Hypothesis Testing

ANISHA JOSEPH

Testing of hypothesis ✓ → Assumption/speculation

→ The process which anable us to decide whether to reject or accept a statement on the basis of sample collected from population.

→ Whether arrested person is annocent or guilty.

→ Casual system is effective then online class system.

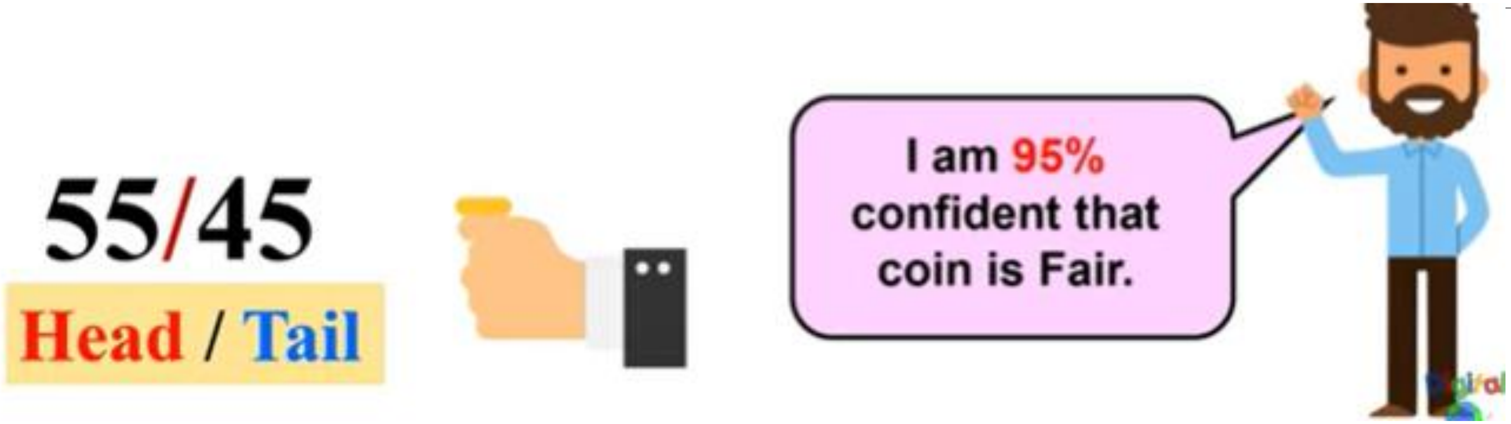
→ Fresh juices are better than soft drinks.

Hypothesis

-A CLAIM OR STATEMENT ABOUT A POPULATION *PARAMETER*

Guide to Hypothesis Testing

- Hypothesis testing is an educated guess to detect significant differences by comparing sample statistics with population parameters.
- Hypothesis testing is used to infer the result of a hypothesis performed on sample data from a larger population.


55/45
Head / Tail
I am 95% confident that coin is Fair.

Null Hypothesis vs Alternate Hypothesis

Null Hypothesis

- Denoted by (Ho).
- States there is no significant difference.
- Refer to as Status Quo.
- Always contains the ' = ' sign $H_0 : \mu_1 = \mu_2$

50/50

Head / Tail

Alternate Hypothesis

- Denoted by (Ha) or (H1).
- (H1) always contradicts the H0
- Challenges the Status Quo.
- Never contains the ' = ' sign $H1: \mu 1 \neq \mu 2$

Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu = 23$	$H_0: \mu \geq 23$	$H_0: \mu \leq 23$
$\boxed{H_1: \mu \neq 23}$	$\boxed{H_1: \mu < 23}$	$\boxed{H_1: \mu > 23}$

Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu = 23$	$H_0: \mu \geq 23$	$H_0: \mu \leq 23$
$H_1: \mu \neq 23$	$H_1: \mu < 23$	$H_1: \mu > 23$

Reject H_0 $\longrightarrow$ Support H_1

Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu = 23$	$H_0: \mu \geq 23$	$H_0: \mu \leq 23$
$H_1: \mu \neq 23$	$H_1: \mu < 23$	$H_1: \mu > 23$

Fail to Reject H_0 $\Rightarrow$ Cannot Support H_1

Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
H_0: $\mu = 23$	H_0: $\mu \geq 23$	H_0: $\mu \leq 23$
H_1: $\mu \neq 23$	H_1: $\mu < 23$	H_1: $\mu > 23$

SIGNIFICANCE LEVEL α

0.05 MOST COMMON

0.10 & 0.01 ALSO USED

Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
H_0: $\mu = 23$	H_0: $\mu \geq 23$	H_0: $\mu \leq 23$
H_1: $\mu \neq 23$	H_1: $\mu < 23$	H_1: $\mu > 23$

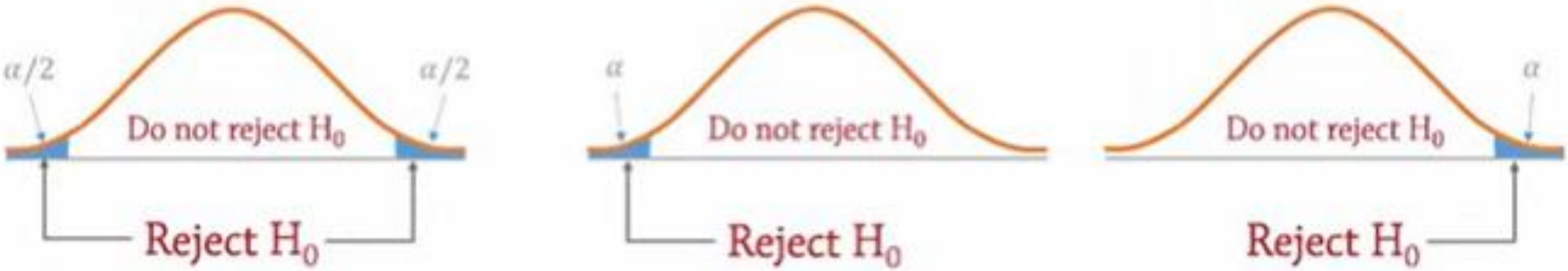
SIGNIFICANCE LEVEL α

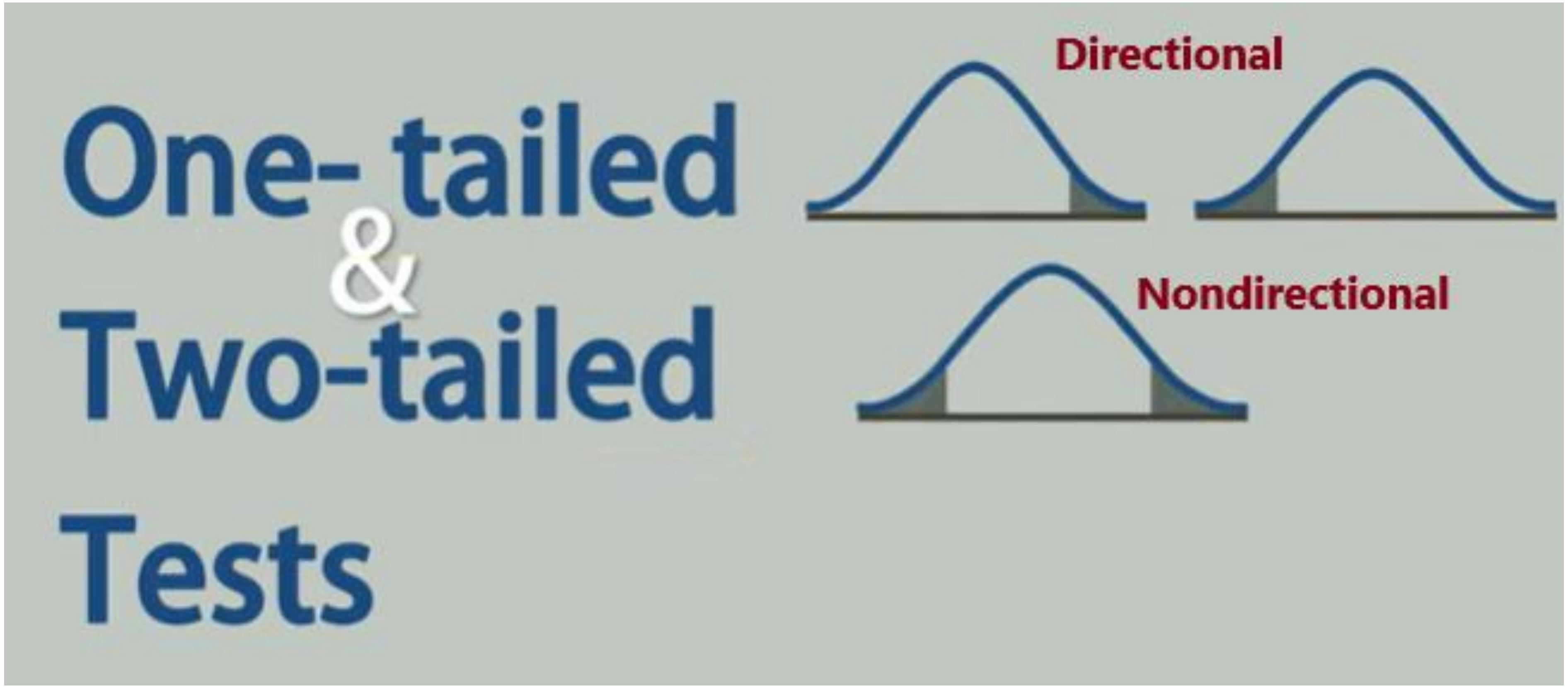
specifies the size of the rejection region

Hypothesis Testing

One-tailed

Two-tailed	Left-tailed	Right-tailed
$H_0: \mu = 23$	$H_0: \mu \geq 23$	$H_0: \mu \leq 23$
$H_1: \mu \neq 23$	$H_1: \mu < 23$	$H_1: \mu > 23$




One- tailed
&
Two-tailed
Tests
Directional
Nondirectional

Right-tailed Test

upper -tailed

$\mu > 20$

population parameter

hypothesized value

$\mu > 20$

increased

more

higher

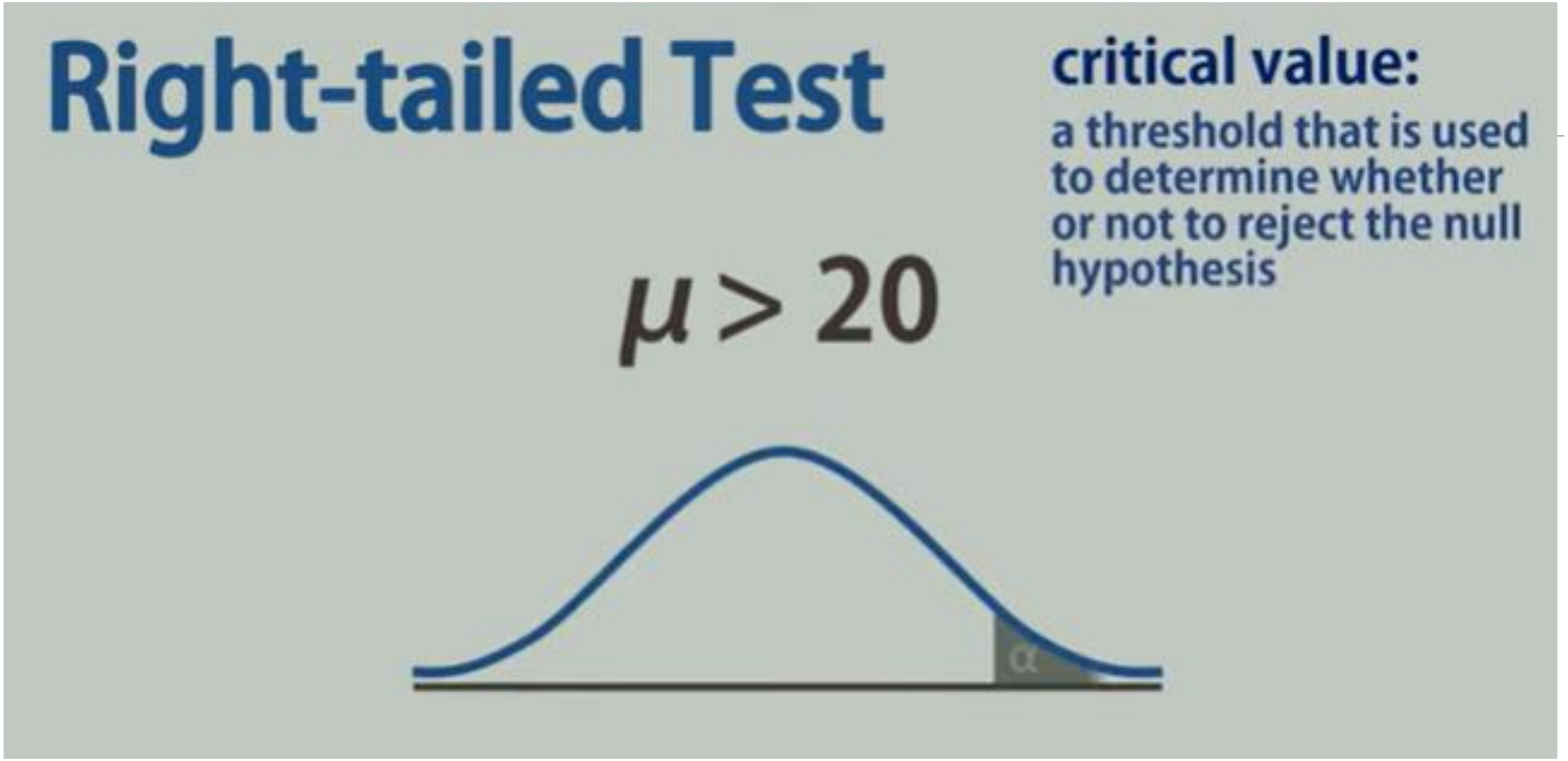
larger

over

exceed

above

beyond


Right-tailed Test
critical value:
a threshold that is used to determine whether or not to reject the null hypothesis
$\mu > 20$
α

Left-tailed Test

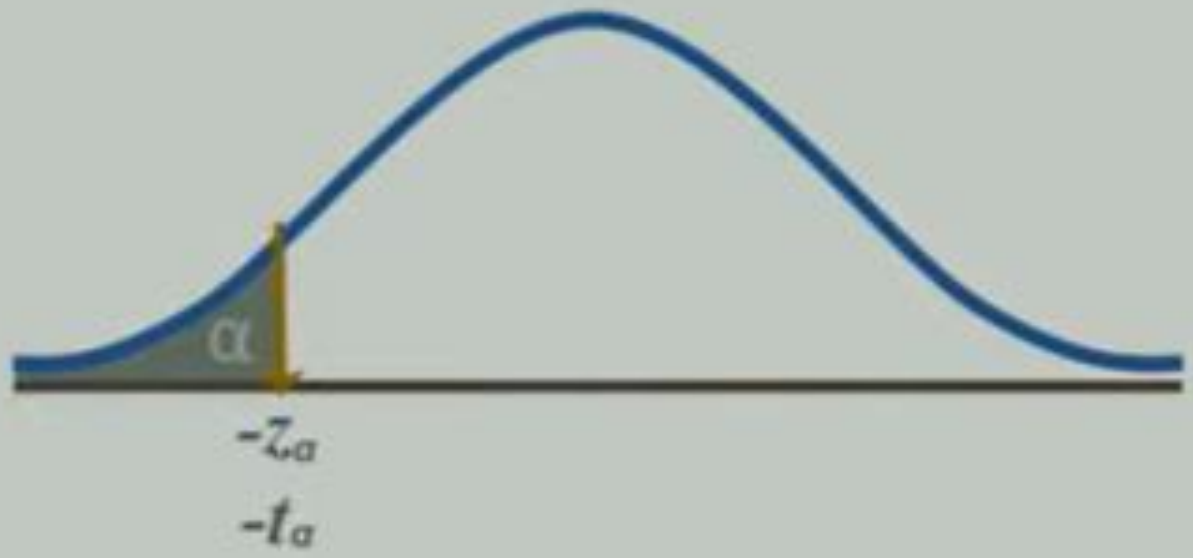
$$H_1: \mu < 40$$

lower smaller decreased

under reduced below

Left-tailed Test

$$H_1: \mu < 40$$



other ambiguous keywords

stronger

weaker

improved

worsened

outperform

underperform

Two-tailed Test

$$H_1: \mu \neq 50$$

$$\Rightarrow \quad \mu > 50 \text{ or } \mu < 50$$

Two-tailed Test

$$H_1: \mu \neq 50$$

different from

changed

deviate from

not the same

inconsistent with

This year

Example: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of $\sigma = 2.4$. Can we infer at $\alpha = 0.05$ that the population mean age has changed?

hypothesis testing: *z*-test

Example: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at α = 0.05 that the population mean age has changed? $n = 42$ $\bar{x} = 23.8$ $\sigma = 2.4$ $\alpha = 0.05$

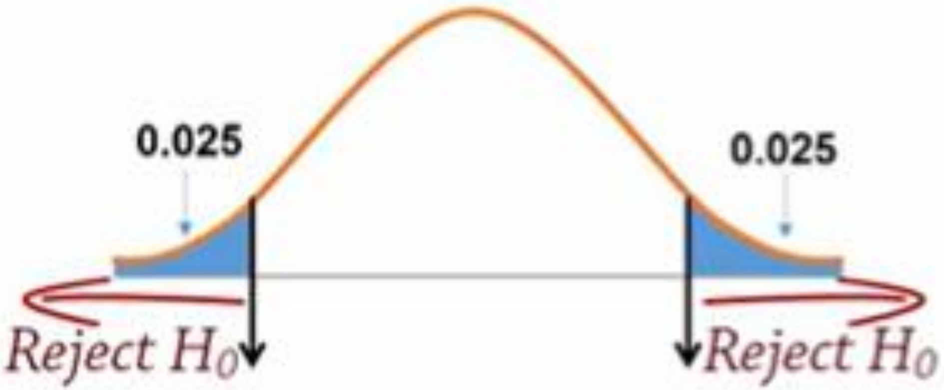
$$H_0\colon\ \mu = 23$$
$$H_1\colon\ \mu \neq 23$$

$$\alpha = 0.05$$

Example: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at α = 0.05 that the population mean age has changed? $n = 42$ $\bar{x} = 23.8$ $\sigma = 2.4$ $\alpha = 0.05$

$$H_0\colon\ \mu = 23$$
$$H_1\colon\ \mu \neq 23$$

$$\alpha = 0.05$$



Test of μ

z-test

or

t-test

Example: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of $\sigma = 2.4$. Can we infer at $\alpha = 0.05$ that the population mean age has changed? $n = 42$ $\bar{x} = 23.8$ $\sigma = 2.4$ $\alpha = 0.05$

$$H_0:\ \mu = 23$$
$$H_1:\ \mu \neq 23$$

$$\alpha = 0.05$$

σ *known* $\quad$ $\boldsymbol{z} - test$

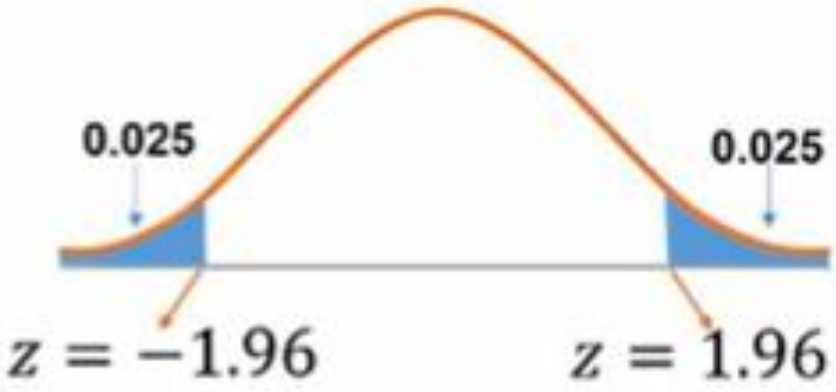
σ *unknown* $\quad$ $\boldsymbol{t} - test$

Central Limit Theorem

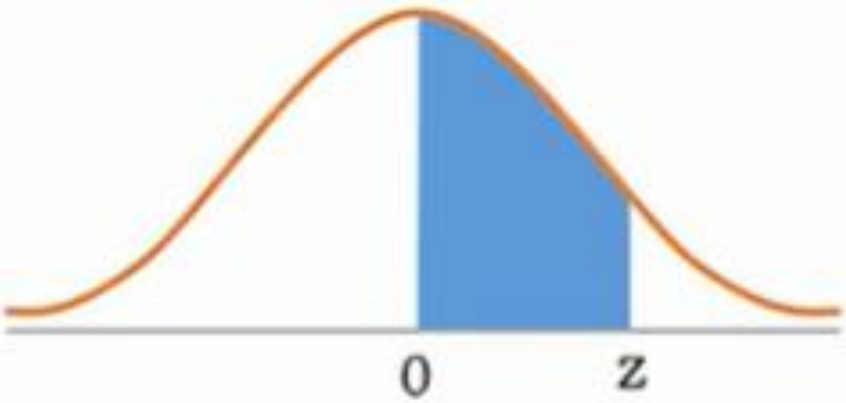
large samples $\quad$ $\boldsymbol{z} - test$

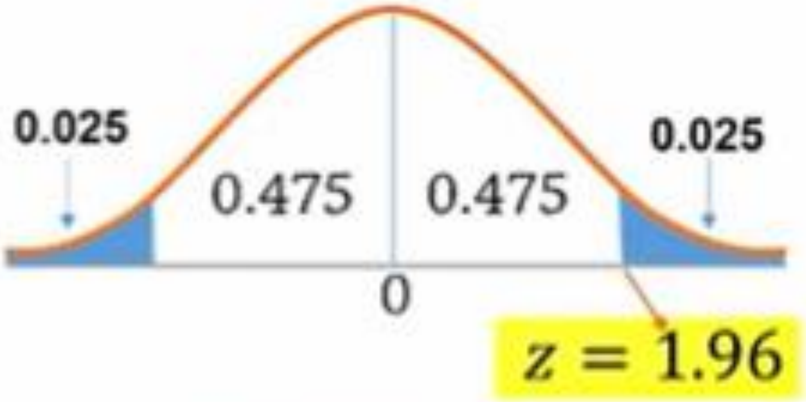
$n \geq 30$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559





z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4951	0.4952

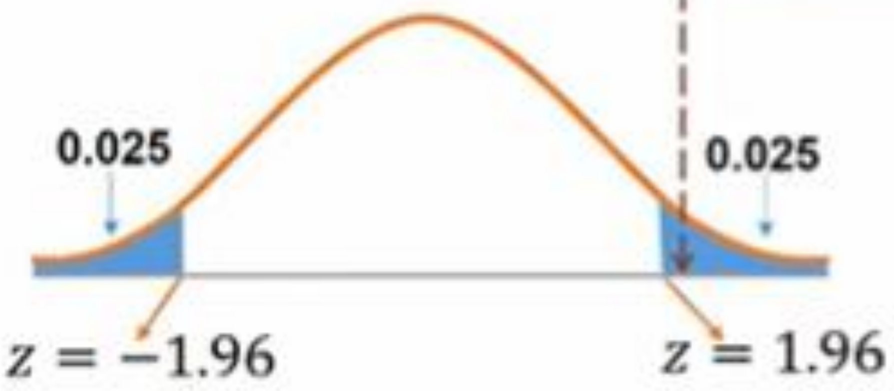
1 tail, α	0.10	0.05	0.025	0.01	0.005
2 tails, α	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
df 1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
Z ∞	**1.282**	**1.645**	**1.960**	**2.326**	**2.576**

Example: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at α = 0.05 that the population mean age has changed?

$n = 42 \quad \bar{x} = 23.8 \quad \sigma = 2.4 \quad \alpha = 0.05$

$$H_0: \mu = 23$$
$$H_1: \mu \neq 23$$

$$\alpha = 0.05$$



Reject H_0 if $z < -1.96$ or $z > 1.96$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{23.8 - 23}{2.4/\sqrt{42}} = 2.16$$

Since $z = 2.16 > 1.96$, reject H_0.

There is enough evidence that the mean age has changed.

at α = 0.05

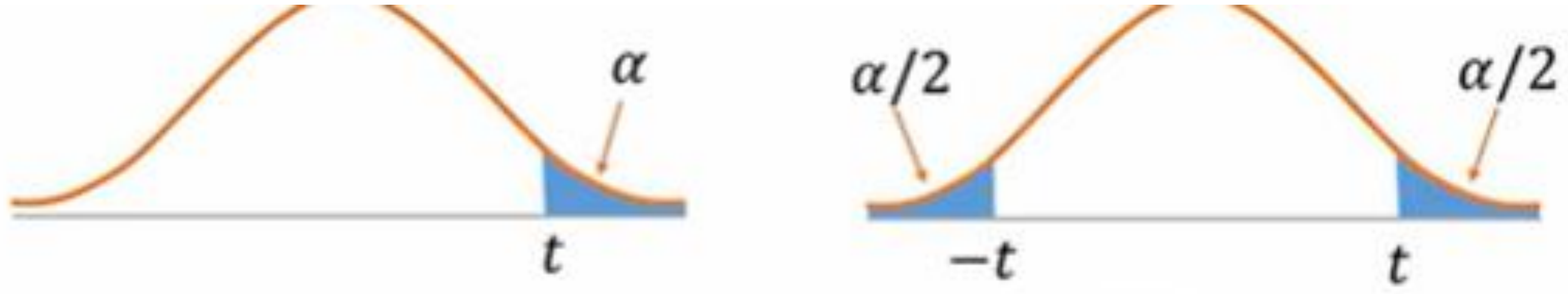
Example 2: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at $\boxed{\alpha = 0.02}$ that the population mean age has changed?

$n = 42 \quad \bar{x} = 23.8 \quad \sigma = 2.4 \quad \alpha = 0.02$

$$H_0: \quad \mu = 23$$
$$H_1: \quad \mu \neq 23$$

Example 2: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at $\boxed{\alpha = 0.02}$ that the population mean age has changed? $n = 42$ $\bar{x} = 23.8$ $\sigma = 2.4$ $\alpha = 0.02$

$$H_0: \quad \mu = 23$$
$$H_1: \quad \mu \neq 23$$



1 tail, α	0.10	0.05	0.025	0.01	0.005
2 tails, α	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
df 1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
Z **∞**	**1.282**	**1.645**	**1.960**	**2.326**	**2.576**

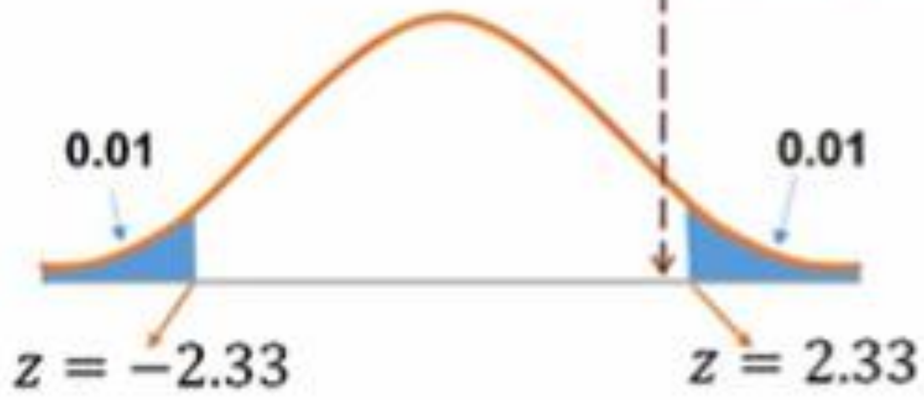
Example 2: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at α = 0.02 that the population mean age has changed? $n = 42 \quad \bar{x} = 23.8 \quad \sigma = 2.4 \quad \alpha = 0.02$

$$H_0: \quad \mu = 23$$
$$H_1: \quad \mu \neq 23$$

$$\alpha = 0.02$$

Example 2: In recent years, the mean age of all college students in city X has been 23. A random sample of 42 students revealed a mean age of 23.8. Suppose their ages are normally distributed with a population standard deviation of σ = 2.4. Can we infer at α = 0.02 that the population mean age has changed? $n = 42 \quad \bar{x} = 23.8 \quad \sigma = 2.4 \quad \alpha = 0.02$

$$H_0: \quad \mu = 23$$
$$H_1: \quad \mu \neq 23$$
$$\alpha = 0.02$$



Reject H_0 if $z < -2.33$ or $z > 2.33$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{23.8 - 23}{2.4/\sqrt{42}} = 2.16$$

Since $z = 2.16 < 2.33$, fail to reject H_0.

Cannot support H_1

At $\alpha = 0.02$

There is not enough evidence that the mean age has changed.

A Telecom service provider claims that individual customers pay on an average 400 rs. per month with standard deviation of 25 rs. A random sample of 50 customers bills during a given month is taken with a mean of 250 and standard deviation of 15. What to say with respect to the claim made by the service provider?

Solution:

First thing first, Note down what is given in the question:

H_0 (Null Hypothesis) : $\mu = 400$

H_1 (Alternate Hypothesis): $\mu \neq 400$ (Not equal means either $\mu > 400$ or $\mu < 400$

Hence it will be validated with **two tailed test**)

$\sigma = 25$ (Population Standard Deviation)

LoS (α) = 5% (Take 5% if not given in question)

$n = 50$ (Sample size)

xbar $\bar{x} = 250$ (Sample mean)

$s = 15$ (sample Standard deviation)

n > = 30 hence will go with z-test

Step 1:

Calculate z using z-test formula as below:

$z = (\bar{x} - \mu) / (\sigma/\sqrt{n})$

$z = (250 - 400) / (25/\sqrt{50})$

$z = -42.42$

Step 2:

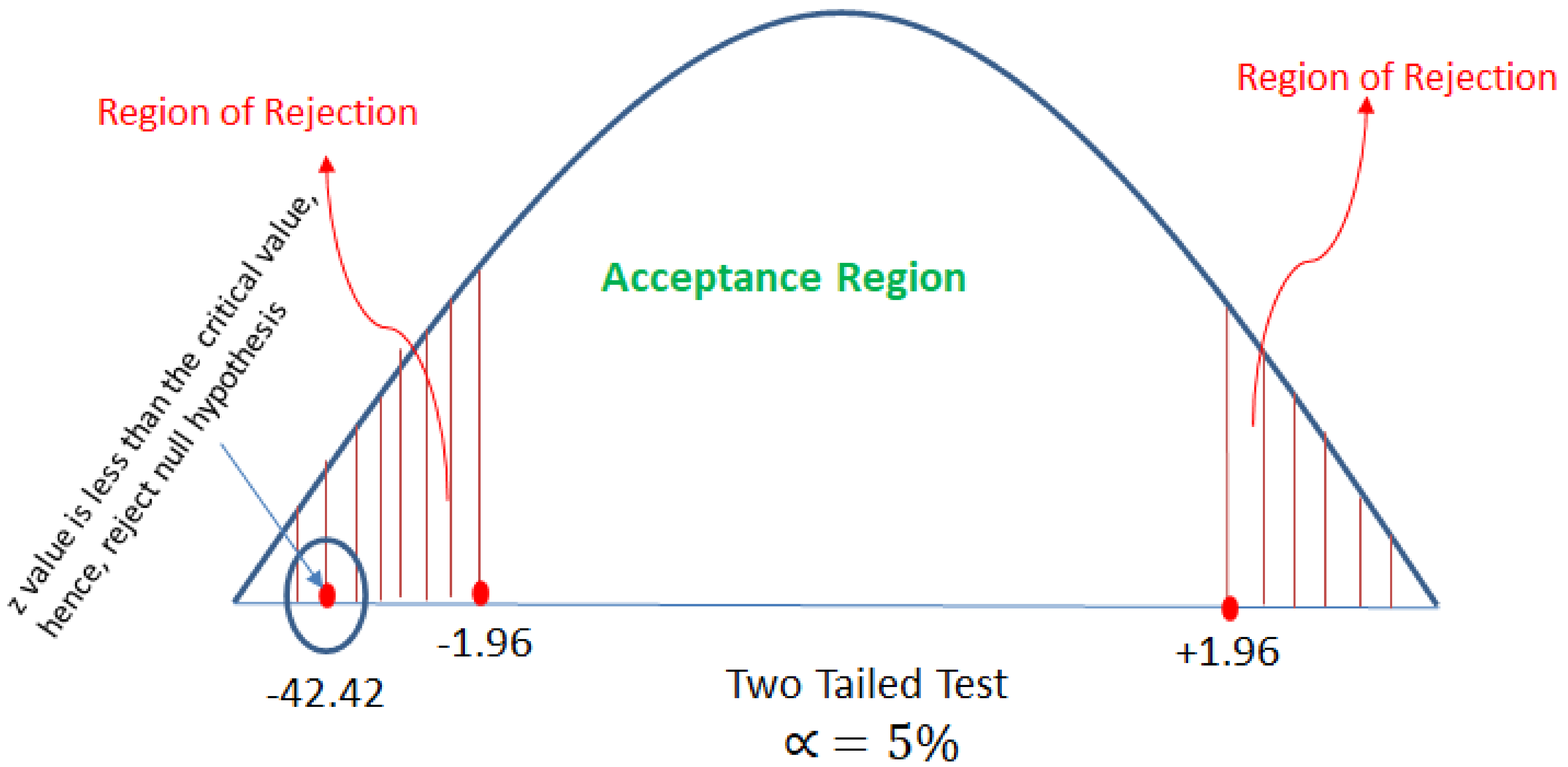
get z critical value from z table for $\alpha = 5\%$

z critical values = $(-1.96, +1.96)$

to accept the claim (significantly), calculated z should be in between

$-1.96 < z < +1.96$

but calculated z (-42.42) < -1.96 which mean reject the null hypothesis



z-test example 1

Hypothesis Testing

A fitness app claims that the average user exercises for 30 minutes or less per day. A random sample of 50 users shows a mean daily exercise time of 32 minutes. The standard deviation of daily exercise time for all users of the app is 6 minutes. At the 3% significance level, can we conclude that the app's claim is incorrect?

Hypothesis Testing

A fitness app claims that the average user exercises for 30 minutes or less per day. A random sample of 50 users shows a mean daily exercise time of 32 minutes. The standard deviation of daily exercise time for all users of the app is 6 minutes. At the 3% significance level, can we conclude that the app's claim is incorrect?

$n = 50$ $\bar{x} = 32$ $\sigma = 6$ $\alpha = 0.03$

more than 30 minutes

A fitness app claims that the average user exercises for 30 minutes or less per day.
A random sample of 50 users shows a mean daily exercise time of 32 minutes.
The standard deviation of daily exercise time for all users of the app is 6 minutes.
At the 3% significance level, can we conclude that the app's claim is incorrect?

$n = 50$	$\bar{x} = 32$	$\sigma = 6$	$\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

One-tailed

right-tailed

σ known → One sample Z-test

$n = 50$ | $\bar{x} = 32$ | $\sigma = 6$ | $\alpha = 0.03$

$$H_0: \mu \le 30$$
$$H_1: \mu > 30$$

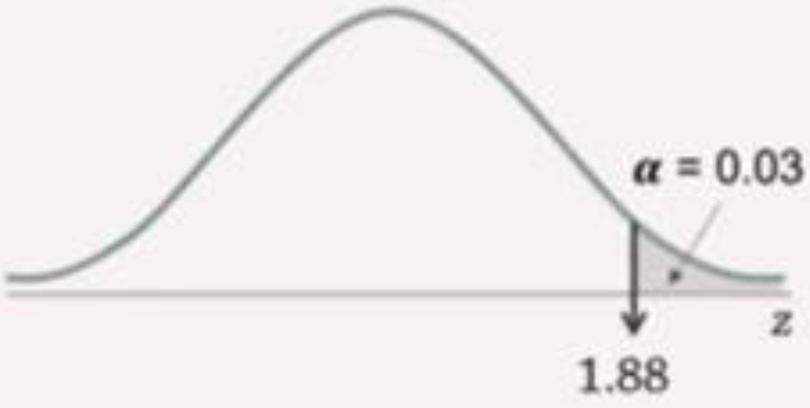
$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{6}{\sqrt{50}}} = \frac{2}{0.8485} \approx 2.36$$

$n = 50$ $\bar{x} = 32$ $\sigma = 6$ $\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{6}{\sqrt{50}}} = \frac{2}{0.8485} \approx 2.36$$



Z	0.00	0.01	...	0.07	0.08	0.09
1.5	0.9332	0.9345	...	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	...	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	...	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	...	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	...	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	...	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	...	0.9850	0.9854	0.9857
...	...	...	...	...	...	...

$n = 50$ $\bar{x} = 32$ $\sigma = 6$ $\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{6}{\sqrt{50}}} = \frac{2}{0.8485} \approx 2.36$$

Since z = 2.36 > 1.88, reject H_0 at α = 0.03

significant result

Reject H_0 if z(obtained) > 1.88

$n = 50$	$\bar{x} = 32$	$\sigma = 6$	$\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{6}{\sqrt{50}}} = \frac{2}{0.8485} \approx 2.36$$

The *P*-value represents the probability of obtaining an observed effect or more extreme results assuming the null hypothesis is true.

$n = 50$ $\bar{x} = 32$ $\sigma = 6$ $\alpha = 0.03$

$$H_0: \mu \leq 30$$
$$H_1: \mu > 30$$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{6}{\sqrt{50}}} = \frac{2}{0.8485} \approx 2.36$$

Z	0.00	0.01	...	0.05	0.06	0.07
2.0	0.9772	0.9778	...	0.9798	0.9803	0.9808
2.1	0.9821	0.9826	...	0.9842	0.9846	0.9850
2.2	0.9861	0.9864	...	0.9878	0.9881	0.9884
2.3	0.9893	0.9896	...	0.9906	0.9909	0.9911
2.4	0.9918	0.9920	...	0.9929	0.9931	0.9932
2.5	0.9938	0.9940	...	0.9946	0.9948	0.9949
2.6	0.9953	0.9955	...	0.9960	0.9961	0.9962
...	...	...	...	...	...	...

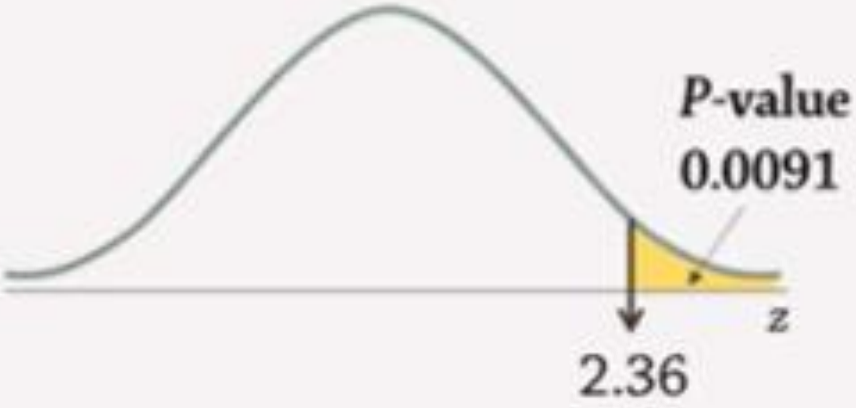
A fitness app claims that the average user exercises for 30 minutes or less per day. A random sample of 50 users shows a mean daily exercise time of 32 minutes. The standard deviation of daily exercise time for all users of the app is 6 minutes. At the 3% significance level, can we conclude that the app's claim is incorrect?

$n = 50$ $\quad \bar{x} = 32$ $\quad \sigma = 6$ $\quad \alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \approx 2.36$$



There is only a 0.91% chance of observing a sample mean of 32 minutes or more if the null hypothesis were true

if *P*-value $\leq \alpha$, reject H_0

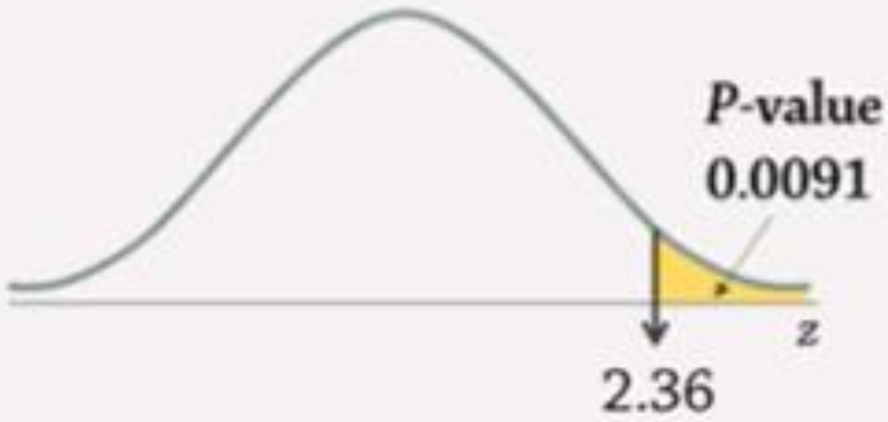
Since *P*-value $= 0.0091 \leq 0.03$, reject H_0

A fitness app claims that the average user exercises for 30 minutes or less per day. A random sample of 50 users shows a mean daily exercise time of 32 minutes. The standard deviation of daily exercise time for all users of the app is 6 minutes. At the 3% significance level, can we conclude that the app's claim is incorrect?

$n = 50$	$\bar{x} = 32$	$\sigma = 6$	$\alpha = 0.03$

$$H_0: \mu \leq 30$$
$$H_1: \mu > 30$$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \approx 2.36$$



Since P-value = $0.0091 \leq 0.03$, reject H_0

There is enough evidence to conclude that the average user of the app exercises for longer than 30 minutes per day

Hypothesis Testing

$n = 50$ $\bar{x} = 32$ $\sigma = 8$ $\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Increasing Variability

Hypothesis Testing

$n = 50$ | $\bar{x} = 32$ | $\sigma = 8$ | $\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{8}{\sqrt{50}}} = \frac{2}{1.1314} \approx 1.77$$

Reject H_0 if $z > 1.88$

Since $z = 1.77 \not> 1.88$ fail to reject H_0

not significant

Hypothesis Testing

$n = 50$ $\bar{x} = 32$ $\sigma = 8$ $\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{8}{\sqrt{50}}} = \frac{2}{1.1314} \approx 1.77$$

z	0.00	0.01	...	0.06	0.07	0.08
1.4	0.9192	0.9207	...	0.9279	0.9292	0.9306
1.5	0.9332	0.9345	...	0.9406	0.9418	0.9429
1.6	0.9452	0.9463	...	0.9515	0.9525	0.9535
1.7	0.9554	0.9564	...	0.9608	0.9616	0.9625
1.8	0.9641	0.9649	...	0.9686	0.9693	0.9699
1.9	0.9713	0.9719	...	0.9750	0.9756	0.9761
2.0	0.9772	0.9778	...	0.9803	0.9808	0.9812
...	...	...	...	...	...	...

fail to reject H_0 at $\alpha = 0.03$

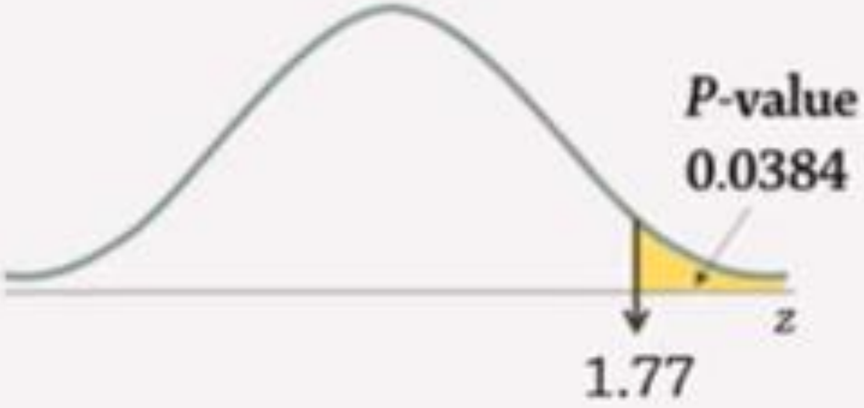
There is insufficient evidence to conclude that the average user exercises for more than 30 minutes per day

$n = 50$	$\bar{x} = 32$	$\sigma = 8$	$\alpha = 0.03$

$H_0: \mu \leq 30$

$H_1: \mu > 30$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{32 - 30}{\frac{8}{\sqrt{50}}} = \frac{2}{1.1314} \approx 1.77$$



Since P-value $= 0.0384 \nless 0.03$
fail to reject H_0 at $\alpha = 0.03$

not significant

WHAT IS Z-TEST FOR TWO-SAMPLE MEAN TEST?

The z-test for two-sample mean test is another parametric test used to compare the means of two independent groups of samples drawn from a normal population, if there are more than 30 samples for every group.

WHEN DO WE USE Z-TEST FOR TWO-SAMPLE MEAN?

- When we compare the means of samples of independent groups taken from a normal population.

WHY DO WE USE THE Z-TEST?

- We use the z-test to find out if there is a significant difference between two populations by only comparing the sample mean of the population.

HOW DO WE USE Z-TEST FOR A TWO-SAMPLE MEAN TEST?

The formula is,

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

$\bar{x}_1$ = the mean of sample 1 $\quad$ $\bar{x}_2$ = the mean of sample 2

s_1^2 = the variance of sample 1 $\quad$ s_2^2 = the variance of sample 2

n_1 = size of sample 1 $\quad$ n_2 = size of sample 2

Example:

An admission test was administered to incoming freshmen in the College of Medical Laboratory and Sciences and College of Radiologic Technology with 100 students each college randomly selected. The mean scores of the given samples were $\bar{x}_1 = 90$ and $\bar{x}_2 = 85$ and the variances of the test scores were 40 and 35 respectively. Is there a significant difference between the two groups? Use .01 level of significance.

SOLVING BY STEPWISE METHOD

I. Problem:

Is there a significant difference between the two groups?

II. Hypotheses:

H_0: There is no significant difference between the two groups. ($\bar{x}_1 = \bar{x}_2$)

H_1: There is a significant difference between the two groups. ($\bar{x}_1 \neq \bar{x}_2$)

IV. Statistics: Z-test for Two-sample Means

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$\bar{x}_1 = 90$ $\bar{x}_2 = 85$

$s_1^2 = 40$ $s_2^2 = 35$

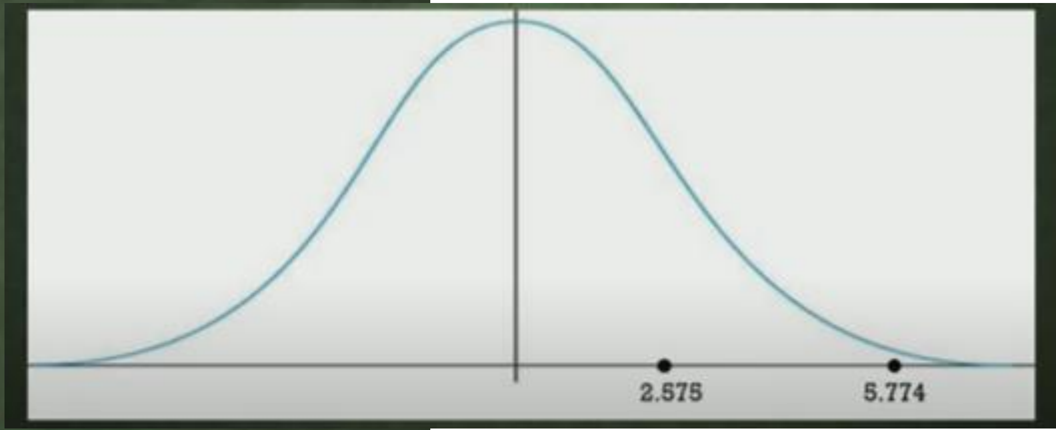
$n_1 = 100$ $n_2 = 100$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{90 - 85}{\sqrt{\frac{40}{100} + \frac{35}{100}}}$$

$$= \frac{5}{\sqrt{\frac{75}{100}}} = \frac{5}{\sqrt{.75}} = \frac{5}{.866} = 5.774$$

III. Level of Significance:

$\alpha = 0.01$



Test	Level of Significance	
	.01	.05
One-tailed	± 2.33	± 1.645
Two-tailed	± 2.575	± 1.96

VI. Conclusion:

Since the z-computed value of 5.774 is greater than the z-tabular value of 2.575 at .01 level of significance, the alternative hypothesis is confirmed which means that there is a significant difference between the two groups. It implies that the incoming freshmen of the College of Medical Laboratory and Sciences are better than the incoming freshmen of the College of Radiologic Technology.

Type I Error: Rejecting the null hypothesis when it is true.

Type 2 Error: Not rejecting the null hypothesis when it is false.

P(type I error / Ho is true) = α

P(type II error / Ho is false) = β

P(rejecting a false Ho) = $1 - \beta$

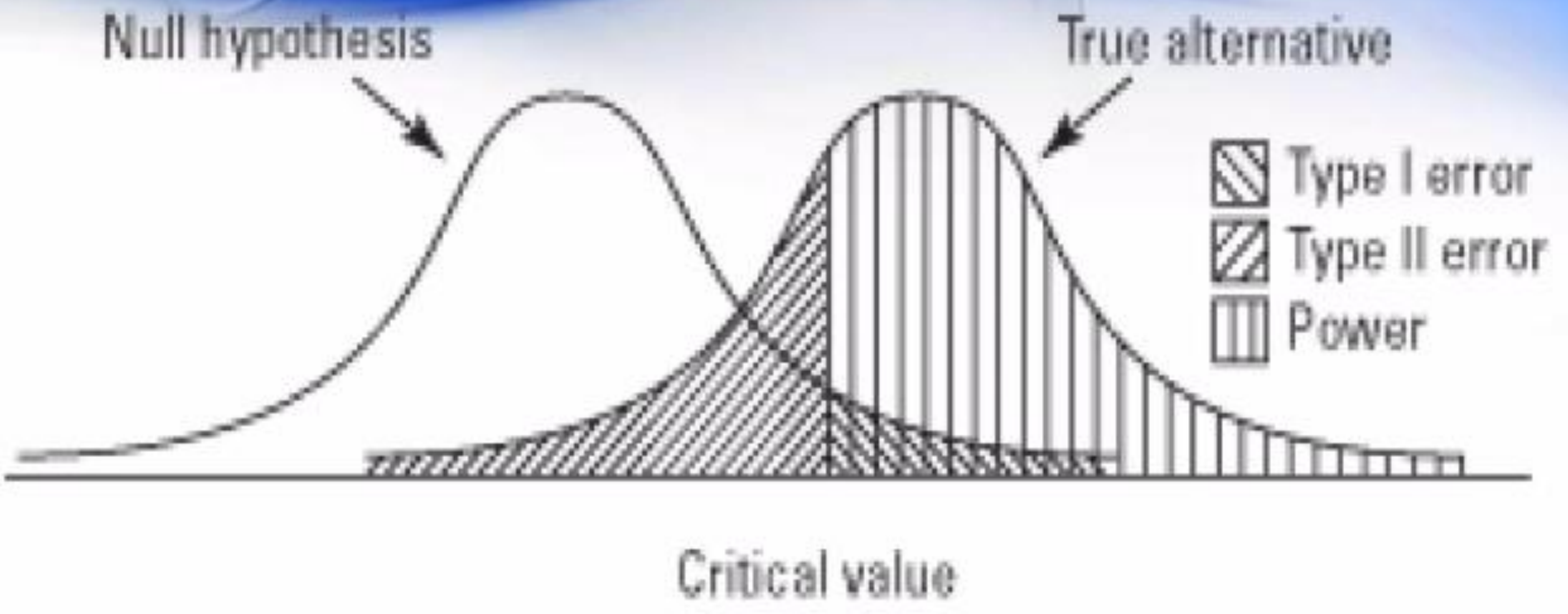
	Ho	
	True	False
Reject Ho	Type I error	✓
Fail to Reject Ho	✓	Type II error

Type I Error

- A **type I error**, also known as an **error of the first kind**, occurs when the null hypothesis (H_0) is true, but is rejected.
- A type I error may be compared with a so called ***false positive.***
- A Type I error occurs when we believe a **falsehood**.
- The rate of the type I error is called the *size* of the test and denoted by **the Greek letter α (alpha).**
- It usually equals the **significance level of a test**.
- If type I error is fixed at 5 %, it means that there are about 5 chances in 100 that we will reject H_0 when H_0 is true.

Type II Error

- **Type II error**, also known as an **error of the second kind**, occurs when the null hypothesis is false, but erroneously fails to be rejected.
- Type II error means accepting the hypothesis **which should have been rejected**.
- A type II error may be compared with a so-called ***False Negative.***
- A Type II error is committed when **we fail to believe a truth.**
- A type II error occurs when one rejects the alternative hypothesis (**fails to reject the null hypothesis**) when the **alternative hypothesis is true.**
- The rate of the type II error is denoted by the **Greek letter β (beta)** and related to the power of a test **(which equals 1-β)**.



Graphical depiction of the relation between Type I and Type II errors

What are the differences between Type 1 errors and Type 2 errors?

Type 1 Error

- A type 1 error is when a statistic calls for the rejection of a null hypothesis which is factually true.
- We may reject H_0 when H_0 is true is known as Type I error .
- A type 1 error is called a **false positive.**
- It denoted by **the Greek letter α (alpha).**
- Null hypothesis and type I error

Type 2 Error

- A type 2 error is when a statistic does not give enough evidence to reject a null hypothesis even when the null hypothesis should factually be rejected.
- We may accept H_0 when infect H_0 is not true is known as Type II Error.
- A type 2 error is a false negative.
- It denoted by the ***Beta***
- Alternative hypothesis and type II error

Reducing Type I Errors

- **Prescriptive testing** is used to increase the level of confidence, which in turn reduces Type I errors. The chances of making a Type I error are reduced by increasing the level of confidence.

Reducing Type II Errors

- **Descriptive testing** is used to better describe the test condition and acceptance criteria, which in turn reduces Type II errors. This *increases the number of times we reject the Null hypothesis* – ***with a resulting increase in the number of Type I errors*** (rejecting H0 when it was really true and should not have been rejected).

Therefore, reducing one type of error comes at the expense of increasing the other type of error! THE SAME MEANS CANNOT REDUCE BOTH TYPES OF ERRORS SIMULTANEOUSLY!

Type III Errors

- Many statisticians are now adopting a third type of error, a type III, which is where the null hypothesis was rejected for the wrong reason.
- In an experiment, a researcher might assume a hypothesis and perform research. After analyzing the results statistically, the null is rejected.
- The problem is, that there may be some relationship between the variables, but it could be for a different reason than stated in the hypothesis. An unknown process may underlie the relationship.

t distribution

Suppose we are about to draw a random sample of n observations from a normally distributed population.

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

has the standard normal distribution

t distribution

Suppose we are about to draw a random sample of n observations from a normally distributed population.

The population standard deviation is almost always unknown

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

has the standard normal distribution

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

The sample standard deviation

t distribution

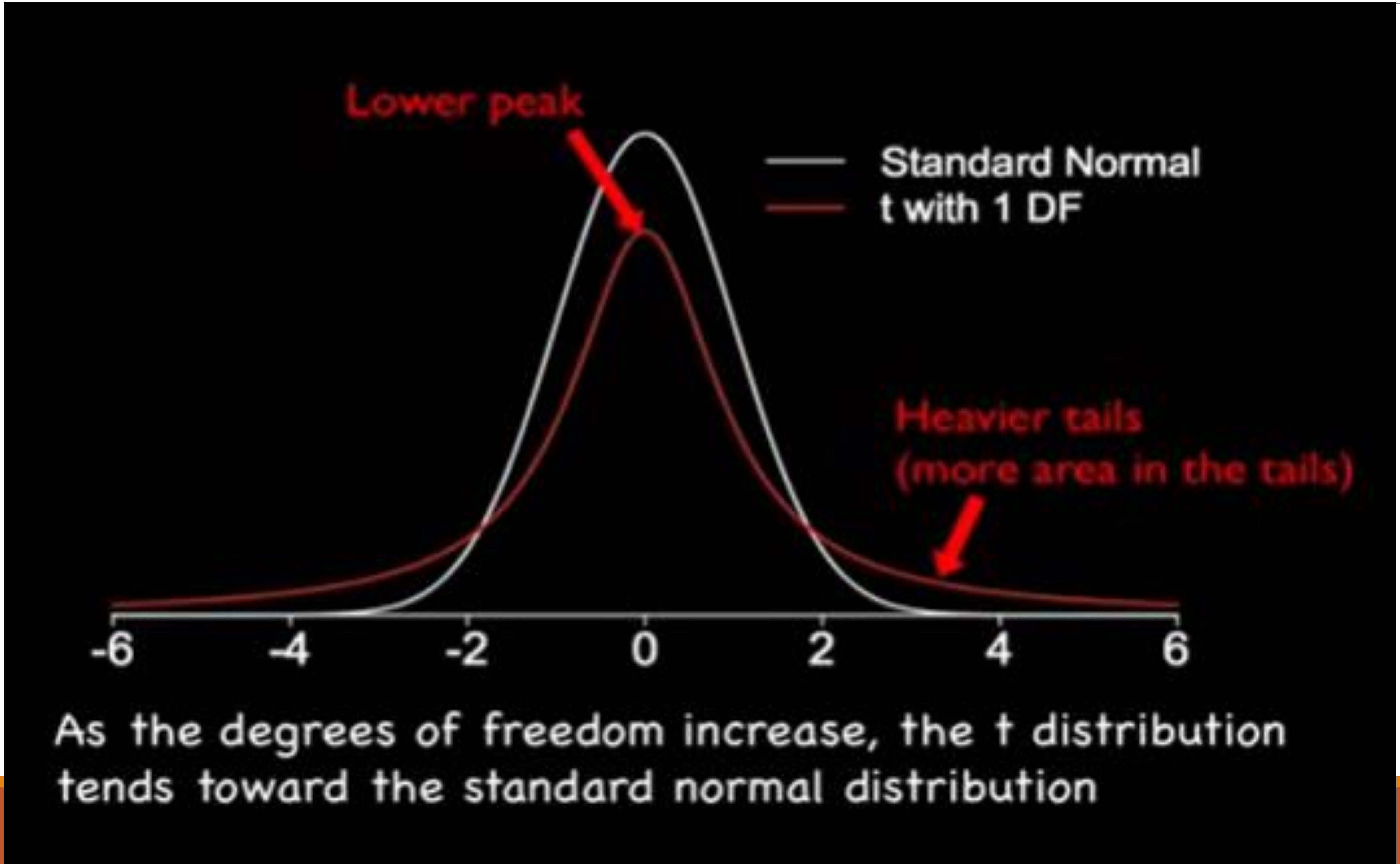
$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

has the t distribution with $n - 1$ degrees of freedom

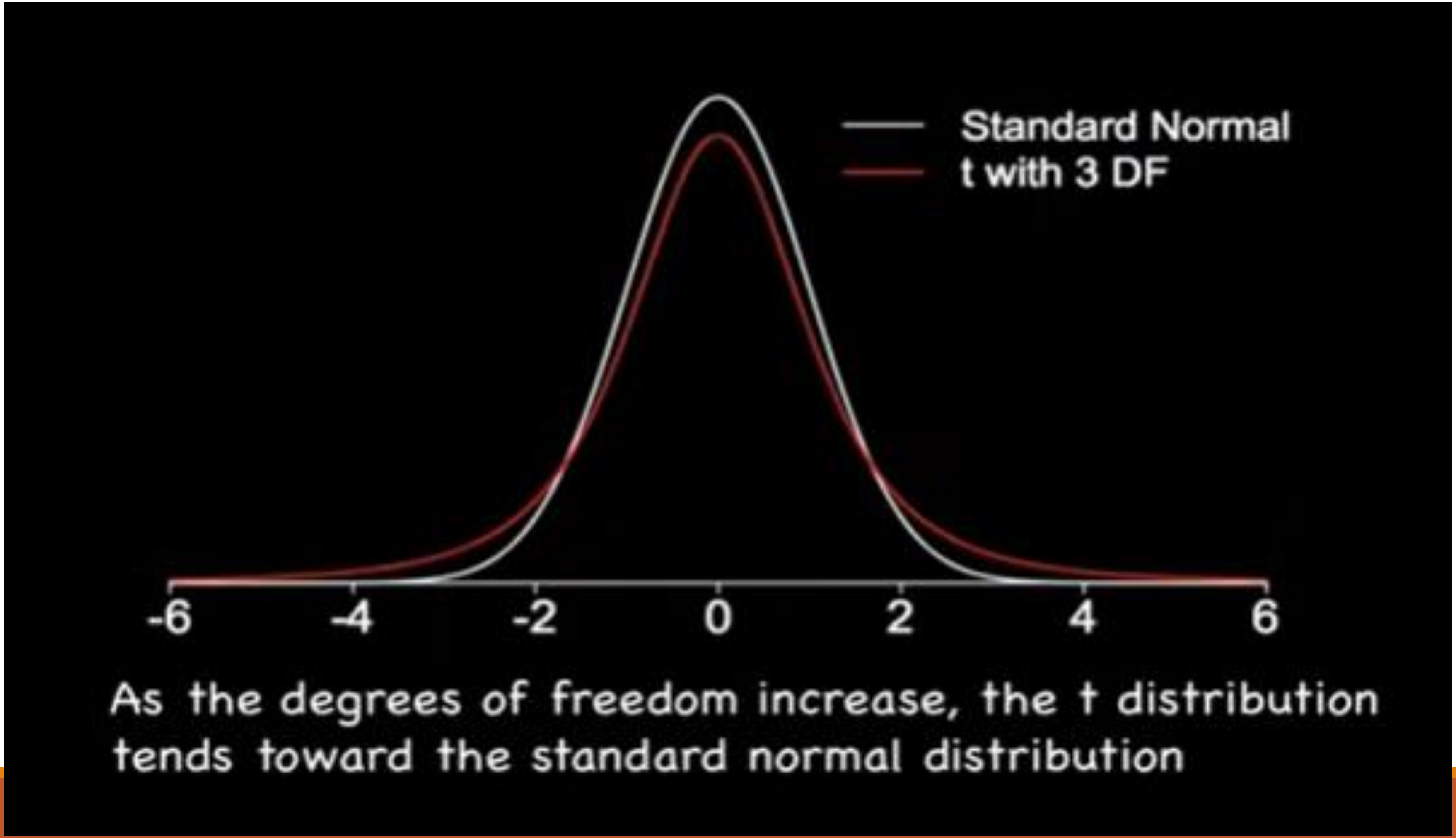
S^2

The denominator of the sample variance is n - 1

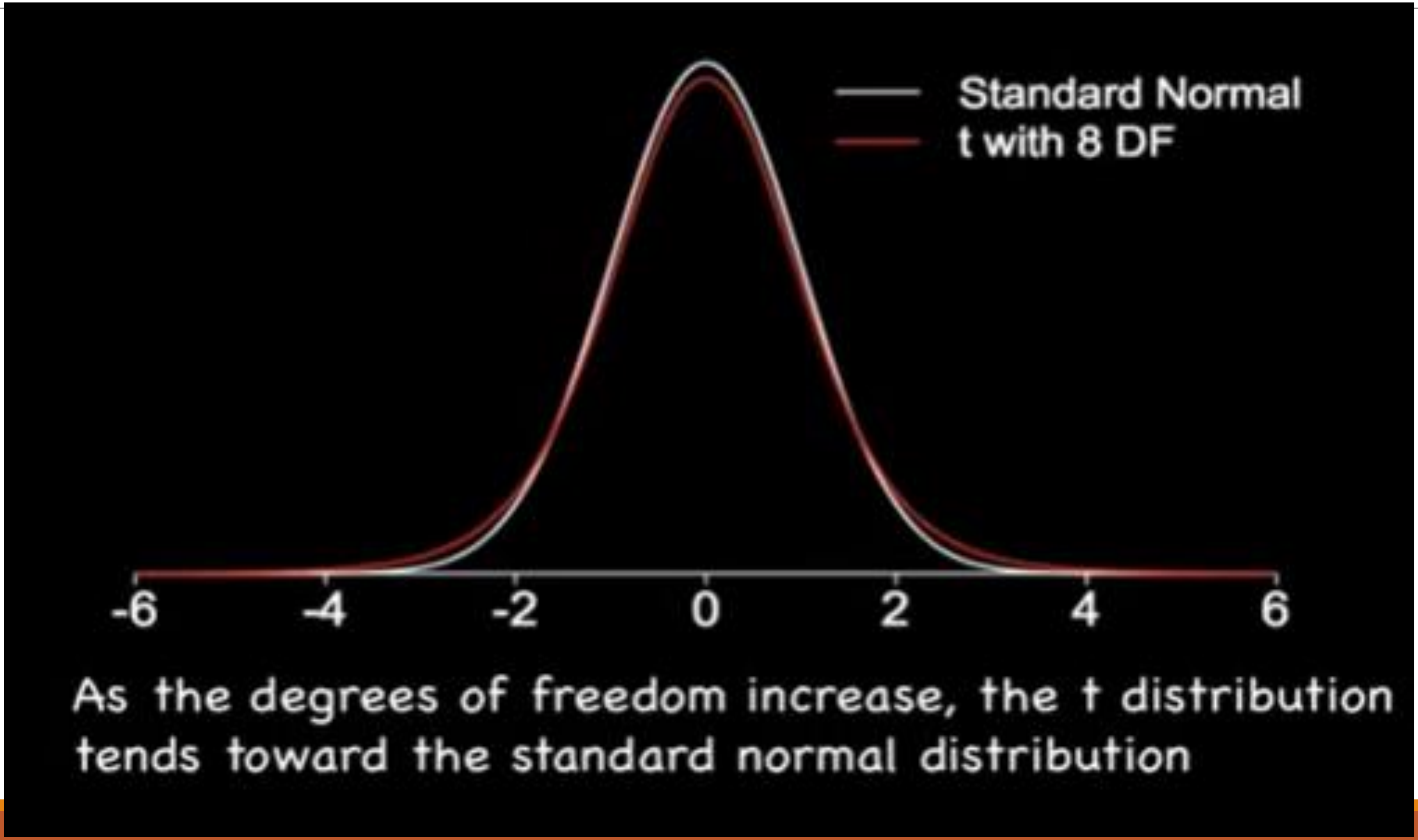
t distribution



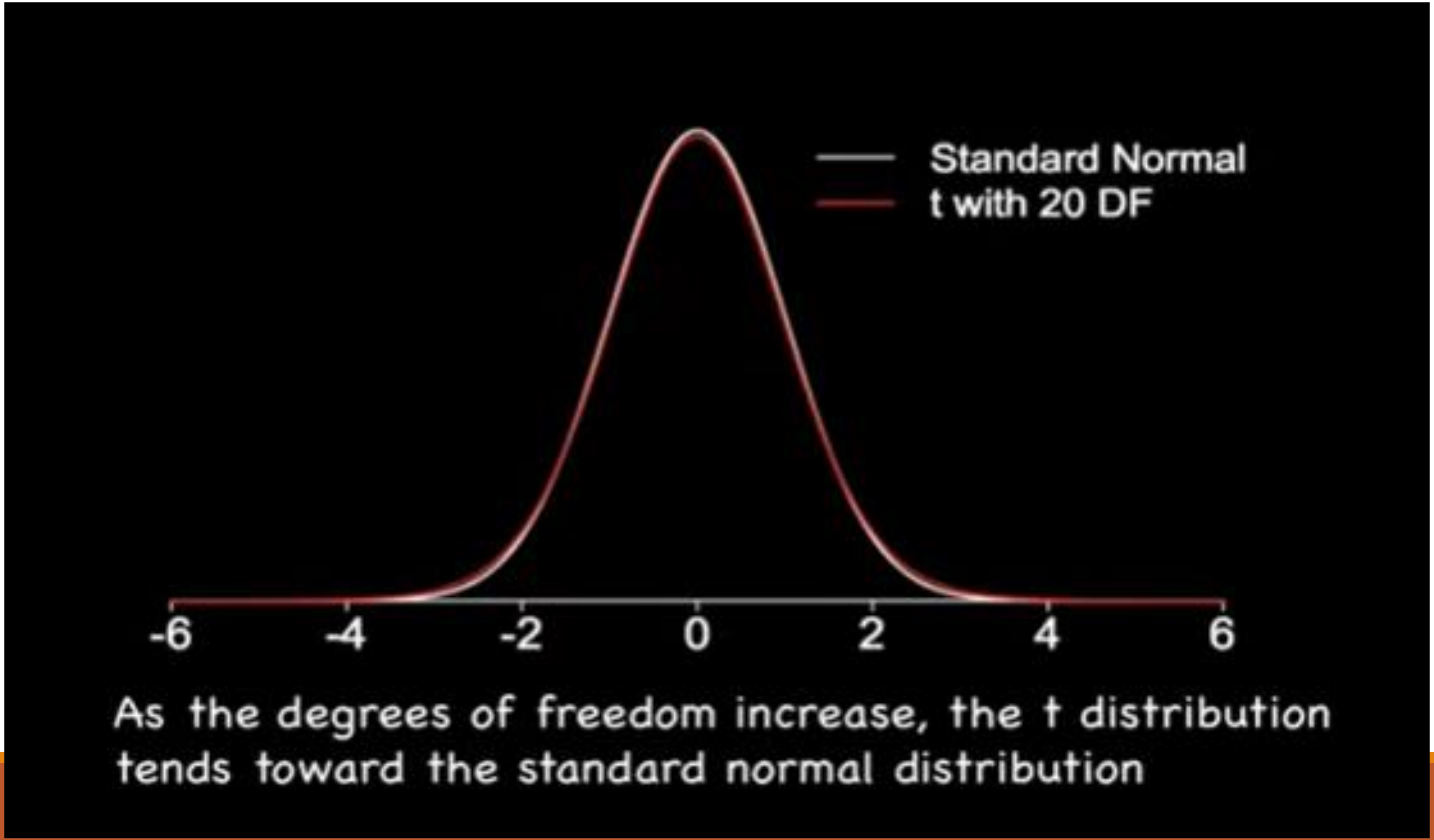
t distribution



t distribution



t distribution



Constructing a 95% confidence interval

(for the population mean)

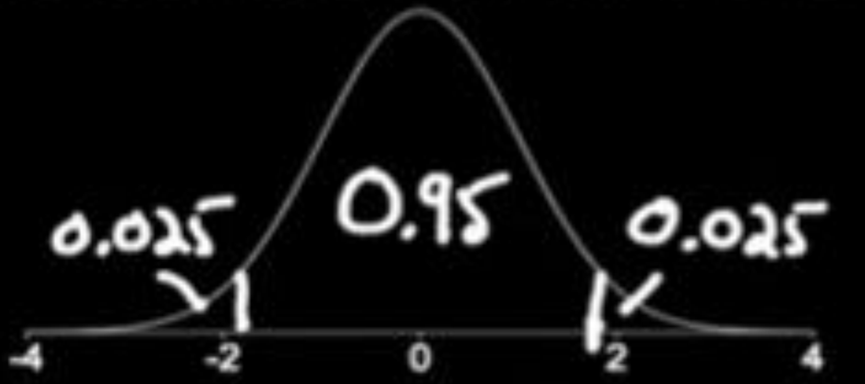
If σ is known:

$$\bar{X} \pm 1.96 \times \frac{\sigma}{\sqrt{n}}$$

If σ is not known:

$$\bar{X} \pm\ ? \times \frac{s}{\sqrt{n}} \leftarrow$$

The standard normal distribution



$z_{0.025} = 1.96$

A t distribution

The t value that has an area of 0.025 to the right $\rightarrow t_{.025} > 1.96$

$t_{.025}$ values (for 95% confidence intervals)

n	df	$t_{.025}$
6	5	2.571
11	10	2.228
31	30	2.042
51	50	2.009
101	100	1.984
∞	∞	1.960

$z_{.025}$

$t_{.025}$ values (for 95% confidence intervals)

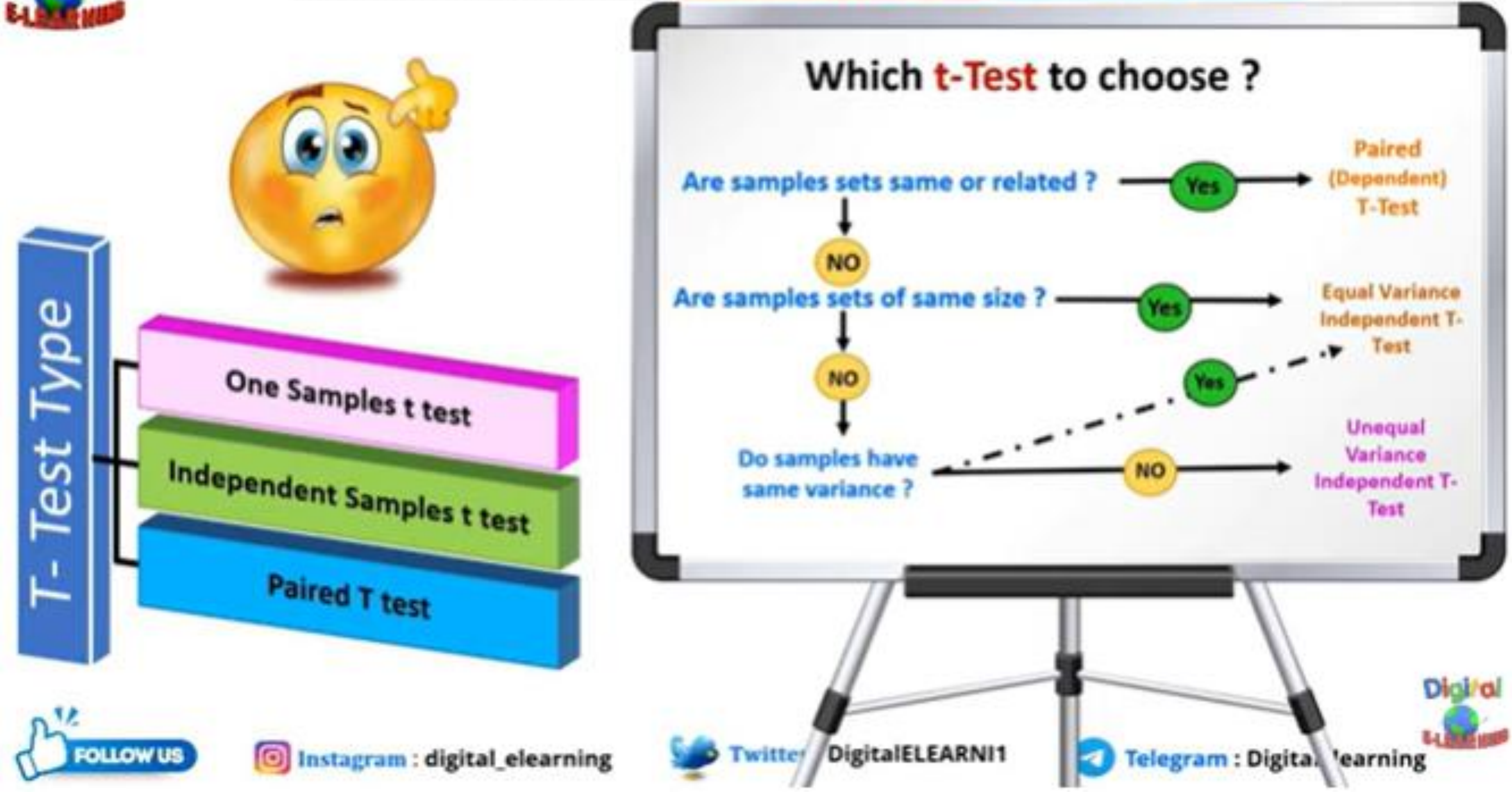
~~Some sources say:
If n > 30 forget about t
and use the standard
normal distribution~~

n	df	$t_{.025}$
6	5	2.571
11	10	2.228
31	30	2.042
51	50	2.009
101	100	1.984
∞	∞	1.960

$z_{.025} \rightarrow \infty$

If we use z when we should be using t, our calculated margin of error will be smaller than it should be

Understanding Student's T- Distribution



One Sample T-test Example

Company XYZ wants to test the claim that their batteries last more than 40 hours. Using a simple random sampling method, we draw 15 batteries with mean of 44.9 hrs. and standard deviation of 8.9 hours. We need to test this claim .

One Sample T-test Example

Company XYZ wants to test the claim that their batteries last more than 40 hours. Using a simple random sampling method, we draw 15 batteries with mean of 44.9 hrs. and standard deviation of 8.9 hours. We need to test this claim .

Ans: $t = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$

$H_o: \mu = 40$

$H_o: \mu > 40$

$\mu = 40$

$\overline{X} = 44.9$ $s = 8.9$

$df = n - 1$ = 15-1 => 14

H0: μ ≤ 40

H1: μ > 40

Right Tailed

$$t_{stat} = \frac{44.9 - 40}{8.9/\sqrt{15}}$$

$$t_{stat} = 2.13$$

One Sample T-tes

Company XYZ wants to test the claim that their bat
random sampling method, we draw 15 batteries wi
hours. We need to test this claim .

Ans: $t = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$

$H_o : \mu = 40$

$\mu = 40$

$df = n - 1 = 15\text{-}1 \Rightarrow 14$

$t_{stat} = \dfrac{44.9 - 40}{8.9/\sqrt{15}}$

$t_{Crit} = 1.761$

$t_{stat} = 2.13$

	P						
one-tail	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	0.2	0.1	0.05	0.02	0.01	0.002	0.001
Df							
1	3.078	6.314	12.706	31.821	63.656	318.289	636.578
2	1.886	2.92	4.303	6.965	9.925	22.328	31.6
3	1.638	2.353	3.182	4.541	5.841	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.61
5	1.476	2.015	2.571	3.365	4.032	5.894	6.869
6	1.44	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.86	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.25	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.93	4.318
13	1.35	1.771	2.16	2.65	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.14
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.12	2.583	2.921	3.686	4.015
17	1.333	1.74	2.11	2.567	2.898	3.646	3.965
18	1.33	1.734	2.101	2.552	2.878	3.61	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.85
21	1.323	1.721	2.08	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.5	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.06	2.485	2.787	3.45	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.689
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.66
30	1.31	1.697	2.042	2.457	2.75	3.385	3.646
60	1.296	1.671	2	2.39	2.66	3.232	3.46
120	1.289	1.658	1.98	2.358	2.617	3.16	3.373
1000	1.282	1.646	1.962	2.33	2.581	3.098	3.3
inf	1.282	1.645	1.96	2.326	2.576	3.091	3.291

One Sample T-test Example

Company XYZ wants to test the claim that their batteries last more than 40 hours. Using a simple random sampling method, we draw 15 batteries with mean of 44.9 hrs. and standard deviation of 8.9 hours. We need to test this claim .

Ans: $t = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$

$H_o : \mu = 40$

$H_o : \mu > 40$

$\mu = 40$

$\overline{X} = 44.9$

$s = 8.9$

$df = n - 1 = 15\text{-}1 \Rightarrow 14$

$t_{stat} = \dfrac{44.9 - 40}{8.9/\sqrt{15}}$

$t_{Crit} = 1.761$

$t_{stat} = 2.13$

Since $t_{Critical} <= t_{stat}$

Reject Null hypothesis

Two Sample - Paired T-test

- Is there a difference two dependent Groups at two point in time .

Two Sample - Independent T-test

- Is there a difference between a two independent Groups.

Understanding Student's T- Distribution

Formulae

One Sample T-test

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Two Sample - Paired T-test

$$t = \frac{\sum (x_1 - x_2) / n}{s_d / \sqrt{n}}$$

Two Sample - Independent T-test

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Where, $df = df1 + df2 = n_1 + n_2 - 2$

Independent Samples t -Test (Equal variances)

$$t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \qquad s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

Where,

$$S_p{}^2 = \frac{S_1{}^2(df1) + S_2{}^2(df2)}{df1 + df2}$$

df1= n_1-1

df2= n_2-1

Where, df=df1+df2= n_1+n_2 -2

Example: A random sample of 27 observations from a large population has a mean of 22 and a standard deviation of 4.8. Can we conclude at $\alpha = 0.01$ that the population mean is significantly below 24?

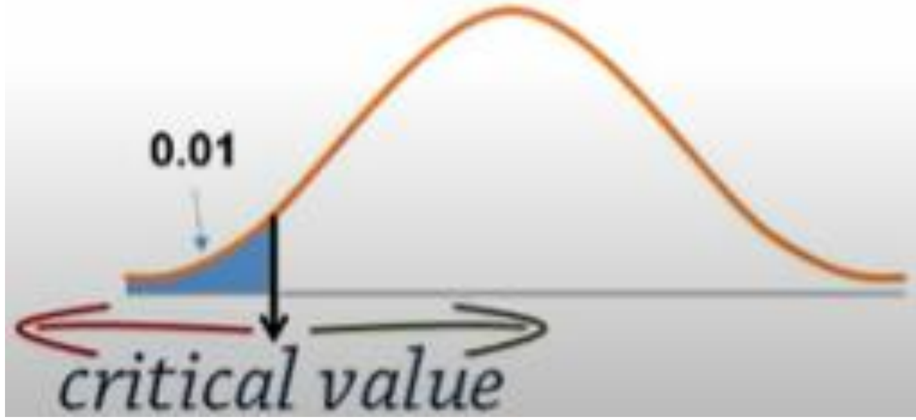
hypothesis testing: *t* - test

Example: A random sample of 27 observations from a large population has a mean of 22 and a standard deviation of 4.8. Can we conclude at α = 0.01 that the population mean is significantly below 24?

$n = 27 \quad \bar{x} = 22 \quad s = 4.8 \quad \alpha = 0.01$

$$H_0: \mu \geq 24$$
$$H_1: \mu < 24$$

$$\alpha = 0.01$$



https://www.youtube.com/watch?v=O7uJoAVl3qo&t=26s

1 tail, α	0.10	0.05	0.025	0.01	0.005
2 tails, α	0.20	0.10	0.05	0.02	0.01
df 1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
∞	**1.282**	**1.645**	**1.960**	**2.326**	**2.576**

Example: A random sample of 27 observations from a large population has a mean of 22 and a standard deviation of 4.8. Can we conclude at α = 0.01 that the population mean is significantly below 24?

$n = 27 \quad \bar{x} = 22 \quad s = 4.8 \quad \alpha = 0.01$

$H_0\text{: } \mu \geq 24$
$H_1\text{: } \mu < 24$

$\alpha = 0.01 \qquad df = 26$

Reject H_0 if $t < -2.479$

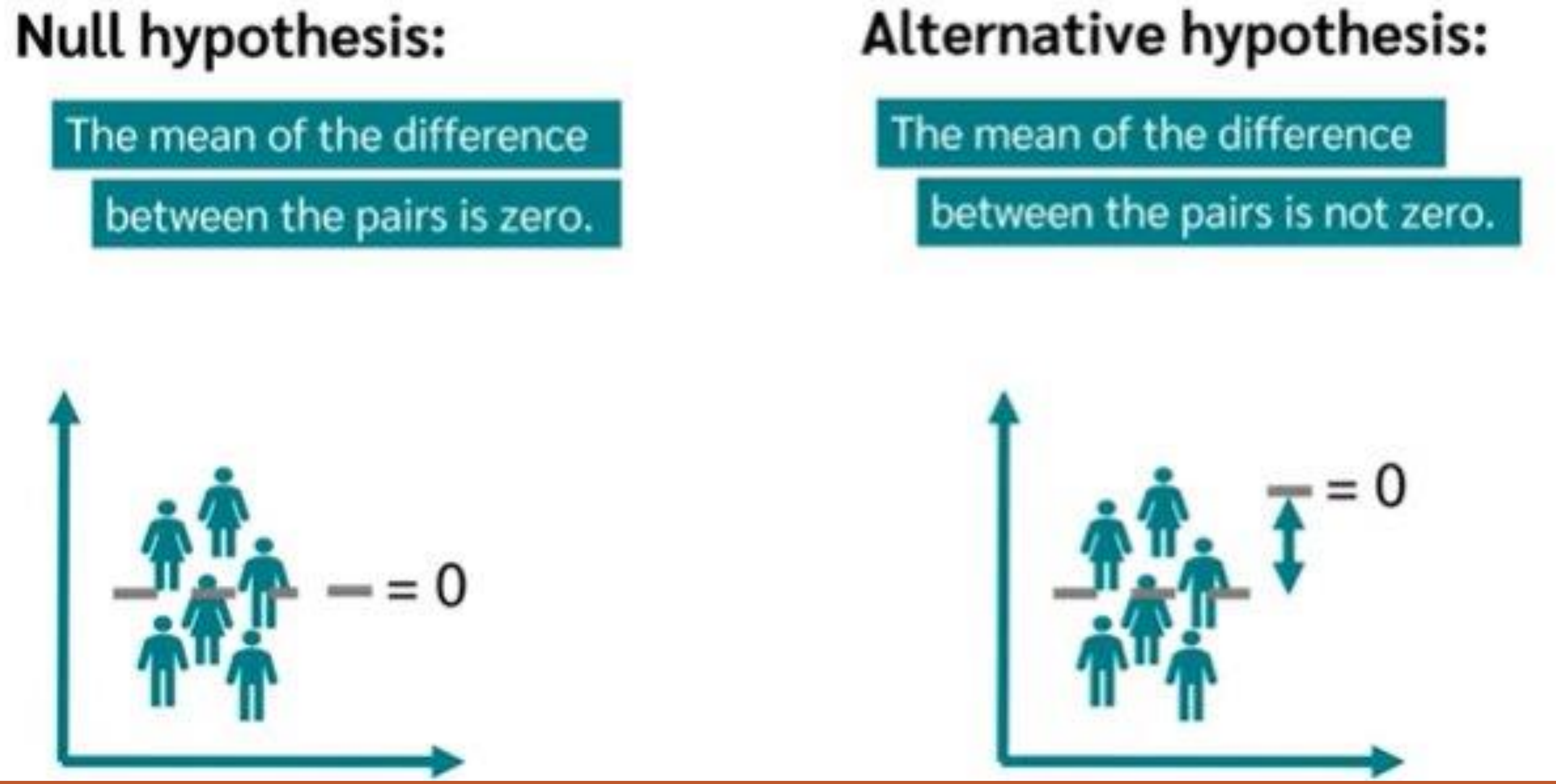
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{22 - 24}{4.6/\sqrt{27}} = -2.165$$

Since t = -2.165 is not less than -2.479,

Fail to Reject H_0

There is not enough evidence that the population mean is less than 24.

Paired t- test



Paired T-test Example

Consider study on two different types of exercise equipment measuring heart rate (beats per min) for individuals who used the two types of equipment A and B

Person	1	2	3	4	5
A	61	72	66	88	75
B	55	75	68	81	78
Di=Ai-Bi	6	-3	-2	7	-3

Ans: $t = \dfrac{\bar{d}}{s_d/\sqrt{n}}$ $t = \dfrac{\Sigma d_i/n}{s_d/\sqrt{n}}$ $t = \dfrac{\Sigma(X_1 - X_2)/n}{s_d/\sqrt{n}}$

Paired T-test Example

Consider study on two different types of exercise equipment measuring heart rate (beats per min) for individuals who used the two types of equipment A and B

Person	1	2	3	4	5
A	61	72	66	88	75
B	55	75	68	81	78
Di=Ai-Bi	6	-3	-2	7	-3

Ans: $t = \dfrac{\bar{d}}{s_d/\sqrt{n}}$ $\quad t = \dfrac{\Sigma d_i/n}{s_d/\sqrt{n}}$ $\quad t = \dfrac{\Sigma(X_1 - X_2)/n}{s_d/\sqrt{n}}$ $\quad \alpha = 0.05$

$H_o: \mu_D = 0 \quad H_a: \mu_D \neq 0 \quad df = n - 1 = 5\text{-}1 \Rightarrow 4$

$\bar{d} = \Sigma d_i/n \quad \bar{d} = 6 + (-3) + (-2) + 7 + (-3)/n \quad \bar{d} = 5/5 \quad \bar{d} = 1$

$S_d = \dfrac{\Sigma\sqrt{(di-\bar{d})^2}}{\sqrt{n-1}} \quad S_d = \dfrac{\Sigma\sqrt{(6-1)^2+(-3-1)^2+(-2-1)^2+(7-1)^2+(-3-1)^2}}{\sqrt{5-1}} \quad \cancel{S_d = 51}$, Sd = 5.05

$t = \dfrac{\bar{d}}{s_d/\sqrt{n}} \quad \cancel{t = \dfrac{1}{51/\sqrt{4}}} \quad t_{stat} = \cancel{0.04}$

$t_{Crit}\,(\alpha/2, 4) = 2.776$

t= 1/(5.05/√5)= 0.442

Paired T-test Example

Consider study on two different types of exercise equipment measuring heart rate (beats per min) for individuals who used the two types of equipment A and B

Person	1	2	3	4	5
A	61	72	66	88	75
B	55	75	68	81	78
Di=Ai-Bi	6	-3	-2	7	-3

Ans: $t = \dfrac{\bar{d}}{s_d/\sqrt{n}}$ $\quad t = \dfrac{\Sigma d_i/n}{s_d/\sqrt{n}}$ $\quad t = \dfrac{\Sigma(X_1 - X_2)/n}{s_d/\sqrt{n}}$

$\alpha = 0.05$

$H_o: \mu_D = 0 \quad H_a: \mu_D \neq 0 \quad df = n - 1 = 5\text{-}1 \Rightarrow 4$

$\bar{d} = \Sigma d_i/n \quad \bar{d} = 6 + (-3) + (-2) + 7 + (-3)/n \quad \bar{d} = 5/5 \quad \bar{d} = 1$

$S_d = \dfrac{\Sigma\sqrt{(di-\bar{d})^2}}{\sqrt{n-1}} \quad S_d = \dfrac{\Sigma\sqrt{(6-1)^2+(-3-1)^2+(-2-1)^2+(7-1)^2+(-3-1)^2}}{\sqrt{5-1}} \quad \cancel{S_d = 51}$, Sd = 5.05

$t = \dfrac{\bar{d}}{s_d/\sqrt{n}} \quad \cancel{t = \dfrac{1}{51/\sqrt{4}}} \quad t_{stat} = \cancel{0.04}$

$t_{Crit\,(\alpha/2,\,4)} = 2.776$

t= 1/(5.05/√5)= $\boxed{0.442}$

-2.776< $\boxed{0.442}$ <2.776

Failed to reject Null Hypothesis

Example 9.2

A consumer rating magazine has developed a test to compare the mean life length of two brands of batteries. Five electronic toys were used to make the comparison. Two of each type of toy were purchased, and one toy of each type was powered by each type of battery (Type 1 and Type 2). The resulting durations of toy usage are given below.

Toy	Battery Type I	Battery Type II
1	52.6 hr	61.4 hr
2	103.4 hr	112.8 hr
3	68.2 hr	67.1 hr
4	88.4 hr	92.3 hr
5	111.6 hr	121.5 hr

a. State the null and alternative hypotheses.

b. Identify the test statistic and the critical region (use $\alpha = 0.05$).

c. State the conclusion resulting from the test and find the P-value for the test.

d. Construct a confidence interval that is relevant to this problem, and use it to verify the conclusion identified in part (c).

Example 9.2

paired t-test

A consumer rating magazine has developed a test to compare the mean life length of two brands of batteries. Five electronic toys were used to make the comparison. Two of each type of toy were purchased, and one toy of each type was powered by each type of battery (Type 1 and Type 2). The resulting durations of toy usage are given below.

Toy	Battery Type I (μ_1)	Battery Type II (μ_2)	D_i
1	52.6 hr	61.4 hr	52.6 − 61.4 = −8.8
2	103.4 hr	112.8 hr	−9.4
3	68.2 hr	67.1 hr	1.1
4	88.4 hr	92.3 hr	−3.9
5	111.6 hr	121.5 hr	−9.9

a. State the null and alternative hypotheses. $\begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D \neq 0 \end{cases}$

b. Identify the test statistic and the critical region (use α = 0.05).

c. State the conclusion resulting from the test and find the P-value for the test.

d. Construct a confidence interval that is relevant to this problem, and use it to verify the conclusion identified in part (c).

Example 9.2 (more space)

$$T_0 = \frac{\bar{D} - \Delta_0}{\frac{S_D}{\sqrt{n}}} = \frac{-6.18 - 0}{4.73/\sqrt{5}} = -2.93$$

$$\bar{D} = \frac{D_1 + D_2 + \cdots + D_5}{5} = -6.18$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{5} (D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{(-8.8+6.18)^2 + (-9.4+6.18)^2 + \cdots}{4}} = 4.73$$

C.R. $\begin{cases} t_{\alpha/2, n-1} = t_{0.025, 4} = 2.776 \\ -t_{\alpha/2, n-1} = -t_{0.025, 4} = -2.776 \end{cases}$

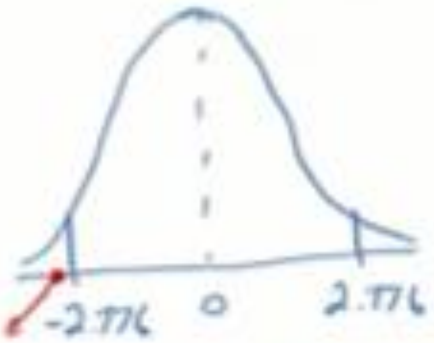
Example 9.2 (more space)

$$T_0 = \frac{\bar{D} - \Delta_0}{\frac{S_D}{\sqrt{n}}} = \frac{-6.18 - 0}{4.73/\sqrt{5}} = -2.93$$

$$\bar{D} = \frac{D_1 + D_2 + \cdots + D_5}{5} = -6.18$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{5}(D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{(-8.8+6.18)^2 + (-9.4+6.18)^2 + \cdots}{4}} = 4.73$$

C.R. $\left\{ \begin{array}{l} t_{\alpha/2, n-1} = t_{0.025,4} = 2.776 \\ -t_{\alpha/2, n-1} = -t_{0.025,4} = -2.776 \end{array} \right.$



Since $T_0 < -2.776$

$-2.93 < -2.776$

Reject H_0

Example 9.2 (more space)

$P\text{-value} = 2P(t_{n-1} > |T_0|) = 2(t_4 > 2.93)$

$P(t_4 > 2.93) \cong 0.02$

Table V Percentage Points $t_{\alpha,\nu}$ of the t Distribution

0.02

ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453
4	.271	.741	1.533	2.132	2.776 2.93	3.747	4.604	5.598
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833

Example 9.2 (more space)

$\alpha = 0.05$

$$P\text{-Value} = 2P(t_{n-1} > |T_0|) = 2(t_4 > 2.93) \cong 2 \times 0.02 = \underline{0.04}$$

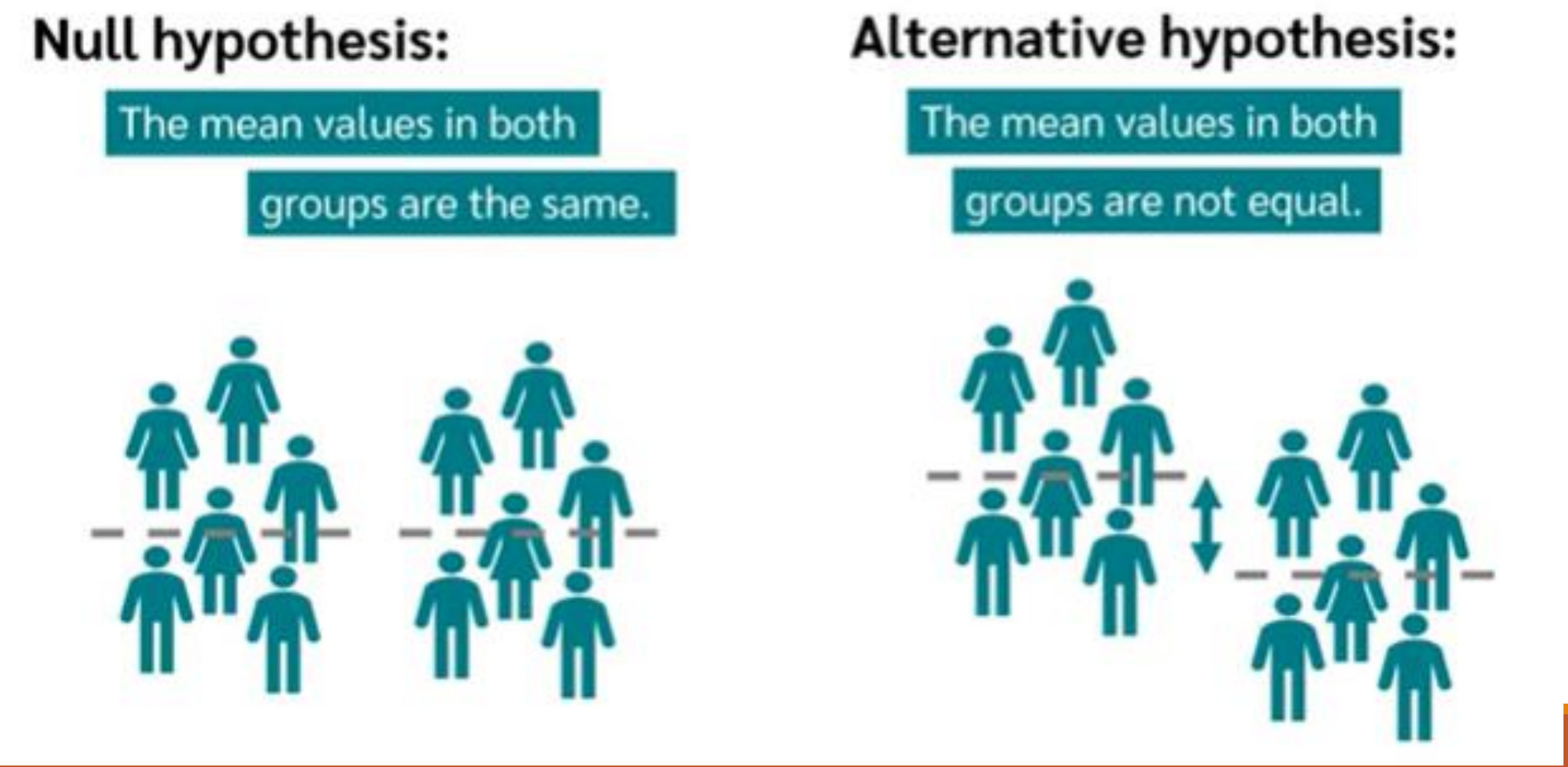
Since $P_V = 0.04 < \alpha = 0.05 \Rightarrow$ $\underline{\text{Reject } H_0}$

d) $$\bar{D} - t_{\alpha/2, n-1} \frac{S_D}{\sqrt{n}} \le \mu_D \le \bar{D} + t_{\alpha/2, n-1} \frac{S_D}{\sqrt{n}}$$

$$-6.18 - 2.776 \frac{4.73}{\sqrt{5}} \le \mu_D \le -6.18 + 2.776 \frac{4.73}{\sqrt{5}}$$

$$\underline{-12.04 \le \mu_D \le -0.32}$$

Independent t Test



Requirements for the independent t-test

There is an independent variable (e.g. gender), which has two characteristics or groups (e.g. male and female). These two groups should be compared in the analysis. The question is thus, is there a difference between the two groups regarding the dependent variable (e.g. income).

1. The two groups or samples must be independent.
2. The variables must be scaled in intervals.
3. The variables must be normally distributed.
4. The variance within the groups should be similar.

Requirements

1. The two groups or samples must be independent

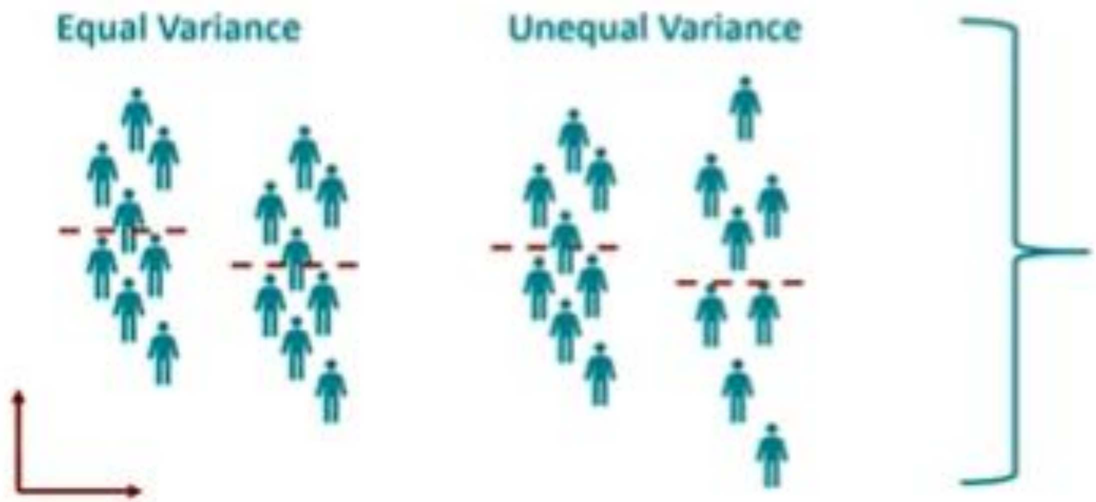
As the name of this t-test suggests, the samples must be independent. This means that a value in one sample must not influence a value in the other sample.

- ✓ Measuring the weight of people who have been on a diet and people who have not been on a diet.
- ✗ Measuring the weight of a person before and after a certain diet.

DATAtab
Webapp for data analysis

Level test for variance equality

How can it now be checked whether the variances are homogeneous, i.e. whether there is variance equality? This is where the **Levene test** helps. The level test checks whether several samples have the same variance.



If the p-value in the level test is greater than 0.05, variance equality can be assumed

Equal Variance

➢ Independent Samples t-Test

A statistics teacher wants to compare his two classes to see if they performed any differently on the tests he gave that semester. Class A had 25 students with an average score of 70, standard deviation 15. Class B had 20 students with an average score of 74, standard deviation 25. Using alpha 0.05, did these two classes perform differently on the tests?

1. **Define Null and Alternative Hypotheses**

$$H_0; \mu_{classA} = \mu_{classB}$$

$$H_1; \mu_{classA} \neq \mu_{classB}$$

2. **State Alpha**

$$\alpha = 0.05$$

3. Calculate Degrees of Freedom

<u>Degrees of Freedom for Independent Samples t-Test</u>

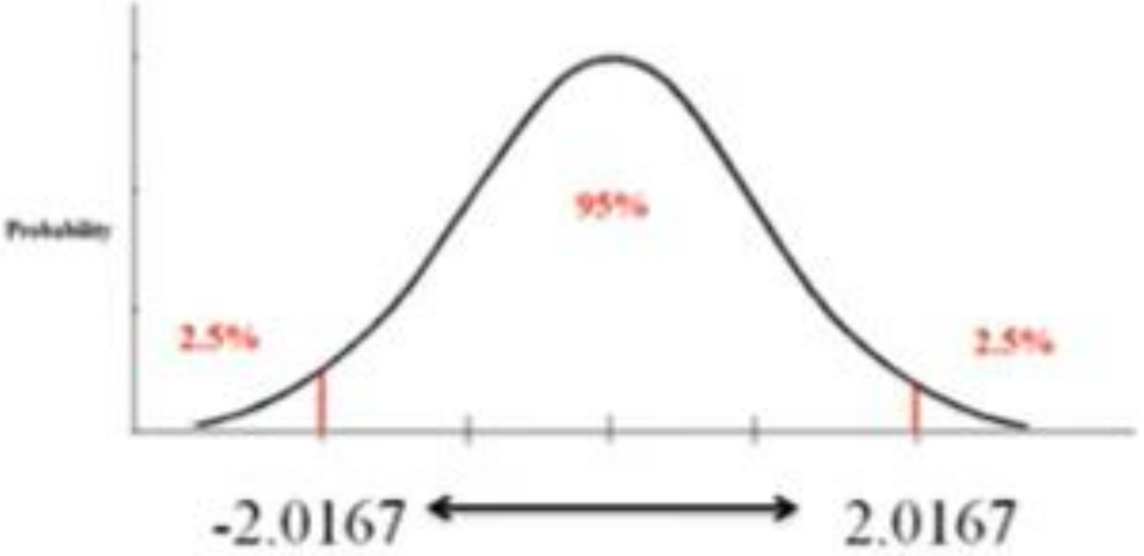
$$df = (n_1 - 1) + (n_2 - 1)$$

$$df = (25 - 1) + (20 - 1) = 43$$

4. State Decision Rule

df	Critical Values				
1-tailed	0.1	0.05	0.025	0.01	0.00
2-tailed	0.2	0.1	0.05	0.02	0.0
32	1.3086	1.6939	2.0369	2.4487	2.73
33	1.3077	1.6924	2.0345	2.4448	2.73
34	1.3070	1.6909	2.0322	2.4411	2.72
35	1.3062	1.6896	2.0301	2.4377	2.72
36	1.3055	1.6883	2.0281	2.4345	2.71
37	1.3049	1.6871	2.0262	2.4314	2.71
38	1.3042	1.6860	2.0244	2.4286	2.71
39	1.3036	1.6849	2.0227	2.4258	2.70
40	1.3031	1.6839	2.0211	2.4233	2.70
41	1.3025	1.6829	2.0195	2.4208	2.70
42	1.3020	1.6820	2.0181	2.4185	2.69
43	1.3016	1.6811	2.0167	2.4163	2.69
44	1.3011	1.6802	2.0154	2.4141	2.69

4. State Decision Rule



If t is less than -2.0167, or greater than 2.0167, reject the null hypothesis.

5. Calculate Test Statistic

$$t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

$$s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

$$df_1 = n_1 - 1 = 25 - 1 = 24$$

$$df_2 = n_2 - 1 = 20 - 1 = 19$$

$$SS_1 = s_1^2(df_1) = (15^2)(24) = 5400$$

$$SS_2 = s_2^2(df_2) = (25^2)(19) = 11875$$

$$s_p^2 = \frac{5400 + 11875}{24 + 19} = 401.74$$

5. Calculate Test Statistic

$$t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

$$s_p^2 = \frac{5400 + 11875}{24 + 19} = 401.74$$

$$t = \frac{(70 - 74)}{\sqrt{\frac{401.74}{25} + \frac{401.74}{20}}} = \frac{-4}{\sqrt{36.16}} = -0.67$$

6. State Results

Decision Rule: If t is less than -2.0167, or greater than 2.0167, reject the null hypothesis.

$t = -0.67$

Result: Do not reject H0.

7. State Conclusion

There was no significant difference between the test performances of Class A and Class B, $t = -0.67$, $p > 0.05$.

Unequal variance

➢ Independent Samples t-Test

A statistics teacher wants to compare his two classes to see if they performed any differently on the tests he gave that semester. Class A had 25 students with an average score of 70, standard deviation 15. Class B had 20 students with an average score of 74, standard deviation 25. Using alpha 0.05, did these two classes perform differently on the tests?

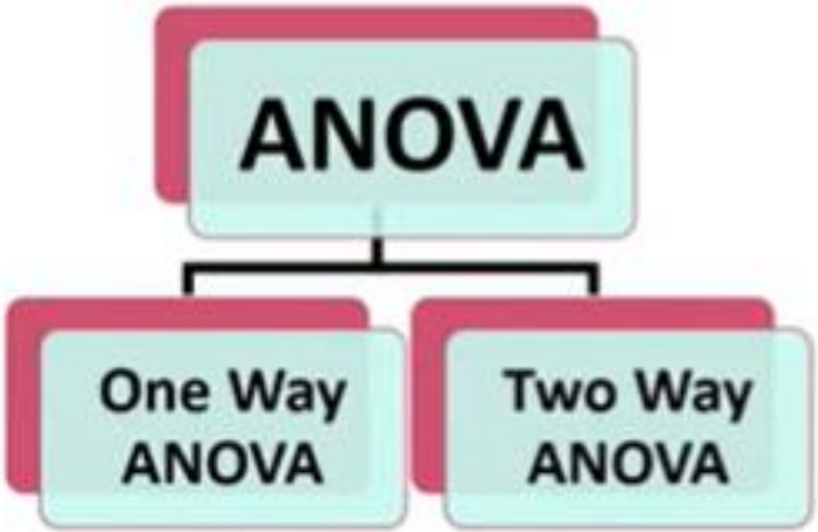
Basics of ANOVA

- ANOVA stands for Analysis of Variance.
- ANOVA enables us to test for significance of difference among more than two sample means.

Extension of t-Test

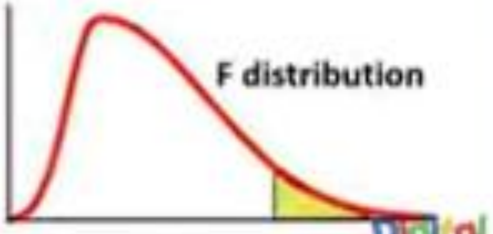
Basics of Anova

- ANOVA stands for **Analysis of Variance.**
- ANOVA enables us to test for significance of difference among more than two sample means.

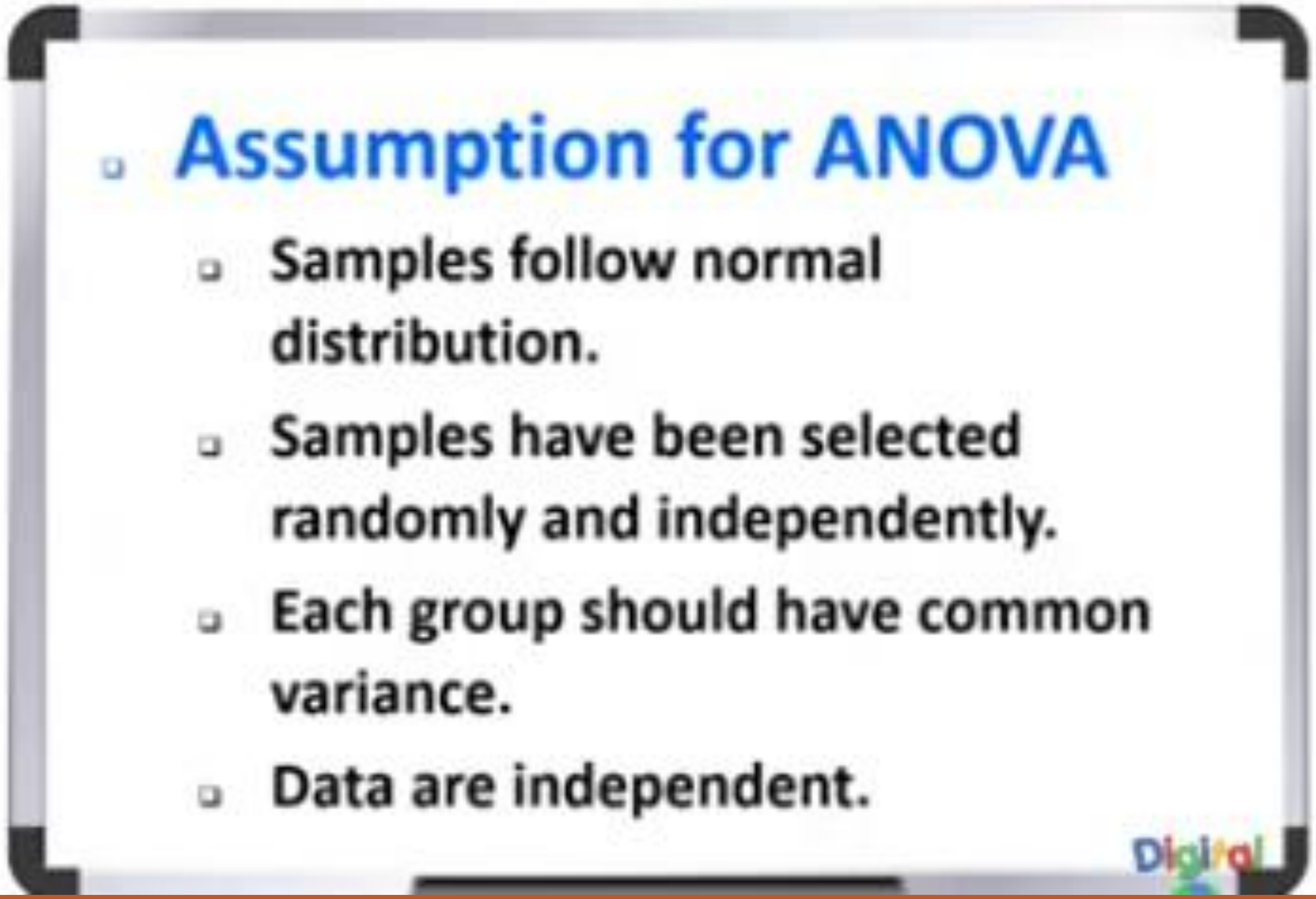


- One way ANOVA
  - One factor or independent variable.
  - Compares three or more levels of one factor.
- Two way ANOVA
  - Extension of One-way Anova
  - More than one factor or independent variable.
  - Compares the effect of multiple levels of two factors.

Test statistics for ANOVA is F-test



Assumptions for ANOVA



Basics of ANOVA

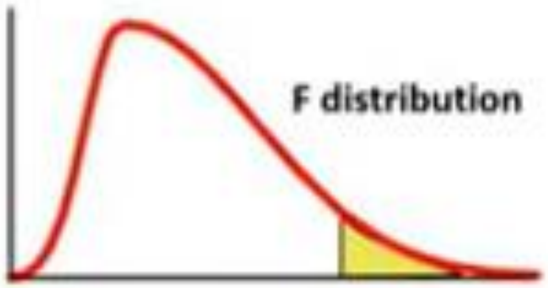
Null Hypothesis – The means for all groups are the same (equal).

$$H_0 : \mu_1 = \mu_2 = \mu_3 \ldots\ldots\ldots\ldots \mu_n$$

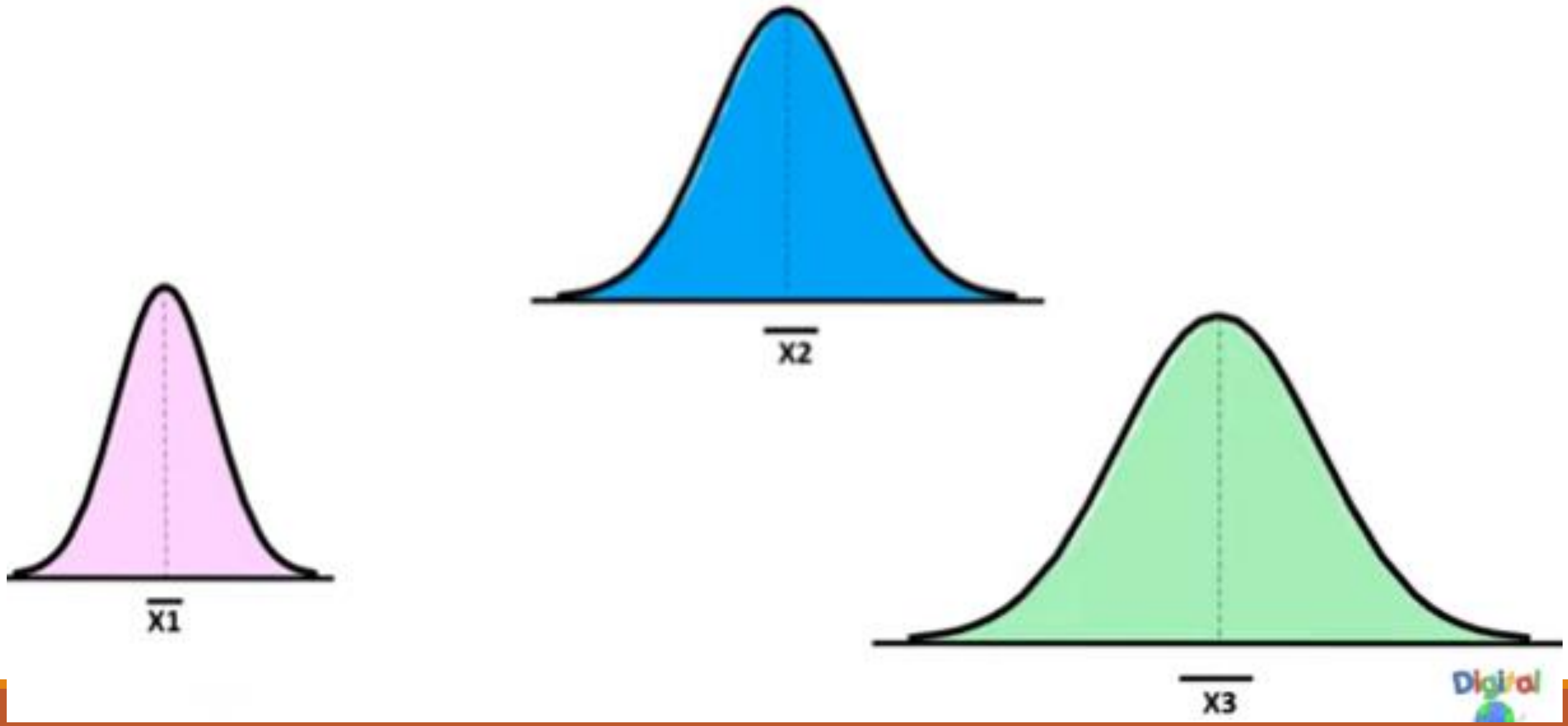
Alternate Hypothesis – The means are different for at least one pair of groups.

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots\ldots\ldots \mu_n$$

$$\text{ANOVA} = \frac{\text{Variance Between}}{\text{Variance Within}}$$

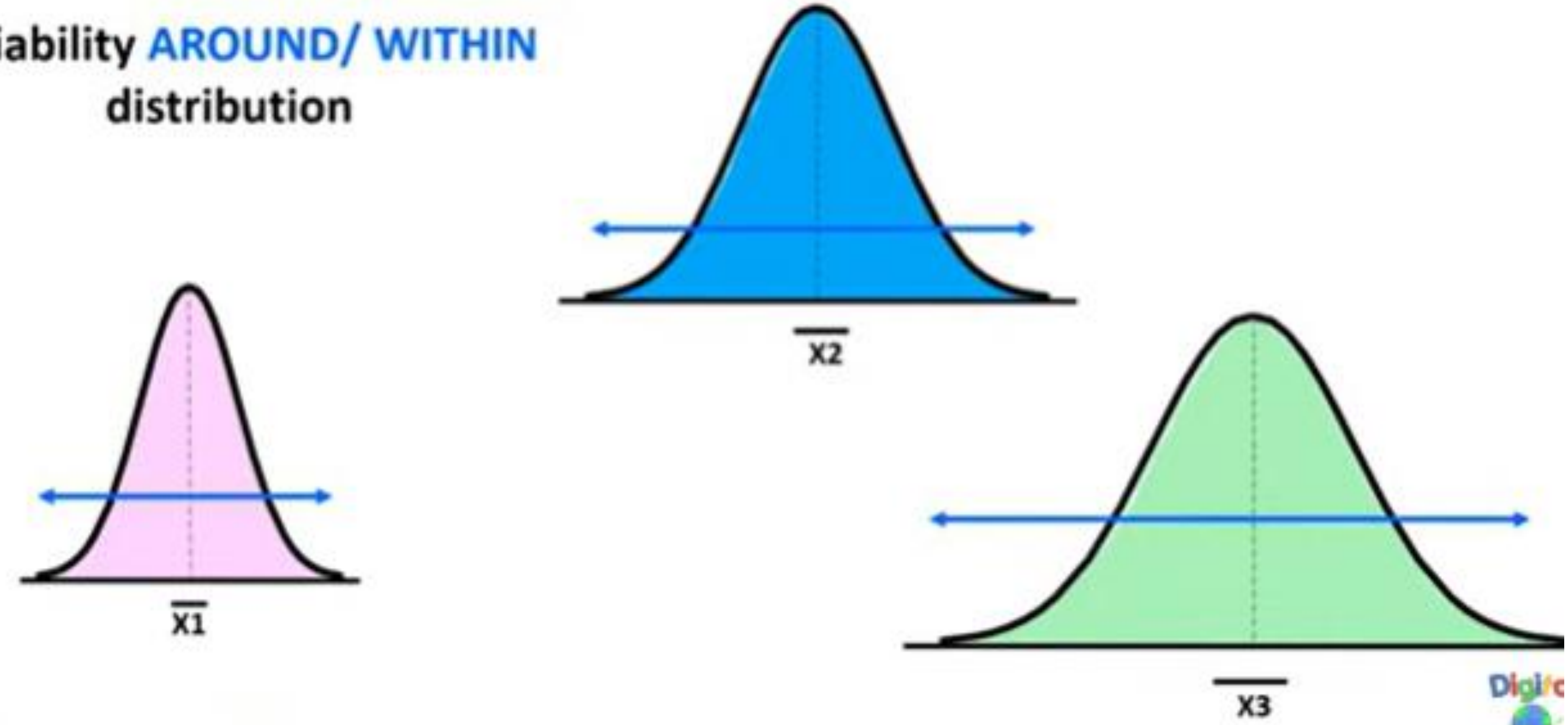


Basics of ANOVA

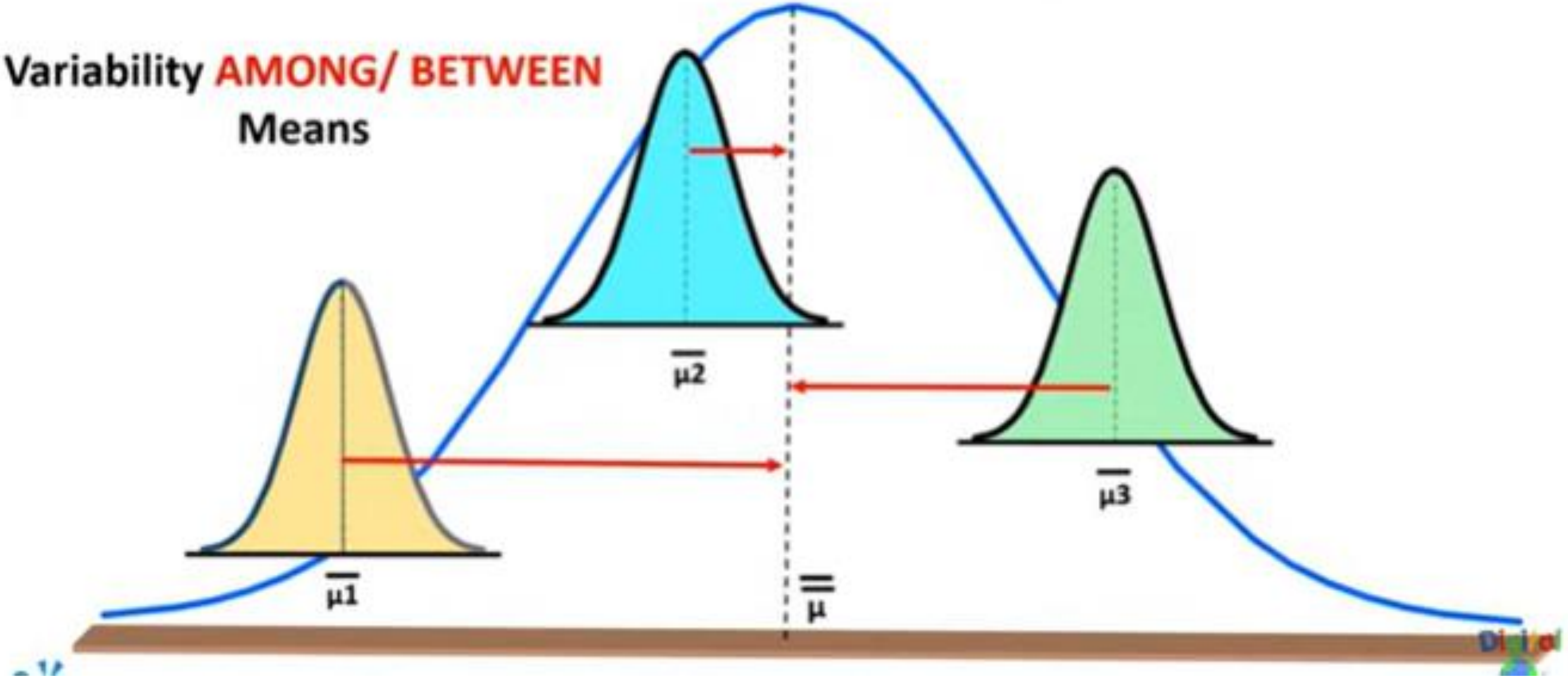


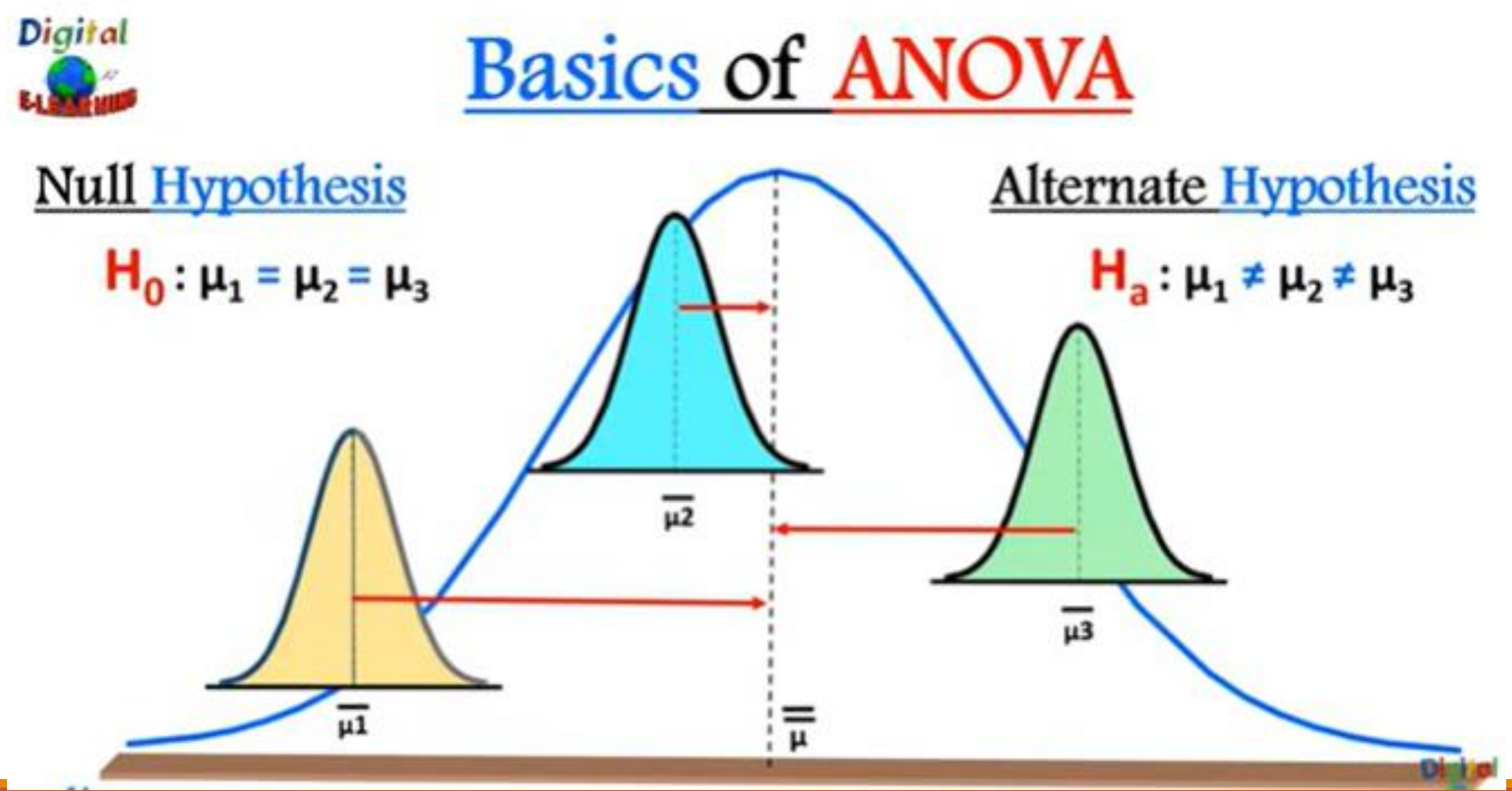
Basics of ANOVA

Variability AROUND/ WITHIN distribution



Basics of ANOVA




Digital E-LEARNING
Basics of ANOVA
Null Hypothesis
$H_0 : \mu_1 = \mu_2 = \mu_3$
Alternate Hypothesis
$H_a : \mu_1 \neq \mu_2 \neq \mu_3$
$\overline{\mu 1}$
$\overline{\mu 2}$
$\overline{\mu 3}$
$\overline{\overline{\mu}}$

Basics of ANOVA

$$\text{ANOVA} = \frac{\text{Variance Between}}{\text{Variance Within}}$$

Total Variance = Variance Between + Variance Within

$$\frac{\text{Variance Between}}{\text{Variance Within}} > 1 \quad \text{Reject } H_0$$

$$\frac{\text{Variance Between}}{\text{Variance Within}} < 1 \quad \text{Fail to Reject } H_0$$

$$\frac{\text{Variance Between}}{\text{Variance Within}} = 1 \quad \text{Fail to Reject } H_0$$

Basics of ANOVA

We want to see if three different studying methods can lead to different mean exam scores or not. To test this, we select 30 students and randomly assign 10 each to use a different studying method.

We will solve this using 2 different method. Let look at Method 1

Basics of ANOVA

Method 1

Basics of ANOVA

Sno	Method A	Method B	Method C
1.	10	8	9
2.	9	9	8
3.	8	10	7
4.	7.5	8	10
5.	8.5	8.5	9
6.	9	7	8
7.	10	9.5	7
8.	8	9	10
9.	8	7	9
10.	9	10	8
Group Mean	8.7	8.6	8.5

Overall Mean 8.6

Between Group Variation =10*(8.7-8.6)^2+10*(8.6- 8.6)^2+10*(8.5-8.6)^2

Between Group Variation = 0.2

Within Group Variation: $\Sigma(X_{ij} - X_j)^2$

Where:

Σ: a symbol that means "sum"

X_{ij}: the i^{th} observation in group j

X_j: the mean of group j

Method A: $(10-8.7)^2 + (9-8.7)^2 + (8-8.7)^2 + (7.5-8.7)^2 + (8.5-8.7)^2 + (9-8.7)^2 + (10-8.7)^2 + (8-8.7)^2 + (8-8.7)^2 + (9-8.7)^2 = 6.6$

Method B: $(8-8.6)^2 + (9-8.6)^2 + (10-8.6)^2 + (8-8.6)^2 + (8.5-8.6)^2 + (7-8.6)^2 + (9.5-8.6)^2 + (9-8.6)^2 + (7-8.6)^2 + (10-8.6)^2 = 10.9$

Method C: $(9-8.5)^2 + (8-8.5)^2 + (7-8.5)^2 + (10-8.5)^2 + (9.5-8.5)^2 + (8-8.5)^2 + (7-8.5)^2 + (10-8.5)^2 + (9-8.5)^2 + (8-8.5)^2 = 10.5$

Within Group Variation: 6.6+10.9+10.5 = 28

Basics of ANOVA

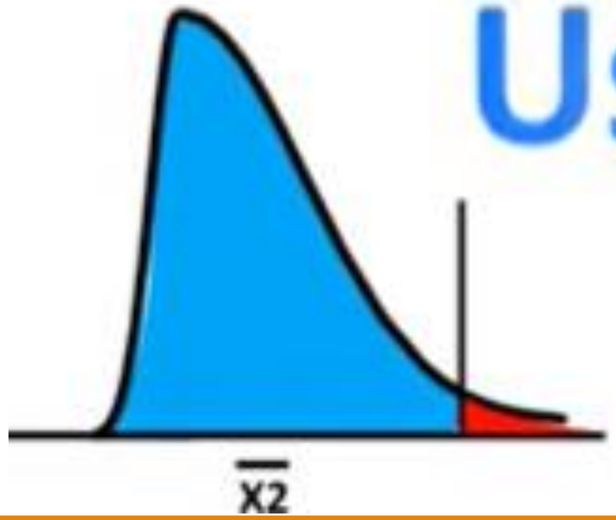
$$\frac{\text{Variance Between}}{\text{Variance Within}} = \frac{0.2}{28} = 0.0071 < 1$$

Fail to Reject H_0

"Means are very close to overall mean and distribution overlap is hard to distinguish".

Basics of ANOVA

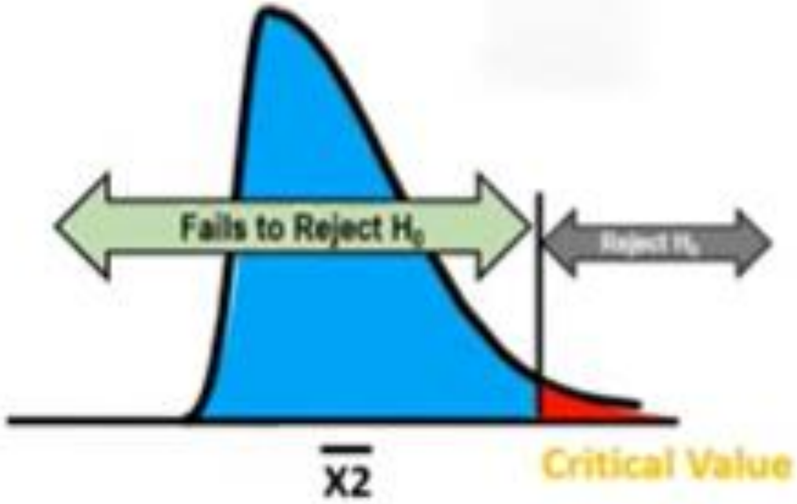
Method 2
Using F-Table



Basics of ANOVA

$F_{Critical} > F_{Stat}$ Fail to Reject H_0

$F_{Critical} < F_{Stat}$ Reject H_0



Assuming $\alpha = 0.05$

$$F_{Stat} = \frac{\text{Variance Between}}{\text{Variance Within}}$$

$$\frac{0.2}{28} = 0.0071$$

$$F_{Critical} = \frac{\text{Numerator Degree of Freedom}}{\text{Denominator Degree of Freedom}}$$

Numerator Degree of Freedom = No. of Samples - 1 = 3-1 = 2

Denominator Degree of Freedom = $\sum(nj-1) = n_T-k = 30 - 3 = 27$

$F_{Critical} = F_{(2,27)} = 3.35$

Most Formal and Accurate Method

Source	SS = Sum of Squares	*df*	MS = Mean Square	F
Between (Factor)	SSB	k – 1	$MSB = \frac{SSB}{k-1}$	$F = \frac{MSB}{MSW}$
Within (Error)	SSW	N – k	$MSW = \frac{SSW}{N-k}$	
Total	SST	N – 1		

➢ One-Way ANOVA

One factor with at least two levels, levels are independent

Researchers want to test a new anti-anxiety medication. They split participants into three conditions (0mg, 50mg, and 100mg), then ask them to rate their anxiety level on a scale of 1-10. Are there any differences between the three conditions using alpha = 0.05?

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

➢ One-Way ANOVA

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

1. **Define Null and Alternative Hypotheses**
2. **State Alpha**
3. **Calculate Degrees of Freedom**
4. **State Decision Rule**
5. **Calculate Test Statistic**
6. **State Results**
7. **State Conclusion**

➢ One-Way ANOVA

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

1. Define Null and Alternative Hypotheses

$$H_0; \mu_{0mg} = \mu_{50mg} = \mu_{100mg}$$

$$H_1; not\ all\ \mu's\ are\ equal$$

➢ One-Way ANOVA

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

3. **Calculate Degrees of Freedom**

Example

Group 1	Group 2	Group 3
3	10	7
2	9	6
1	9	7
1	8	6
4	7	5
2	8	4
4	6	3
3	7	6

Mean values

$$G = \frac{\sum x}{N}$$

$$G = \frac{3 + 10 + \ldots + 7 + 6}{24} = 5.3$$

$$M_i = \frac{\sum^{Group} x}{n_{Group}}$$

$$M_1 = \frac{3 + 2 + \ldots + 4 + 3}{8} = 2.5$$

$$M_2 = \frac{10 + 9 + \ldots + 6 + 7}{8} = 8$$

$$M_3 = \frac{7 + 6 + \ldots + 3 + 6}{8} = 5.5$$

Sum of squares

$$SS_{btw} = \sum^{Groups} n_i (M_i - G)^2$$
$$= 8 \cdot ((2.5 - 5.3)^2 + (8 - 5.3)^2 + (5.5 - 5.3)^2)$$
$$= 121.3$$

$$SS_{wi} = \sum^{Groups} \sum^{Group} (x_{mi} - M_i)^2$$
$$= (3 - 2.5)^2 + (2 - 2.5)^2 + \ldots$$
$$+ (10 - 8)^2 + (9 - 8)^2 + \ldots$$
$$+ (7 - 5.5)^2 + (6 - 5.5)^2 + \ldots$$
$$= 36$$

Mean squares

$$MS_{btw} = \frac{SS_{btw}}{df_{btw}} = \frac{121.33}{2} = 60.67$$

$$MS_{wi} = \frac{SS_{wi}}{df_{wi}} = \frac{36}{21} = 1.71$$

F-value

$$F = \frac{MS_{btw}}{MS_{wi}} = \frac{60.67}{1.71} = 35.39$$

Example for One way ANOVA

Low Noise			Medium Noise			Loud Noise		
Student	Questions (X)		Student	Questions (X)		Student	Questions (X)	
1	10		5	8		9	4	
2	9		6	4		10	3	
3	6		7	6		11	6	
4	7		8	7		12	4	

Data for one way ANOVA

Example for Two way ANOVA

Students	Low Noise	Medium Noise	Loud Noise	
Male Students	10 12 11 9	7 9 8 12	4 5 6 5	
Female Students	12 13 10 13	13 15 12 12	6 6 4 4	

Does Noise has an effect on the marks a student scores
Does Gender has an effect on the marks a student scores
Does Gender effects how a students reacts to Noise

Two way ANOVA

A two-way ANOVA is a **statistical method** used to test the effect of **two categorical variables** on a **continuous variable**.

Two way ANOVA

A two-way ANOVA is a **statistical method** used to test the effect of **two categorical variables** on a **continuous variable.**

The categorical variables are the independent variables

for example, the variable drug type with drug A and B

and gender with female and male.

Two way ANOVA

A two-way ANOVA is a **statistical method**

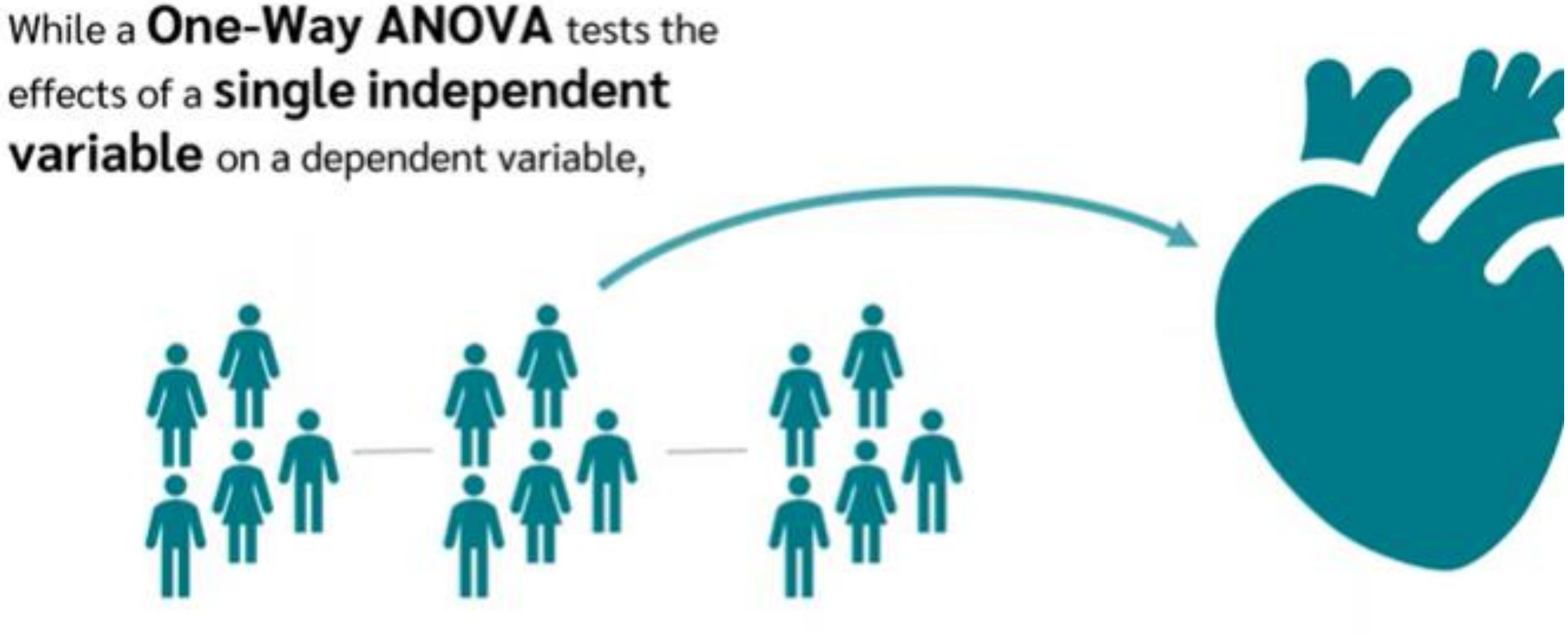
used to test the effect of **two categorical variables**

on a **continuous variable**.

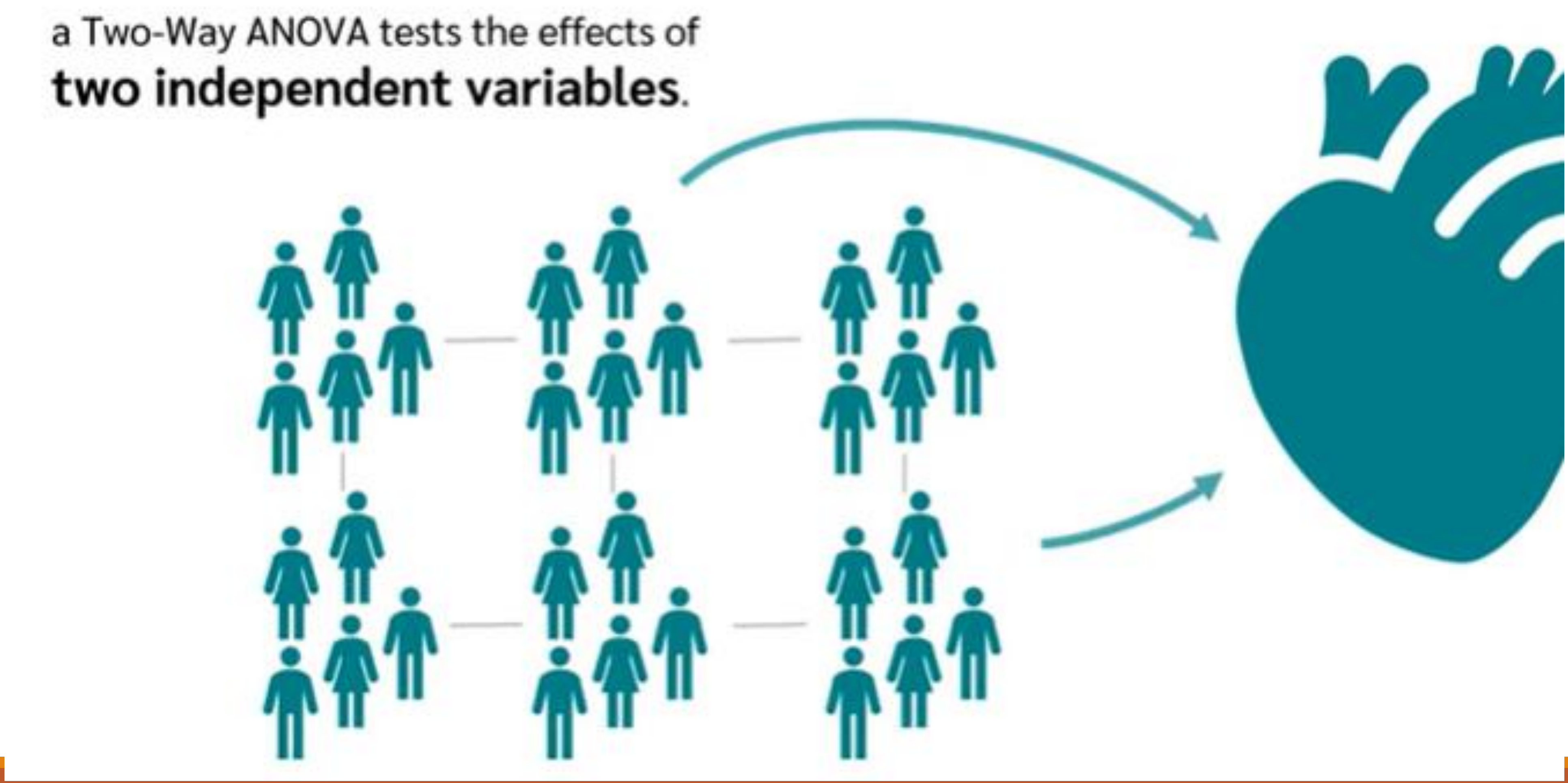
And the continuous variable is the dependent variable

for example, the reduction in blood pressure.

Two way ANOVA

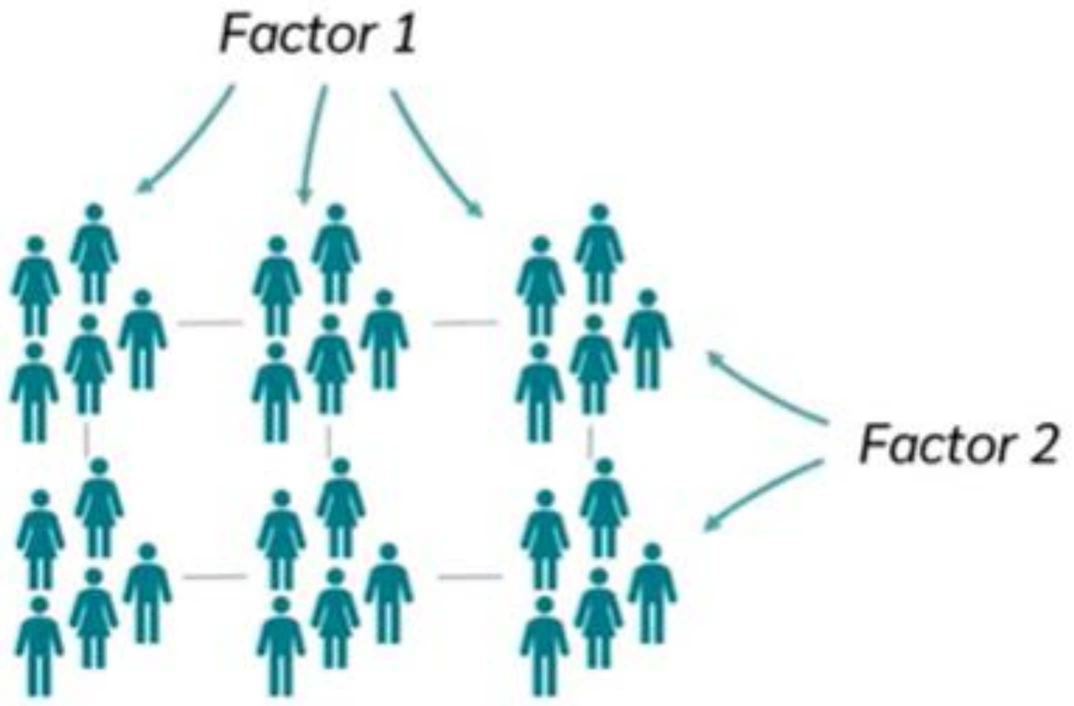


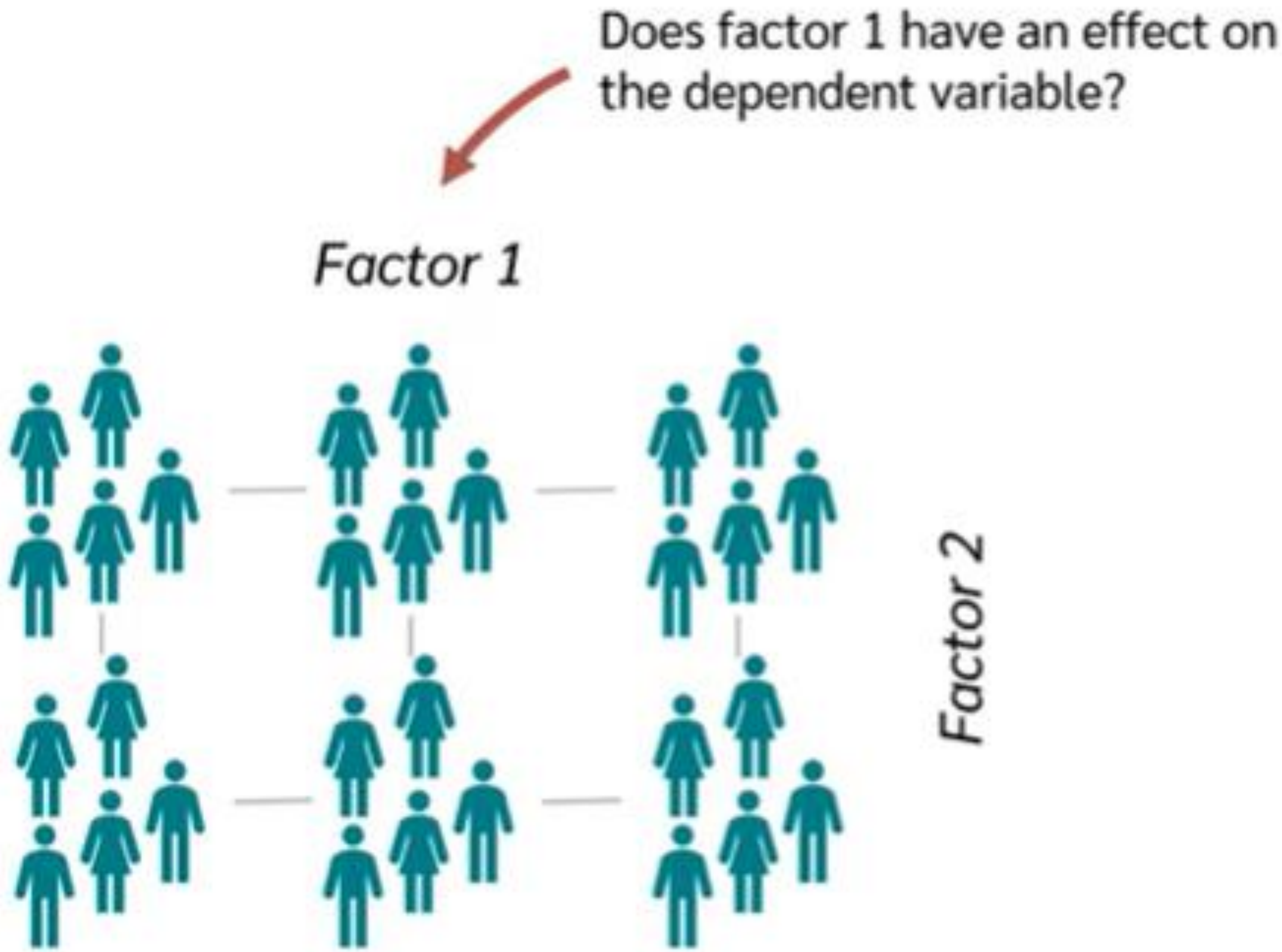
Two way ANOVA

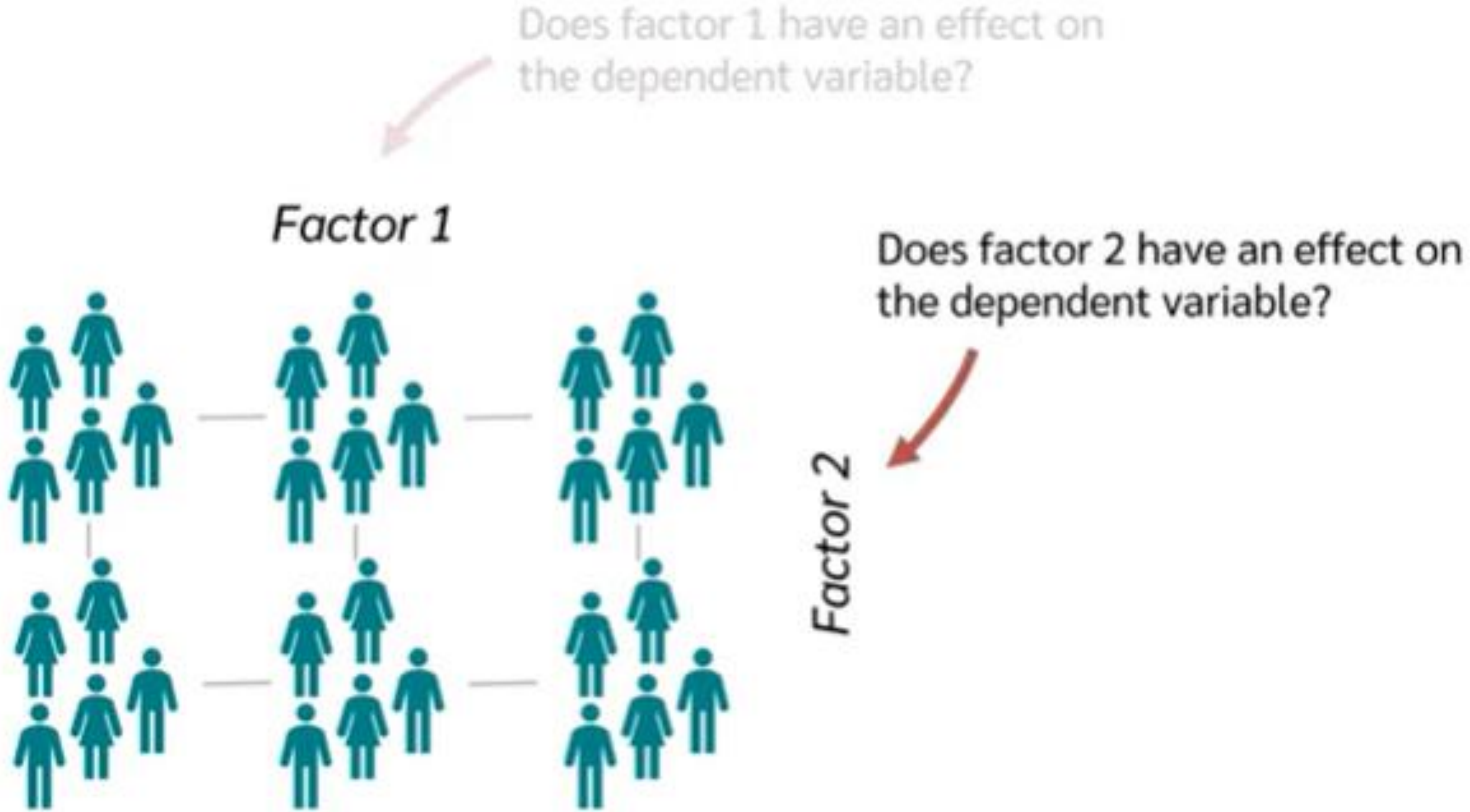


Two way ANOVA

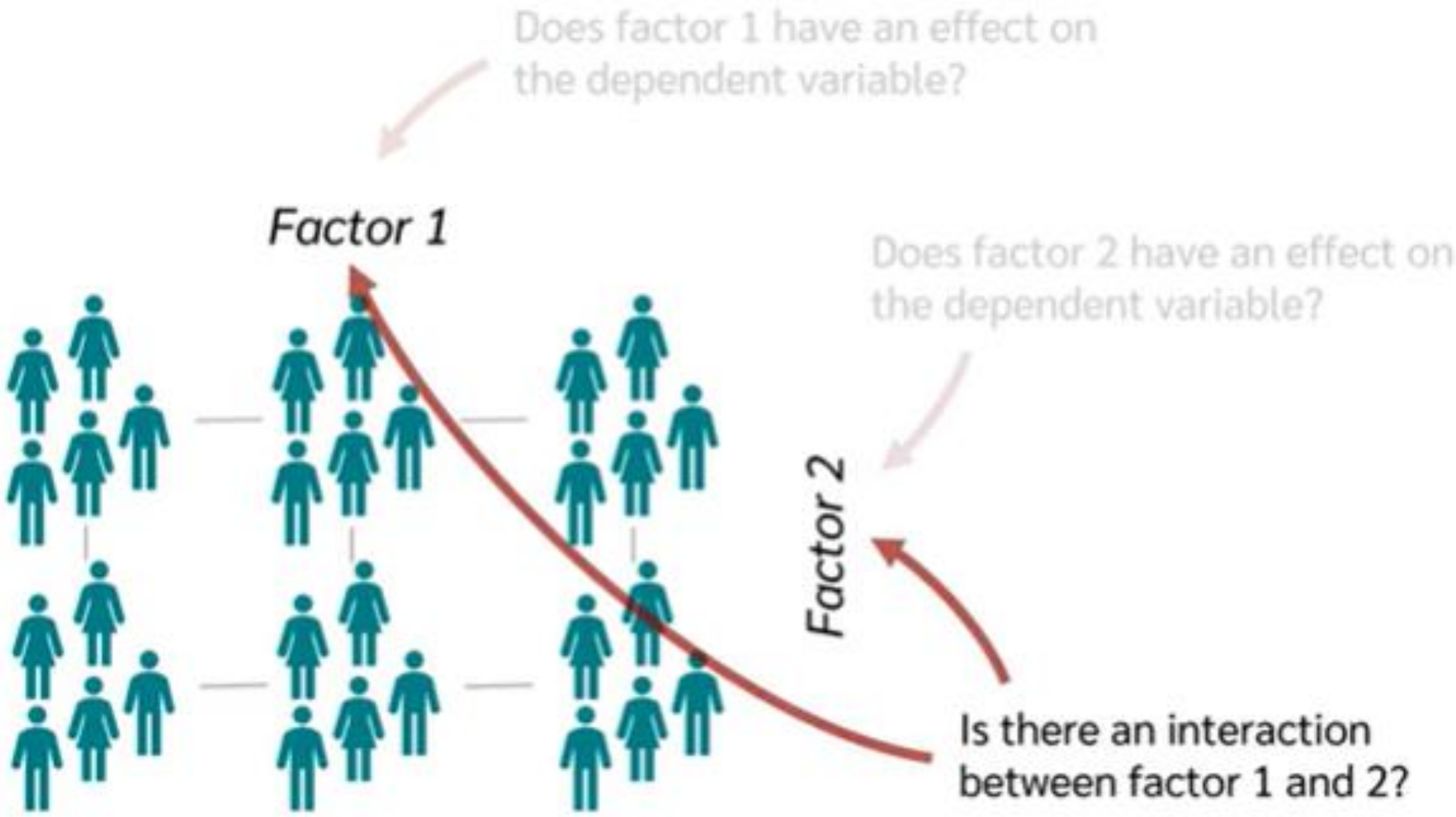
The independent variables are called **factors**.




Does factor 1 have an effect on
the dependent variable?
Factor 1
Factor 2


Does factor 1 have an effect on the dependent variable?
Factor 1
Does factor 2 have an effect on the dependent variable?
Factor 2

Two way ANOVA



Two way ANOVA

Null hypothesis H0:

There is no significant difference between the groups of the first factor.

There is no significant difference between the groups of the second factor.

One factor has no effect on the effect of the other factor.

Alternative hypothesis H1:

There is a significant difference between the groups of the first factor.

There is a significant difference between the groups of the second factor.

One factor has an influence on the effect of the other factor.

Two way ANOVA

We randomly assigned patients to the **treatment combinations** and **measured** their **reduction in blood pressure** after a month.

Drug	Gender	Reduc. BP
A	male	6
A	male	4
A	male	5
A	female	3
A	female	4
A	female	3
B	male	5
B	male	9
B	male	2
...	...	...

Two way ANOVA

Now let us answer the questions:

Is there a main effect of **drug type** on the reduction in **blood pressure**?

Drug	Gender	Reduc. BP
A	male	6
A	male	4
A	male	5
A	female	3
A	female	4
A	female	3
B	male	5
B	male	9
B	male	2
...	...	...

Two way ANOVA

Now let us answer the questions:

Is there a main effect of **gender** on the reduction in **blood pressure**?

Drug	Gender	Reduc. BP
A	male	6
A	male	4
A	male	5
A	female	3
A	female	4
A	female	3
B	male	5
B	male	9
B	male	2
...	...	...

Two way ANOVA

Now let us answer the questions:

Is there an interaction effect between **drug type** and **gender** on the reduction in **blood pressure**?

X

Drug	Gender	Reduc. BP
A	male	6
A	male	4
A	male	5
A	female	3
A	female	4
A	female	3
B	male	5
B	male	9
B	male	2
...	...	...

Calculate mean values

	Drug A	Drug B	
Male	6	4	
	4	5	
	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
Female	8	3	
	3	5	
	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

Sum Of Squares

$$SS_{tot} = SS_A + SS_B + SS_{AB} + SS_{err}$$

Total variance of the dependent variable

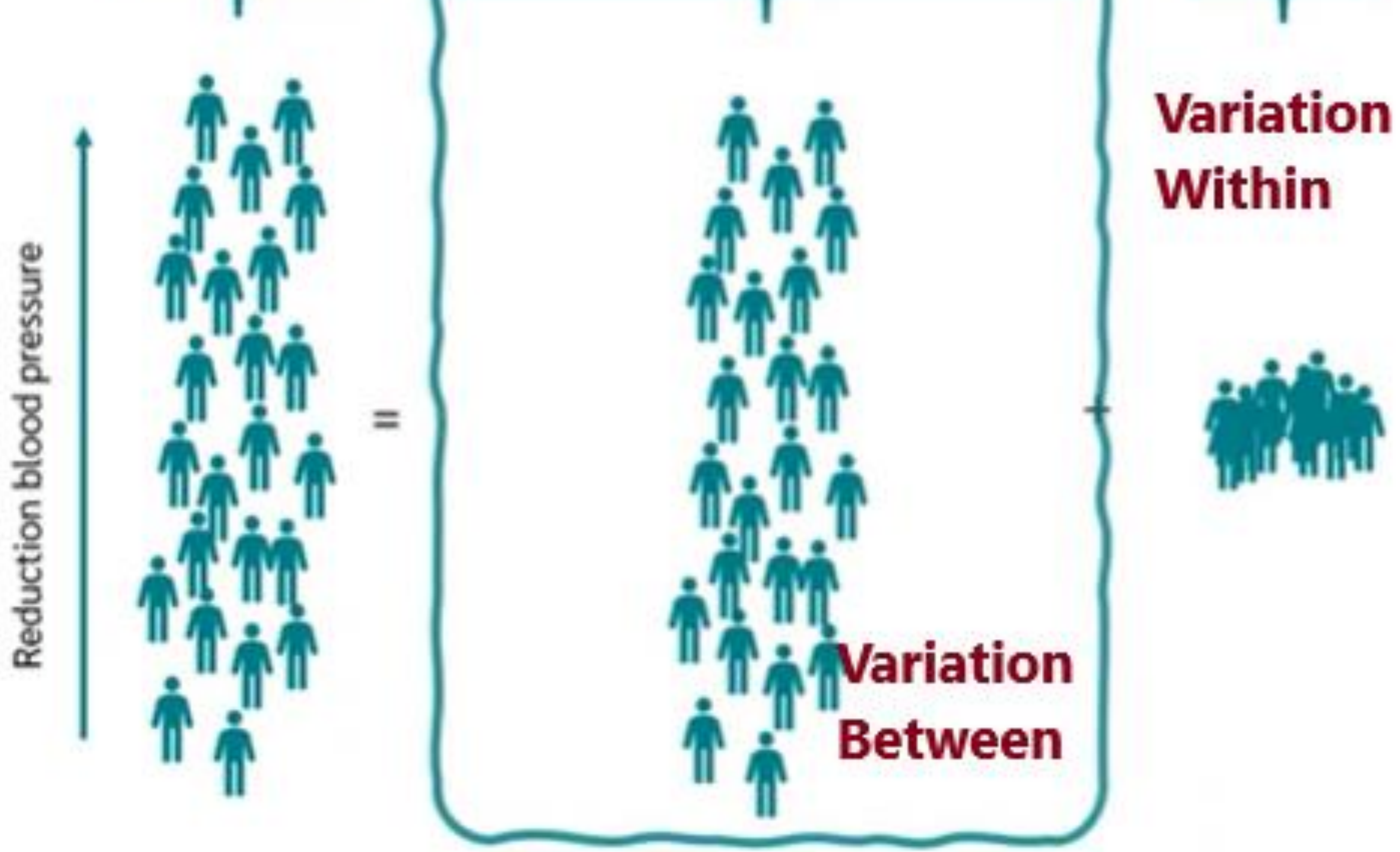
Variance that can be explained by Factor A,

Variance that can be explained by Factor B,

Variance that can be explained by the interaction of A and B,

the error variance.

$$\underbrace{SS_{tot}} = \underbrace{\boxed{SS_A + SS_B + SS_{AB}}} + \underbrace{SS_{err}}$$



Mean value of each group

Total mean

$$SS_{btw} = n \cdot \sum\sum\left(\overline{AB}_{ij} - \bar{G}\right)^2$$
$$= 5 \cdot \left((5.8 - 5.4)^2 + (5.4 - 5.4)^2 + \cdots + (4.4 - 5.4)^2\right)$$
$$= 7.6$$

$$df_{btw} = p \cdot q - 1$$
$$= 2 \cdot 2 - 1$$
$$= 3$$

$$\sigma^2_{btw} = \frac{SS_{btw}}{df_{btw}} = \frac{7.6}{3} = 2.53$$

	Drug A	Drug B	
Male	6	4	
	4	5	
	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
Female	8	3	
	3	5	
	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

$$SS_{tot} = \sum\sum\sum (x_{mij} - \bar{G})^2$$
$$= (6 - 5.4)^2 + (4 - 5.4)^2 + \cdots + (3 - 5.4)^2$$
$$= 84.8$$

Number of people per group

Number of groups Factor A

Number of groups Factor B

$$df_{tot} = n \cdot p \cdot q - 1$$
$$= 5 \cdot 2 \cdot 2 - 1$$
$$= 19$$

$$\sigma^2_{tot} = \frac{SS_{tot}}{\boxed{df_{tot}}} = \frac{84.8}{19} = 4.46$$

	Drug A	Drug B	
Male	6	4	
	4	5	
	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
Female	8	3	
	3	5	
	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

Mean value of the groups of factor A

$$SS_A = n \cdot q \sum (\bar{A}_i - \bar{G})^2$$
$$= 5 \cdot 2\left((5.9 - 5.4)^2 + (4.9 - 5.4)^2\right)$$
$$= 5$$

$$df_A = p - 1$$
$$df_A = 2 - 1$$
$$= 1$$

$$\sigma_A^2 = \frac{SS_A}{df_A} = \frac{5}{1} = \boxed{5}$$

	Drug A	Drug B	
Male	6	4	
	4	5	
	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
Female	8	3	
	3	5	
	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

Mean value of the groups of factor B

$$SS_B = n \cdot p \sum (\bar{B}_i - \bar{G})^2$$
$$= 5 \cdot 2\left((5.6 - 5.4)^2 + (5.2 - 5.4)^2\right)$$
$$= 0{,}8$$

$$df_{\mathrm{B}} = \mathrm{q} - 1$$
$$df_{\mathrm{B}} = 2 - 1$$
$$= 1$$

$$\sigma_B^2 = \frac{SS_B}{df_B} = \frac{0.8}{1} = 0.8$$

	Drug A	Drug B	
	6	4	
	4	5	
Male	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
	8	3	
	3	5	
Female	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

$$SS_{AB} = SS_{btw} - SS_A - SS_B$$

$$= 7.6 - 5 - 0.8$$

$$= 1.8$$

$$df_{AB} = (p-1)\cdot(q-1)$$

$$df_{AB} = 1 \cdot 1$$

$$= 1$$

$$\sigma^2_{AB} = \frac{SS_{AB}}{df_{AB}} = \frac{1.8}{1} = \boxed{1.8}$$

	Drug A	Drug B	
Male	6	4	
	4	5	
	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
Female	8	3	
	3	5	
	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

Variation Within

Mean value of the groups

Individual values

$$SS_{err} = \sum\sum\sum(x_{mij} - \overline{AB}_{ij})^2$$
$$= (6 - 5.8)^2 + (4 - 5.4)^2 + \cdots + (3 - 4.4)^2$$
$$= 77{,}2$$

$$df_{err} = (n-1)\cdot p \cdot q$$
$$= 4 \cdot 2 \cdot 2$$
$$= 16$$

$$\sigma^2_{err} = \frac{SS_{err}}{df_{err}} = \frac{77.2}{16} = 4.83$$

	Drug A	Drug B	
Male	6	4	
	4	5	
	7	6	
	9	7	
	3	5	
Mean	5.8	5.4	5.6
Female	8	3	
	3	5	
	5	9	
	8	2	
	6	3	
Mean	6	4.4	5.2
	5.9	4.9	5.4

Fstat values and Fcritical Values

$$F_A = \frac{\boxed{\sigma_A^2}}{\sigma_{err}^2} = \frac{5}{4.83} = 1.04 \qquad F_B = \frac{\sigma_B^2}{\sigma_{err}^2} = \frac{0.8}{4.83} = 0.17 \qquad F_{AB} = \frac{\sigma_{AB}^2}{\sigma_{err}^2} = \frac{1.8}{4.83} = 0.37$$

FAcrit=F(dfA,dferror)
FBcrit=F(dfA,dferror)
FABcrit=F(dfAB,dferror)

Students	Low Noise	Medium Noise	Loud Noise	
Male Students	10 12 11 9	7 9 8 12	4 5 6 5	
Female Students	12 13 10 13	13 15 12 12	6 6 4 4	

Does Noise has an effect on the marks a student scores
Does Gender has an effect on the marks a student scores
Does Gender effects how a students reacts to Noise

Students	Low Noise	Medium Noise	Loud Noise	Row Total
Male Students	10 12 11 9	7 9 8 12	4 5 6 5	$R_1 =$ 98
Female Students	12 13 10 13	13 15 12 12	6 6 4 4	$R_2 =$ 120
Column Total	C_1 =90	C_2 =88	C_3 =40	

Source	Degrees of freedom (d.f.)	Sum of Squares(SS)	Mean of Sum of Squares $MSS_x = \frac{SS}{df}$	F Ratio $F = \frac{MSS_x}{MSS_E}$
Noise	$(C-1) = 2$	200		
Gender	$(R-1) = 1$	20		
Interaction	$(C-1) \times (R-1) = 2$	16.33		
Residual	$C \times R \times (n-1) = 18$	37		
Total	$(N-1) = 23$	59.25		

$C - No.\,of\ columns\ (Noise\ Categories) = 3$
$R - No.\,of\ rows\ (Student\ Categories) = 2$
$N - Total\ number\ of\ students = 24$
$n - No.\,of\ students\ in\ a\ group = 4$

Source	Degrees of freedom (d.f.)	Sum of Squares(SS)	Mean of Sum of Squares $MSS_x = \frac{SS}{df}$	F Ratio $F = \frac{MSS_x}{MSS_E}$
Noise	$(C - 1) = 2$	200	$MSS_C = 100$	
Gender	$(R - 1) = 1$	20	$MSS_R = 20$	
Interaction	$(C - 1) \times (R - 1) = 2$	16.33	$MSS_g = 8.167$	
Residual	$C \times R \times (n - 1) = 18$	37	$MSS_E = 2.06$	
Total	$(N - 1) = 23$	59.25		

$C - No. of\ columns\ (Noise\ Categories) = 3$
$R - No. of\ rows\ (Student\ Categories) = 2$
$N - Total\ number\ of\ students = 24$
$n - No. of\ students\ in\ a\ group = 4$

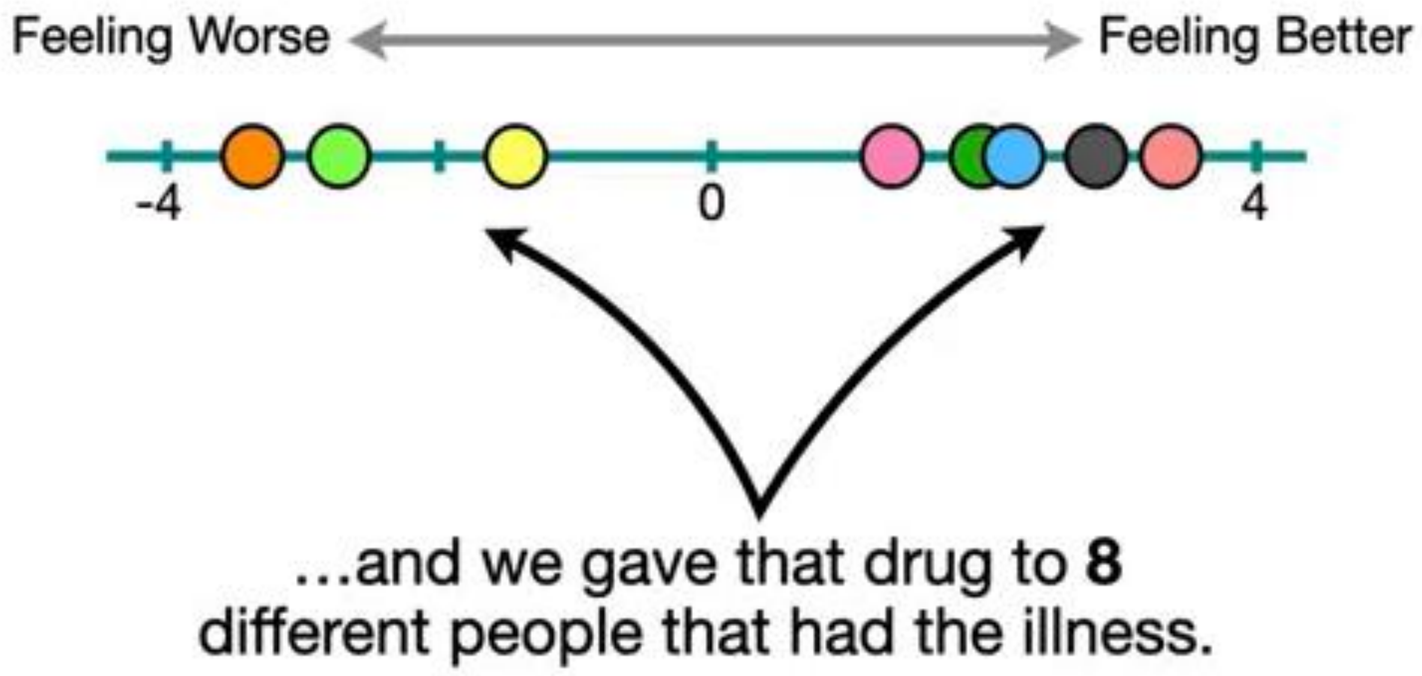
Vectors

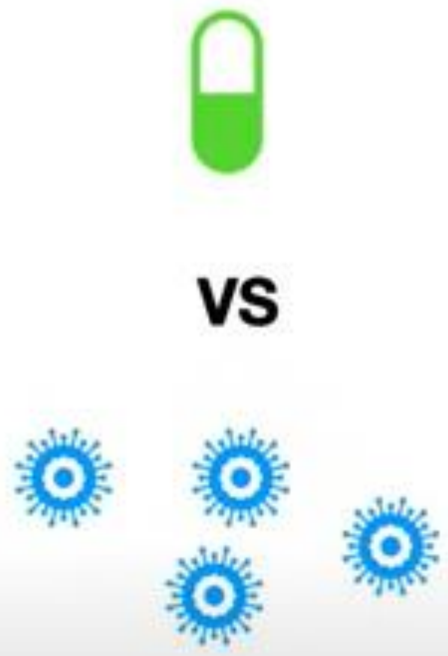
Source	Degrees of freedom (d.f.)	Sum of Squares(SS)	Mean of Sum of Squares $MSS_x = \frac{SS}{df}$	F Ratio $F = \frac{MSS_x}{MSS_E}$
Noise	$(C-1)=2$	200	$MSS_C = 100$	**48.73**
Gender	$(R-1)=1$	20	$MSS_R = 20$	**9.81**
Interaction	$(C-1)\times(R-1)=2$	16.33	$MSS_g = 8.167$	**3.97**
Residual	$C\times R\times(n-1)=18$	37	$MSS_E = 2.06$	
Total	$(N-1)=23$	59.25		

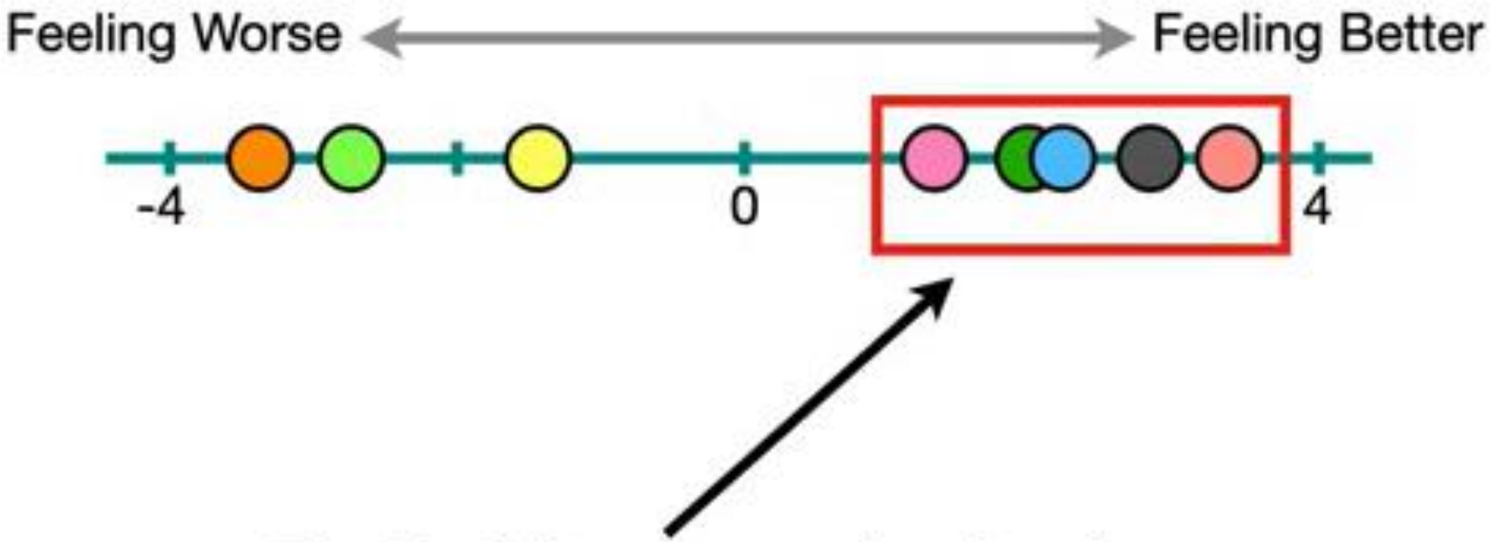
$C - No.of\ columns\ (Noise\ Categories) = 3$
$R - No.of\ rows\ (Student\ Categories) = 2$
$N - Total\ number\ of\ students = 24$
$n - No.of\ students\ in\ a\ group = 4$

Source	Degrees of freedom (d.f.)	F Ratio $F = \frac{MSS_x}{MSS_E}$	$F_{critical}$ at $\alpha = 0.05$
Noise	$(C-1) = 2$	48.73	$F_{(2,18)} = 3.55$
Gender	$(R-1) = 1$	9.81	$F_{(1,18)} = 4.41$
Interaction	$(C-1) \times (R-1) = 2$	3.97	$F_{(2,18)} = 3.55$
Residual	$C \times R \times (n-1) = 18$		
Total	$(N-1) = 23$		

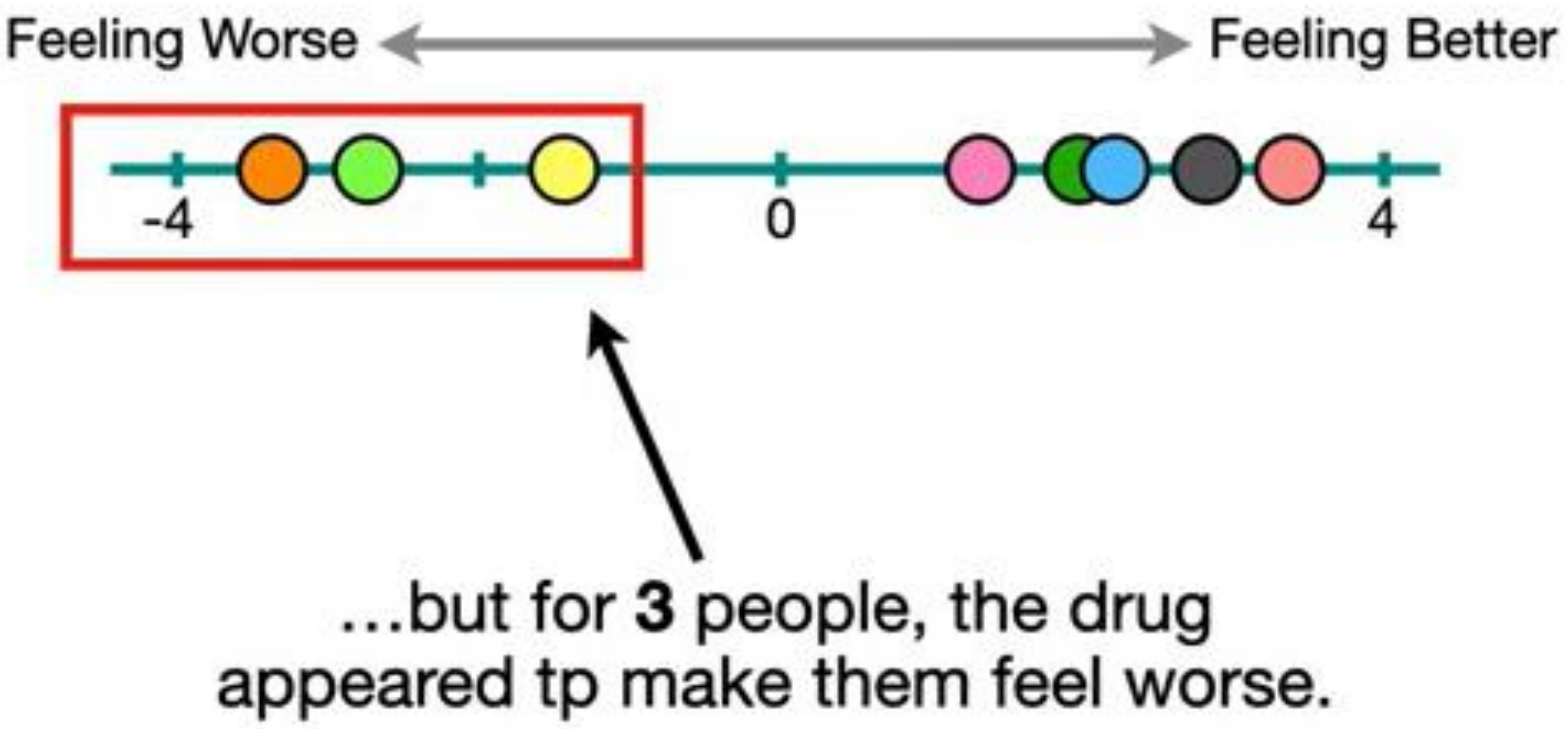
Bootstrapping

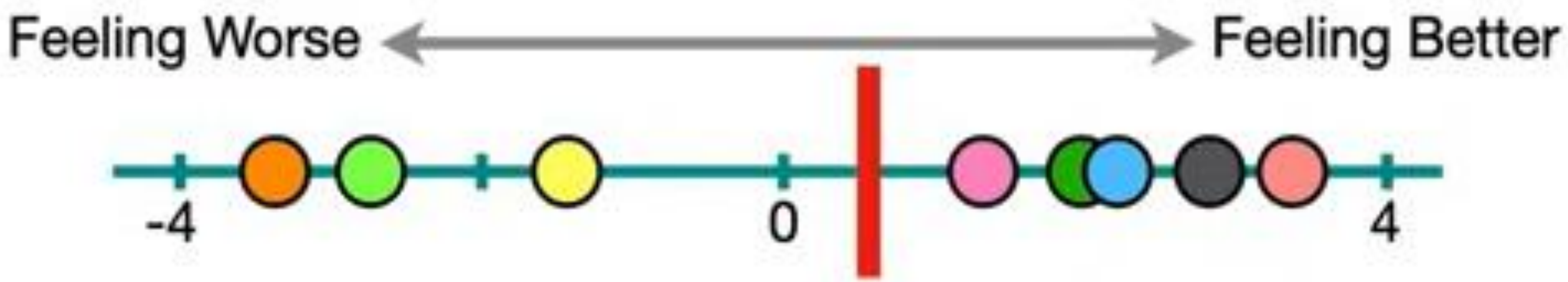

Feeling Worse
Feeling Better
-4
0
4
...and we gave that drug to 8
different people that had the illness.


VS

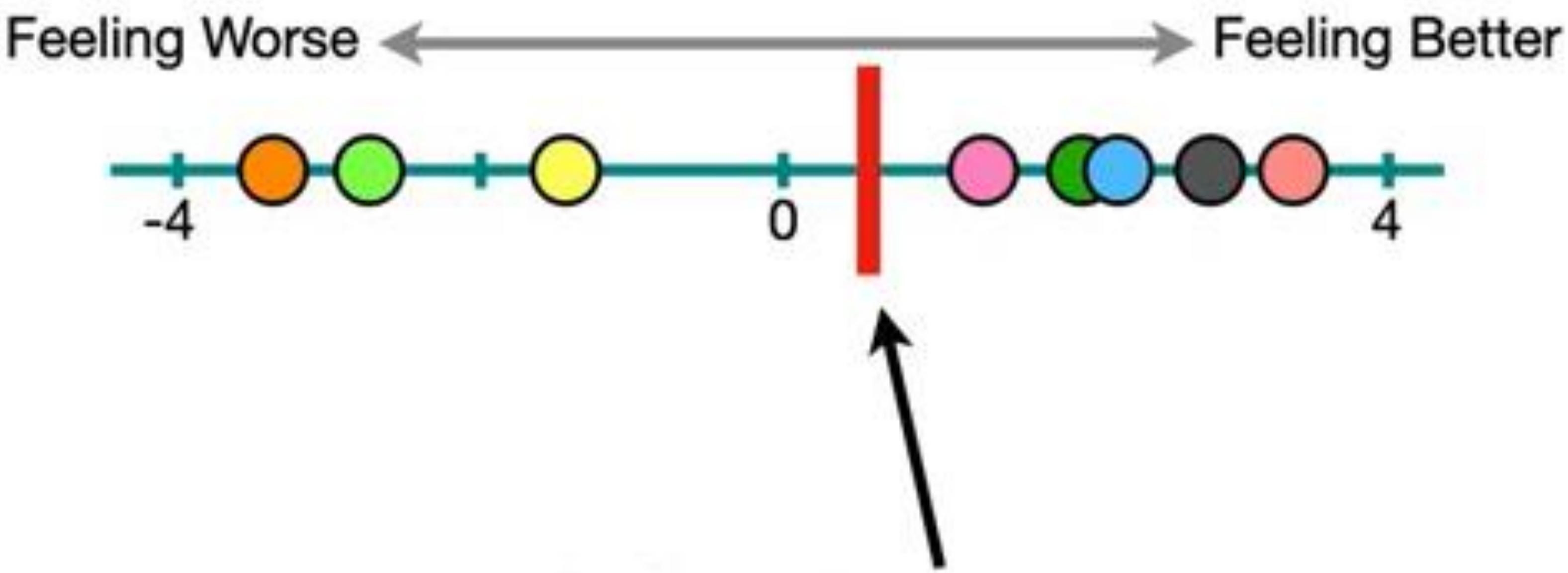


For **5** of those people, the drug appeared to help them feel better…

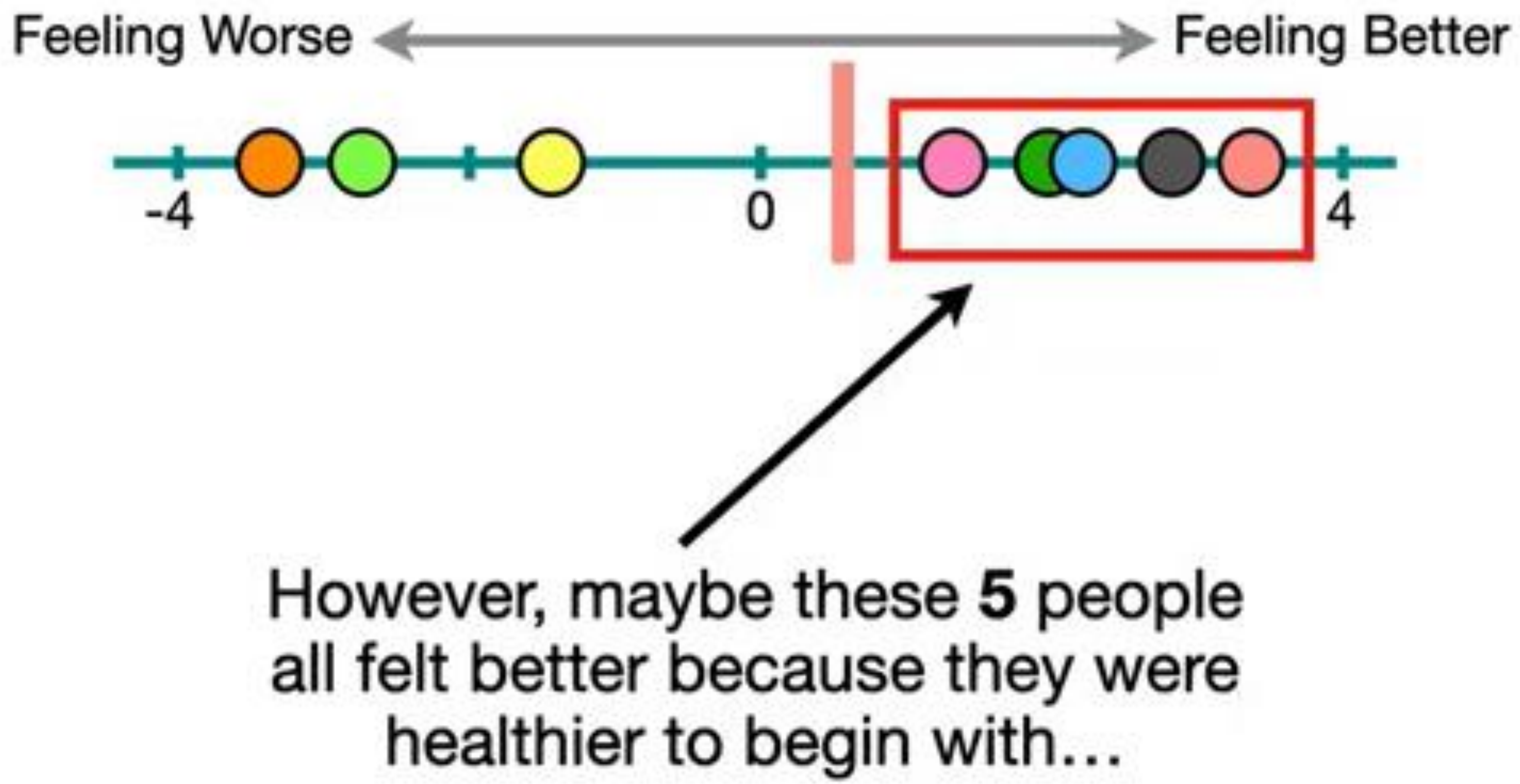

Feeling Worse
Feeling Better
-4
0
4
...but for **3** people, the drug
appeared tp make them feel worse.

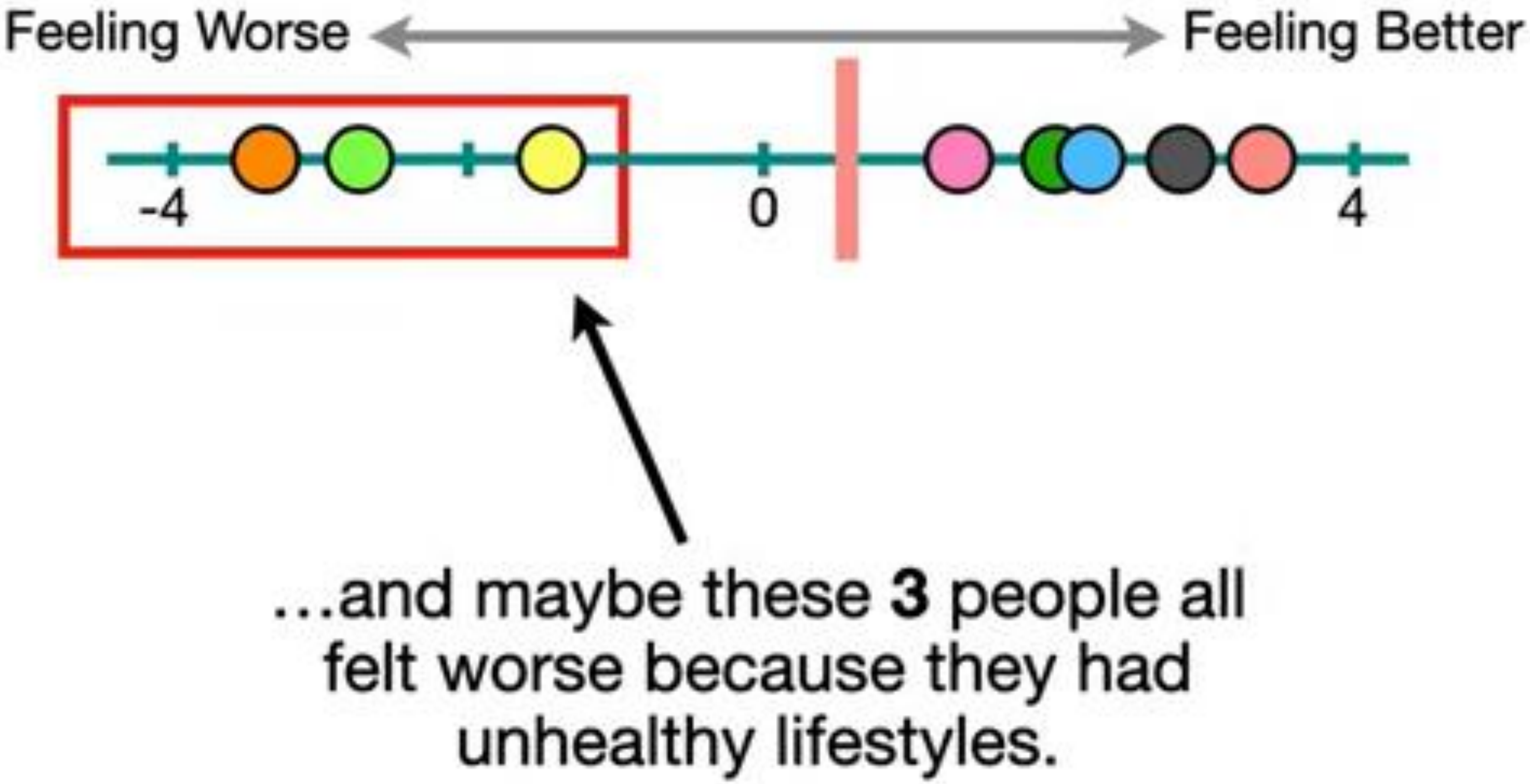


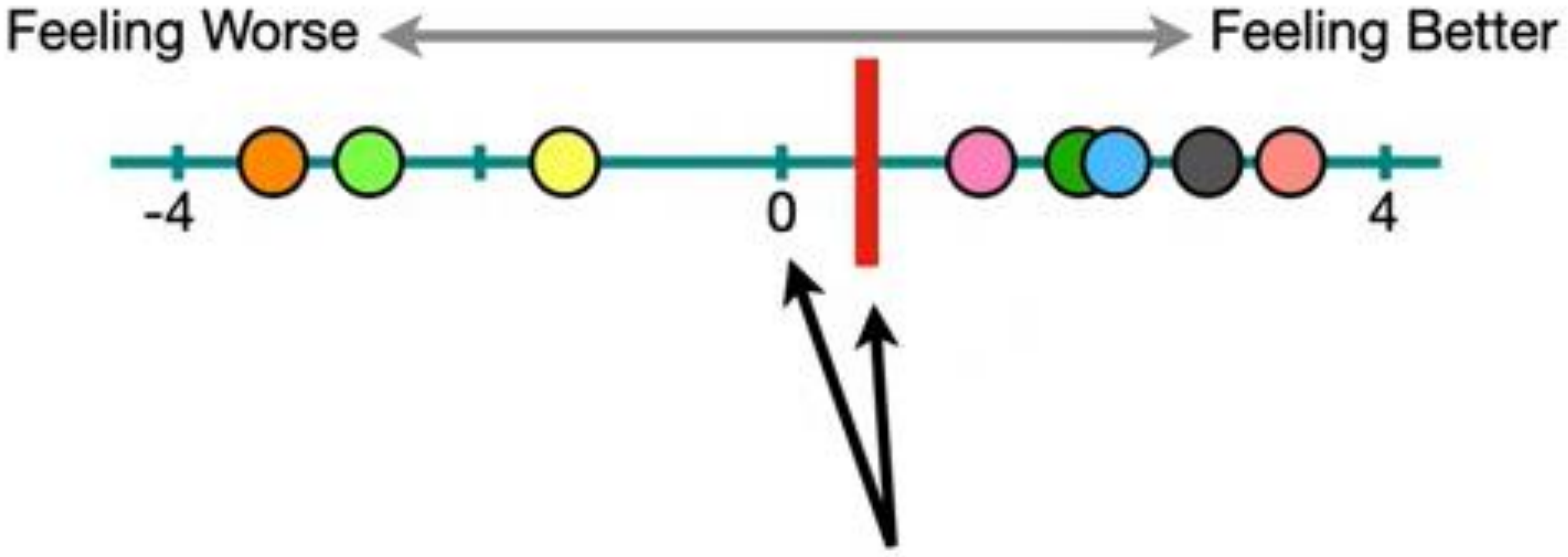
If we calculate the mean of the response to the drug we get **0.5**.



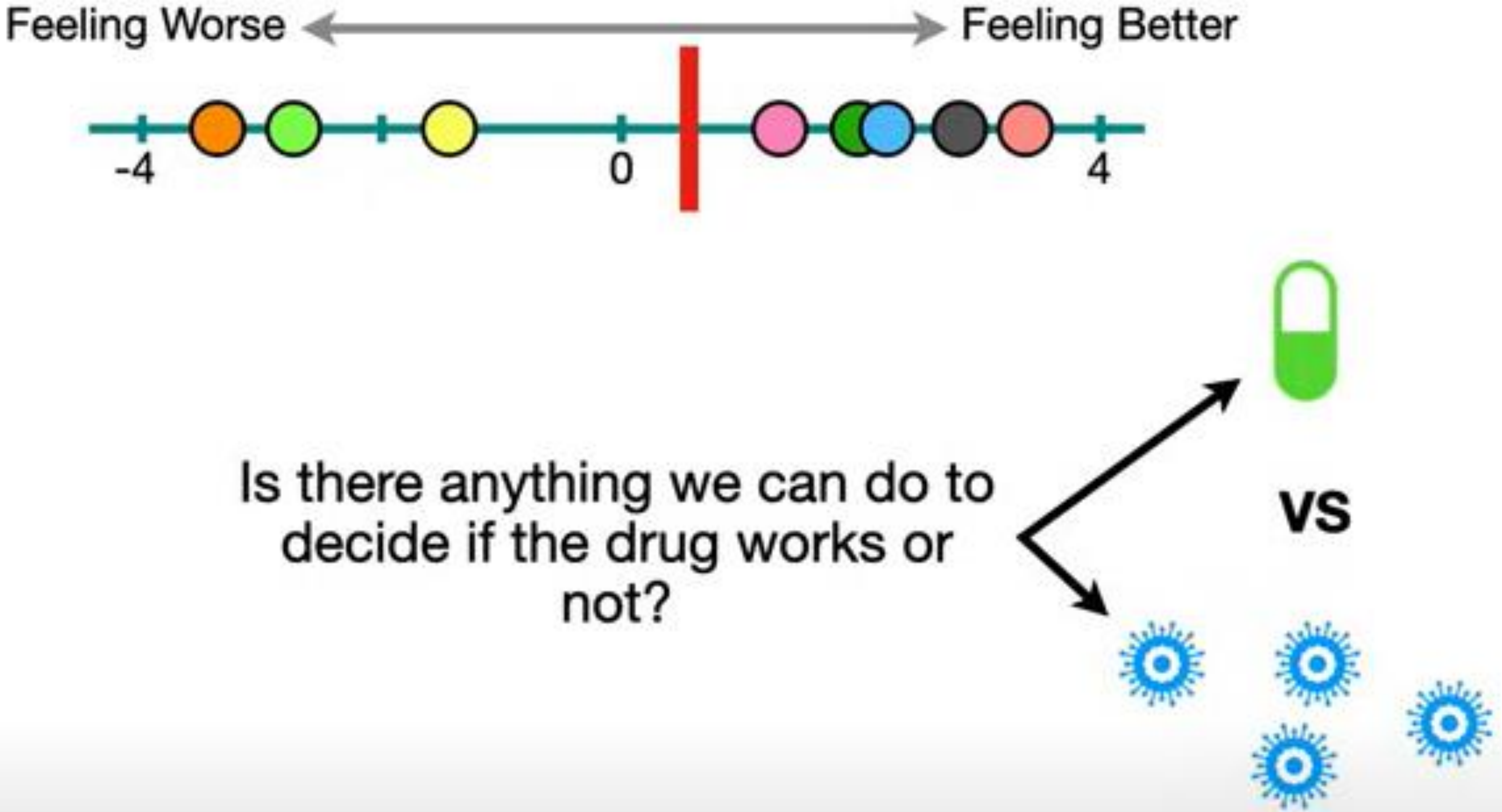
0.5 is not a huge improvement, but, since most of the people, **5** of **8**, improved, maybe this drug is better than using no drug at all.

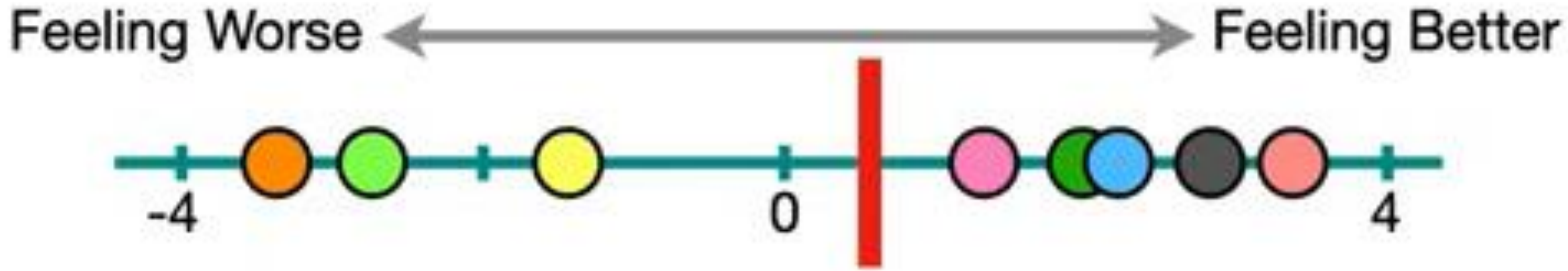

Feeling Worse
Feeling Better
-4
0
4
However, maybe these **5** people all felt better because they were healthier to begin with…


Feeling Worse
Feeling Better
-4
0
4
…and maybe these **3** people all felt worse because they had unhealthy lifestyles.



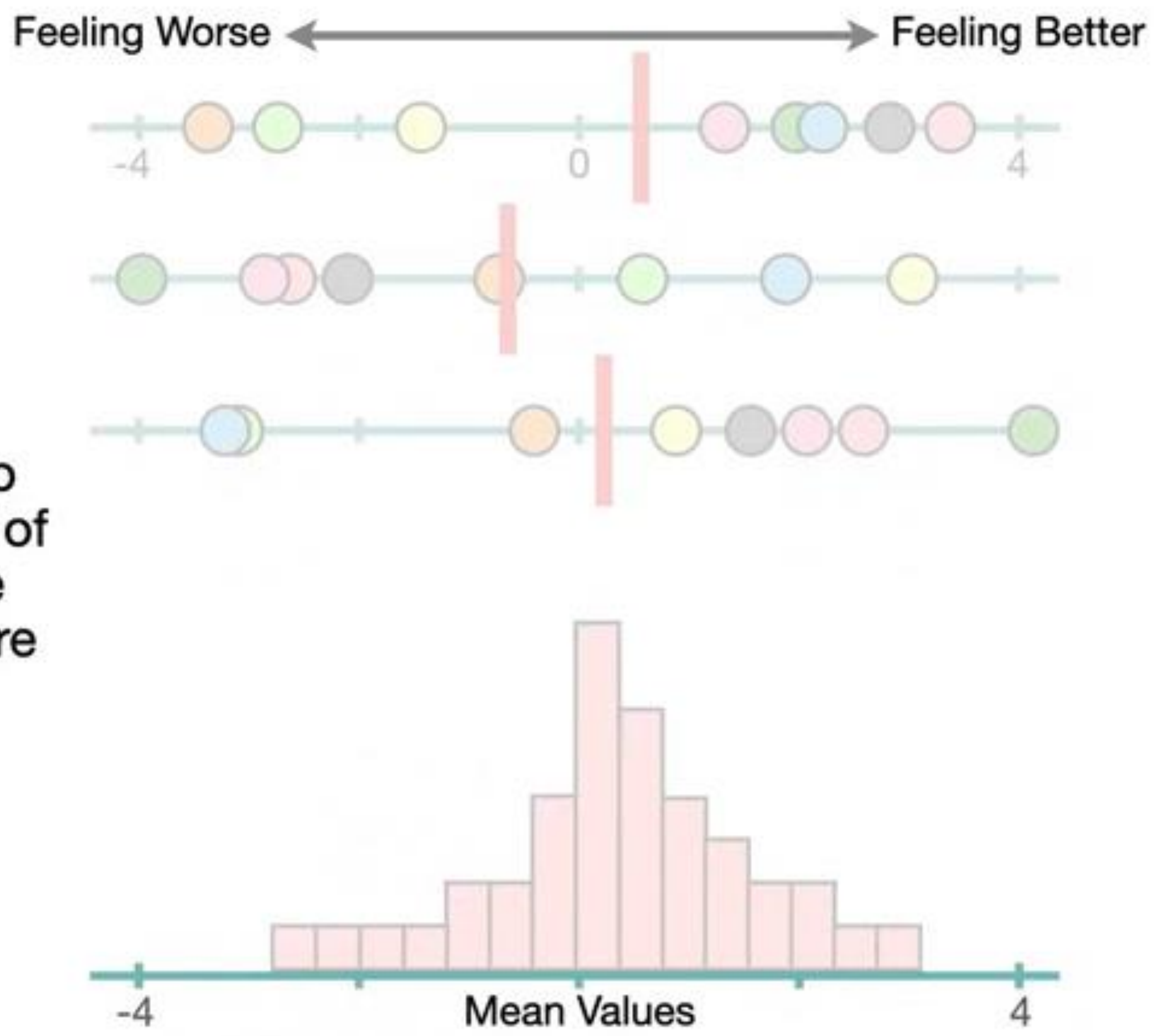
So it is possible that the reason we got a mean value = **0.5** instead of **0** is because of random things that we can't control.

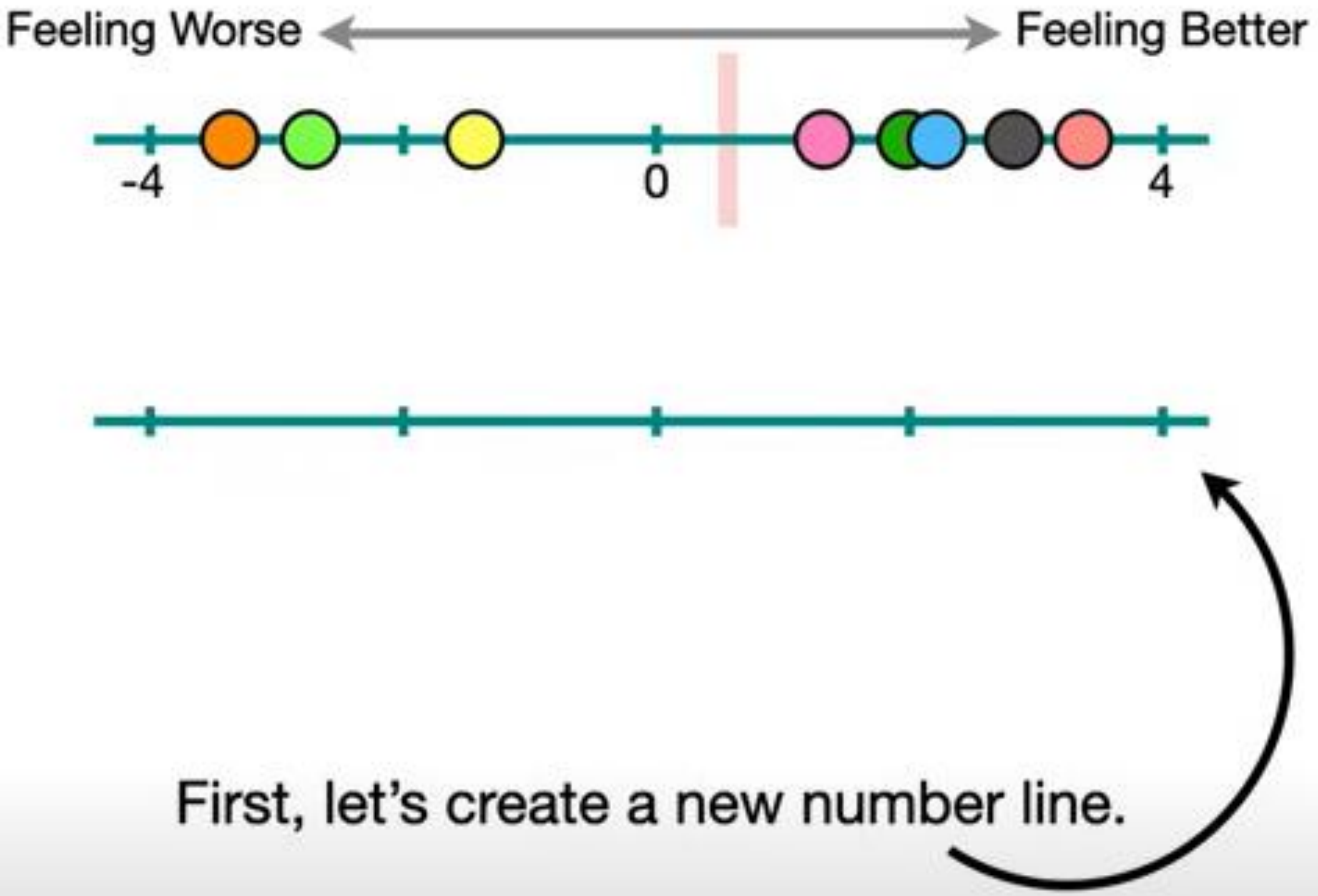

Feeling Worse
Feeling Better
-4
0
4
Is there anything we can do to decide if the drug works or not?
vs



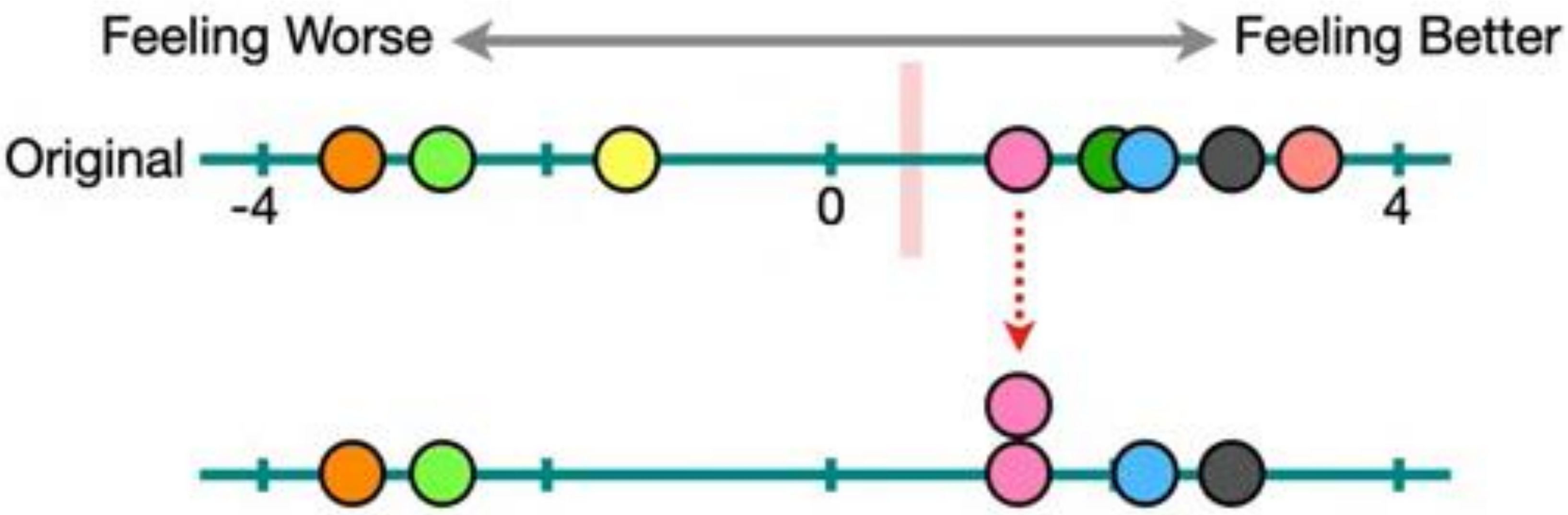
One ***expensive*** and ***time consuming*** option would be to replicate the experiment a bunch of times.

So let's use **Bootstrapping** to get a better sense of which results are likely and which are rare.

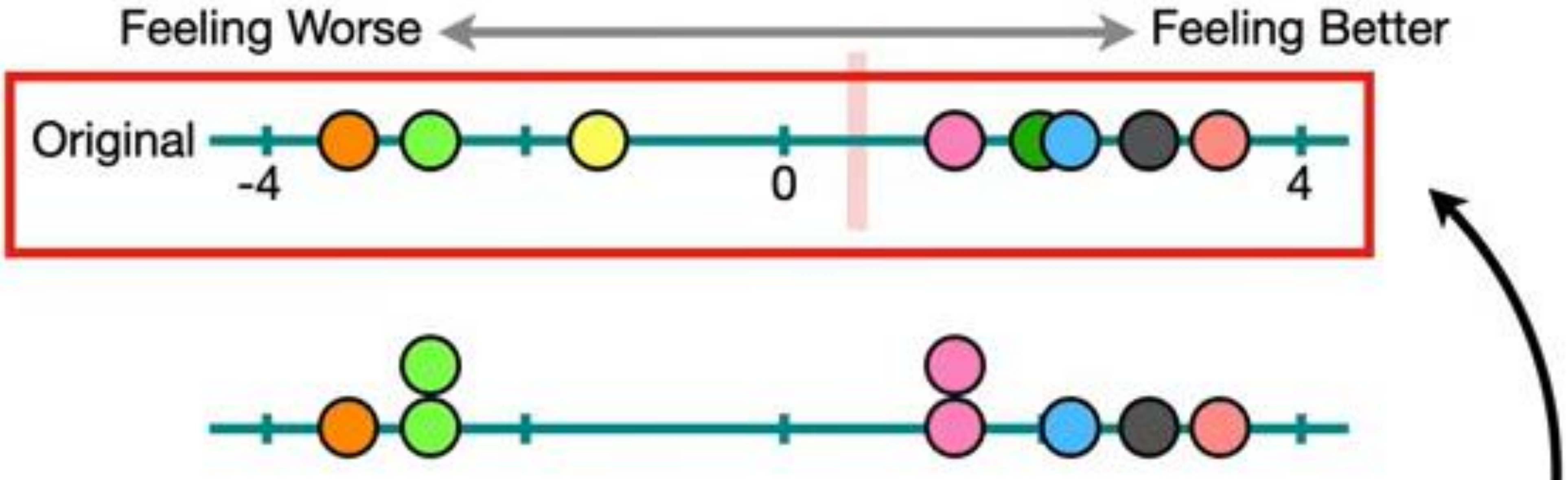




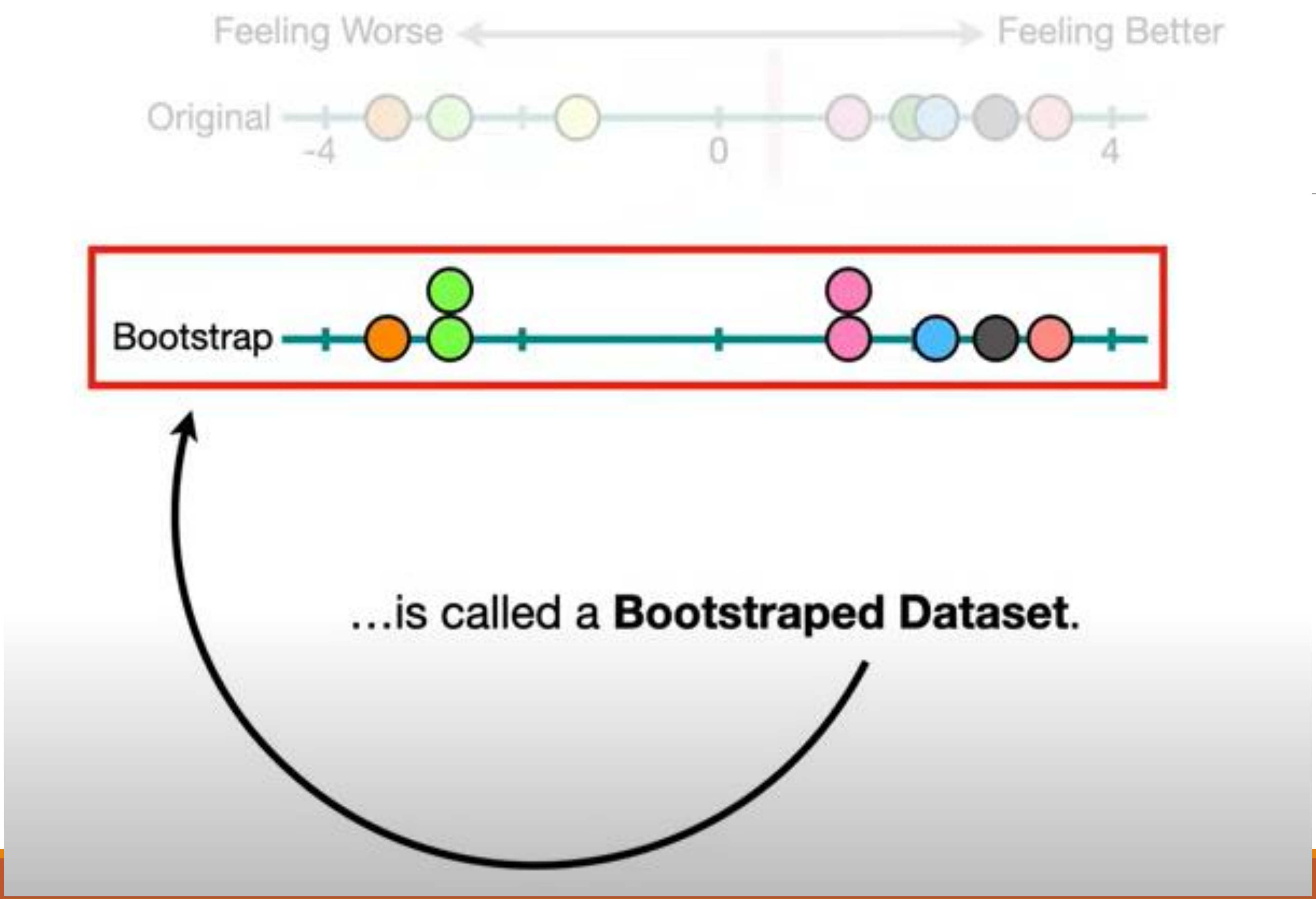
First, let's create a new number line.

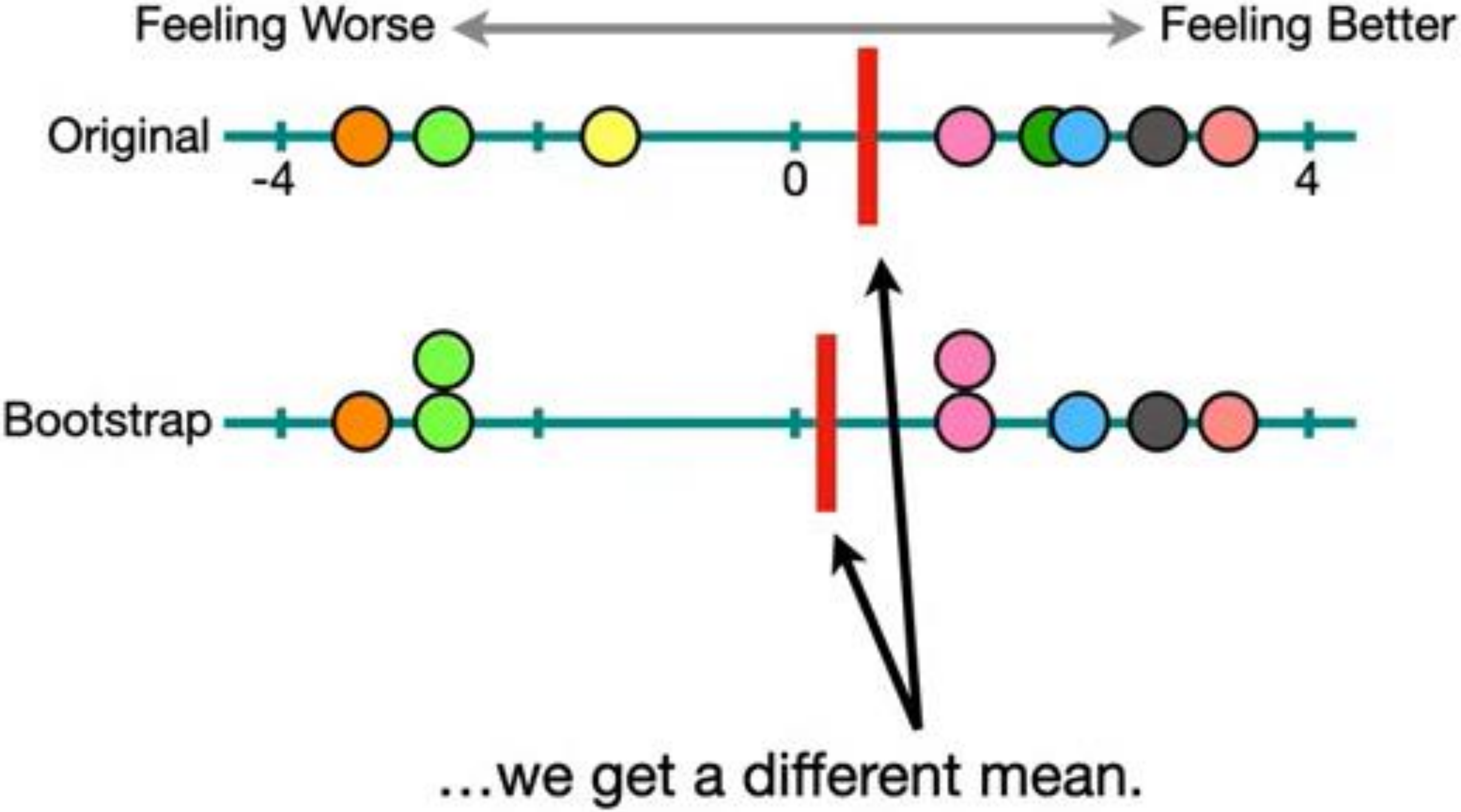


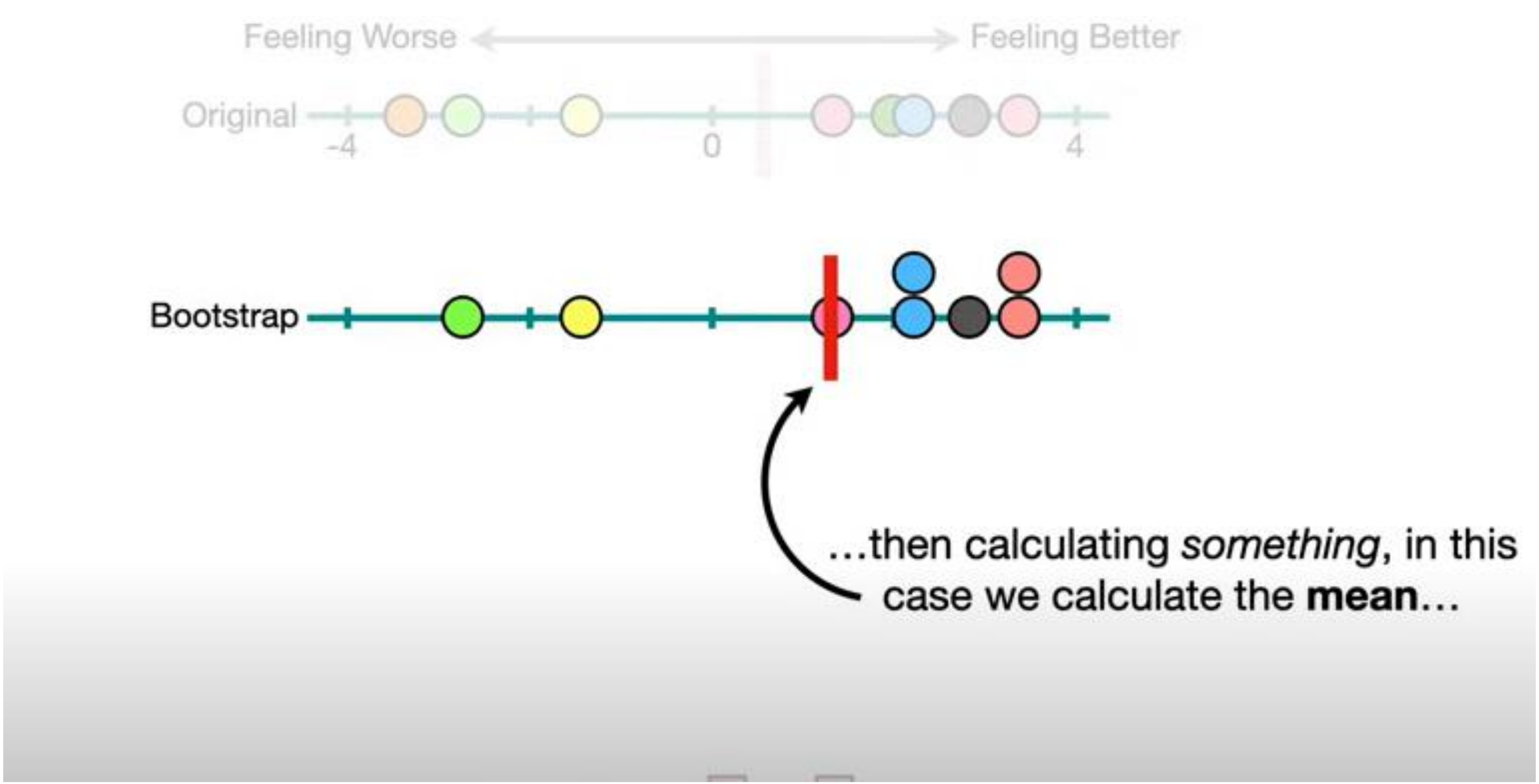
Randomly selecting data and allowing for duplicates is called **Sampling With Replacement**.

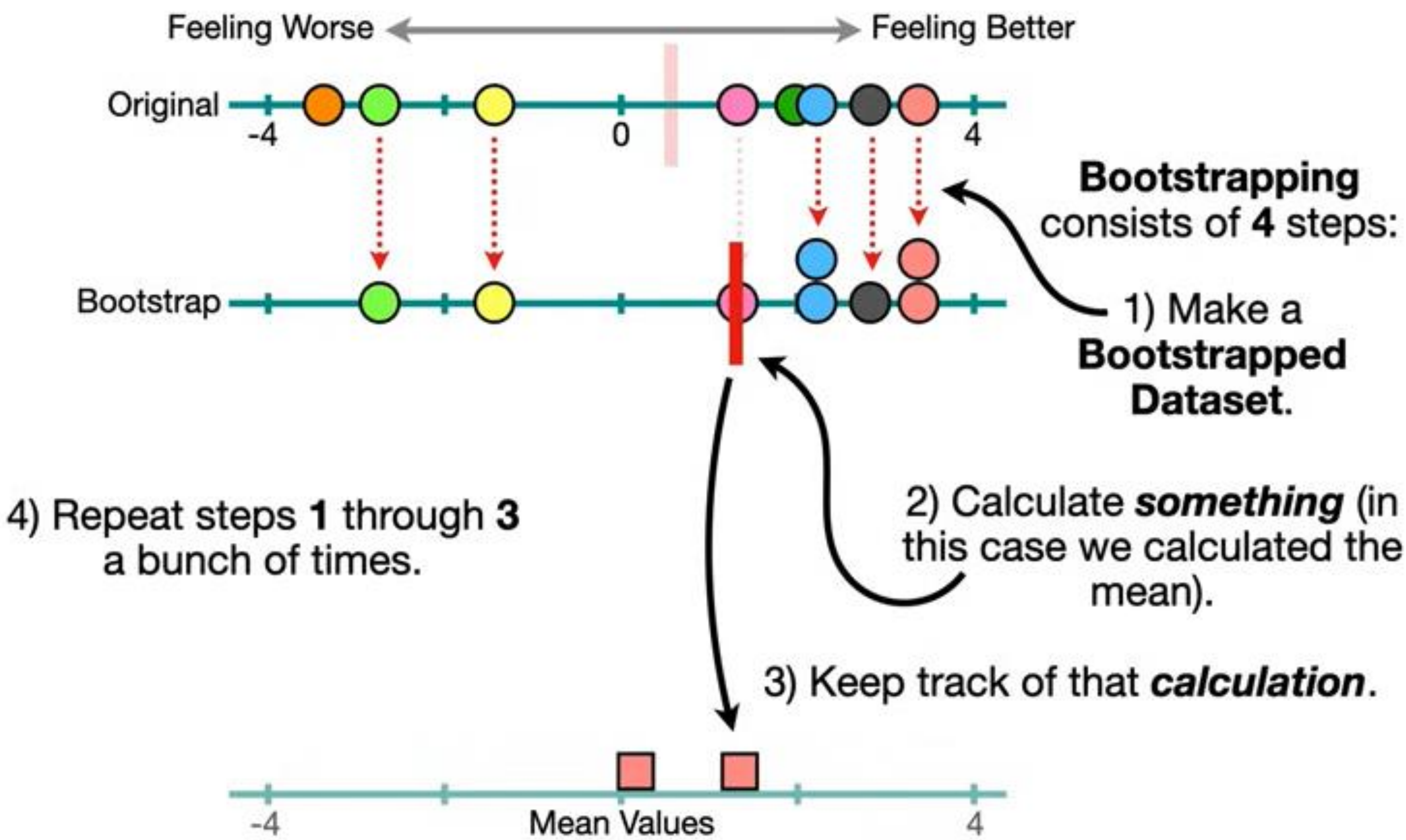


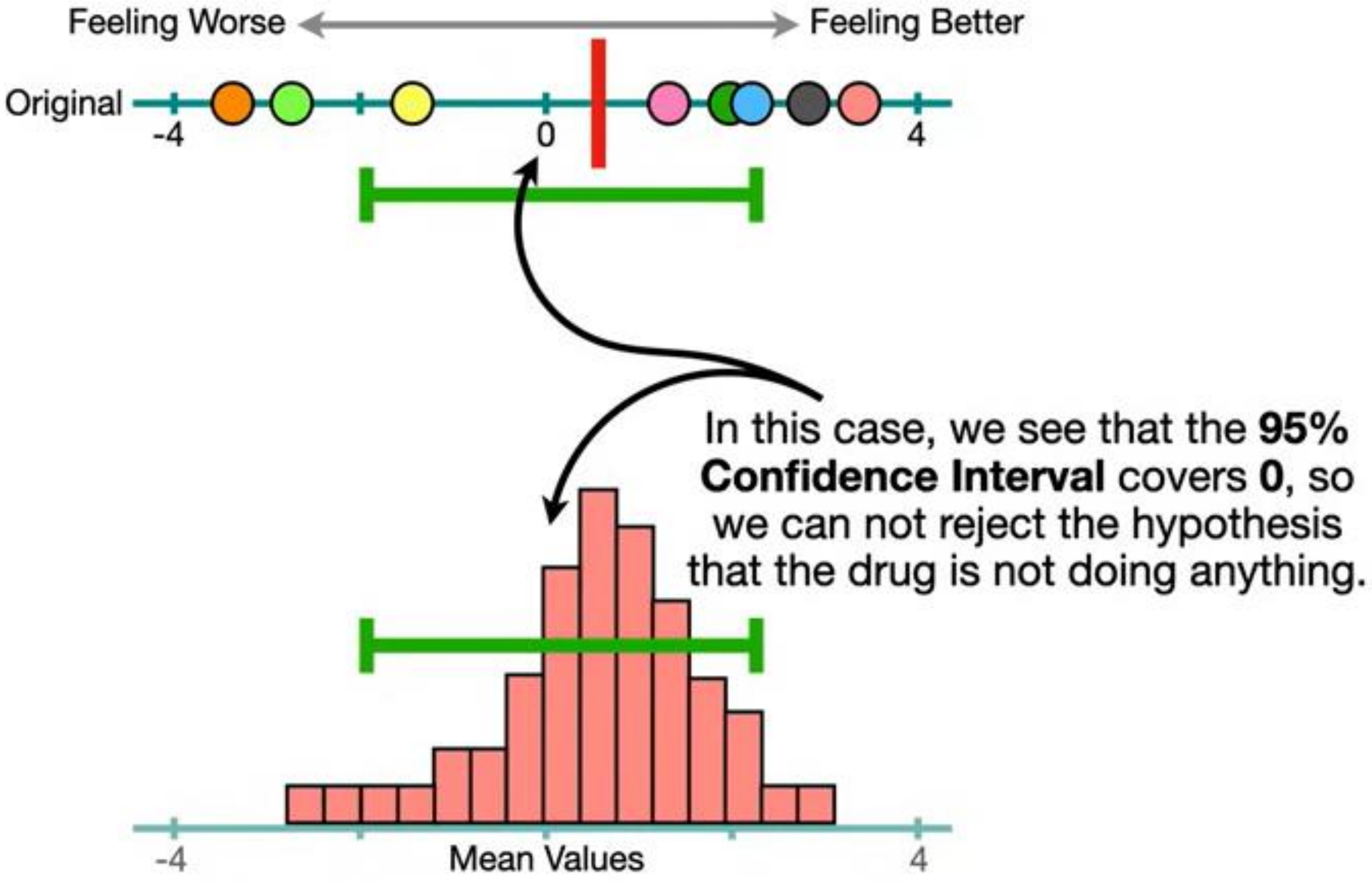
…is because the original dataset
that we are sampling from
contains **8** measurements.


Feeling Worse
Feeling Better
Original
-4
0
4
Bootstrap
...is called a **Bootstraped Dataset**.

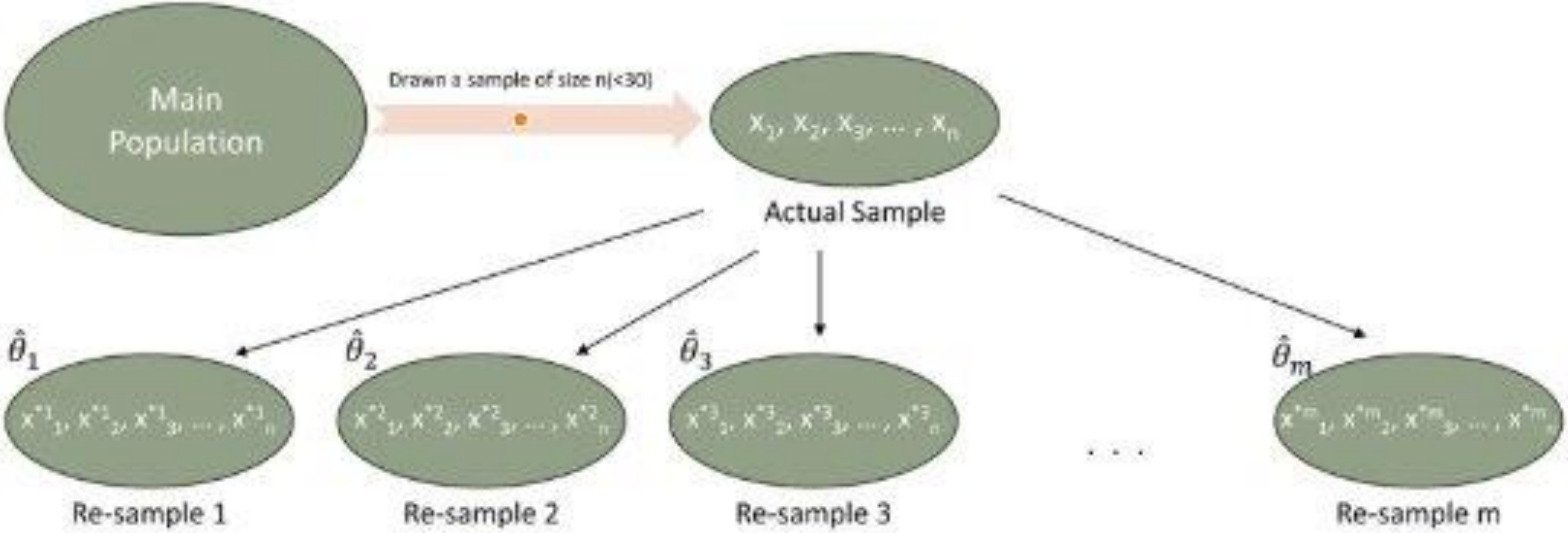

Feeling Worse
Feeling Better
Original
-4
0
4
Bootstrap
...we get a different mean.


Feeling Worse
Feeling Better
Original
-4
0
4
Bootstrap
…then calculating *something*, in this case we calculate the **mean**…


Feeling Worse
Feeling Better
Original
-4
0
4
Bootstrap
Bootstrapping consists of **4** steps:
1) Make a **Bootstrapped Dataset.**
2) Calculate ***something*** (in this case we calculated the mean).
3) Keep track of that ***calculation***.
4) Repeat steps **1** through **3** a bunch of times.
-4
Mean Values
4


Feeling Worse
Feeling Better
Original
-4
0
4
In this case, we see that the **95% Confidence Interval** covers **0**, so we can not reject the hypothesis that the drug is not doing anything.
-4
Mean Values
4

How Bootstrapping Statistics Works?



Here $\hat{\theta}_*$ represents the estimate of the model parameters

Differences between Bootstrapping and Traditional Hypothesis Testing

1) **Sampling Method**:
   - Traditional Hypothesis Testing: Relies on theoretical distributions and assumptions. It often assumes that the sample is a random representation of the population, and the analysis is based on predefined statistical distributions (e.g., normal distribution).
   - Bootstrapping: Involves resampling with replacement from the observed data. Instead of assuming the distribution, it uses the sample to estimate the sampling distribution.
2) **Parameter Estimation**:
   - Traditional Hypothesis Testing: Involves estimating parameters of the population based on the sample and using them to make inferences.
   - Bootstrapping: Estimates parameters by repeatedly resampling from the observed data, creating a distribution of the parameter not assuming a specific distribution.

Differences between Bootstrapping and Traditional Hypothesis Testing

3) **Sample Size Requirements**:
   - Traditional Hypothesis Testing: May require a sufficiently large sample size to meet the assumptions of the chosen statistical test.
   - Bootstrapping: Can be more robust with smaller sample sizes, as it generates its own "virtual samples" through resampling.

4) **Statistical Inference**:
   - Traditional Hypothesis Testing: Involves comparing a test statistic (calculated from the sample) to a critical value from a theoretical distribution to make inferences about the population parameter.
   - Bootstrapping: Constructs confidence intervals and makes inferences based on the distribution of the parameter obtained from resampling.