SOME CONCENTRATION INEQUALITIES

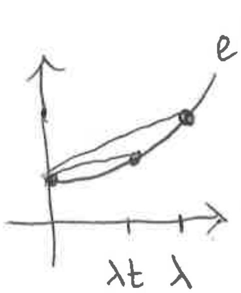
Thm 1 (Chernoff bound) Let $X_1, \dots, X_n$ be indep. r.vs with $0 \le X_i \le 1$ $\forall i \le n$. Let $S = X_1 + \dots + X_n$, $\mu = \mathbb{E}S$. Then,

$$\mathbb{P}(S > \mu + t) \le e^t \left(\frac{\mu}{\mu+t}\right)^{\mu+t}, \qquad 0 \le t \le n - \mu.$$

Proof For $\lambda > 0$, we have

$$\mathbb{P}(S > \mu+t) = \mathbb{P}(\lambda S > \lambda(\mu+t)) \le e^{-\lambda(\mu+t)} \mathbb{E} e^{\lambda S}$$

$$= e^{-\lambda(\mu+t)} \prod_{i=1}^n \mathbb{E} e^{\lambda X_i}.$$



By convexity, $\forall\ 0 \le t \le 1$

$$\frac{e^{\lambda t} - 1}{\lambda t} \le \frac{e^\lambda - 1}{\lambda}$$

$$e^{\lambda t} \le 1 + t(e^\lambda - 1),$$

so $\mathbb{E} e^{\lambda X_i} \le 1 + (e^\lambda - 1)\mathbb{E}X_i$, so

$$\prod_{i=1}^n \mathbb{E} e^{\lambda X_i} \le \prod_{i=1}^n \left(1 + (e^\lambda - 1)\mathbb{E}X_i\right)$$

$$\underset{\text{AM-GM}}{\le} \left(\frac{\sum [1 + (e^\lambda - 1)\mathbb{E}X_i]}{n}\right)^n$$

$$= \left(1 + \frac{e^\lambda - 1}{n}\mu\right)^n,$$

$$\mathbb{P}(S > \mu+t) \le e^{-\lambda(\mu+t)} \left(1 + \frac{e^\lambda - 1}{n}\mu\right)^n.$$

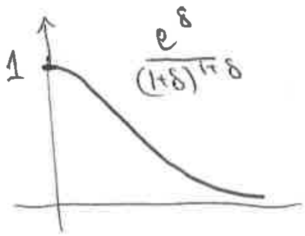
Choosing $\lambda > 0$ which minimises the RHS yields ...

We choose $e^{-\lambda} = \frac{\mu}{\mu+t}$, which gives

$$\mathbb{P}(S > t+\mu) \leq \left(\frac{\mu}{\mu+t}\right)^{\mu+t}\left(1+\frac{\mu}{n}\left(\frac{\mu+t}{\mu}-1\right)\right)^{n}$$

$$= \left(\frac{\mu}{\mu+t}\right)^{\mu+t}\left(1+\frac{t}{n}\right)^{n} \leq e^{t}\left(\frac{\mu}{\mu+t}\right)^{\mu+t}. \quad \square$$

⚠ 1) The Chernoff bound is usually written as (set $t = \delta\mu$)



$$\mathbb{P}(S > (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}, \quad \delta > 0$$

$$\leq \begin{cases} e^{-\frac{\delta^2\mu}{3}}, & 0 \leq \delta \leq 1 \\ e^{-\frac{\delta\mu}{3}}, & \delta \geq 1. \end{cases}$$

2) The same arguments give bounds for lower tails

$$\mathbb{P}(S < \mu - t) \overset{\lambda<0}{=} \mathbb{P}(\lambda S > \lambda(\mu-t)) \leq e^{-\lambda(\mu-t)}\mathbb{E}e^{\lambda S}$$

$$\leq e^{-t}\left(\frac{\mu}{\mu-t}\right)^{\mu-t}, \quad \mu \geq t > 0.$$

3) By $x + \frac{x^2}{2(1+x/3)} \leq (1+x)\log(1+x)$, $x>0$, we also get

$$P(S > \mu+t) \leq \exp\left\{-\frac{t^2}{2(\mu+t/3)}\right\}, \quad t>0,$$

which is another consequence of the Chernoff bound, useful in applications.

E.g. 1) $S \sim \mathrm{Bin}(\frac{1}{2}, n)$

$$\mathbb{P}\left(S > (1+\delta)\frac{n}{2}\right) \le \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{n/2}$$

$\delta = \frac{1}{2}$

$$\mathbb{P}\left(S > \frac{3n}{4}\right) \le \left(\frac{e^{1/2}}{(3/2)^{3/2}}\right)^{n/2}$$

2) $S \sim \mathrm{Bin}(\frac{1}{n}, n)$ $\quad$ (Poiss (1))

$$\mathbb{P}(S > 1+\delta) \le \frac{e^{\delta}}{(1+\delta)^{1+\delta}}$$

$$\mathbb{P}(S > k) \le \frac{1}{e}\left(\frac{e}{k}\right)^k$$

$$\left(\mathbb{P}(\mathrm{Poiss}(1) = k) = \frac{1}{e}\frac{1}{k!} \sim \frac{1}{e}\left(\frac{e}{k}\right)^k \frac{1}{\sqrt{2\pi k}} \right)$$

Thm 2 (Bernstein bound) Let $X_1, \dots, X_n$ be indep. r.vs with $|X_i| \le 1$, $\mathbb{E}X_i = 0$. Let $S = X_1 + \dots + X_n$, $\sigma^2 = \mathrm{Var}(S)$. Then,

$$\mathbb{P}(S > t) \le \exp\left\{ -\frac{t^2}{2(\sigma^2 + t/3)} \right\}, \quad t > 0.$$

Proof For $\lambda > 0$,
$$\mathbb{P}(S > t) = \mathbb{P}(\lambda S > \lambda t) \le e^{-\lambda t}\, \mathbb{E}e^{\lambda S} = e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}e^{\lambda X_i}.$$

We have,
$$\mathbb{E}e^{\lambda X_i} = \mathbb{E}\left(1 + \lambda X_i + \frac{1}{2!}(\lambda X_i)^2 + \dots\right)$$
$$\le 1 + \sum_{k=2}^{\infty} \frac{\lambda^k}{k!} \mathbb{E}|X_i|^k$$

For $k \geq 2$, $\mathbb{E}|X_i|^k = \mathbb{E}\underbrace{|X_i|^{k-2}}_{\leq 1}|X_i|^2 \leq \mathbb{E}X_i^2$, so

$$\mathbb{E}e^{\lambda X_i} \leq 1 + \left(\sum_{k=2}^{\infty} \frac{\lambda^k}{k!}\right)\mathbb{E}X_i^2$$

$$= 1 + (e^\lambda - \lambda - 1)\,\mathbb{E}X_i^2$$

$$\leq \exp\{(e^\lambda - \lambda - 1)\,\mathbb{E}X_i^2\},$$

$$\mathbb{P}(S > t) \leq e^{-\lambda t}\exp\{(e^\lambda - \lambda - 1)\sigma^2\} = \exp\{-\lambda(t+\sigma^2) + e^\lambda\sigma^2 - \sigma^2\}.$$

Set $e^\lambda = 1 + \frac{t}{\sigma^2}$, to get

$$\mathbb{P}(S > t) \leq \exp\left\{-(t+\sigma^2)\log\left(1 + \frac{t}{\sigma^2}\right) + t\right\}$$

$$= e^t\left(\frac{\sigma^2}{\sigma^2 + t}\right)^{t+\sigma^2}.$$

As in ⚠3) we get the assertion. □