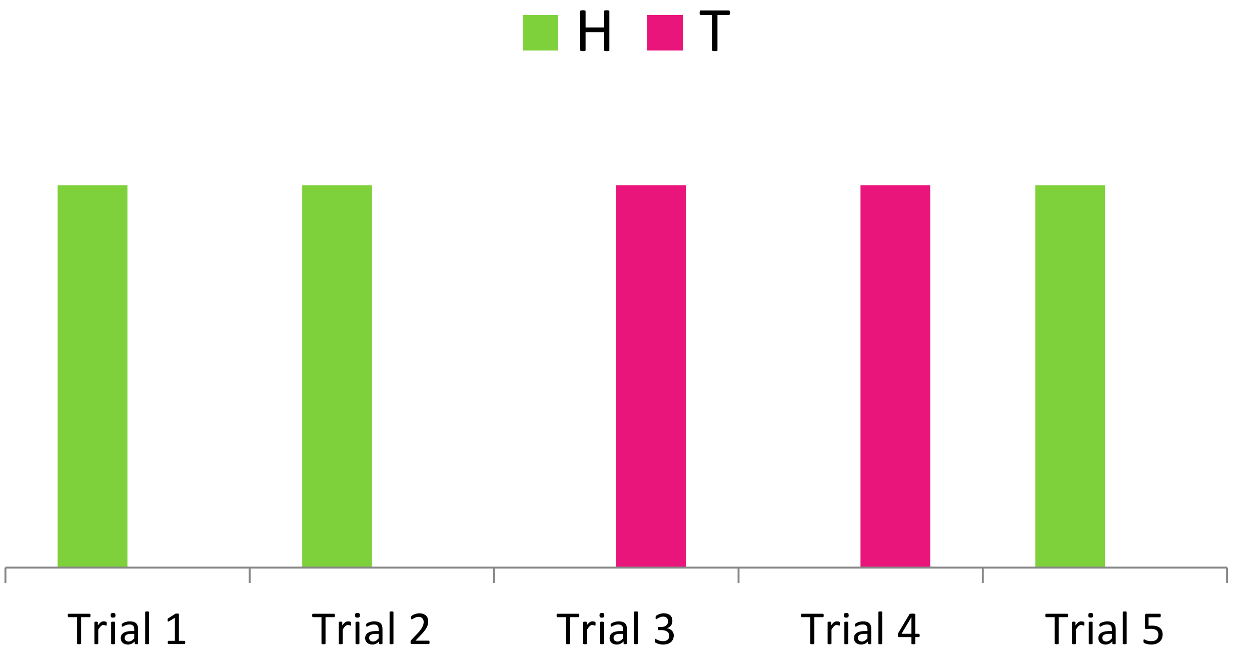
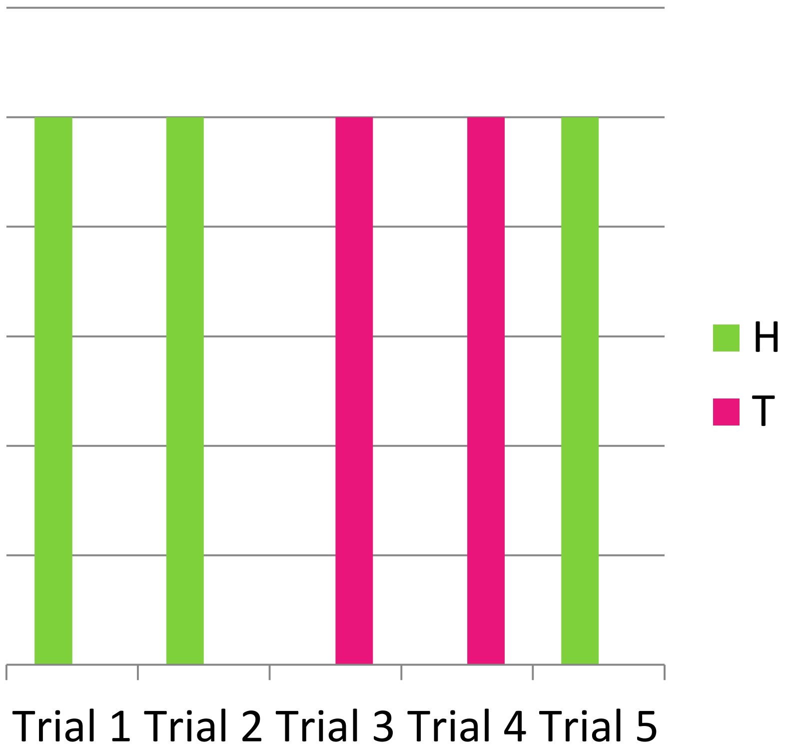
MARGIN OF ERROR

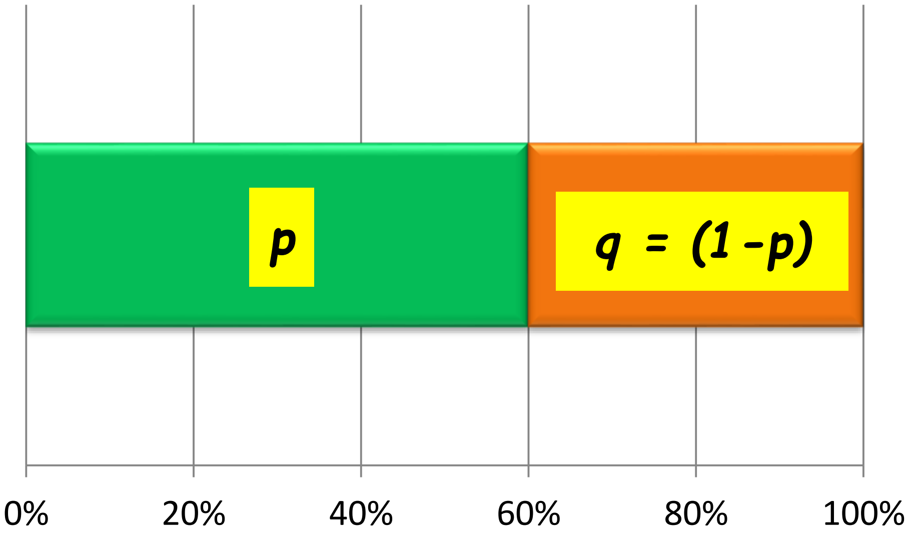
Amit Bhola

PROBABILITY



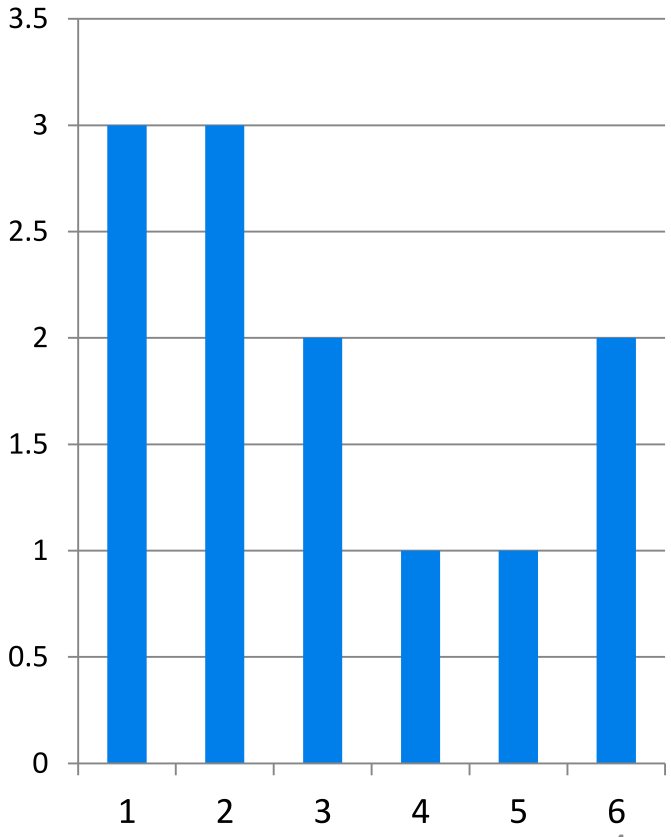
PROBABILITY



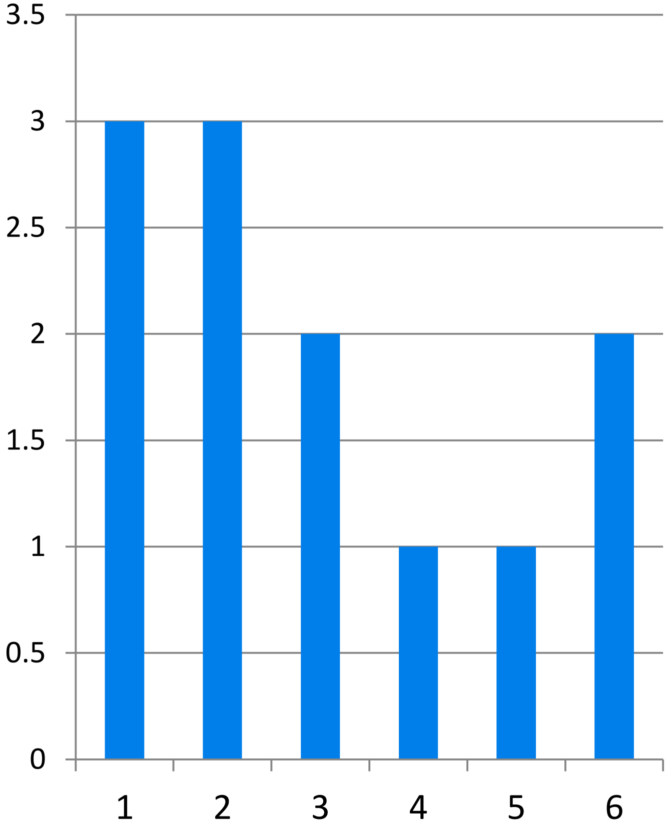


PROBABILITY

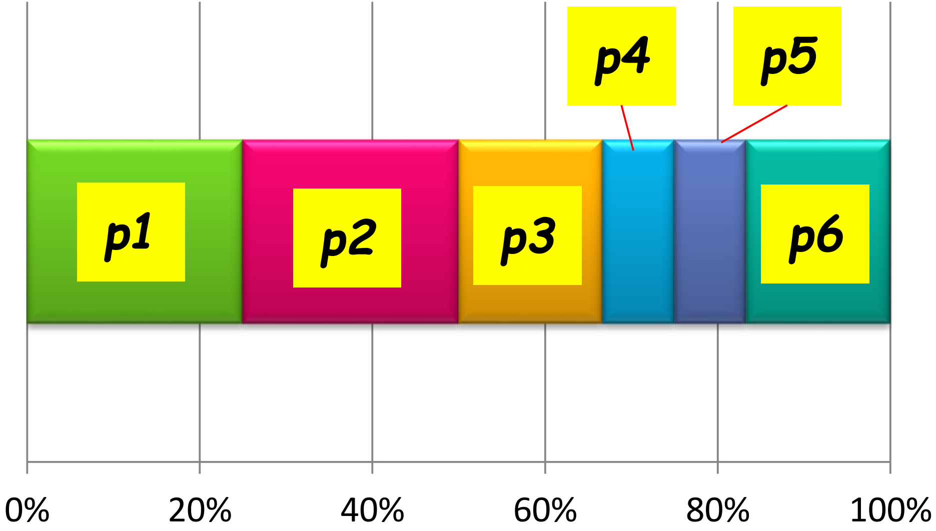
	1	2	3	4	5	6
Trial 1	●					
Trial 2		●				
Trial 3		●				
Trial 4	●					
Trial 5			●			
Trial 6			●			
Trial 7					●	
Trial 8		●				
Trial 9	●					
Trial 10						●
Trial 11				●		
Trial 12						●
	3	3	2	1	1	2



PROBABILITY



p1 + p2 + p3 + p4 + p5 + p6 = 1



PROBABILITY

Lucky No?

Pollution Level?

Height?

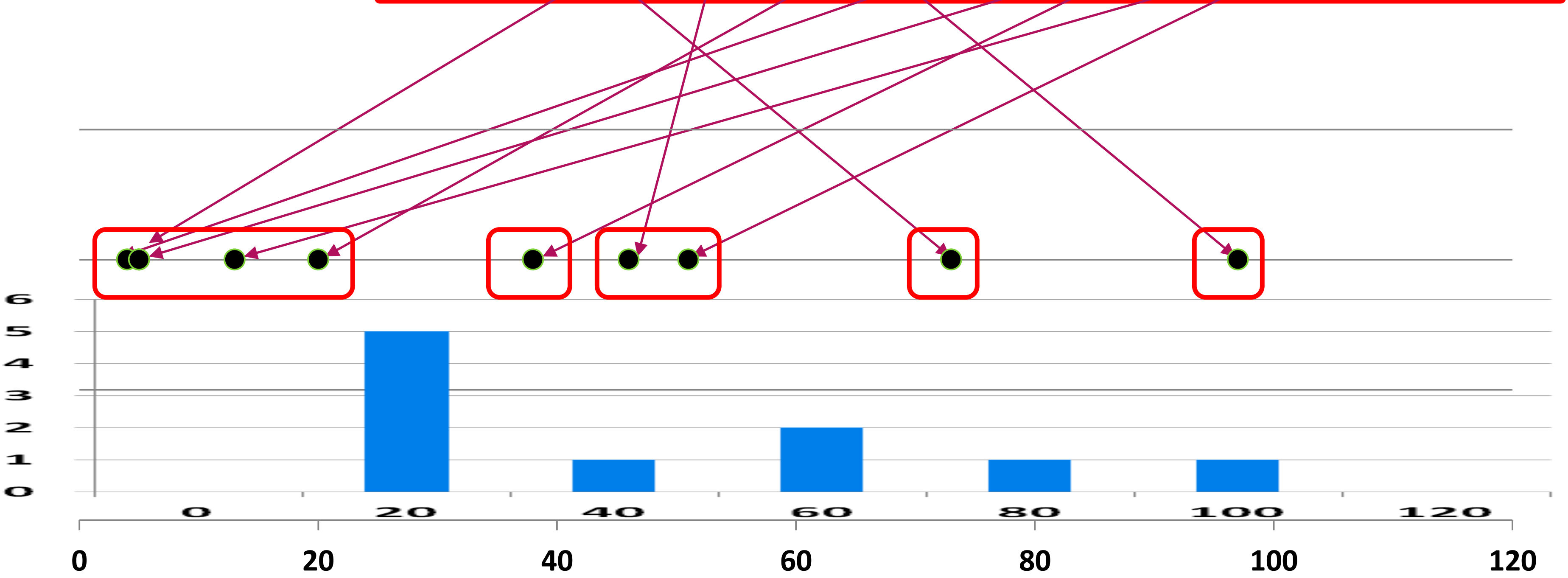
Person	1	2	3	4	5	6	7	8	9	10	11	12	...	...
Response	5	73	46	20	4	97	5	38	13	51	...	...	...	

The response / outcome possible is **NOT** fixed / restricted.

Theoretically **EVERY** real no. has a Non Zero Probability

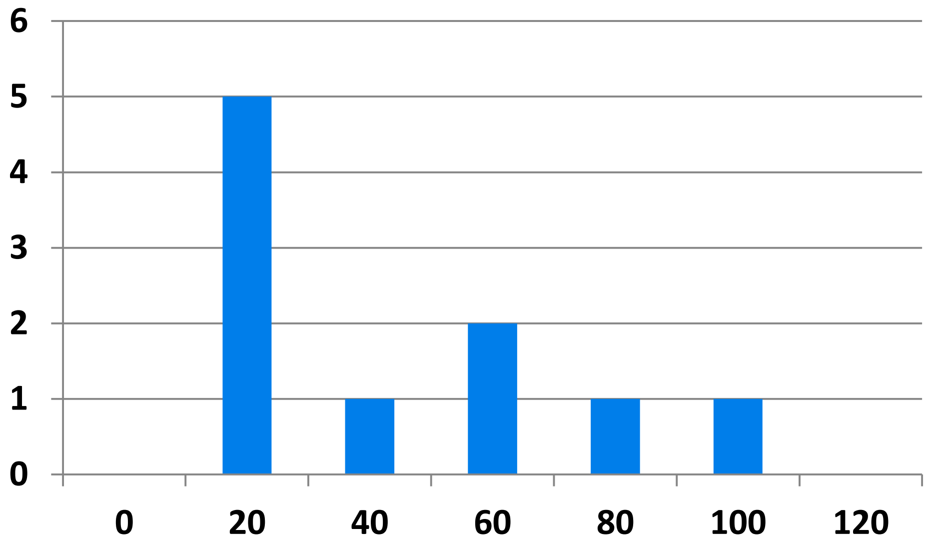
PROBABILITY

Person	1	2	3	4	5	6	7	8	9	10	11	12	…	…
Response	5	73	46	20	4	97	5	38	13	51	…	…	…	



PROBABILITY

Person	1	2	3	4	5	6	7	8	9	10	11	12	…	…
Response	5	73	46	20	4	97	5	38	13	51	…	…	…	



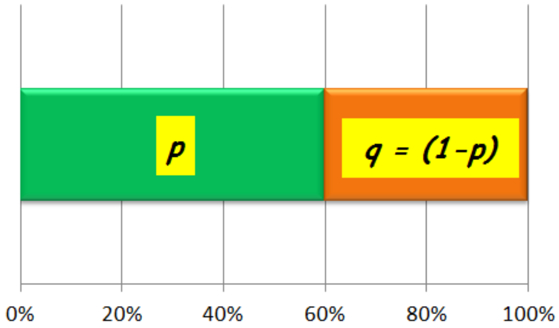
HISTOGRAMS

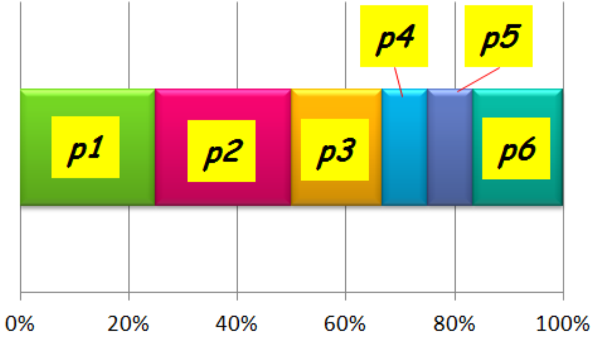
TELL

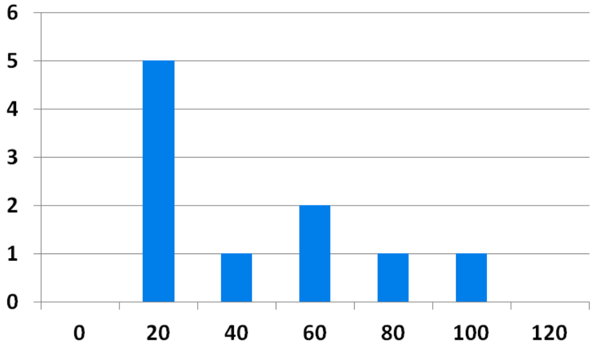
PROBABILITY

WHAT NEXT ?

PROBABILITY

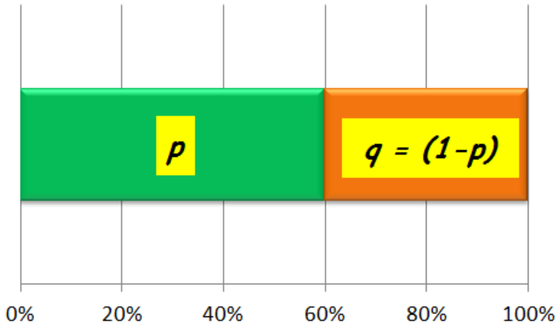


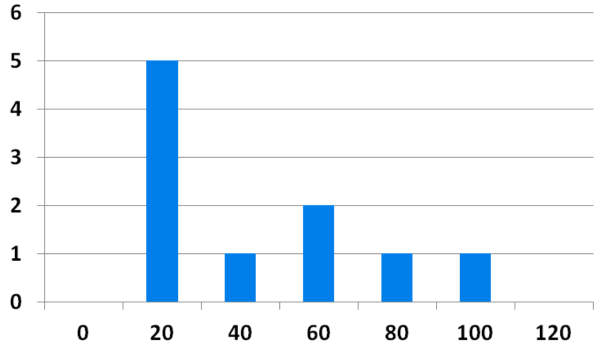




BINOMIAL	MULTINOMIAL	CONTINUOUS
p	p1	p (R1) = f(R1)
q=1-p	p2 ...	p (R2) = f(R2) ...
p+q=1	p1+p2+p3+...+p**n**= 1	∑ p (all R) = 1

PROBABILITY



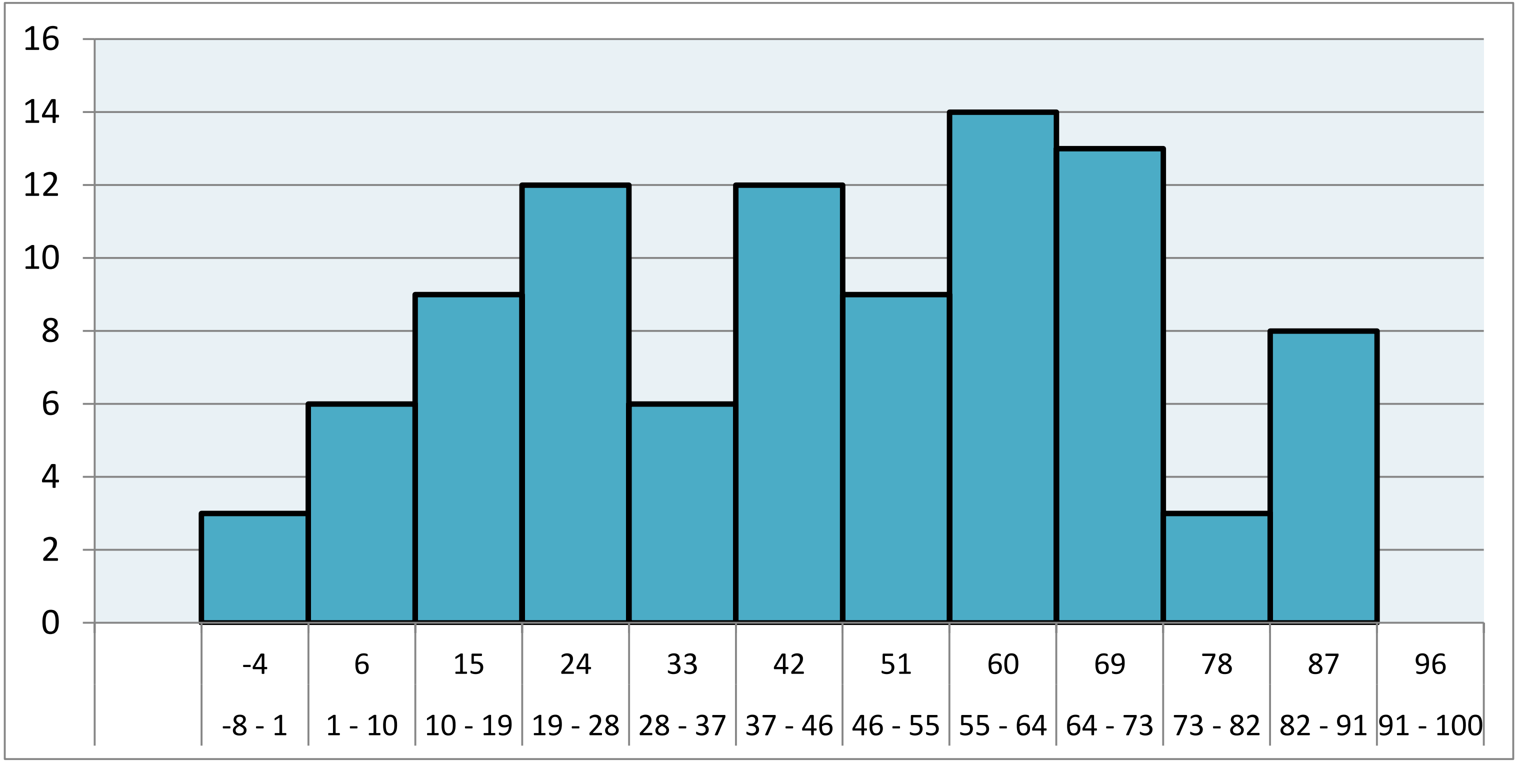


RARE
AND
COMPLEX

OUT
OF
SCOPE

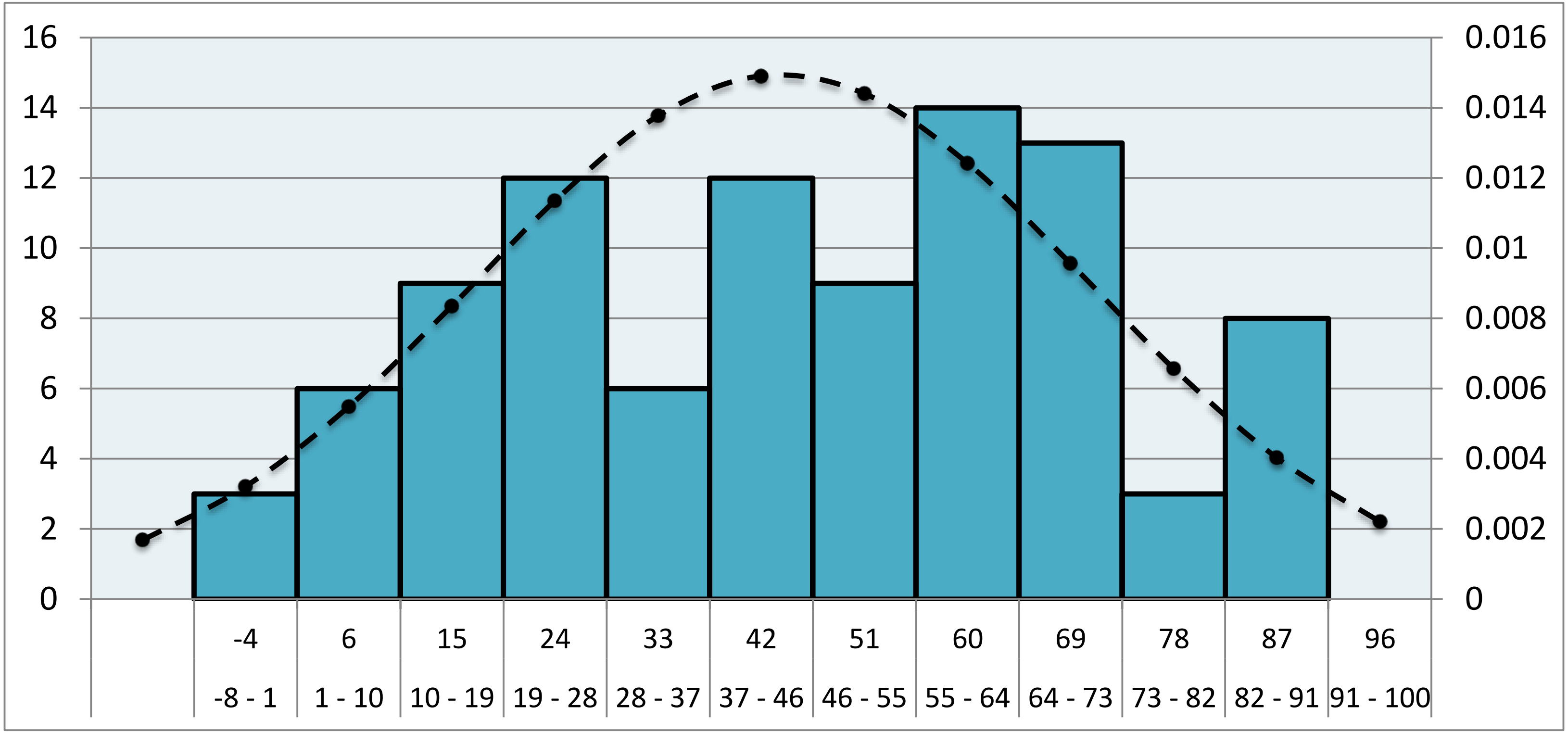
BINOMIAL	CONTINUOUS
Formulae 1	Formulae 2

HISTOGRAMS

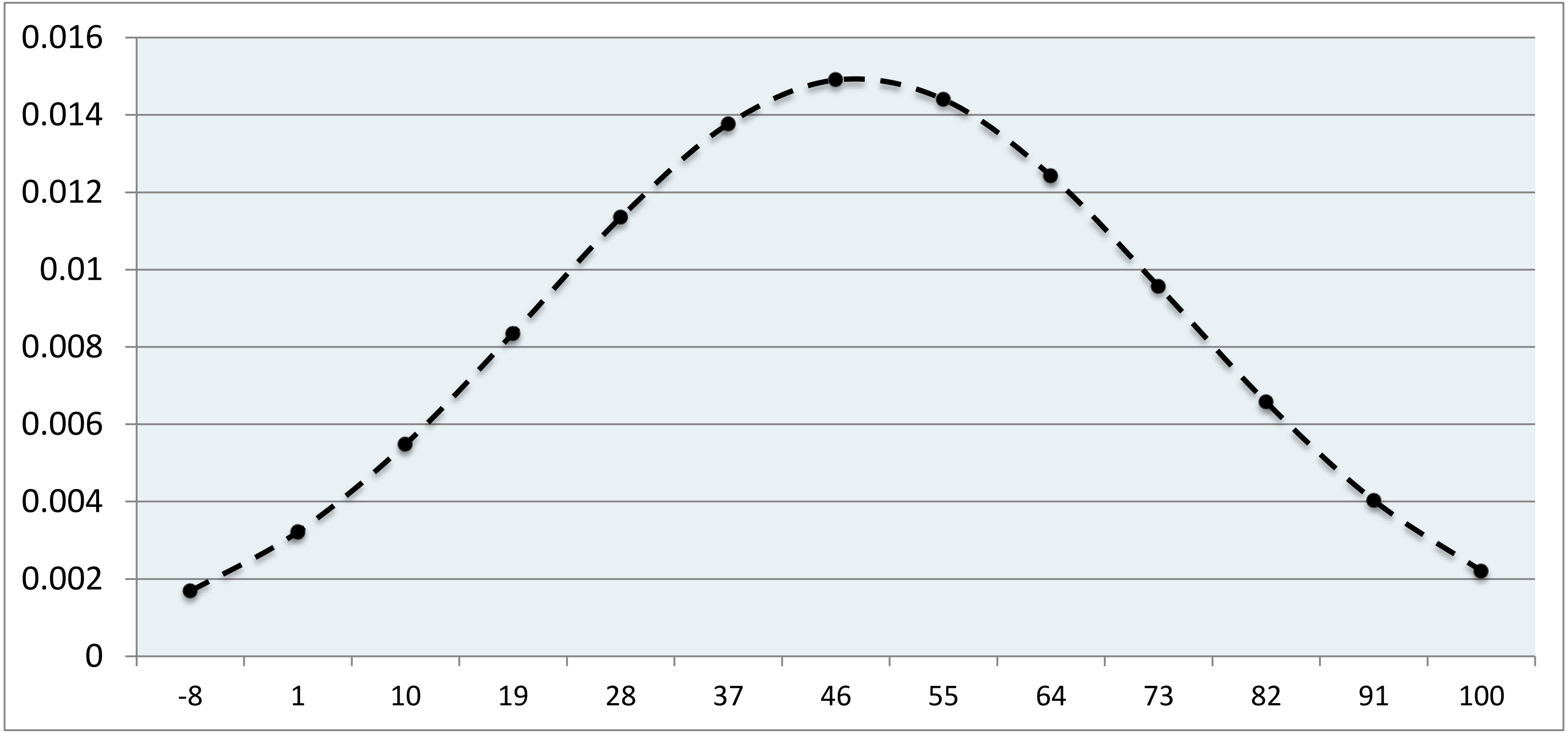


16
14
12
10
8
6
4
2
0
-4
6
15
24
33
42
51
60
69
78
87
96
-8 - 1
1 - 10
10 - 19
19 - 28
28 - 37
37 - 46
46 - 55
55 - 64
64 - 73
73 - 82
82 - 91
91 - 100

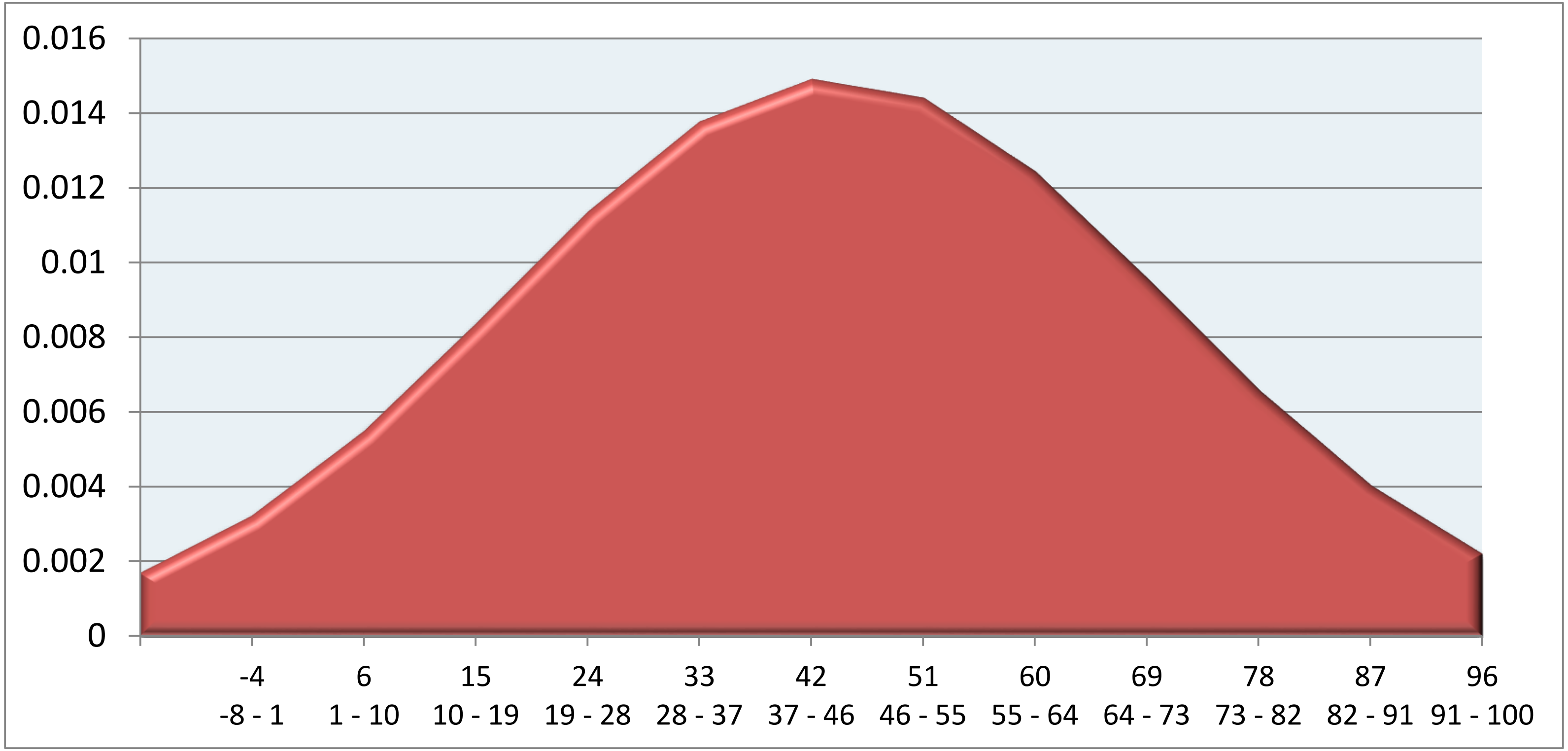
PROBABILITY (DENSITY) FUNCTION



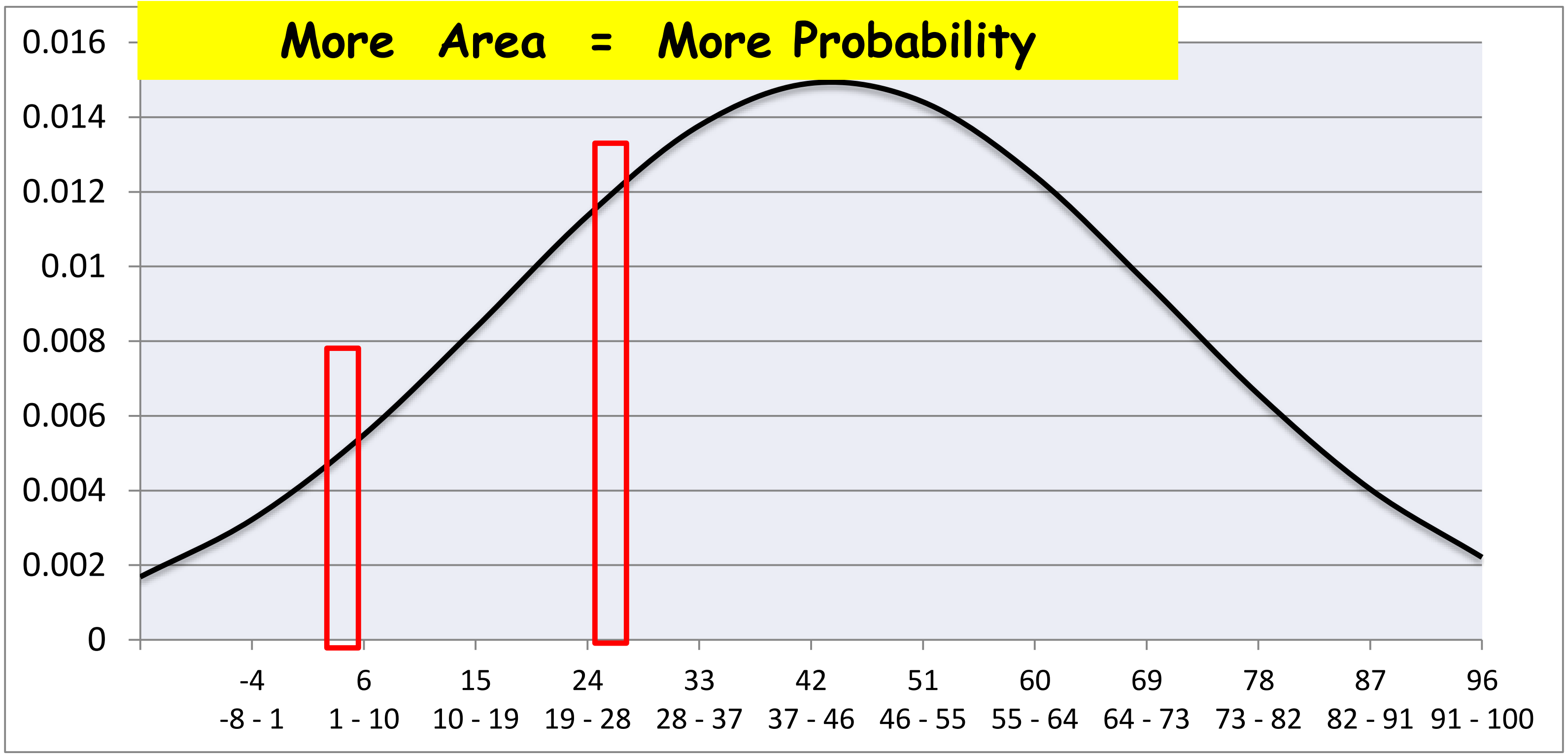
PROBABILITY FUNCTION



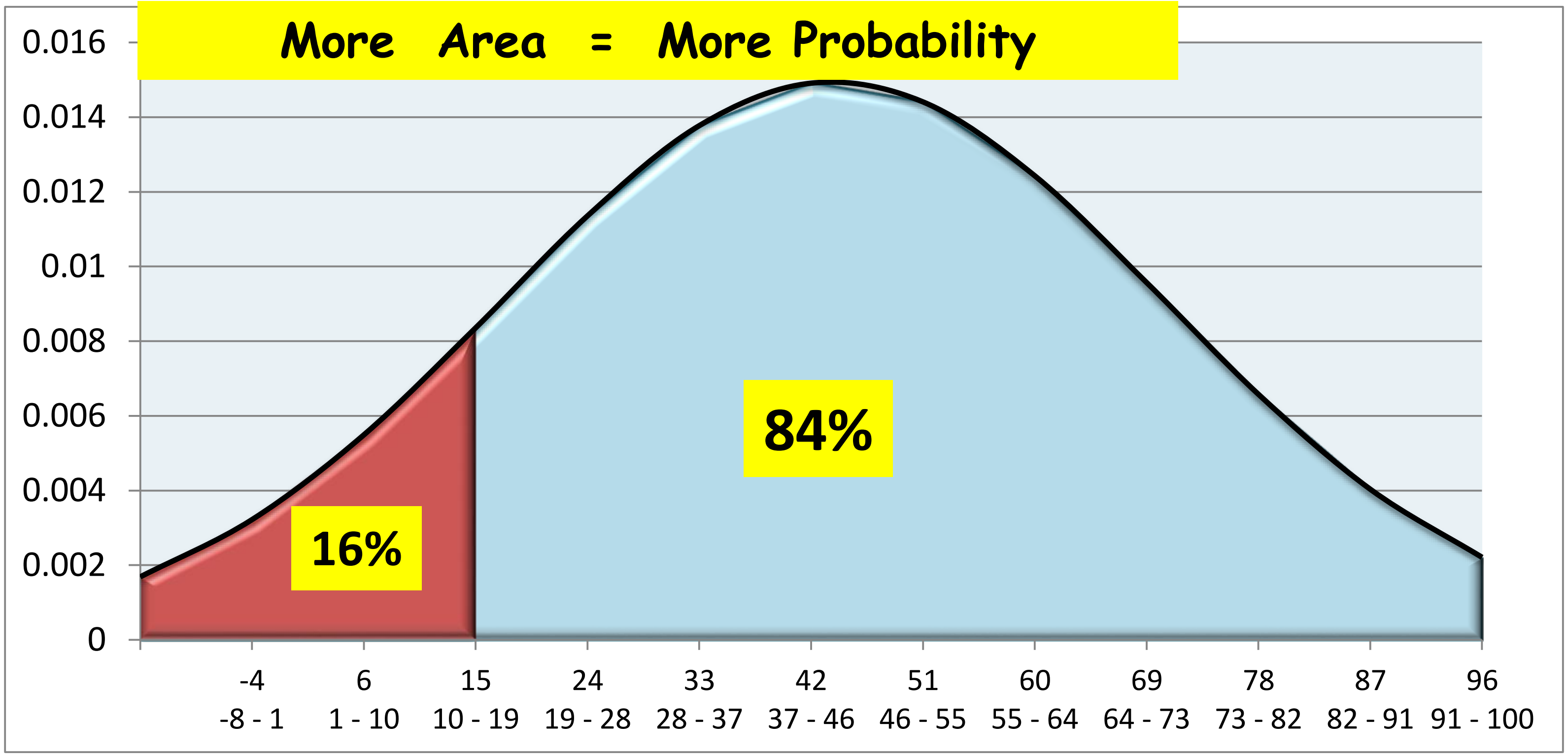
CUMULATIVE DENSITY FUNCTION



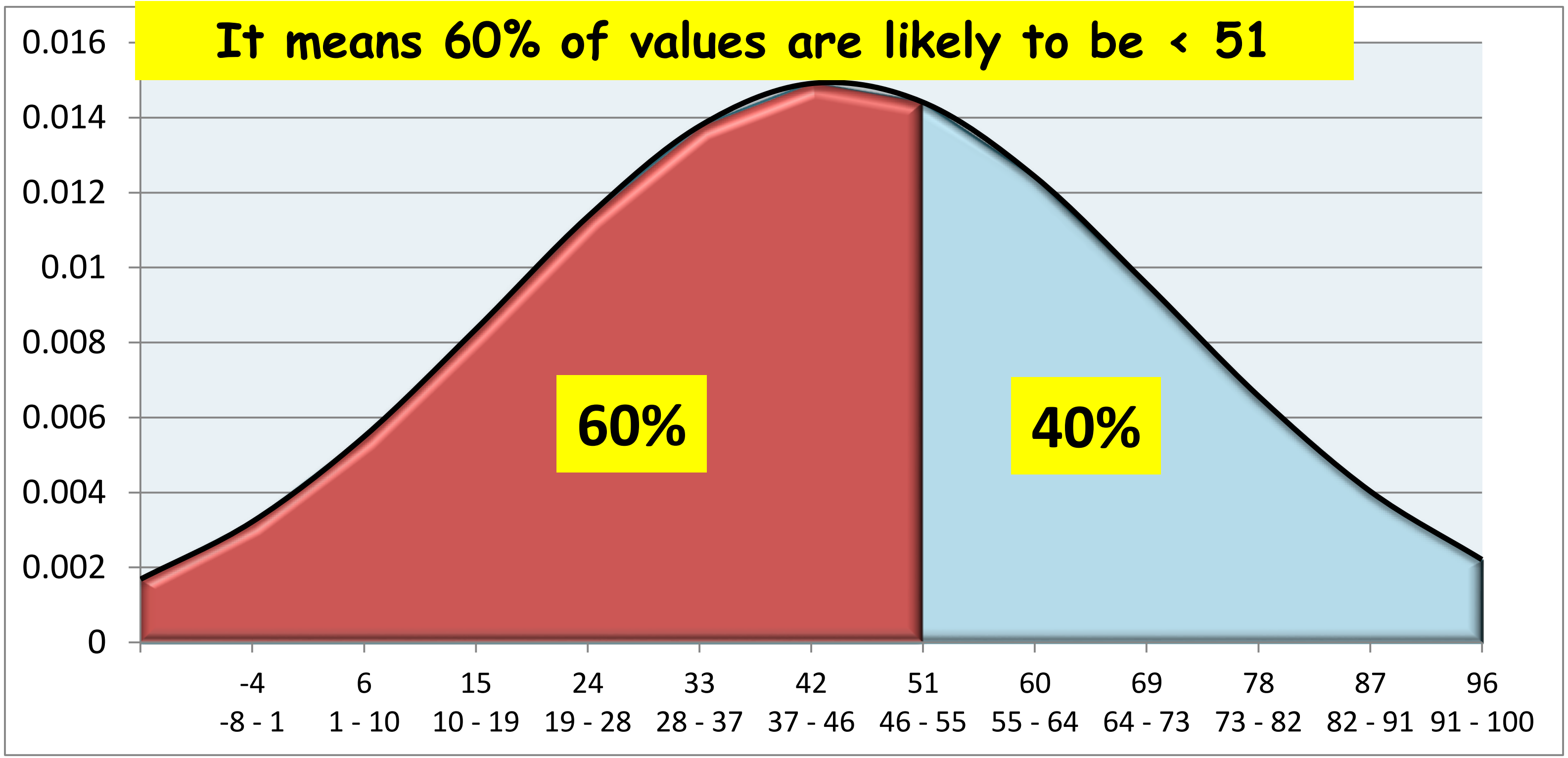
P.D.F. AS PROBABILITY INDICATOR



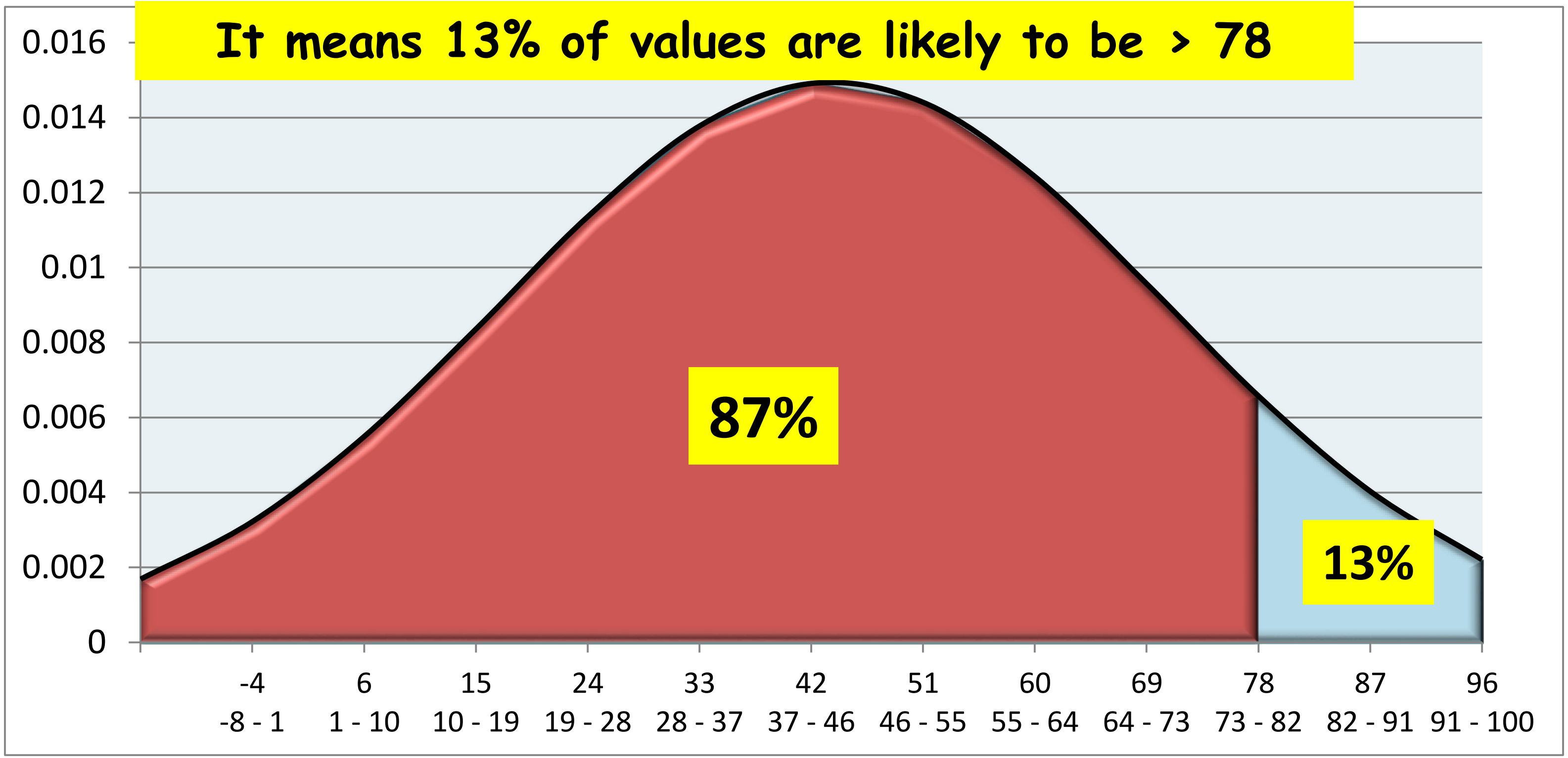
RANGE PROBABILITY FROM C.D.F.



RANGE PROBABILITY FROM C.D.F.

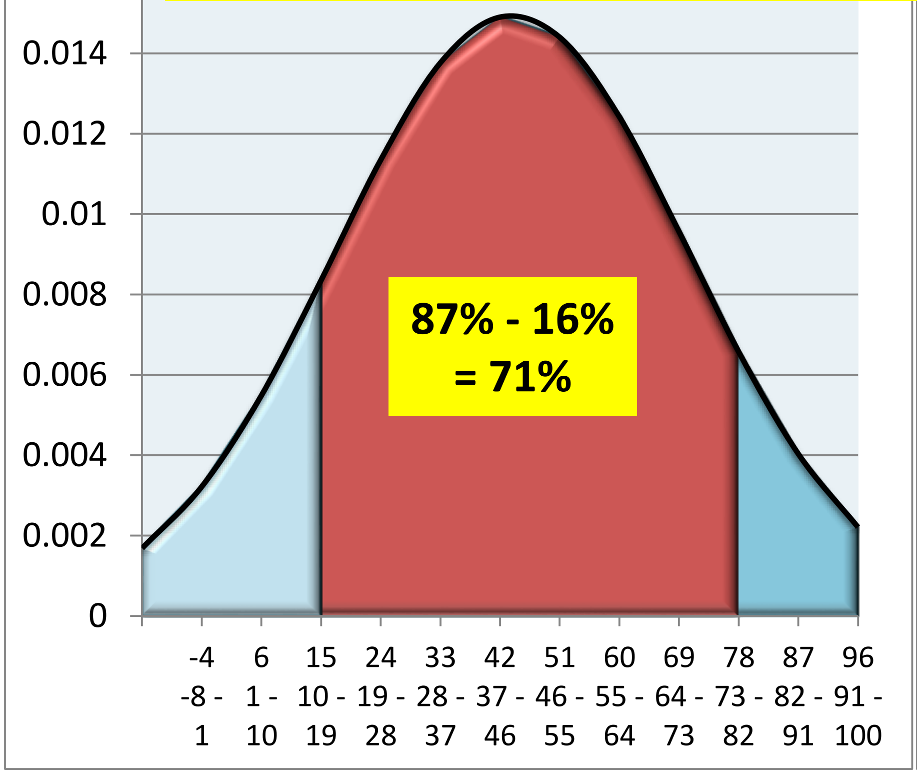


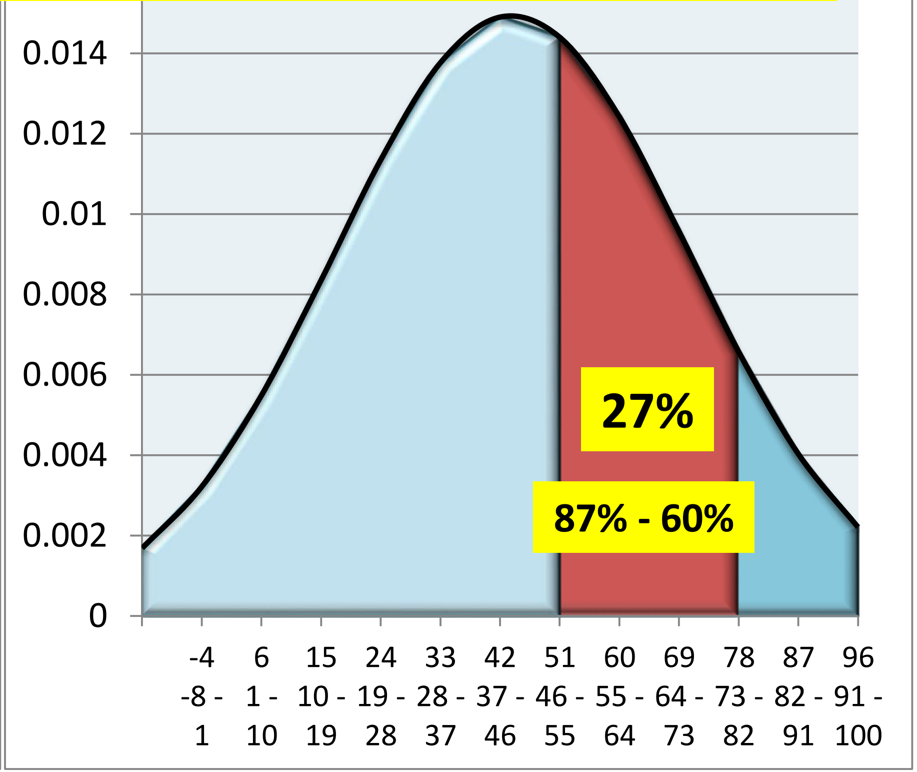
RANGE PROBABILITY FROM C.D.F.



RANGE PROBABILITY FROM C.D.F.

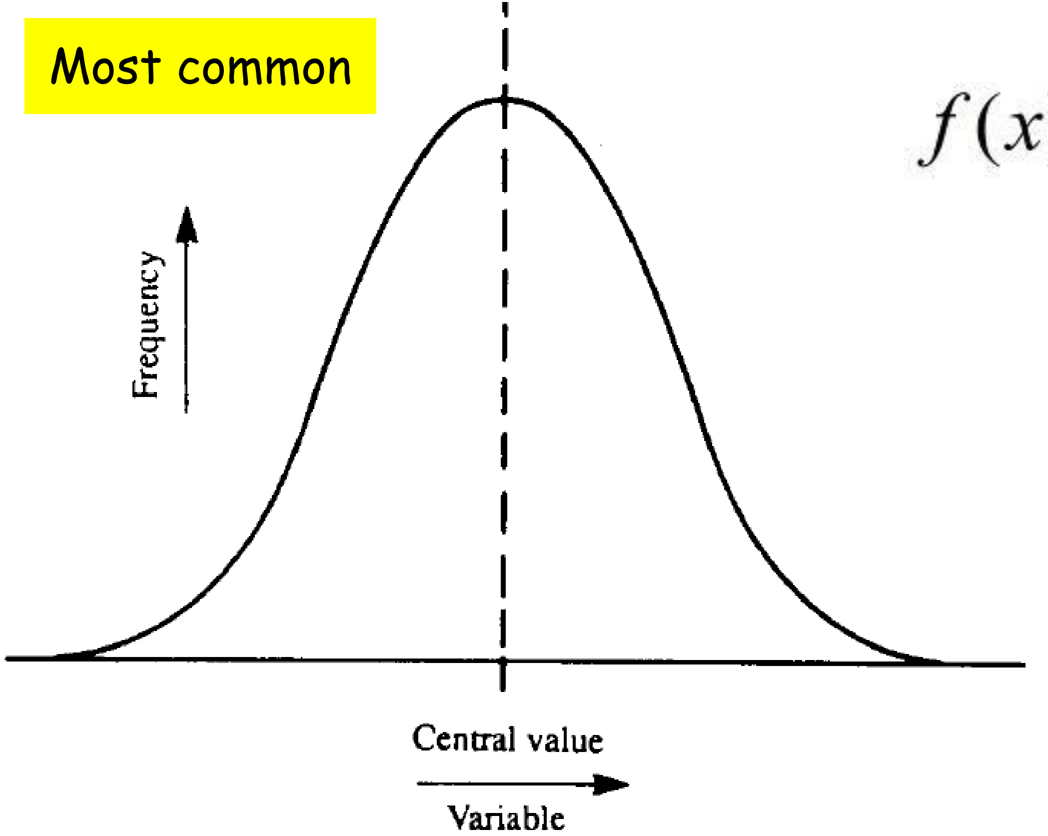
It means 27% of values are likely to be between 51 and 78





NORMAL DISTRIBUTION

Most common

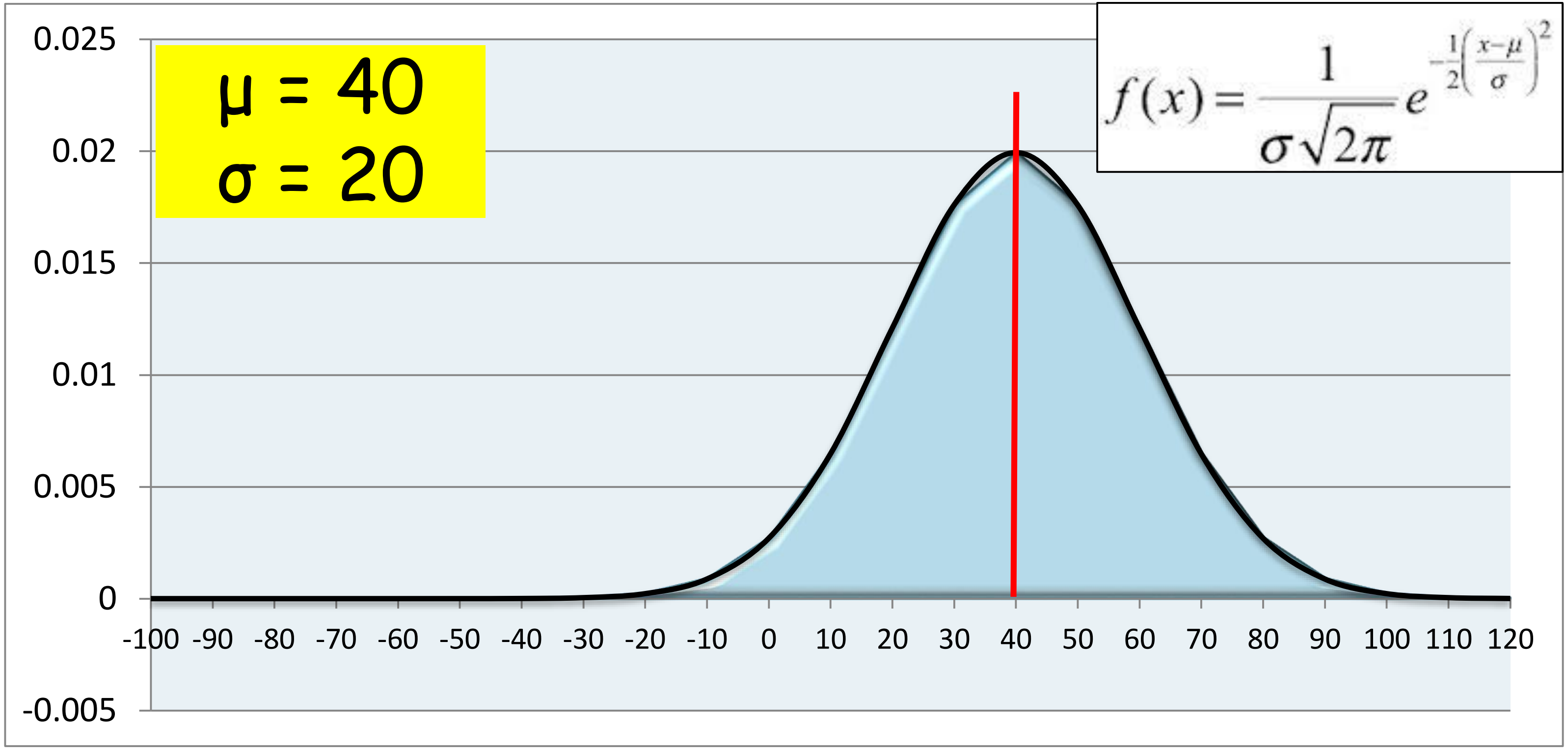


$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

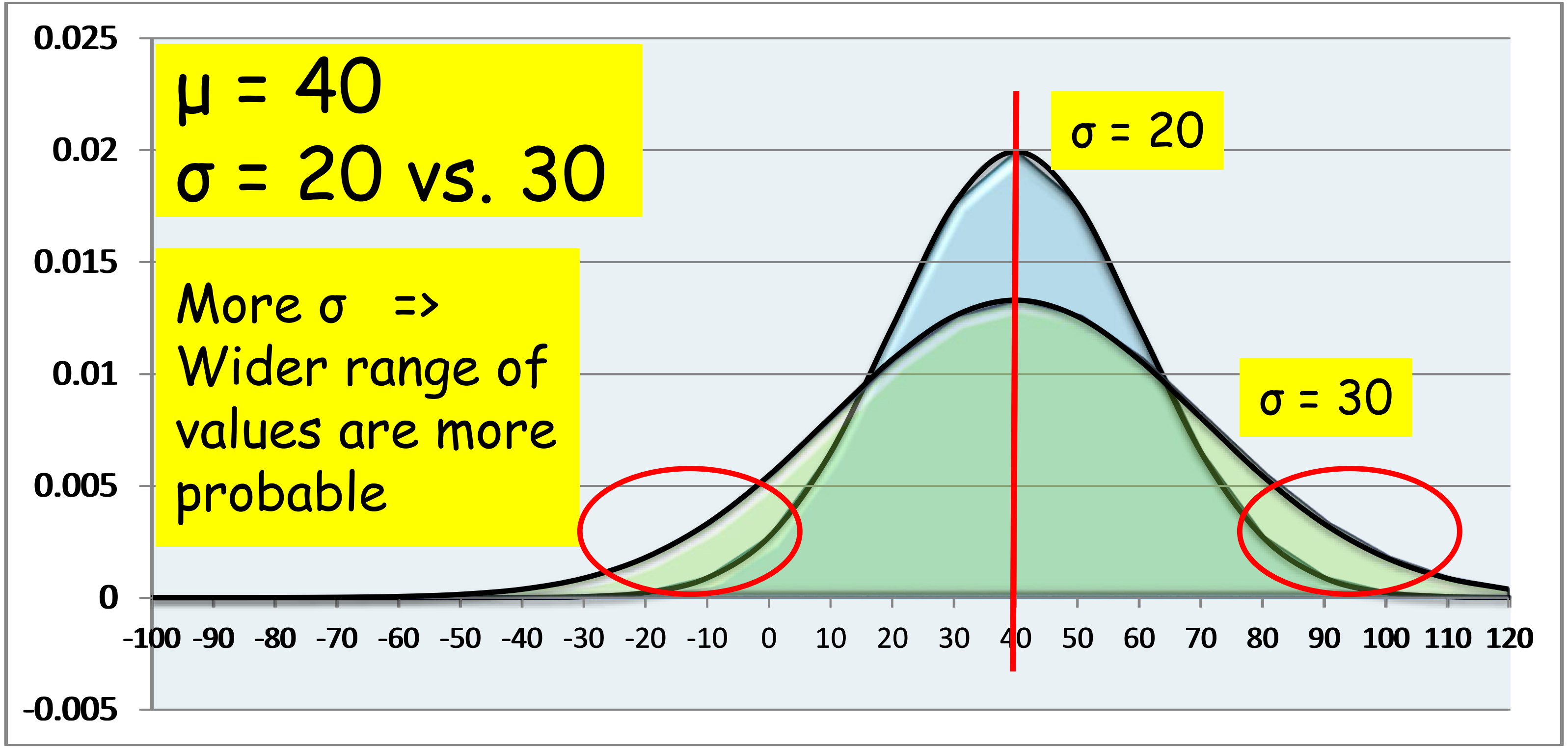
It means:-

Knowing **μ and σ** of a **normally** distributed variable, one can determine how much **probable** is it to **lie between a range**.

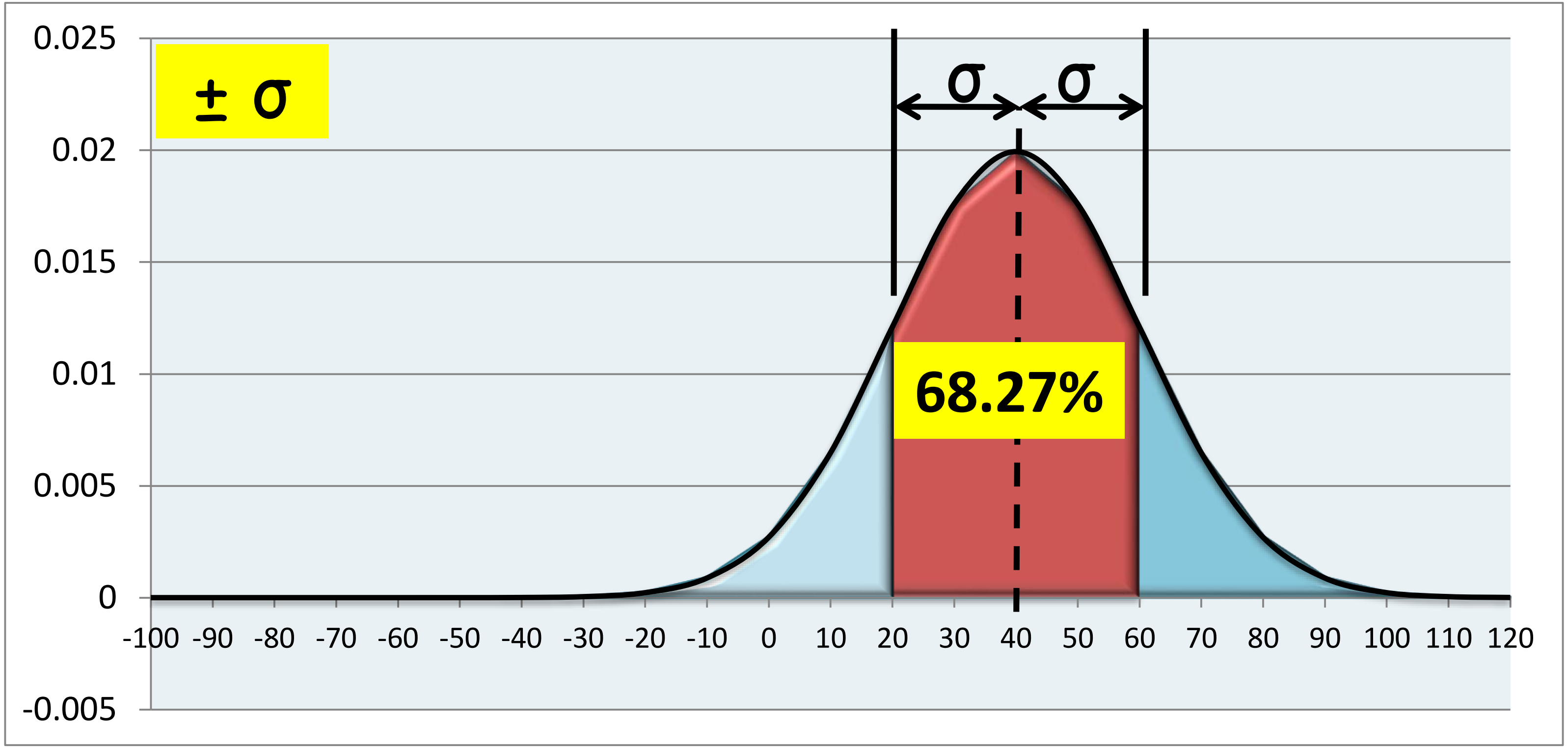
MEAN AND STANDARD DEVIATION



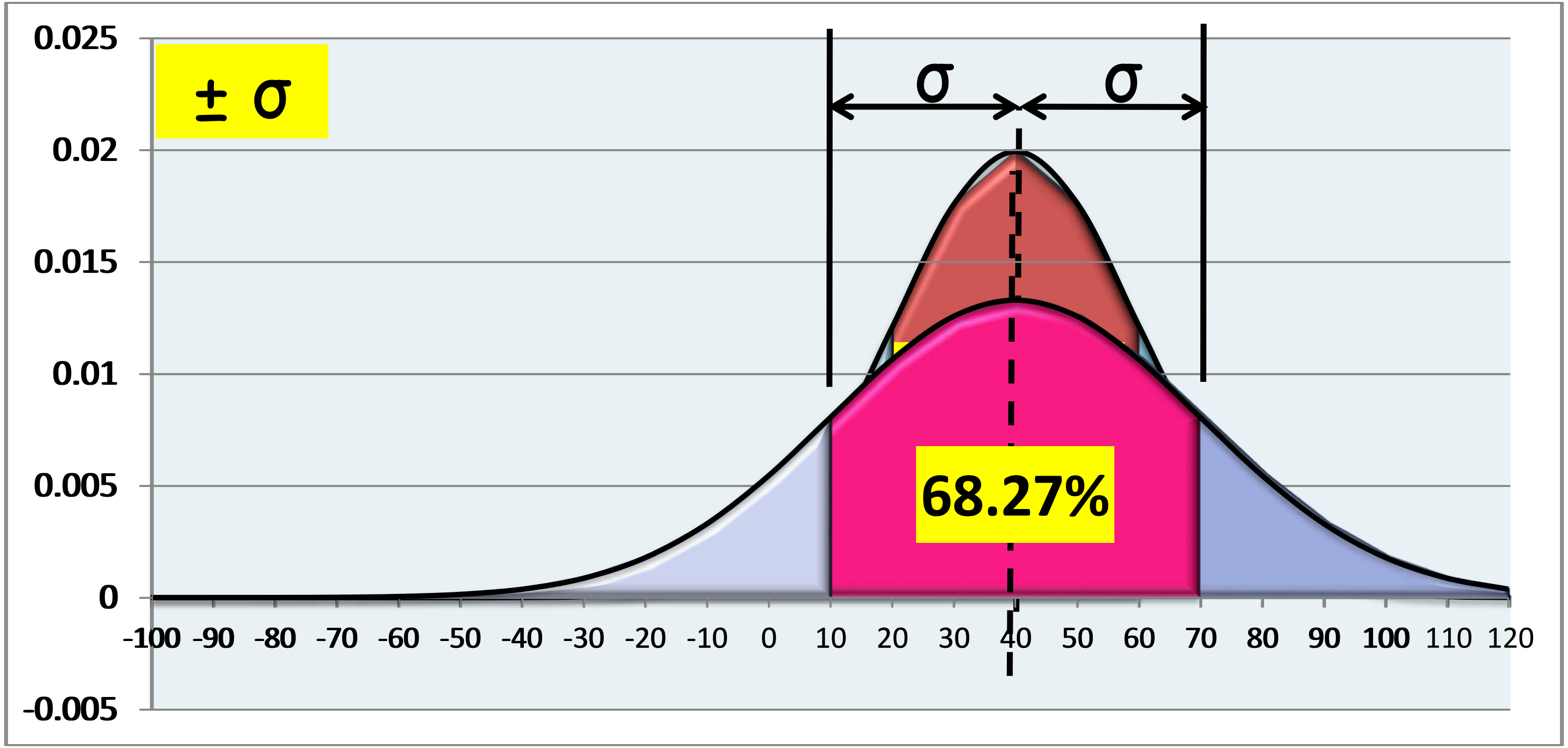
MEAN AND STANDARD DEVIATION



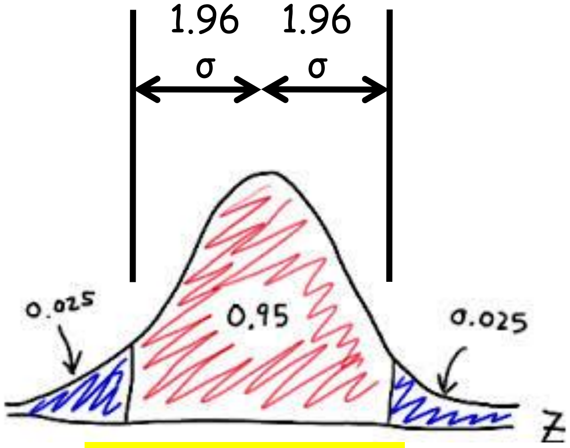
AREA COVERED BETWEEN STD DEV



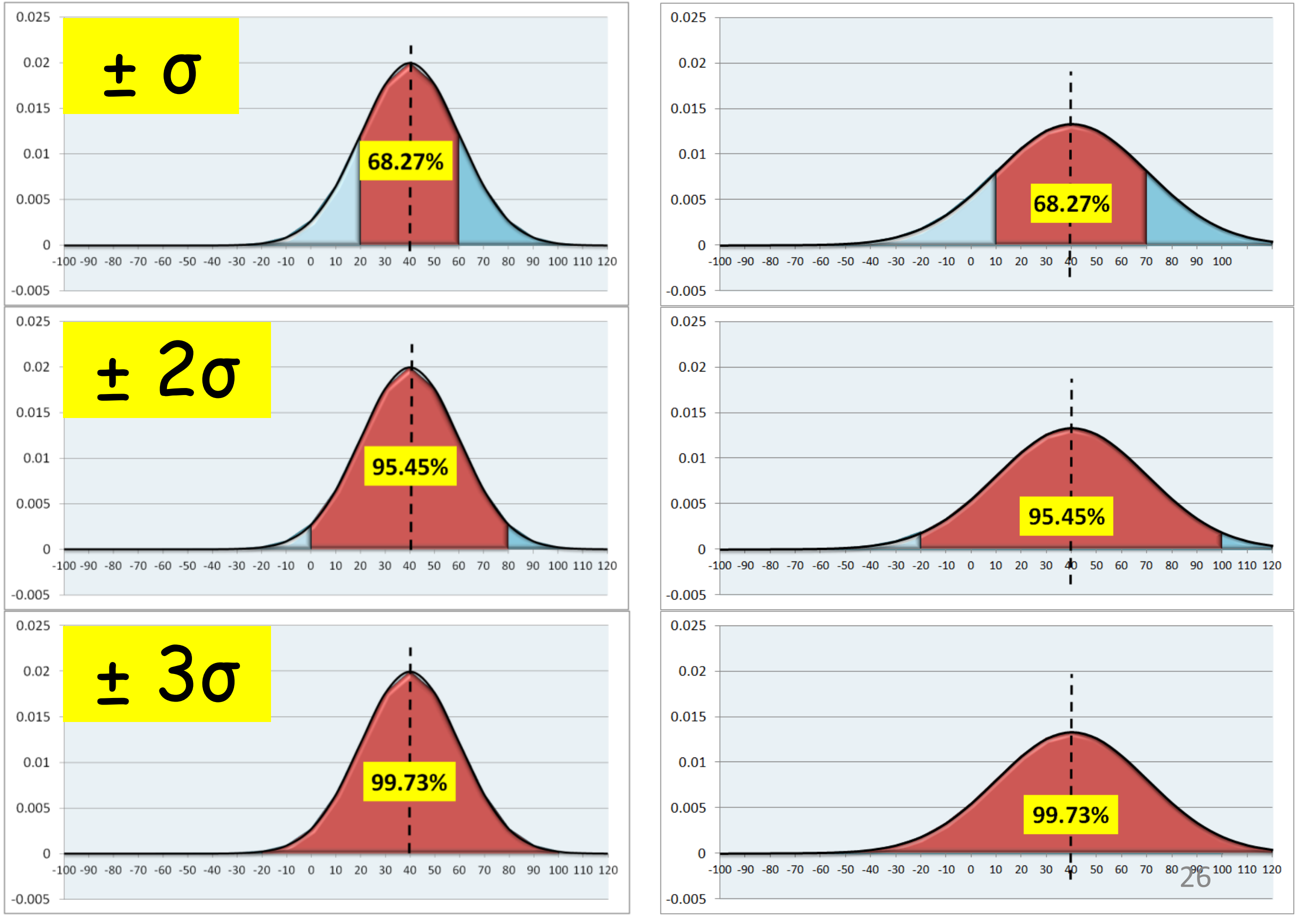
AREA COVERED BETWEEN STD DEV



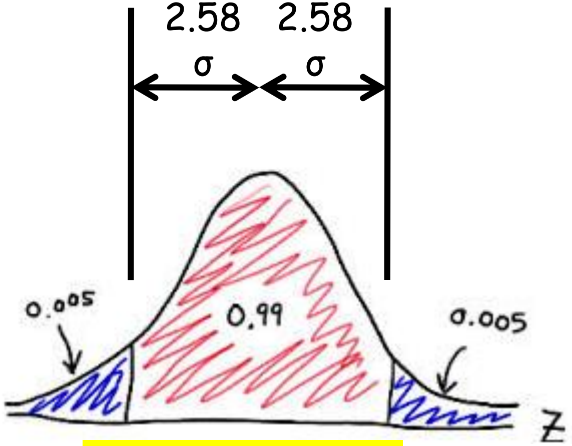
AREA COVERED BETWEEN STD DEV



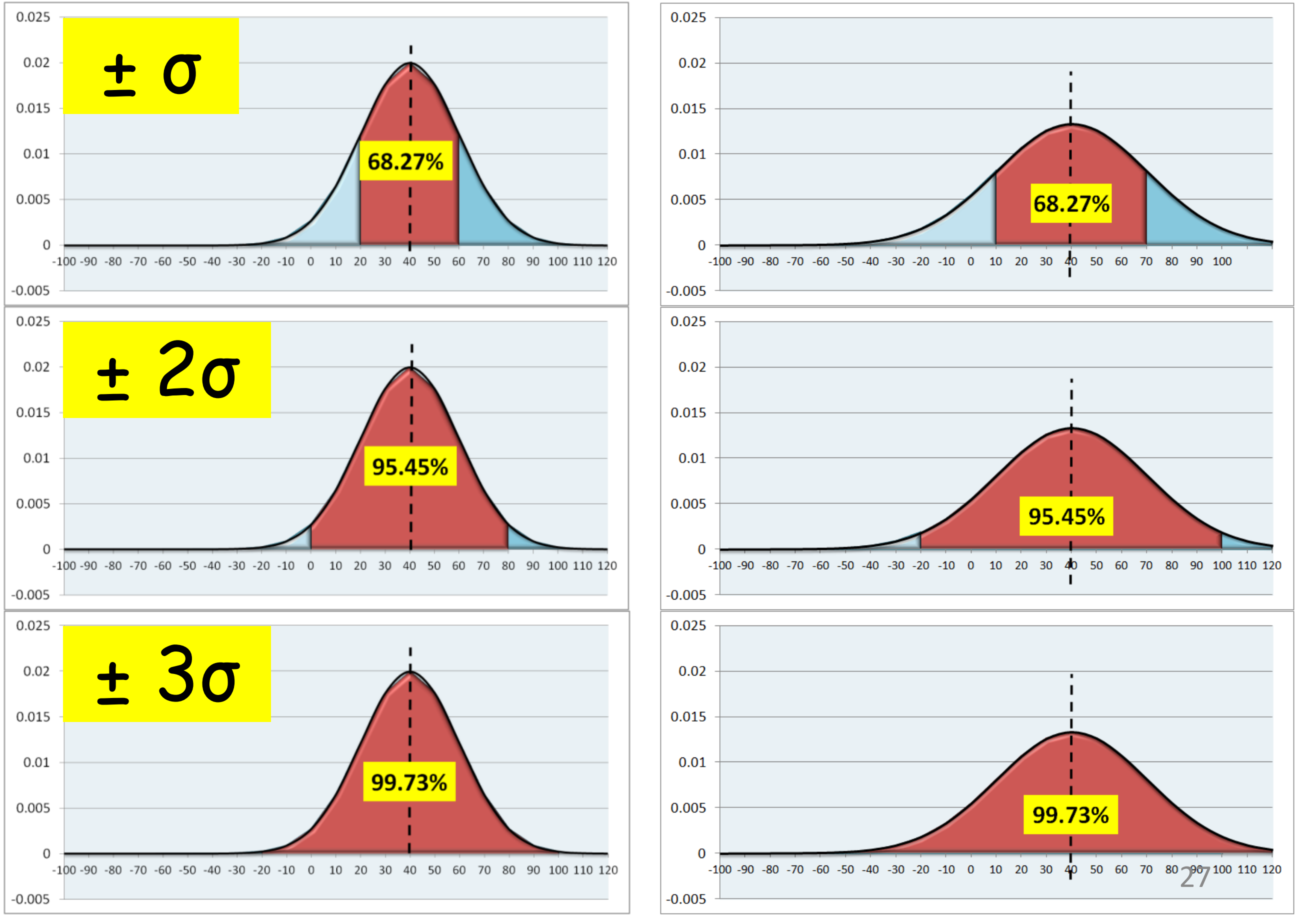
± 1.96σ



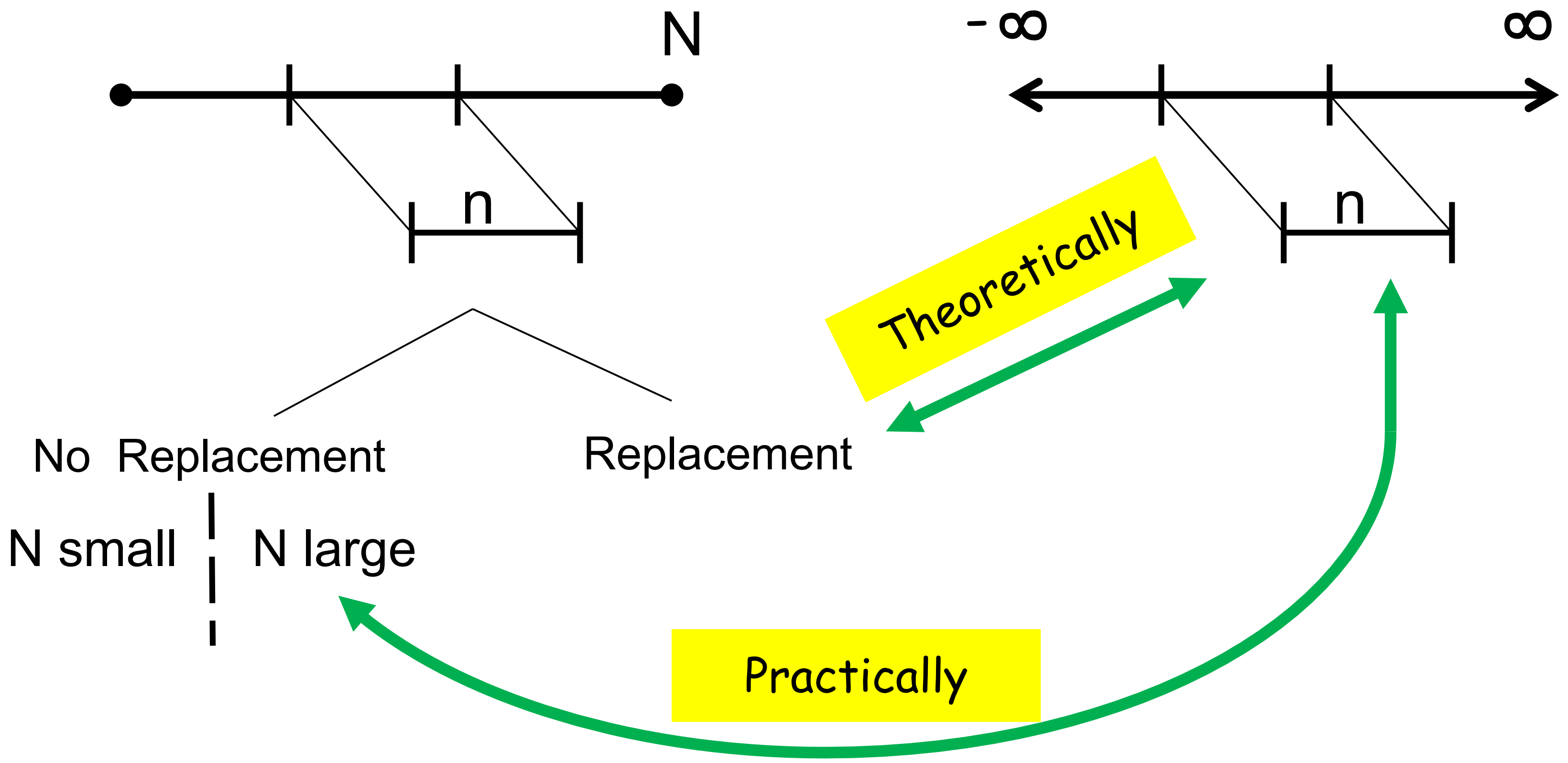
AREA COVERED BETWEEN STD DEV



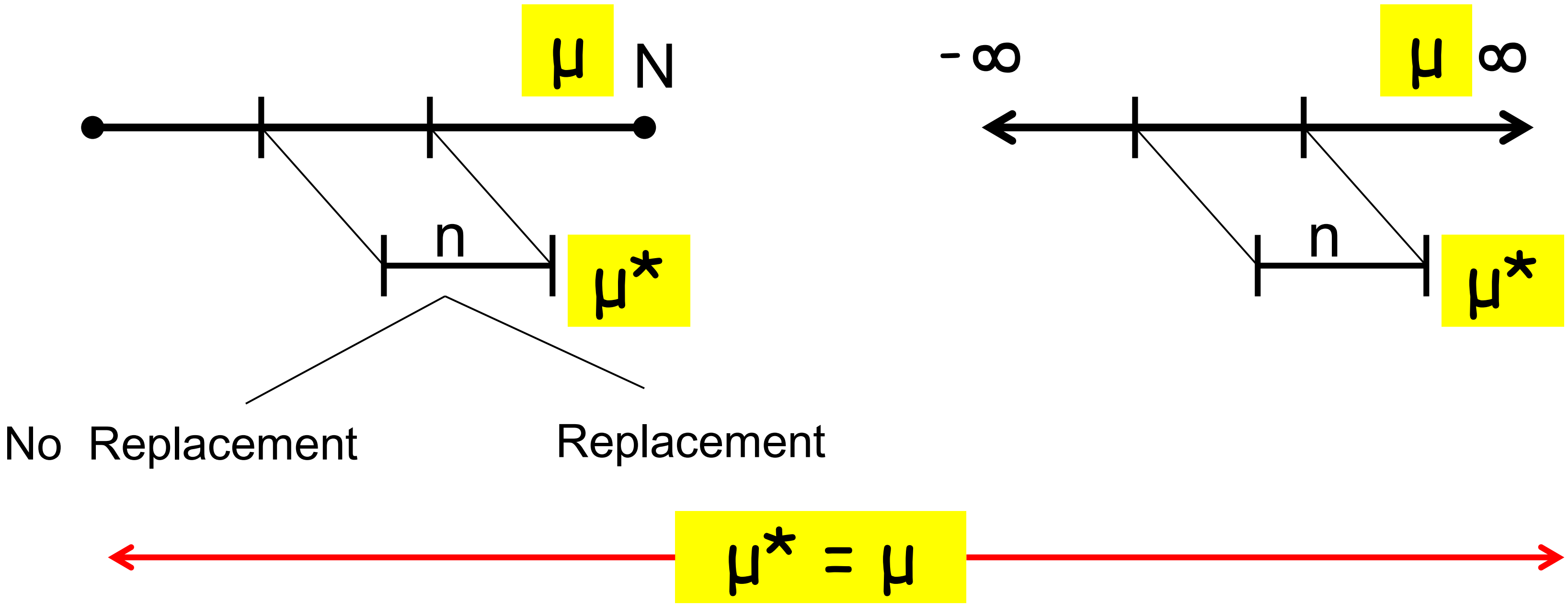
± 2.58σ



SAMPLING THEORY



SAMPLING INFERENCE - MEAN



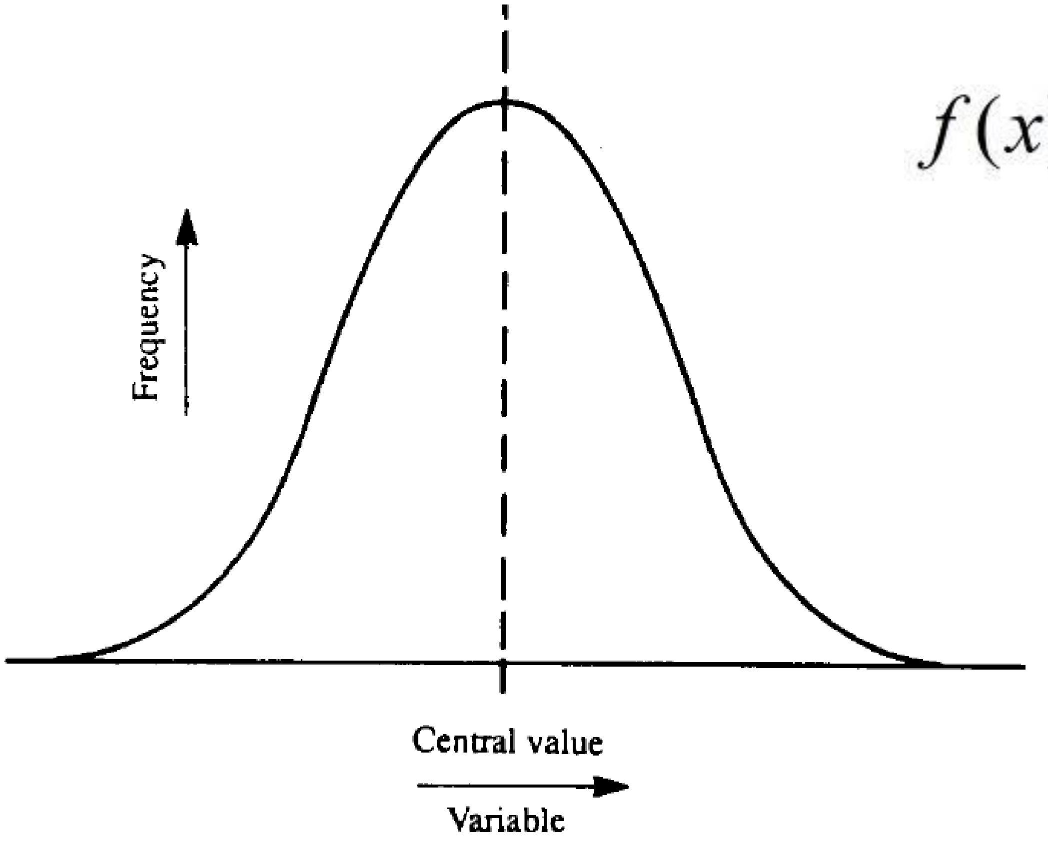
Sample mean = Population mean
With some error

SAMPLING INFERENCE - MEAN

I sampled the ABC tax payments. I conclude that Avg. tax as 1.2 MRs

How much sure are you about your inference ?

NORMAL DISTRIBUTION

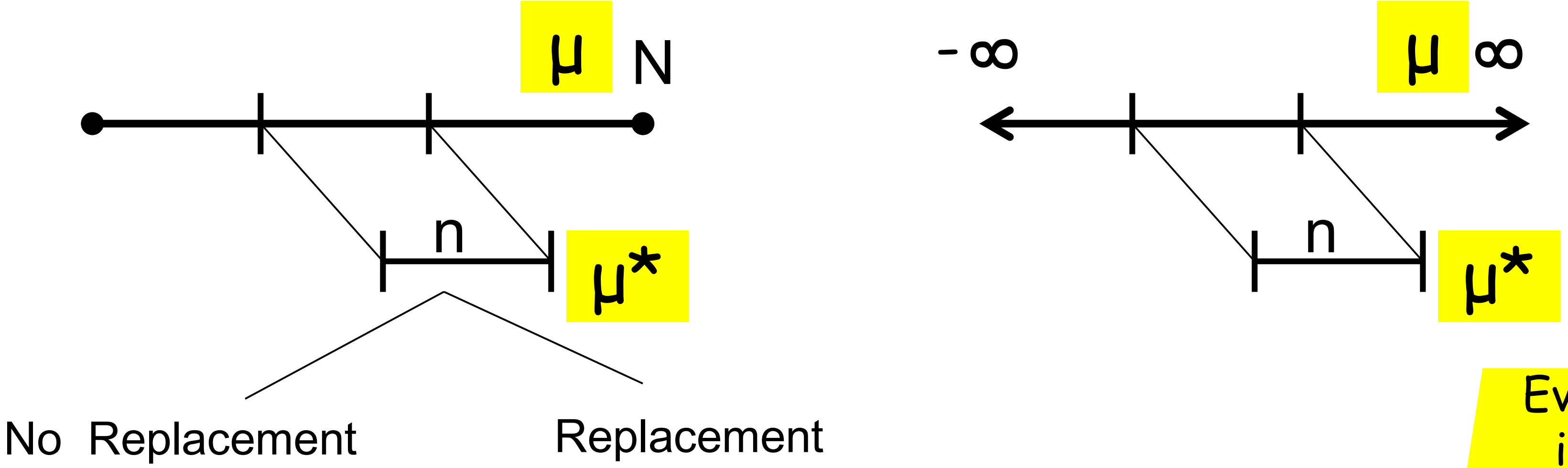


$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

It means:-

Knowing **μ and σ** of a **normally** distributed variable, one can determine how much **probable** is it to **lie between a range**.

SAMPLING INFERENCE – MEAN's DISTRIBU.

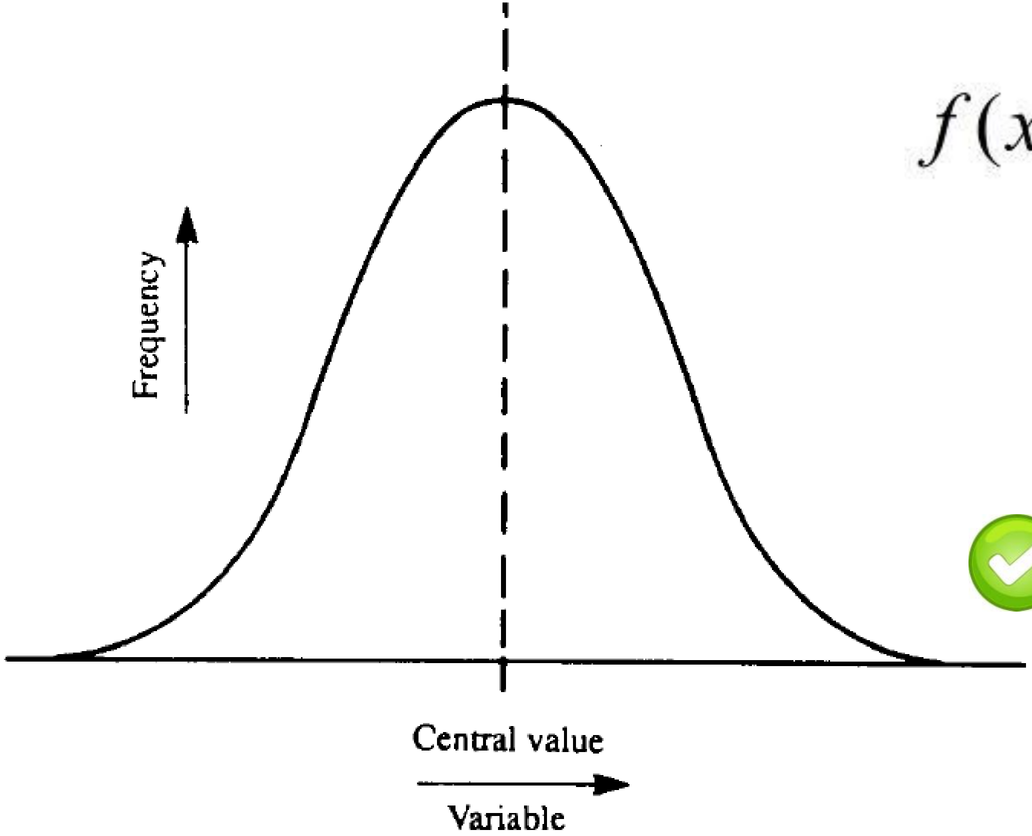


Sampling	Sampling 1	Sampling 2	Sampling 3	Sampling 4	
Sample size	n (say 100)	n	n	n	
Sample mean μ^*	μ_1	μ_2	μ_3	μ_4	

Even if population is not normal !

μ^* = <u>sample mean</u> is <u>normally</u> distributed $N(\mu^*, \sigma^*)$

NORMAL DISTRIBUTION

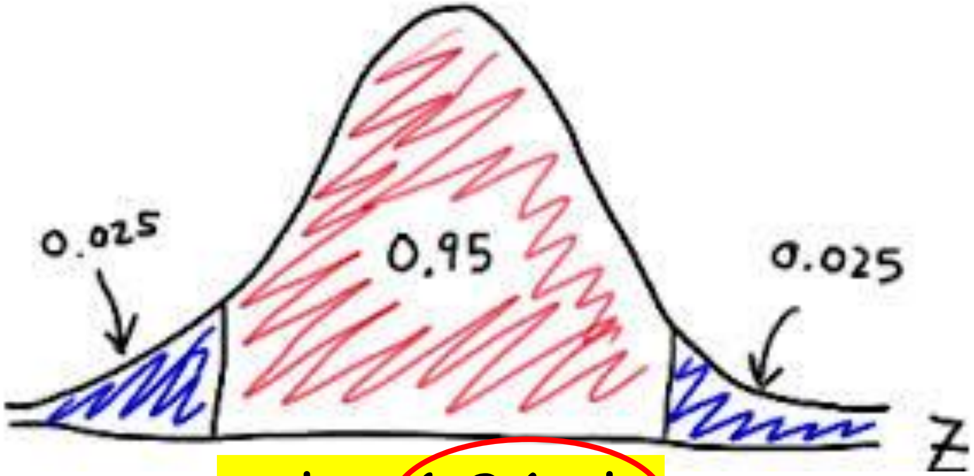


$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

It means:-

Knowing **μ and σ** of a **normally** distributed variable, one can determine how much **probable** is it to **lie between a range**.

MARGIN OF ERROR

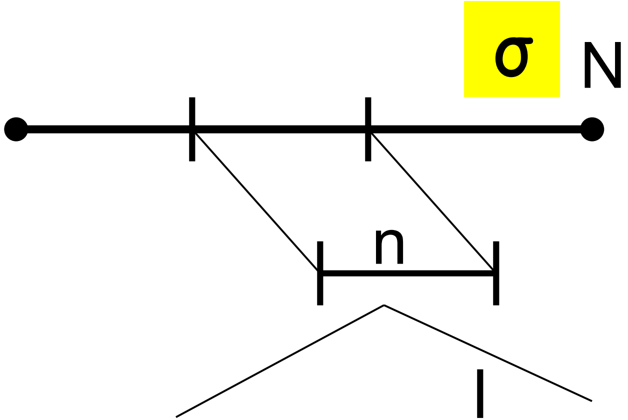


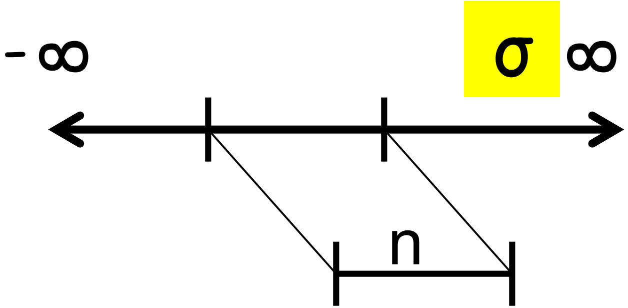
$\mu^* \pm 1.96\sigma^*$

This value is Margin of error at 95% confidence

σ^* has special calculation

MARGIN OF ERROR - CALCULATION





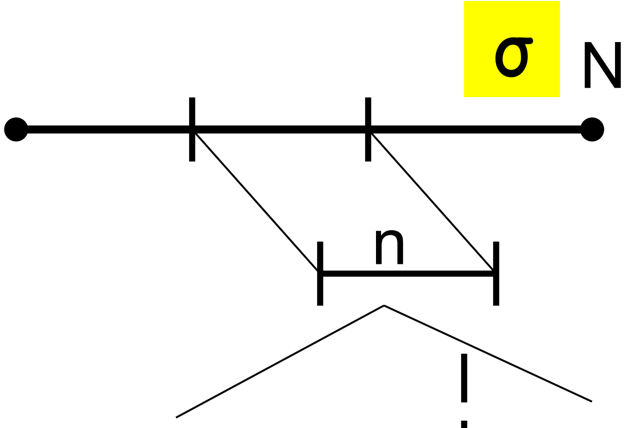
No Replacement

$$\sigma^* = \frac{\sigma \sqrt{(N-n)}}{\sqrt{n\,(N-1)}}$$

Replacement

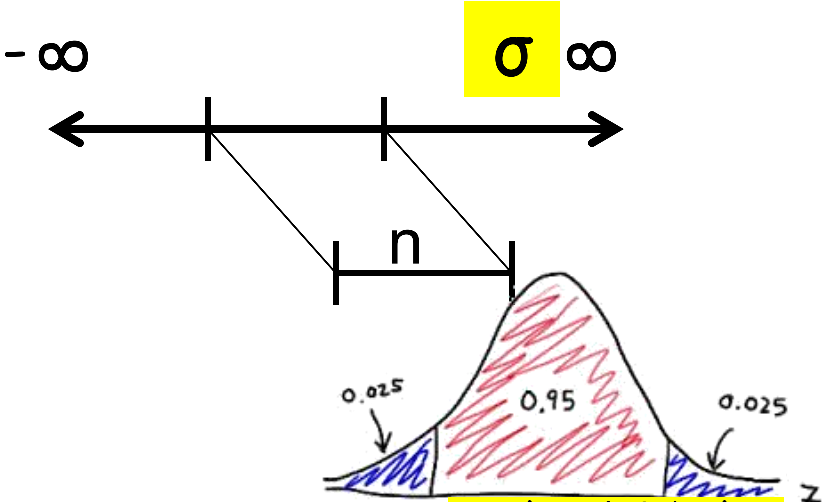
$$\sigma^* = \frac{\sigma}{\sqrt{n}}$$

MARGIN OF ERROR - CALCULATION



No Replacement

Replacement

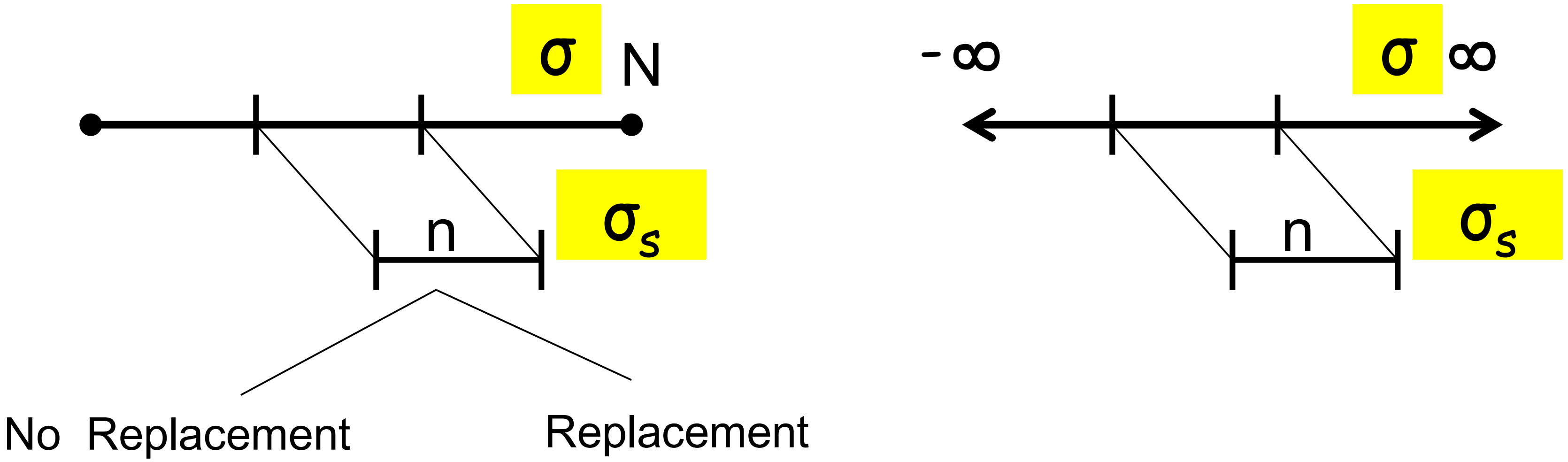


μ* ± 1.96σ*

$$E = \frac{1.96\sigma\sqrt{(N-n)}}{\sqrt{n\,(N-1)}}$$

$$E = \frac{1.96\ \sigma}{\sqrt{n}}$$

POPULATION STD DEV



Sample σ_s ≈ Population σ

If $n \geq 100$

SAMPLING INFERENCE - MEAN

I sampled 240 firms so n=240.

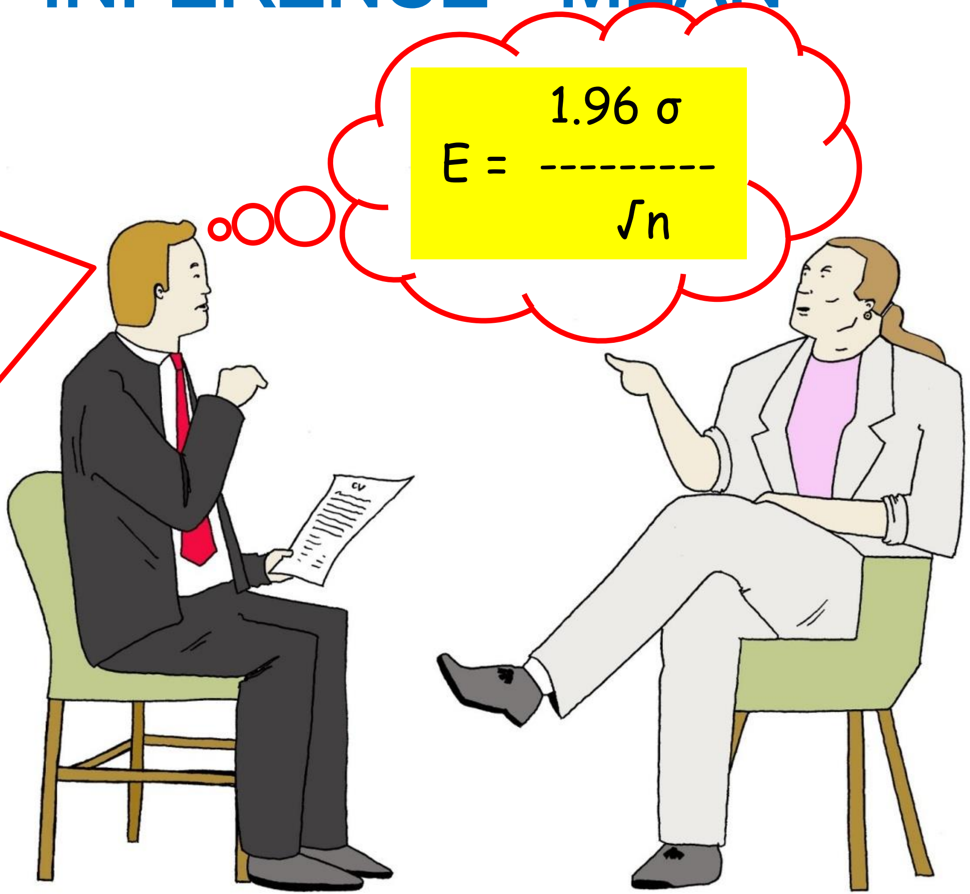
Also 240>100, so 'σ' can be taken = σ_s

$$E = \frac{1.96\,\sigma}{\sqrt{n}}$$

SAMPLING INFERENCE - MEAN

In my sample σ_s came out to be 0.8 MRs

$\sigma_s \approx \sigma \approx 0.8$



SAMPLING INFERENCE - MEAN

$$E = \frac{1.96\,\sigma}{\sqrt{n}}$$

So I can say that Margin of Error = ± E

= ± 1.96*(0.8) / (240^0.5)

≈ 0.1

SAMPLING INFERENCE - MEAN

Thus tax Avg.
=1.2±E
=1.2±0.1

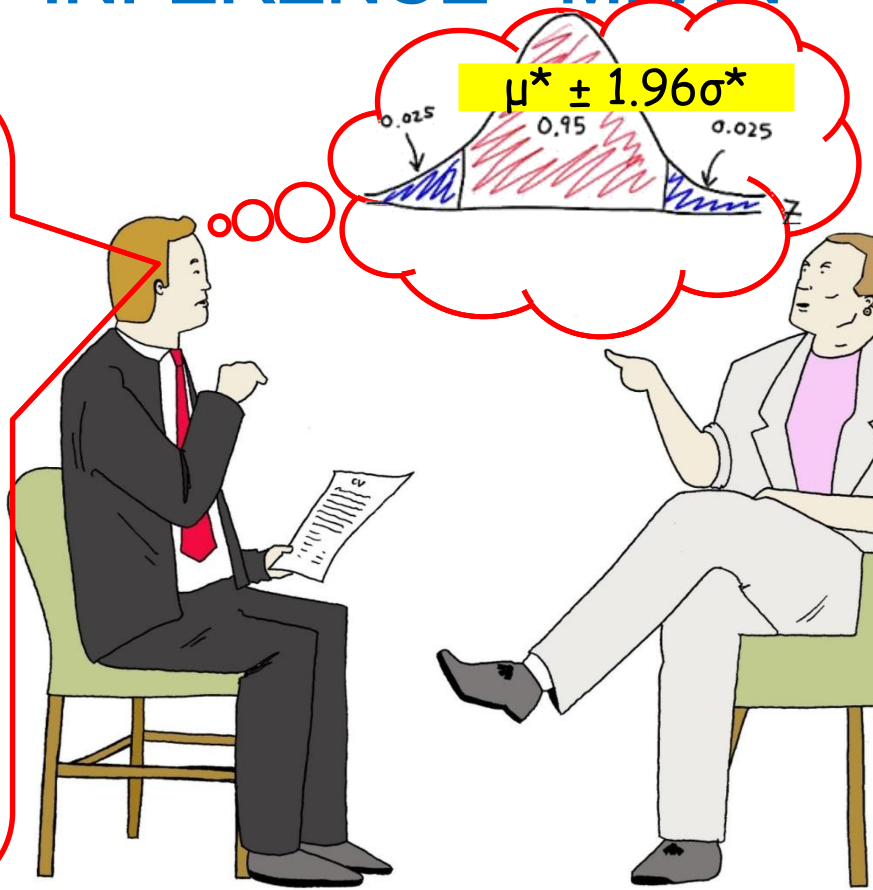
It lies b/w
1.1 to 1.3
MRs

$\mu^* \pm 1.96\sigma^*$

Always
?

SAMPLING INFERENCE - MEAN

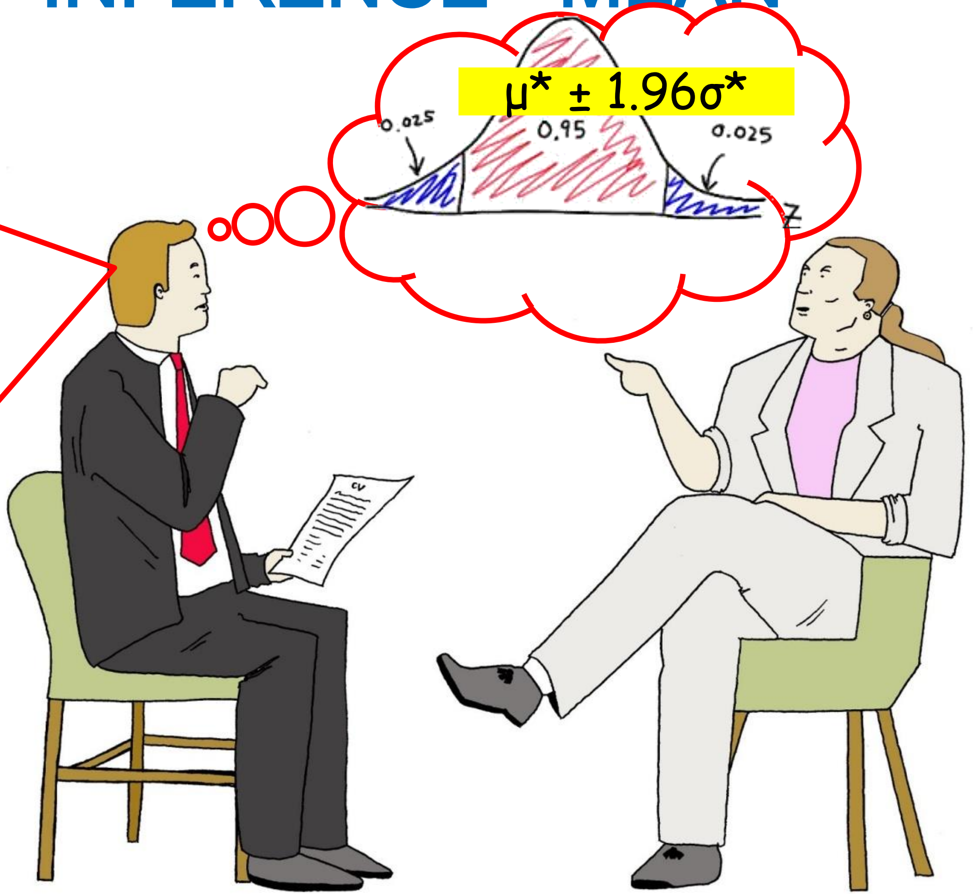
No,
Not always,
but (almost)
95% times
you sample
n=240, it
would lie b/w
1.1 to 1.3
MRs



So, tax
would lie
b/w 1.1
~ 1.3
95% of
times..

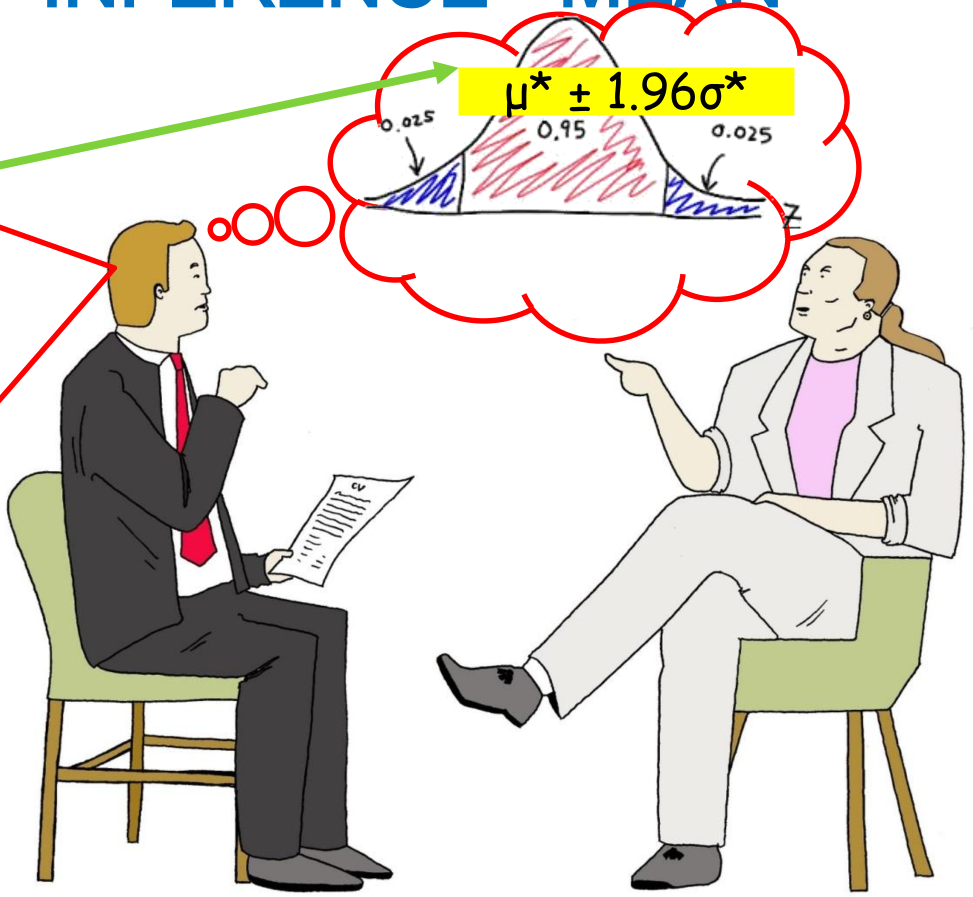
SAMPLING INFERENCE - MEAN

Strictly speaking - No, Not the tax.
We are talking about a particular sample statistic here…



SAMPLING INFERENCE - MEAN

We are discussing the statistic "AVG."
So there's 95% chance that sample AVG. is 1.1~1.3



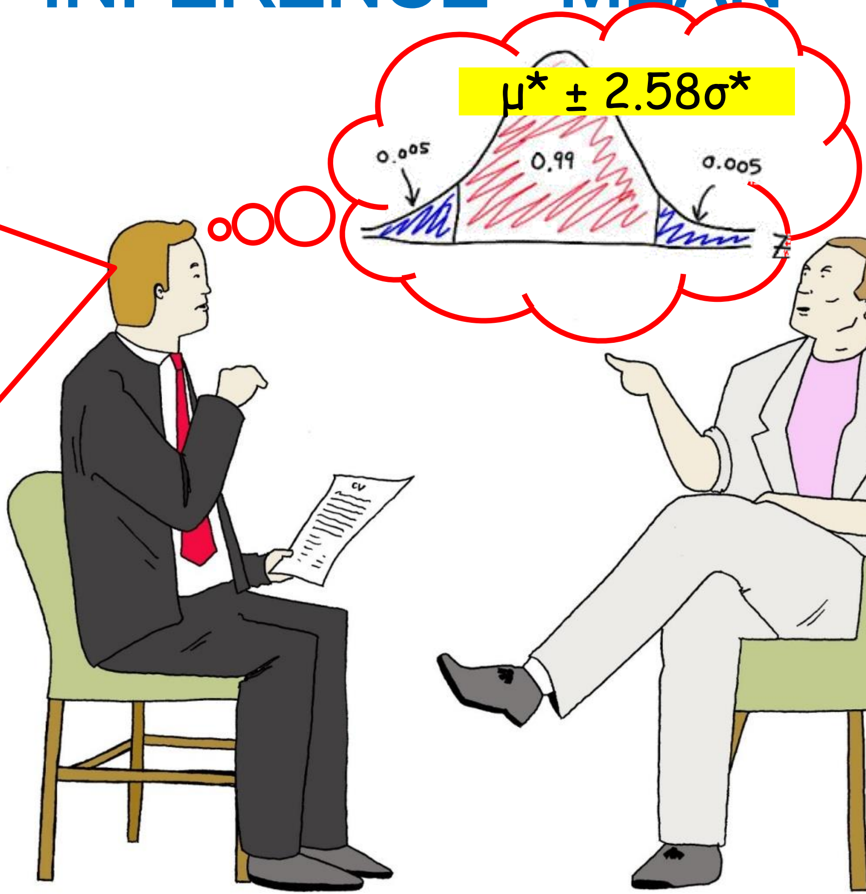
SAMPLING INFERENCE - MEAN

If you ask me 99% confidence level, my range would be

E= ± 2.58*(0.8) / (240^0.5)

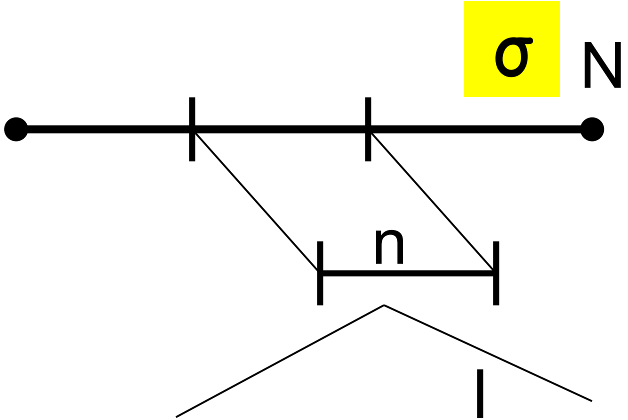
Avg. =1.2 ± 0.133

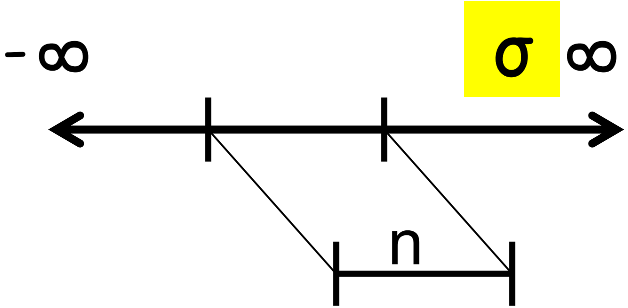
b/w 1.067 to 1.333



How to further reduce the margin of Error E ?

SAMPLE SIZE - CALCULATION





Calculate n using the desired E

No Replacement

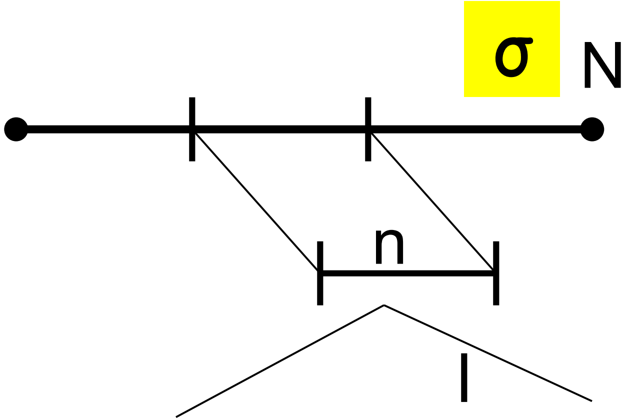
$$E = \frac{1.96\sigma\sqrt{(N-n)}}{\sqrt{n\,(N-1)}}$$

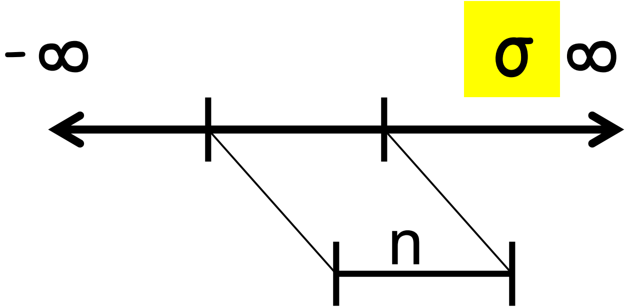
Replacement

$$E = \frac{1.96\,\sigma}{\sqrt{n}}$$

$$n = \left[\frac{1.96\,\sigma}{E}\right]^2$$

SAMPLE SIZE - CALCULATION





No Replacement

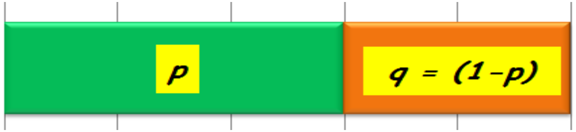
$$E = \frac{1.96\sigma \sqrt{(N-n)}}{\sqrt{n\,(N-1)}}$$

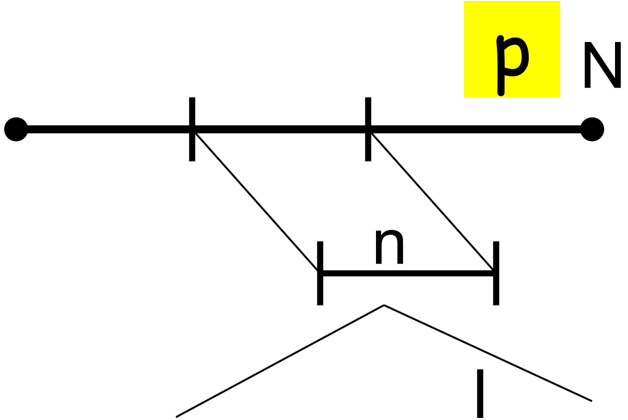
Replacement

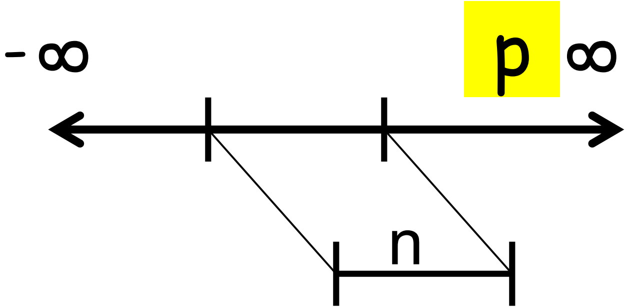
N large

$$n = \left[\frac{1.96\,\sigma}{E}\right]^2$$

MARGIN OF ERROR - BINOMIAL





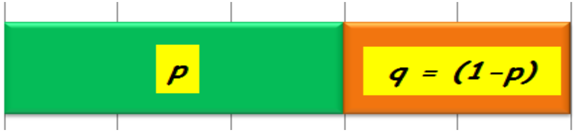


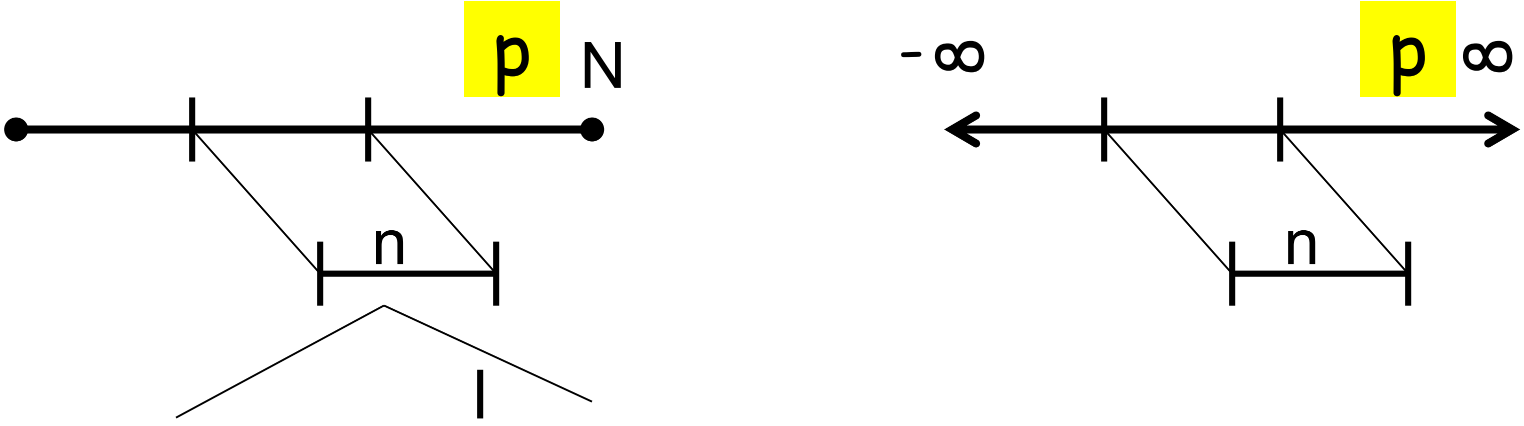
No Replacement | Replacement

$$\sigma^* = \frac{\sqrt{pq\ (N-n)}}{\sqrt{n\ (N-1)}}$$

$$\sigma^* = \frac{\sqrt{pq}}{\sqrt{n}}$$

MARGIN OF ERROR - BINOMIAL



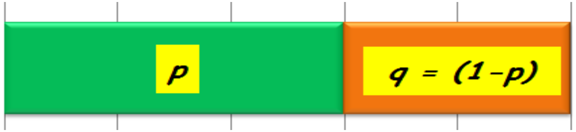


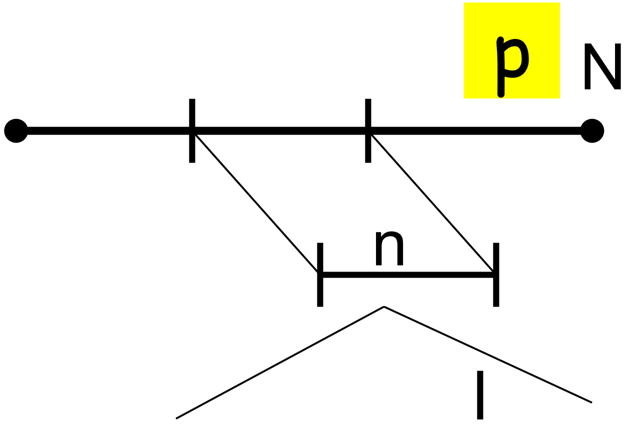
No Replacement | Replacement

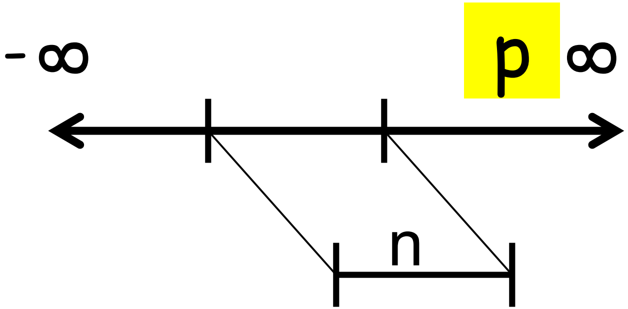
$$E = \frac{1.96\sqrt{pq\ (N-n)}}{\sqrt{n\ (N-1)}}$$

$$E = \frac{1.96\sqrt{pq}}{\sqrt{n}}$$

SAMPLE SIZE - BINOMIAL







Calculate n using the desired E

No Replacement

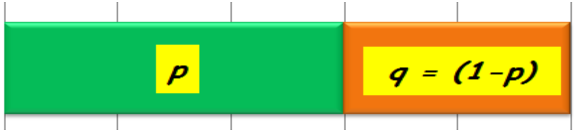
$$E = \frac{1.96\sqrt{pq\,(N-n)}}{\sqrt{n\,(N-1)}}$$

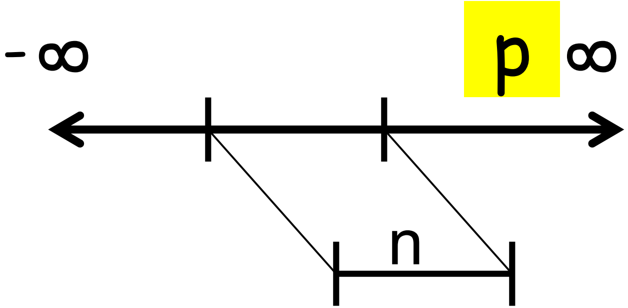
Replacement

$$E = \frac{1.96\sqrt{pq}}{\sqrt{n}}$$

$$n = \left[\frac{1.96\sqrt{pq}}{E}\right]^2$$

SAMPLE SIZE - BINOMIAL



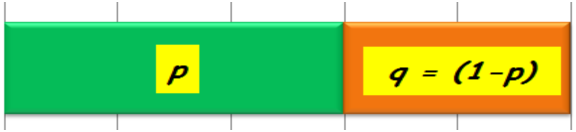


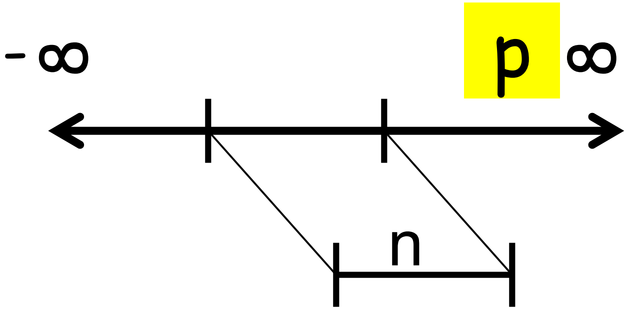
p & q are expected ideal population proportions here

When not known beforehand, assumed p = q = 0.5

$$n = \left[\frac{1.96\sqrt{pq}}{E}\right]^2$$

SAMPLE SIZE - BINOMIAL



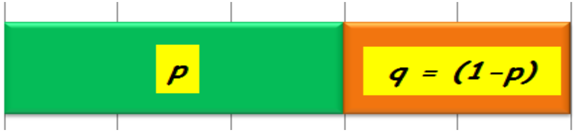


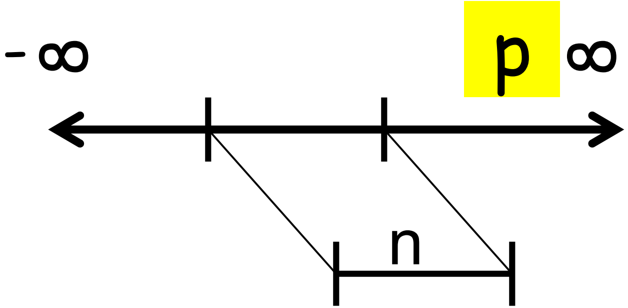
Note that in case of proportions, the E is also in terms of proportions, e.g. 5%

So, in the formula it is entered in %age e.g. 0.05 for ±5% error

$$n = \left[\frac{1.96\sqrt{pq}}{E}\right]^2$$

SAMPLE SIZE - BINOMIAL

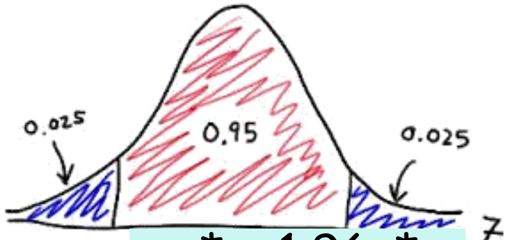




So, example calculation for E=5% and <u>unknown</u> p & q, & 95% confidence

$$n = \left[\frac{1.96\sqrt{0.5*0.5}}{E}\right]^2 \Rightarrow n = \left[\frac{1.96\sqrt{0.5*0.5}}{0.05}\right]^2 \Rightarrow n = 384$$

CAUTION ! APPLICABILITY



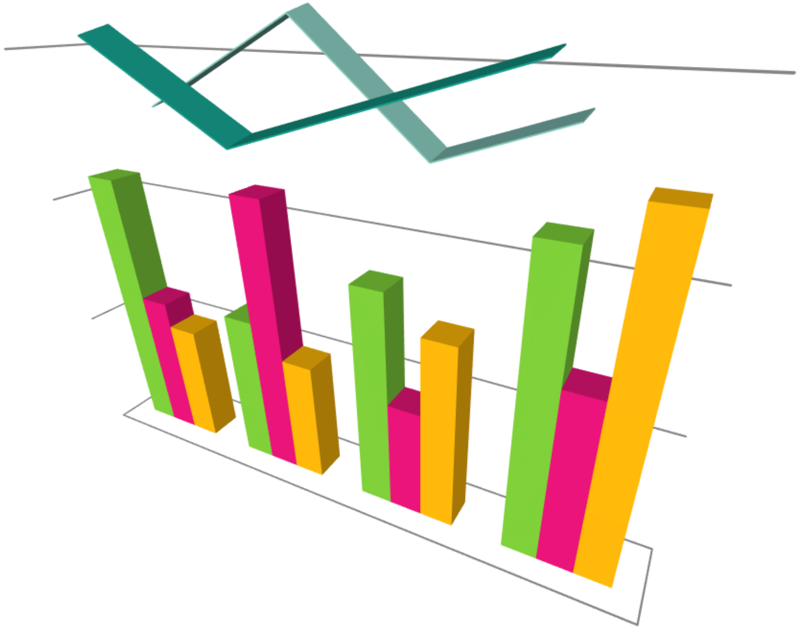
$\mu^* \pm 1.96\sigma^*$

The Formulas of Margin of Error E studied are valid only for the sampling statistic = AVERAGE μ^*
And not for other sampling statistics like STD DEV. etc.

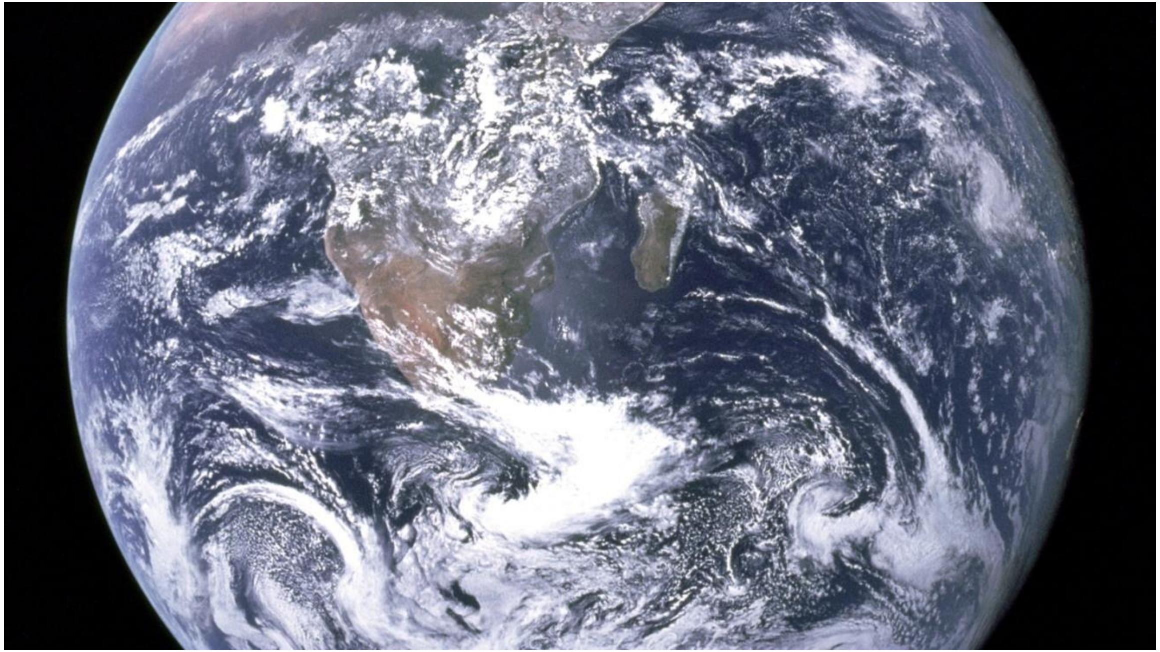
Sampling	**Sampling 1**	**Sampling 2**	**Sampling 3**	**Sampling 4**	**…**
Sample size	n (say 100)	n	n	n	…
Sample mean μ^*	μ_1	μ_2	μ_3	μ_4	…

μ^* = sample mean is normally distributed $N(\mu^*, \sigma^*)$

REPORTING STATISTICS



PARAMETER MENTIONED IN SAMPLING REPORT	CHECK
Basic Assumptions about Population	✓
Sampling Technique used	✓
Sample Size	✓
Sampling Inference (outcome, eg. μ)	✓
Margin of Error (± E)	✓
Confidence Level (eg. 95%)	✓



Thanks