example:

Take 100 steps with probability forward & backward being $\frac{1}{2}$

$X \sim \begin{cases} +1 & \text{wp } \frac{1}{2} \\ -1 & \text{wp } \frac{1}{2} \end{cases}$ What's the probability you are more than 10 steps away from starting position after 100 steps

$T = X_1 + \cdots + X_{100} \overset{d}{\underset{③}{\approx}} N(n\mu, (\sigma\sqrt{n})^2)$

$= N(0, 10^2) = P(|T| \geq 10)$

$\boxed{11/29}$

How to use CLT If $X_1, \ldots, X_n \overset{iid}{\sim}$ and n "large"

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

③ $T \overset{d}{\approx} N(n\mu, (\sigma\sqrt{n})^2)$

example:

$$X_1, \dots, X_{100} \overset{iid}{\sim} \begin{cases} 1 & wp\ 1/2 \\ -1 & wp\ 1/2 \end{cases} \Rightarrow \boxed{\mu = 0}, \boxed{\sigma^2 = 1}$$

$$\boxed{\sigma = 1}$$

What is the probability of being more than or equal 10 steps away from the origin after 100 steps?

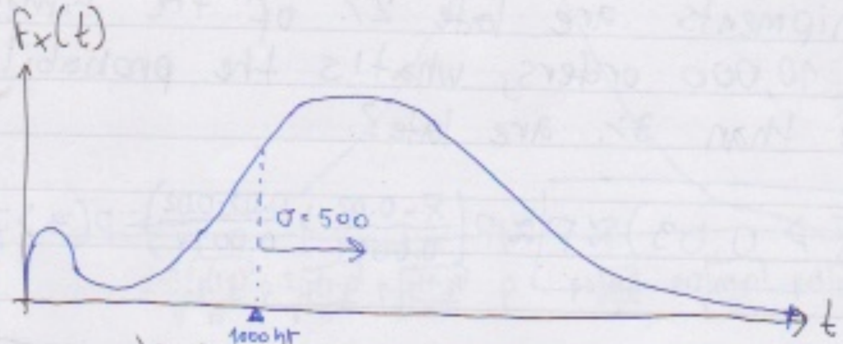
$$T = X_1 + X_2 + \dots + X_{100} \approx N(n\mu, (\sqrt{n}\sigma)^2) = N(0, 10^2)$$

$$P(T \geq 10 \text{ or } T \leq -10) = P(T \geq 10) + P(T \leq -10)$$

$$= P(|T| \geq 10) \underset{\substack{\uparrow \\ \text{standardise}}}{=} P\left(\frac{T-0}{10} \geq \frac{10-0}{10}\right) + P\left(\frac{T-0}{10} \leq \frac{-10-0}{10}\right) =$$

$$= P(Z \geq 1) + P(Z \leq -1) = 16\% + 16\% = \boxed{32\%}$$

example:



X= Lightbulb Survival

If you get 50 bulbs. What's the probability the average lifetime is more than 1300 hr?

$$P(\bar{X} > 1300) \approx P\left(\frac{\bar{X} - 1000}{70.7} > \frac{1300 - 1000}{70.7}\right) = P(Z > 4.24) \approx \boxed{0}$$

$$\bar{X} \overset{d}{\approx} N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) = N\left(1000, \left(\frac{500}{\sqrt{50}}\right)^2\right) = N\left(1000, 70.7^2\right)$$

by CLT

more example next page

example:

Shipments are late 2% of the times. In 10,000 orders, what's the probability more than 3% are late?

$$\boxed{P(\bar{X} > 0.03) \approx P} \approx P\left(\frac{\bar{X}-0.02}{0.0014} > \frac{0.03-0.02}{0.0014}\right) = P(Z > 7.14) \approx \boxed{0}$$

$$X_1, \ldots, X_{10{,}000} \overset{iid}{\sim} \text{Bern}(0.02) \Rightarrow \boxed{\mu = 0.02} \quad \boxed{\begin{aligned}\sigma &= \sqrt{0.02(1-0.02)} \\ &= 0.14\end{aligned}}$$

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

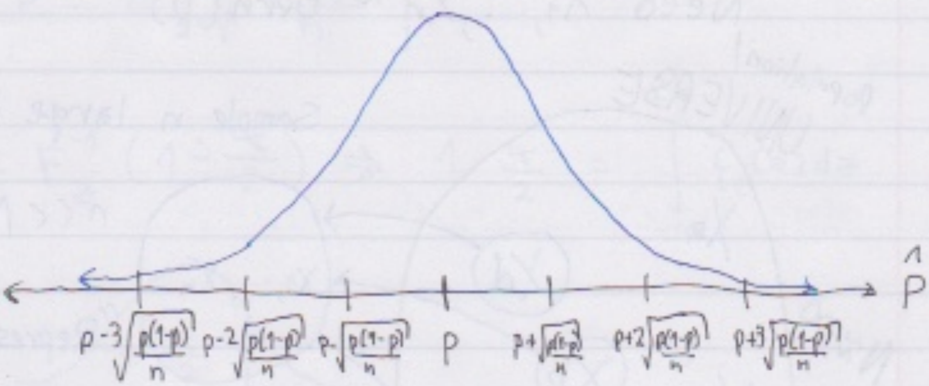
$$= N\left(0.02, \left(\frac{0.14}{\sqrt{10000}}\right)^2\right) = N(0.02, 0.0014^2)$$

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

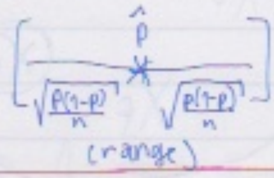
$$\text{estimate} \rightarrow \hat{p} = \underset{\text{little}}{\bar{X}} = \frac{1+1+0+0+0}{5} = 0.4 \approx p$$

sample proportion

$$\boxed{\hat{p} \overset{d}{\approx} N\Bigg(\underset{\text{centered at}}{p}, \Bigg(\underset{\text{standard deviation}}{\sqrt{\frac{p(1-p)}{n}}}\Bigg)^2\Bigg)} \leftarrow \text{Important}$$



Probability ↑



Statistics ↓

Just dealing with Bern now

$\hat{p} \overset{d}{\sim} N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$ ← This is boss

Statistical Inference

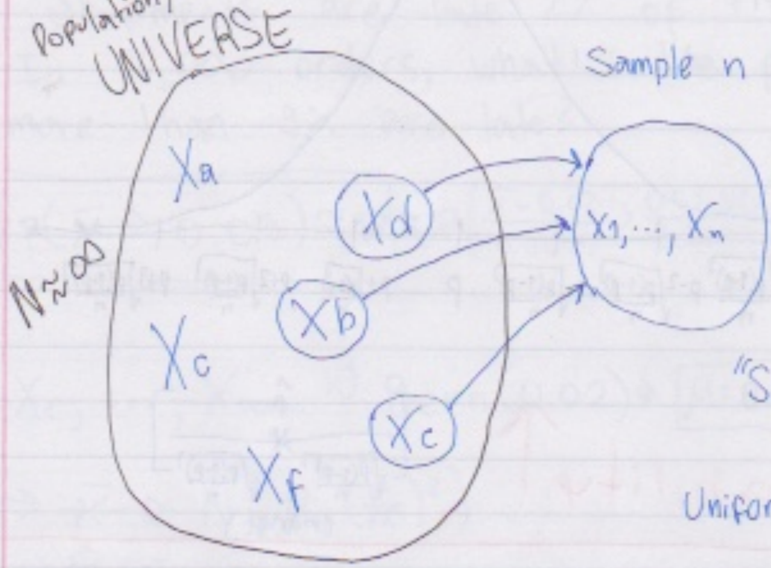
$\boxed{\text{3 Goals}}$:

I Estimate best guess of p ($\hat{p}$)

II Estimate range/window of p (Confidence interval)

III Test theories about p. (Hypothesis Testing)

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



Sample n large << ∞

$n << N \approx \infty$

"Representative Sample"

"Simple Random Sample"

prob.
Uniform discrete $\frac{1}{N}$

$$P\left(p \in \left[\hat{p} \pm \sqrt{\frac{p(1-p)}{n}}\right]\right)$$

$$= P\left(\hat{p} - \sqrt{\frac{p(1-p)}{n}} \leq p \leq \hat{p} + \sqrt{\frac{p(1-p)}{n}}\right)$$

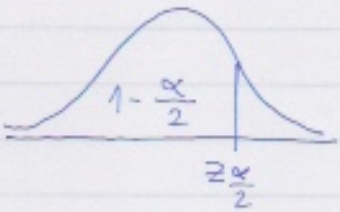
$$= P\left(-\sqrt{\frac{p(1-p)}{n}} \leq p - \hat{p} \leq \sqrt{\frac{p(1-p)}{n}}\right)$$

$$= P\left(-1 \leq \frac{p-\hat{p}}{\sqrt{\frac{p(1-p)}{n}}} \leq 1\right) = P(1 \geq Z \geq -1) = P(Z \in [-1,1])$$

$$= \boxed{.68}$$

$$\left[\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]$$

$$z_{\frac{\alpha}{2}} := F_Z^{-1}\left(1-\frac{\alpha}{2}\right) \Rightarrow 1-\frac{\alpha}{2} = \int_{-\infty}^{z_{\frac{\alpha}{2}}} f_Z(z)\,dz$$



$$\alpha = 5\% \Rightarrow \frac{\alpha}{2} = 2.5\% \Rightarrow 1-\frac{\alpha}{2} = 97.5\% \Rightarrow Z_{2.5\%} = Z$$

$$97.5\% = \int_{-\infty}^{Z_{2.5\%} = Z} f_Z(z)\,dz$$

conference Interval

$CI_{p,1-\alpha} :=$

100yr of debate

unknown

$$\left[\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}} \right] \approx \left[\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

When repeated ... gives you $1-\alpha$ "coverage" of p

2-sided 1-proportion confidence interval