

# Math 180A: Introduction to Probability

Lecture A00 (Au)

[math.ucsd.edu/~bau/w21.180a](http://math.ucsd.edu/~bau/w21.180a)

Lecture B00 (Nemish)

[math.ucsd.edu/~ynemish/teaching/180a](http://math.ucsd.edu/~ynemish/teaching/180a)

Today: ASV 9.2, 9.3

Video: Prof. Todd Kemp, Fall 2019

Next: Additional office hours on Wed/Fri

Week 10: Quiz 5 (Wednesday, Mar 10 on Lectures 17-20)

Homework 8 (due Friday, Mar 12)

## Reminder: Chebychev's Inequality

9.1

For any random variable  $X$  with finite

$$\mathbb{E}(X) = \mu \quad \text{Var}(X) = \sigma^2$$

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

$$\Rightarrow \mathbb{P}(|X - 1000| \geq 3\sigma) \leq \frac{1}{3^2} = \frac{1}{9}$$

We proved this using the fact that  $\mathbb{E}$  is monotone:

$$\begin{aligned} X \leq Y &\Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y) \\ \text{Pf. } \mathbb{E}(Y) - \mathbb{E}(X) &= \mathbb{E}(Y) + \mathbb{E}(-X) \\ &= \mathbb{E}(Y - X) \end{aligned}$$

$\Rightarrow Y - X \geq 0$  or  $\int_{-\infty}^{\infty}$   
 $\therefore \mathbb{E}(Y - X) \geq 0$

E.g. Ramen Menya Ultra has, on average, 1000 customers/day, with a standard deviation of 15. Estimate the probability that today they will have between 956 and 1044 customers. #customers =  $X$

$$\begin{aligned} \mathbb{P}(956 \leq X \leq 1044) &= \mathbb{P}(955 < X < 1045) \Rightarrow \mathbb{P}(|X - 1000| < 3\sigma) \\ &= \mathbb{P}(-4\sigma < X - 1000 < 4\sigma) \Rightarrow 1 - \mathbb{P}(|X - 1000| \geq 3\sigma) \geq 1 - \frac{1}{9} = \frac{8}{9} \end{aligned}$$

## (Weak) Law of Large Numbers

Let  $X_1, X_2, X_3, \dots, X_n, \dots$  be an infinite sequence of i.i.d. random variables, each with  $\mathbb{E}(X_j) = \mu$  and  $\text{Var}(X_j) = \sigma^2$  finite.

Let  $S_n = X_1 + X_2 + \dots + X_n$ . Then for any fixed  $\varepsilon > 0$ ,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = O\left(\frac{1}{n}\right)$$

Pf. Let  $\bar{X}_n = S_n/n$ .

$$\begin{aligned} \mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) &\stackrel{\text{want}}{\leq} \frac{1}{k^2} \\ \mathbb{P}(|\bar{X}_n - \mu| \geq k \cdot \sqrt{\frac{\sigma^2}{n}}) &\stackrel{\text{II}}{\leq} \frac{1}{k^2} \\ \frac{1}{(k \cdot \sqrt{\frac{\sigma^2}{n}})^2} &= \frac{\frac{1}{n}}{\varepsilon^2} \end{aligned}$$

$$\begin{aligned} \mathbb{E}(\bar{X}_n) &= \mathbb{E}\left(\frac{1}{n} S_n\right) = \frac{1}{n} \mathbb{E}(X_1 + \dots + X_n) \\ &= \frac{1}{n} (\mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n)) \\ &= \frac{1}{n} (n\mu + \mu + \dots + \mu) \\ &= \mu. \\ \text{S.D.}(\bar{X}_n) &= \sqrt{\text{Var}(\bar{X}_n)} = \sqrt{\frac{1}{n^2} \text{Var}(S_n)} \\ &\stackrel{\text{II}}{=} \sqrt{\frac{1}{n^2} \cdot \sigma^2 n} \\ &= \frac{\sigma}{\sqrt{n}}. \\ \text{Var}(\bar{X}_n) &= \text{Var}\left(\frac{1}{n} S_n\right) = \frac{1}{n^2} \text{Var}(S_n) \\ &\stackrel{\text{II}}{=} \frac{1}{n^2} \text{Var}(X_1 + X_2 + \dots + X_n) \\ &= \frac{1}{n^2} (\text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)) \\ &= \frac{1}{n^2} (\sigma^2 + \sigma^2 + \dots + \sigma^2) \end{aligned}$$

E.g. The Large Hadron Collider was built to detect and measure the mass of the Higgs Boson. Call the mass  $M$ .

For theoretical reasons, it is known that  $M \leq 1.78 \times 10^{-23} \text{ g}$ .

How many trials do the LHC physicists need to do to estimate the correct mass (via sample mean) within  $10^{-24} \text{ g} \leq \varepsilon$ , with probability  $\geq 95\%$ ?

Trial measurements  $M_1, M_2, M_3, \dots, M_n, \dots$   $\bar{M}_n = \frac{M_1 + \dots + M_n}{n}$

Want  $P(|\bar{M}_n - \mu| < \varepsilon) \geq 95\%$

$$\text{WLLN: } P(|\bar{M}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \cdot \frac{1}{n}.$$

$$\therefore P(|\bar{M}_n - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2} \cdot \frac{1}{n}. \quad \begin{matrix} \text{want} \\ \geq 95\% \end{matrix}$$

$$\therefore \frac{\sigma^2}{\varepsilon^2} \cdot \frac{1}{n} \leq 5\% \quad n \geq \frac{\sigma^2}{\varepsilon^2} \cdot 20.$$

Fact: (HW)

$$\text{If } |X| \leq C$$

$$\Rightarrow \sqrt{\text{var}(X)} \leq C$$

$$\geq \frac{(1.78 \times 10^{-23})^2}{(10^{-24})^2} \cdot 20 \\ = 6337$$

## Strong Law of Large Numbers

Let  $X_1, X_2, X_3, \dots, X_n, \dots$  be an infinite sequence of i.i.d. random variables each with  $E(X_j) = \mu$ .

Let  $\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ . Then

$$P\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1.$$

Beautiful Proof uses existence of  $E(X_n^4) < \infty$ .

The Law of Large Numbers says that if  
 $X_1, X_2, X_3, \dots$  are iid. with mean  $\mu$ , and  $S_n = X_1 + \dots + X_n$ ,  
then

$$\frac{S_n}{n} \rightarrow \mu \quad \text{i.e.} \quad \frac{S_n - n\mu}{n} \rightarrow 0$$

$$\text{i.e. } S_n - n\mu = o(n).$$

At what exact rate  $n^\alpha$  does  $S_n - n\mu = O(n^\alpha)$ ?

## Central Limit Theorem

Let  $X_1, X_2, \dots, X_n, \dots$  be iid random variables with  $E(X_j) = \mu$ ,  $\text{Var}(X_j) = \sigma^2$ .

Then  $S_n - n\mu = O(\sqrt{n})$ , and for  $-\infty < a \leq b < \infty$ ,

$$\lim_{n \rightarrow \infty} P\left(\underbrace{\frac{S_n - n\mu}{\sigma\sqrt{n}}}_{\sim N(0, 1)} \in [a, b]\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \Phi(b) - \Phi(a).$$

$$\sigma\sqrt{n} = \sqrt{\sigma^2 n} = \sqrt{\text{Var}(S_n)} = \sqrt{\text{Var}(S_n - n\mu)}$$

Proof. Let  $Y_n = (S_n - n\mu) / \sqrt{n}$ . We will show that

$$M_{Y_n}(t) \rightarrow M_{N(0,1)}(t) = e^{t^2/2} \text{ for all } t \in \mathbb{R}.$$

$$\begin{aligned} M_{Y_n}(t) &= E(e^{tY_n}) = E\left(e^{t \cdot \frac{1}{\sqrt{n}}(S_n - n\mu)}\right) \\ &= E\left(e^{\frac{t}{\sqrt{n}}(\overset{\circ}{X}_1 + \overset{\circ}{X}_2 + \dots + \overset{\circ}{X}_n)}\right) \\ &= E\left(e^{\frac{t}{\sqrt{n}}\overset{\circ}{X}_1} e^{\frac{t}{\sqrt{n}}\overset{\circ}{X}_2} \dots e^{\frac{t}{\sqrt{n}}\overset{\circ}{X}_n}\right) \\ &\quad \uparrow \quad \uparrow \quad \dots \quad \uparrow \\ &\quad \text{independent.} \\ &= E\left(e^{\frac{t}{\sqrt{n}}\overset{\circ}{X}_1}\right) \dots E\left(e^{\frac{t}{\sqrt{n}}\overset{\circ}{X}_n}\right) \end{aligned}$$

$$\begin{aligned} S_n - n\mu &= (\overset{\circ}{X}_1 + \overset{\circ}{X}_2 + \dots + \overset{\circ}{X}_n) - (n\mu) \\ &= \overset{\circ}{X}_1 + \overset{\circ}{X}_2 + \dots + \overset{\circ}{X}_n \\ \overset{\circ}{X}_j &= X_j - \mu, \\ i.e., \quad E(\overset{\circ}{X}_1) &= 0 \end{aligned}$$

$$\begin{aligned} M_{\overset{\circ}{X}_1}(s) &= 1 + \cancel{E(\overset{\circ}{X}_1)} s + \frac{E(\overset{\circ}{X}_1^2)}{2!} s^2 + \frac{E(\overset{\circ}{X}_1^3)}{3!} s^3 + \dots \\ &= M_{\overset{\circ}{X}_1}\left(\frac{s}{\sqrt{n}}\right)^n \end{aligned}$$

$$\begin{aligned} M_{\overset{\circ}{X}_1}\left(\frac{s}{\sqrt{n}}\right) &= 1 + \frac{\sigma^2}{2} \frac{s^2}{n} + O\left(\frac{1}{n^{3/2}}\right) \Rightarrow M_{\overset{\circ}{X}_1}\left(\frac{s}{\sqrt{n}}\right)^n = \frac{1 + \frac{\sigma^2}{2} s^2 + O(s^3)}{\left(1 + \frac{\sigma^2}{2n} + O(n^{-3/2})\right)^n} \end{aligned}$$