Lecture 19

Professor Virginia R. Young
Transcribed by Hao Chen

October 24, 2022

Limit Theorems

Defn: Markov's Inequality

If X is a random variable with finite mean, then for any $a > 0$,

$$\mathbb{P}(|X| \geq a) \leq \frac{1}{a}\mathbb{E}(|X|)$$

Proof: Let $A = \{\omega : |X(\omega) \geq a|\}$

$$|X| \geq a1_A$$

where the right hand side is given by rhs $= \begin{cases} a & \text{if } |X| \geq a \\ 0 & \text{if } |X| < a \end{cases}$

$$\begin{aligned} \implies \mathbb{E}(|X|) &\geq a\mathbb{E}(1_A) \\ &= a\mathbb{P}(A) \\ &= a\mathbb{P}(|X| \geq a) \end{aligned}$$

$$\implies \mathbb{P}(|X| \geq a) \leq \frac{1}{a}\mathbb{E}(|X|)$$

Defn: Chebyshev's inequality

If X is a random variable with finite mean μ and variance σ^2, then for any $k > 0$,

$$\mathbb{P}(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Proof: Apply Markov's inequality to $(X-\mu)^2$ with $a = k^2$:

$$\mathbb{P}((X-\mu)^2 \geq k^2) \leq \frac{1}{k^2}\mathbb{E}((X-\mu)^2)$$

$$\implies \mathbb{P}(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

E.g.: Consider three trials(not necessarily iid): $X_i = 1$ if the i^{th} trial is a success, 0, otherwise. Let $X = X_1 + X_2 + X_3$ to be the total number of successes. Suppose $\mathbb{E}X = 1.8$.

1. What is the largest possible value of $\mathbb{P}(X = 3)$?
   Solution:

   $$\begin{aligned} \mathbb{P}(X = 3) &= \mathbb{P}(X \geq 3) \\ &\leq \frac{1}{3}\mathbb{E}X = 0.6 \end{aligned}$$

   Achieve the bound as follows:
   $X_i \sim \text{Bern}(0.6)$, $i = 1, 2, 3$
   $X_1 = X_2 = X_3 \implies X \sim 3 \cdot \text{Bern}(0.6)$

   $$X = \begin{cases} 3 & \text{wp } 0.6 \\ 0 & \text{wp } 0.4 \end{cases}$$

2. What is the smallest possible value of $\mathbb{P}(X = 3)$?
   Solution: 0

   Let $U \sim \text{Unif}(0, 1)$, we define X_1, X_2 as following

   $$X_1 = \begin{cases} 0 & \text{if } 0 < U \leq 1/2 \\ 1 & \text{if } 1/2 < U \leq 1 \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{if } 0 < U \le 1/2 \\ 0 & \text{if } 1/2 < U \le 1 \end{cases}$$

such that $X_1 + X_2 = 1$ and $\mathbb{E}(X_1 + X_2) = 1$.

$$X_3 = \begin{cases} 1 & \text{if } 0 < U \le 0.8 \\ 0 & \text{if } 0.8 < U \le 1 \end{cases}$$

where $\mathbb{E}X_3 = 0.8 \implies \mathbb{E}X = 1.8$.

Thm: Strong law of large number(pf in Section 2.8)

Let $X_1, X_2, \dots$ be a sequence of iid r.v.s with common, finite mean μ. Then

$$\mathbb{P}(\{\omega : \lim_{n\to\infty} \bar{X}_n = \mu\}) = 1.$$

where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Often written $\bar{X}_n \to \mu$ as $n \to \infty$ w.p. 1.

Thm: Central Limit Theorem(CLT)

Let $X_1, X_2, \dots$ be a sequence of iid r.v.s with common, finite mean μ and variance σ^2. Then, the distribution of

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

approaches that of $N(0,1)$ as $n \to \infty$.

$$\mathbb{E}(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}X_i = \frac{n\mu}{n} = \mu$$

$$\begin{aligned} \mathrm{Var}(\bar{X}_n) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^n X_i\right) \\ &\overset{\text{ind't}}{=} \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\, X_i = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \le z\right) = \Phi(z)$$

or

$$\lim_{n\to\infty} \mathbb{P}\left(\bar{X}_n \le \mu + z \cdot \frac{\sigma}{\sqrt{n}}\right) = \Phi(z)$$

Aside: Usual proof involves the characteristic function of X_i, $\Phi(t) = \mathbb{E}(e^{iXt})$. Instead, we will assume $M_X(t) = \mathbb{E}(e^{Xt})$ exists in an open interval about 0.

Let $Y_i = \frac{X_i - \mu}{\sigma}$

$$\begin{aligned} M_{Y_i}(t) &= \mathbb{E}(e^{Y_i t}) \\ &= \mathbb{E}(1 + Y_i t + \frac{1}{2}(Y_i t)^2 + \frac{1}{6}(Y_i t)^3 + \dots) \\ &= 1 + t\mathbb{E}Y_i + \frac{1}{2}t^2\mathbb{E}(Y_i^2) + O(t^3) \\ &= 1 + \frac{1}{2}t^2 + O(t^3) \end{aligned}$$

where $O(h)$ means that $\lim_{h\to 0} O(h) = 0$.

Let $U_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$

$$\begin{aligned} M_{U_n}(t) &= \mathbb{E}\left(e^{t\cdot\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}\right) \\ &= \mathbb{E}\left(e^{t\cdot\frac{\sqrt{n}\bar{X}_n - \sqrt{n}\mu}{\sigma}}\right) \\ &= \mathbb{E}\left(e^{\frac{t}{\sqrt{n}}\cdot\frac{n\bar{X}_n - n\mu}{\sigma}}\right) \\ &= \mathbb{E}\left(e^{\frac{t}{\sqrt{n}}\cdot\frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma}}\right) \\ &= \mathbb{E}\left(e^{\frac{t}{\sqrt{n}}\cdot\frac{X_1 - \mu}{\sigma}} e^{\frac{t}{\sqrt{n}}\cdot\frac{X_2 - \mu}{\sigma}} \dots e^{\frac{t}{\sqrt{n}}\cdot\frac{X_n - \mu}{\sigma}}\right) \\ &= \prod_{i=1}^n \mathbb{E}\left(e^{\frac{t}{\sqrt{n}}\cdot\frac{X_i - \mu}{\sigma}}\right) \\ &= \prod_{i=1}^n M_{Y_i}\left(\frac{t}{\sqrt{n}}\right) \\ &= \left(1 + \frac{1}{2}\cdot\frac{t^2}{n} + O\left(\frac{t^3}{n^{3/2}}\right)\right)^n \end{aligned}$$

$$n \to \infty \implies M_{U_n}(t) = e^{\frac{1}{2}t^2}$$

which is the mgf of the $N(0,1)$.

Aside:

$$(1+ax)^{\frac{1}{x}} \xrightarrow{x \to 0} e^a$$

E.g.: 2.50: X is the number of times a fair coin flipped 40 times, lands heads up. Find the probability that $X = 20$ via the normal approximation.

$X = n\bar{X}_n \sim$ sum of n iid Bern(p)

$$\begin{aligned} &\overset{\text{CLT}}{\implies} \frac{\bar{X}_n - p}{\sqrt{p(1-p)/n}} \to N(0,1) \\ &\implies \frac{X - np}{\sqrt{np(1-p)}} \to N(0,1) \end{aligned}$$

$$\begin{aligned} \mathbb{P}(X = 20) &\approx \mathbb{P}(19.5 < X \leq 20.5) \\ &\approx \mathbb{P}\left(\frac{19.5 - 20}{\sqrt{10}} < Z \leq \frac{20.5 - 20}{\sqrt{10}}\right) \\ &= \Phi(0.1581) - \Phi(-0.1581) = 0.1256 \end{aligned}$$

where the true probability is 0.1268.