13 Doob's Maximal Inequality

The following result is known either as Doob's Maximal Inequality or as Doob's Submartingale Inequality. It resembles the well-known Markov inequality, except that there is a "max" in the left hand side.

Theorem 13.1. *Suppose $(X_n)_{n\geq 1}$ is a submartingale with $X_n \geq 0$ almost surely for all n. Then for all $a > 0$ we have*

$$\mathbb{P}[\max_{1\leq i\leq n}(X_i) \geq a] \leq \frac{\mathbb{E}[X_n]}{a}$$

Proof. Take the Doob decomposition $X_n = M_n + A_n$. Set $T = \min\{k : X_k \geq a\}$, with $T = +\infty$ if $X_k < a$ for all k. Then we have almost surely that

$$X_{T\wedge n} \geq aI_{\{T\leq n\}}$$

and taking expectations gives

$$a\mathbb{P}[T \leq n] \leq \mathbb{E}[X_{T\wedge n}] = \mathbb{E}[M_{T\wedge n} + A_{T\wedge n}].$$

By Lemma 7.2, $\mathbb{E}[M_{T\wedge n}] = \mathbb{E}[M_1] = \mathbb{E}[M_n]$, while since A_n is a.s. nondecreasing in n we have $A_{T\wedge n} \leq A_n$ a.s. so that $\mathbb{E}[A_{T\wedge n}] \leq \mathbb{E}[A_n]$. Thus

$$a\mathbb{P}[\max_{1\leq i\leq n}(X_i) \geq a] = a\mathbb{P}[T \leq n] \leq \mathbb{E}[M_n] + \mathbb{E}[A_n] = \mathbb{E}[X_n]. \qquad \square$$

The next result is a version of what is known in the literature either as Azuma's inequality or as the Azuma-Hoeffding inequality.

Theorem 13.2. *Suppose $(M_n)_{n\geq 0}$ is a martingale with $M_0 = 0$, and suppose for some constnat c that for all $n \geq 0$ we have $|M_{n+1} - M_n| \leq c$ almost surely. Then for any $a > 0$ we have*

$$\mathbb{P}[\sup_{1\leq k\leq n}(M_k) > a] \leq \exp\left(-\frac{a^2}{2nc^2}\right).$$

Proof. In two steps.

Step (i): Suppose X is a random variable with $|X| \leq c$ almost surely and $\mathbb{E}[X] = 0$. Let $\theta > 0$. Then the function $f(x) = e^{\theta x}$ is a convex function of x, so for $|x| \leq c$ it is bounded above by the linear function $g(x)$ which satisfies $g(x) = f(x)$ for $x = -c$ and $x = c$. By the two assumptions on X we then obtain

$$\mathbb{E}[f(X)] \leq \mathbb{E}[g(X)] = \cosh(\theta c) \leq \exp((1/2)\theta^2 c^2)$$

where the last bound comes from comparing terms in the power series.

Step (ii): Set $X_n = f(M_n) = e^{\theta M_n}$, which is nonnegative, and is a submartingale because of the convexity of f. By the maximal inequality,

$$\mathbb{P}[\sup_{1\leq k\leq n}(M_n) \geq a)] = \mathbb{P}[\sup_{1\leq k\leq n}(X_n) \geq f(a)] \leq \mathbb{E}[X_n]/f(a) \tag{13.1}$$

and (by the tower and TOWIK properties)

$$\mathbb{E}[X_n] = \mathbb{E}[\mathbb{E}[X_n|\mathcal{F}_{n-1}|\mathcal{F}_{n-1}]] = \mathbb{E}[e^{\theta M_{n-1}}\mathbb{E}[e^{\theta(M_n - M_{n-1})}|\mathcal{F}_{n-1}]].$$

Since $\mathbb{E}[M_n - M_{n-1}|\mathcal{F}_n] = 0$ and $|M_n - M_{n-1}| \leq c$ almost surely, we can apply Step (i) to the conditional distribution of $M_n - M_{n-1}$ given $\mathcal{F}_{n-1}$ to deduce that $\mathbb{E}[X_n] \leq \mathbb{E}[X_{n-1}\exp((1/2)\theta^2 c^2)]$. Repeating this argument n times yields

$$\mathbb{E}[X_n] \leq \exp((n/2)\theta^2 c^2)\mathbb{E}[X_0] = \exp((n/2)\theta^2 c^2).$$

Substituting into (13.1) yields

$$\mathbb{P}[\sup_{1\leq k\leq n}(M_n) \geq a)] \leq \exp((n/2)\theta^2 c^2 - \theta a)$$

and taking $\theta = a/(nc^2)$ gives the result. $\square$

14 Lévy's upward/downward theorems

Theorem 11.4 shows that every uniformly integrable martingale is of the type that data about some (hidden) random variable is accumulated. Conversely, every martingale of this type is uniformly integrable and convergent to the hidden variable. This is the content of the following Lévy's upward theorem.

Theorem 14.1 (Lévy's upward theorem). *Let X be an integrable random variable and $\{\mathcal{F}_n\}$ be a filtration. Define $M_n = \mathbb{E}[X|\mathcal{F}_n]$, and let $\mathcal{F} = \sigma(\cup_{n=1}^{\infty}\mathcal{F}_n)$ be the σ-algebra generated by $\bigcup_n \mathcal{F}_n$. Then $\{M_n\}$ is a uniformly integrable martingale and*

$$\lim_{n\to\infty} M_n = \mathbb{E}[X|\mathcal{F}] \qquad \textit{almost surely and in } L^1.$$

In particular, if X is measurable with respect to $\mathcal{F}$ then $M_n \to X$ almost surely and in L^1.

The key to the proof of the theorem is the following lemma.

Lemma 14.2. *Let X be an integrable random variable and $\{\mathcal{F}_n\}$ be a sequence of σ-algebras. If $M_n = \mathbb{E}[X|\mathcal{F}_n]$, then the sequence $\{M_n\}$ is uniformly integrable.*

Proof. For $\varepsilon > 0$, choose $\delta > 0$ such that, for $F \in \mathcal{A}$,

$$P(F) < \delta \quad \Rightarrow \quad \int_F |X|\, dP < \varepsilon.$$

This is possible, see Exercise 4 in the problem sheet, applied to $|X|$.

Having δ at our disposal, we choose $K > \frac{1}{\delta}\mathbb{E}[|X|]$. By considering positive and negative parts of X separately (or using Jensen's inequality), we obtain $|M_n| = |\mathbb{E}[X|\mathcal{F}_n]| \leq \mathbb{E}[|X|\,|\mathcal{F}_n]$ almost surely. Hence, by Chebyshev's inequality,

$$P(|M_n| > K) \leq \frac{\mathbb{E}[|M_n|]}{K} \leq \frac{\mathbb{E}\big[\mathbb{E}[|X|\,|\mathcal{F}_n]\big]}{K} = \frac{\mathbb{E}[|X|]}{K} < \delta.$$

Note that this event is in $\mathcal{F}_n$. From the definition of conditional expectation, we obtain that, for all n,

$$\int_{\{|M_n|>K\}} |M_n|\, dP \leq \int_{\{|M_n|>K\}} \mathbb{E}[|X|\,|\mathcal{F}_n]\, dP = \int_{\{|M_n|>K\}} |X|\, dP < \varepsilon.$$

This finishes our proof. $\blacksquare$

Proof of Theorem 14.1. We already know that $\{M_n\}$ is a martingale and is uniformly integrable. Hence by Theorem 11.4, there is a random variable M_∞ such that

$$\lim_{n\to\infty} M_n = M_\infty \qquad \text{almost surely and in } L^1.$$

We have to show that $M_\infty = \mathbb{E}[X|\mathcal{F}]$ almost surely. We may assume that $X \geq 0$, which also implies $M_n \geq 0$ and hence $M_\infty \geq 0$ almost surely. Observe that

$$\mathbb{E}[X] = \mathbb{E}[M_n] \to \mathbb{E}[M_\infty].$$

Define the measures μ and ν on $\mathcal{F}$ (defined to be the σ-algebra generated by $\bigcup_n \mathcal{F}_n$) by

$$\mu(A) = \int_A X dP, \qquad\qquad \nu(A) = \int_A M_\infty dP.$$

Now $\bigcup_n \mathcal{F}_n$ is a π-system (a family of subsets of Ω stable under *finite* intersection; see Williams, Chapter 1.6) which generates $\mathcal{F}$. If $A \in \mathcal{F}_n$, then, for all $m \geq n$,

$$\mu(A) = \int_A X dP = \int_A M_m dP \to \int_A M_\infty dP = \nu(A).$$

The uniqueness theorem (i.e., two measures that agree on a π-system agree on the σ-algebra generated by that π-system; see Williams, Chapter 1.6) now yields $\mu(A) = \mu(A)$ for any $A \in \mathcal{F}$. Hence $\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[M_\infty \mathbf{1}_A]$ for any $A \in \mathcal{F}$, and since also M_∞ is $\mathcal{F}$-measurable this shows that $M_\infty = \mathbb{E}[X|\mathcal{F}]$.

■

Theorem 14.3 (Lévy's downward theorem). *Suppose that $\{\mathcal{G}_{-n} : n \in \mathbb{N}\}$ is a collection of σ-algebras such that*

$$\mathcal{G}_{-\infty} := \bigcap_{k=1}^{\infty} \mathcal{G}_{-k} \subset \cdots \subset \mathcal{G}_{-n} \subset \cdots \subset \mathcal{G}_{-2} \subset \mathcal{G}_{-1}.$$

Let X be an integrable random variable and define

$$X_{-n} = \mathbb{E}[X|\mathcal{G}_{-n}].$$

Then,

$$\lim_{n\to\infty} X_{-n} = \mathbb{E}[X|\mathcal{G}_{-\infty}] \qquad \textit{almost surely and in } L^1.$$

Proof. Fix a positive integer N. We look at the filtration $\{\mathcal{F}_n\}$ given by

$$\mathcal{F}_n = \mathcal{G}_{(n-N)\wedge(-1)}$$

and the adapted process $Y_n = X_{(n-N)\wedge(-1)}$. By the tower property,

$$\mathbb{E}[Y_n|\mathcal{F}_{n-1}] = \mathbb{E}\Big[\mathbb{E}\big[X\big|\mathcal{G}_{(n-N)\wedge(-1)}\big]\Big|\mathcal{G}_{(n-1-N)\wedge(-1)}\Big] = \mathbb{E}\big[X\big|\mathcal{G}_{(n-1-N)\wedge(-1)}\big] = Y_{n-1},$$

which is a martingale. From Doob's upcrossing lemma,

$$(b-a)\,\mathbb{E}\big[U_N[a,b]\big] \leq \mathbb{E}[(a-Y_N)\vee 0] = \mathbb{E}[(a-X_{-1})\vee 0].$$

Letting $N \to \infty$ shows that the total number of downcrossings of $[a,b]$ by the process $\{X_{-n}\}$ is finite almost surely. Now one has this simultaneously for all rationals $a < b$, and we can argue for $\{X_{-n}\}$ as in the martingale convergence theorem. Hence, $\lim_{n\to\infty} X_{-n} = X_{-\infty}$ exists almost surely. By Lemma 14.2, the sequence is even uniformly integrable and hence convergence holds in L^1. To see that $X_{-\infty} = \mathbb{E}[X|\mathcal{G}_{-\infty}]$, first observe that, for all m and $G \in \mathcal{G}_{-\infty} \subset \mathcal{G}_{-m}$,

$$\int_G X dP = \int_G X_{-m}\, dP,$$

and then let $m \to \infty$ to see that $X_{-\infty}$ satisfies the properties of a conditional probability of X given $\mathcal{G}_{-\infty}$. ∎

15 Classical probability theorems via martingales

The **Kolmogorov zero-one law** is an extremely useful tool to see that certain events necessarily have probability zero or one.

Theorem 15.1. *Suppose that $\{X_n\}$ is a sequence of independent random variables, and let*

$$\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \dots), \qquad \mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{T}_n,$$

where $\mathcal{T}$ is called the tail-σ-algebra. *Then (i) every integrable $\mathcal{T}$-measurable random variable X is a.s. constant (i.e. satisfies Var$[X] = 0$), and (ii) every event $A \in \mathcal{T}$ has $P(A) = 0$ or $P[A] = 1$*

Examples of tail events $A \in \mathcal{T}$ are

- $\{X_n \in A_n$ for infinitely many $n\}$,
- $\{\lim_{n\to\infty} X_n$ exists$\}$, or
- $\{\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n X_i = \mu\}$.

Proof. Let $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$ and $\mathcal{F}_\infty = \sigma(X_1, X_2, \dots)$ $(= \sigma(\bigcup_n \mathcal{F}_n)$, the σ-algebra generated by $\bigcup_n \mathcal{F}_n)$. Note that $\mathcal{T} \subset \mathcal{F}_\infty$, though $\mathcal{T}$ is independent of $\mathcal{F}_n$ for any n.

Let X be an integrable $\mathcal{T}$-measurable random variable. Then X is $\mathcal{F}_\infty$-measurable. By Lévy's upward theorem,

$$X = \lim_{n\to\infty} \mathbb{E}[X|\mathcal{F}_n] \qquad \text{almost surely.}$$

Since X is $\mathcal{T}$-measurable and independent of $\mathcal{F}_n$, the right-hand side equals $\mathbb{E}[X]$ so X is a.s. constant. Part (ii) follows by taking $X := \mathbb{1}_A$. ∎

We now give a very short proof of Kolmogorov's **strong law of large numbers** under minimal moment conditions. Martingale theory will serve as the major tool.

Theorem 15.2. *Suppose that $\{X_n\}$ is a sequence of i.i.d. integrable random variables, with $\mu = \mathbb{E}[X_1]$. Then,*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n X_k = \mu \qquad \textit{almost surely and in } L^1.$$