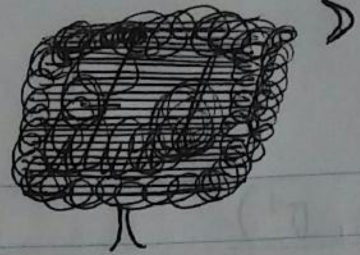


November 29, Lecture 20

- Defined moment generating function for random variable X:
  $\underline{M_X(t) = E[e^{tX}]}$, recall $e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \cdots$

  Taylor Series Expansion

Properties

Ⅰ If $M_X(t) = M_Y(t)$ this means $X \overset{d}{=} Y$

Ⅱ $M_X^{(k)}(0) = E[X^k]$ (aka k^{th} moment)

Ⅲ $Y = aX + c \Rightarrow$ ~~$M_Y(t) = e^{tc}$ M~~ $M_Y(t) = e^{tc} M_X(at)$

Ⅳ If X, Y independent
$M_{X+Y}(t) \overset{\downarrow}{=} M_X(t) M_Y(t) = (M_X(t))^2$

Ⅴ $\lim_{n\to\infty} M_{X_n}(t) = M_Y(t) \Leftrightarrow X \to Y$
This random variable becoming more like Y

* If n large $\Rightarrow X_n \overset{d}{\approx} Y$
  (X_n is approx. equally distributed as Y)

- $X \sim Bern(p) \Rightarrow M_X(t) = 1 - p + pe^t$
  $X \sim Binomial(n, p) \Rightarrow M_X(t) = (1 - p + pe^t)^n$ [comes from fact #4]
  $X \sim Geometric(p) \Rightarrow M_X(t) = \frac{pe^t}{1 - e^t(1-p)}$ [comes from definition] if $t < \ln\left(\frac{1}{1-p}\right)$

$X \sim Exp(\lambda) \Rightarrow M_X(t) = \frac{\lambda}{\lambda - t}$ if $t < \lambda$

$Z \sim Normal(0,1) \Rightarrow$ ~~$M_X(t) = e^{t^2}$~~ $M_X(t) = e^{t^2/2}$

$X \sim Normal(\mu, \sigma^2) \Rightarrow M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$
$X \sim Deg(c) = M_X(t) = e^{tc}$

(LLN): $\bar{X}_n$ average random variable converges towards the mean

- If $X_1, \ldots, X_n$ idd some distribution with mean μ, then $\bar{X}_n \to \mu$. *

* $\mu \sim Deg(\mu)$

$\rightarrow \lim_{n\to\infty} M_{\bar{X}_n}(t) = e^{t\mu}$

→ Proof:

$$M_{\bar{X}_n}(t) = \underbrace{M_{\frac{T}{n}}(t)}_{\text{got this from fact III.}} = M_T\left(\frac{t}{n}\right)$$

$$= \underbrace{\prod_{i=1}^{n} M_{X_i}\left(\frac{t}{n}\right)}_{\text{fact IV}} = \underbrace{\left(M_X\left(\frac{t}{n}\right)\right)^n}_{\text{fact IV}} = \underbrace{\left(E\left[e^{\frac{t}{n}X}\right]\right)^n}_{\text{definition of mgf}} = \underbrace{\left(E\left[1+\frac{tX}{n}+\frac{t^2X^2}{2!n^2}+\frac{t^3X^3}{3!n^3}+\ldots\right]\right)^n}_{\text{Taylor series expansion } e^{tX}}$$

♥ We say $f(n) = o(g(n))$ "little o" if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$ [$g(n)$ gets larger "faster" than $f(n)$]

$n^2 = o(n^2)$ $\lim \frac{n^2}{n^3} = \lim \frac{1}{n} = 0$ ✓

$$\frac{t^2X^2}{2!n^2} + \frac{t^3X^3}{3!n^3} + \ldots = o\left(\frac{1}{n}\right)$$

Proof: $\lim \dfrac{\frac{t^2X^2}{2!n^2} + \frac{t^3X^3}{3!n^3} + \ldots}{\frac{1}{n}} = \lim \frac{t^2X^2}{2!n} + \frac{t^3X^3}{3!n^2} + \ldots = 0$

$$\left(E\left[1+\frac{tX}{n}+\frac{t^2X^2}{2!n^2}+\frac{t^3X^3}{3!n^3}+\ldots\right]\right)^n = \left(E\left[1+\frac{tX}{n}+o\left(\frac{1}{n}\right)\right]\right)^2 = \left(1+\frac{t\mu}{n}+E\left[o\left(\frac{1}{n}\right)\right]\right)^n$$

$$= \left(1+\frac{t\mu}{n}+o\left(\frac{1}{n}\right)\right)^n$$

Want to show $\lim_{n\to\infty}\left(1+\frac{t\mu}{n}+o\left(\frac{1}{n}\right)\right)^n = e^{t\mu}$. Remember $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$ & $\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$. Does $\underbrace{\lim\left(1+\frac{a}{n}+o\left(\frac{1}{n}\right)\right)^n = e^a}_{\text{yes}}$?

$\lim_{n\to\infty}\left(1+\frac{1}{n}+\frac{1}{n^{1.1}}\right)^n$ $n = 1$ billion $\lim = 3.08 \neq 2.718$

$n = 1$ trillion $\lim = 2.90 \neq 2.718$

↑ will get there very slowly, but will get there, $n^{1.1}$ does affect how fast limit grows.

$\frac{1}{n+1} = O\left(\frac{1}{n}\right) \Rightarrow \lim \frac{\frac{1}{n+1}}{\frac{1}{n}} = \lim \frac{n'}{n''} = \lim \frac{1}{n+1} = 0$ ✓

- let $X_1, \ldots, X_n \overset{iid}{\sim}$ with mean μ & Standard error σ. Consider

$C_n = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ Why is this an important fraction? $\bar{X}$ standardized to have the mean 0, Standard error 1.

$$* \; C_n = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \Rightarrow \frac{\sqrt{n}\left(\frac{X_1 + \ldots + X_n}{n} - \mu\right)}{\sigma} \Rightarrow \frac{\sqrt{n}\left(\frac{X_1 + \ldots X_n}{n} - \frac{n\mu}{n}\right)}{\sigma}$$

$$\Downarrow$$

$$* \; \frac{1}{\sqrt{n}}\left(\frac{X_1-\mu}{\sigma} + \frac{X_2-\mu}{\sigma} + \ldots\right) \Leftarrow \frac{(X_1-\mu)+(X_2-\mu)+\ldots(X_n-\mu)}{\sigma\sqrt{n}} \Leftarrow \frac{\sqrt{n}\,\frac{X_1-\mu+X_2-\mu+\ldots X_n-\mu}{n}}{\sigma}$$

let Z_i be $\frac{X_i-\mu}{\sigma}$

$$\boxed{\frac{1}{\sqrt{n}}(Z_1 + Z_2 + \ldots + Z_i)}$$

$E[Z_i] = 0$
$SE[Z_i] = 1$

When you take a random variable & standardize it, always get mean 0 & standard error 1

- $\lim_{n\to\infty} M_{C_n}(t) =$

$M_{C_n}(t) =$ ~~$M_{\frac{1}{\sqrt{n}}(Z_1+\ldots+Z_n)}t$~~

$$M_{\frac{1}{\sqrt{n}}(Z_1 + \ldots + Z_n)}(t) \underset{\text{by III}}{=} M_{Z_1+\ldots+Z_n}\left(\frac{t}{\sqrt{n}}\right) \underset{\text{by IV}}{=} M_{Z_1}\left(\frac{t}{\sqrt{n}}\right) * \ldots * M_{Z_n}\left(\frac{t}{\sqrt{n}}\right) =$$

$$= \left(M_Z\left(\tfrac{t}{\sqrt{n}}\right)\right)^n = \left(E\left[e^{\frac{t}{\sqrt{n}}Z}\right]\right)^n = \left(E\left[1 + \frac{tZ}{\sqrt{n}} + \underbrace{\frac{t^2Z^2}{2!\,n}}_{\substack{\text{goes to 0}\\ \text{as a constant of } \frac{1}{n}}} + \underbrace{\frac{t^3Z^3}{3!\,n^{3/2}} + \frac{t^4Z^4}{4!\,n^2} + \cdots}_{\substack{\text{goes to 0 quicker than } \frac{1}{n}\\ o\left(\frac{1}{n}\right)}}\right]\right)^n$$

$$= \left(E\left[1 + \frac{tZ}{\sqrt{n}} + \frac{t^2Z^2}{2n} + o\left(\frac{1}{n}\right)\right]\right)^n$$

$$\Downarrow$$

$$\left(1 + \left(E\left[\frac{tZ}{\sqrt{n}}\right] + E\left[\frac{t^2Z^2}{2n}\right] + E\left[o\left(\frac{1}{n}\right)\right]\right)\right)^n = \left(1 + \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

$$\frac{t}{\sqrt{n}}E[Z]=0 \qquad \frac{t^2}{n}E[Z^2]=1$$

$$\lim_{n\to\infty} M_{Cn}(t) = \lim_{n\to n}\left(1 + \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

$$\boxed{= e^{t^2/2}} \Rightarrow C_n \to N(0,1)$$

the mgf of the standard normal

$Var[X] = E[X^2] - \mu^2$

if $\mu^2 = 0$

$Var[X] = E[X^2]$

if $Var(X) = 1$

$E[X^2] = 1$

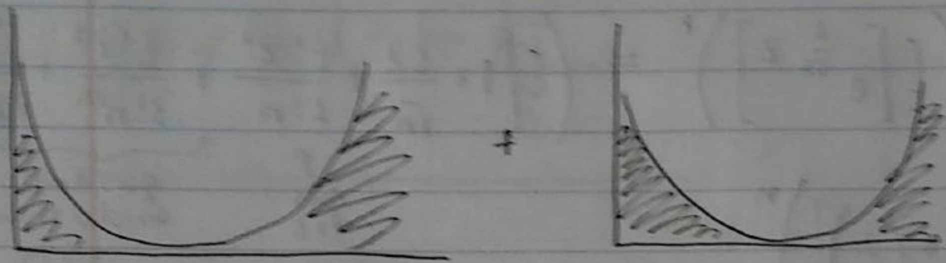
- CLT (Central Limit Theorem): (I) If $X_1, \ldots X_n \overset{iid}{\sim}$ with mean μ & s.e. σ

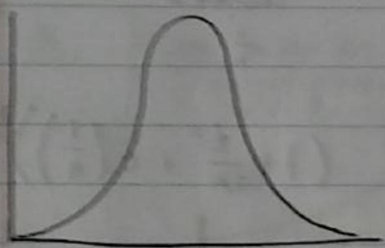
then $\dfrac{\bar{X}_1 - \mu}{\frac{\sigma}{\sqrt{n}}} \longrightarrow Z \sim N(0,1)$.

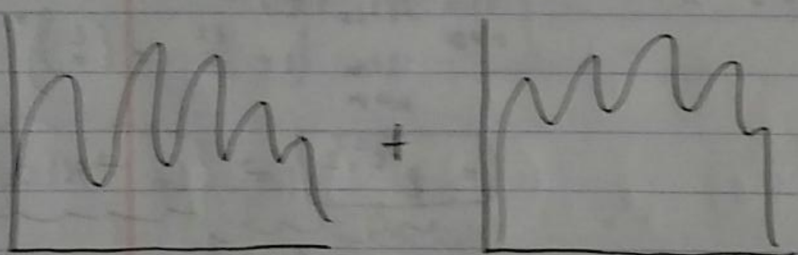
* If n is large, then $\underbrace{\dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \overset{d}{\approx} Z \sim N(0,1)}_{\text{II}} \Rightarrow \underbrace{\bar{X} \approx N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)}_{\text{III}}$ $(\bar{X}=$

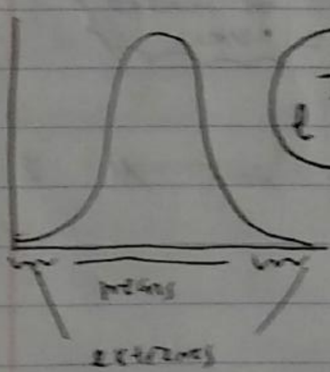
$$\Rightarrow \underbrace{T \overset{d}{\approx} N\left(n\mu, (\sqrt{n}\,\sigma)^2\right)}_{\text{IV}}$$

$$\boxed{X = \mu + \sigma Z \sim N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}(X-\mu)^2}}$$

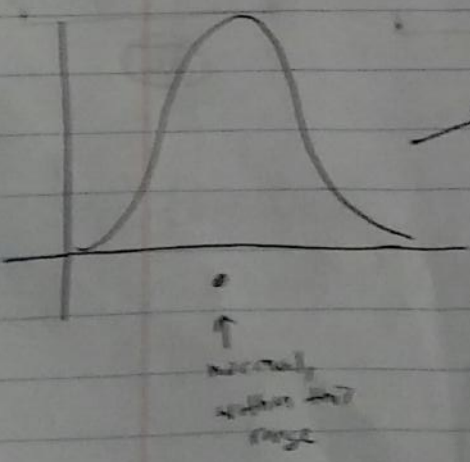
-  + ...

$\approx$  $e^{-\frac{x^2}{2}}$

-  + . . . $\approx$



$e^{-\frac{x^2}{2}}$ ← fn. able to harmonize means & extremes

- 

→ $\boxed{\text{called normal b/c defines what normality is}}$

· ← abnormal

• Example: $X_1, \dots, X_{30} \overset{iid}{\sim}$ Geom(p), $p = \frac{1}{2}$, What is the probability that the average realization is more than 2.75?

$$P(\bar{X} > 2.75)$$

$$\bar{X} \overset{d}{\approx} N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$$

$$\mu = \frac{1}{p} = \underline{\underline{2}} \leftarrow \left(\frac{1}{\frac{1}{2}} = 2\right)$$

$$\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{3} = 1.414$$

$$\frac{\sigma}{\sqrt{n}} = \frac{1.414}{\sqrt{30}} \approx .258$$

by the CLT,

$$\bar{X} \overset{d}{\approx} N(2, .258^2)$$

$$\Downarrow$$

$$\approx P\left(\frac{\bar{X} - 2}{.258} > \frac{2.75 - 2}{2.58}\right)$$

$$\Downarrow$$

$$\approx P(Z > 3)$$

$$\Downarrow$$

$$\boxed{0.15\%}$$

• Example: Random walk 100 steps, what is probability you are more than 10 steps away from where you started?

Let $T = X_1 + X_2 + \dots + X_{100}$

where $X_1 \dots X_{100} \overset{iid}{\sim} \begin{cases} 1 & \text{wp } \frac{1}{2} \\ -1 & \text{wp } \frac{1}{2} \end{cases}$

From the CLT we know that

$$T \overset{d}{\approx} N(n\mu, (\sqrt{n}\,\sigma)^2)$$

$$\mu = 0$$

$$\sigma = \underline{1}$$

T is now $\Rightarrow T \overset{d}{\approx} N(0, 10^2)$

$$P(|T| > 10) = P(T < -10) + P(T > 10)$$

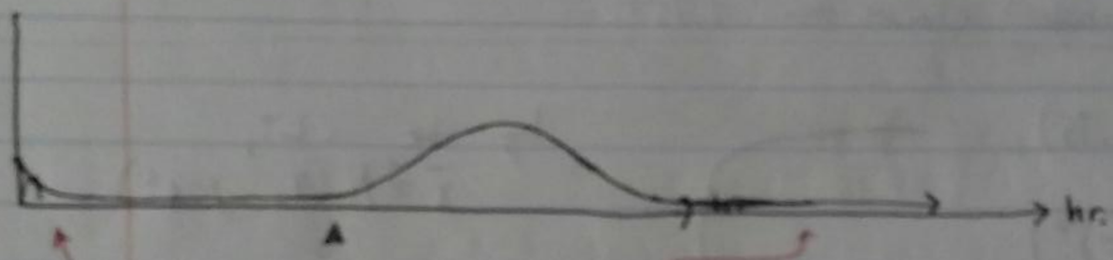
$$\cancel{= P\left(\frac{T-0}{10} < \frac{10-0}{10}\right)} +$$

$$= P\left(\frac{T-0}{10} < \frac{-10-0}{10}\right) + P\left(\frac{T-0}{10} > \frac{10-0}{10}\right)$$

$$= P(Z < -1) + P(Z > 1) = P(Z \notin [-1, 1]) = \boxed{32\%}$$

Key words usually average or total

• <u>Example</u>: Lightbulb lifetime



$\mu = 1000$ hr

$\sigma = 5000$ hr

density function : $f(t)$

This is not normal

long right tail

We buy 50 bulbs. What is probability the $\boxed{\text{average}}$ bulb lasts more than 1700 hours?

$$\bar{X} \approx N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \quad \text{by CLT}$$

$$P(\bar{X} > 1300\text{hrs}) = P\left(\frac{\bar{X} - 1000}{\frac{500}{\sqrt{50}}} > \frac{1300 - 1000}{\frac{500}{\sqrt{50}}}\right)$$