1 Differences bounded by 1

Let $\{W_n\}$ be a martingale with respect to the filtration $\{\mathcal{F}_n\}$ and assume $W_0 = 0$. Define $\Delta_n = W_n - W_{n-1}$ and suppose $|W_n - W_{n-1}| \leq 1$ for all n.

Theorem 1 (Azuma's inequality). *Let $u > 0$ be any constant. Then*

$$\mathbf{P}(W_N \geq u) \leq e^{-u^2/2N} .$$

PROOF: The first step is to show that for $\lambda > 0$,

$$M_n := \exp\left(\lambda W_n - \lambda^2 \frac{N}{2}\right)$$

is a supermartingale. Let ν be the conditional law of Δ_n given $\mathcal{F}_{n-1}$ and let Z have law ν. The supermartingale property follows from

$$\mathbf{E}e^{\lambda Z} \leq \cosh(\lambda)$$

which holds for any random variable with mean zero in $[-1, 1]$ (proof: the worst case is $Z = \pm 1$, in which case equality holds). Furthermore, we note that $\cosh(x) \leq e^x$ for all $x > 0$, as follows from comparing Taylor coefficients.

Having verified that M_n is a supermartingale, the rest of the argument is brief. We have

$$\mathbf{E}\exp\left(\lambda W_N - \lambda^2 \frac{N}{2}\right) \leq 1 ,$$

hence

$$\mathbf{E}\exp(\lambda W_N) \leq e^{N\lambda^2/2} .$$

Markov's inequality gives

$$\mathbf{P}(W_N \geq u) \leq \exp(\lambda^2 N/2 - \lambda u),$$

and setting $\lambda = u/N$ gives

$$\mathbf{P}(W_N \geq u) \leq \exp(-u^2/(2N)) .$$

$\square$

2 Conditional quadratic L^∞ version

Let $\{W_n\}$ be a martingale with respect to the filtration $\{\mathcal{F}_n\}$ and assume $W_0 = 0$. Define $\Delta_n = W_n - W_{n-1}$ and let D_n be the conditional essential supremum of $|\Delta_n|$ given $\mathcal{F}_{n-1}$. Let $Q_n = \sum_{k=1}^n D_n^2$.

Theorem 2 (Azuma inequality with conditional bound). *With W_n, Δ_n, D_n, Q_n as above, let $K > 0$ and $u > 0$ be any constants. Then*

$$\mathbf{P}(W_N \geq u) \leq e^{-u^2/2K} + \mathbf{P}(Q_N > K).$$

Proof: The first step is to show that for $\lambda > 0$,

$$M_n := \exp(\lambda W_n - \lambda^2 Q_n/2)$$

is a supermartingale. To see this, first observe that if $\mathbf{E}Z = 0$ and $|Z| \leq L$ then $\mathbf{E}e^Z \leq \cosh(L)$; this follows from the e fact that $\mathbf{E}\phi(Z) \leq [\phi(-L) + \phi(L)]/2$ for every convex function ϕ. Also, $\cosh(z) \leq \exp(z^2/2)$ as may be seen by a glance at the power series. It remains only to observe that $D_n \in \mathcal{F}_{n-1}$ and to compute

$$\begin{aligned} \mathbf{E}(M_n \mid \mathcal{F}_{n-1}) &= M_{n-1} \exp(-\lambda^2 D_n^2/2) \mathbf{E}(e^{\lambda \Delta_n} \mid \mathcal{F}_{n-1}) \\ &\leq M_{n-1} \exp(-\lambda^2 D_n^2/2) \cosh(\lambda \Delta_n) \\ &\leq M_{n-1}. \end{aligned}$$

Now let $\tau = \sup\{n : Q_n \leq K\}$. Since $Q_n \in \mathcal{F}_{n-1}$, τ is a stopping time. Applying the above claim to the supermartingale $\{W_{\tau \wedge n}\}$ shows that

$$\mathbf{E} \exp(\lambda W_{\tau \wedge N} - \lambda^2 Q_{\tau \wedge N}/2) \leq 1.$$

Since $Q_{\tau \wedge N} \leq K$, this gives

$$\mathbf{E} \exp(\lambda W_{\tau \wedge N}) \leq e^{\lambda^2 K/2}.$$

Markov's inequality gives

$$\mathbf{P}(W_{\tau\wedge N} \geq u) \leq \exp(\lambda^2 K/2 - \lambda u),$$

and setting $\lambda = u/K$ gives

$$\mathbf{P}(W_{\tau\wedge N} \geq u) \leq \exp(-u^2/2K).$$

Since $W_N = W_{\tau\wedge N}$ except on the event $Q_N > K$, this proves the theorem. $\square$