Lecture 4: Chernoff Bound

→ (Markov inequality) For a non-negative R.V X,

$$\Pr(X \geq a) \leq \frac{E(X)}{a}, \text{ for all } a \geq 0 \Rightarrow \text{증명}$$

의미: X → distribution ⇒ 예측, 정보, bound   평균을 알고있으면 분포에 대한 정보

어렵다. $P(X \leq a) = 1 - \Pr(X \geq a)$ tail probability

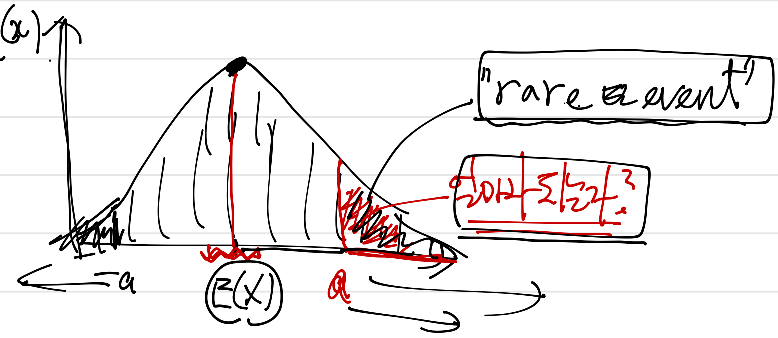
→ (Chebyshev's inequality) For any $a > 0$,

$$\boxed{\Pr\left(|X - E[X]| \geq a\right) \leq \frac{Var[X]}{a^2}} \Rightarrow \text{증명}$$

평균에서 얼마나 많이 떨어져 있는 부분이 많을까?

X: 평균에서 벗어난 구간에서는 어떤 행동?

Continuous R.V X $f_X(x)$



의미: Variance를 알면, 평균에서 얼마나 떨어져있는가 → 어떤식의 추측?

Corollary) For any $t > 1$,

① $\Pr\left(|X - E(X)| \geq t \cdot \sigma(X)\right) \leq \frac{Var(X)}{t^2 \cdot \sigma^2(X)} = \frac{1}{t^2}$ ($\sigma(X)$: 표준편차; $t \cdot \sigma(X) = a$)

② $\Pr\left(|X - E(X)| \geq t E(X)\right) \leq \frac{Var(X)}{t^2 E^2(X)}$ ($t E(X) = a$)

→ second moment

(Ex) – n fair coin flips → Prob(more than $\frac{3}{4}n$ heads)

Let X be the R.V that represents # of heads in n fair coin flips

(i) MI: $\Pr\left(X \geq \frac{3}{4}n\right) \leq \frac{E(X)}{\frac{3}{4}n} = \frac{\frac{n}{2}}{\frac{3}{4}n} = \frac{4}{6} = \frac{2}{3}$

(ii) CI: $\Pr\left(X \geq \frac{3}{4}n\right) \leq \Pr\left(|X-E(X)| \geq \frac{n}{4}\right) \leq \frac{Var(X)}{\left(\frac{n}{4}\right)^2} = \frac{\frac{n}{4}}{\left(\frac{n}{4}\right)^2} = \frac{4}{n}$

$E(X) = \frac{1}{2}n$

n이 충분히 크면 CI가 훨씬 tight

$Var(X) = Var\left(\sum_{i=1}^{n} Y_i\right) = nVar(Y_i)$

X: Binomial (n,p) $\underline{(n,\frac{1}{2})}$ $\to np(1-p) \to \frac{n}{4}$

Y_i: Bernoulli with p → $\frac{1}{4}$

(Chernoff Bound) Given i.i.d RVs, $X_1, \dots, X_n$, for any $a \in \mathbb{R}$,

$$\Pr\left(\sum_{i=1}^{n} X_i \geq na\right) \leq e^{-nh(a)}$$

⇐ 관심있는 regime $a > E(X_1)$

where $h(a) = \sup_{\theta \geq 0}\left[\theta a - \log E\left(e^{\theta X_1}\right)\right]$.

(의미) $\Pr\left(\frac{\sum_{i=1}^{n} X_i}{n} \geq a\right)$, $\frac{\sum_{i=1}^{n} X_i}{n} = Y_n$

(Question) $n \to \infty$, "얼마나 빨리" $\frac{\sum X_i}{n}$ →

CB → exponential하게 빨리.

조건 $a > E(X_1)$

$a < E(X)$ trivial

그경우에 $h(a) < 0$

대수의 법칙 Law of Large Numbers (LLN)

$$\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{n\to\infty} E(X_1)$$

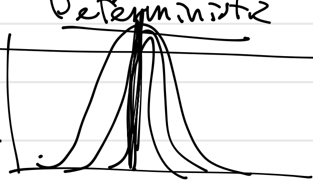
R.V (확률적 세상) Probabilistic

상수 (결정적 세상) Deterministic

→ 수렴으로 2가지 있다. (why?)

(hint) [illegible]

① SLLN
② WLLN



Proof) For all $\theta \geq 0$, ~~by MI~~

$$\Pr\left(\sum_{i=1}^{n} X_i \geq na\right) = \Pr\left(e^{\theta \sum_{i=1}^{n} X_i} \geq e^{\theta na}\right)$$

$$\leq e^{-\theta na} E\left(\left(e^{\theta \sum_{i=1}^{n} X_i}\right)\right) \quad (\text{from M.I})$$

$$= e^{-\theta na} \cdot E\left(e^{\theta X_1} \cdot e^{\theta X_2} \cdot e^{\theta X_3} \cdots \cdot e^{\theta X_n}\right)$$

$$= e^{-\theta na} \cdot \boxed{\left(E\left(e^{\theta X_1}\right)\right)^n} \rightarrow e^{n \times (\log E(e^{\theta X_1}))}$$

$$= e^{-n\left(\theta a - \log E\left(e^{\theta X_1}\right)\right)}$$

$$\leq \sup_{\theta \geq 0} e^{-n \underbrace{\left(\theta a - \log E\left(e^{\theta X_1}\right)\right)}_{h(a)}}.$$

$$\boxed{h(a)} = \sup_{\theta \geq 0}\left(\theta a - \log E\left[e^{\theta X_1}\right]\right) = \sup_{\theta \geq 0}\left(\theta a - \log MGF\right)$$

rate function

X_1의 MGF

log MGF

Homework

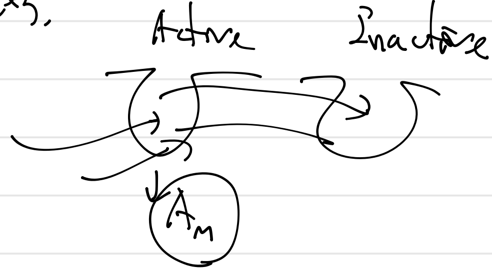
(1) $X \sim$ Gaussian with mean μ, var: σ^2 $\quad h(a) = \boxed{\frac{(a-\mu)^2}{2\sigma^2}}$

(2) $X \sim \exp(\mu)$, $h(a) = ?$

(3) $X \sim \text{poisson}(\lambda)$, $h(a) =$

(4) $X \sim \text{Bernoulli}(p)$, $h(a) = ?$

1.5 ~~Top~~ Total population 의 bound 계산하기.



Recall : $A_m = 1 + \boxed{\sum_{i=1}^{n} \xi_i} - m$

$$P(X \geq k) = \Pr(A_1 > 0, A_2 > 0, \ldots, A_k > 0)$$

$$\leq \Pr(A_k > 0)$$

$$= \Pr\left(1 + \sum_{i=1}^{k} \xi_i - k > 0\right) = \Pr\left(\sum_{i=1}^{k} \xi_i > k-1\right)$$

$$= \Pr\left(\sum_{i=1}^{k} \xi_i \geq k\right) = \Pr\left(\frac{\sum_{i=1}^{k} \xi_i}{k} \geq 1\right) \quad E(\xi_i) < 1$$

$$\leq \exp(-k\,h(1)),$$ where $h(x)$는 ξ_i의 rate function,

$$h(x) \triangleq \sup_{\theta \geq 0} \left[\theta x - \log E(\exp \theta \xi)\right].$$

∎

<Chernoff Bound for independent Poisson trials>

→ 앞에서는: i.i.d $X_1, X_2, \ldots, X_n$ $\Pr\left(\frac{\sum^{n} X_i}{n} > a\right)$

generalize

→ 여기서는 " ~~identical~~ (independent, not necessarily identical)

$X_i = 0, 1$: poisson trial, Bernoulli trial

Let $\Pr(X_i = 1) \stackrel{\circ}{=} P_i$ $\quad P_i \neq P_j$.

$\hookrightarrow \Pr\left(\frac{\sum_{i=1}^{n} X_i}{n} \geq a\right)$ $\quad // \quad \frac{1}{n} E\left(\sum X_i\right) = \frac{1}{n} \sum P_i$

→ $a > E\left(\frac{\sum X_i}{n}\right)$ 일때 의미가 있다.

$\Pr\left(\sum_{i=1}^{n} X_i \geq na\right)$ $\quad a > \frac{\sum P_i}{n} \Rightarrow na > \sum P_i$

Let $X = \sum_{i=1}^{n} X_i$ $\quad$ Let $\mu \stackrel{\circ}{=} E(X) = \sum P_i$.

$\Pr\left(\frac{X}{n} \geq \frac{(1+\delta)\mu}{n}\right)$??? $\Rightarrow$ $\boxed{\Pr\left(X \geq (1+\delta)\mu\right)}$

Thm $\Pr(X \geq (1+\delta)\mu) \leq e^{-\mu h(\delta)}$, where

$h(\delta) = (1+\delta)\log(1+\delta) - \delta$ $\quad$ for any $\delta \geq 0$

$$e^{-\mu h(\delta)} = e^{-\mu\left((1+\delta)\log(1+\delta) - \delta\right)} = e^{-\mu(1+\delta)\log(1+\delta)} \cdot e^{\mu\delta}$$

$$= \frac{e^{\mu\delta}}{\left(\left(e^{\log(1+\delta)}\right)^{1+\delta}\right)^{\mu}}$$

$$= \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$

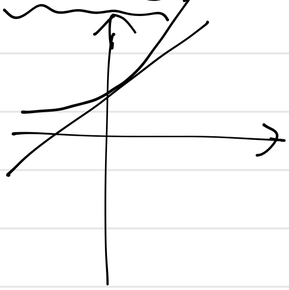
(Proof) For any $\theta \geq 0$,

$$\Pr(X-\mu \geq \delta\mu) \leq E\left(e^{\theta(\overbrace{X}^{\sum X_i}-\overbrace{\mu}^{\sum p_i})}\right)\cdot e^{-\theta\delta\mu} \quad (\text{from M.I})$$

$$= e^{-\theta\delta\mu}\prod_{i=1}^{n} E\left(e^{\theta(X_i-p_i)}\right)$$

linear $\rightarrow$ exp

$1+x \leq e^{x}$ $(x \geq 0)$



$$= e^{-\theta\delta\mu}\cdot e^{-\theta\mu}\prod_{i=1}^{n}\boxed{E\left(e^{\theta X_i}\right)} = (1-p_i)\cdot 1 + p_i e^{\theta} = 1+p_i(e^{\theta}-1)$$

$$= e^{-\theta\mu(1+\delta)}\prod_{i=1}^{n}\left(1+p_i(e^{\theta}-1)\right)$$

이걸생략

$$\leq e^{-\theta\mu(1+\delta)}\prod_{i=1}^{n} e^{p_i(e^{\theta}-1)}$$

$$= e^{-\theta\mu(1+\delta)}\cdot\left(e^{\mu}\right)^{(e^{\theta}-1)} = e^{-\mu\left(\theta(1+\delta)+1-e^{\theta}\right)}$$

$$\leq e^{-\mu\left(\theta^{*}(1+\delta)+1-e^{\theta^{*}}\right)}$$

$f(\theta) = \theta(1+\delta)+1-e^{\theta}$

$f'(\theta) = 1+\delta - e^{\theta} = 0$ $\quad e^{\theta} = 1+\delta$

$\theta^{*} = \log(1+\delta)$

(Homework) 책에

(1) If $0 < \delta \leq 1$, $\Pr(X \geq (1+\delta)\mu) \leq e^{-\mu\frac{\delta^2}{3}}$

(hint) $\left(\dfrac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu} \leq e^{-\mu\frac{\delta^2}{3}}$ $\quad 0 < \delta \leq 1$

(2) $R \geq 6\mu$ $\quad P(X \geq R) \leq 2^{-R}$

$\boxed{\Pr(X \geq (1+\delta)\mu)}$

$\boxed{\delta \geq 5}$