FEATURE ENGINEERING-3

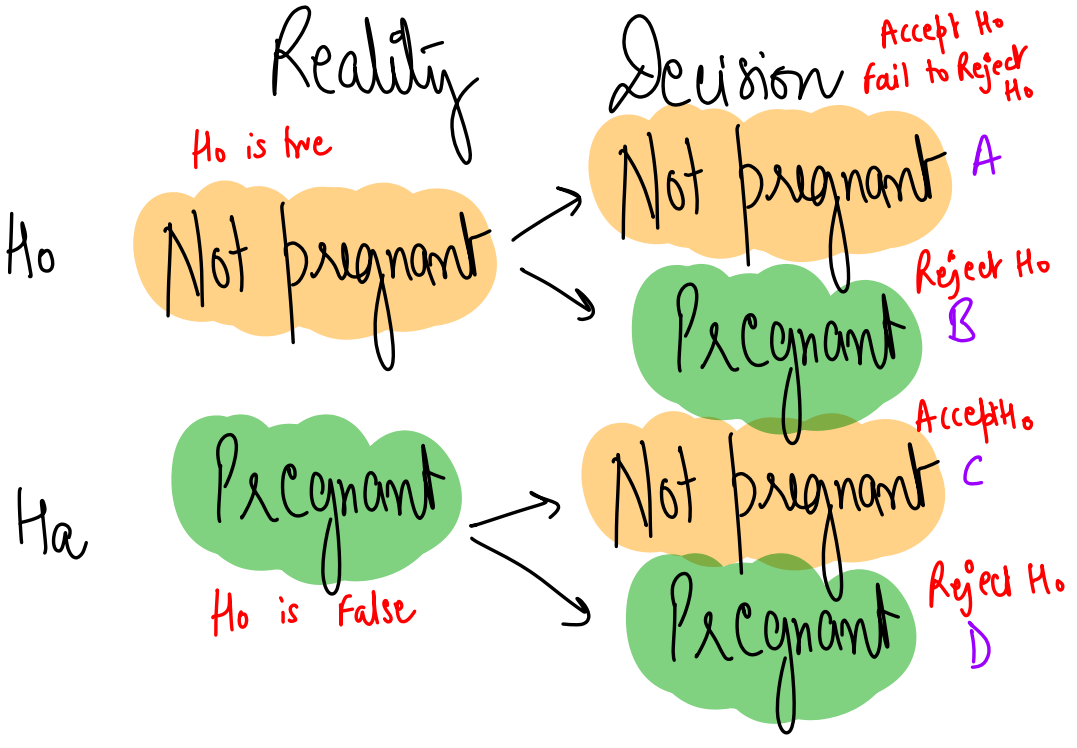
FORTIS

H_0: NOT PREGANT

H_a: PREGNANT

* Reject H_0

* Fail to Reject H_0 → Accept H_0



TRUE NEGATIVE

FALSE POSITIVE
(Type I error)

FALSE NEGATIVE
(Type II errors)

TRUE POSITIVE

Height of normal people

$\mu = 65$

$\sigma = 4$

Height of NBA players

$\mu = 72$

$\sigma = 3$

(-ve) H_0 : (Normal) $\mu = 65$

(+ve) H_a : (NBA players) $\mu > 65$

"Right tailed Test"

99% Confidence

$\alpha = 0.01$

① Sample 2 Individual ⇒ Avg = 70 dof = 1

$H_0 = 65$

pvalue = 0.16

$\alpha = 0.01$

pvalue > α

$$\text{pvalue} = 1 - t.cdf\left(\frac{70-65}{4/\sqrt{2}}\right)$$

Fail to Reject H_0

② Sample 10 Individuals ⇒ Avg = 10

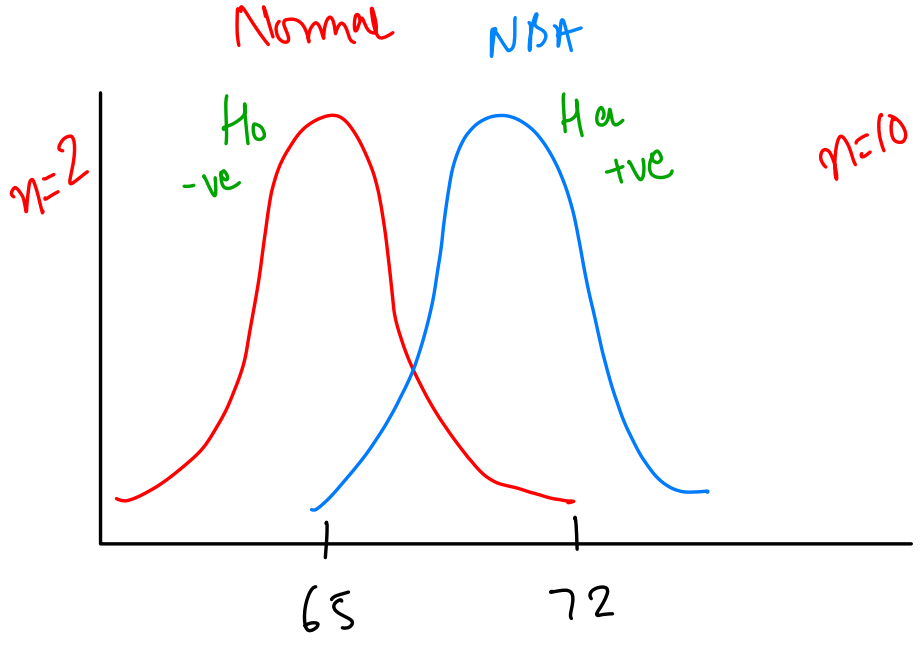
pvalue = 0.001

$\alpha = 0.01$

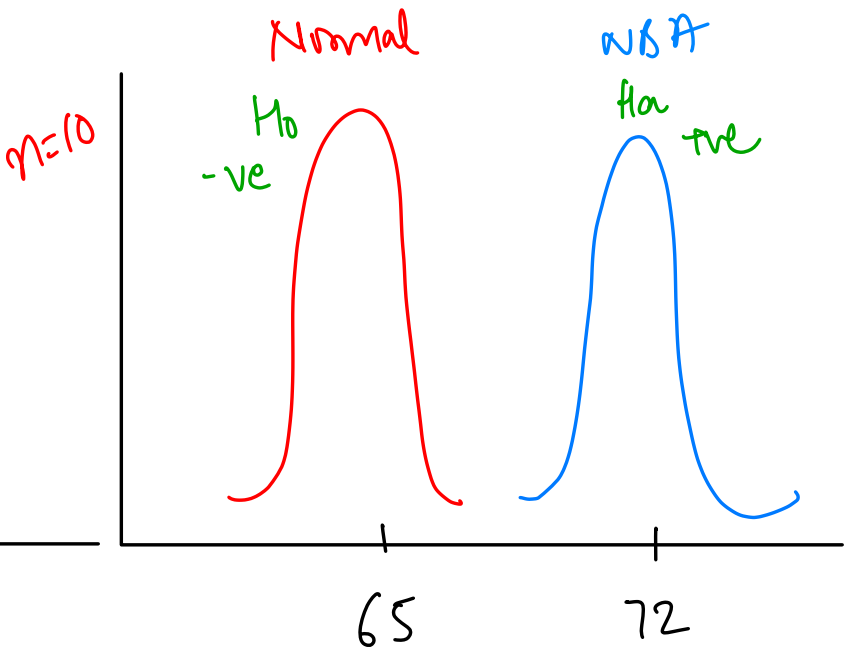
pvalue < α

$$\text{pvalue} = 1 - t.cdf\left(\frac{70-65}{4/\sqrt{10}}\right) \quad dof = 9$$

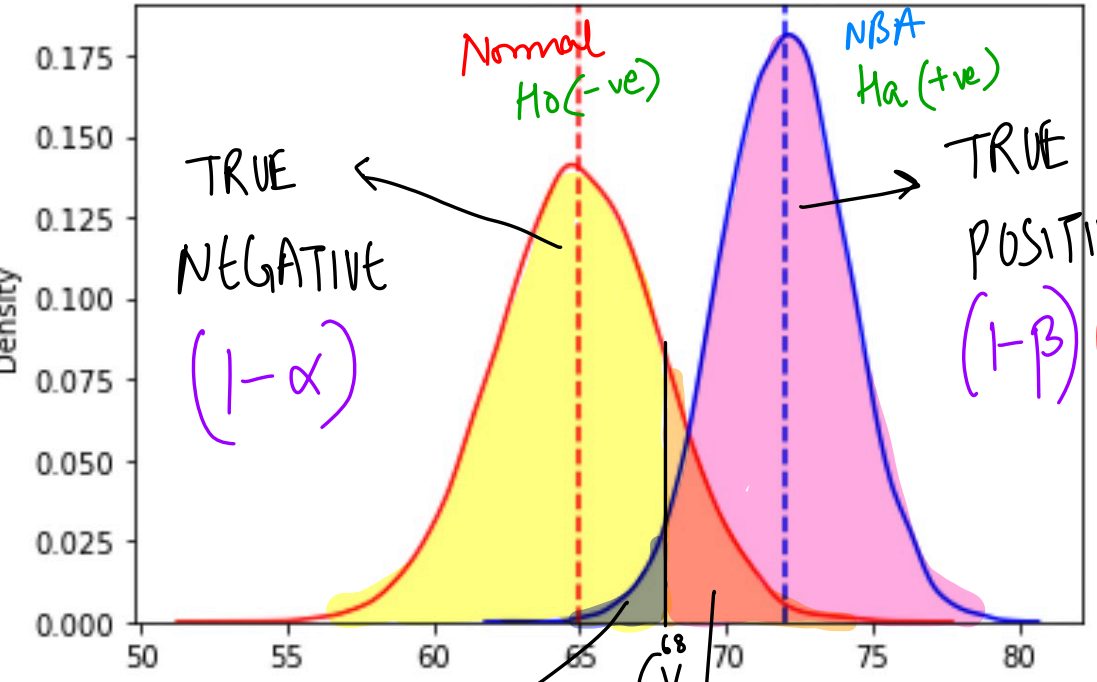
Reject H_0



more overlap



less overlap



Confidence 99%

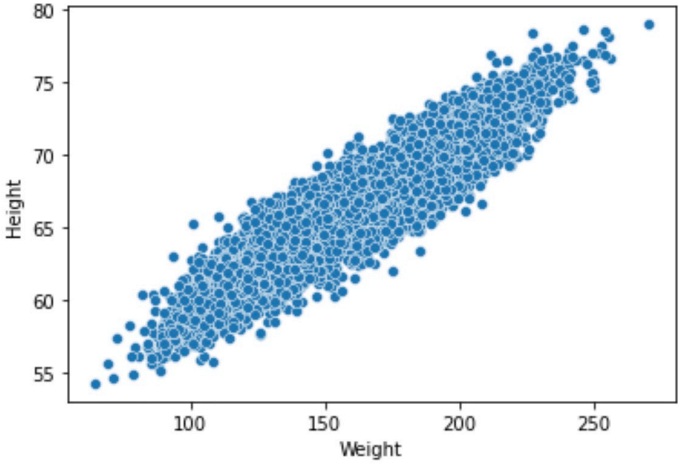
$\alpha = 0.01$

		Decision: Accept	Decision: Reject
H_0	True	True negative $1-\alpha$ Confidence level	False positive α Significance level
	False	False negative β	True positive $1-\beta$ Power

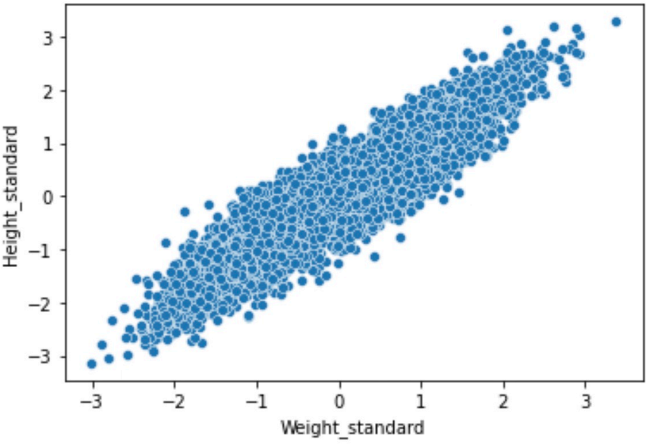
Normalisation / Standardisation

(FOOT) Heights	(gms) Weights	Target	$Z = \frac{x-\mu}{\sigma}$	Standardized Form	Normalisation $\frac{x - min}{max - min}$
1	1000	Cat	$\frac{1-3.5}{\sigma}$	→	
2	3000	dog	$\frac{2-3.5}{\sigma}$	→	
3	4060	elephant	$\frac{3-3.5}{\sigma}$	→	
4	2200	Cat	$\frac{4-3.5}{\sigma}$	→	
5	5000	dog	$\frac{5-3.5}{\sigma}$	↗	
6	3900	Cat	$\frac{6-3.5}{\sigma}$	↗	

ORIGINAL

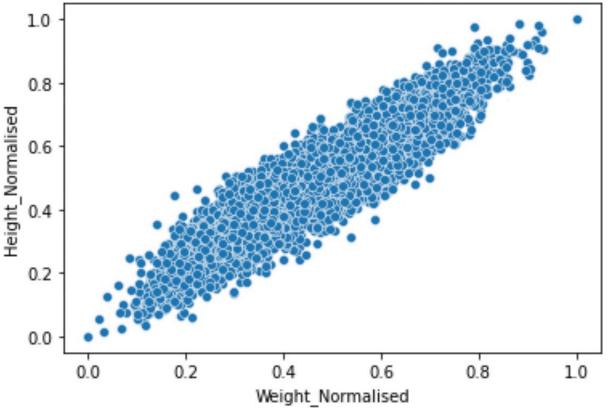


STANDARD



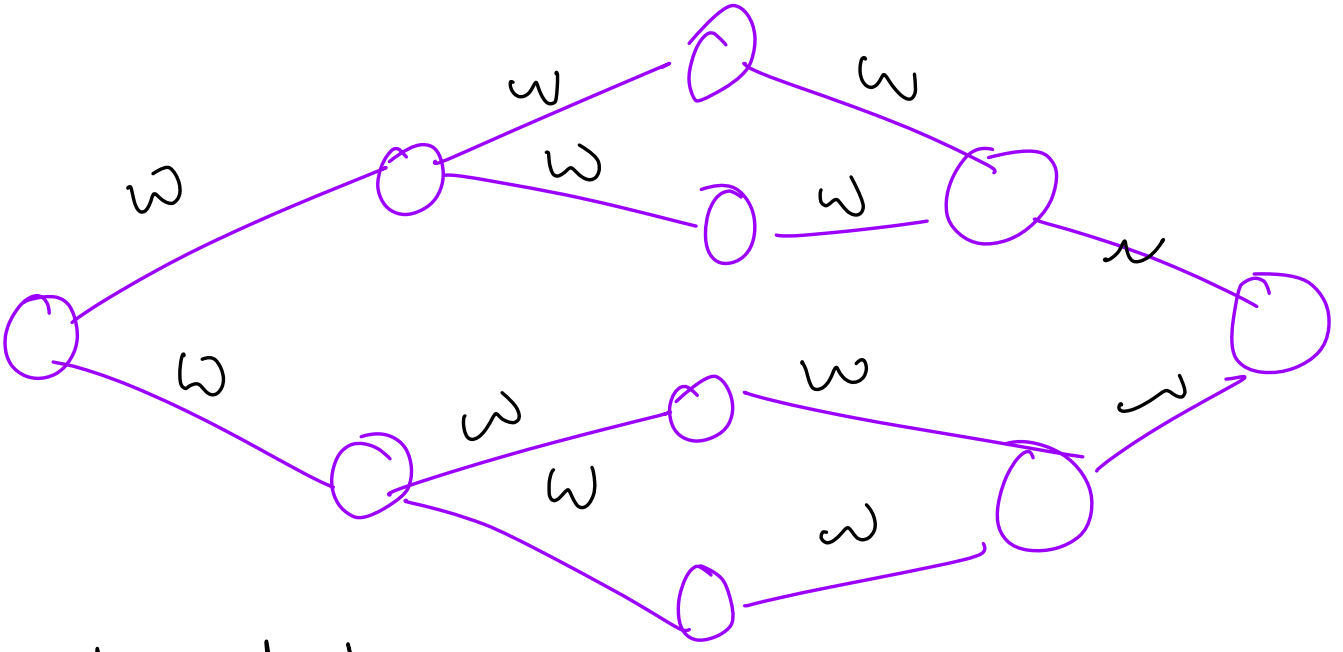
RANGE = Unknown

$$\frac{x - \min(x)}{\max(x) - \min(x)}$$



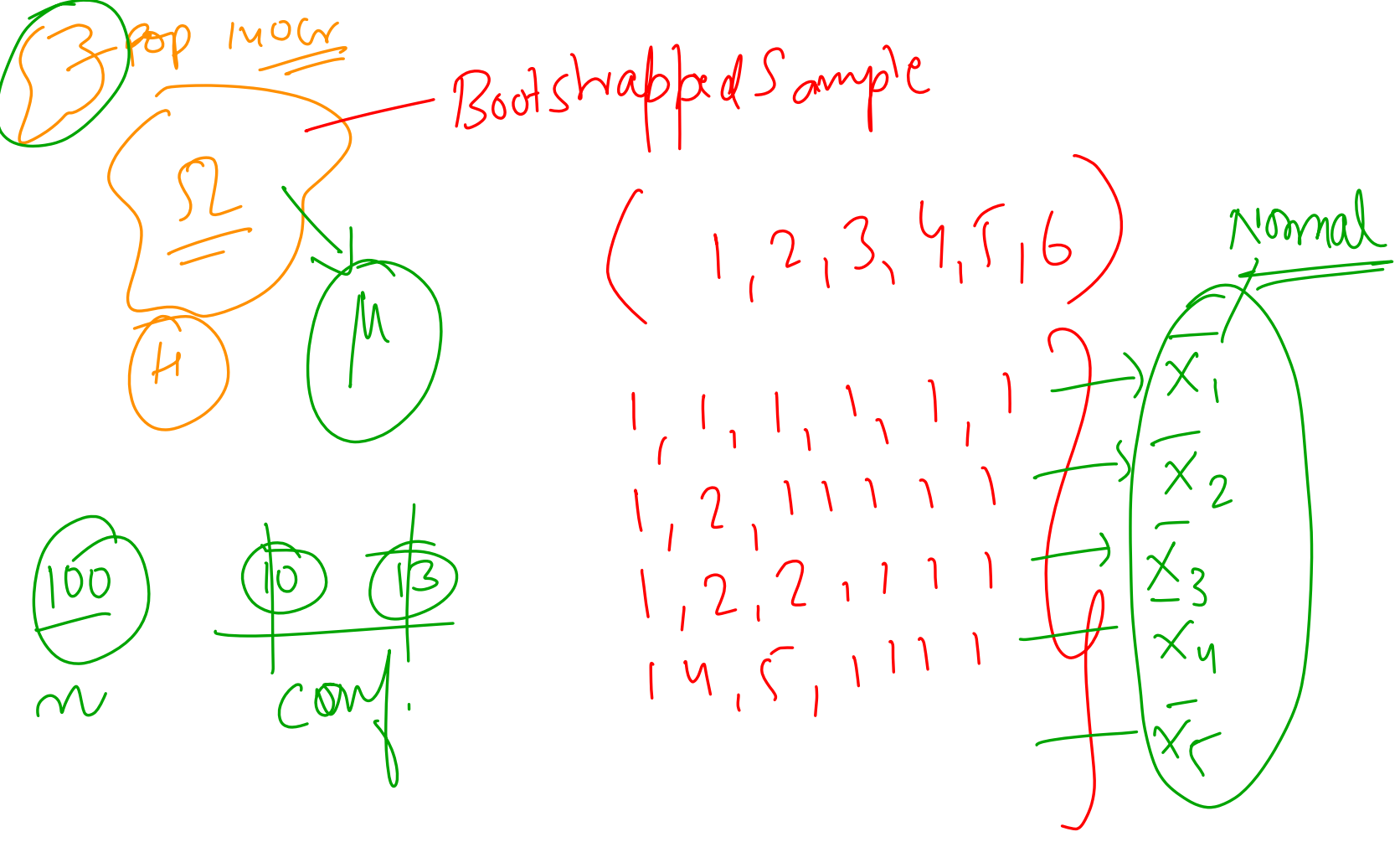
NORMALISATION

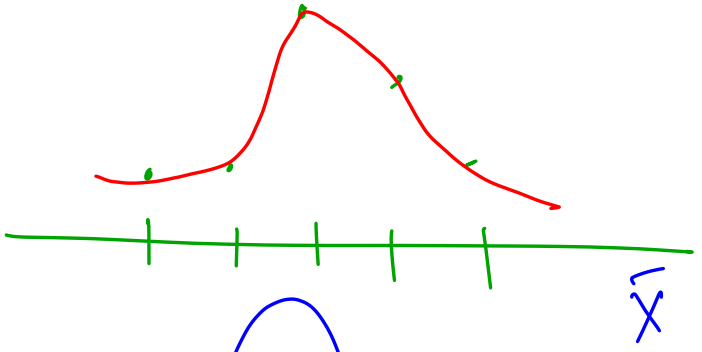
Range · known

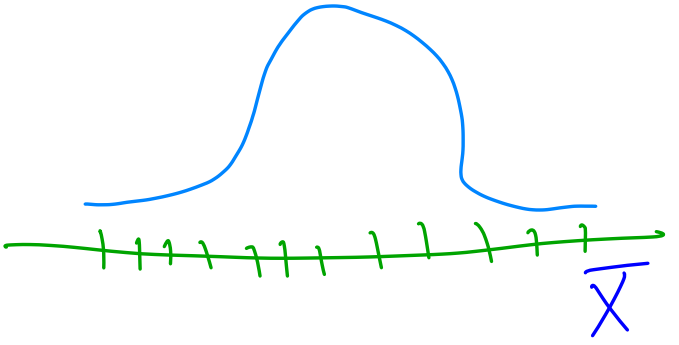


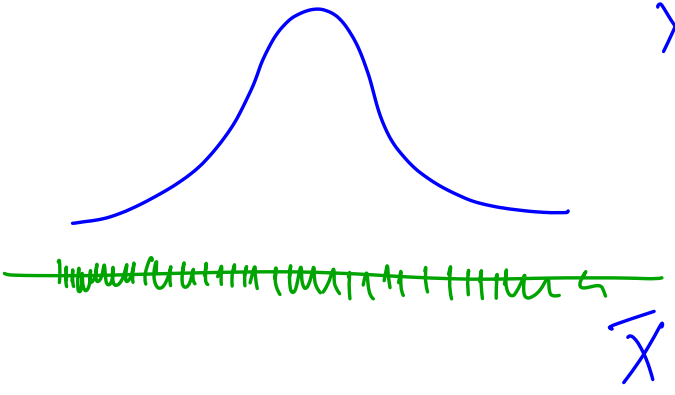
Batch Normalisation

→ Standardisation


Pop mean
Bootstrapped Sample
Normal
100
10
13
conf.


$\bar{X}$


$\bar{X}$


$\bar{X}$