# 1 Concentration Inequalities

(Ref. Bertsekas Chapter 5)

① **Markov inequality:** Let $X$ be a non-negative random variable. Then

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

$A = \{X \geq a\}$

$$E[X] = E[X \mid A]\, P(A) + \underbrace{P(A^c)}_{\geq 0}\, \underbrace{E[X \mid A^c]}_{\geq 0}$$

total law of expectation

as $X$ is non-negative R.V.

$$\geq E[X \mid A]\, P(X \geq a).$$

$Y$ is a R.V st $Y \geq a$ then $E[Y] \geq a$.

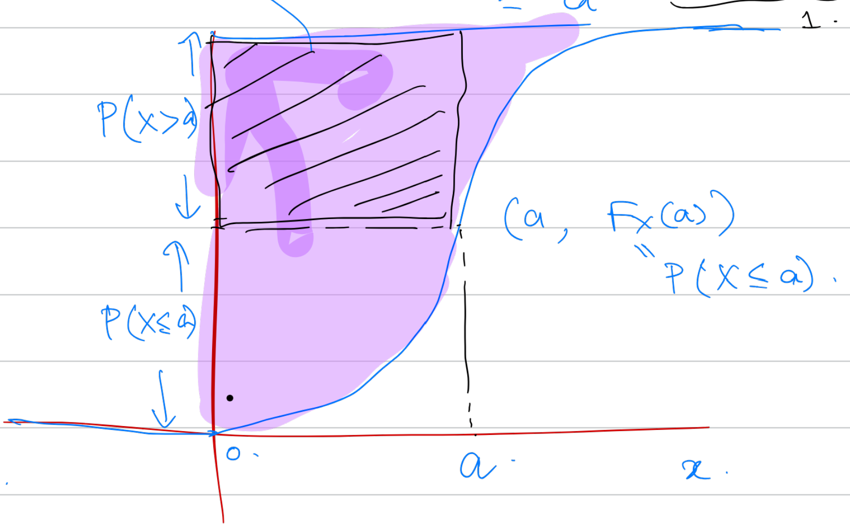
$$\geq a\, P(X \geq a).$$

$$E[X \mid X \geq a] = \int_{x \in \mathbb{R}} x\, f_{X|A}(x)\, dx = \int_a^\infty x\, f_{X|A}(x)\, dx \geq a \underbrace{\int_a^\infty f_{X|A}(x)\, dx}_{1} = a$$

**Visual Representation of Markov Inequality**

"Stanley Chan".

$E[X]$.




$$E[X] = \int_0^\infty (1 - F_X(x))\, dx.$$

$X \in \{0, 1, \ldots n\}$ $\Rightarrow$ $X$ is non-negative

**Examples** (a) $X \sim \text{Binomial}(n, p)$, let $p < \alpha < 1$.

$$P(X > \alpha n) \leq \frac{E[X]}{\alpha n}$$

$$= \frac{np}{n\alpha}$$

$$P(X > \alpha n) \leq \frac{p}{\alpha} = \frac{1/2}{3/4} = \frac{2}{3}.$$ Suppose $p = 1/2$, $\alpha = 3/4$.

(b) $X \sim \text{Uniform}[0, 4]$. Use Markov inequality to upper bound $P(X \geq 2)$, $P(X \geq 3)$, $P(X \geq 4)$.

$= \frac{1}{2}$, $= \frac{1}{4}$, $= 0$

$$P(X \geq 2) \leq \frac{E[X]}{2} = \frac{2}{2} = 1.$$

$$P(X \geq 3) \leq \frac{E[X]}{3} = 2/3$$

$$P(X \geq 4) \leq \frac{2}{4} = 1/2.$$

(2) **Chebyshev Inequality**: Let $X$ be a R.V with mean $\mu$ and variance $\sigma^2$. Then for any $\epsilon > 0$

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

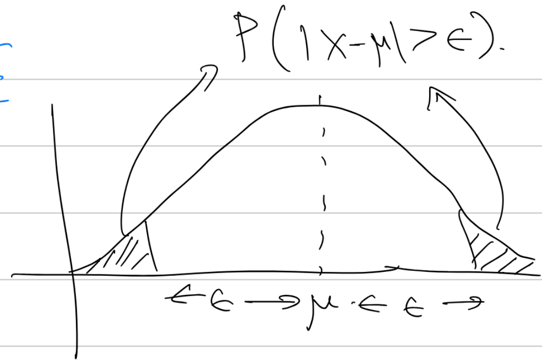
$$P(|X - \mu| > \epsilon) = P((X - \mu)^2 > \epsilon^2).$$

$Y = (X - \mu)^2$ is a non-negative R.V

$$P(|X-\mu| > \epsilon) = P(Y > \epsilon^2) \leq \frac{E[Y]}{\epsilon^2}$$ from Markov inequality.

$$= \frac{E[(X - \mu)^2]}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}.$$

$$P(|X-\mu|>\epsilon) \le \frac{\sigma^2}{\epsilon^2}$$




**Bounded R.V, $Y \in [a,b]$.**

Let $X$ be a R.V st $X$ takes values in $[0, c]$

$\mathrm{Var}(X) \le c^2/4$.

$$\begin{aligned}
\mathrm{Var}(X) &= E[X^2] - (E[X])^2. \\
&\le E[cX] - (E[X])^2 = c\,E[X] - (E[X])^2 \\
&= E[X](c - E[X]). \\
&\le \frac{c}{2}\,\frac{c}{2} = c^2/4.
\end{aligned}$$

$f(z) = zc - z^2$

$\max f(z)$

$f'(z^*) = c - 2z^* = 0$

$\Rightarrow z^* = c/2$.

Let $X = Y - a \in [0, b-a]$.

$$\mathrm{Var}(X) \le \frac{(b-a)^2}{4}.$$

$$\mathrm{Var}(Y) = \mathrm{Var}(X) \le (b-a)^2/4.$$

$$P(|Y - E[Y]|^2 > \epsilon) \le \frac{\mathrm{Var}(Y)}{\epsilon^2} \le \frac{(b-a)^2}{4\epsilon^2}.$$

**Examples:** (a) $X \sim \mathrm{Binomial}(n,p)$, $p < \alpha < 1$

$$\begin{aligned}
P(X > \alpha n) &= P(X - np > \alpha n - np) \\
&\le P(|X - np| > (\alpha - p)n). \\
&\le \frac{\mathrm{Var}(X)}{(\alpha-p)^2 n^2} = \frac{\cancel{n}\,p(1-p)}{(\alpha-p)^2 n^{\cancel{2}}}
\end{aligned}$$

Made with Goodnotes

If $p = 1/2$, $\alpha = 3/4$

$$P(X > \alpha n) \leq \frac{1/4}{\left(\frac{1}{4}\right)^2 n} = \frac{4}{n}.$$

(b) **Weak law of large numbers:** Let $X_1, X_2, \dots X_n$ iid R.Vs. with finite variance $Var(X_i) = \sigma^2$, mean $E[X_i] = \mu$

$$P\left( \left| \frac{\sum_{i=1}^{n} X_i}{n} - \mu \right| > \epsilon \right) \longrightarrow 0 \quad \text{as } n \to \infty$$

$$S_n = \sum_{i=1}^{n} X_i, \quad E\left[\frac{S_n}{n}\right] = E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right].$$

$$= \frac{1}{n} n E[X_i] = \mu$$

$$Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2} Var(S_n).$$

$$= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$$

$X_i$'s are independent.

$$= \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i). = \frac{\cancel{n}\, Var(X_i)}{n^{\cancel{2}}}$$

$$P\left( \left| \frac{S_n}{n} - \mu \right| > \epsilon \right) \leq \frac{Var\left(\frac{S_n}{n}\right)}{\epsilon^2} = \frac{\sigma^2}{n \epsilon^2}$$

$$\lim_{n \to \infty} P\left( \left| \frac{S_n}{n} - \mu \right| > \epsilon \right) \leq 0.$$

$$\Rightarrow \lim_{n \to \infty} P\left( \left| \frac{S_n}{n} - \mu \right| > \epsilon \right) = 0$$

as $P\left( \left| \frac{S_n}{n} - \mu \right| > \epsilon \right) \geq 0.$

Come up with an $\epsilon$ value that gaurantees that

$\frac{S_n}{n} \in [\mu - \epsilon, \mu + \epsilon]$ with probability $1 - 10^{-6}$.

## ③ Chernoff bound :

$$P(X>a) = P(e^{Xt} > e^{at}).$$

for any $t>0$.

Let $Y = e^{Xt}$, then $Y$ is non negative

$$= P(Y > e^{at})$$

Markov inequality

$$\le \frac{E[Y]}{e^{at}} = \frac{E[e^{Xt}]}{e^{at}}$$

$$P(X>a) \le \frac{M_X(t)}{e^{at}} \quad \text{for any } t>0.$$

$$P(X>a) \le \inf_{t>0} \frac{M_X(t)}{e^{at}}$$

$$= \inf_{t>0} e^{-(at - \ln M_X(t))}$$

$$= e^{-\sup_{t>0}(at - \ln M_X(t))}.$$

Let $f(a) = \sup_{t>0} \underbrace{(at - \ln M_X(t))}_{f_a(t)}$.

$$P(X>a) \le e^{-f(a)}.$$

Lemma : If $a > E[X]$ then $f(a) > 0$.

$$f_a(0) = a \cdot 0 - \ln M_X(0) = 0.$$

$$\frac{\partial f_a(t)}{dt} = a - \frac{1}{M_X(t)} \frac{\partial M_X(t)}{\partial t}$$

Made with Goodnotes

$$\left.\frac{\partial f_a(t)}{dt}\right|_{t=0} = a - \frac{1}{M_x(0)} \cdot \left.\frac{\partial M_x(t)}{dt}\right|_{t=0}$$

$$= a - E[x] > 0.$$

$\Rightarrow$ $f_a(t)$ is increasing around $t=0$.

$f_a(t) > 0$ for $t > 0$ and small $t$.

$\Rightarrow$ $f(a) > 0$.

Examples ① $X \sim N(0,1)$.

$$P(X \geq a) \leq e^{-f(a)}$$

$$f(a) = \sup_{t>0} (at - \ln M_x(t)).$$

$$f_a(t) = at - \ln M_x(t).$$

$$M_x(t) = E[e^{xt}]$$

$$= \int_{-\infty}^{\infty} e^{xt} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx.$$

$$= \int_{-\infty}^{\infty} \frac{e^{-(x-t)^2/2}}{\sqrt{2\pi}} e^{t^2/2} dx$$

$$= e^{t^2/2}.$$ MGF of Standard normal R.V.

$$f_a(t) = at - t^2/2.$$

$$\frac{\partial f_a(t)}{dt} = a - t, \quad f_a''(t) < 0.$$

$$f(a) = \sup_{t>0} f_a(t) = a \times a - a^2/2 = a^2/2.$$

$$P(X>a) \le e^{-f(a)} = e^{-a^2/2}.$$

If Chebyshev inequality is used we would get

$$P(X>a) \le \frac{1}{a^2}.$$

Compare both (Excercise).

② $X \sim \text{Binomial}(n, p)$ $\quad p < \alpha < 1$

$$P(X > \alpha n) \le \underbrace{\inf_{t>0} \overbrace{M_X(t)\, e^{-\alpha n t}}^{g_{\alpha,n}(t)}}_{g(\alpha,n)}$$

$$M_X(t) = \left((1-p) + p e^t\right)^n$$

$$g_{\alpha,n}(t) = \left((1-p) + p e^t\right)^n e^{-\alpha n t}$$

$$\frac{\partial g_{\alpha,n}(t)}{\partial t} = n\left((1-p)+pe^t\right)^{n-1} p e^t e^{-\alpha n t} + \left((1-p)+pe^t\right)^n e^{-\alpha n t}(-\alpha n)$$

$$= 0.$$

$$pe^t - \alpha(1-p+pe^t) = 0.$$

$$p(1-\alpha)e^t = \alpha(1-p).$$

$$e^t = \frac{\alpha}{p}\frac{(1-p)}{(1-\alpha)}.$$

$$g(\alpha,n) = \left((1-p) + \frac{\alpha(1-p)}{(1-\alpha)}\right)^n \left(\frac{p(1-\alpha)}{\alpha(1-p)}\right)^{\alpha n}$$

$$= \left(\frac{(1-p)}{1-\alpha}\right)^{n} \cdot \left(\frac{p\,(1-\alpha)}{\alpha\,(1-p)}\right)^{\alpha n}$$

$$= \left(\frac{p}{\alpha}\right)^{\alpha n} \left[\frac{(1-p)}{(1-\alpha)}\right]^{n(1-\alpha)}$$

For $p = 1/2, \quad \alpha = 3/4$

$$= \left(\frac{1/2}{3/4}\right)^{\frac{3n}{4}} \cdot \left(\frac{1/2}{1/4}\right)^{\frac{n}{4}}$$

$$= \left(\frac{2}{3}\right)^{\frac{3n}{4}} (2)^{n/4}$$

$$= \left[\left(\frac{2}{3}\right)^{3} \times 2\right]^{n/4}$$

$$= \left(\frac{16}{27}\right)^{n/4}$$

$$= e^{-n/4 \ln \frac{27}{16}}.$$

faster than $\frac{1}{n}$.

$$P\left(X > \frac{3}{4} n\right) \leq e^{-n/4 \ln \frac{27}{16}}.$$

$P\left(X > \frac{3}{4} n\right) \leq 2/3$ (Markov)

$P\left(X > \frac{3}{4} n\right) \leq \frac{4}{n}$ (Chebyshev)

$P\left(X > \frac{3}{4} n\right) \leq e^{-\frac{n}{4} \ln \frac{27}{16}}$ (Chernoff).