$P\left\{\max_{i\in[n]} |X_i| > a\right\} \leqslant \frac{\mathbb{E}|X_n|}{a}, \qquad P\left\{\max_{i\in[n]} |X_i| > a\right\} \leqslant \frac{\mathbb{E}X_n^2}{a^2}.$$

Proof. Since X is a martingale adapted to $\mathcal{F}_\bullet$, and maps $x \mapsto |x|$ and $x \mapsto x^2$ are convex functions, it follows that $(|X_n| : n \in \mathbb{Z}_+)$ and $(X_n^2 : n \in \mathbb{Z}_+)$ are nonnegative submartingales adapted to $\mathcal{F}_\bullet$. The result follows from and Kolmogorov's inequality for submartingales. $\square$

Theorem 2.3 (Strong law of large numbers). *Let $S : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a random walk with* i.i.d. *step size X having finite mean μ. If the moment generating function $t \mapsto M(t) \triangleq \mathbb{E}e^{tX_1}$ exists for all $t \in \mathbb{R}_+$, then $\lim_{n\in\mathbb{N}} \frac{S_n}{n} = \mu$ almost surely.*

Proof. We fix $\epsilon > 0$.

(a) We define the following two maps for all $t \in \mathbb{R}_+$,

$$t \mapsto g(t) \triangleq \frac{e^{t(\mu+\epsilon)}}{M(t)}, \qquad t \mapsto f(t) \triangleq \frac{e^{t(\mu-\epsilon)}}{M(t)}.$$

Since $M(0) = 1$, it follows that $g(0) = f(0) = 1$. From the fact that $M(0) = 1$ and $M'(0) = \mu = \mathbb{E}X_1$, we obtain

$$g'(0) = \frac{M(0)(\mu+\epsilon) - M'(0)}{M^2(0)} = \epsilon > 0, \qquad f'(0) = -\epsilon < 0.$$

Hence, there exists $t_0, t_1 > 0$ such that $g(t_0) > 1 > f(t_1)$.

(b) We define random sequences $Y, Z : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ for each $n \in \mathbb{N}$ as

$$Y_n \triangleq \frac{e^{t_0 S_n}}{M^n(t_0)} = \prod_{i=1}^n \frac{e^{t_0 X_i}}{M(t_0)}, \qquad Z_n \triangleq \frac{e^{t_1 S_n}}{M^n(t_1)} = \prod_{i=1}^n \frac{e^{t_1 X_i}}{M(t_1)}.$$

We observe that Y_n and Z_n are a product of independent nonnegative random variables with unit mean, and hence are nonnegative product martingales with $\lim_n \mathbb{E}Y_n = \lim_n \mathbb{E}Z_n = 1$.

(c) We next show that $P\{S_n/n > \mu + \epsilon \text{ i.o.}\} = 0$. To this end, note that

$$\left\{\frac{S_n}{n} \geqslant \mu + \epsilon\right\} \subseteq \left\{\frac{e^{t_0 S_n}}{M(t_0)^n} \geqslant g(t_0)^n\right\} = \{Y_n \geqslant g(t_0)^n\}. \tag{1}$$

By martingale convergence theorem, the limit $\lim_{n\in\mathbb{N}} Y_n$ exists, is finite, and has mean unity. Since $g(t_0) > 1$, it follows from (1) that

$$0 = P\{Y_n \geqslant g(t_0)^n \text{ infinitely often}\} \geqslant P\{S_n/n \geqslant \mu + \epsilon \text{ infinitely often}\} \geqslant 0.$$

(d) Similarly, we can show that $P\{S_n < n(\mu - \epsilon) \text{ i.o.}\} = 0$. We note that $\{S_n/n \leqslant \mu - \epsilon\} \subseteq \{Z_n \leqslant f(t_1)^n\}$. By martingale convergence theorem, the limit $\lim_{n\in\mathbb{N}} Z_n$ exists, is finite, and has mean unity. Since $f(t_1) > 1$, it follows that

$$0 = P\{Z_n \leqslant f(t_1)^n \text{ infinitely often}\} \geqslant P\{S_n/n \leqslant \mu - \epsilon \text{ infinitely often}\} = 0.$$

(e) The result follows from combining both these results, and taking the limit of arbitrary $\epsilon \downarrow 0$.

$\square$

2.1 Generalized Azuma inequality

Lemma 2.4. *Consider a random variable $X : \Omega \to [-\alpha,\beta]$ with mean $\mathbb{E}X = 0$, and any convex function $f : \mathbb{R} \to \mathbb{R}$. Then,*

$$\mathbb{E}f(X) \leqslant \frac{\beta}{\alpha+\beta} f(-\alpha) + \frac{\alpha}{\alpha+\beta} f(\beta).$$

Proof. From convexity of f, any point (X,Y) on the line joining points $(-\alpha, f(-\alpha))$ and $(\beta, f(\beta))$ is

$$Y = f(-\alpha) + (X+\alpha)\frac{f(\beta)-f(-\alpha)}{\beta+\alpha} \geqslant f(X).$$

Result follows from taking expectations on both sides. □

Lemma 2.5. *For $\theta \in [0,1]$ and $\bar{\theta} \triangleq 1-\theta$, we have $\theta e^{\bar{\theta}x} + \bar{\theta}e^{-\theta x} \leqslant e^{x^2/8}$ for all $x \in \mathbb{R}$.*

Proof. We define $\alpha \triangleq 2\theta - 1, \beta \triangleq \frac{x}{2}$, and a map $f(\alpha,\beta) \triangleq \cosh\beta + \alpha\sinh\beta - e^{\alpha\beta+\beta^2/2}$ for all $\alpha \in [-1,1]$ and $\beta \in \mathbb{R}$. We observe that $f(\alpha,0) = 0$ and

$$\theta e^{\bar{\theta}x} + \bar{\theta}e^{-\theta x} - e^{x^2/8} = \frac{(1+\alpha)}{2}e^{(1-\alpha)\beta} + \frac{(1-\alpha)}{2}e^{-(1+\alpha)\beta} - e^{\beta^2/2} = e^{-\alpha\beta}f(\alpha,\beta).$$

We observe that $f(\alpha,\beta) < 0$ for $|\alpha| = 1$ and sufficiently large β. We will show that there are no stationary points for $f(\alpha,\beta)$ other than $\alpha = 0$ or $\beta = 0$, which implies that $f(\alpha,\beta) \leqslant 0$ for all $\alpha \in [-1,1]$ and $\beta \in \mathbb{R}$. To find the stationary points of $f(\alpha,\beta)$, we take the partial derivative with respect to variables α and β, equate them to zero, and obtain

$$\sinh\beta + \alpha\cosh\beta = (\alpha+\beta)e^{\alpha\beta+\beta^2/2}, \qquad \sinh\beta = \beta e^{\alpha\beta+\beta^2/2}.$$

If $\beta \neq 0$, then the stationary point satisfies $1 + \alpha\coth\beta = 1 + \frac{\alpha}{\beta}$, with the only solution being $\beta = \tanh\beta$ for $\alpha \neq 0$. By Taylor series expansion, it can be seen that $\beta = 0$ is the unique solution to this equation. □

Proposition 2.6. *Let $\alpha,\beta,a,b > 0$. If X a zero mean martingale adapted to filtration $\mathcal{F}_\bullet$ such that $X_n - X_{n-1} \in [-\alpha,\beta]$ for each $n \in \mathbb{N}$, then*

$$P\Big(\cup_{n\in\mathbb{N}}\{X_n \geqslant a+bn\}\Big) \leqslant e^{-\frac{8ab}{(\alpha+\beta)^2}}. \tag{2}$$

Proof. Let $X_0 = 0$ and $c > 0$.

(a) We define a random sequence $W : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ for each $n \in \mathbb{Z}_+$ as $W_n \triangleq e^{c(X_n - a - bn)}$. We will show that W is a supermartingale adapted to the filtration $\mathcal{F}_\bullet$. We fix $n \in \mathbb{N}$. Since X is adapted to $\mathcal{F}_\bullet$ and W_n is a deterministic function of X_n, it follows that $\sigma(W_n) \subseteq \sigma(X_n) \subseteq \mathcal{F}_n$. We observe that $|W_n| = W_n$ and $\mathbb{E}W_0 = e^{-ca} < 1$. Writing $W_n = W_{n-1}e^{-cb}e^{c(X_n - X_{n-1})}$, and taking conditional expectation on both sides, we obtain

$$\mathbb{E}[W_n \mid \mathcal{F}_{n-1}] = W_{n-1}e^{-cb}\mathbb{E}[e^{c(X_n-X_{n-1})} \mid \mathcal{F}_{n-1}]. \tag{3}$$

Applying Lemma 2.4 to the convex function $f(x) = e^{cx}$ and conditionally zero mean random variable $X_n - Xn - 1 \in [-\alpha,\beta]$, replacing expectation with conditional expectation, the fact that $\mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] = 0$, and applying Lemma 2.5 to $\theta \triangleq \frac{\alpha}{(\alpha+\beta)} \in [0,1]$ and $x \triangleq c(\alpha+\beta) \in \mathbb{R}_+$, we obtain that

$$\mathbb{E}[e^{c(X_n-X_{n-1})}|\mathcal{F}_{n-1}] \leqslant \frac{\beta e^{-c\alpha} + \alpha e^{c\beta}}{\alpha+\beta} = \bar{\theta}e^{-c(\alpha+\beta)\theta} + \theta e^{c(\alpha+\beta)\bar{\theta}} \leqslant e^{c^2(\alpha+\beta)^2/8}. \tag{4}$$

Setting the value $c \triangleq \frac{8b}{(\alpha+\beta)^2}$, and substituting (4) in (3), we obtain $\mathbb{E}[W_n \mid \mathcal{F}_{n-1}] \leqslant W_{n-1}e^{-cb+\frac{c^2(\alpha+\beta)^2}{8}} = W_{n-1}$. Thus, we have shown that W is a supermartingale adapted to $\mathcal{F}_\bullet$.

(b) Since W is a nonnegative random sequence and a supermartingale, we recursively obtain that $\mathbb{E}|W_n| = \mathbb{E}W_n < \mathbb{E}W_0 \leqslant 1$ for all $n \in \mathbb{N}$. For a fixed positive integer k, we define the following bounded stopping time $\tau \triangleq \inf\{n \in \mathbb{N} : X_n \geqslant a+bn\} \wedge k$, such that $\{X_\tau \geqslant a+b\tau\} = \cup_{n=1}^k\{X_n \geqslant a+bn\}$. Thus, from the definition of W, application of Markov inequality and optional stopping theorem, and supermartingale property of W, we get

$$P\Big(\cup_{n=1}^k\{X_n \geqslant a+bn\}\Big) = P\{X_\tau \geqslant a+b\tau\} = P\{W_\tau \geqslant 1\} \leqslant \mathbb{E}[W_\tau] \leqslant \mathbb{E}[W_0] = e^{-ca} = e^{-\frac{8ab}{(\alpha+\beta)^2}}.$$

Since, the choice of k was arbitrary, the result follow from letting $k \to \infty$.

□

Theorem 2.7 (Generalized Azuma inequality). *Let $\alpha, \beta, c > 0$ and $m \in \mathbb{Z}_+$. If X is a zero mean martingale adapted to $\mathcal{F}_\bullet$ such that $X_n - X_{n-1} \in [-\alpha, \beta]$ for each $n \in \mathbb{N}$, then*

$$P\Big(\cup_{n \geqslant m} \{X_n \geqslant nc\} \Big) \leqslant e^{-\frac{2mc^2}{(\alpha+\beta)^2}}, \qquad P\Big(\cup_{n \geqslant m} \{X_n \leqslant -nc\} \Big) \leqslant e^{-\frac{2mc^2}{(\alpha+\beta)^2}}.$$

Proof. We fix $m \in \mathbb{Z}_+, n \geqslant m$ and define $a \triangleq \frac{mc}{2}, b \triangleq \frac{c}{2}$. It follows that $\{X_n \geqslant nc\} \subseteq \{X_n \geqslant a + bn\}$. Applying Proposition 2.6 to zero mean martingale X, we get

$$P\Big(\cup_{n \geqslant m} \{X_n \geqslant nc\} \Big) \leqslant P\Big(\cup_{n \geqslant m} \{X_n \geqslant a + bn\} \Big) \leqslant P\Big(\cup_{n \in \mathbb{N}} \{X_n \geqslant a + bn\} \Big) \leqslant e^{-\frac{8(\frac{mc}{2})(\frac{c}{2})}{(\alpha+\beta)^2}} = e^{-\frac{2mc^2}{(\alpha+\beta)^2}}.$$

This proves first inequality, and second inequality follows by considering the martingale $-X$. □

A Uniform integrability

Definition A.1. A random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ with distribution function $F_n \triangleq F_{X_n}$ for each $n \in \mathbb{N}$, is said to be **uniformly integrable** if for every $\epsilon > 0$, there is a y_ϵ such that for each $n \in \mathbb{N}$

$$\mathbb{E}[|X_n| \mathbb{1}_{\{|X_n| > y_\epsilon\}}] = \int_{|x| > y_\epsilon} |x| \, dF_n(x) < \epsilon.$$

Lemma A.2. *If a random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ is uniformly integrable then there exists a finite M such that $\mathbb{E}|X_n| < M$ for all $n \in \mathbb{N}$.*

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

$$\mathbb{E}|X_n| = \int_{|x| \leqslant y_1} |x| dF_n(x) + \int_{|x| > y_1} |x| dF_n(x) \leqslant y_1 + 1.$$

□