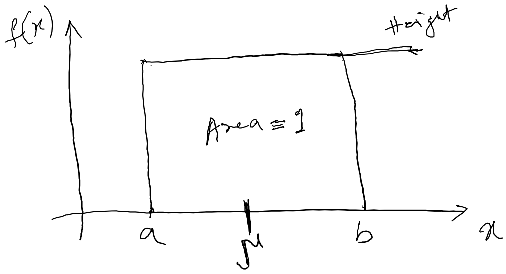
Continuous → Measured → Uniform, Normal (Z dist), Exponential

→ Area under the curve = 1 → $p(x)$

→ It can have ∞ values in a given Interval

Uniform



Area = $(b-a) \times$ Height

$1 = (b-a) \times$ Height

$$\boxed{\text{Height} = f(x) = \frac{1}{b-a}}$$

$$\boxed{\mu = \frac{a+b}{2}}$$

$$\boxed{SD = \sigma = \frac{b-a}{\sqrt{12}}}$$

$\mu = E(x)$

$\sigma^2 = E\left((x-\mu)^2\right)$

$= \boxed{E(x^2)} - \mu^2$

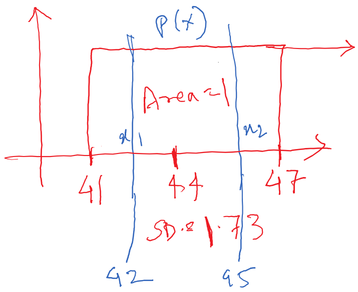
$E(x^2) = \int_a^b \frac{x^2}{b-a} \, dx$

$$\boxed{\Rightarrow \frac{1}{3}\left(a^2+ab+b^2\right)}$$

Derivation

$\sigma^2 = \frac{1}{3}\left(a^2+b^2+ab\right) - \left(\frac{a-b}{2}\right)^2$

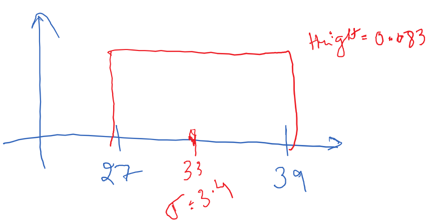
$\sigma^2 = \frac{1}{2}(b-a)^2 \Rightarrow \boxed{\sigma = \frac{b-a}{\sqrt{12}}}$



$$\boxed{P(x) = \frac{x_2 - x_1}{b-a}}$$

$a \leq x_1 \leq x_2 \leq b$

$P(x) = \boxed{42 \,\&\, 45} = \frac{42-42}{47-41} = \frac{3}{6} = 0.5$

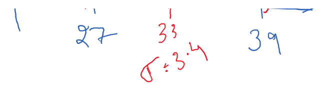


(1) Describe

$f(x) = 0.083$

$\mu = 33$

$SD = 3.4$



S.D = 3.4

③ Prob (≤ 30)

$$P = \frac{30-27}{39-27} = \frac{3}{12} = 0.25$$

② Prob (30 to 35)

$$P = \frac{35-30}{39-27} = 0.416$$

Normal Dist

→ Continuous dist

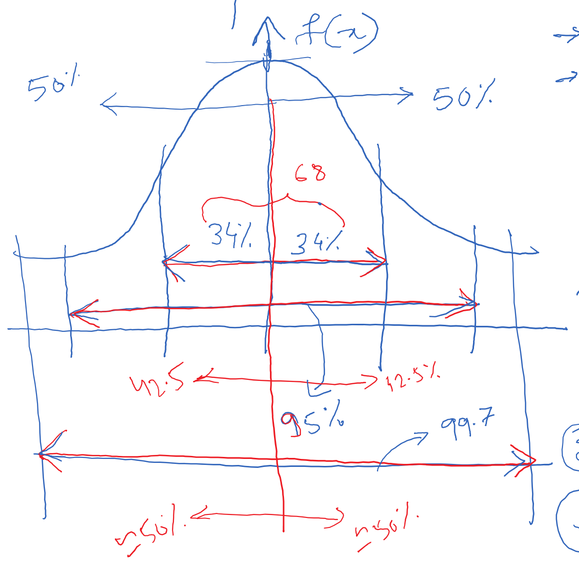
→ Symmetrix → μ

→ Asymptotic to X axis
−∞ to +∞

$\boxed{\mu = M_D = \text{Mode}}$

→ Area under curve = 1

→ Unimodal → single mode



Chebisher's ⇒ Inequality

① 68% 1 SD from mean

34 34

② 95% → 2 SD from mean

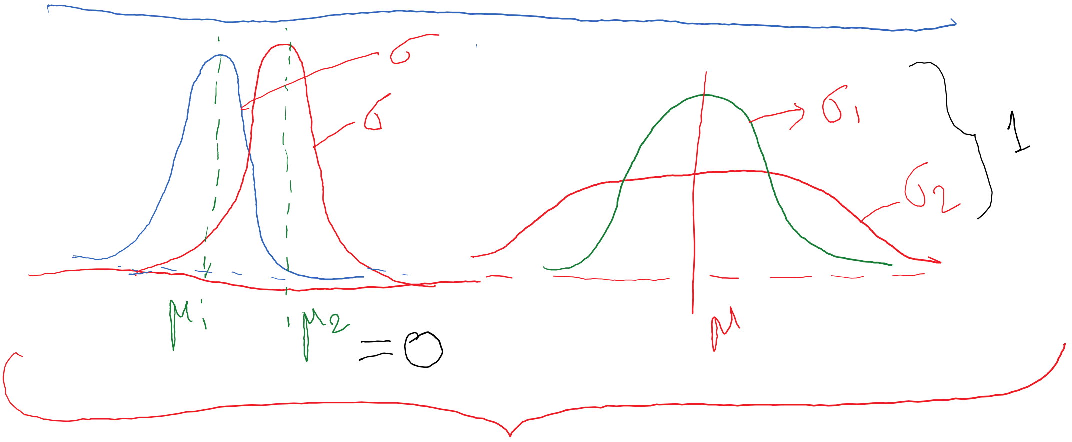
③ 99.7% → 3D from μ

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left[(x-\mu)/2\right]^2}}$$

$\pi = 3.14$

$e = 2.713$

Not practically usefull   σ, μ



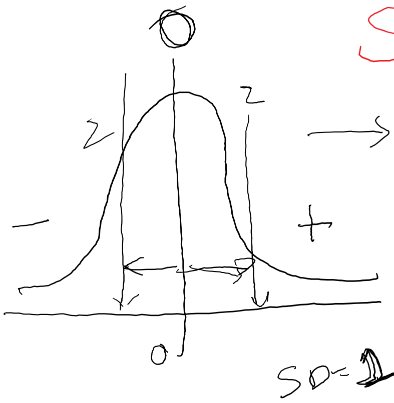
= 0

Standard

0
1

$x - \mu$

Standard



$\longrightarrow$ Z - Distribution

$$\text{Z Score} = \frac{x-\mu}{\sigma}$$

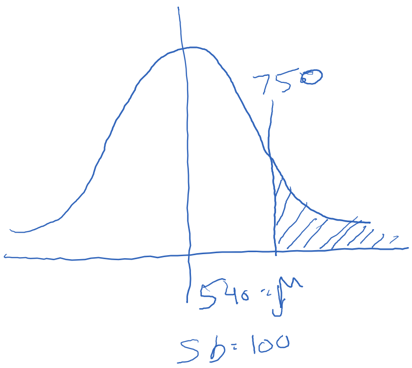
$= Z = 1.25$

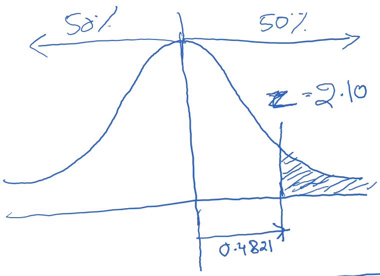
1.2 5

Probability $\longleftarrow$

$P(>750)$? $\mu = 540$, $SD = 100$

$$Z = \frac{x-\mu}{\sigma} = \frac{750-540}{100} = 2.10$$

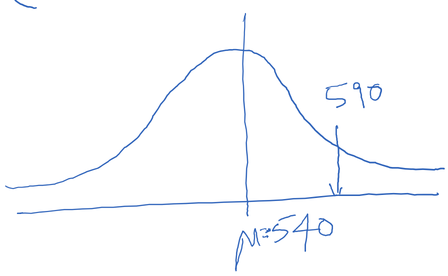




$$\begin{array}{r} 0.5 \\ -\,0.4821 \\ \hline 0.0179 \times 100 \\ = 1.79\% \end{array}$$

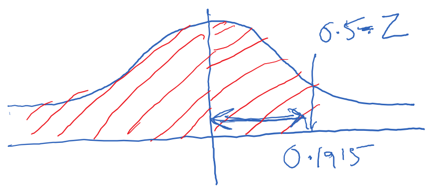
$P(\leq 590)$, $\mu = 540$, $SD = 100$

$$Z = \frac{x-\mu}{\sigma} = \frac{590-540}{100} = \frac{50}{100} = 0.5$$



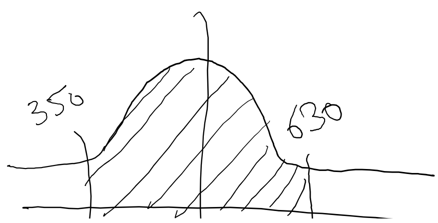
$Z = 0.1915 + 0.5$

$$P(\leq 590) \ \boxed{= 0.6915}$$



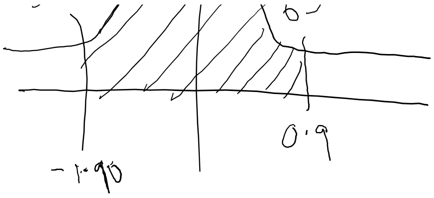
$P(350 < x < 630)$

$$Z = \frac{630-540}{\quad} = \frac{90}{\quad} = 0.90,$$



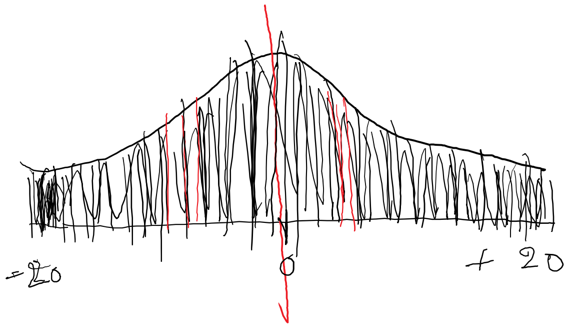
$Z = \frac{630-540}{100} = \frac{90}{100} = \underline{0.90}$

$Z = \frac{550-540}{100} = \frac{-190}{100} = \underline{-1.90}$



$0.4713 + 0.3159$

$\boxed{= 0.7872}$



$n = 10$
$n = 20$
$n = 50$
$n = 100$
$n = 1000$
$n = 1L$

Binom ⇒ Normal

$n >>>>$

1 Parameter Conversion

Binomial = n, p

$\boxed{p + q = 1}$

Normal = μ, σ

$\boxed{\boxed{\mu = n \cdot p} \quad \sigma = \sqrt{n \cdot p \cdot q}}$

Remember → 68%, 95%, 99.7%

$\boxed{O \,\&\, N \to \text{Binomial}}$

$\mu \pm 3\sigma$ → Normal dist

2 Interval check

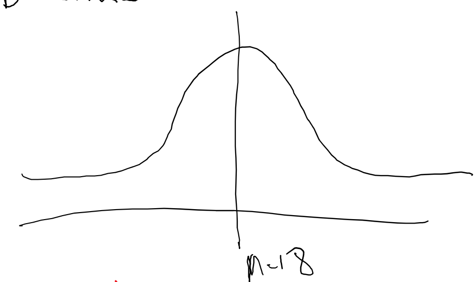
$\boxed{0 < -3SD < \mu < 3SD < n}$

$P(x \geq 25 \mid n = 60 \,\&\, p = 30\%) \to$ Binomial

Param Conversion 1

$\mu = n \cdot p = 60 \times 0.3 = \underline{18}$

$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{60 * 0.3 \times 0.7} = \underline{3.55}$

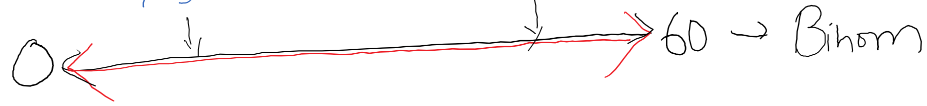


$P(x \geq 25 \mid \mu = 18, \sigma = 3.55) \to$ Normal

→ $\boxed{\mu - 3\sigma}$ & $\boxed{\mu + 3\sigma}$

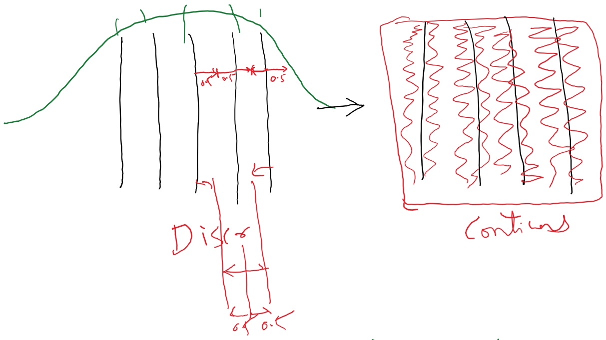
Interval

→ $\boxed{\mu - 3\sigma}$ & $\boxed{\mu + 3\sigma}$

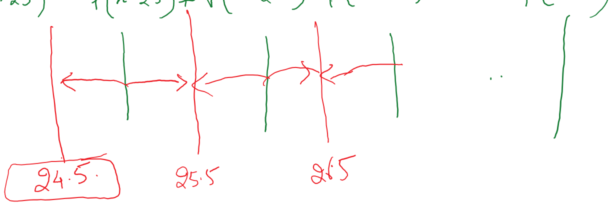


Interval check (2)

(3) Correction for Continuity



$P(x \geq 25) = P(x=25) + P(x=26) + P(x=27) + \cdots P(x=60)$

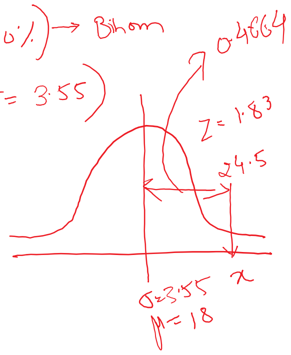


⇒ $P(x \geq 25 \mid n = 60 \ \& \ p = 30\%)$ → Binom

⇒ $P(x \geq 24.5 \mid \mu = 18 \ \& \ \sigma = 3.55)$

$Z = \frac{24.5 - 18}{3.55} = 1.83$

Normal → $\boxed{\begin{array}{l} \geq 0.5 - 0.4664 \\ = 0.0336 \end{array}}$



Value to be determined	Correction
$x >$	$+.50$
$x \geq$	$-.50$
$x <$	$-.50$
$x \leq$	$+.50$
$\leq x \leq$	$-.5$ & $+.5$
$< x <$	$+.5$ & $-.5$
$x =$	$-.5$ & $+.5$

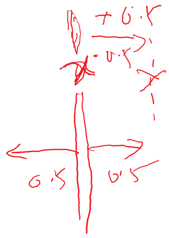
→ Reduced

Real

= → Extend

No = Reduced

9611757941

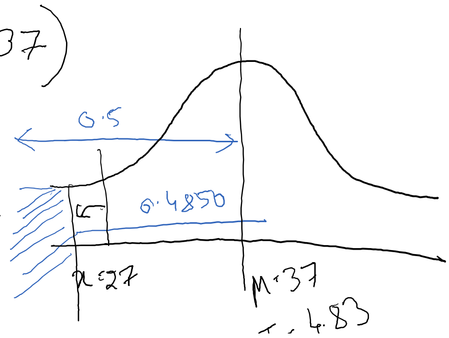


$P(x < 27 \mid n = 100 \ \& \ p = 0.37)$

(1) parameter

$\mu = n.p = 100 \times 0.37 = 37$

$\sigma = \sqrt{n.p.q} = 4.83$



26.5 $\sigma = 4.83$

② Test for Interval

$\mu \pm 3\sigma$

$\mu + 3\sigma$ $\mu - 3\sigma$

22.51 51.49

0 ⟷ 100

③ $Z = \frac{x-\mu}{\sigma} = \frac{26.5-37}{4.83} = -2.17$

$P(<27) \simeq P(<26.5)$

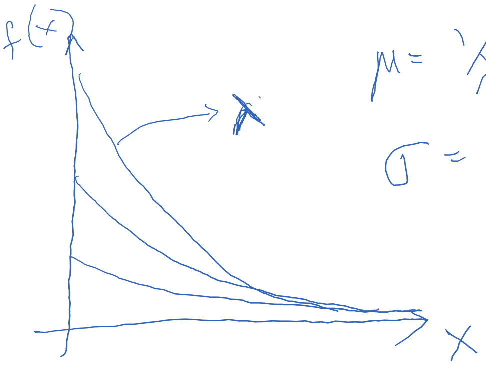
$\boxed{\Rightarrow 0.5 - 0.4850 = 0.015}$

Exponential Dist

$f(x) = \lambda e^{-\lambda x}$

$\mu = 1/\lambda$

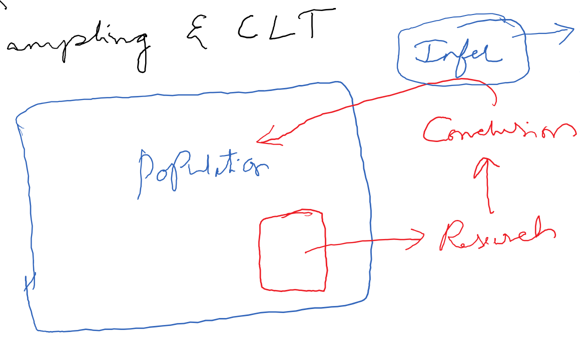
$\sigma = 1/\lambda$



Sampling & CLT

Infer → Inferential stats

ANMP



Population → Parameter

Sample → statistic

$a\ \bar{x}\ p\ v$

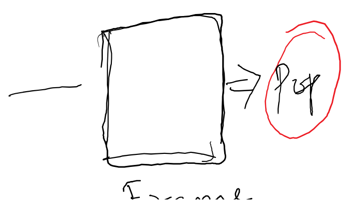
---

1) Budget
2) Time
3) Research → Destructive
4) Population is Not accessible

Why do we need Sampling

Frame Theory

Frame = Population



Frame

Practical

Frame $\neq$ Population $\rightarrow$ Pop

Sampling

Random $\rightarrow$ Equal probability

- Simple
- Stratified
- Systematic
- Cluster

Non-Random $\rightarrow$ Right place Right time

- Convenience
- Judgement
- Quota
- Snowball

Simple

1 person

100

10 people

Assign numbers

1 –	11 –		91 –
2 –			
3 –			
10 –	20 –		100 –

Generate Random Numbers $\rightarrow$ 32 – 41 – 36 – 89 – $\rightarrow$ Sample

1

Stratified Random

Sample

Age

20-30 Homogenous Alike	
40-50 Homogen Alike	
60-70 Homogen Alike	

Heterogenous Different

Heterogenous Different

Systematic Sampling

1^{st}, 2^{nd} . . . . K^{th} n

0 10^{th} 20^{th} 30^{th} . . . . 1000^{th}

N = 1000

"n" = 100 "N"

$$K^{th} = \frac{N}{n} = \frac{1000}{100} = \boxed{10}$$

Cluster → Groups that are natural

Natural Heterogenous

E-N

E-D

E-G

→ Strata

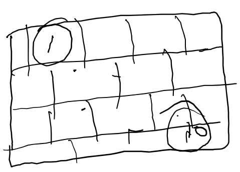
E-ng

N D G | M / F

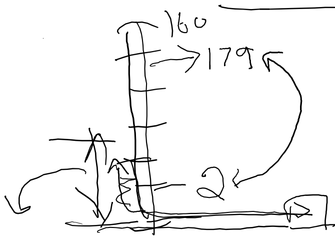
Stratified

Non-Random Sampling → Without Random sampling process

1) Convenience Sampling →

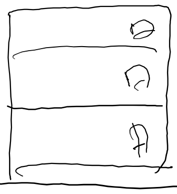


Burj → 160 Floors



Judgemental Sample → Researcher →

Quota :



Stratified + Non-Random

Snowball

Sampling error

Non-sampling error.

→ Missing values

→ Reading error →

→ Processing error

→ Analysis error

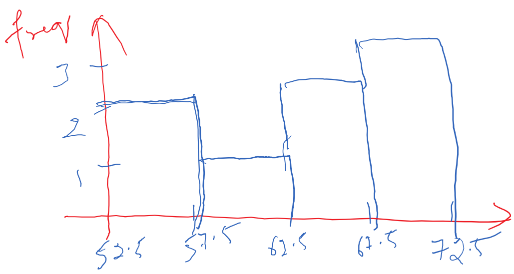
Sampling distribution of $\bar{x}$ ($\bar{x}$ → sample mean)

→ Know how my samples are distributed → $\bar{x}$

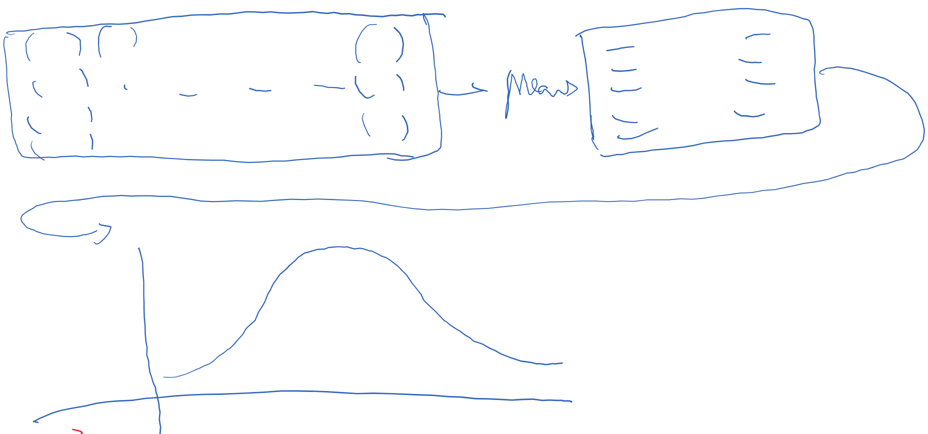
69 54 55 59 63 64 68 70

$N = 8$

$n = 2$



x	$f(x)$
52.5	2
57.5	1
62.5	2
67.5	3
72.5	



CLT → If samples of n are drawn randomly from a population that has μ & $SD = \sigma$

Population that has μ & $SD = \sigma$

→ The sample means ($\bar{x}$) approximate to a normal distribution ($n \geq 30$) → Irrespective of population

$\boxed{\mu_{\bar{x}} = \mu}$ & $\boxed{\sigma_{\bar{x}} = \sigma/\sqrt{n}}$

$$\boxed{Z = \frac{x-\mu}{\sigma} \Rightarrow \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x}-\mu_{\bar{x}}}{\sigma/\sqrt{n}}}$$

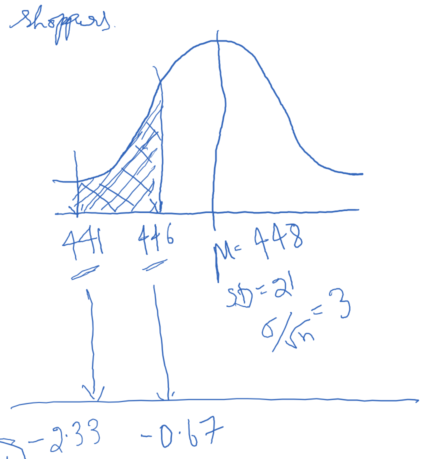
$\boxed{n \uparrow}$ $\sigma_{\bar{x}} \downarrow \Rightarrow Z_{\bar{x}}$ approximates to Z

Any hour → no of shoppers → 448 (Avg), SD = 21

What is the prob that a random sample of 49 hours will give me the sample mean btn 441 & 446 shoppers.

$$Z_{446} = \frac{\bar{x}-\mu_{\bar{x}}}{\sigma/\sqrt{n}} = -0.67 = 0.2486$$

$$Z_{441} = -2.333 \Rightarrow 0.4901$$



$$\Rightarrow \begin{array}{r} 0.4901 \\ -0.2486 \\ \hline \end{array}$$

$\boxed{\simeq 0.25} = \boxed{25\%}$

Sample proportion $\hat{p} = x/n$

$\boxed{\mu = p}$ $\boxed{\sigma = \sqrt{\frac{p \cdot q}{n}}}$ $\boxed{Z = \hat{p} - p \Big/ \sqrt{\frac{p * q}{n}}}$

→ 60% → Lenovo laptops.

→ 120 Random samples

→ < 50% → using lenovo

$p = 0.6$

$q = 0.4$

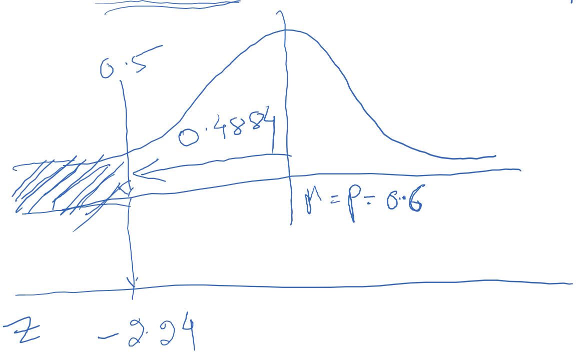
$\hat{p} = 0.5$

$n = 120$

$$Z = \frac{0.5 - 0.6}{\sqrt{\frac{0.6 * 0.4}{120}}} = \frac{-0.10}{0.04472}$$

→ 120 Random samp

→ < 50% → Using lenovo



$\hat{p} = 0.5$

$n = 120$

$= \dfrac{-0.10}{0.04472}$

$\boxed{z = -2.27}$

$P(<0.5) = 0.5 - 0.4884$

≈ 0.012