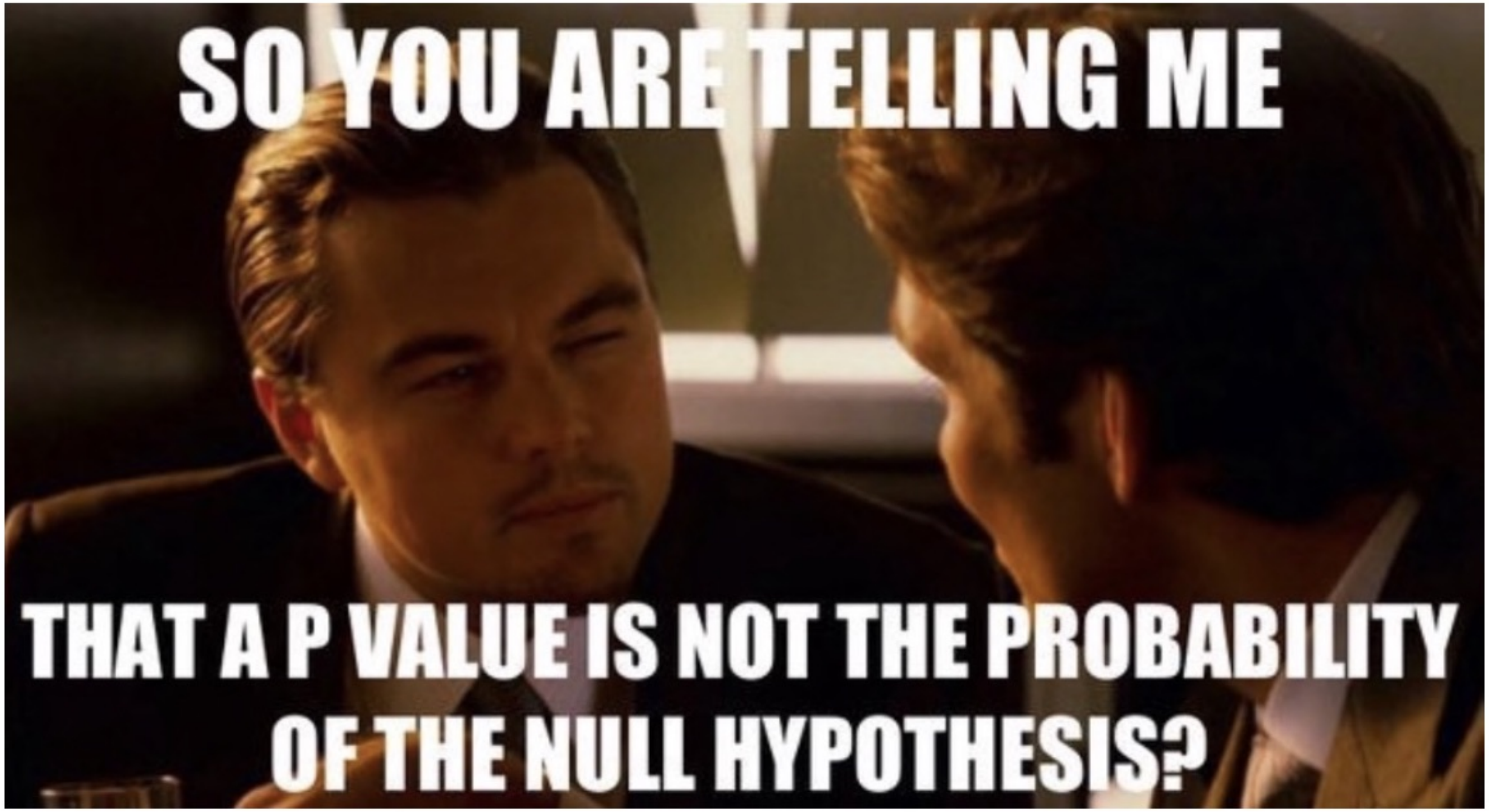
24th June '23



Pvalue — Prob. of Ho being True

✓ ✗

- Recap framework
- Recap z test

left   Two tailed

- 1 sample z test
- 1 sample T-test
- 2 sample T-test.

Right, Left, Two-tailed *

Framework

(i) H_0 & H_A

(ii) Test stat & $\underline{db}^n$

(iii) which tail test.

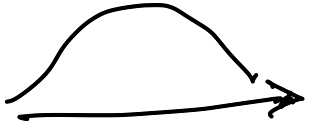
(iv) compute pvalue

(v) compare pvalue with α

Sig. level.

If $p < \alpha \rightarrow$ Reject $\underline{H_0}$

* Premature Childrens :



IQ

Avg IQ = 100 → population

std dev = 15 → σ.

· Medical Researchers wants to check.

↓

lower IQ.

@ 5% significance ~~level~~

$\alpha = 0.05$

50 premature child

↓

finding → Avg. IQ = 95

$n = 50$

(i) $H_0 : \mu = 100$

$H_A : \mu < 100$

(ii) Test stat.

'm' → Sample mean

↓

R.V → normal

(iii) which Tail Test $\longrightarrow$ Left tail.

$E(m) = 100$

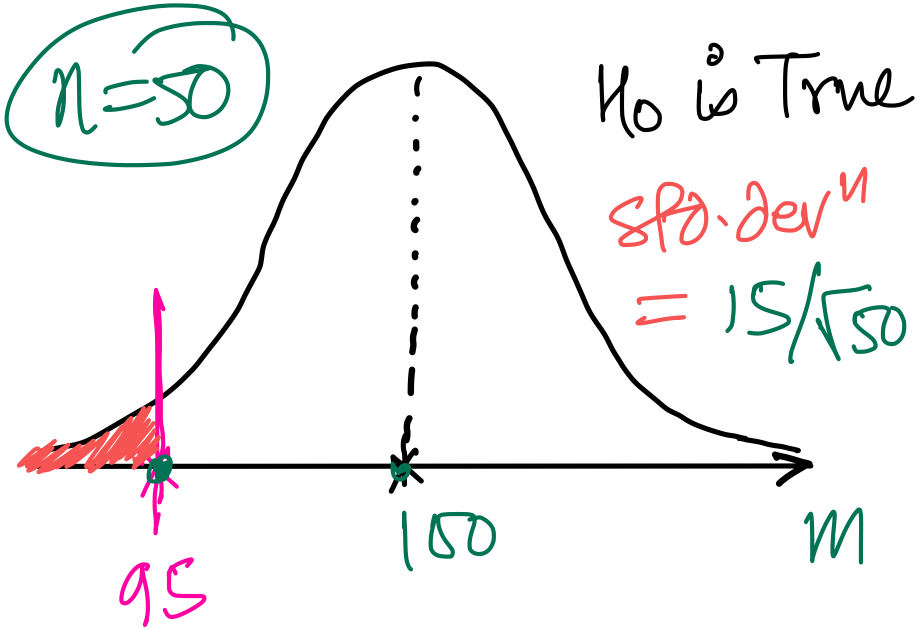
(iv) p-value.

$\downarrow$

$P\left[\, m \leq 95 \mid H_0 \text{ is True} \right]$

$= \text{norm.cdf}\left[\left(\dfrac{95 - 100}{15/\sqrt{50}}\right)\right]$

$= 0.009$



z_{95} ? $\rightarrow$ norm.cdf (z_{95}).

pvalue $< \alpha$ (0.05) $\rightarrow$ reject H_0

Premature babies have lower IQ.

II nd

lower IQ $\longleftrightarrow$ higher IQ.

99 % confidence.

1 % Significance.

$\alpha = 0.01$

$\mu = 100$
$\sigma = 15$ } Population.

· 50 premature child. $\rightarrow$ 94

obs. mean.

m $\rightarrow$ Sample mean

framework :

· $H_0 : \mu = 100$
$H_A : \mu \neq 100$

$H_A : < 100 \rightarrow$ left.

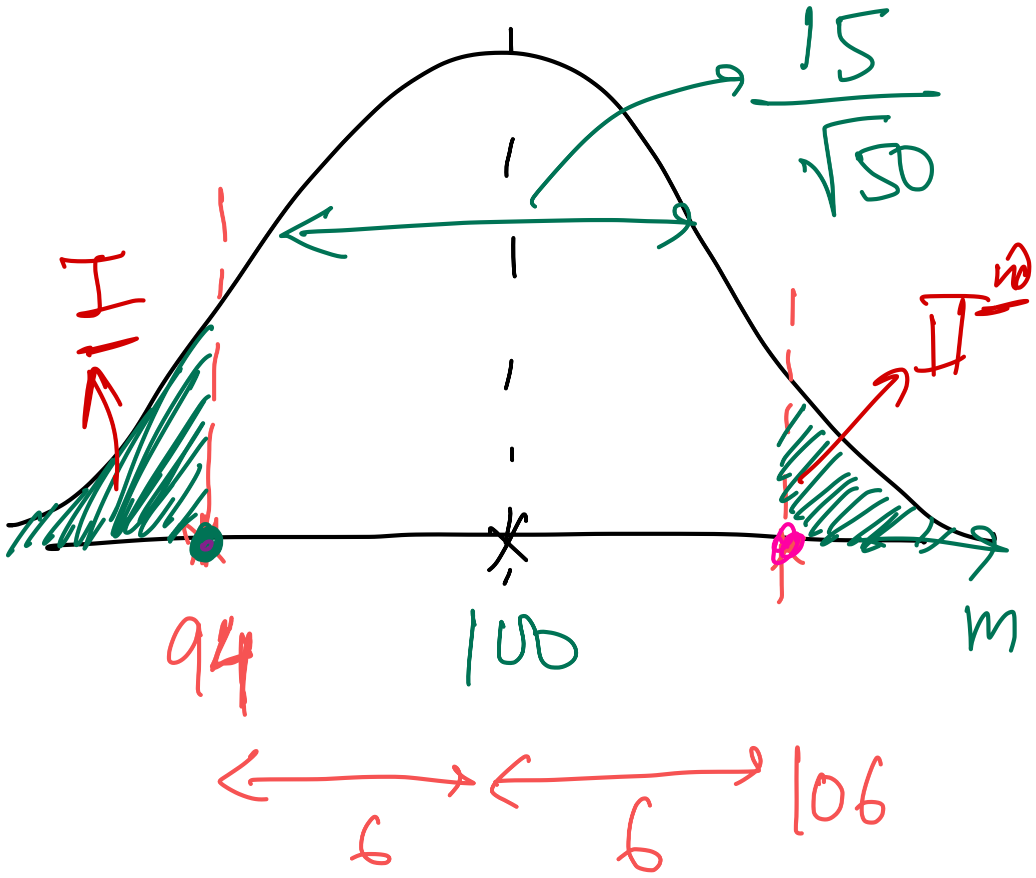
$H_A : > 100 \rightarrow$ Right.

$H_A : \neq 100 \rightarrow$ Two tailed Test.

(2) Sample mean → m

$m \sim$ Normal dbⁿ.

$m_{obs} = 94$



(3) Two tailed Test

(4) P value

Ist : norm.cdf$\left[\left(\frac{94-100}{15/\sqrt{50}}\right)\right] = 0.0023$

IInd : $1-$ norm.cdf$\left[\frac{106-100}{15/\sqrt{50}}\right] = 0.0023$

p value = I + IInd

= 2 x 0.0023

$\boxed{\text{p value} = 0.0046}$

$\alpha = 0.01$

Reject H_o

$p < \alpha$

premature child have significatly diff. ID than normal babies at 99% conf. or 1% sig.

8:15

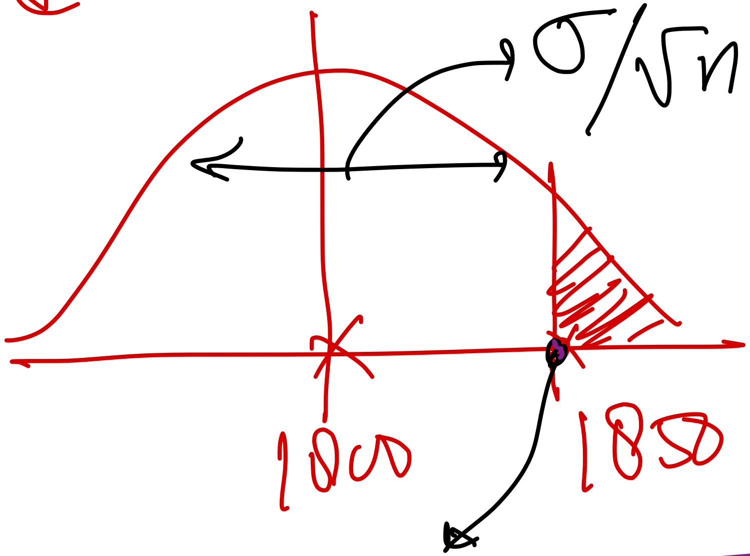
break !

σ : known

n : Sample size

Popl^n mean → claims

$m \to$ sample mean $\sim N(\mu, \sigma/\sqrt{n})$



$$Z\_stat = \left(\frac{1850-1800}{\sigma/\sqrt{n}}\right)$$

σ : is unknown

$m \sim T$ distribution

n

$n<30$

$m \sim T$ db^n

$n>30$

$m \sim T$ db^n / Normal

Popl^n std - dev^n

σ → known → 'n' doesn't matter. Z test Normal distb^n

σ → unknown. → n

$n < 30$ → T distb^n T test

$n > 30$ T test / Z test almost same

IQ Pill

$\mu = 100$

σ is known.

Sample of people = { , , , , , , , }

$\rightarrow n$

σ is Unknown

$\Rightarrow$ test statistic

$$Z\_stat = \left(\frac{x - 100}{\sigma / \sqrt{n}} \right)$$

$\downarrow$

Normal

$$t\_stat = \frac{x - 100}{(s / \sqrt{n})}$$

$\downarrow$

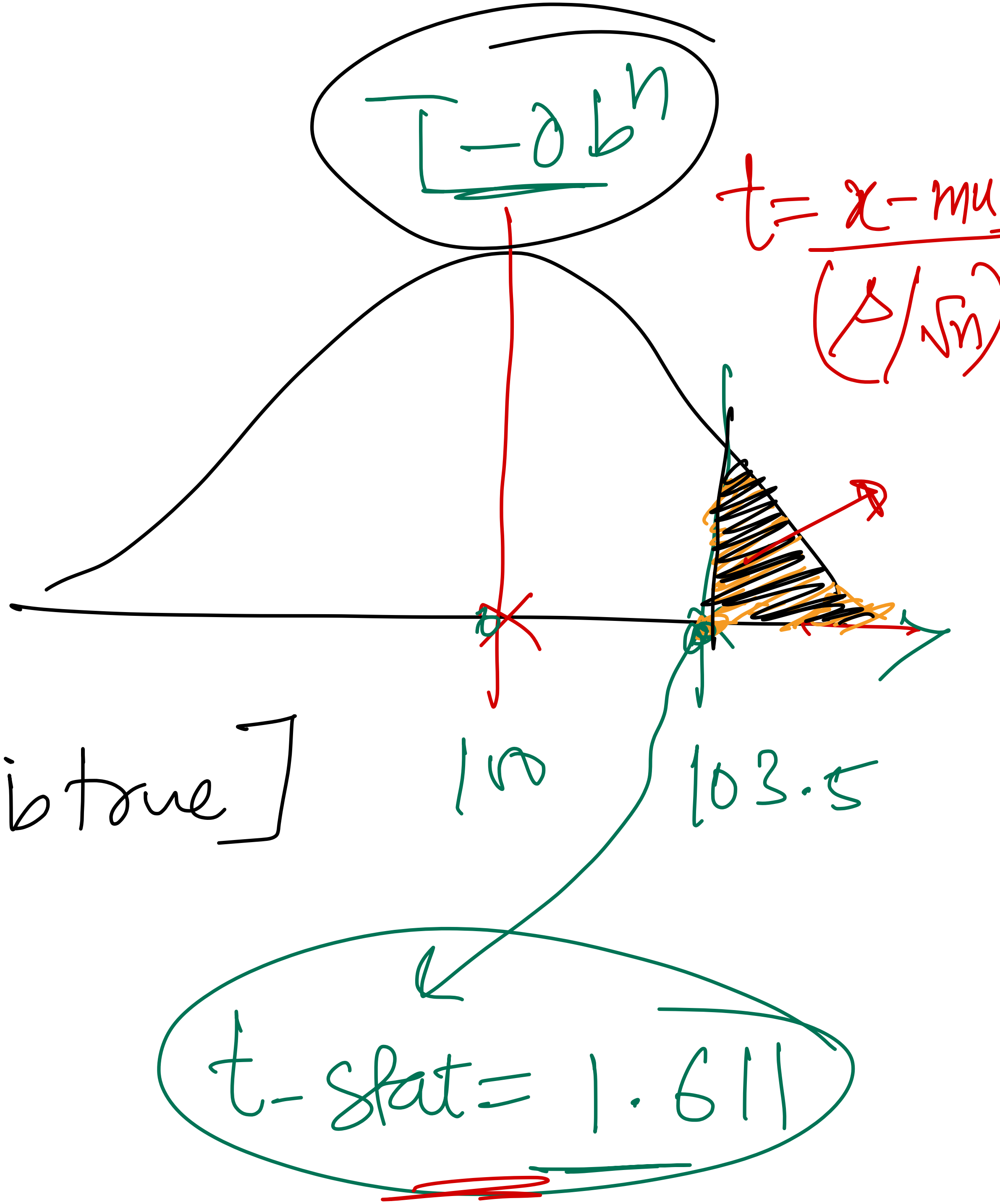
t distribution

$\Rightarrow$ Avg IQ = 100

$H_0 = \mu = 100$

$H_A: \mu > 100$

$P[t\_stat \geq 1.611 \mid H_0 \text{ is true}]$



CLT $\quad \bar{x} \longrightarrow \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right).$

$n < 30$

t dbⁿ

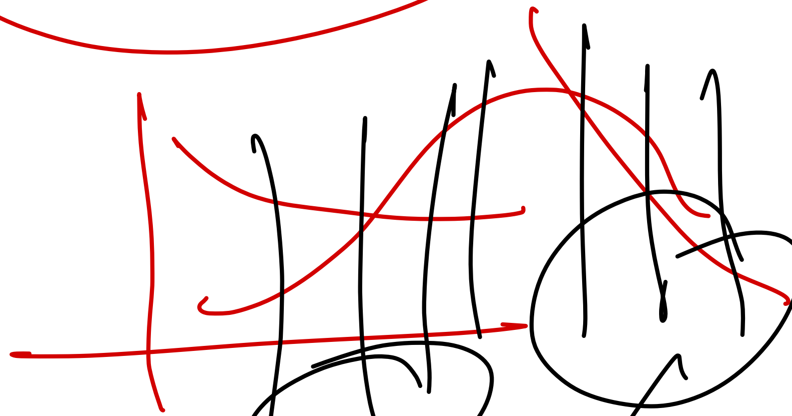
Normal dbⁿ

σ is known

normal

'n' doesn't matter

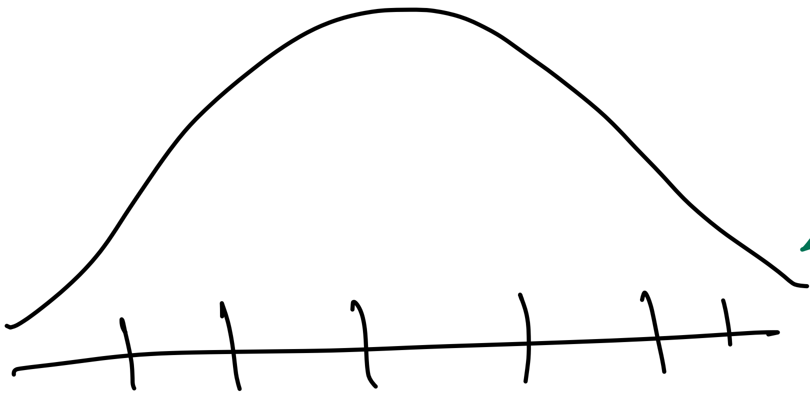
not normal



m_1

m_2

normal



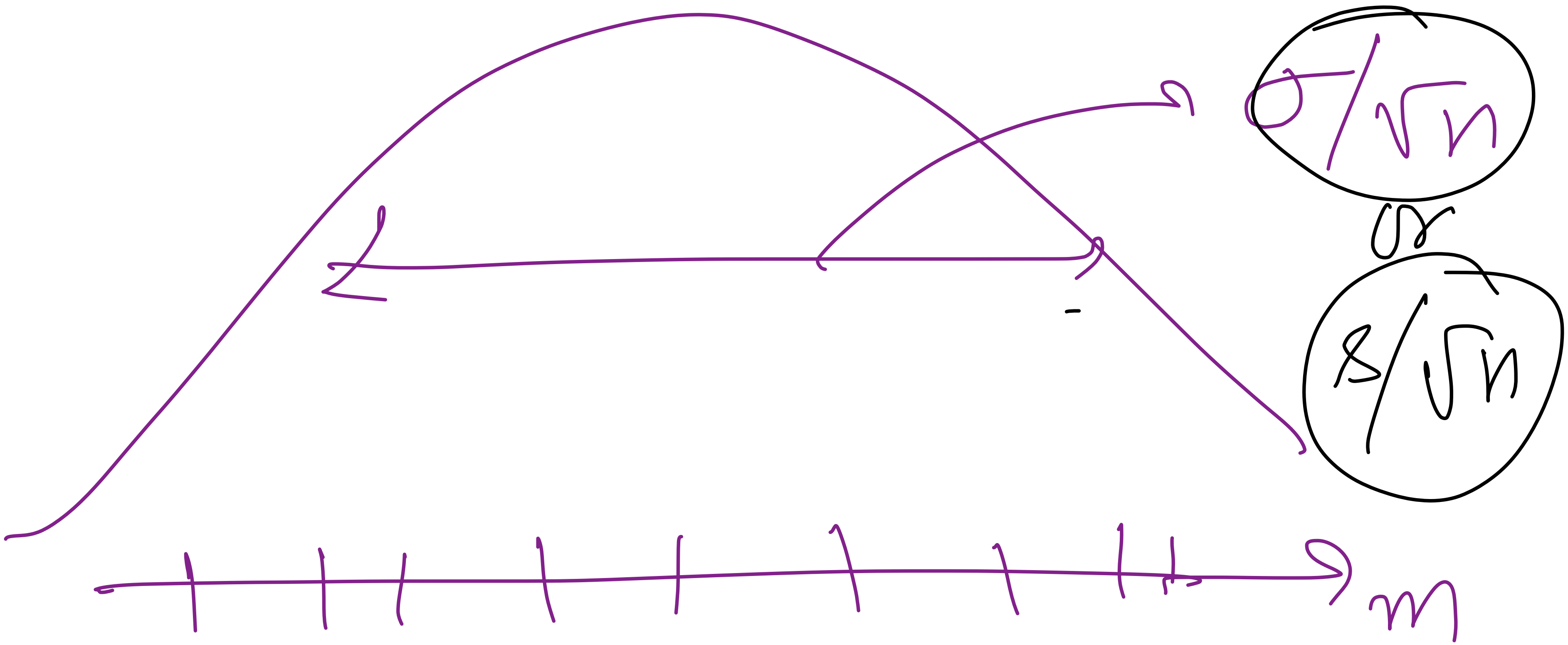
$n < 30$

σ = pop^n std. dev^n

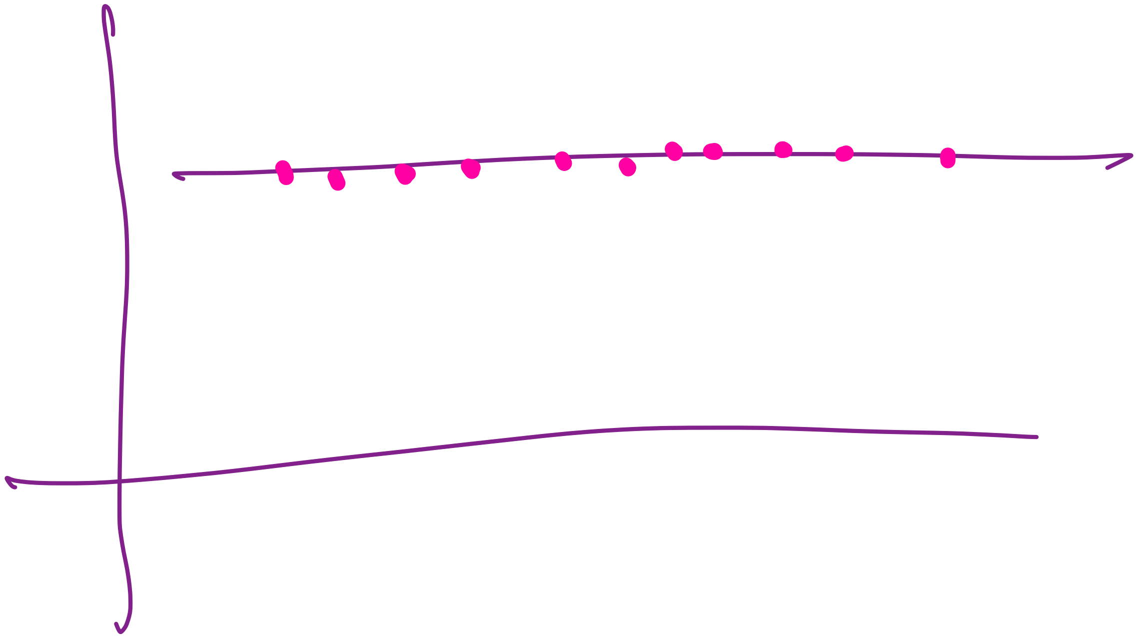
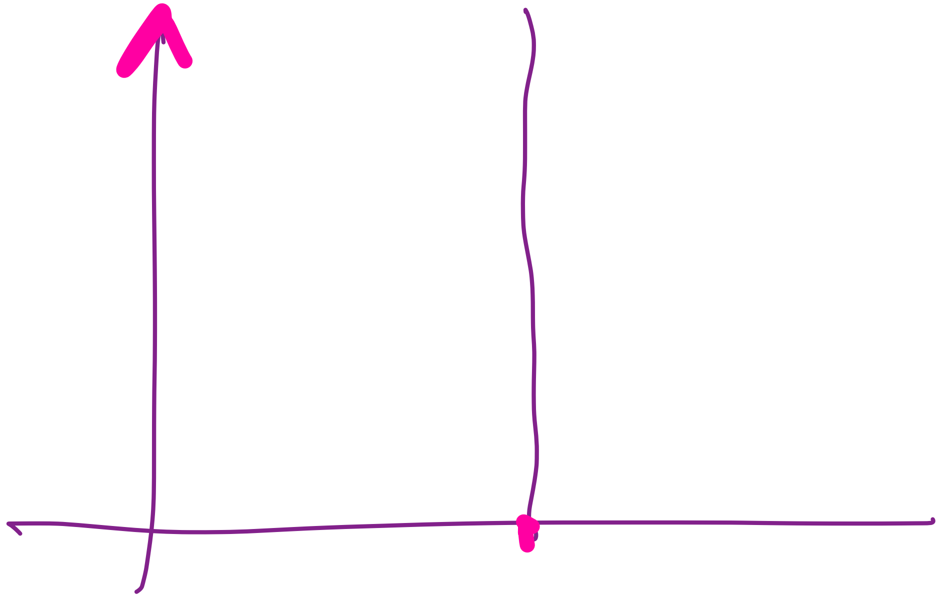
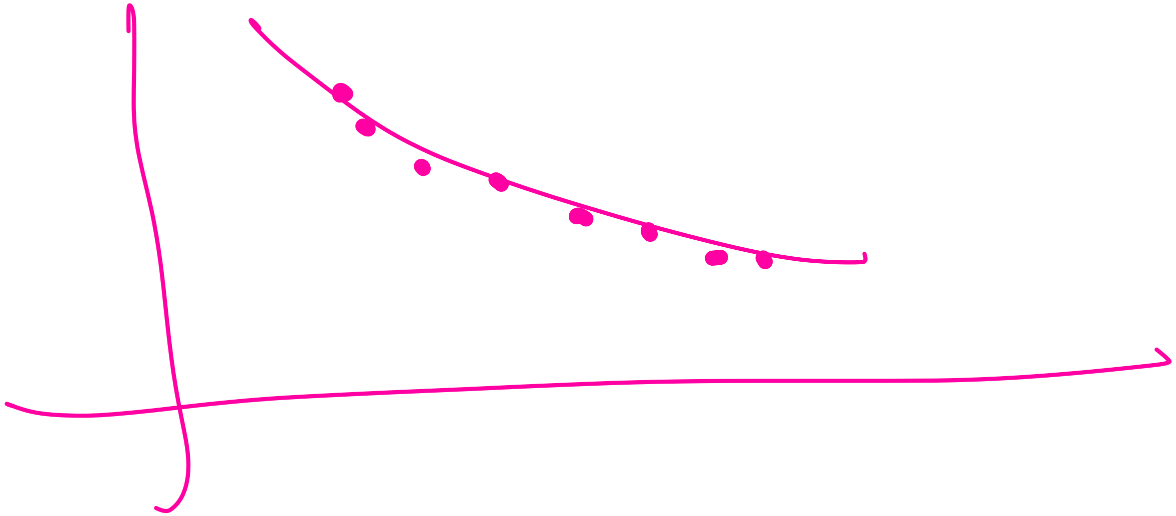
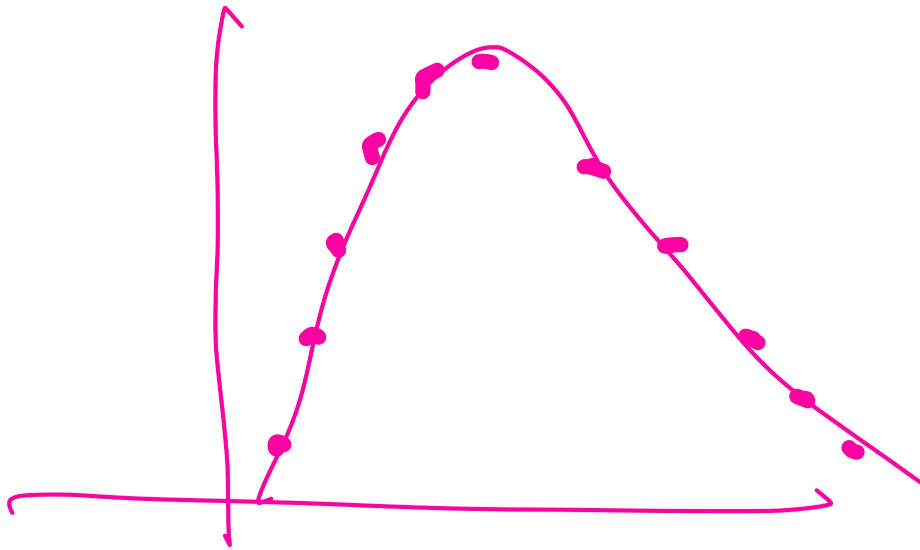
Population variability

CLT → m → Normal

s → Sample std. dev^n



σ/√n
or
s/√n
m

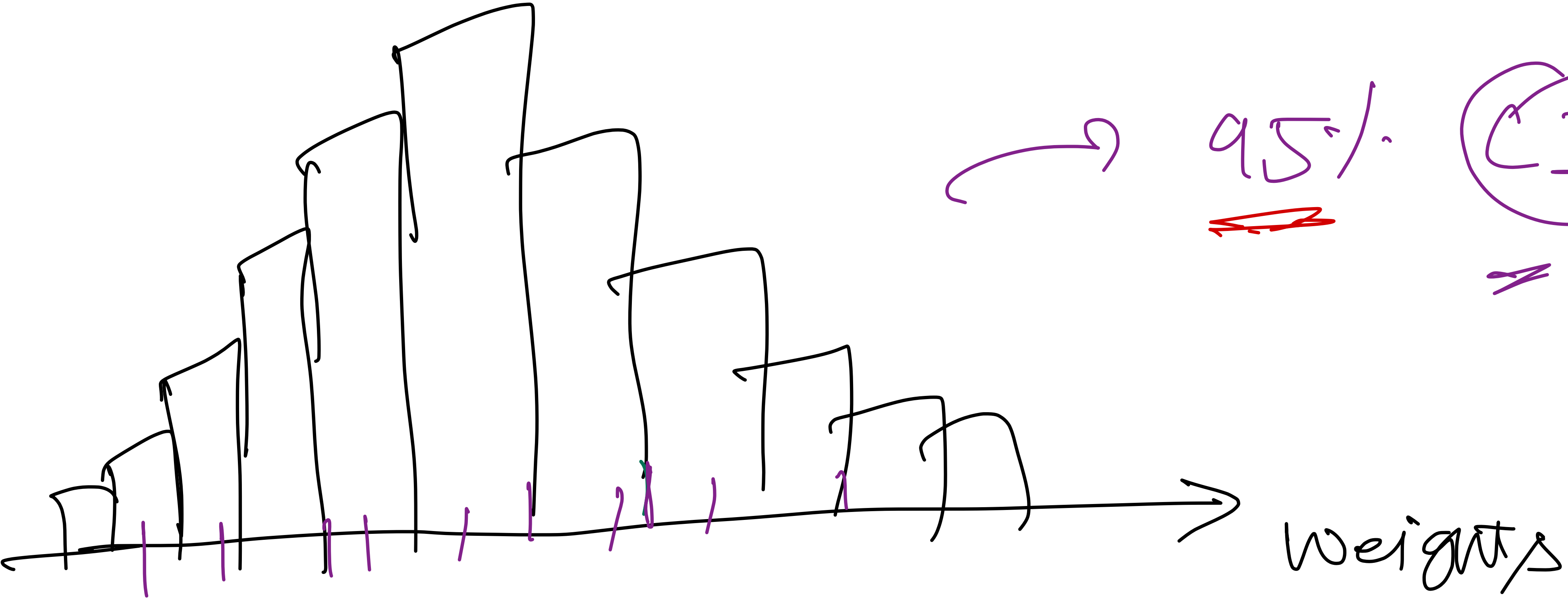





Theoretical

utilized the prob. of normal db^n

Emperical

Bootstrapping.



→ 95% CI

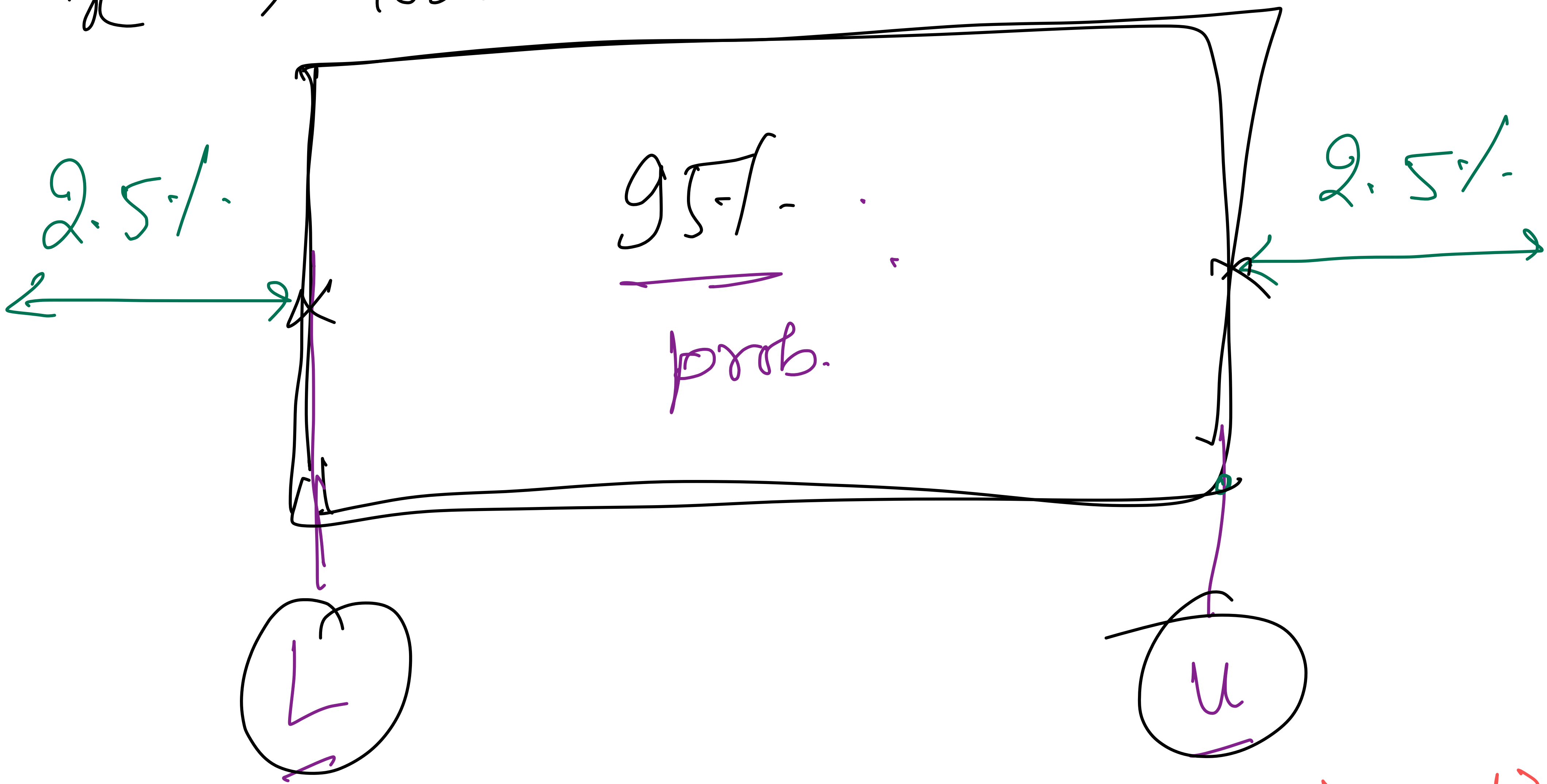
$w_1, \ldots\ldots w_i, \ldots\ldots w_n$

↓

Individual w_i

Range in which 95% weight lies.

$x \longrightarrow$ 1000 wts.



np.percentile(x, 2.5)

np.percentile = (x, 97.5)

weights = popl^n mean is unknown.

$n = 200$ → $\bar{x}$ = 65kg

1 time activity

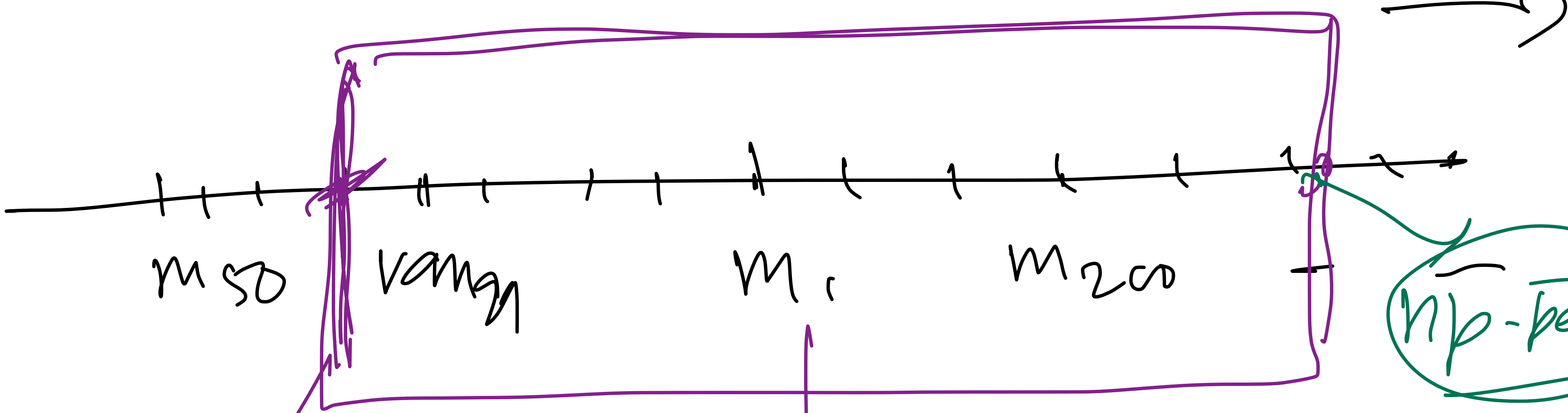
$n = 200 \Rightarrow S_1 \Rightarrow m_1$

$= 200 \Rightarrow S_2 \Rightarrow m_2$

⋮

$S_{1000} \Rightarrow m_{1000}$

Sort thes 1000 means.

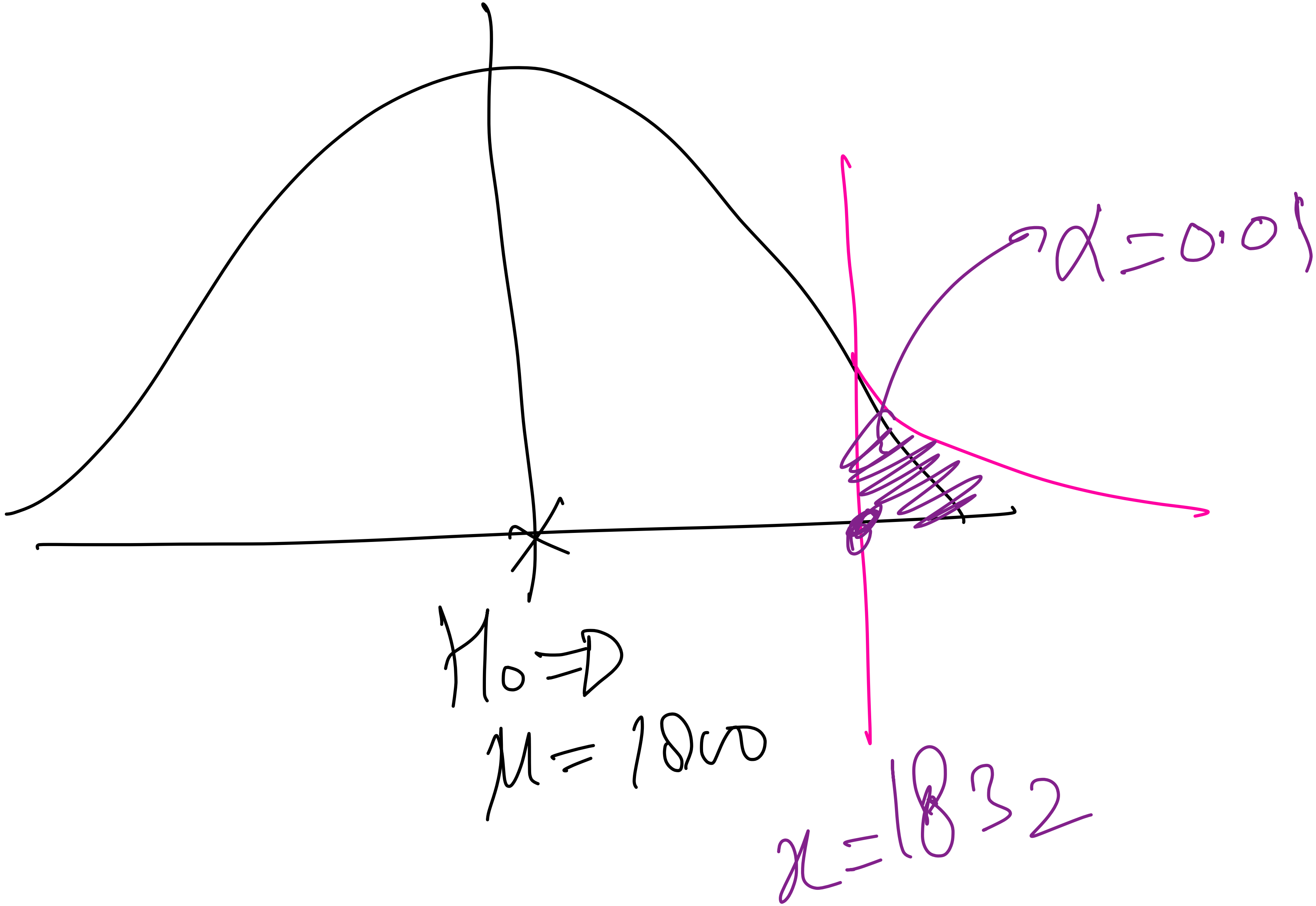


np.percen

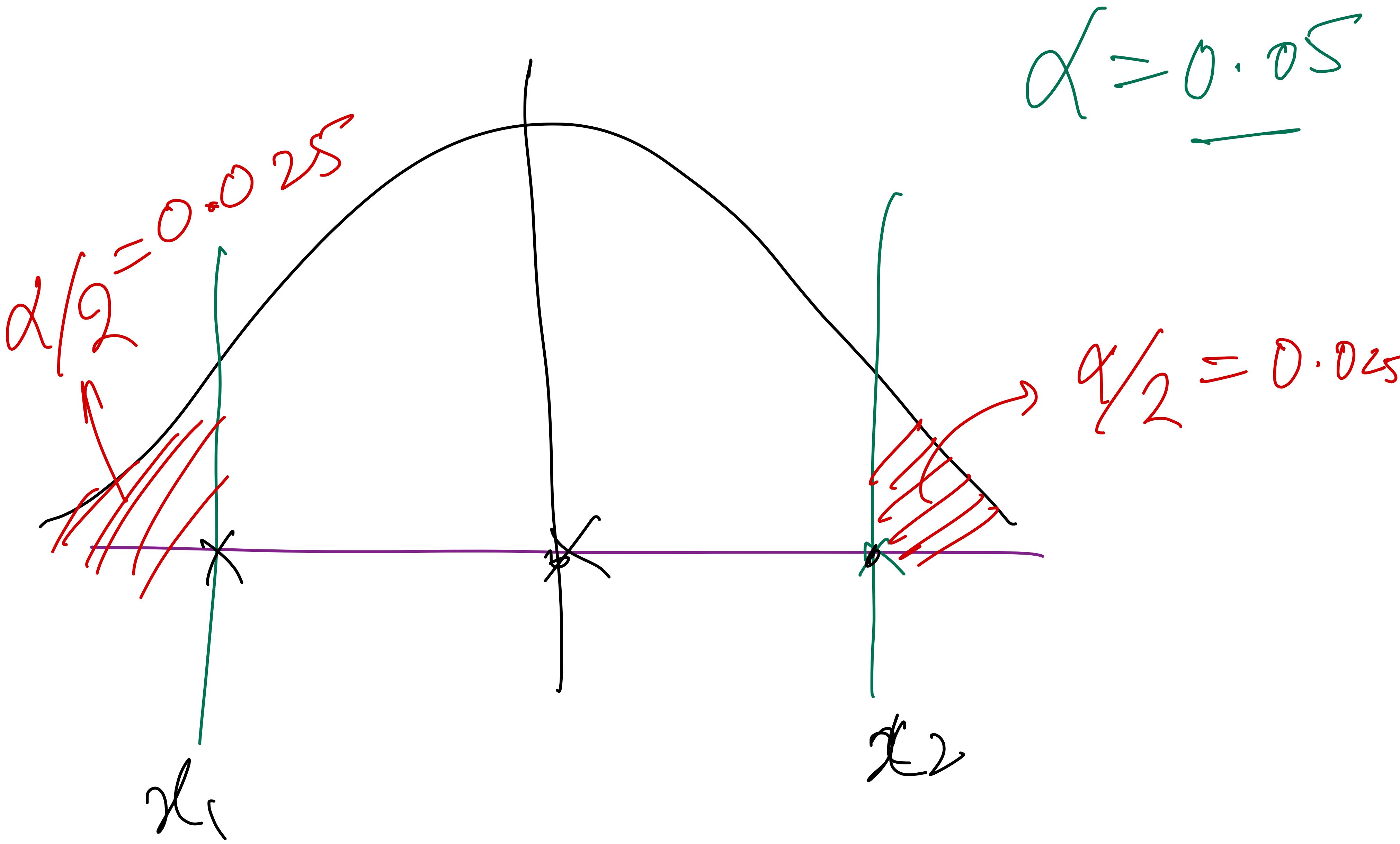
Interval contain 95% of the means

[63.2 , 67.5]

95% confidence



$\alpha = 0.01$
$H_0 \Rightarrow$
$\mu = 1800$
$x = 1832$



$\alpha = 0.05$
$\alpha/2 = 0.025$
$\alpha/2 = 0.025$
x_1
x_2