

Chernoff bound

Theorem: Let X_1, \dots, X_n be independent random variables that lie in interval $[0, 1]$ with probability one. Let $p = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i]$.

For any $\epsilon > 0$,

$$\Pr \left[\frac{1}{n} \sum_{i=1}^n X_i \geq (1+\epsilon)p \right] \leq \left[\frac{e^\epsilon}{(1+\epsilon)^{1+\epsilon}} \right]^{np}$$

Proof: Let $p_i = \mathbb{E}[X_i]$

Chernoff bounding technique.

$$\Pr \left[\frac{1}{n} \sum_{i=1}^n X_i \geq (1+\epsilon)p \right]$$

$$= \Pr \left[e^{s \sum_{i=1}^n X_i} \geq e^{(1+\epsilon)snp} \right]$$



$$\leq \frac{\mathbb{E} \left[e^{s \sum_{i=1}^n X_i} \right]}{e^{(1+\epsilon) sn}}$$

$$= \frac{\prod_{i=1}^n \mathbb{E} \left[e^{s X_i} \right]}{e^{(1+\epsilon) sn}} \quad (\text{by independence})$$

$$\leq \frac{\prod_{i=1}^n (1 - p_i + p_i e^s)}{e^{(1+\epsilon) sn}} \quad (\text{HW \#1})$$

$$= \frac{\prod_{i=1}^n (1 + p_i (e^s - 1))}{e^{(1+\epsilon) sn}}$$

$$\leq \frac{\prod_{i=1}^n e^{p_i (e^s - 1)}}{e^{(1+\varepsilon) \sum p_i}}$$

$$1+z \leq e^z$$

$$= \frac{e^{(e^s - 1) \sum_{i=1}^n p_i}}{e^{(1+\varepsilon) \sum p_i}}$$

$$= \frac{e^{(e^s - 1) \sum p_i}}{e^{(1+\varepsilon) \sum p_i}}$$

$$= \left(e^{(e^s - 1) - (1+\varepsilon)s} \right)^{\sum p_i}$$

Find s that minimizes the

last expression.

$$f(s) = e^s - 1 - (1+\varepsilon)s$$

$$f'(s) = e^s - (1+\varepsilon)$$

Solve $f'(s) = 0$:

$$e^s - (1+\varepsilon) = 0$$

$$s = \ln(1+\varepsilon)$$

$$\Pr[\dots] \leq \left(e^{(e^s-1)} - (1+\varepsilon)s \right)^{np}$$

$$= \left(e^{(1+\varepsilon-1)} - (1+\varepsilon)\ln(1+\varepsilon) \right)^{np}$$

$$= \left(\frac{e^\varepsilon}{(1+\varepsilon)^{1+\varepsilon}} \right)^{np}$$

