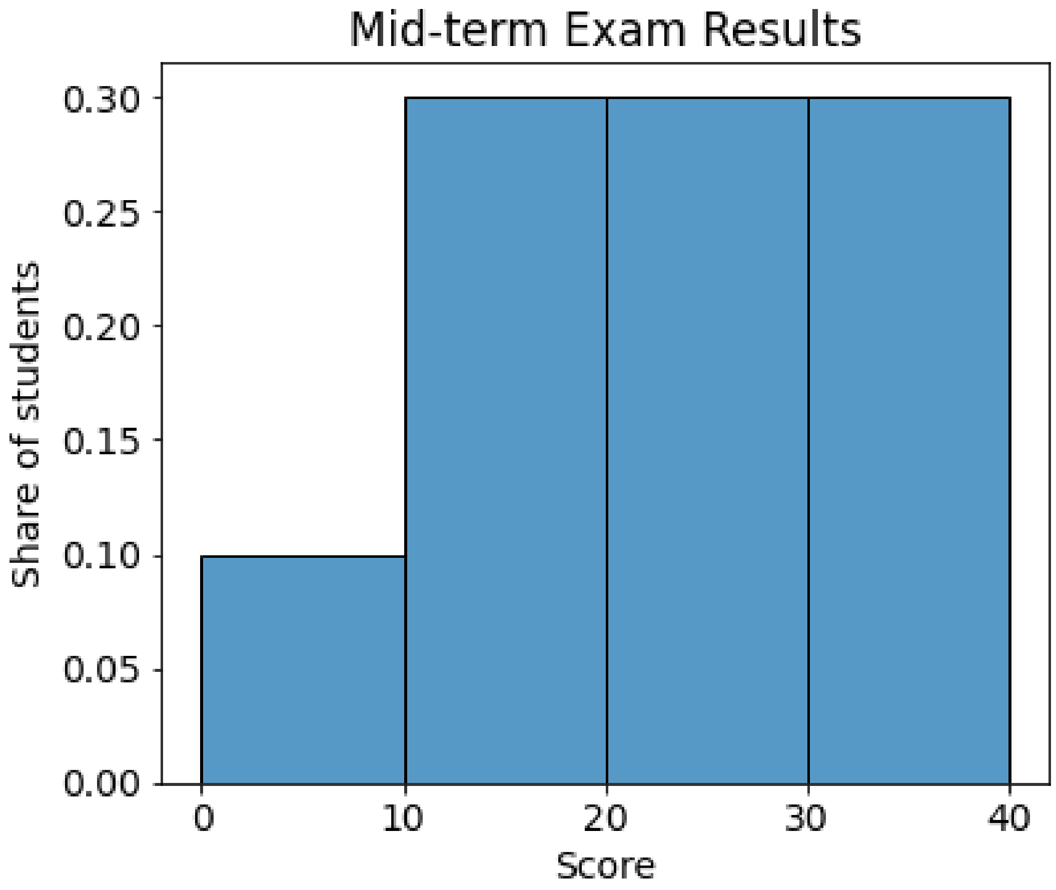
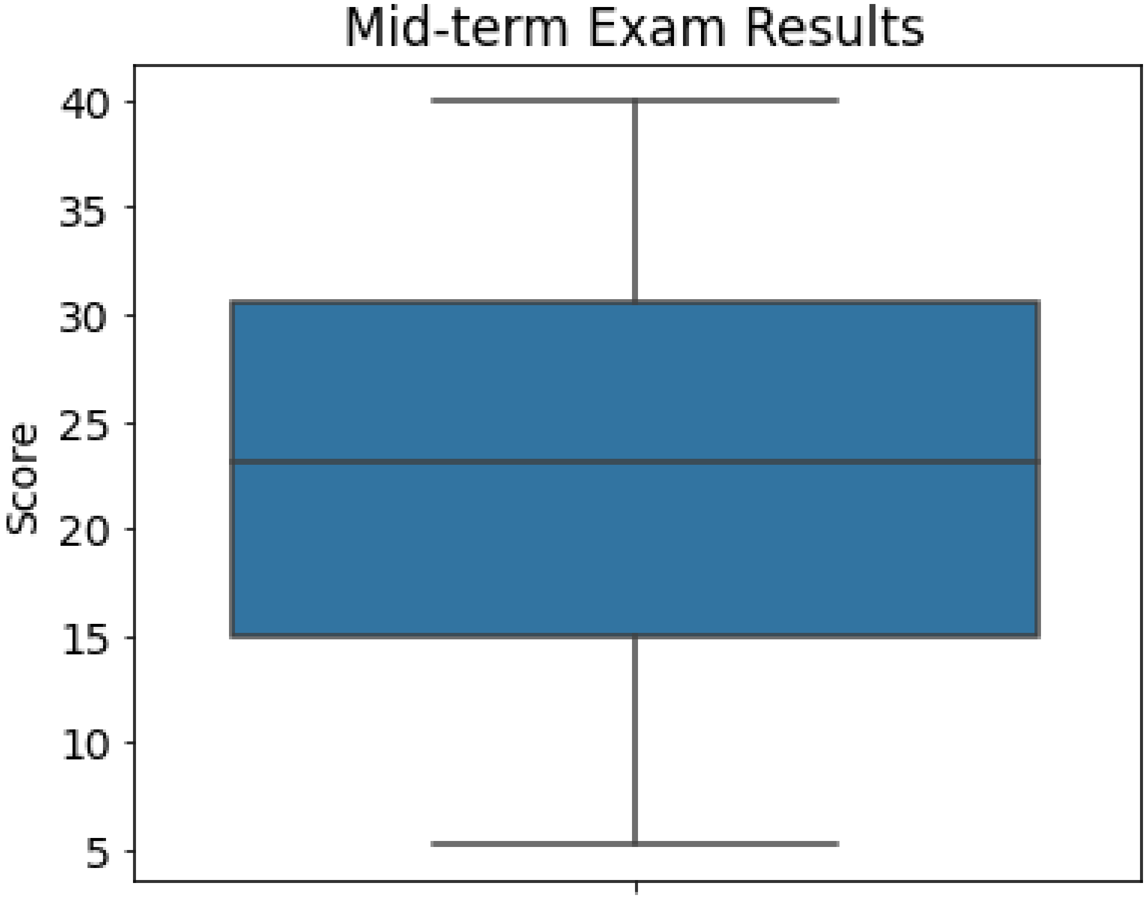
INTRODUCTION TO STATISTICS

LECTURE 10

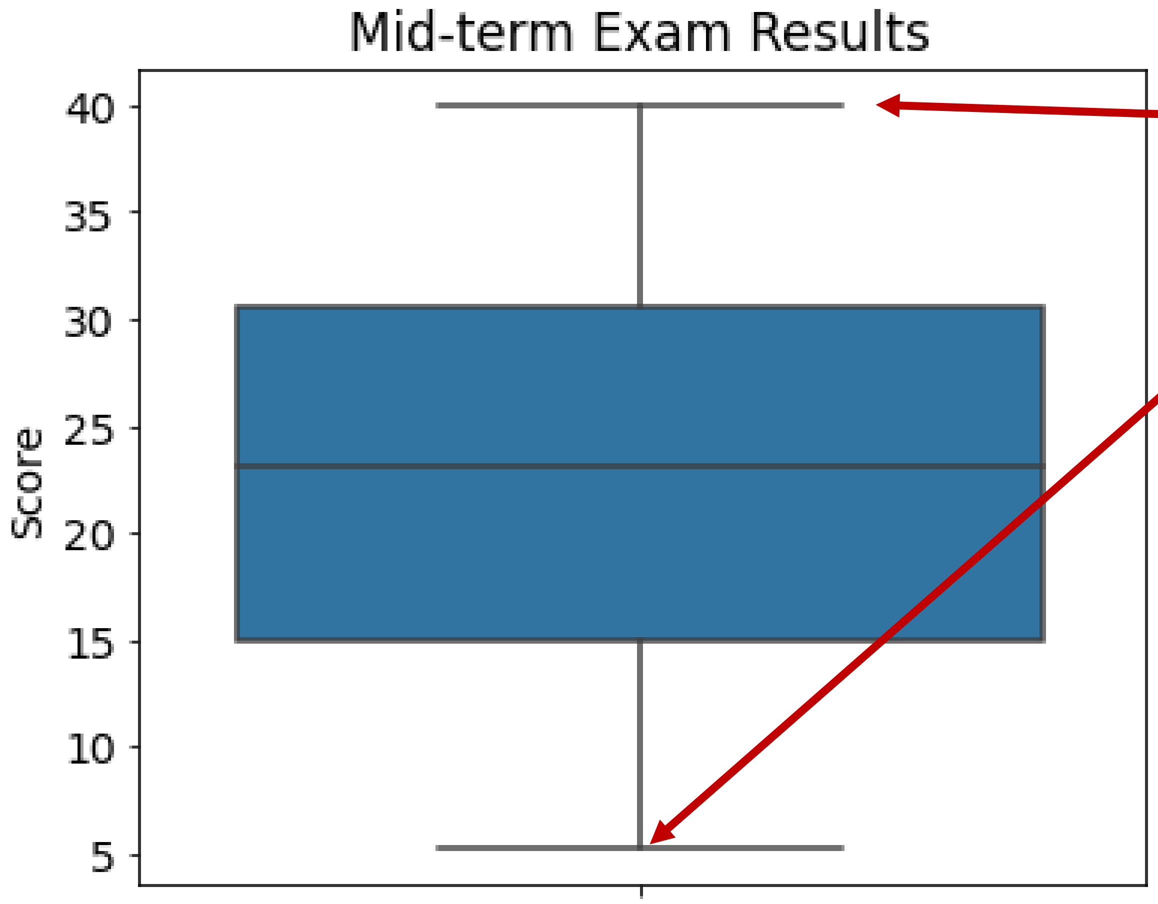
MID-TERM RESULTS: AN OVERVIEW



MID-TERM RESULTS: AN OVERVIEW

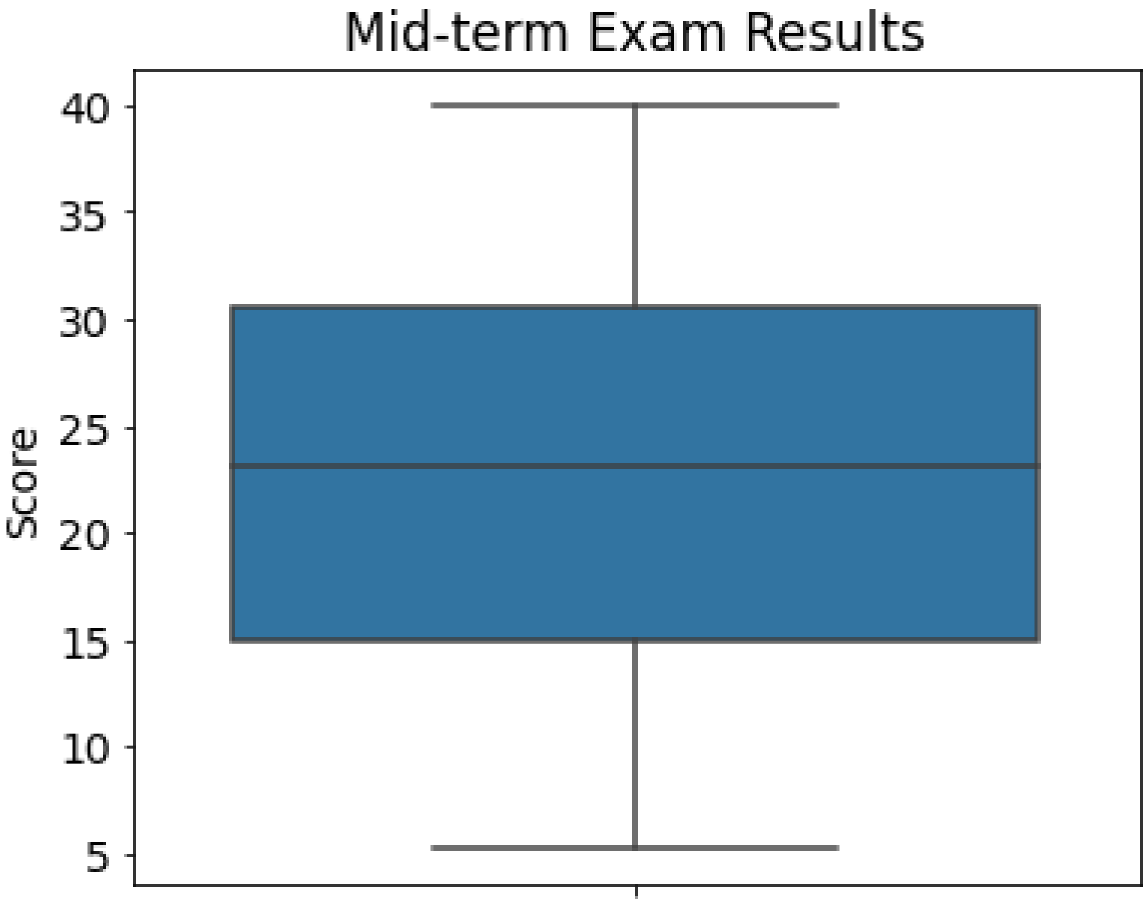


MID-TERM RESULTS: AN OVERVIEW



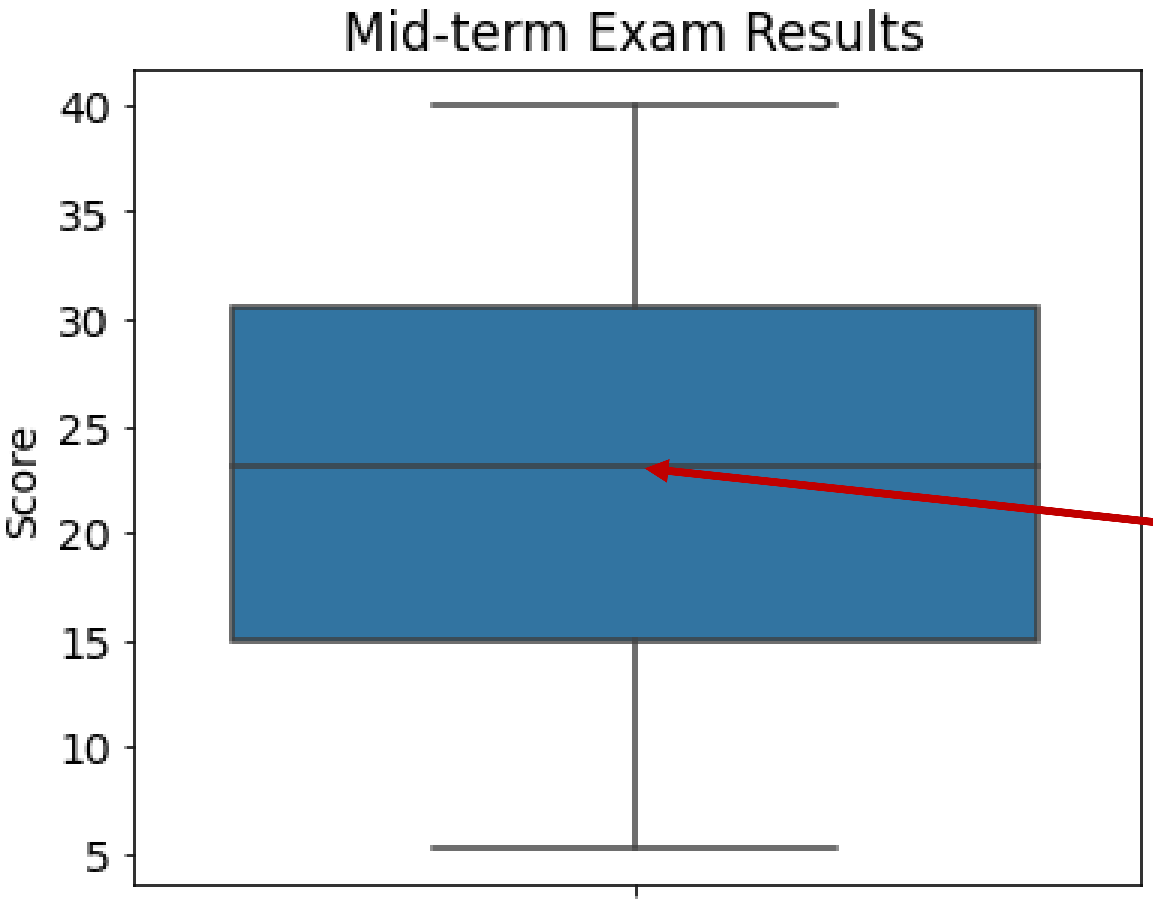
- MIN and MAX scores:
  - 5.25 and 40

MID-TERM RESULTS: AN OVERVIEW



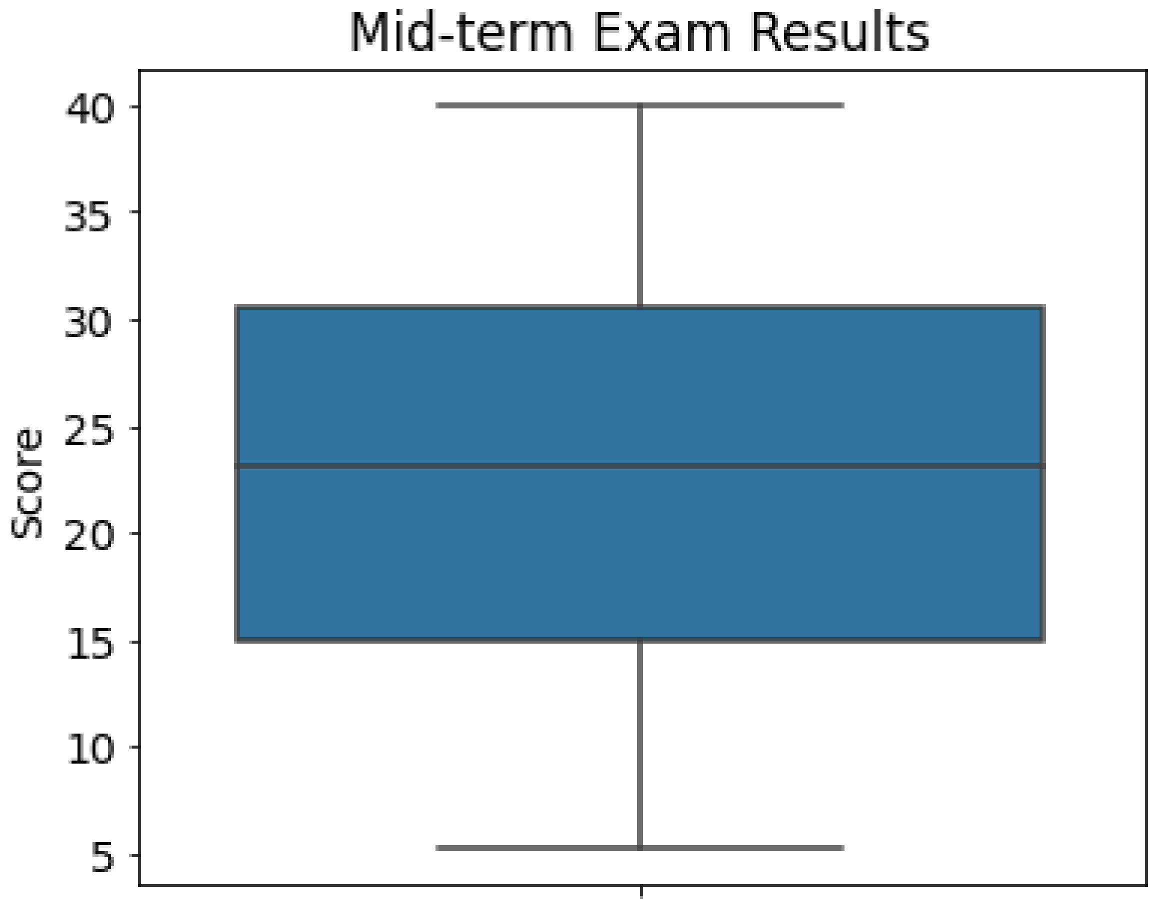
- MIN and MAX scores:
  - 5.25 and 40
- Mean score:
  - 23 points

MID-TERM RESULTS: AN OVERVIEW



- MIN and MAX scores:
  - 5.25 and 40
- Mean score:
  - 23 points
- Median score:
  - 23.1 points

MID-TERM RESULTS: AN OVERVIEW



- MIN and MAX scores:
  - 5.25 and 40
- Mean score:
  - 23 points
- Median score:
  - 23.1 points
- Standard deviation:
  - 10.5 points

PLANS FOR THE WEEK

- MONDAY
  - CI recap
  - Hypothesis testing I
- TUESDAY
  - Hypothesis testing II
- WEDNESDAY
  - Covariance and correlation
  - Linear regression
- THURSDAY
  - Big recap
- FRIDAY
  - Final exam

PLANS FOR THE WEEK

- MONDAY
  - CI recap
  - Hypothesis testing I

- Assignment 4, part I is due 23:59
- Assignment 4, part II is out

- TUESDAY
  - Hypothesis testing II

- WEDNESDAY
  - Covariance and correlation
  - Linear regression

- Assignment 4, part II is due 23:59
- Assignment 5 is out

- THURSDAY
  - Big recap

- FRIDAY
  - Final exam

- Assignment 5 is due 23:59

LAST TIME

- Confidence intervals:
  - z-interval;
  - t-interval;
  - χ^2-interval;
  - some examples.

- Hypothesis testing
  - *The Lady Tasting Tea experiment;*
  - *an introduction to hypothesis testing.*

LAST TIME

- Confidence intervals:
  - z-interval;
  - t-interval;
  - χ^2-interval;
  - some examples.

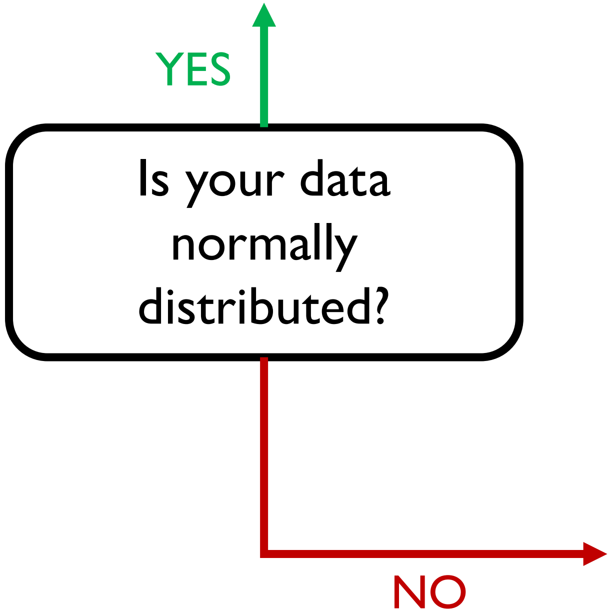
CONFIDENCE INTERVALS

A BRIEF RECAP

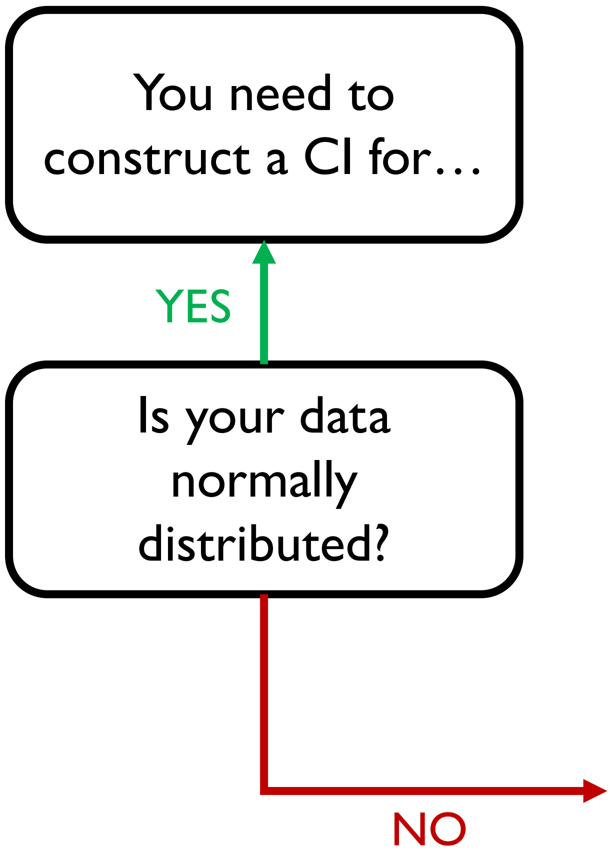
CI: A BRIEF RECAP

Is your data
normally
distributed?

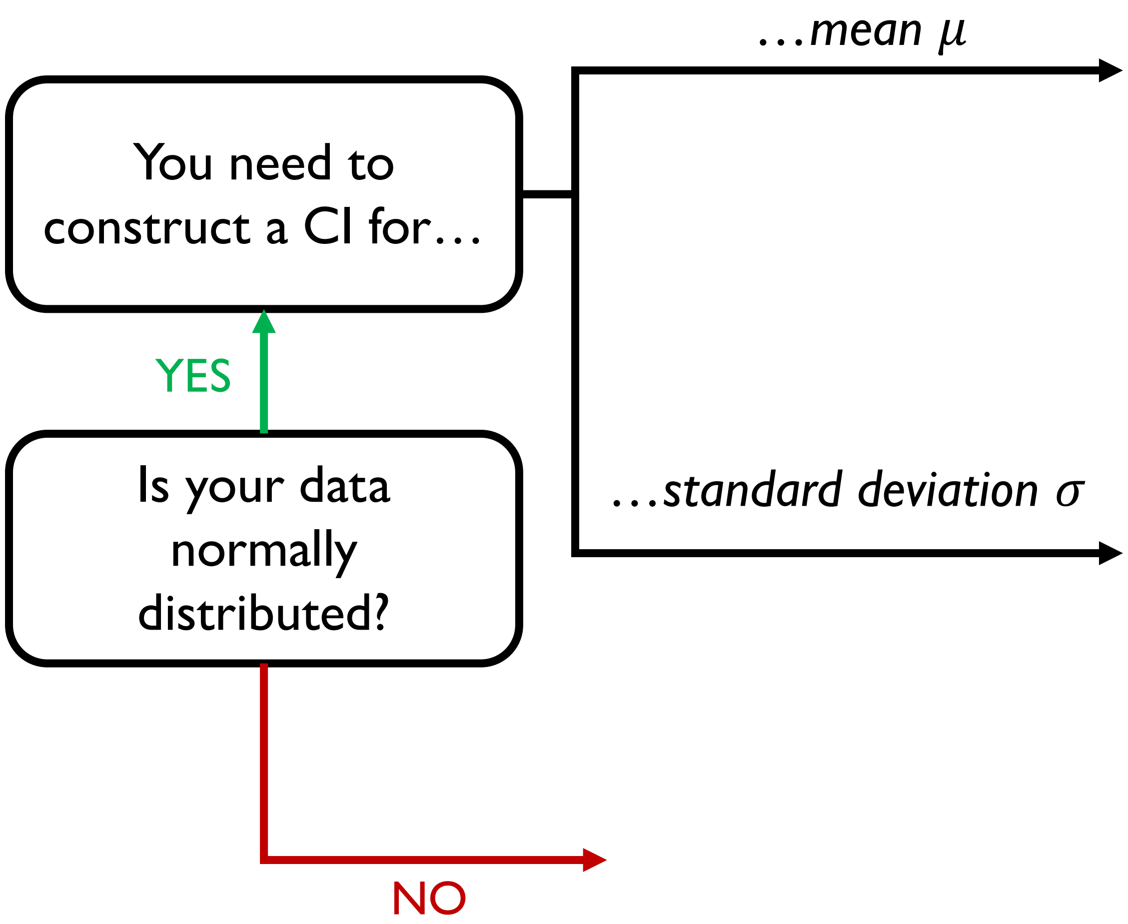
CI: A BRIEF RECAP



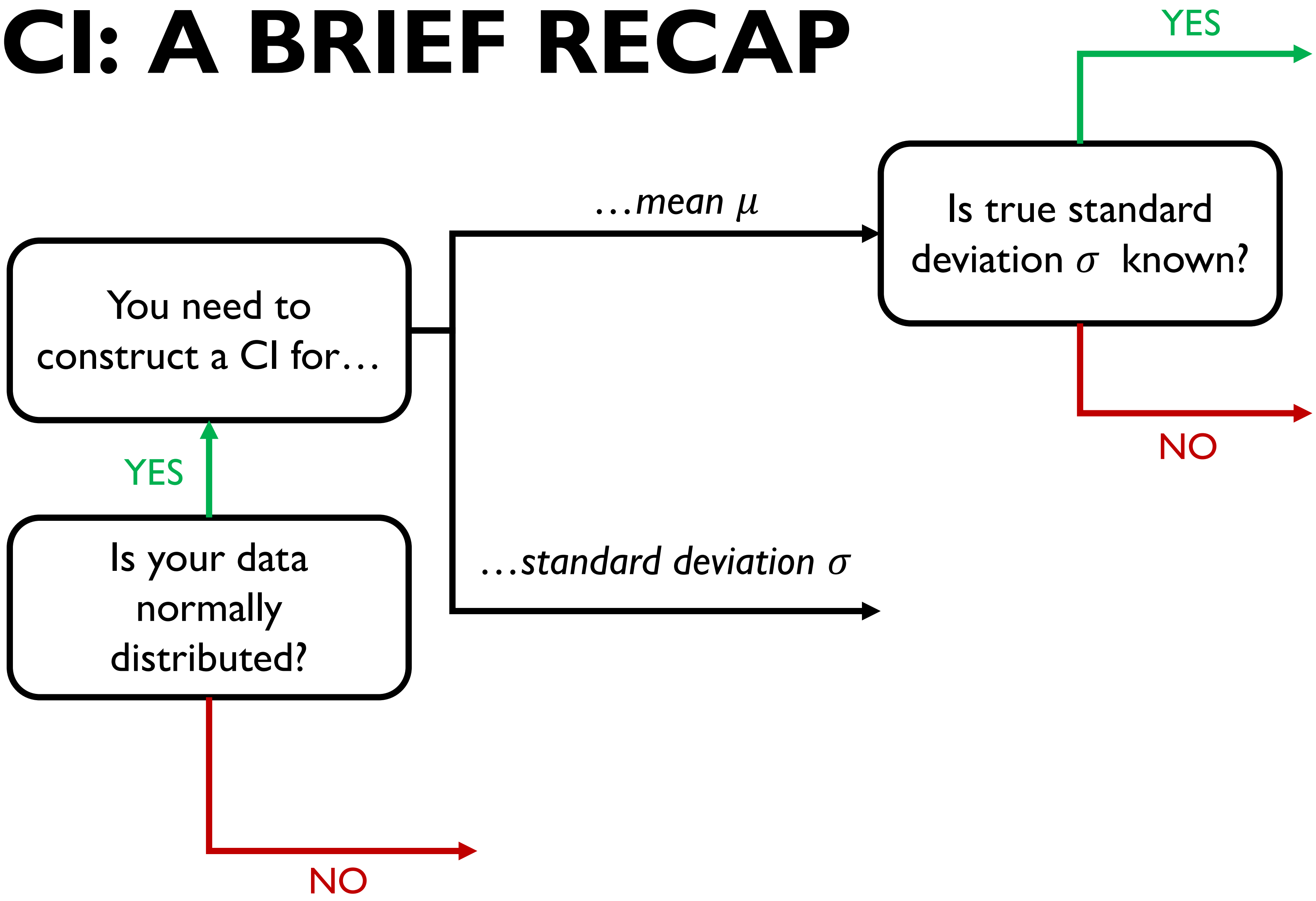
CI: A BRIEF RECAP



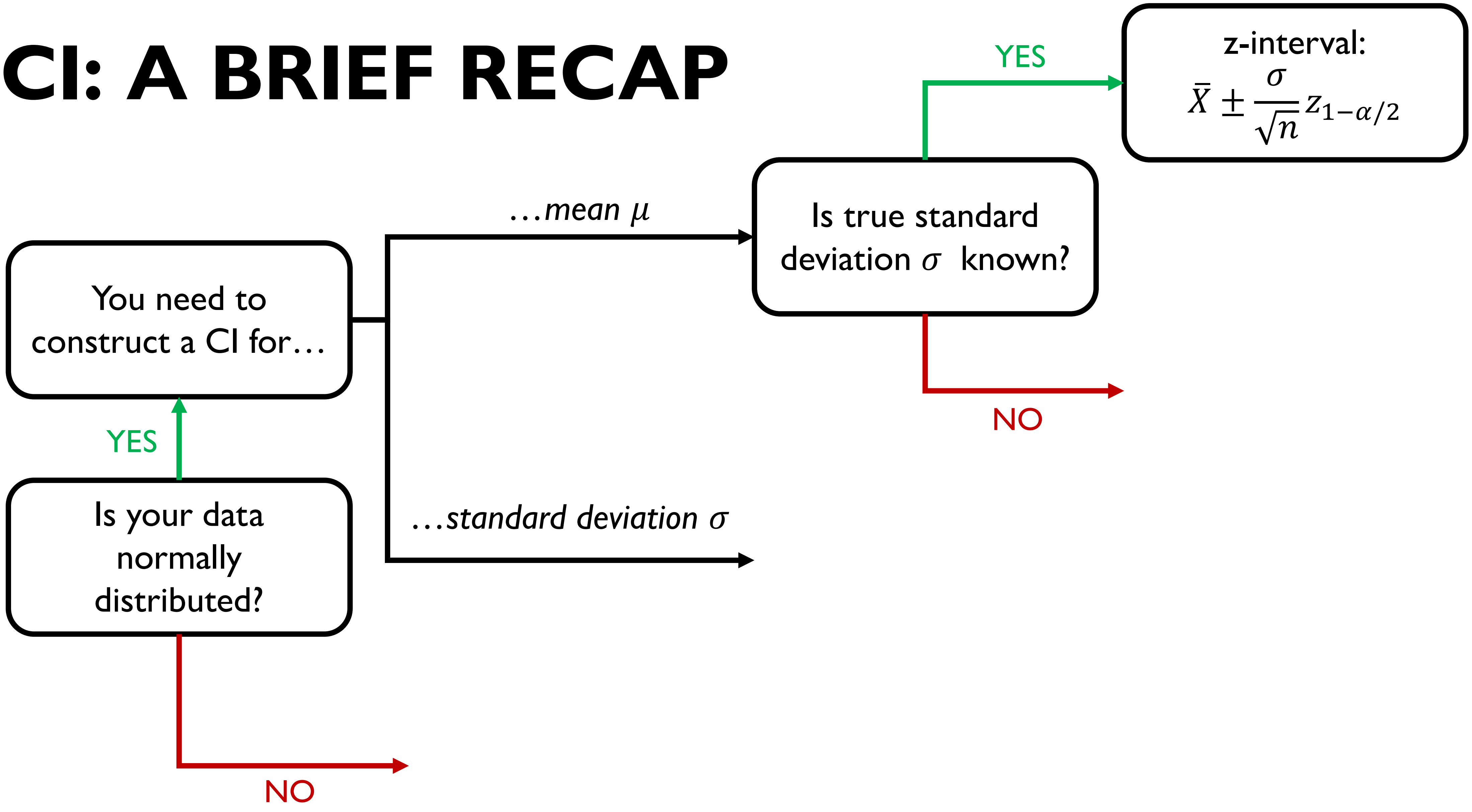
CI: A BRIEF RECAP



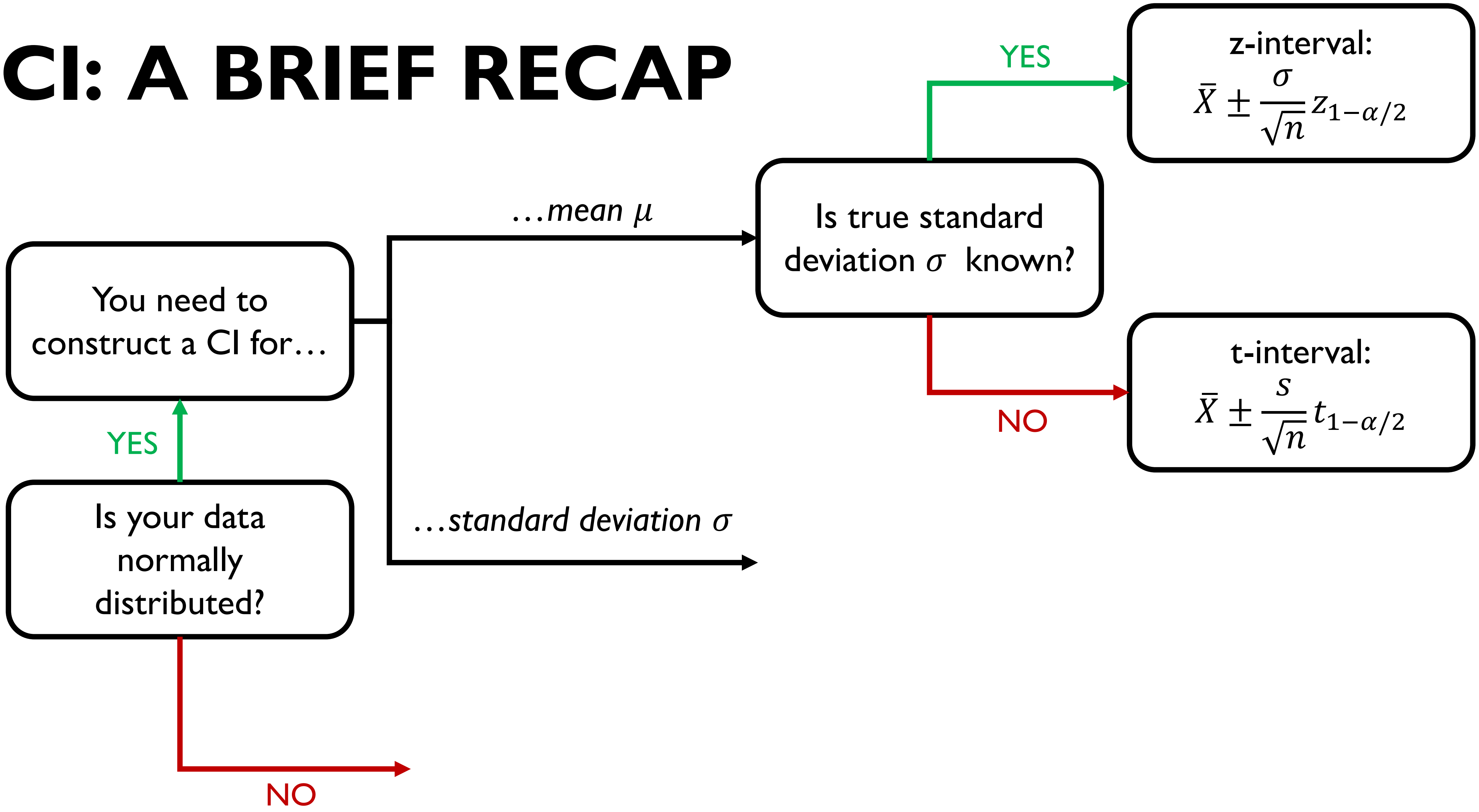
CI: A BRIEF RECAP



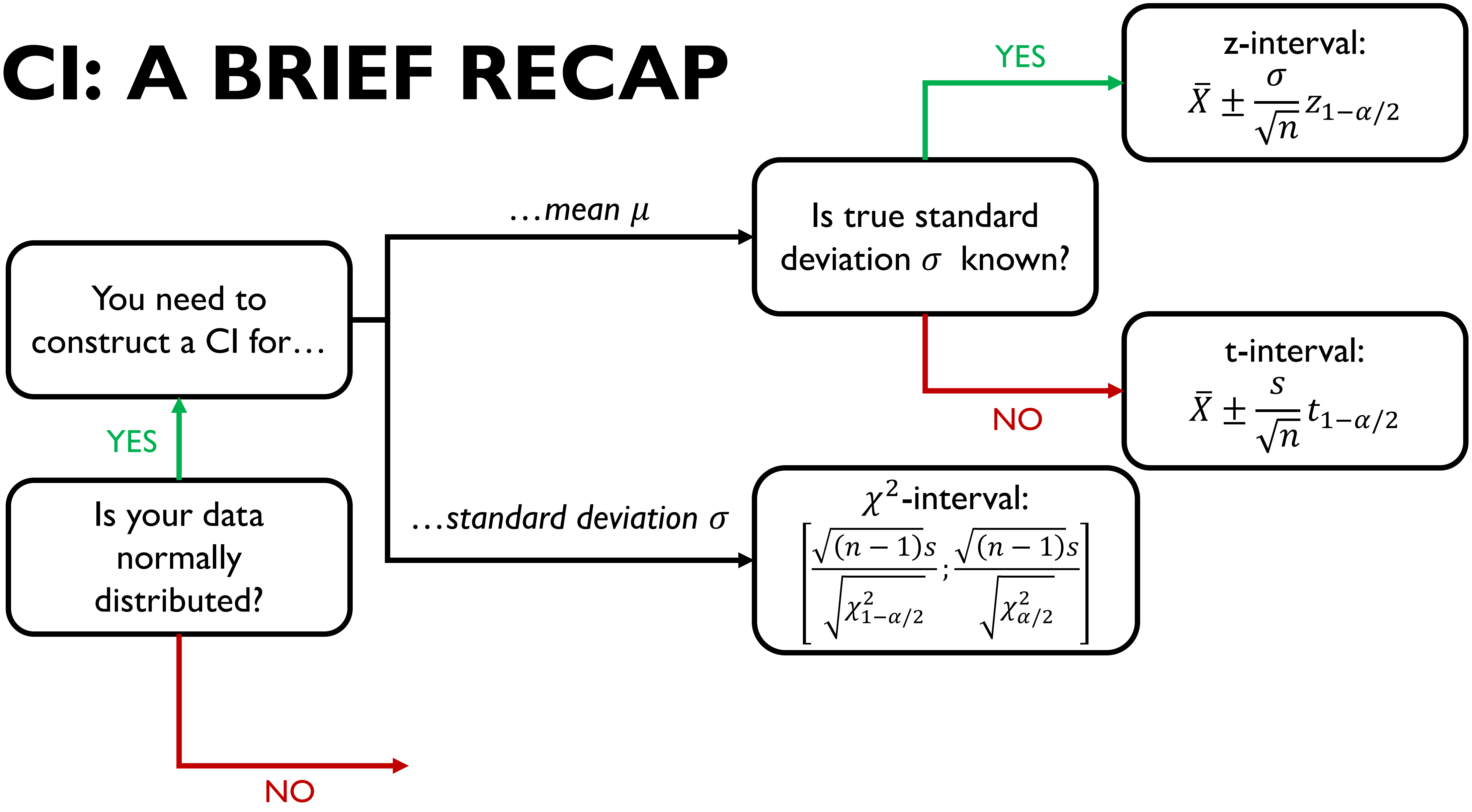
CI: A BRIEF RECAP



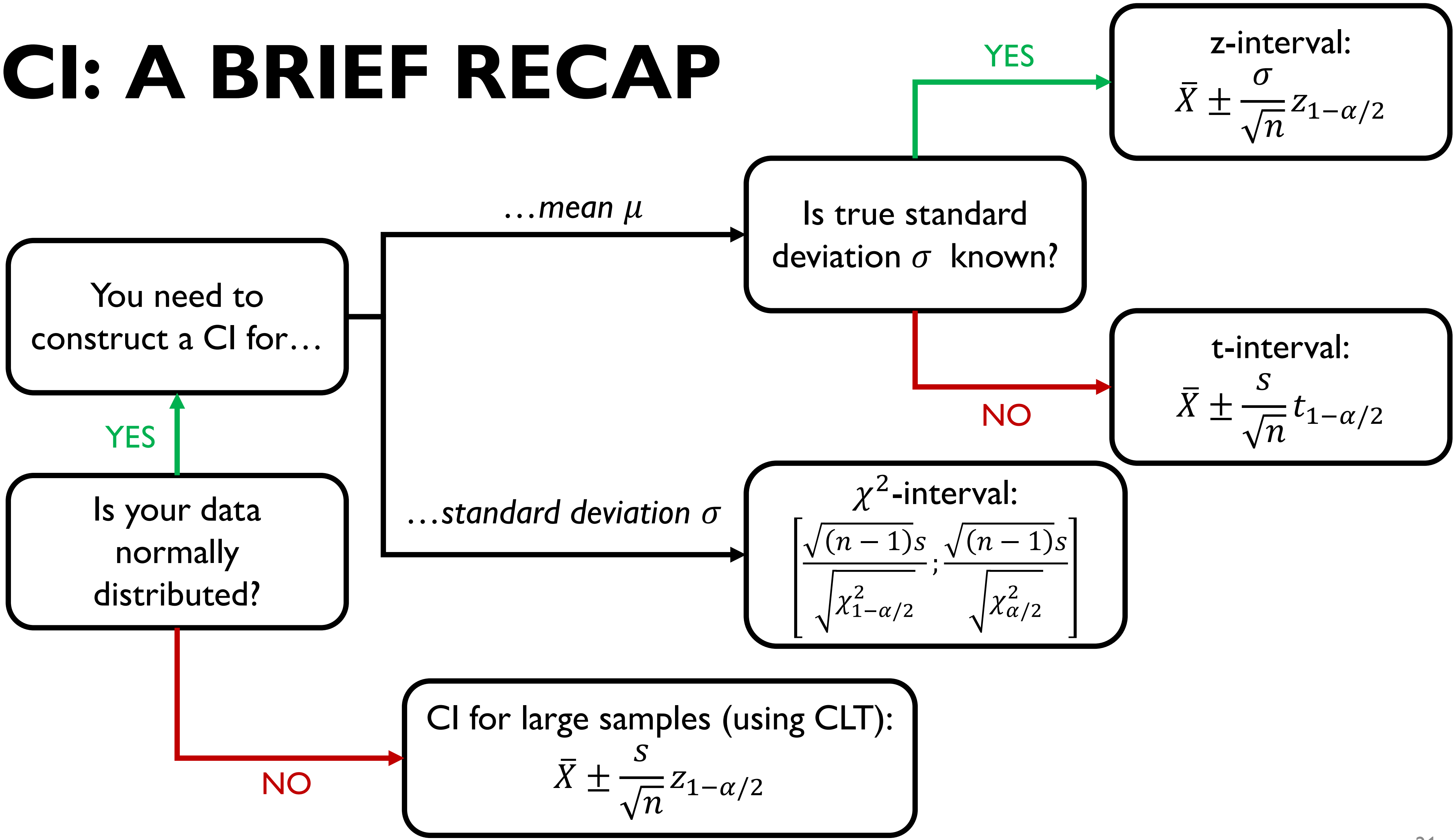
CI: A BRIEF RECAP



CI: A BRIEF RECAP



CI: A BRIEF RECAP



QUICK QUIZ

CI: A BRIEF RECAP

A machine fills in wine bottles with a random amount of wine that follows normal distribution with unknown mean μ and unknown standard deviation σ.

You check the last 100 bottles and see that on average they were filled with $\bar{X} = 705$ ml of wine, with sample std $s = 3$ ml.

Which interval would you use to construct a 95%-CI for μ?

CI: A BRIEF RECAP

In a study on cholesterol levels a sample of 1000 patients was chosen.

The average plasma cholesterol levels subjects was $\bar{X} = 6$ mmol/L, with sample std $s = 0.4$ mmol/L.

Which interval would you use to construct a 95%-CI for the true mean?

CI: A BRIEF RECAP

Height of a female student is a random variable following normal distribution with unknown mean μ and standard deviation $\sigma = 5$ cm.

The average height of 100 female students $\bar{X} = 165$ cm, and you sample std $s = 4$ cm

Which interval would you use to construct a 95%-CI for μ?

CI: A BRIEF RECAP

On a candy factory, a machine fills packs with random number of candies which follows normal distribution with unknown mean μ and unknown standard deviation σ.

In the last 50 packs, there were on average $\bar{X} = 90$ g of sweets, with sample std $s = 7$ g.

Which interval would you use to construct a 95%-CI for σ?

HYPOTHESIS TESTING

HYPOTHESIS TESTING

- Test whether a default assumption about the data is plausible.

- Key idea: you can't *prove* something, but you can disprove it.

HYPOTHESIS TESTING

- Test whether a default assumption about the data is plausible.

HYPOTHESIS TESTING

ASSUMPTION:
ALL THE SWANS ARE WHITE

HYPOTHESIS TESTING

ASSUMPTION:
ALL THE SWANS ARE WHITE

EVIDENCE:

HYPOTHESIS TESTING

ASSUMPTION:
ALL THE SWANS ARE WHITE

EVIDENCE:
NOT CONVINCING…

HYPOTHESIS TESTING

ASSUMPTION:
ALL THE SWANS ARE WHITE

EVIDENCE:
NOW IT's CONVINCING!

HYPOTHESIS TESTING

ASSUMPTION:
ALL THE SWANS ARE WHITE

EVIDENCE:

HYPOTHESIS TESTING

ASSUMPTION:
THERE’RE NO ALIENS ON THE MOON

HYPOTHESIS TESTING

ASSUMPTION:
THERE'RE NO ALIENS ON THE MOON

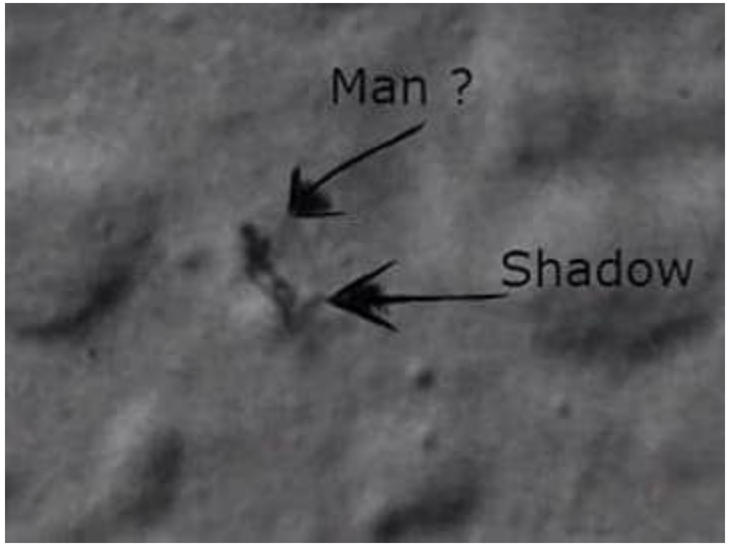
EVIDENCE:

HYPOTHESIS TESTING

ASSUMPTION:
THERE’RE NO ALIENS ON THE MOON

EVIDENCE:

STILL NO REASON TO BELIEVE THERE ARE ALIENS

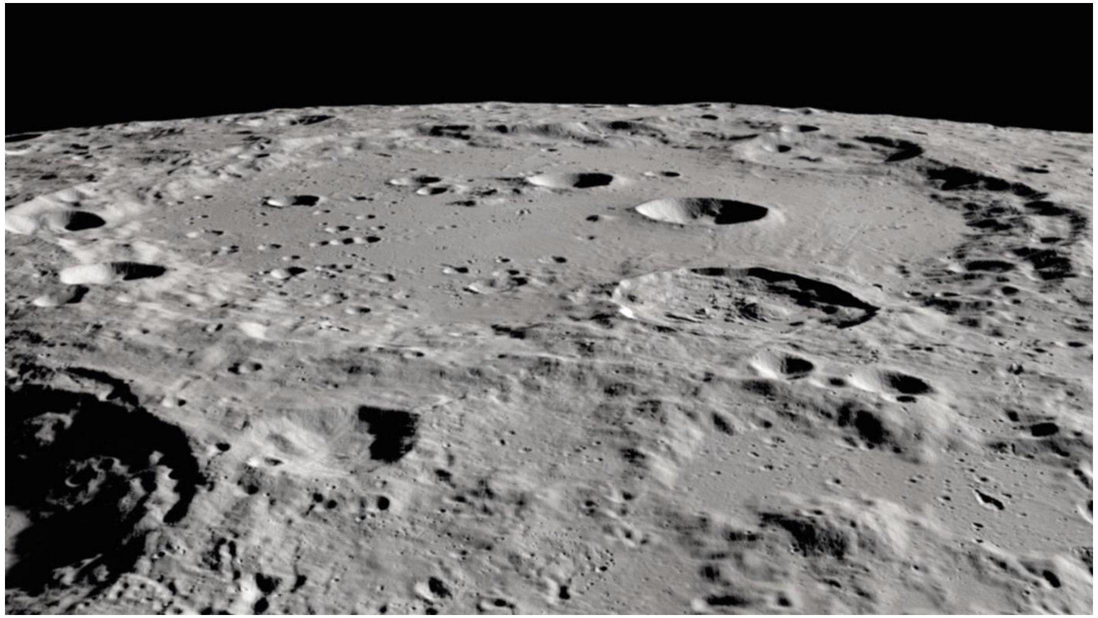


HYPOTHESIS TESTING

ASSUMPTION:
THERE'RE NO ALIENS ON THE MOON

EVIDENCE:

STILL NO REASON TO BELIEVE THERE ARE ALIENS



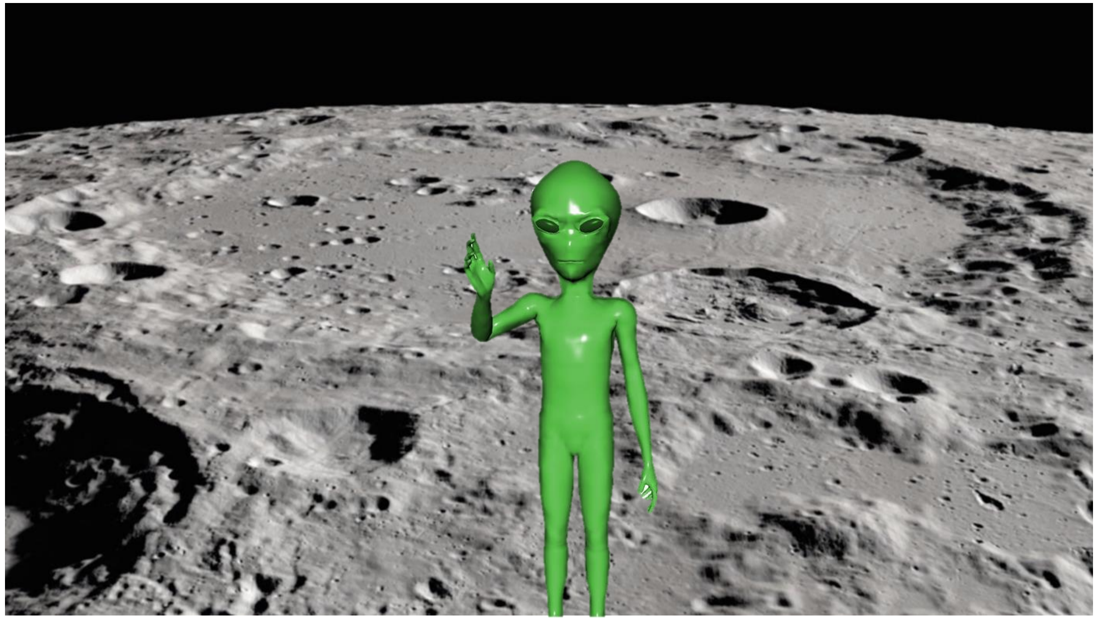
HYPOTHESIS TESTING

ASSUMPTION:

THERE'RE NO ALIENS ON THE MOON

EVIDENCE:

THERE ARE ALIENS!!!

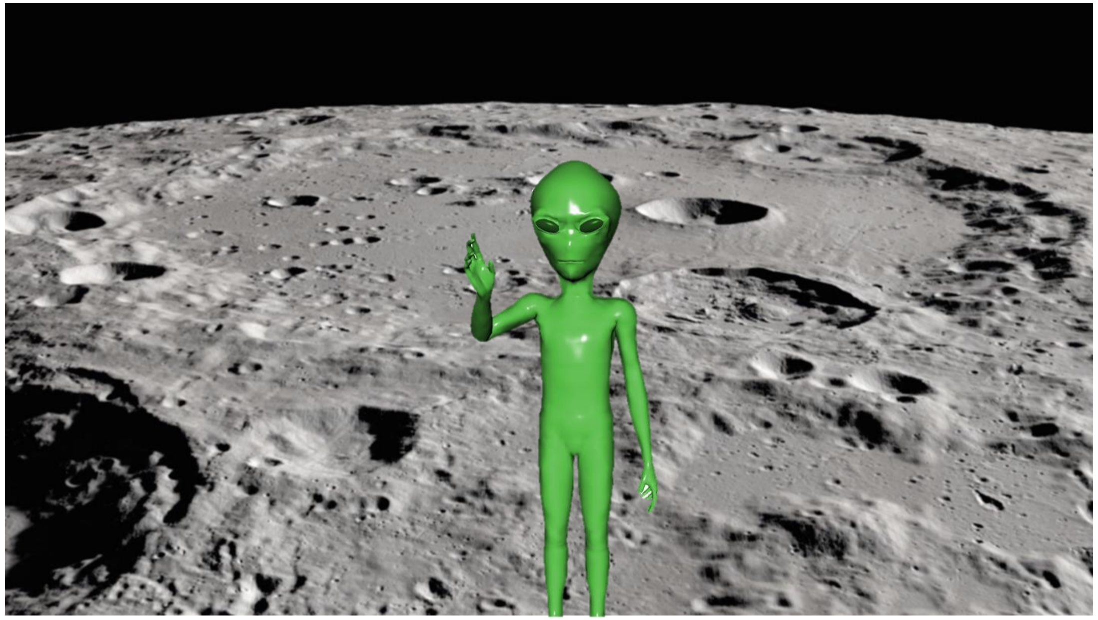


HYPOTHESIS TESTING

ASSUMPTION:
THERE'RE NO ALIENS ON THE MOON

EVIDENCE:

THERE ARE ALIENS!!!



HYPOTHESIS TESTING

- Note:
  - Default assumption = "status quo".

HYPOTHESIS TESTING

- Note:
    - Default assumption = “status quo”.
    - Alternative: something we can find evidence for.

HOW TO TEST A HYPOTHESIS

BASIC INGREDIENTS

RUNNING EXAMPLE

How to check if a coin is a fair one?

RUNNING EXAMPLE

We flip a coin 10 times to test if it’s a fair one.

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis** $\boldsymbol{H_0}$: *the default assumption for the model generating the data.*

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\colon\ p = 0.5 \ \text{‘the coin is fair’}$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\text{: } p = 0.5 \text{ ‘the coin is fair’}$$

- **The alternative hypothesis $\boldsymbol{H_1}$**: *if we reject H_0, we accept H_1 as the best explanation for the data.*

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\text{: } p = 0.5 \text{ ‘the coin is fair’}$$

- **The alternative hypothesis $\boldsymbol{H_1}$**: *if we reject H_0, we accept H_1 as the best explanation for the data.*

$$H_1\text{: } p \neq 0.5 \text{ ‘the coin is not fair’}$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\colon\ p = 0.5\ \text{'the coin is fair'}$$

- **The alternative hypothesis $\boldsymbol{H_1}$**: *if we reject H_0, we accept H_1 as the best explanation for the data.*

$$H_1\colon\ p \neq 0.5\ \text{'the coin is not fair'}$$

$$H_1\colon p > 0.5$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\colon\ p = 0.5 \ \text{‘the coin is fair’}$$

- **The alternative hypothesis $\boldsymbol{H_1}$**: *if we reject H_0, we accept H_1 as the best explanation for the data.*

$$H_1\colon\ p \neq 0.5 \ \text{‘the coin is not fair’}$$

$$H_1\colon p > 0.5 \quad \text{or} \quad H_1\colon p < 0.5$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\!: p = 0.5 \text{ ‘the coin is fair’}$$

- **The alternative hypothesis $\boldsymbol{H_1}$**: *if we reject H_0, we accept H_1 as the best explanation for the data.*

Two-sided: $H_1\!: p \neq 0.5$ ‘the coin is not fair’

One-sided: $H_1\!: p > 0.5$ or $H_1\!: p < 0.5$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

- **The null hypothesis $\boldsymbol{H_0}$** : *the default assumption for the model generating the data.*

$$H_0\colon\ p = 0.5 \ \text{'the coin is fair'}$$

- **The alternative hypothesis $\boldsymbol{H_1}$**: *if we reject H_0, we accept H_1 as the best explanation for the data.*

Two-sided: $H_1\colon\ p \neq 0.5$ 'the coin is not fair'

One-sided: $H_1\colon p > 0.5$ or $H_1\colon p < 0.5$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

X – *number of H*

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

  X – *number of H*

- **Null distribution**: *the probability distribution of* X *assuming* H_0

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

X − *number of H*

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim Binomial(\quad\quad)$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim Binomial(10, 0.5)$$

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim Binomial(10, 0.5)$$

X											
$P(X\|H_0)$											

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim Binomial(10, 0.5)$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$											

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of X assuming H_0*

$$X \sim Binomial(10, 0.5)$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001										0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim Binomial(10, 0.5)$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01								0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

- **Test statistic**: *computed from the data*

$$X - \textit{number of H}$$

- **Null distribution**: *the probability distribution of* X *assuming* H_0

$$X \sim Binomial(10, 0.5)$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X \mid H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

- **Rejection region R**: if X is in the rejection region, we reject H_0 in favor of H_1.

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

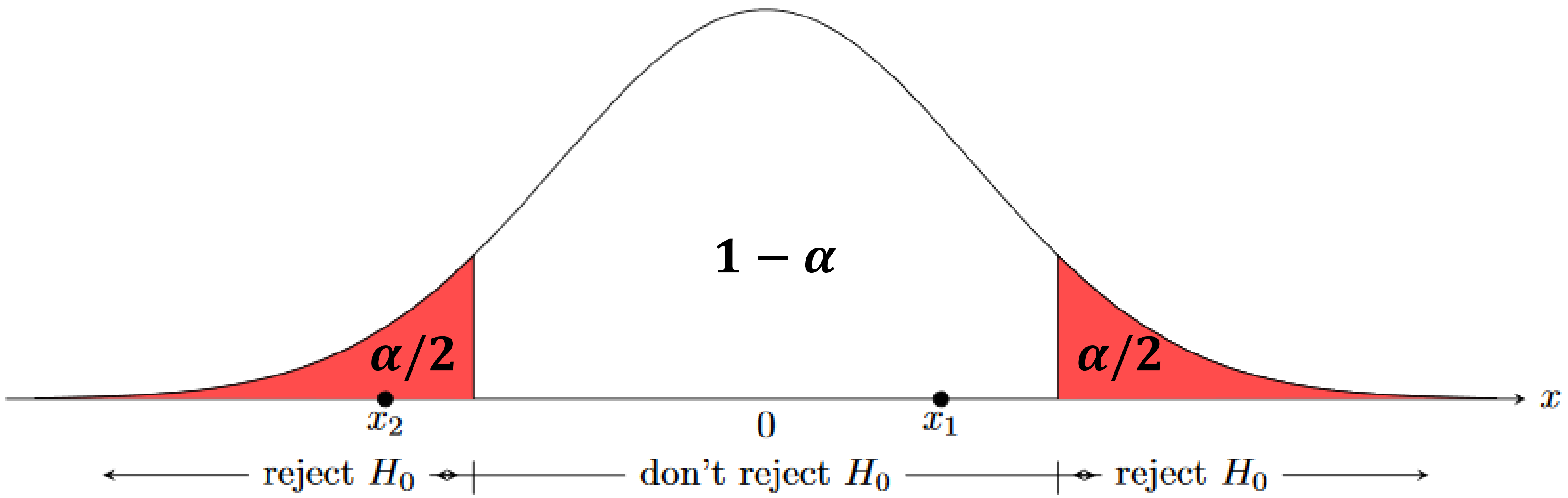
H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

- **Rejection region R**: if X is in the rejection region, we reject H_0 in favor of H_1.
- **Significance level** α **:** $P(X \in R|H_0) \leq \alpha$
  - Typically chosen in advance (common values are 0.1, 0.05, 0.01.)

INGREDIENTS

- Distribution of the test statistic under H_0, two-sided alternative:



INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

$$\boldsymbol{\alpha = 0.05}$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

$$\boldsymbol{\alpha = 0.05}$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

$$\boldsymbol{\alpha = 0.2}$$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

$$\boldsymbol{\alpha = 0.2}$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

REMARKS

- **The null hypothesis** is the *cautious default*
  - we won’t claim the coin is unfair unless we have compelling evidence.

REMARKS

- **The null hypothesis** is the *cautious default*
  - we won't claim the coin is unfair unless we have compelling evidence.

- **Rejection region**: data that is *extreme* <u>under the null hypothesis</u> = outcomes in the tail(s) of the null distribution.
  - depends on the significance level α of the test.

REMARKS

H_0: $p = 0.5$ ‘the coin is fair’, H_1: $p \neq 0.5$ ‘the coin is not fair’ $\alpha = \mathbf{0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

REMARKS

H_0: $p = 0.5$ ‘the coin is fair’, H_1: $p \neq 0.5$ ‘the coin is not fair’ $\alpha = \mathbf{0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

- If we get *0, 1, 9, 10* H
  - Reject H_0

REMARKS

H_0: $p = 0.5$ ‘the coin is fair’, H_1: $p \neq 0.5$ ‘the coin is not fair’ $\boldsymbol{\alpha = 0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

- If we get *0, 1, 9, 10* H
  - Reject H_0

- If we get 2, *3, 4, 5, 6, 7, 8* H, then the test statistic is in the non-rejection region.

REMARKS

H_0: $p = 0.5$ ‘the coin is fair’, H_1: $p \neq 0.5$ ‘the coin is not fair’ $\boldsymbol{\alpha = 0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

- If we get *0, 1, 9, 10* H
  - Reject H_0
- If we get 2, *3, 4, 5, 6, 7, 8* H, then the test statistic is in the non-rejection region.
  - Interpretation: the data ‘does not support rejecting the null hypothesis’.

REMARKS

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair' $\boldsymbol{\alpha = 0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

- If we get *0, 1, 9, 10* H
  - Reject H_0
- If we get 2, *3, 4, 5, 6, 7, 8* H, then the test statistic is in the non-rejection region.
  - Interpretation: the data 'does not support rejecting the null hypothesis'.
  - **Never claim that the data proves the null hypothesis is true.**

EXPERIMENT

EXPERIMENT

- Toss a coin 10 times, record the number of heads you got.

EXPERIMENT

- Toss a coin 10 times, record the number of heads you got.

- Based on the result, do you reject the null hypothesis?

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair', $\alpha = 0.2$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INCORRECTLY REJECTING H_0

H_0: $p = 0.5$ ‘the coin is fair’, H_1: $p \neq 0.5$ ‘the coin is not fair’

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X \mid H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

What is the probability to incorrectly reject H_0?

INCORRECTLY REJECTING H_0

H_0: $p = 0.5$ 'the coin is fair', H_1: $p \neq 0.5$ 'the coin is not fair'

X	0	1	2	3	4	5	6	7	8	9	10
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

What is the probability to incorrectly reject H_0?

$$P(Reject\ H_0|H_0) = P(X \in R|H_0) = \alpha -$$

significance level

ONE-SIDED ALTERNATIVES

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p > 0.5}$ 'the biased towards H'

$\boldsymbol{\alpha = 0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p > 0.5}$ 'the biased towards H'

$\boldsymbol{\alpha = 0.05}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p > 0.5}$ 'the biased towards H'

$\boldsymbol{\alpha = 0.1}$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p > 0.5}$ 'the biased towards H'

$$\boldsymbol{\alpha = 0.1}$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

$H_0\colon p = 0.5,$ $\boldsymbol{H_1\colon p < 0.5}$ 'the biased towards T'

$\boldsymbol{\alpha = 0.05}$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p < 0.5}$ 'the biased towards T'

$$\boldsymbol{\alpha = 0.05}$$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p < 0.5}$ 'the biased towards T'

$\boldsymbol{\alpha = 0.1}$

X	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$P(X\|H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

INGREDIENTS

We flip a coin 10 times to test if it's a fair one.

H_0: $p = 0.5$, $\boldsymbol{H_1}$: $\boldsymbol{p < 0.5}$ 'the biased towards T'

$\boldsymbol{\alpha = 0.1}$

X	0	1	2	3	4	5	6	7	8	9	10
$P(X \mid H_0)$	0.001	0.01	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.01	0.001

Z-TEST

When the data is coming from normal distribution
and variance is known

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.
- We suspect that HarbourSpace students are different. How do we test it?

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.

- We suspect that HarbourSpace students are different. How do we test it?

$$H_0: \ X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} = \mu_0$$

'HarbourSpace students' IQs are the same as those of the general population'

$$H_1: X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} > \mu_0$$

'HarbourSpace students' IQs are higher than those of the general population'

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.

- We suspect that HarbourSpace students are different. How do we test it?

$$H_0: \; X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} = \mu_0$$

'HarbourSpace students' IQs are the same as those of the general population'

$$H_1: X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} \neq \mu_0$$

'HarbourSpace students' IQs aren't the same as than those of the general population'

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.

- We suspect that HarbourSpace students are different. How do we test it?

$$H_0:\ X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} = \mu_0$$

'HarbourSpace students' IQs are the same as those of the general population'

$$H_1: X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} < \mu_0$$

'HarbourSpace students' IQs are lower than those of the general population'

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.

- We suspect that HarbourSpace students are different. How do we test it?

$$H_0:\ X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} = \mu_0$$

'HarbourSpace students' IQs are the same as those of the general population'

$$H_1: X \sim N(\mu_{HS}, \sigma^2), \qquad \mu_{HS} \neq \mu_0$$

'HarbourSpace students' IQs aren't the same as than those of the general population'

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.
- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.
- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\qquad$ $H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Average IQ of n students: $\bar{X}$

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.
- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Average IQ of n students: $\bar{X}$
- Significance level α

z-test

- IQ follows $N(\mu_0, \sigma^2)$ where σ is known.
- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\qquad$ $H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Average IQ of n students: $\bar{X}$
- Significance level α
- Can we reject H_0?

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$

- Test statistic:

z-test

- $H_0 \colon X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \qquad H_1 \colon X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n} \sum\nolimits_{i=1}^{n} X_i$$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Assuming H_0, $X_i \sim$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\quad$ $H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Assuming H_0, $X_i \sim N(\mu_0, \sigma^2)$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Assuming H_0, $X_i \sim N(\mu_0, \sigma^2)$

$$\Rightarrow \qquad\qquad \sim N(0,1)$$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Assuming $H_0,\ \ X_i \sim N(\mu_0, \sigma^2)$

$$\Rightarrow \qquad \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1\,)$$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Assuming $H_0, \quad X_i \sim N(\mu_0, \sigma^2)$

$$\Longrightarrow T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1\ )$$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \qquad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Assuming H_0, $\ X_i \sim N(\mu_0, \sigma^2)$

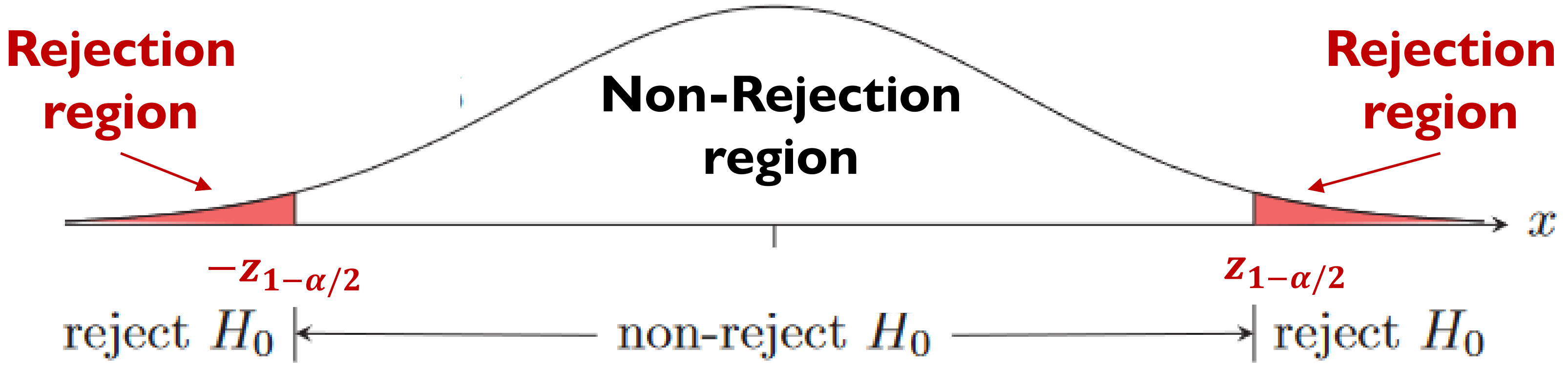
$$\Rightarrow T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1\ )$$

z-test

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \quad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic: $T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1\ )$
- Significance level α
- (non-) Rejection region:

z-test

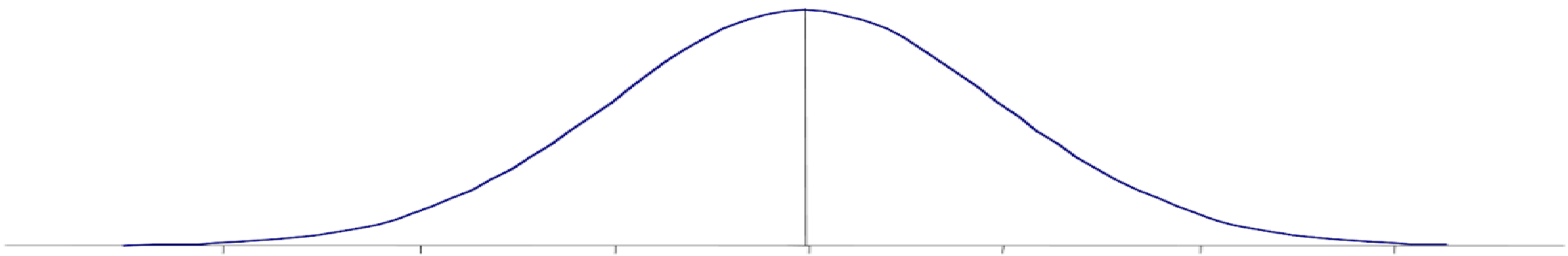
- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \quad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu \neq \mu_0$
- Test statistic: $T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1)$
- Significance level α
- (non-) Rejection region:



ONE-SIDED ALTERNATIVES

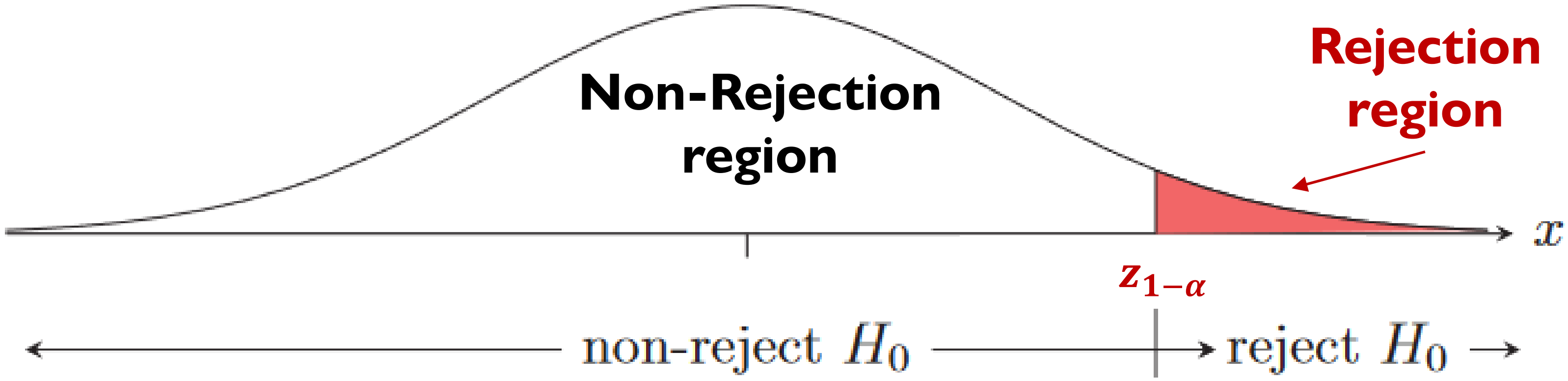
z-test: ONE-SIDED

- $H_0 : X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0 \quad H_1 : X \sim N(\mu_{HS}, \sigma^2),\ \mu > \mu_0$
- Test statistic: $T(X_1, \ldots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1)$
- Significance level α
- (non-) Rejection region:



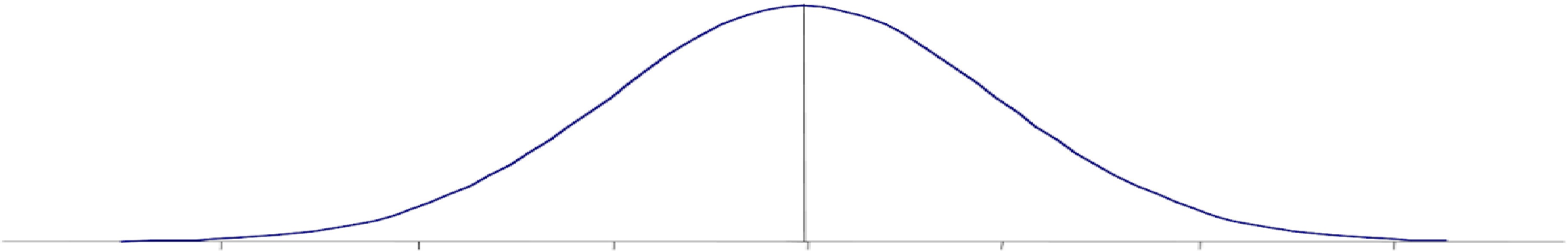
z-test: ONE-SIDED

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\quad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu > \mu_0$
- Test statistic: $T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1)$
- Significance level α
- (non-) Rejection region:



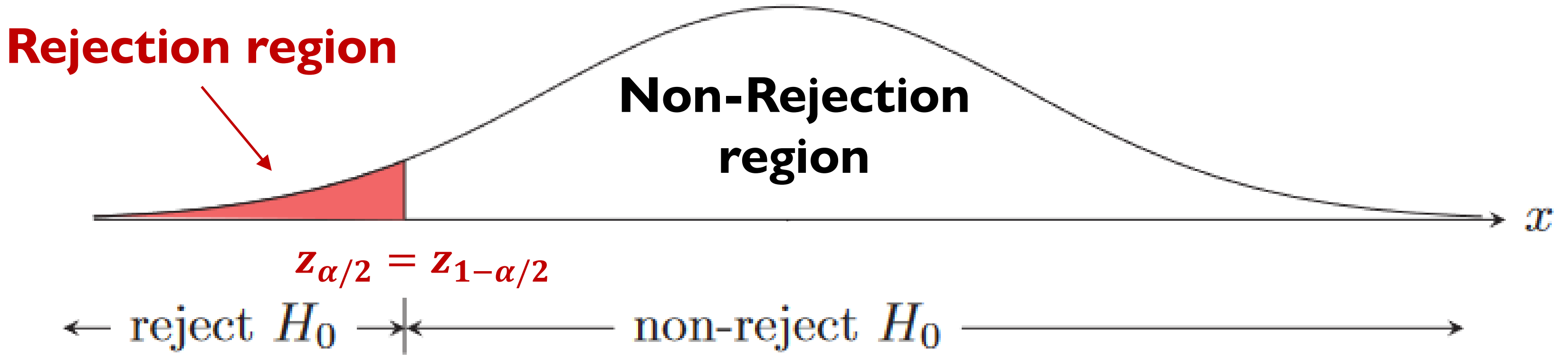
z-test: ONE-SIDED

- $H_0: X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\quad H_1: X \sim N(\mu_{HS}, \sigma^2),\ \mu < \mu_0$
- Test statistic: $T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1)$
- Significance level α
- (non-) Rejection region:



z-test: ONE-SIDED

- $H_0 : X \sim N(\mu_{HS}, \sigma^2),\ \mu = \mu_0$ $\quad H_1 : X \sim N(\mu_{HS}, \sigma^2),\ \mu < \mu_0$
- Test statistic: $T(X_1, \ldots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sigma} \sim N(0,1)$
- Significance level α
- (non-) Rejection region:



Z-TEST: EXAMPLE 1

- Weight loss is described by $X \sim N(5, 7^2)$.
- $n = 30$ patients followed an experimental diet.
- Average weight loss $\bar{X} = 6.1$ kg.

Does the diet make any difference?

Z-TEST: EXAMPLE 2

- A bakery supplies loaves of bread to supermarkets. Not every loaf weighs the same: loaf weight $X \sim N(\mu_0, 0.1^2)$, supermarket expects $\mu_0 = 2$ kg.
- A supermarket draws a sample of $n = 20$ loaves: average weights is $\bar{X} = 1.97$ kg.
- The supermarket wants to be sure that the weights are, on average, not lower than 2 kg.

Is there evidence against this?

T-TEST

When the data is coming from normal distribution
and variance is unknown

T-TEST

- Hypothesis about μ from $X \sim N(\mu_0, \sigma^2)$, σ is unknown.

T-TEST

- Hypothesis about μ from $X \sim N(\mu_0, \sigma^2), \quad \sigma$ is unknown.
- Unknown variance is estimated from the sample:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

T-TEST

- Hypothesis about μ from $X \sim N(\mu_0, \sigma^2), \quad \sigma$ is unknown.
- Unknown variance is estimated from the sample:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- The test statistic is therefore:

$$T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{s} \sim$$

T-TEST

- Hypothesis about μ from $X \sim N(\mu_0, \sigma^2)$, $\quad \sigma$ is unknown.
- Unknown variance is estimated from the sample:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- The test statistic is therefore:

$$T(X_1, \dots, X_n) = \frac{(\bar{X} - \mu_0)\sqrt{n}}{s} \sim t(n-1)$$

T-TEST: EXAMPLE 1

- A bakery supplies loaves of bread to supermarkets. Not every loaf weighs the same: loaf weight $X \sim N(\mu_0, \sigma^2)$, supermarket expects $\mu_0 = 2$ kg.

- A supermarket draws a sample of $n = 20$ loaves: average weights is $\bar{X} = 1.9668$ kg, $s^2 = 0.0927^2$

- The supermarket wants to be sure that the weights are, on average, not lower than 2 kg.

Is there evidence against this?

T-TEST: EXAMPLE 2

- Sales at your company: $X \sim N(\mu_0, \sigma^2)$, $\mu_0 = 100$ dollars per transaction.

- You hire new employees. Based of a sample of $\mathrm{n} = 25$ salesmen, average sale is $\bar{X} = 130$ dollars, with sample standard deviation $\mathrm{s} = 15$.

Are the new employees better than the old ones?
Test your hypothesis at $\boldsymbol{\alpha = 0.05}$.

TO SUM UP

- Hypothesis testing ingredients
  - H0 and H1
  - Test statistic
  - Null distribution
  - Rejection region
  - Two- and one-sided alternatives

TO SUM UP

- Hypothesis testing ingredients
  - H0 and H1
  - Test statistic
  - Null distribution
  - Rejection region
  - Two- and one-sided alternatives

- One-sample tests:
  - Z-test: normal distribution, variance known
  - T-test: normal distribution, variance unknown