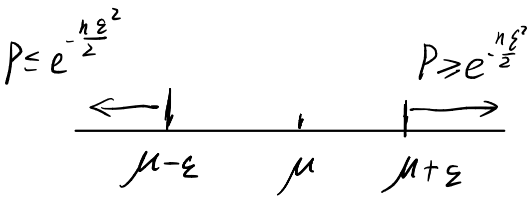
Chernoff inequality

Chernoff inequality:

Let $X_1, \cdots, X_n$ be independent r.v. Each X_i follows a probability distribution $P_i \in \mathcal{E}_{[0,1]}$, with mean $E[X_i] = \mu_i \in [0,1]$. Denote $\mu = \frac{1}{n}\sum_{i=1}^{n}\mu_i$. for any $\varepsilon > 0$,
$(\sum P_i = 1)$

it holds that

$$\boxed{\begin{aligned} &Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq \mu + \varepsilon\right] \leq \exp\left(-\frac{n\varepsilon^2}{2}\right) \\ &Pr\left[\frac{1}{n}\sum_{i=1}^{n} X_i \leq \mu - \varepsilon\right] \leq \exp\left(-\frac{n\varepsilon^2}{2}\right)\end{aligned}}$$



Proof:

denote $Y_i = X_i - \mu_i$ for each $i \in \{1, 2, \cdots, n\}$. Let $\lambda > 0$ be a constant.

(step 1) we have that

$$\begin{aligned} &Pr\left[\frac{1}{n}\sum X_i \leq \mu - \varepsilon\right] \\ &= Pr\left[-\frac{1}{n}\sum Y_i \geq \varepsilon\right] \\ &= Pr\left[-\lambda \cdot \sum Y_i \geq \lambda \cdot n \cdot \varepsilon\right] \\ &= Pr\left[\exp(-\lambda \cdot \sum Y_i) \geq \exp(\lambda \cdot n \cdot \varepsilon)\right] \\ &\leq \exp(-\lambda \cdot n \cdot \varepsilon) \cdot E\left[\exp(-\lambda \cdot \sum Y_i)\right] \\ &= \exp(-\lambda \cdot n \cdot \varepsilon) \cdot \prod_{i=1}^{n} E\left[\exp(-\lambda \cdot Y_i)\right] \end{aligned}$$

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

(step 2) $E(Y_i) = E(X_i) - E(\mu_i) = 0$

we note that $f(y) = \exp(-\lambda y) = e^{-\lambda y}$ is convex in y. for each $y \in [-1, 1]$,

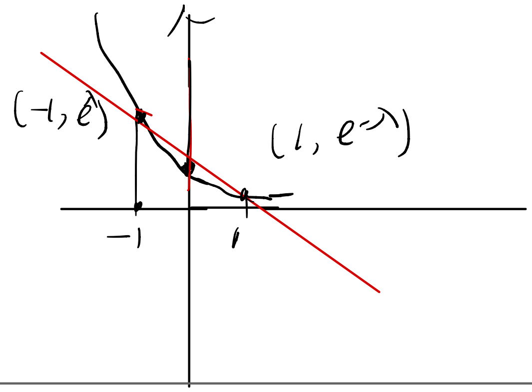
the function is upper bounded by a straigh line crossing point $(1, e^{-\lambda})$ and $(-1, e^{\lambda})$. The line is govemed by the equation

$$z = \frac{e^{-\lambda} - e^{\lambda}}{2} Y_i + \frac{e^{-\lambda} + e^{\lambda}}{2}$$

$$\frac{z-e^{\lambda}}{e^{-\lambda}-e^{\lambda}} = \frac{y+1}{2} \qquad \frac{y-y_2}{y_1-y_2} = \frac{x-x_2}{x_1-x_2}$$

$$2z-2e^{\lambda} = (e^{-\lambda}-e^{\lambda})\,y + (e^{-\lambda}-e^{\lambda})$$

$$z = \frac{e^{-\lambda}-e^{\lambda}}{2}\,y + \frac{e^{-\lambda}+e^{\lambda}}{2}$$



---

so $f(y) = \exp(-\lambda y_i) \le \frac{e^{-\lambda}-e^{\lambda}}{2} y_i + \frac{e^{-\lambda}+e^{\lambda}}{2}$, $y \in [-1, 1]$

$$\Rightarrow E[\exp(-\lambda y_i)] \le \frac{e^{-\lambda}+e^{\lambda}}{2}$$

Using Taylor expansion of $e^{\lambda} = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}$

$$\frac{1}{2}(e^{\lambda}+e^{-\lambda}) = \sum_{i=0}^{\infty} \frac{\lambda^{2i}}{(2i)!} \le \sum_{i=0}^{\infty} \frac{\lambda^{2i}}{2^i \cdot i!} = \sum_{i=0}^{\infty} \frac{(\lambda^2/2)^i}{i!} = \exp\left(\frac{\lambda^2}{2}\right)$$

(Step3)

$$\Pr\left[\frac{1}{n}\sum X_i \le \mu - \varepsilon\right] \le \exp(-\lambda \cdot n \cdot \varepsilon) \cdot \prod_{i=1}^{n} E[\exp(-\lambda \cdot Y_i)]$$

$$\le \exp(-\lambda \cdot n \cdot \varepsilon) \cdot \exp\left(\frac{n \cdot \lambda^2}{2}\right)$$

$$= \exp\left[n \cdot \left(-\lambda \varepsilon + \frac{\lambda^2}{2}\right)\right]$$

$$\overset{\text{Let } \lambda = \varepsilon}{=\!=\!=} \exp\left[\frac{-n\varepsilon^2}{2}\right]$$