Moments and deviation

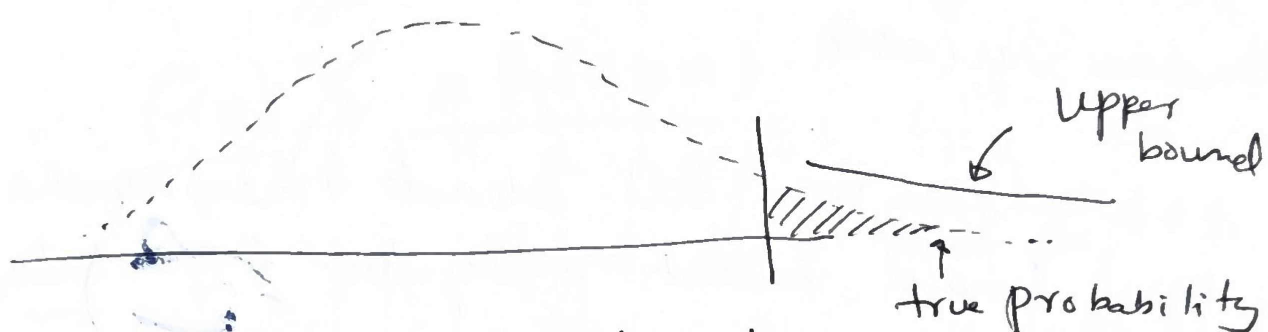
1. Randomised Algorithm / Randomised proposal for Compression needs to be analysed probabilistically.

Thus e.g. $X =$ Running time of an Algorithm.

$X =$ fraction of failures

there need to be a way to understand how big these can go.

Thus tail probability.



$$\Pr(X \geq a) \leq \underbrace{\text{Some bound}}_{\text{Some ~~estimate~~ upper bound}}$$

If we know distribution of X and it is easy to analyse

then $\cancel{\Phi^{-1}}$ $\underline{1 - F(a)}$

Otherwise: Tail bounds

Example:

$$X = X_1 + X_2 + \dots X_n$$

$X_1 \downarrow$ Coin toss with P_1

$Ber(P_1)$ $Ber(P_2)$ ...

Q: What is the probability that #heads $\geq \frac{3n}{4}$.

$P_i = P \; \forall i \Rightarrow$ Binomial distribution.

if not $\Rightarrow$ even more difficult to manipulate

In such a case, there are some interesting things that we can measure about a random variable X

e.g.

$$E(X) = \sum p_i$$

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

$$= \sum p_i(1-p_i) + \left(\sum p_i\right)^2$$

$\vdots$

So on:

Q: Moments are a natural ~~...~~ analysis of the random variable.

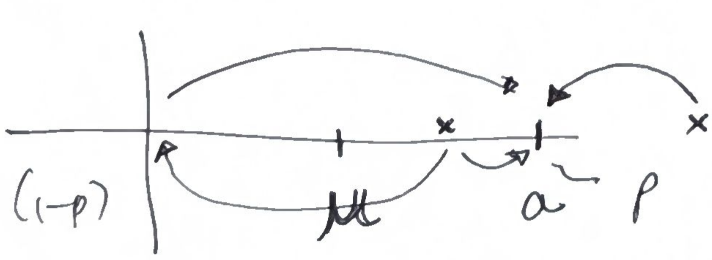
Q: What can you tell about tail bounds if you know about only the expectation $E(X)$

Markov Inequality:

$$P(X \geq a) \leq \frac{E(X)}{a} \quad \text{if } X \text{ takes only } > 0 \text{ values}$$

How did Markov come up with this inequality.

* $\mu = E(X)$ $\quad$ $X > 0$ ↑ necessary $\quad$ || Push prob mass indefinitely



Anything in $[0, a)$ push to "0"

$[a, \infty)$ push to a

X for max prob mass @ a is Bernoulli

$\geq$

$ap = \mu$ or $p = \frac{\mu}{a}$

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

arkov Inequality

$$Pr(X \geq a) \leq \frac{E(X)}{a}$$ if X takes all ~~positive~~ non-neg. value

Proof:→

$$E(X) = \underbrace{\sum_{x} x \, Pr(x)}$$

$$\underbrace{\sum_{x<a}}_{\geq 0} \qquad \sum_{x \geq a} x \, Pr(X)$$

$$\geq a \, Pr(X \geq a)$$

$$\frac{E(x)}{a} \geq \frac{a \, Pr(x \geq a)}{a}$$

Is it tight? Yes. What does it mean for it to be tight? $\exists X$

Ex. $X = X_1 + X_2 + \ldots X_n$ | Bin Unbiased

$E(X) = \Sigma P_i$ | $nP. = \frac{n}{2}$

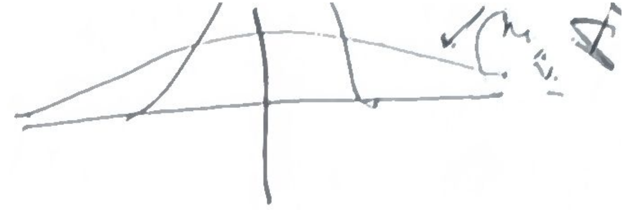
$$Pr\left(X \geq \frac{3n}{4}\right) \leq \frac{n/2}{\frac{3n}{4}}$$

$$Pr\left(X \geq \frac{3n}{4}\right) \leq \frac{2}{3}$$

What if we know higher moments?

$$\text{Variance}(X) = E((X-\mu)^2) = E(X^2) - E^2(X)$$

k^{th} Moment: $E(X^k)$



$$\text{Var}(X) = E(X^2) - E^2(X)$$

Chebyshev:

$$\Pr(|X-\mu| \geq a) \leq$$

$$= \Pr((X-\mu)^2 \geq a^2) \leq \frac{E((X-\mu)^2)}{a^2} = \frac{\text{Var}(X)}{a^2}$$

Proof? Markov.

Tight? Yes. Is there a X s.t equality holds

$$Y = (X-\mu)^2 \leftarrow \text{Markov in } Y$$

$$Y = 0 \quad \cancel{p}\ (1-p)$$

$$a^2 \quad \cancel{(1-p)}(p)$$

$$X - \mu = 0 \ (1-p)$$

$$= -a \ (p/2)$$

$$+a \ (p/2)$$

} distaqual

$E((X-\mu)^2)$

$= a^2 \cdot p$

$$X = \mu \quad (1-\text{Var})$$

$$\mu - a \quad \text{Var}/2$$

$$\mu + a \quad \text{Var}/2.$$

$E(X) = \mu = 0$

$$E(X^2) = \mu^2(1-v) + (\mu-a)^2\left(\frac{v}{2}\right) + (\mu+a)^2\frac{v}{2}$$

$$a^2\frac{v}{2} + a^2\frac{v}{2} = a^2 v$$

Apply to our binomial example

$X = X_1 + X_2 + \dots X_n \qquad X_i \in Burn(p)$

$E(X) = np \quad = n/2$

$Var(X) = np(1-p) = n/4$

$$Pr\left(\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right) \leq \frac{n/4}{n^2/16} = \frac{4}{n}$$

$$p\left(X - \frac{n}{2} \geq \frac{n}{4}\right)$$

$$p\left(X \geq \frac{3n}{4}\right) \leq \frac{4}{n}\cdot\frac{1}{2} = \frac{2}{n}$$

What if we knew more moments.

$$P(|X-\mu| \geq a) \leq \frac{E((X-\mu)^{2k})}{a^{2k}}$$

higher "k" you use better the bound.

Chernoff:

Moment Generating function:

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

We are mainly interested about existence & properties of $M_X(t)$ around $t=0$

* MGF: → used to generate Moment

$$M_X^{(n)}(t) = E(X^n e^{tX})$$

$$M_X^{(n)}(0) = E(X^n)$$

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

$$= \sum_i e^{tx_i} Pr(x_i)$$

$$M'_X(t) = \sum_i x_i e^{t} Pr(x_i)$$

$$= E(x_i e^{t})$$

Mild condition under which we can interchange E and $\frac{d}{dt}$

MGF: captures all the moments.

---

* If you know all the Moments, it uniquely defines the distribution of X

MGF uniquely defines X

If $M_X(t) = M_Y(t)$ for $\forall t \in (-\delta, \delta)$ for some $\delta > 0$ then

X, Y have same distribution.

→ All moments exist and finite $\not\Rightarrow$ ~~they~~ $X \stackrel{d}{=} Y$ and match for X & Y

$MGF_X(t) = MGF_Y(t)$ in $(-\delta, \delta)$ for some δ

then $X \stackrel{d}{=} Y$

Th: If X and Y are independed then

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

Deriving Chernoff bound

$$Pr(X \geq a)$$

$$= Pr(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}}$$

if $t > 0$

true $\forall t$

$$\leq \min_{t>0} \frac{E[e^{tX}]}{e^{ta}}$$

$t < 0$

$$Pr(X \leq a)$$

$$Pr(e^{tX} \geq e^{ta}) \leq \underbrace{\frac{E[e^{tX}]}{e^{ta}}}_{\substack{\min \\ t<0}}$$

R.H.S : $\dfrac{MGF_X(t)}{e^{ta}}$

Apply chernoff to poisson trials. (5)

$$X = X_1 + X_2 + \dots X_n \qquad X_i < \begin{matrix} 0 & (1-P_i) \\ 1 & P_i \end{matrix}$$

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= \Pi\, E(e^{tX_i}) \\ &= \Pi \left[e^t \cdot P_i + (1-P_i) \right] \\ &= \Pi \left[1 + P_i(e^t - 1) \right] \\ &\leq \Pi \left[e^{(e^t-1)P_i} \right] \\ &= e^{(e^t-1)\mu} \end{aligned} \qquad \mu = \cancel{n} \sum P_i$$

Theorem: $X_1, X_2 \dots X_n$ are poisson trials.

$Pr(X_i) = P_i \quad \mu = \sum P_i$

then

$\delta > 0 \qquad Pr(X \geq (1+\delta)\mu) < \left(\dfrac{e^\delta}{(1+\delta)^{(1+\delta)}} \right)^\mu$

$0 < \delta < 1 \qquad Pr(X \geq (1+\delta)\mu) \leq e^{-\mu\delta^2/3}$

$R \geq 6\mu \qquad Pr(X \geq R) \leq 2^{-R}$

Proof $\Pr(X \geq (1+\delta)\mu) \leq \dfrac{E[e^{tX}]}{e^{(1+\delta)\mu}}$

$$\leq \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}$$

$\underline{\delta > 0}$

$t = \ln(1+\delta)$

$$\leq \frac{e^{\delta\mu}}{e^{t(1+\delta)\mu}}$$

t^*

$$= \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

$0 < \delta < 1$ $\quad \dfrac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \leq e^{-\delta^2/3}$

You can verify it

Similarly $\quad R \geq 6\mu \quad | \; R = (1+\delta)\mu$

$\delta \geq 5$

$$\left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$$

$$= \left(\frac{e}{6}\right)^{R}$$

$$\leq 2^{-R}$$

Binomial :

$$\Pr\left(X \geq \frac{3n}{4}\right)$$

$$\left(1+\frac{1}{2}\right)\mu$$

μ: $\frac{n}{2}$.

$\delta = 0.5$

(we did a lot of approx)

$$\Pr\left(X \geq \frac{3n}{4}\right) \leq \underbrace{e^{-\mu\delta^2/3}}$$

$$e^{-\mu/12} \qquad \mu = \frac{n}{2}$$

$$\boxed{e^{-n/24}.} \leftarrow \text{exponential bound.}$$

4

$$\Pr(X \geq (1+\delta)\mu) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

Show desmos.

→ $\min_{t>0}$ best value.

Gen : ~~if you~~

More informed ⟹ better bounds.

→ Analytically looking at bounds.