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

$M_Z(t) = e^{\frac{x^2}{2}}$

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

$Z_i = \frac{x_i - M}{\sigma} \rightarrow E(Z) = 0$

$SE(Z) = 1 \Rightarrow var[Z] = 1 \quad E(Z^2) - M^2 = 1$

$X_1 \dots X_n \overset{iid}{\sim}$ something with M, σ^2

$M_X(t) = E(e^{tX}) = E\underbrace{\left[1 + \frac{tE(X)}{1!} + \frac{t^2 E(X^2)}{2!}\right]}_{\text{taylor ser. of } e^{tx}}$

$M_{C_n}(t) = \left(M_{\frac{Z}{\sqrt{n}}}(t)\right) = \left(M_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n =$

$= \left(1 + \frac{t}{\sqrt{n}} \cdot \frac{E(Z)}{1!} + \frac{t^2}{n} \cdot \frac{E(Z^2)}{2!} + \frac{t^3}{n^{\frac{3}{2}}} \cdot \frac{E(Z^3)}{3!} + \frac{t^4}{n^2} \frac{E(Z^4)}{4!} + \dots\right)$

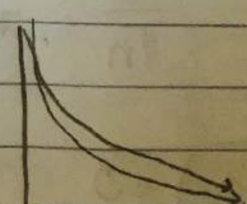
$= \left(1 + \frac{t^2/2}{n} + tail\right)^n \underset{\text{if } \lim_{n \to \infty}}{=} \left(1 + \frac{t^2/2}{n} + o\left(\frac{1}{n}\right)\right)^n \rightarrow = e^{\frac{t^2}{2}} \Rightarrow$

$\Rightarrow \boxed{C_n \overset{approx}{\sim} N(0,}$

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

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

$\frac{1}{n} \rightarrow 0$



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

- $$\lim_{n\to\infty} \frac{\frac{t^3}{n^{3/2}} E(Z^3) + \frac{t^4}{n^2} E(Z^4) + \cdots}{\frac{1}{n}} = \lim_{n\to\infty} \frac{t^3}{\sqrt{n}} E(Z^3) + \frac{t^4}{n} E(Z^4) + \cdots = 0$$

$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ same as $\left(1 + \frac{1}{n} + \frac{1}{n^2}\right)^n$ limit still be e

if it becomes $\frac{1}{n^{1.5}}$ & closer to $\frac{1}{n}$, it will affect e

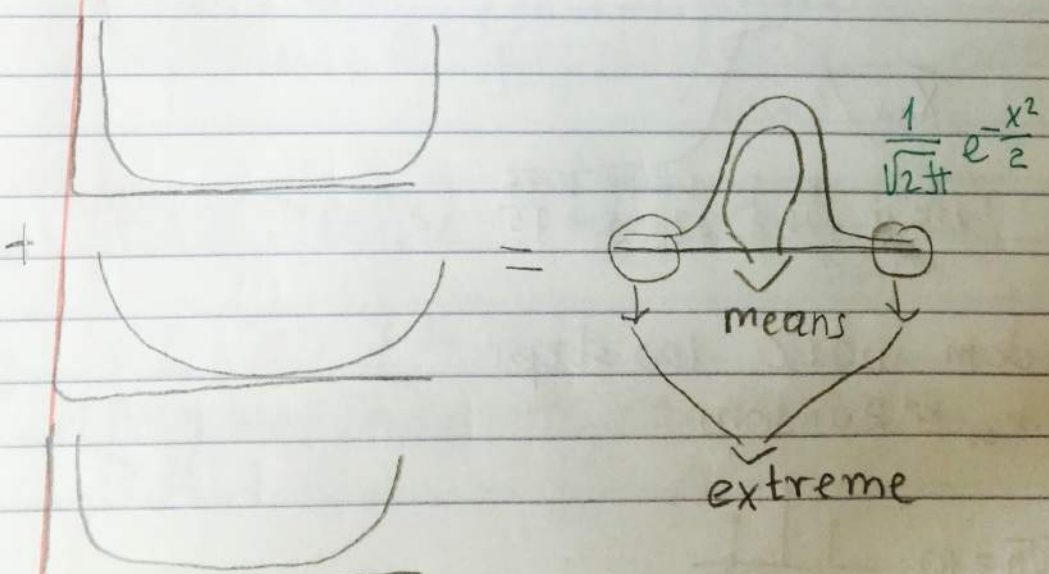
Everything that is slower than $\frac{1}{n}$ wont affect e

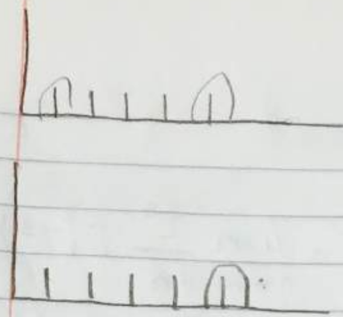
Central limit Theorem

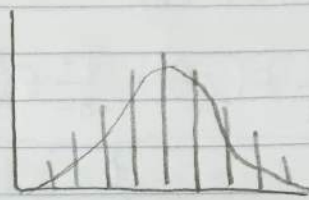
$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \text{ if n large}$$

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

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







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

- Prob of waiting more than 2.75 seconds on average?

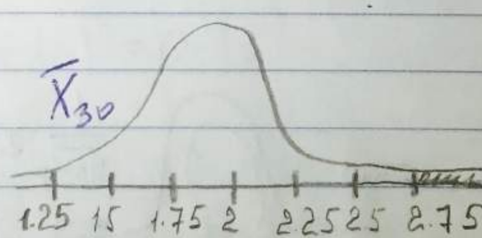
by the CLT

$\bar{X}_{30} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$

$P(\bar{X}_{30} > 2.75) = P\left(\frac{\bar{X} - 2}{.250} > \frac{2.75 - 2}{.258}\right) = P(Z > 3) = 0.0015$

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

$\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1-\frac{1}{2}}{\left(\frac{1}{2}\right)^2}} \rightarrow \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{2}{30}} = 2.58$

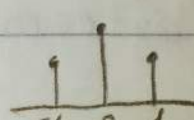


- Random walk 100 steps

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

$\mu = 0$

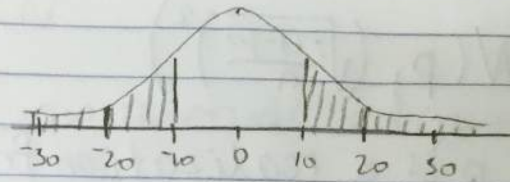
$\sigma = 1 \Rightarrow \sigma\sqrt{n} = 10$



Probab more than 10 steps from starting position?

$T = \sum X_i$ $\qquad T \overset{\text{by CLT}}{\sim} N(n\mu, (\sqrt{n}\,\sigma)^2) = N(0, 10^2)$

$P(|T| > 10)$



$P(|T|>10) = 2P(T>10) = 2 \cdot P\left(\frac{T-0}{10} > \frac{10-0}{10}\right) = 2P(Z>1) =$

$= 2 \cdot .16 = \boxed{0.32}$

- Shipments are late 2% of the time. What is probab more than 3% are late (on average) if 10,000 orders are shipped?

$X_1, \dots, X_{10000} \overset{iid}{\sim} \text{Bern}(0.02)$

$P(\bar{X} > 0.03)$

$\mu \overset{p=}{=} .02, \quad \sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.02 \cdot .30}{10000}} = .0014$

$P(\bar{X} > 0.03) = P\left(\frac{\bar{X} - .02}{.0014} > \frac{.03 - .02}{.0014}\right)$

$= P(Z > 7.14) \approx 0$

numbers > 4 or -4 are absurd

$N(0,1)$

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

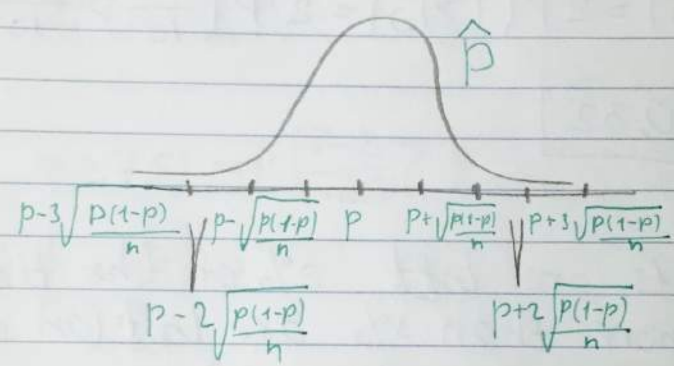
"a proportion"

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

$\hat{p}$ $\qquad \hat{p}$

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

p is a realization from



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

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

- What is p?

the "true" percentage of mushroom lovers or the "population proportion".

- Do we Know p? No. We cannot query the entire population.

But we can "estimate" or "infer" P from a sample.

P is Known as a "parameter"
Best "point estimate" is $p = \hat{p}$

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

What about an interval estimate?