Inequalities

- many important models are too complex to analyze exactly
- help us prove important limiting theorems

We will discuss $\underbrace{\text{Markov, Chebyshev}}_{\text{EE178}}$ & Hoeffding's bound

simplest example of a concentration ineq.

Markov's Inequality:

For any non-negative r.v X and $t > 0$, we have

$$\mathbb{P}\{X \geq t\} \leq \frac{\mathbb{E}[X]}{t}$$

↳ tail probability (one-sided)

e.x. average height of Stanford students is 165cm
5 feet 5 inches.

fraction of students that are above 180 cm
5 feet 11 inches

$$P\{X \geq 180\} \leq \frac{165}{180}$$

$$\mathbb{1}_{\{x \geq t\}} = \begin{cases} 1 & x \geq t \\ 0 & x < t \end{cases}$$

$$X = X \mathbb{1}\{X \geq t\} + X \mathbb{1}\{X \leq t\}$$

$$\mathbb{E}[X] = \mathbb{E}\left[X \mathbb{1}\{X \geq t\}\right] + \mathbb{E}\left[X \mathbb{1}\{X \leq t\}\right]$$

$$= \int x \mathbb{1}\{x \geq t\} f_X(x)\,dx + \int x \mathbb{1}\{x \leq t\} f_X(x)\,dx$$

$$= \int_t^\infty x f_X(x)\,dx + \int_0^t x f_X(x)\,dx.$$

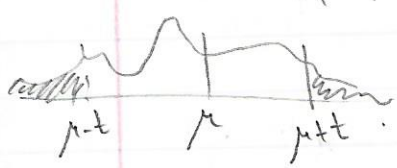
$$\geq t \int_t^\infty f_X(x)\,dx = t\,P\{X \geq t\}.$$

$$\Rightarrow P\{X \geq t\} \leq \frac{\mathbb{E}[X]}{t}$$

Chebyshev's inequality:
- better quadratic dependence on t.
- quantifies concentration of X about mean, (controls both tails).

Let X be a r.v with mean μ and variance σ^2. Then for any $t > 0$, we have

$$\mathbb{P}\{|X-\mu| \geq t\} = P\{(X-\mu)^2 \geq t^2\}$$



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

Limit Theorems:

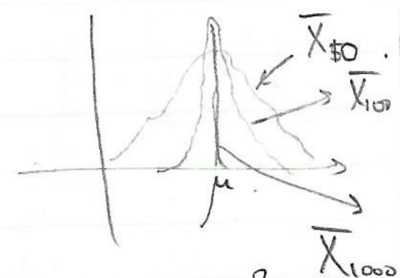
1) The Weak Law of Large Numbers (WLLN).

Let $X_1, X_2, \ldots X_n$ iid r.v's with mean $\mu = \mathbb{E}[X_1] = \ldots = \mathbb{E}[X_n]$

$S_n = \sum_{i=1}^{n} X_i \quad \bar{X}_n = \frac{1}{n} S_n. \quad \bar{X}_n \to \mu$ as $n \to \infty$. LLN

WLLN:

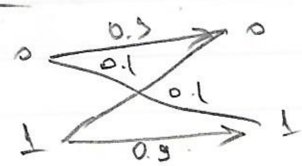
$$\forall \varepsilon > 0 \quad P\{|\bar{X}_n - \mu| > \varepsilon\} \xrightarrow{n \to \infty} 0.$$



Ex: X_i independent flips of a coin with $P\{\underset{\downarrow H}{X_i = 1}\} = 0.3$

Ex: Statistics / survey sampling.

Ex: Reliable communication.



Codebook = { a collection of binary strings that are well separated from each other } codewords

$C = \{ 00000, 00111, 11100, 11011 \}$.

use a codebook with codewords of length n for large n and make sure all codewords are seperated from each other with a distance $2(0.1+\varepsilon)n$ for some small ε

$$\frac{\#\text{ of errors}}{n} \longrightarrow 0.1.$$

e.x. Monte Carlo Integration.

$$I(f) = \int_0^1 f(x)\,dx \longrightarrow$$ numerical int. rieman sum.

generate $X_1, \dots X_n$ iid uniform $[0,1]$.

$$\hat{I}(f) = \frac{1}{n}\sum_{i=1}^{n} f(X_i).$$

$$\hat{I}(f) \underset{n\to\infty}{\longrightarrow} \mathbb{E}[f(X_1)] = \int_0^1 f(x)\,dx$$

Proof: $Var(\bar{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

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

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mu$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)$$ because $X_1, \dots X_n$ are ind.

$$= \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}.$$

Chebyshev's Inequality.

$$P\{ |\bar{X}_n - \mu| > \varepsilon \} \le \frac{Var(\bar{X}_n)}{\varepsilon^2}.$$

$$= \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0 \text{ as } n \to \infty$$

$$S_n = \sum_{i=1}^{n} X_i$$

$$\frac{S_n}{n} - \mu = \frac{S_n - n\mu}{n}$$

$$Var\left[\frac{S_n - n\mu}{n}\right] = \frac{1}{n}\sigma^2$$

$$W_n = \frac{\sqrt{n}}{\sigma}\left(\frac{S_n - n\mu}{n}\right) \qquad Var(W_n) = n \cdot \frac{1}{n}\frac{\sigma^2}{\sigma^2} = 1$$

$$= \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{n\to\infty} N(0,1)$$

$$E[W_n] = 0$$

$$Var(W_n) = 1$$

<u>Central Limit Theorem:</u> Let $X_1, X_2, \ldots X_n$ iid with mean μ and variance σ^2.

Let $W_n = \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma}$ be their standardized sum r.v's

then

$$F_{W_n}(w) = P\{W_n \le w\} \xrightarrow{n\to\infty} P\{W \le w\}$$

cdf of W_n — convergence in distribution.

for every w; where $W \sim N(0,1)$.

CLT: some justification for Gaussian models.

noise = sum of many independent small noise sources.

ex. thermal noise.