Last time: Toss a fair coin for $n \geq 5$ times. Let X be the # of switches from H to T. $Var(X) = ?$ $X = X_1 + X_2 + \cdots X_{n-1}$

$\cdots H T \cdots H T$

$$X_i = \begin{cases} 1 & \text{if ith H, (i+1)th T} \\ 0 & \text{o.w.} \end{cases}$$

$X = X_1 + X_2 + \cdots + X_{n-1}$

$$X_i = \begin{cases} 1 & \text{if ith H, (i+1)th T} \\ 0 & \end{cases}$$

$$Var(X) = \sum_{i=1}^{n-1} Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j)$$

$$Var(X) = \sum_{i=1}^{n-1} Var(X_i) + \sum_{i \neq j} Cov(X_i, X_j)$$

$E(X_i) = \frac{1}{4}$, $E(X_i^2) = \frac{1}{4}$

$\rightarrow Var(X_i) = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$

H T _

1,2 2,1 ... n-1,n-2

2,3 ...

$Cov(X_i, X_j)$ $2(n-1)-2 = 2n-4$

if $|i-j| \geq 2$, the $Cov(X_i, X_j) = 0$

If $|i-j| > 1$, $HT \cdots HT$, then $Cov(X_i, X_j) = 0$. b/c of X_i and X_j are indep when $|i-j|>1$

$Cov(X_i, X_{i+1}) = E(X_i X_{i+1}) - E(X_i)E(X_{i+1}) = 0 - \frac{1}{4} \times \frac{1}{4} = \frac{-1}{16}$

if $|i-j| = 1$,

$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$

$= 0 - \frac{1}{16}$

$= \frac{-1}{16}$

$P(X_i = 1) = \frac{1}{4}$, then we have:

$Var(X_i) = E(X_i^2) - E^2(X_i)$

$= E(X_i) - E^2(X_i)$

$= \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$

and we have: $(2n-4)$ legal pairs

So $\sum_{i \neq j} Cov(X_i, X_j) = \frac{-1}{16}(2n-4)$

Hence $Var(X) = (n-1) \times \frac{3}{16} - \frac{1}{16}(2n-4)$

$= \frac{1}{16}n + \frac{1}{16}$

$= \frac{1}{16}(n+1)$

$2(n-2)$ non-zero covariances. Treat X_i, X_{i+1} and X_{i+1}, X_i as diff pairs.

$$Var(X) = (n-1) \times \frac{3}{16} - 2(n-2) \times \frac{1}{16} = \frac{n+1}{16}$$

Variance from the Generating Functions

If X takes values $0, 1, 2, \cdots$

$$G_X(s) = \sum_{k=0}^{\infty} P(X=k) \cdot s^k$$

$$G_X'(s) = \sum_{k=0}^{\infty} P(X=k) \cdot k s^{k-1}$$

$$G_X''(s) = \sum_{k=0}^{\infty} P(X=k)\, k(k-1)\, s^{k-2}$$

$$G_X(s) = \sum_{k=0}^{\infty} P(X=k) \cdot s^k$$

$$G_X'(s) = \sum_{k=0}^{\infty} P(X=k) \cdot k s^{k-1}$$

$$G_X'(1) = \sum_{k=0}^{\infty} P(X=k) \cdot k = E(X)$$

$$G_X''(s) = \sum_{k=0}^{\infty} P(X=k) \cdot k(k-1) s^{k-2}$$

$$G_X''(1) = \sum_{k=0}^{\infty} P(X=k) \cdot (k^2 - k)$$

$$G_X'(1) = \sum_{k=0}^{\infty} P(X=k)\cdot k = E(X)$$

$$G_X''(1) = \sum_{k=0}^{\infty} P(X=k)\, k(k-1)$$

$$= \sum_{k=0}^{\infty} k^2 P(X=k) - \sum_{k=0}^{\infty} k P(X=k)$$

$$= E(X^2) - E(X)$$

legal, if the radius of convergence >1

$$= E(X^2) - E(X)$$

$$E(X^2) = G_X''(1) + E(X) = G_X''(1) + G_X'(1)$$

$$Var(X) = E(X^2) - E^2(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2$$

$$Var(X) = E(X^2) - E^2(X)$$

$$= G_X''(1) + G_X'(1) - (G_X'(1))^2$$

$$Var(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2$$

$$G_X(s) = \sum_{k=0}^{\infty} P(X=k)\cdot s^k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} s^k$$

$$= \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} e^{-\lambda}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!}$$

$$= e^{-\lambda}\cdot e^{\lambda s} = e^{\lambda(s-1)}$$

ex. X is Poisson with $E(X) = \lambda$. $Var(X) = ?$

$G_X(s) = e^{\lambda(s-1)}$ ← compute some time ago

$G_X'(s) = \lambda e^{\lambda(s-1)}$, $G_X''(s) = \lambda^2 e^{\lambda(s-1)}$

$$Var(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

$$G_X'(s) = e^{\lambda(s-1)}\lambda$$

$$G_X''(s) = e^{\lambda(s-1)}\lambda^2$$

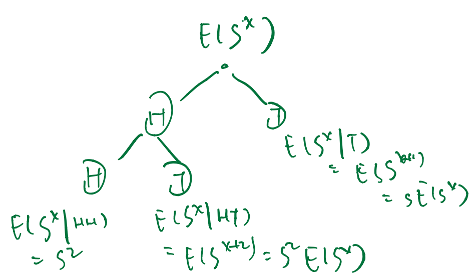
$$G_X'(1) = \lambda,\quad G_X''(1) = \lambda^2$$

$$\to Var(X) = \lambda^2 + \lambda - \lambda^2 = \underline{\lambda}$$

ex. Toss a fair coin till HH appears.

Let X be the # of tosses.

Q: $Var(X) = ?$



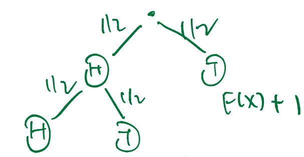
In the HW, we compute $G_X(s) = \dfrac{s^2}{4-2s-s^2} = s^2(4-2s-s^2)^{-1}$

$G_X'(s) = 2s(4-2s-s^2)^{-1} + (2+2s)s^2(4-2s-s^2)^{-2}$

$G_X''(s) = 2(4-2s-s^2)^{-1} + 2s(2+2s)(4-2s-s^2)^{-2} + 2s(2+2s)(4-2s-s^2)^{-2}$
$+ 2s^2(4-2s-s^2)^{-2} + 2s^2(2+2s)^2(4-2s-s^2)^{-3}$

$G_X'(1) = 2+4 = 6$, $G_X''(1) = 2+8+8+2+32 = 52$

So $\text{Var}(X) = 56 + 6 - 6^2 = 22$



$E(X|HH) = 2$ $E(X|HT) = \cancel{E(X)+2} = E(X)+2$

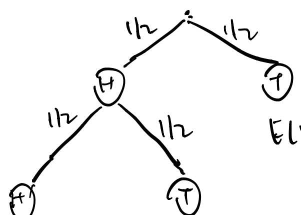
$\frac{1}{4}\times 2 + \frac{1}{4}E(X) + \frac{1}{2} + \frac{1}{2}E(X) + \frac{1}{2} = E(X)$

$\frac{1}{4}E(X) = \frac{3}{2}$

$E(X) = 6$

By conditioning,

First $E(X)$



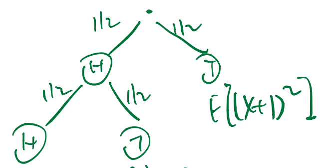
$E(X|T) = E(X)+1$

$E(X|HH) = 2$ $E(X|HT) = E(X)+2$

So $E(x) = \frac{1}{4}\times 2 + \frac{1}{4}(E(X)+2) + \frac{1}{2}(E(X)+1)$

$= \frac{3}{4}E(X) + \frac{3}{2}$

$\to E(X) = 6$



$E(X^2|HH) = 4$ $E(X^2|HT) = E[(X+2)^2]$ $E[(X+1)^2]$

$E(X^2) = \frac{1}{4}\times 4 + \frac{1}{4}E(X+2)^2 + \frac{1}{2}E(X+1)^2$

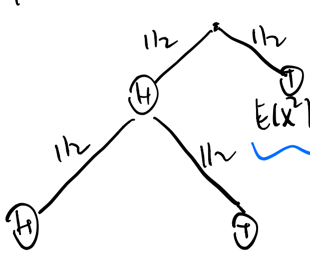
$E(X^2) = 1 + \frac{1}{4}E(X^2) + 1 + E(X) + \frac{1}{2}E(X^2) + \frac{1}{2} + E(X)$

$\frac{1}{4}E(X^2) = \frac{5}{2} + \underbrace{2E(X)}_{12} = \frac{29}{2}$

$E(X^2) = 58$

$\text{Var}(X) = E(X^2) - E^2(X) = 58 - 6^2 = \boxed{22}$

Then $E(X^2)$



$E(X^2|T) = E(X+1)^2$

$E(X^2|HH) = 4$ $E(X^2|HT) = E(X+2)^2$

So $E(X^2) = \frac{1}{4}\times 4 + \frac{1}{4}E(X+2)^2 + \frac{1}{2}E(X+1)^2$

$$= \frac{1}{2} E(X^2+2X+1) + 1 + \frac{1}{4} E(X^2+4X+4)$$

$$= \frac{3}{4} E(X^2) + 2\underbrace{E(X)}_{6} + \frac{5}{2}$$

$$\rightarrow \frac{1}{4} E(X^2) = 12 + \frac{5}{2} = \frac{29}{2}$$

$$E(X^2) = \frac{29}{2} \times 4 = 58$$

Hence $Var(X) = E(X^2) - E^2(X) = 58 - 36 = 22$

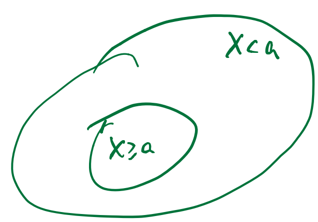
$$P\{X \geq a\} \leq \frac{E(X)}{a}$$

Concentration

Markov Inequality:

Let $X: \Omega \to \mathbb{R}$ be a r.v. Such that $P\{w : X(w) \geq 0\} = 1$, and $E(X)$ exists.

Then: For $a > 0$, $P\{w : X(w) \geq a\} \leq \frac{E(X)}{a}$



Why? Define $Y(w) = \begin{cases} a, & \text{if } X(w) \geq a \\ 0, & \text{if } X(w) < a \end{cases}$

$$\left.\begin{array}{l} \text{Then } Y \leq X \Rightarrow E(Y) \leq E(X) \\ E(Y) = a \times P(X \geq a) \end{array}\right\} \Rightarrow a P(X \geq a) \leq E(X)$$

$$\rightarrow P(X \geq a) \leq \frac{E(X)}{a}$$

$$P\{w : X(w) \geq a\} \leq \frac{E(X)}{a}$$

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Chebyshev Inequality $P(|X(\omega)-E(x)|\geq a)\leq \frac{Var(X)}{a^2}$

let $X:\Omega\to\mathbb{R}$ be a r.v. s.t. $E(x)$ and $Var(x)$ exist.

Then $P\{\omega: |X(\omega)-E(x)|\geq a\}\leq \frac{Var(X)}{a^2}$

$P\{\omega: |X(\omega)-E(x)|\geq a\}\leq \frac{Var(X)}{a^2}$

Here is why:

Define $Z=(X-E(x))^2$, then $Z\geq 0$.

$P\{\omega: |X(\omega)-E(x)|\geq a\}\leq \frac{Var(X)}{a^2}$

$E(Z)=E(X-E(x))^2=Var(X)$

$\underbrace{np(1-p)}$

and $|X-E(x)|\geq a \Rightarrow Z\geq a^2$

$E(X)=20\times\frac{1}{5}=4$

By Markov,

$P(X\geq 16)=P(X\geq 16)+P(X\leq -8)$

$$P(Z\geq a^2)\leq \frac{E(Z)}{a^2}=\frac{Var(X)}{a^2}$$

$=P(|X-4|\geq 12)$

$\leq \frac{Var(X)}{12^2}=\frac{20\times 0.2\times 0.8}{12^2}$

ex (Very Typical)

Roll a fair die 10,000 times,

let X be the total # of $\underbrace{\text{points}}_{\text{the \# we get}}$.

$X=X_1+X_2+\cdots X_{10,000}$, where X_k is the # shown on the k-th roll.

$E(x)=E(X_1)+\cdots E(X_{10,000})$

$P(|X-35,000|\geq 1000)\leq \frac{Var(X)}{1000^2}$

$E(X_k)=\frac{1+2+3+4+5+6}{6}=\frac{21}{6}=\frac{7}{2}$

$Var(X)=\sum_{i=1}^{10,000} Var(X_i)=10,000\times\frac{35}{12}=\frac{350,000}{12}$

So $E(X)=\frac{7}{2}\times 10,000=35,000$

$\therefore P(\cdots)\leq \frac{350,000/12}{\cancel{1,000,000}}=\frac{3.5/12}{10}$

$=\frac{3.5}{120}$

Q: Using Chebyshev Inequality to bound the prob that:

$P(|X-35,000|\geq 1000)$

$$\mathrm{Var}(X) = \sum_{i=1}^{10{,}000} \mathrm{Var}(X_i)$$

$X_1, \dots, X_{10{,}000}$ are indep

$$P(|X-E(X)| \geq a) \leq \frac{\mathrm{Var}(x)}{a^2}$$

$$\mathrm{Var}(X_k) = E(X_k^2) - E(X_k)^2$$

$$= \frac{1^2+2^2+3^2+4^2+5^2+6^2}{6} - \left(\frac{1+2+3+4+5+6}{6}\right)^2$$

$$= \frac{35}{12}$$

So $\mathrm{Var}(x) = 10{,}000 \times \frac{35}{12} = \frac{350{,}000}{12}$

$$P(|X - 35{,}000| \geq 1000) \leq \frac{\mathrm{Var}(X)}{(1{,}000)^2} = \frac{350{,}000/12}{1{,}000{,}000} = \frac{35}{1200} = \frac{7}{240} \approx 0.029$$

$|X_k| \leq 1$, $E(X_k) = 0$ Show that $\frac{e^t + e^{-t}}{2} \leq e^{\frac{t^2}{2}}$

Bernstein Inequality: $X = X_1 + \cdots + X_n$ $P(X \geq a) \leq e^{-a^2/2n}$ $P(X \leq -a) \leq e^{-a^2/2n}$

let $X_1, \dots X_n$ be indep r.v. s.t.:

① $|X_k| \leq 1$ for $k = 1, 2, \dots n$

② $E(X_k) = 0$ for $k = 1, 2, \dots n$

expect should be shifted to 0.

Then for $X = X_1 + \cdots X_n$ and $a > 0$,

$$\left.\begin{aligned} &P(X \geq a) \leq e^{-a^2/2n} \\ &P(X \leq -a) \leq e^{-a^2/2n} \end{aligned}\right\} \text{Strong Inequalities}$$

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

We'll prove: $P(X \geq a) \leq e^{-a^2/2n}$

$P(X \leq -a) \leq e^{-a^2/2n}$ follows from that by replacing X_k with $-X_k$.

Pick a $t > 0$ (to specify later)

$P(X \geq a) = P(tX \geq ta) = P(e^{tX} \geq e^{ta})$,

Use Taylor expansion to expand e^t:

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}, \quad e^{-t} = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!}$$

So $$\frac{e^t + e^{-t}}{2} = \sum_{n=0}^{\infty} \frac{t^n + (-t)^n}{2(n!)} = \sum_{n=0}^{\infty} \frac{2t^{2n}}{2 \times (2n)!} = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!}$$

$$e^{\frac{t^2}{2}} = \sum_{n=0}^{\infty} \frac{t^{2n}}{2^n \cdot n!}$$

Since when $n \geq 1$, $\frac{(2n)!}{2^n \cdot n!} = \frac{\cancel{n!}\,(n+1) \times \cdots (2n)}{\cancel{n!} \times 2^n} \geq \frac{2^n}{2^n} = 1$

when $n = 0$, $2^n \cdot n! = (2n)! = 1$, so $(2n)! \geq 2^n \cdot n!$

Hence $\frac{e^t + e^{-t}}{2} \leq e^{\frac{t^2}{2}}$.

$|X_k| \le 1$ $\quad X = X_1 + \cdots X_n$

$E\, X_k = 0$ $\quad a \ge 0$

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

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

$$\underset{\substack{\uparrow \\ \text{Markov}}}{\le} e^{-ta} E(e^{tX}) = e^{-ta} \left(\frac{e^t + e^{-t}}{2}\right)^n \le e^{-ta} \left(e^{\frac{t^2}{2}}\right)^n = e^{-ta} e^{\frac{nt^2}{2}}$$

Compare Taylor Series

Choose $t = a/n$,

$$\le e^{-a^2/n} \cdot e^{a^2/2n} = e^{-a^2/2n}$$

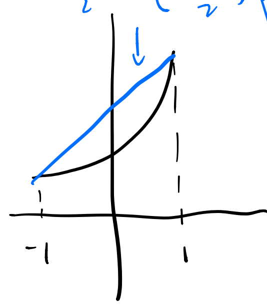
$$E(e^{tX}) = E(e^{tX_1 + \cdots + tX_n})$$

$$= E(e^{tX_1} \cdots e^{tX_n})$$

independence

$$\overset{\Downarrow}{=} E(e^{tX_1}) \cdots E(e^{tX_n})$$

the ftn: $x \mapsto e^{tx}$ is convex.

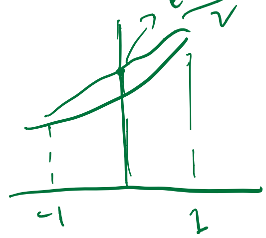


If $-1 \le x \le 1$:

then $e^{tx} \le \frac{e^t + e^{-t}}{2} + \left(\frac{e^t - e^{-t}}{2}\right) x$

$$\text{So } E(e^{tX_k}) \underset{\substack{\uparrow \\ \text{use that } -1 \le X_k \le 1}}{\le} E\left(\frac{e^t + e^{-t}}{2} + \frac{e^t - e^{-t}}{2} X_k\right)$$

$$= \frac{e^t + e^{-t}}{2} + \frac{e^t - e^{-t}}{2} E(X_k) = \frac{e^t + e^{-t}}{2}$$



$t = a/n$, $\quad e^{-ta} \le e^{-a^2/n} \times e^{a^2/2n}$

$= e^{-a^2/2n}$

$X = X_1 + \cdots X_n$

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

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

$$\overline{a > 0}$$

$$P(X \ge a) \le \frac{E(X)}{a}$$

$$P(e^{tX} \ge e^{ta}) \le \frac{E(e^{tX})}{e^{ta}}$$

$$= e^{-ta} E(e^{tX})$$

$$E(e^{tX}) = E(e^{t(X_1 + \cdots X_n)})$$

$$= E(e^{tX_1} \times e^{tX_2} \times \cdots e^{tX_n})$$

$$= E(e^{tX_1}) \times \cdots E(e^{tX_n})$$

ftn: $x \mapsto e^{tx}$ is convex

$$e^{tx} \le \frac{e^t + e^{-t}}{2} + \left(\frac{e^t - e^{-t}}{2}\right) x$$

$$E(e^{tX_k}) \le E\left(\frac{e^t + e^{-t}}{2} + \frac{e^t - e^{-t}}{2} X_k\right)$$

$$= \frac{e^t + e^{-t}}{2} + \frac{e^t - e^{-t}}{2} E(X_k) = \frac{e^t + e^{-t}}{2}$$