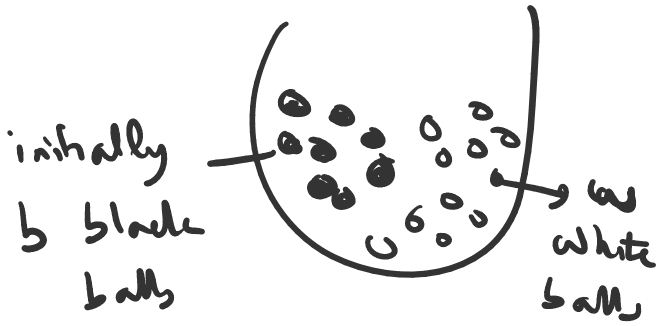
14 March 2021 23:59

Martingales continued

Example for martingale

Polya's Urn Scheme:



Perform a sequence of random selection, where at each step the chosen ball is replaced by $c+1$ balls of the same color. Let X_i denote the fraction of black balls in the urn after the i^{th} trial. Show that the sequence $X_0, X_1, \ldots,$ is a martingale.

Clearly $E[|X_i|] < \infty$. (ii)

Let Y_i be the R.V. that denotes the no. of black balls in the urn after the i^{th} iteration.

the i^{th} iteration.

So, $Y_0 = b$.

$$X_i = \frac{Y_i}{b+w+ci}.$$

(i) X_i is fully determined by $Y_0, Y_1, \ldots, Y_i$

To show:

(iii) $E[X_{n+1} \mid Y_0, \ldots, Y_n] = X_n$ for any $n \geq 0$

$$\text{LHS} = E[X_{n+1} \mid Y_0, \ldots, Y_n]$$

$$= E\left[\frac{Y_{n+1}}{b+w+c(n+1)} \,\middle|\, Y_0, \ldots, Y_n\right]$$

$$= \frac{1}{b+w+c(n+1)} E[Y_{n+1} \mid Y_0, \ldots, Y_n]$$

$$= \frac{1}{b+w+c(n+1)} \left(Y_n + \frac{Y_n c}{b+w+cn}\right)$$

drawn 1^{st} trial → $b+w$ balls in urn

$(n+1)^{th}$ trial → $b+w+cn$ balls

$$= \frac{1}{\cancel{b+w+c(n+1)}} \frac{\cancel{(b+w+c(n+1))}\, Y_n}{b+w+cn}$$

$$= Y_n$$

$$= \frac{Y_n}{b + w + cn}$$

$$= X_n$$

We have concluded that $X_0, X_1, \ldots$ is a martingale w.r.t. $Y_0, Y_1, \ldots$.

Proposition: If $Z_0, Z_1, \ldots Z_n$ is a martingale w.r.t. $X_0, \ldots, X_n$ then $Z_0, \ldots, Z_n$ is a martingale w.r.t. itself.

Proof Outline:

Given: $E[Z_{n+1} | X_0, \ldots, X_n] = Z_n$.

To show: $E[Z_{n+1} | Z_0, \ldots, Z_n] = Z_n$.

↳ $n=0$.

Given: $\underbrace{E[Z_1|X_0]}_{f(X_0)} = Z_0$, Z_0 is fully determined by X_0.

$f(X_0) = g(Z_0)$

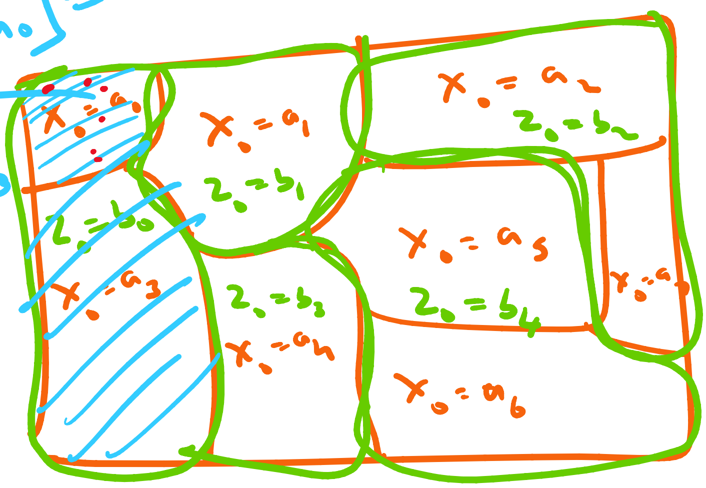
To show: $\underbrace{E[Z_1|Z_0]}_{} = Z_0$

$\sum_{z_1} z_1 \Pr[Z_1 = z_1 | X_0 = a_0]$

$\parallel$

$E[Z_1|X_0=a_0] = b_0$

$E[Z_1|X_0=a_3] = b_0$



Orange blocks show the partitions of the sample space given by R.V. X_0.

Green blocks 〃 〃 〃 〃 Z_0.

Tail Inequalities for Martingales..

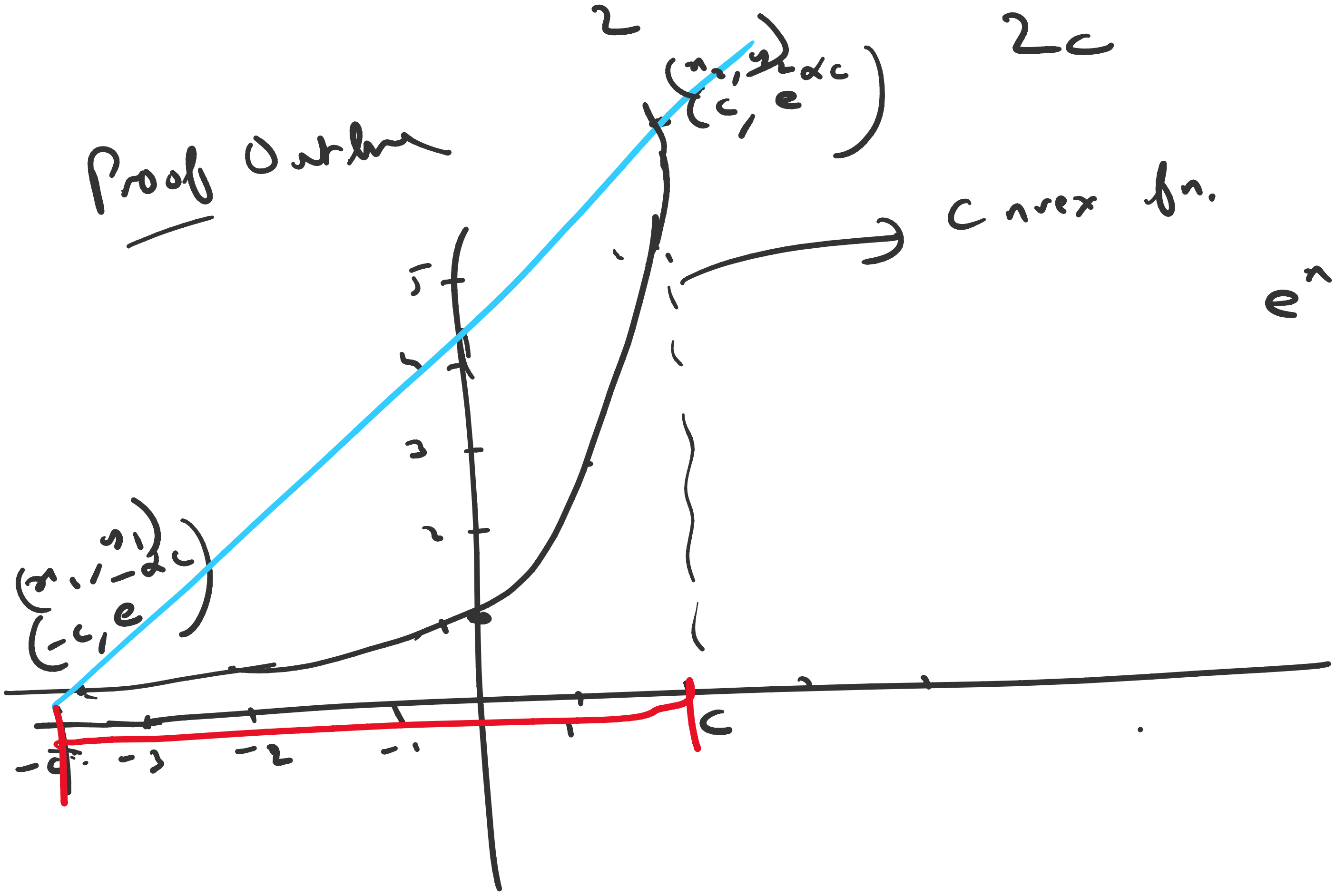
↳ Azuma-Hoeffding Inequality

Proposition 1 : For any $x \in [-c, +c]$ and for an arbitrary α,

$$e^{\alpha x} \le h(x), \quad \text{where}$$

$$h(x) = \frac{e^{\alpha c} + e^{-\alpha c}}{2} + \frac{e^{\alpha c} - e^{-\alpha c}}{2c} x .$$

Proof Outline



$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

$$\frac{y - e^{-\alpha c}}{\quad} = \frac{e^{\alpha c} - e^{-\alpha c}}{\quad}$$

$$\frac{y-e}{x+c} = \frac{e-e}{c+c}$$

$$y = \frac{e^{\alpha c} + e^{-\alpha c}}{2} + \frac{\left(e^{\alpha c} - e^{-\alpha c}\right)x}{2c} = h(x)$$

$\square$

Proposition 2: $\cosh(\alpha) := \dfrac{e^{\alpha} + e^{-\alpha}}{2} \leq e^{\frac{\alpha^2}{2}}$

$$1 + \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} + \frac{\alpha^6}{6!} + \dots$$

$$1 + \alpha^2/2 + \frac{(\alpha^2/2)^2}{2!} + \frac{(\alpha^2/2)^3}{3!} \dots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} \dots$$

$\square$

Theorem [Azuma-Hoeffding Inequality]

Theorem [Azuma–Hoeffding Inequality]

Let $Z_0, Z_1, \ldots, Z_n$ be a martingale with respect to itself with

$$|Z_i - Z_{i-1}| \leq c_i, \quad \forall i \in [n] \to \{1, \ldots, n\}$$

Then, for any $\lambda > 0$ and $t \geq 1$,

$$\Pr\left(|Z_t - Z_0| \geq \lambda\right) \leq \frac{2}{e^{\left(\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}\right)}}$$

Proof: Set $Y_i := Z_i - Z_{i-1}$, $\forall i \in [n]$.

We have,

(i) $|Y_i| \leq c_i$,

(ii) $$E\left[Y_i \mid Z_0, Z_1, \ldots, Z_{i-1}\right]$$

$$= E\left[Z_i - Z_{i-1} \mid Z_0, Z_1, \ldots, Z_{i-1}\right]$$

$$= \underbrace{E\left[Z_i \mid Z_0, \ldots, Z_{i-1}\right]} - \underbrace{E\left[Z_{i-1} \mid Z_0 \ldots Z_{i-1}\right]}$$

$$= Z_{i-1} - Z_{i-1}$$

$$= 0 \quad //$$

For any α, we have

$$E\left[e^{\alpha Y_i} \,\middle|\, Z_0, \ldots, Z_{i-1}\right]$$

$$\leq E\left[h(Y_i) \,\middle|\, Z_0, \ldots, Z_{i-1}\right] \quad \text{(from Prop 1)}$$

where

$$h(Y_i) = \underbrace{\frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2}}_{\cosh(\alpha c_i)} + \frac{e^{\alpha c_i} - e^{-\alpha c_i}}{2c} Y_i$$

$$\leq \cosh(\alpha c_i) + 0$$

$$\leq e^{\frac{\alpha^2 c_i^2}{2}} \quad // \qquad \text{(from Prop 2)}$$

Now,

$$Z_t - Z_0 = (Z_t - Z_{t-1}) + (Z_{t-1} - Z_{t-2}) + \ldots$$

$$\ldots + (Z_1 - Z_0)$$

$$= \sum_{k=1}^{t} Y_k \quad //$$

Moment generating function

... generating function,

X

$$M_X(t) = E[e^{tX}]$$

$$E\left[e^{\alpha(Z_t - Z_0)}\right] = E\left[e^{\alpha \sum_{k=1}^{t} Y_k}\right]$$

$$= E\left[\prod_{k=1}^{t} e^{\alpha Y_k}\right]$$

$Y_i = Z_i - Z_{i-1}$

$$= E\left[E\left[\prod_{k=1}^{t} e^{\alpha Y_k} \,\middle|\, Z_0, \dots, Z_{t-1}\right]\right]$$

$$= E\left[\prod_{k=1}^{t-1} e^{\alpha Y_k} E\left[e^{\alpha Y_t} \,\middle|\, Z_0, \dots, Z_{t-1}\right]\right]$$

$$\le E\left[e^{\frac{\alpha^2 c_t^2}{2}} \cdot \prod_{k=1}^{t-1} e^{\alpha Y_k}\right]$$

$$= e^{\frac{\alpha^2 c_t^2}{2}} E\left[\prod_{k=1}^{t-1} e^{\alpha Y_k}\right]$$

(repeating the same approach)

$$\vdots$$

$$\le e^{\frac{\alpha^2 \sum_{k=1}^{t} c_k^2}{2}}$$

Now, for any $\alpha > 0$,

(use $\alpha < 0$, when proving ②)

$$\Pr\left((Z_t - Z_0) \ge \lambda\right) = \Pr\left(e^{\alpha(Z_t - Z_0)} \ge e^{\alpha\lambda}\right)$$

$$\le \dots E\left[\, e^{\alpha(Z_t - Z_0)} \,\right] \dots$$

$$\leq \frac{E\left[e^{\alpha(Z_t - Z_0)}\right]}{e^{\alpha\lambda}}$$ (by Markov Ineq.)

$$\leq e^{\frac{\alpha^2\left(\sum_{k=1}^{t} c_k^2\right)}{2} - \alpha\lambda}$$

for (B), $\alpha = \frac{-\lambda}{\sum c_k^2}$

Substituting with $\alpha = \frac{\lambda}{\sum_{k=1}^{t} c_k^2}$, we get

$$= e^{\left(\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2} - \frac{\lambda^2}{\sum_{k=1}^{t} c_k^2}\right)}$$

$$= \frac{1}{e^{\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}}} \longrightarrow A$$

Similarly, we can show that

$$P_r\left((Z_t - Z_0) \leq -\lambda\right) \leq \frac{1}{e^{\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}}} \text{ --- } B$$

Combining (A) and (B), we get

$$\Pr\left[|Z_t - Z_0| \geq \lambda\right] = \Pr\left[((Z_t - Z_0) \geq \lambda) \text{ or } ((Z_t - Z_0) \leq -\lambda)\right]$$

$$\leq \Pr((Z_t - Z_0) \geq \lambda) + \Pr((Z_t - Z_0) \leq -\lambda)$$

$$\leq (A) + (B)$$

$$= \frac{2}{e^{\frac{\lambda^2}{2\sum_{k=1}^{t} c_k^2}}}$$

$\square$

<u>Corollary</u>: Let $Z_0, \ldots, Z_n$ be a martingale such that, for all $i \geq 1$,

$$|Z_i - Z_{i-1}| \leq c.$$

Then, for all $t \geq 0$ and $\lambda > 0$,

$$\Pr\left[|Z_t - Z_0| \geq \lambda c\sqrt{t}\right] \leq \frac{2}{e^{\frac{\lambda^2}{2}}}$$

A more general form is below.

<u>Theorem</u> [Azuma - Hoeffding Thm]

Let $Z_0, Z_1, \ldots, Z_n$ be a martingale such that

$$B_k \leq Z_k - Z_{k-1} \leq B_k + d_k,$$

for some constant d_k and for some R.V.s B_k that may be functions of $Z_0, \ldots, Z_{k-1}$. Then, for all $t > 0$ and any $\lambda > 0$,

... $c > 0$ that any $\lambda > 0$,

$$\Pr\left(|Z_t - Z_0| \geq \lambda\right) \leq \frac{2}{e^{\frac{2\lambda^2}{\sum_{k=1}^{t} d_k^2}}}$$

Putting $B_k = -c_k$ and $d_k = 2c_k$ gives the earlier theorem.

Lipschitz Condition . .

Let $f : D_1 \times D_2 \times \cdots \times D_n \rightarrow \mathbb{R}$ be a function with n arguments. We say f satisfies the Lipschitz condition with bound c if for any $x_1 \in D_1, x_2 \in D_2, \ldots, x_n \in D_n$, any $i \in \{1, \ldots, n\}$, and any $y_i \in D_i$,

$$\left| f(x_1, x_2, \ldots, x_{i-1}, \underline{\underline{x_i}}, x_{i+1}, \ldots, x_n) - f(x_1, x_2, \ldots, x_{i-1}, \underline{\underline{y_i}}, x_{i+1}, \ldots, x_n) \right| \leq c.$$

$\leq c.$

Let

$$Z_0 = E[f(X_1, X_2, \ldots, X_n)]$$

$\vdots$

$$Z_k = E[f(X_1, \ldots, X_n) \mid X_1, X_2, \ldots, X_k]$$

$$Z_n = E[f(X_1 \ldots X_n) \mid X_1 \ldots X_n] = f(X_1 \ldots X_n).$$

This sequence $Z_0, Z_1, \ldots, Z_n$ is a Doob martingale.

$$|Z_i - Z_{i-1}| \leq c$$

can one show this?

Theorem [McDiarmid's Theorem] Let f be a function that satisfies the Lipschitz condition with bound c. Let $X_1, \ldots X_n$ be independent random variables whose range are in $D_1, D_2, \ldots D_n$. Then,

$f: D_1 \times D_2 \times \cdots \times D_n \to \mathbb{R}$

$$P_r\left(\left|\overset{\to Z_n}{f(X_1, \ldots, X_n)} - \overset{\to Z_0}{E[f(X_1 \ldots, X_n)]}\right| \geq \lambda\right) \leq \frac{2}{e^{\frac{2\lambda^2}{nc^2}}}$$

Proof:

Proof:

To prove this, it is enough to show

$$B_k \le Z_k - Z_{k-1} \le B_k + c \quad , \forall k \in [n]$$

We use S_k for $X_1, X_2, \ldots, X_k$.

$(x_1, x_2, \ldots, x_n)$

That is, $E[f(\vec{X}) \mid X_1, \ldots, X_k] = E[f(\vec{X}) \mid S_k]$

$$f_k(\bar{X}, x) = f(X_1, X_2, \ldots X_{k-1}, x, X_{k+1}, \ldots, X_n)$$

$$f_k(\bar{z}, x) = f(z_1, z_2, \ldots, z_{k-1}, x, z_{k+1}, \ldots, z_n)$$

We have

$$Z_k - Z_{k-1} = E[f(\bar{X} \mid S_k)] - E[f(\bar{X}) \mid S_{k-1}]$$

upper bound $\le \sup_x E[f(\bar{X}) \mid S_{k-1}, X_k = x] - E[f(\bar{X}) \mid S_{k-1}]$

lower bound $\ge$

$$\geq \inf_y E\left[f(\bar{X}) \mid S_{k-1}, X_k = y\right] - E\left[f(\bar{X}) \mid S_{k-1}\right]$$

Here,

$$B_k = \inf_y E\left[f(\bar{X}) \mid S_{k-1}, X_k = y\right] - E\left[f(\bar{X}) \mid S_{k-1}\right]$$

To show:

$$\sup_x E\left[f(\bar{X}) \mid S_{k-1}, X_k = x\right] - \inf_y E\left[f(\bar{X}) \mid S_{k-1}, X_k = y\right] \leq c.$$

$$\text{LHS} = \sup_{x,y} \Big(E\left[f(\bar{X}) \mid S_{k-1}, X_k = x\right] - E\left[f(\bar{X}) \mid S_{k-1}, X_k = y\right] \Big)$$

$$= \sup_{x,y} E\left[f_k(\bar{X}, x) - f_k(X, y) \mid S_{k-1} \right]$$

Since, $X_1, X_2, \ldots, X_n$ are independent R.V,
for any $x, y, z_1, z_2, \ldots, z_{k-1}$,

$$E\left[f_k(\bar{X}, x) - f_k(X, y) \mid X_1 = z_1, \ldots, X_{k-1} = z_{k-1} \right]$$

is equal to

is equal to

$$\sum_{z_{k+1},\dots,z_n} \underbrace{\Pr\big((X_{k+1}=z_{k+1})\cap\dots\cap(X_n=z_n)\big)}\cdot\underbrace{\big(f_k(\bar z,x)-f_k(\bar z,y)\big)}$$

$\bar z = (z_1, z_2, \dots, z_n)$

$$\le c \sum_{z_{k+1},\dots,z_n} \Pr\big((z_{k+1}=z_{k+1})\cap\dots\cap(z_n=z_n)\big) \le c$$

$\le c$

Now, the thm follows from the Azuma-Hoeffding proved above

$\square$.