Markov's inequality: For $X \geq 0$ (always takes non-neg. values)

for any $a \geq 0$,

$$\Pr(X \geq a) \leq \frac{\mu}{a}$$

where $\mu = E[X]$.

Chebyshev's: For r.v. X with $\mu = E[X]$ & $\sigma^2 = Var(X)$,

For any $k > 0$,

$$\Pr(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

or $$\Pr(|X-\mu| \geq r) \leq \frac{Var(X)}{r^2}$$

Proof:

Let $Y = (X-\mu)^2$ & $a = (k\sigma)^2$ (Note $Y \geq 0$)

Then, $$\Pr(|X-\mu| \geq k\sigma) = \Pr(Y \geq a^2) \leq \frac{E[Y]}{a^2}$$

$$= \frac{E[(X-\mu)^2]}{a^2} = \frac{Var(X)}{a^2}$$

$\blacksquare$

Example: Suppose $X_i = \begin{cases} 1 & \text{with prob. } 1/2 \\ 0 & \text{"} \end{cases}$

Let $X = \sum_{i=1}^{n} X_i = \text{Binomial}(n, 1/2)$

Then, $E[X] = \frac{n}{2}$, $Var(X) = \frac{n}{4}$, $\sigma = \text{stddev} = \frac{\sqrt{n}}{2}$

From Chebyshev's, $Pr(|X - \frac{n}{2}| \geq 5\sqrt{n})$

$$= Pr(|X - E[X]| \geq 10\sigma) \leq \frac{1}{10^2} = \frac{1}{100}$$

Hence, $Pr(\frac{n}{2} - 10\sigma \leq X \leq \frac{n}{2} + 10\sigma) \geq .99$

~~but $Bin(n, \frac{1}{2})$ is within a few~~

What about bigger deviations? $\gg 10\sigma$?

Say $X \sim Bin(1000, 1/2)$

Then, by Chebyshev's $Pr(X \geq 750) = Pr(|X - 500| \geq 250)$

$$\leq \frac{250}{250^2} = \frac{1}{250} = .004$$

but in fact: $Pr(X \geq 750) = \sum_{v=750}^{1000} \binom{1000}{v} 2^{-1000} \approx .67 \times 10^{-58}$

Can we get a better bound than using that X is the sum of indpt random variables.

Chernoff: For $X \sim Bin(n, \frac{1}{2})$, for $0 \le t \le \sqrt{n}$,

$$Pr\left(X \ge \frac{n}{2} + t\frac{\sqrt{n}}{2}\right) \le e^{-t^2/2}$$

$$Pr\left(X \le \frac{n}{2} - t\frac{\sqrt{n}}{2}\right) \le e^{-t^2/2}$$

First proof sketch based on Ryan O'Donnell's nice notes

Then, $Pr\left(Bin(1000, \frac{1}{2}) \ge 750\right)$

$$= Pr\left(X \ge 500 + t\frac{\sqrt{1000}}{2}\right) \le e^{-\left(\frac{\sqrt{1000}}{2}\right)^2/2} = e^{\frac{-1000}{8}} = .52 \times 10^{-54}$$

for $t = \frac{\sqrt{1000}}{2}$ so $\frac{t\sqrt{1000}}{2} = 250$

Let's do a change of variables so that the mean is 0.

Let $Y_i = -1 + 2X_i$

thus, $Y_i = \begin{cases} 1 & \text{w.p. } 1/2 \\ -1 & \text{"} \end{cases}$

& let $Y = \sum_{i=1}^{n} Y_i = -n + 2X$

Then, $E[Y]=0$

& $X \geq \frac{n}{2} + \frac{t\sqrt{n}}{2} \iff -n+2X \geq -n+2\left(\frac{n}{2}+\frac{t\sqrt{n}}{2}\right)$

$Y \geq \frac{t\sqrt{n}}{2}$

$Var(Y) = n, \ \sigma = \sqrt{n}$

Think of $Y_1 + \cdots + Y_{i-1}$ & adding Y_i
with prob. $\frac{1}{2}$ go $+1$ & prob. $\frac{1}{2}$ go -1
So it's an unbiased random walk.

We're trying to bound prob. $|Y|$ is large
but can't use Markov's since not nonnegative.

Instead of adding ± 1 at each step,
let's multply or divide by $1+\lambda$ at each step
for some tiny λ.

Note, $(1+\lambda_1)\times(1+\lambda_2) = 1+\lambda_1+\lambda_2+\lambda_1\lambda_2$

$\approx 1+\lambda_1+\lambda_2$ for tiny ~~&~~ λ_1, λ_2

or $(1+\lambda_1)\times(1+\lambda_2) \approx e^{\lambda_1+\lambda_2} \approx 1+\lambda_1+\lambda_2$

So, Position $(1+\lambda)^U$ in new walk

$\approx$ position U in old walk

Set $\lambda = \frac{1}{\sqrt{n}}$

then $\left(1+\frac{1}{\sqrt{n}}\right)^{Y_1}\left(1+\frac{1}{\sqrt{n}}\right)^{Y_2}\cdots\left(1+\frac{1}{\sqrt{n}}\right)^{Y_n} = \left(1+\frac{1}{\sqrt{n}}\right)^{Y}$

if $Y=\sqrt{n}$, then $\left(1+\frac{1}{\sqrt{n}}\right)^{\sqrt{n}} \approx e$

but if $Y=5\sqrt{n}$ then $\left(1+\frac{1}{\sqrt{n}}\right)^{5\sqrt{n}} \approx e^{5} \approx 150$

& if $Y=100\sqrt{n}$ then $\left(1+\frac{1}{\sqrt{n}}\right)^{100\sqrt{n}} \approx e^{100}$

which is large so we can use Markov's ineq.

Let $Z_i = (1+\lambda)^{Y_i}$ & we'll choose λ later

$$Z_i = \begin{cases} 1+\lambda & \text{w.p. } 1/2 \\ \frac{1}{1+\lambda} & \text{"} \end{cases}$$

Let $Z = Z_1 \times Z_2 \times \cdots \times Z_n = (1+\lambda)^Y$

& since the Y_i's are indp't. so are the Z_i's.

Note, $Z \geq 0$ so Markov's ineq. applies.

$$\& \; E[Z] = E[Z_1 \times Z_2 \times \cdots \times Z_n]$$
$$= E[Z_1] \times E[Z_2] \times \cdots \times E[Z_n]$$

Since ___ are indp't.

$$E[Z_i] = \frac{1}{2}(1+\lambda) + \frac{1}{2}\left(\frac{1}{1+\lambda}\right) \cancel{\leq 1+}$$
$$= \frac{1}{2}\left(\frac{1+\lambda+\lambda+\lambda^2+1}{1+\lambda}\right) = \frac{1}{2}\left(2+\frac{\lambda^2}{1+\lambda}\right) = 1+\frac{\lambda^2}{2+2\lambda}$$
$$\leq 1+\lambda^2/2$$

Hence, $E[Z] \leq \left(1+\frac{\lambda^2}{2}\right)^n$

$$X \geq \frac{n}{2} + t\sqrt{n} \iff Y \geq t\sqrt{n}$$

$$\iff (1+\lambda)^{Y} \geq (1+\lambda)^{t\sqrt{n}}$$

$$\iff Z \geq (1+\lambda)^{t\sqrt{n}}$$

$$\Pr\left(X \geq \frac{n}{2} + t\sqrt{n}\right) = \Pr\left(Z \geq (1+\lambda)^{t\sqrt{n}}\right)$$

$$\leq \frac{E[Z]}{(1+\lambda)^{t\sqrt{n}}}$$

$$\leq \frac{\left(1+\lambda^2/2\right)^n}{(1+\lambda)^{t\sqrt{n}}}$$

Fix $\lambda = \frac{t}{\sqrt{n}}$

Then:

$$\Pr\left(X \geq \frac{n}{2} + t\sqrt{n}\right) \leq \frac{\left(1+\frac{t^2}{2n}\right)^n}{\left(1+\frac{t}{\sqrt{n}}\right)^{t\sqrt{n}}}$$

not correct!

$$\leq \frac{\left(e^{t^2/2n}\right)^n}{\cdots} = \frac{e^{t^2/2}}{e^{t^2}} = e^{-t^2/2}$$

$\square$

To formalize, use

$$E[Z_i] = 1 + \frac{\lambda^2}{2+2\lambda}$$

instead of

$$E[Z_i] \leq 1 + \frac{\lambda^2}{2}$$

& set $\lambda = e^{t/\sqrt{n}} - 1$

instead of $\lambda = \frac{t}{\sqrt{n}}$.

General form of Chernoff bounds: [Bernstein '24] [Chernoff '52]

Let $X_1, \ldots, X_n$ be independent random variables where $0 \leq X_i \leq 1$.

Let $X = \sum_{i=1}^{n} X_i$ & $\mu = E[X]$

For all $0 < \epsilon \leq 1$,

$$\Pr(X \geq (1+\epsilon)\mu) \leq e^{-\left(\frac{\epsilon^2}{3}\right)\mu}$$

$$\Pr(X \leq (1-\epsilon)\mu) \leq e^{-\left(\frac{\epsilon^2}{2}\right)\mu}$$

$$\text{or } \Pr(|X-\mu| \geq \epsilon\mu) \leq 2e^{-\mu\epsilon^2/3}$$

Proof:

$$\Pr(X \geq (1+\epsilon)\mu) = \Pr(e^X \geq e^{(1+\epsilon)\mu})$$

$$= \Pr(e^{tX} \geq e^{t(1+\epsilon)\mu})$$

$$\leq \frac{E[e^{tX}]}{e^{t(1+\epsilon)\mu}}$$

Suppose $X_i = \text{Bernoulli}(p_i)$: $Pr(X_i = 1) = p_i$

$Pr(X_i = 0) = 1 - p_i$

Then,

$$M_{X_i}(t) = E\left[e^{tX_i}\right] = p_i e^t + (1 - p_i)$$

$$= 1 + p_i(e^t - 1)$$

$$\leq e^{p_i(e^t - 1)} \quad \text{since } 1 + y \leq e^y$$

$$\& \; M_X(t) = \prod_{i=1}^{n} M_{X_i}(t) \quad \text{since } X_i\text{'s are indep't.}$$

$$\leq \prod_{i=1}^{n} e^{p_i(e^t - 1)}$$

$$= e^{\sum_{i=1}^{n} p_i(e^t - 1)}$$

$$= e^{\mu(e^t - 1)}$$

$$\Pr(X \geq (1+\epsilon)\mu) \leq \frac{E[e^{tX}]}{e^{t(1+\epsilon)\mu}}$$

$$\leq \left(\frac{e^{e^t - 1}}{e^{t(1+\epsilon)}} \right)^{\mu}$$

Set $t = \ln(1+\epsilon)$

$$\leq \left(\frac{e^{\epsilon}}{(1+\epsilon)^{(1+\epsilon)}} \right)^{\mu} \leq e^{-(\epsilon^2/3)\mu}$$

Calculus:

we have $e^{\epsilon - (1+\epsilon)\ln(1+\epsilon)} \leq e^{\epsilon - \epsilon \mp \epsilon^2/3} = e^{-\epsilon^2/3}$

note: $\ln(1+\epsilon) = \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3} \pm \cdots$

$$(1+\epsilon)\ln(1+\epsilon) = \epsilon - \frac{\epsilon^2}{2} + \epsilon^2 + \frac{\epsilon^3}{3} - \frac{\epsilon^3}{2} \pm \cdots$$

$$\geq \epsilon + \frac{\epsilon^2}{2} - \frac{\epsilon^3}{6}$$

$$\geq \epsilon + \frac{\epsilon^2}{2} - \frac{\epsilon^2}{6} = \epsilon + \frac{\epsilon^2}{3}$$