Theorem (Hoeffding-Azuma inequality)

Let $(X_t)_{t\in\mathbb{N}^*}$ be a $(\mathcal{F}_t)_{t\geq 1}$-adapted real-valued process in L^1.
Assume that for all $t\geq 1$, there exists a $\mathcal{F}_{t-1}$-measurable r.v. A_t and a positive constant c_t such that $X_t \in [A_t ; A_t + c_t]$ a.s.
Then, the martingale $(S_t)_{t\geq 1}$ defined by $S_t := \sum_{s=1}^{t}(X_s - \mathbb{E}[X_s | \mathcal{F}_{s-1}])$ satisfies:

$$\forall t\geq 1, \quad \forall \lambda\in\mathbb{R}, \quad \mathbb{E}\left[e^{\lambda S_t}\right] \leq \exp\left(\frac{\lambda^2}{8}\sum_{s=1}^{t} c_t^2\right) \qquad (1)$$

In other words, S_t is subgaussian with variance factor $\frac{1}{4}\sum_{s=1}^{t} c_s^2$. As a consequence,

$$\forall x>0, \quad \mathbb{P}\left(\max_{1\leq t'\leq t} S_{t'} \geq x\right) \leq \exp\left(-\frac{2x^2}{\sum_{s=1}^{t} c_s^2}\right) \qquad (2)$$

Proof:

- We first prove (1). Note that, for all $t\geq 1$ and $\lambda\in\mathbb{R}$,

$$\mathbb{E}\left[e^{\lambda S_t}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{\lambda S_{t-1}} \cdot e^{\lambda \underbrace{(X_t - \mathbb{E}(X_t | \mathcal{F}_{t-1}))}_{=: Y_t}} \,\middle|\, \mathcal{F}_{t-1}\right]\right]$$

$$= \mathbb{E}\left[e^{\lambda S_{t-1}} \cdot \underbrace{\mathbb{E}\left[e^{\lambda Y_t} \,\middle|\, \mathcal{F}_{t-1}\right]}_{\leq e^{\frac{\lambda^2}{2}\cdot\frac{c_t^2}{4}}}\right]$$

$\leq e^{\frac{\lambda^2}{2}\cdot\frac{c_t^2}{4}}$ a.s. by the Hoeffding lemma (conditionally on $\mathcal{F}_{t-1}$, Y_t lies in a fixed interval of width c_t).

$$\leq e^{\frac{\lambda^2}{8}c_t^2} \cdot \mathbb{E}\left[e^{\lambda S_{t-1}}\right]$$

By induction, we get $\mathbb{E}\left[e^{\lambda S_t}\right] \leqslant e^{\frac{\lambda^2}{8}\sum_{s=1}^{t} c_s^2} \cdot \mathbb{E}\left[e^{\lambda \cdot \underbrace{S_0}_{=0}}\right] = e^{\frac{\lambda^2}{8}\sum_{s=1}^{t} c_s^2}$,

which proves (1).

- To derive (2), first note that $\left(e^{\lambda S_t}\right)_{t \geqslant 1}$ is a nonnegative submartingale ($e^{\lambda S_t}$ is a non-decreasing and convex function of S_t, and $(S_t)_{t\geqslant 1}$ is a martingale). We can thus use the following result:

Lemma (Doob's inequality for nonnegative submartingales)

Let $(Z_t)_{t\geqslant 1}$ be a nonnegative submartingale. Then, for all $u > 0$,

$$\mathbb{P}\left(\max_{1\leq t' \leq t} Z_{t'} \geqslant u\right) \leqslant \frac{\mathbb{E}\left[Z_t\right]}{u}$$

Ref: cf. Williams, Probability with Martingales.

Now we use the above lemma with $Z_t = e^{\lambda S_t}$. It yields:

$$\mathbb{P}\left(\max_{1\leq t'\leq t} S_{t'} \geqslant x\right) = \mathbb{P}\left(\max_{1\leq t'\leq t} e^{\lambda S_{t'}} \geqslant e^{\lambda x}\right) \leqslant \frac{\mathbb{E}\left[e^{\lambda S_t}\right]}{e^{\lambda x}}$$

$$\leqslant e^{-\left(\lambda x - \frac{\lambda^2}{8}\sum_{s=1}^{t} c_s^2\right)} \quad \text{by (1)}$$

Optimizing in λ $\left(\lambda^* = \frac{4x}{\sum_{s=1}^{t} c_s^2}\right)$ we get:

$$\mathbb{P}\left(\max_{1\leq t'\leq t} S_{t'} \geqslant x\right) \leqslant \exp\left(-\frac{2x^2}{\sum_{s=1}^{t} c_s^2}\right),$$

which concludes the proof. ∎