last time:

Basic Example

We roll a fair die 10,000 times.

let X be the total # of points. then $E(X) = 35{,}000$.

Chebyshev Inequality:

$$P\{|X - E(X)| \geq 1{,}000\} \leq \frac{7}{240} \approx 0.029$$

last time:

Bernstein Inequality: If $X_1, \cdots X_n$ are indep r.v. s.t.

① $E(X_k) = 0$, $k = 0, 1, \cdots n$

② $|X_k| \leq 1$, $k = 0, 1, \cdots n$

$X = X_1 + \cdots X_n$

For $X = X_1 + \cdots X_n$ and $a \geq 0$, we have:

$$P(X \geq a) \leq e^{-a^2/2n}$$

$$P(X \leq -a) \leq e^{-a^2/2n}$$

$P(X \geq a) \leq e^{-a^2/2n}$

$P(X \leq -a) \leq e^{-a^2/2n}$

$E(X_k) = 0$

$|X_k| \leq 1$

Very Basic Example, Continues

$P(X \geq a) \leq e^{-a^2/2n}$, $P(X \leq -a) \leq e^{-a^2/2n}$

$E(X_k) = \frac{7}{2}$. $E\left(X_k - \frac{7}{2}\right) = 0$

$-\frac{5}{2} \leq X_k - \frac{7}{2} \leq \frac{5}{2}$, so $-1 \leq Y_k = \frac{2}{5}\left(X_k - \frac{7}{2}\right) \leq 1$

$X = X_1 + \cdots + X_{10{,}000}$

X_k is the # shown on the k-th row

$$E(X_k) = \frac{1+2+3+4+5+6}{6} = \frac{7}{2}, \quad \text{with } 1 \leq X_k \leq 6$$

$P(|X - 35000| \geq 1{,}000)$

$= P(X \geq 36{,}000) + P(X \leq 34{,}000)$

Scaling X_k: consider $X_k - \frac{7}{2}$, then $E\left(X_k - \frac{7}{2}\right) = 0$

$Y_k = \frac{2}{5}\left(X_k - \frac{7}{2}\right)$

$$-\frac{5}{2} = 1 - \frac{7}{2} \leq X_k - \frac{7}{2} \leq 6 - \frac{7}{2} = \frac{5}{2}$$

$\frac{5}{2} Y_k = X_k - \frac{7}{2}$, $\rightarrow X_k = \frac{5}{2} Y_k + \frac{7}{2}$

let $Y_k = \frac{2}{5}\left(X_k - \frac{7}{2}\right) = \frac{2}{5} X_k - \frac{7}{5}$, $X_k = \left(Y_k + \frac{7}{5}\right) \times \frac{5}{2} = \frac{5}{2} Y_k + \frac{7}{2}$

$\frac{16{,}000}{20{,}000}$

Now: $E(Y_k)=0$, $|Y_k|\le 1$

For $Y=Y_1+\dots+Y_{10{,}000}$, we have:

$a\ge0$, $P(Y\ge a)\le e^{-\frac{a^2}{2n}}=e^{-\frac{a^2}{20{,}000}}$

$P(Y\le -a)\le e^{-\frac{a^2}{2n}}=e^{-\frac{a^2}{20{,}000}}$

$$P(X\ge 36{,}000)=P\left(\tfrac{5}{2}Y+35{,}000\ge 36{,}000\right)$$
$$=P\left(\tfrac{5}{2}Y\ge 1{,}000\right)$$
$$=P(Y\ge 400)$$
$$\le e^{-\frac{400^2}{20{,}000}}=e^{-8}$$

$$P(X\le 34{,}000)=P\left(\tfrac{5}{2}Y+35{,}000\le 34{,}000\right)$$
$$=P\left(\tfrac{5}{2}Y\le -1{,}000\right)$$
$$=P(Y\le -400)$$
$$\le e^{-\frac{(-400)^2}{20{,}000}}=e^{-8}$$

This shows: $P(|X-E(X)|\ge \underbrace{1{,}000}_{}) \le 2e^{-8}\approx 0.00067$ ✓

When this # is large enough, Bernstein beats Chebyshev

So $X=\frac{5}{2}Y+\frac{7}{2}\times 10{,}000$

$=\frac{5}{2}Y+35{,}000$

$P(X\ge a)\le e^{-a^2/2n}$

$P(X\le -a)\le e^{-a^2/2n}$

$P\left(\frac{5}{2}Y+35{,}000\ge 36{,}000\right)$

$=P\left(\frac{5}{2}Y\ge 1{,}000\right)=P(Y\ge 400)\le e^{-\frac{160000}{20{,}000}}=e^{-8}$

$P\left(\frac{5}{2}Y+35{,}000\le 34{,}000\right)$

$\frac{160000}{20{,}000}$

$=P\left(\frac{5}{2}Y\le -1{,}000\right)$

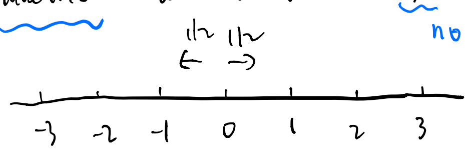
$=P(Y\le -400)$

$\le e^{-\frac{400^2}{20{,}000}}=e^{-\frac{160000}{20000}}=e^{-8}$

$P(\dots)\le 2e^{-8}$

ex. (Symmetric random walk with ~~x~~ barriers, "discrete diffusion")

no



Start at 0, At each step move left | right one unit with prob $\frac{1}{2}$. Each way indep of previous moves.

Let X_n be the position after n-th step.

$X_n = Y_1 + \cdots Y_n$, where $Y_k = \begin{cases} 1 & \text{with prob } \frac{1}{2} \\ -1 & \text{with prob } \frac{1}{2} \end{cases}$

$X_n = Y_1 + \cdots Y_n$

$P(X_n \geq a) \leq e^{-a^2/2n}$

$P(X_n \leq -a) \leq e^{-a^2/2n}$

$P(X_n \geq a) \leq e^{-\frac{a^2}{2n}}$, $P(X_n \leq -a) \leq e^{-\frac{a^2}{2n}}$

$a = 4\sqrt{n}$

$P(X_n \geq 4\sqrt{n}) \leq e^{-16n/2n} = e^{-8}$

Choose $a = 4\sqrt{n}$

So $P(X_n \geq 4\sqrt{n}) \leq e^{-8}$, $P(X_n \leq -4\sqrt{n}) \leq e^{-8}$

$P(X_n \leq -4\sqrt{n}) \leq e^{-8}$

So $P(|X_n| \geq 4\sqrt{n}) \leq 2e^{-8} \approx 0.00067$

$P(|X_n| \geq 4\sqrt{n}) \leq 2e^{-8}$

The bulk doesn't go farther away than $4\sqrt{n}$.

St Petersburg paradox

We are ~~offered~~ to play the following game (one round).

We toss a fair coin until it lands tails.

If we use n tosses in total, we get $\$2^n$.

Q. What is a fair entrance fee to play this game.

e.g. T $\to$ get \$2

HT $\to$ \$4

HHT $\to$ \$8

$\vdots$

let X be the payoff, then:

$E(X) = 2 \times \frac{1}{2} + 4 \times \frac{1}{4} + 8 \times \frac{1}{8} + \cdots = +\infty$

Restate the question:

What if we play it a large # n rounds.

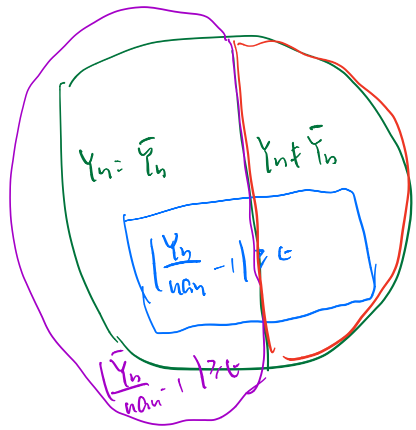
What would a fair fee a_n per round be?

let X_k be the payoff at the k-th round. let $Y_n = X_1 + \cdots X_n \leftarrow$ total pay off when we have n rounds

We would like to find a_n.

So then for any $\epsilon > 0$,

$$P\left(\left|\frac{Y_n}{n a_n} - 1\right| \geq \epsilon\right) \to 0 \text{ as } n \to \infty$$



$x_k = 2^k \leq 2^{a_n}$

$\downarrow$

$k \leq a_n$

Truncate:

$$\bar{X}_k = \begin{cases} X_k, & \text{if } X_k \leq 2^{a_n} \\ 0, & \text{if } X_k > 2^{a_n} \end{cases}$$

let $\bar{Y}_n = \sum_{k=1}^{n} \bar{X}_k$

draw the cloud

$$P\left(\left|\frac{Y_n}{n a_n} - 1\right| \geq \epsilon\right) \leq \underbrace{P\left(\left|\frac{\bar{Y}_n}{n a_n} - 1\right| \geq \epsilon\right)}_{A} + \underbrace{P(Y_n \neq \bar{Y}_n)}_{B}$$

?? LHS is bounded by A or B

We will show if we choose a_n correctly, we can make both terms go 0.

$$E(\bar{X}_k) = \sum_{i=1}^{a_n} 2^i \cdot 2^{-i} = a_n$$

x	2	2^2	2^3	....
tosses	T	HT	HHT	
p(x)	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	

$P(X_k = 2)$

+1: m $\quad m + n = k$

-1: n $\quad m - n = 2$

$m = \frac{k+2}{2}$

$n = \frac{k-2}{2}$

$$P(X_k = 2) = \binom{k}{m}\left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^n = \binom{k}{m}\left(\frac{1}{2}\right)^k = \frac{k!}{m!(k-m)!}\left(\frac{1}{2}\right)^k = \frac{k!}{(\frac{k}{2}+1)!(\frac{k}{2}-1)!}\left(\frac{1}{2}\right)^k$$

$$E(\bar{Y}_n) = n a_n$$

$$\mathrm{Var}(\bar{X}_k) \leq E(\bar{X}_k^2) = \sum_{i=1}^{a_n} 2^{2i} \cdot 2^{-i} = \sum_{i=1}^{a_n} 2^i \leq 2^{a_n + 1}$$

$$\mathrm{Var}(\bar{Y}_n) = \mathrm{Var}(\bar{X}_1) + \cdots + \mathrm{Var}(\bar{X}_n) \leq n \times 2^{a_n+1}$$

So $E\left(\frac{\bar{Y}_n}{n a_n}\right) = 1$, $\mathrm{Var}\left(\frac{\bar{Y}_n}{n a_n}\right) \leq \frac{n \times 2^{a_n+1}}{n^2 \times a_n^2} = \frac{2^{a_n+1}}{n a_n^2}$

Chebyshev Inequality:

$$P\left(\left|\frac{\bar{Y}_n}{n a_n} - 1\right| \geq \epsilon\right) \leq \frac{2^{a_n+1}}{n a_n^2 \epsilon^2}$$

$$P(Y_n \neq \bar{Y}_n) \leq \sum_{k=1}^{n} P(X_k \neq \bar{X}_k)$$

$$\leq n 2^{-a_n}$$

since $Y_n \neq \bar{Y}_n \to \exists\, k$ s.t. $X_k \neq \bar{X}_k$

Goal: Choose a_n s.t. $\frac{2^{a_n+1}}{n a_n^2} \to 0$, and $n 2^{-a_n} \to 0$

A: $a_n = \lfloor \log_2 n \rfloor + \varepsilon_n$, $\varepsilon_n \to \infty$ slower than $\log_2 \log_2 n$