$$Z \sim N(0,1) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$m_Z(t) = e^{t^2/2}$$

$$C_n := \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \ldots = \frac{Z_1}{\sqrt{n}} + \ldots + \frac{Z_n}{\sqrt{n}}$$

$X_1, \ldots, X_n \overset{iid}{\sim}$ Something w/ μ, σ^2

n large, how's C_n distributed

$$Z_i := \frac{X_i - \mu}{\sigma} = E[Z] = 0 \qquad E[Z^2] - \cancel{\mu^2}^{\,0} = 1$$

$$SE[Z] = 1 \Rightarrow Var[Z] = 1$$

$$M_{C_n}(t) \overset{\text{rule\#2}}{=} \left(M_{\frac{Z}{\sqrt{2}}}(t)\right)^n \overset{\text{rule\#3}}{=} \left(M_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n = *$$

sum

moment generate function

$$M_X(t) := E[e^{tX}] = E\Big[1 + \underbrace{\frac{tE[X]}{1!} + t^2\frac{E[X^2]}{2!} + \ldots}_{\text{Taylor exp. of } e^{tX}}\Big]$$

$$* = \left(1 + \frac{t}{\sqrt{n}}\frac{E[Z]}{1!} + \frac{t^2}{n}\frac{E[Z^2]}{2!} + \underbrace{\frac{t^3}{n^{3/2}}E[Z^3] + \frac{t^4}{n^2}\frac{E[Z^4]}{4!} + \ldots}_{\text{"tail(n)"}}\right)^n$$

$$= \left(1 + \frac{t^2/2}{n} + \text{tail}(n)\right)^n = \left(1 + \frac{t^2/2}{n} + o\left(\frac{1}{n}\right)\right)^n$$

"o"

If a $f(n) = o\left(\frac{1}{n}\right)$ "little-o" this mean

$\lim_{n\to\infty} \frac{f(n)}{\frac{1}{n}} = 0 \Rightarrow f(n)$ goes to 0 "quicklier" than $\frac{1}{n} \to 0$

$tail(n) = o\left(\frac{1}{n}\right)$

$$\lim_{n\to\infty} \frac{\frac{t^3}{n^{3/2}} E[Z^3] + \frac{t^4}{n^2} \frac{E[Z^4]}{4!} + \cdots}{\frac{1}{n}}$$

$$= \lim_{n\to\infty} \frac{t^3}{\sqrt{n}} E[Z^3] + \frac{t^4}{n} E[Z^4] + \cdots = 0$$

$e := \lim_{n\to\infty} \left(1 + \frac{1}{n} + \frac{1}{n^2}\right)^n$ if put things inside the lim it would be still the same.

└→ the lower it goes, it would affect even at $\frac{1}{n^{1.1}}$ is still e

anything lower than $\left(\frac{1}{n}\right)$ still gets e.

$$\lim_{n\to\infty} \left(1 + \frac{t^2/2}{n} + o\left(\frac{1}{n}\right)\right)^n = e^{t^2/2} \Rightarrow \boxed{C_n \overset{\text{appox}}{\sim} N(0,1)}$$

most important thing!

Property 0 match !
Same DNA, same person !

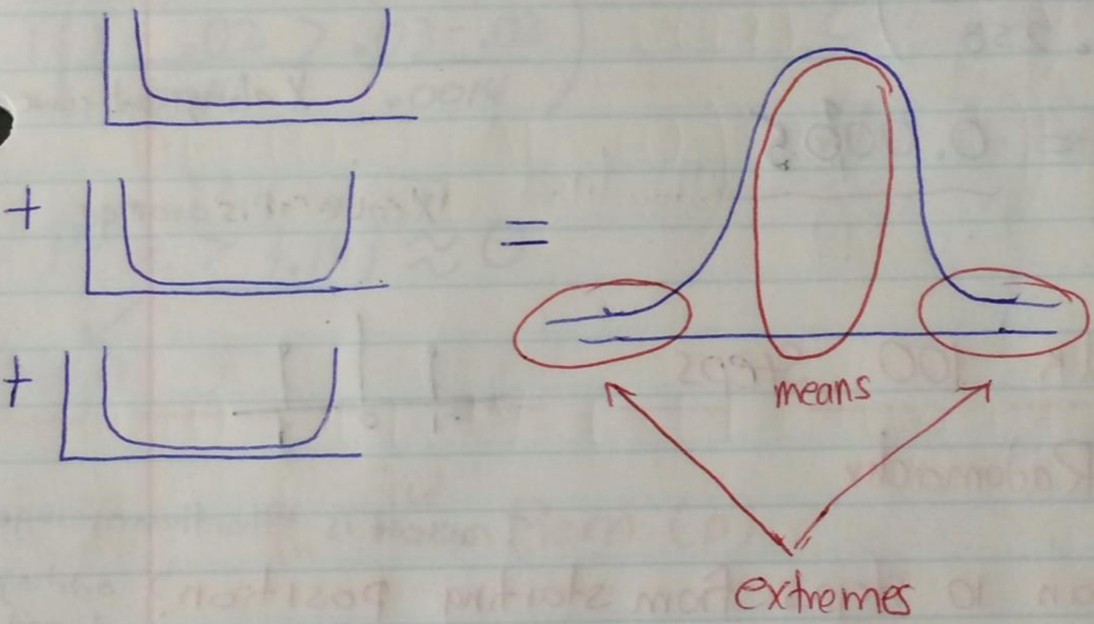
standarized $\bar{X} = N(0,1)$

Central Limit Theorm (CLT)

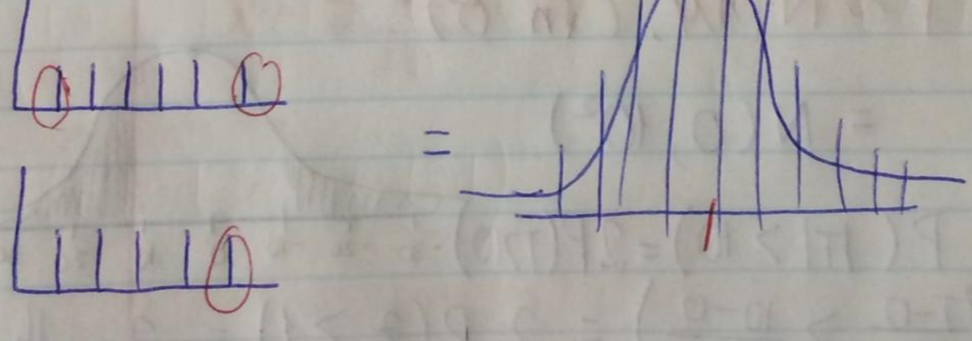
$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \quad \text{if } n \text{ large}$$

$$\bar{X} = \frac{\sigma}{\sqrt{n}} C_n + \mu \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$$

$$T = n\bar{X} \sim N\left(n\mu, (\sqrt{n}\,\sigma)^2\right)$$



$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$



Test Questions

- $X_1, \dots, X_{30} \overset{iid}{\sim} \text{Geo}\left(\frac{1}{2}\right)$

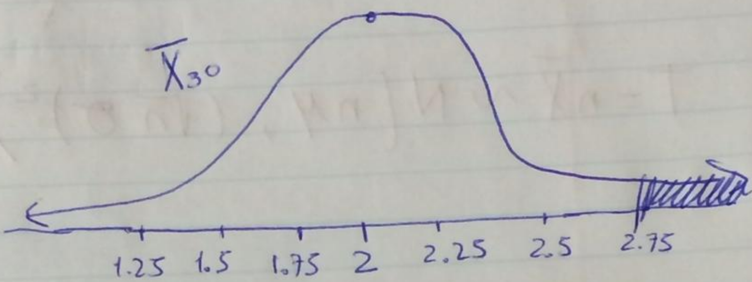
Geometric
$$\sigma = \sqrt{\frac{1 \cdot p}{p^2}} = \sqrt{\frac{1-\frac{1}{2}}{\left(\frac{1}{2}\right)^2}} = \sqrt{2}$$

for geometric
$$\mu = \frac{1}{p} = \frac{1}{\frac{1}{2}} = 2$$

$$\sigma/\sqrt{n} = \sqrt{2/30} \approx .258$$

Prob. of waiting more than 2.75 seconds on avarage?

$$\underbrace{\bar{X}_{30} \sim}_{\text{by CLT}} N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) = N(2, .258^2)$$



$$P(\bar{X}_{30} > 2.75)$$

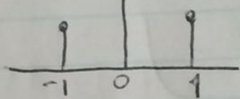
$$= P\left(\frac{\bar{X} - 2}{.258} > \frac{2.75 - 2}{.258}\right)$$

$$= P(Z > 3) = \underline{0.0015}$$

$\bar{X}$ always bell curve

$\bar{X}$ cause it is avarage

- Random Walk 100 Steps



$X_1, \dots, X_{100} \overset{iid}{\sim}$ Radomachy

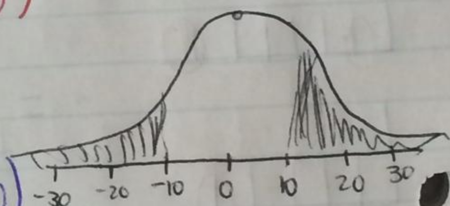
T cause it is other than average adding all + and − steps.

Prob. more than 10 steps from starting position?

$$T \sim N(n\mu, (\sqrt{n}\,\sigma)^2)$$

$$T = \sum X_i$$

$$= N(0, 10^2)$$



$$P(|T| > 10)$$

$$P(|T| > 10) = 2P(T > 10)$$

$$\mu = 0$$
$$\sigma = 1 \Rightarrow \sigma\sqrt{n} = 10$$

$$= 2P\Big(\underbrace{\frac{T-0}{10} > \frac{10-0}{10}}_{\text{standarized}}\Big) = 2P(Z > 1) = 2 \times .16 = \boxed{.32}$$

• Shipments are late 2% at the time. What is prob more than 3% are late (on avg) if 10,000 orders are shipped?

late or not late = Bernuli

$$X_1, \dots, X_{10000} \overset{iid}{\sim} \text{Bern}(.02)$$

$$P(\bar{X} > .03)$$

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

Var for bernuly

$$\mu = .02$$

$$\sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.2 \times .98}{10000}}$$

$$\sigma = .0014$$

$$= N(.02, .0014^2)$$

standarized both sides because want to get in terms of Z. To answer close to the rule he gave.

$$P\left(\frac{\bar{X} - .02}{.0014} > \frac{.03 - .02}{.0014}\right)$$

$$= \underset{N(0,1)}{P(Z > 7.14)} \approx 0$$

$$X_1, \dots, X_n \overset{iid}{\sim} \text{Bern}(p)$$

$\bar{X}$ $\quad \bar{x} = \frac{\sum 1\text{'s}}{n}$

"a proportion"

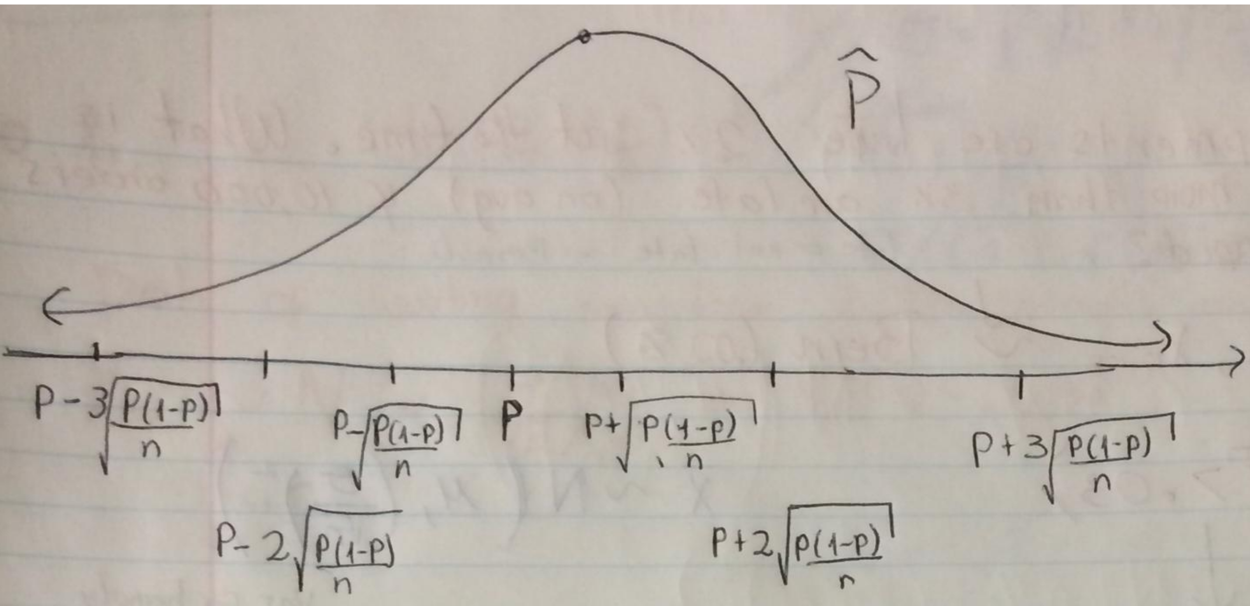
$\bar{X}$ $\quad \bar{x}$

$\hat{P}$ $\quad \hat{p}$

p btw 0 and 1

$$\hat{P} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$$

$\hat{p}$ is a realization from



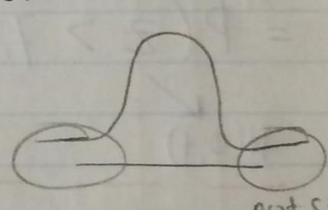
Moving to Statistics

Who likes mushrooms?

$X_1, \ldots, X_{10} \overset{iid}{\sim} \text{Bern}(p)$

$\hat{p} = \frac{8}{10} = .8$

Big things



not so likely to come from.

What is p?

The "true" percentage of mushroom lovers for the "Population proportion" of mushroom lovers.

Do we know p?
never know p

No. We cannot query the entire population

But we can "estimate" or "inter" p from a sample

p is know as a "parameter"

Best "point estimate" is $p = \hat{p}$

$\bar{X} \approx \mu$ by LLN

What about an integral estimate?