

Def (Discrete-time Martingale). Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a prob. space.

sample space — all possible outcomes  
 $\sigma$ -algebra — parts of event  
 $\mathbb{P}$  — assigns prob. to each event

$\{\mathbf{X}_i, \mathcal{F}_i\}_{i=0}^n$  is a Martingale if

random variable       $\sigma$ -algebra      w.r.t. discrete time indexation

1.  $\mathcal{F}_i$ 's form a "filtration" i.e.  $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \dots \subseteq \mathcal{F}_n \subseteq \mathcal{F}$

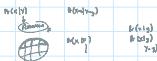
$\mathcal{F}_0$  or algebra      full  $\sigma$ -algebra  
 $\mathcal{F}_1, \mathcal{F}_2, \dots$

2.  $\mathbf{X}_i \in L^1(\Omega, \mathcal{F}_i, \mathbb{P}) \quad \forall i \in \{0, \dots, n\}$

i.e.  $\mathbf{X}_i$  is defined on the same sample space  $\Omega$   
 $\mathbb{E}(|\mathbf{X}_i|) = \int_{\Omega} |\mathbf{X}_i(\omega)| \mathbb{P}(d\omega) < \infty$ .

3.  $\forall i \in \{1, \dots, n\}$

$$\boxed{\mathbf{X}_{i+1} = \mathbb{E}(\mathbf{X}_i | \mathcal{F}_{i-1})}$$



holds almost surely.

NB:  $\mathbf{X}_j = \mathbb{E}[\mathbf{X}_i | \mathcal{F}_j] \quad \forall i > j$

$$\begin{aligned} \boxed{\mathbf{X}_{i-1} = \mathbb{E}[\mathbf{X}_i | \mathcal{F}_{i-1}] \rightsquigarrow} \\ \mathbf{X}_{i-2} = \mathbb{E}[\mathbf{X}_{i-1} | \mathcal{F}_{i-2}] = \mathbb{E}[\mathbb{E}[\mathbf{X}_i | \mathcal{F}_{i-1}] | \mathcal{F}_{i-2}] \\ = \mathbb{E}[\mathbf{X}_i | \mathcal{F}_{i-2}] \quad \text{high-res} \\ \vdots \\ \mathbf{X}_j = \mathbb{E}[\mathbf{X}_i | \mathcal{F}_j] \quad \text{low-res} \end{aligned}$$

E.g. Arbitrary filtration  $\{\mathcal{F}_i\}_{i=0}^n$

$$\mathbf{x}_i := \mathbb{E}(\mathbf{x} | \mathcal{F}_i) \quad \forall i \in \{0, 1, \dots, n\}$$

Then,  $\{\mathbf{x}_i, \mathcal{F}_i\}_{i=0}^n$  is a martingale.

Proof:

$$\begin{aligned} \mathbf{x}_{i-1} &\stackrel{\text{assmt}}{=} \mathbb{E}[\mathbf{x}_i | \mathcal{F}_{i-1}] = \mathbb{E}[\mathbb{E}[\mathbf{x} | \mathcal{F}_i] | \mathcal{F}_{i-1}] \\ &= \mathbb{E}[\mathbf{x} | \mathcal{F}_{i-1}] = \mathbf{x}_{i-1} \end{aligned}$$

NB:  $\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_n = \Omega$

$$\mathbf{x}_0 = \mathbb{E}[\mathbf{x} | \mathcal{F}_0] = \mathbb{E}[\mathbf{x}]$$

$$\mathbf{x}_n = \mathbb{E}[\mathbf{x} | \mathcal{F}_n] = \mathbf{x}$$

Theorem 2.2.1 (The Azuma-Hoeffding inequality).

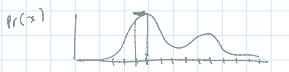
Let  $\{\mathbf{x}_k\}_{k=0}^n$  (real-valued) martingale s.t.

assume  $d_1, \dots, d_n$  s.t.  $|\mathbf{x}_k - \mathbf{x}_{k-1}| \leq d_k \quad \forall k \in \{1, \dots, n\}$

then  $\forall \lambda > 0$ ,

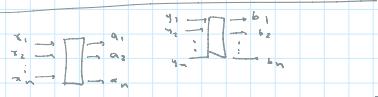
$$\mathbb{P}(|\mathbf{x}_n - \mathbf{x}_0| \geq \lambda) \leq e^{-\lambda^2 / (2 \sum_{k=1}^n d_k^2)}.$$

$x \longrightarrow x \longrightarrow x$



$\rightarrow$  guess  $x$ :

$$H_{\min} = -\log \max_x P(x)$$



$$s = (x, y^n) \quad z = (a^n, b^n)$$

$$H_{\infty}(R|S) = -\log \max_x P(z|x)$$

$$= \min_{a^n b^n} (-\log P(a^n b^n | x^n y^n))$$

Claim: lower bound  $-\log P(a^n b^n | x^n y^n) \leq a^n b^n \leq x^n y^n$

$$\text{assert: } -\log_p P(a^n b^n | x^n y^n) = \sum_{i=1}^n -\log P(a_i b_i | x_i y_i w^i)$$

$$w^i := (a_i^{-1} b_i^{-1} x_i^{i-1} y_i^{i-1})$$

$$a^i := (a_1, \dots, a_j)$$

$$b^i := (b_1, \dots, b_j)$$

$$x^i := (x_1, \dots, x_j)$$

$$y^i := (y_1, \dots, y_j)$$

Narrator: Define can be characterized by

$$P(a_i b_i | x_i y_i w^i)$$

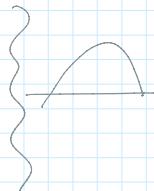
$$I(w^i) = \sum_{a b x y} c_{a b x y} \dots$$

"Then":  $\exists f$  ct.

a convex fn

$$-\log_p P(a^n b^n | x^n y^n) \geq \sum_{i=1}^n f(I(w^i))$$

$$\geq n f\left(\underbrace{\frac{1}{n} \sum_{i=1}^n I(w^i)}_{\text{indicator}}\right)$$



$$\theta f(x) + (1-\theta)f(y) \geq f(\theta x + (1-\theta)y)$$

Def<sup>n</sup>:

$$\hat{I}_i := \sum_{a b x y} c_{a b x y} \frac{\chi(a_i = a, b_i = b, x_i = x, y_i = y)}{P(x, y)}$$

Claim:  $I(w^i) = E(\hat{I}_i | w^i)$

Def<sup>n</sup>:

$$\hat{I} = \frac{1}{n} \sum_{i=1}^n \hat{I}_i$$

Narrator:  $\hat{I}$  is an estimator for  $\frac{1}{n} \sum_{i=1}^n I(w^i)$

Def<sup>n</sup>:  $z^0 := \min_{x, y} P(x, y) = g$

Assum<sup>n</sup>:  $g > 0$ .

Def<sup>n</sup>:  $z^k := \sum_{i=1}^k (\hat{I}_i - \underbrace{I(w^i)}_{E(\hat{I}_i | w^i)})$

$$E(\hat{I}_i | w^i)$$

Claim:  $\{z^k, w^k\}$  is a martingale seq.  
understood to be the  $\sigma$ -algebra generated by  $w^k$ .

Proof:

$$z^0 := 0$$

$$z^{i+1} = E(z^i | w^i)$$

$$(a^{i+1}, b^{i+1}, x^{i+1}, \dots)$$

$$z^0 = E(z^1 | w^1) = E(z^1) = E(\hat{I}_1) - E(\hat{I}_1 | w^1) = 0$$

$$\neq$$

$$(1) E(|z^k|) < \infty \quad \# k$$

(2) ...

$$(3) E(z^k | w^1 \dots w^j) = E(z^k | \underbrace{w^j}_{f^{j-1}}) = z^{j-1} \quad \forall j \leq k.$$

$$\begin{aligned} z^j &= \underbrace{\sum_{i=1}^{j-1} \hat{I}_i - E(\hat{I}_i | w^i)}_{z^{j-1}} + \hat{I}_j - E(\hat{I}_j | w^j) \\ z^{j-1} &\stackrel{\text{assert}}{=} E(z^j | w^j) = \underbrace{\sum_{i=1}^{j-1} E(\hat{I}_i | w^j)}_{\vdots i \leq j} - \underbrace{E(E(\hat{I}_i | w^i) | w^j)}_{+ E(\hat{I}_j | w^j) - E(\hat{I}_j | w^j)} \\ &\quad \vdots i \leq j \\ E(\hat{I}_i | w^j) &= \hat{I}_i \quad \text{reduces ambiguity for } i < j \\ &= \sum_i \hat{I}_i - E(\hat{I}_j | w^j) = z^{j-1} \end{aligned}$$

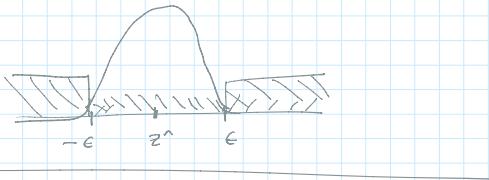
NB:  $|z^i - z^{i-1}| \leq |\hat{I}_i - I(w^i)| \leq \frac{1}{n} + I_i$

$\xrightarrow{\quad}$  max Bell violation for DM.

$\therefore x$  would be 1 for exactly one term in any experiment  
c's are bounded by 1.

Using Azuma-Hoeffding,

$$\frac{1}{n} \sum_{i=1}^n I(w^i) \approx_{\text{exp}} \frac{1}{n} \sum_{i=1}^n I_i \leftarrow$$



$$\begin{array}{c} y_1 \\ s_1 \\ \vdots \\ y_n \\ s_n \end{array} \quad \begin{array}{c} b_1 \\ a_1 \\ \vdots \\ b_n \\ a_n \end{array}$$

$$(x_1, x_2, \dots, x_n) \sim \Delta(\dots)$$

$x_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ game won} \\ 0 & \text{else.} \end{cases}$

Suppose:  $p$  is chosen randomly.  
 $\{p\}$  is a permutation on  $\{1, 2, \dots, n\}$ .

$$\text{Claim: } x_{p(1)} \cdot x_{p(2)} \cdots \cdot x_{p(n-1)} = 1,$$

$$\text{estimate } E(x_{p(n)} | x_{p(1)} \cdot \dots \cdot x_{p(n-1)} = 1)$$

$$\text{Claim: } \sum_{p \in S_n} E(x_{p(n)} | x_{p(1)} \cdots x_{p(n-1)} = 1) \approx_{\text{exp}} 1$$

$$z_i := x_{p(1)} \cdots x_{p(i-1)} - E(x_{p(1)} \cdots x_{p(i-1)})$$

$$z_0 = 0$$

$$y_i := (1 - x_{p(i)}) \quad ;$$

$$\epsilon_0 = 0$$

$$y_i := (-x_{p(i)}) ;$$

$$z_l := \sum_{i=1}^{l-1} y_i - E(y_i | F^{l-1})$$

$$\sum_{i=1}^l y_i - \sum_{i=1}^l E(y_i | F^{i-1})$$

$$z_{l-1} \stackrel{\text{required}}{=} E(z_l | F^{l-1}) = \sum_{i=1}^{l-1} \underbrace{(y_i | F^{l-1})}_{-E(y_i | F^{l-1})} - E(y_l | F^{l-1}) = z_l$$

$$E(y_i | F^{i-1})$$

$$\begin{aligned} & X_i \\ & \left( \sum_{i=1}^m I_i \approx \exp \sum_{i=1}^{l-1} \hat{I}_i \right) \\ & \text{iff} \quad \left( \sum_{i=1}^m I_i \approx \exp \sum_{i=1}^{l-1} \hat{I}_i \right) \leftarrow \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \quad \left( \frac{I_n}{\hat{I}_n} \right) \text{ don't want to measure} \\ & I_n \approx \hat{I}_n \end{aligned}$$

$$P_k \left[ \frac{1}{l} \left[ \sum_{k=1}^l E(x_k | F^{k-1}) - \sum_{k=1}^l x_k \right] < \lambda \right]$$

$$P_k \left[ \frac{1}{l} \sum_{k=1}^m x_k + \underbrace{\sum_{k=m+1}^l x_k}_{\text{upper}} - \sum_{k=1}^m x_k < \lambda + \sum_{k=m+1}^l x_k \right]$$