How To Determine the Result of the Test

Plenty of evidence against H_0 → Reject H_0 (and accept H_1)

How To Determine the Result of the Test

Plenty of evidence against H_0 → Reject H_0 (and accept H_1)

"Earn extra cash"

"Risk free"

"Dear friend"

"Act immediately"

"Apply now"

"Winner"

How To Determine the Result of the Test

Plenty of evidence against H_0 → Reject H_0 (and accept H_1)

"Earn extra cash"

"Risk free"

"Dear friend"

"Act immediately"

"Apply now"

"Winner"

Very unlikely under H_0 (regular email) → Reject H_0 and send email to spam

DeepLearning.AI

Hypothesis Testing

Type I and Type II errors

Sometimes Things Go Wrong…

Sometimes Things Go Wrong…

What if I make the wrong decision?

Sometimes Things Go Wrong…

What if I make the wrong decision?

Type I error
(False positive)

Sometimes Things Go Wrong…

What if I make the wrong decision?

Type I error
(False positive)

Type II error
(False negative)

Sometimes Things Go Wrong…

What if I make the wrong decision?

Type I error
(False positive)

Type II error
(False negative)

Sometimes Things Go Wrong…

What if I make the wrong decision?

Type I error
(False positive)

Type II error
(False negative)

Type I and Type II Errors

Type I and Type II Errors

Decision	Reality	

DeepLearning.AI

Type I and Type II Errors

Decision	Reality	
	H_0 True (Innocent)	
Reject H_0 (Decide Guilty)	Type I error	

Type I and Type II Errors

Decision	Reality	
	H_0 True (Innocent)	
Reject H_0 (Decide Guilty)	Type I error	
Don't reject H_0 (Decide not guilty)	Correct	

Type I and Type II Errors

Decision	**Reality**	
	H_0 True (Innocent)	H_0 False (Guilty)
Reject H_0 (Decide Guilty)	Type I error	Correct
Don’t reject H_0 (Decide not guilty)	Correct	

Type I and Type II Errors

Decision	Reality	
	H_0 True (Innocent)	H_0 False (Guilty)
Reject H_0 (Decide Guilty)	Type I error	Correct
Don’t reject H_0 (Decide not guilty)	Correct	Type II error

Significance Level

Significance Level

Type I error

Type II error

Significance Level

The presumption of innocence implies that sending an innocent person to prison is worse than letting a criminal walk

Type I error

Type II error

Significance Level

The presumption of innocence implies that sending an innocent person to prison is worse than letting a criminal walk

Type I error

Type II error

What is the greatest probability of type I error you are willing to tolerate?

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Type I error

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Type I error

Innocent person determined guilty

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Significance level

Type I error

Innocent person determined guilty

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Significance level (α)

Type I error

Innocent person determined guilty

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Significance level (α)

Type I error

Innocent person determined guilty

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Significance level (α)

Type I error

Innocent person determined guilty

0 — 1 α

Defendants are always considered 'not guilty' → No Type I error

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Significance level (α)

Type I error

Innocent person determined guilty

Significance Level

What is the greatest probability of type I error you are willing to tolerate?

Type I error

Innocent person determined guilty

Significance level (α)

Every defendant is considered guilty

Every time the defendant is innocent you make a Type I error

Significance Level

Significance level (α)

Defendants are always considered 'not guilty'

No Type I error

Every defendant is considered guilty

Every time the defendant is innocent you make a Type I error

Significance Level

Significance level (α)

0 ——— 1 → α

Defendants are always considered 'not guilty'

↓

No Type I error

Both options are awful decision makers

Every defendant is considered guilty

↓

Every time the defendant is innocent you make a Type I error

Significance Level

Significance level (α)

$\alpha = 0.05$

0 — 1 — α

Defendants are always considered 'not guilty'

No Type I error

Both options are awful decision makers

You want a small α, but not zero

Every defendant is considered guilty

Every time the defendant is innocent you make a Type I error

Significance Level

Significance level (α)

Defendants are always considered 'not guilty'

No Type I error

Both options are awful decision makers

You want a small α, but not zero

Catch:

↓Type I error ↑Type II error

Every defendant is considered guilty

Every time the defendant is innocent you make a Type I error

Significance Level

Significance Level

Type I error

Innocent person determined guilty

Significance Level

Type I error

Innocent person determined guilty

$$\alpha = \max \mathbf{P}(\text{Type I error})$$

Significance Level

Type I error

Innocent person determined guilty

$$\alpha = \max \mathbf{P}(\text{Type I error})$$
$$= \max \mathbf{P}(\text{Reject } H_0 \mid H_0)$$

Significance Level

Type I error

Innocent person determined guilty

$$\alpha = \max \mathbf{P}(\text{Type I error})$$
$$= \max \mathbf{P}(\text{Reject } H_0 \,|\, H_0)$$

The value of α is your criteria for designing your test

Significance Level

Type I error

Innocent person determined guilty

$$\alpha = \max \mathbf{P}(\text{Type I error})$$
$$= \max \mathbf{P}(\text{Reject } H_0 \,|\, H_0)$$

The value of α is your criteria for designing your test

For a given sample, α will determine if you reject H_0 or not

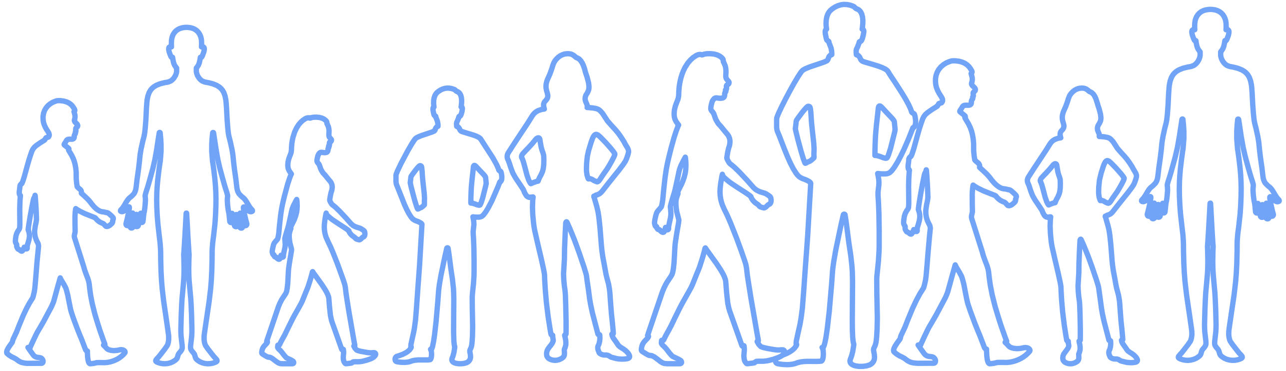
DeepLearning.AI

Hypothesis Testing

Right-Tailed, Left-Tailed and Two-Tailed tests

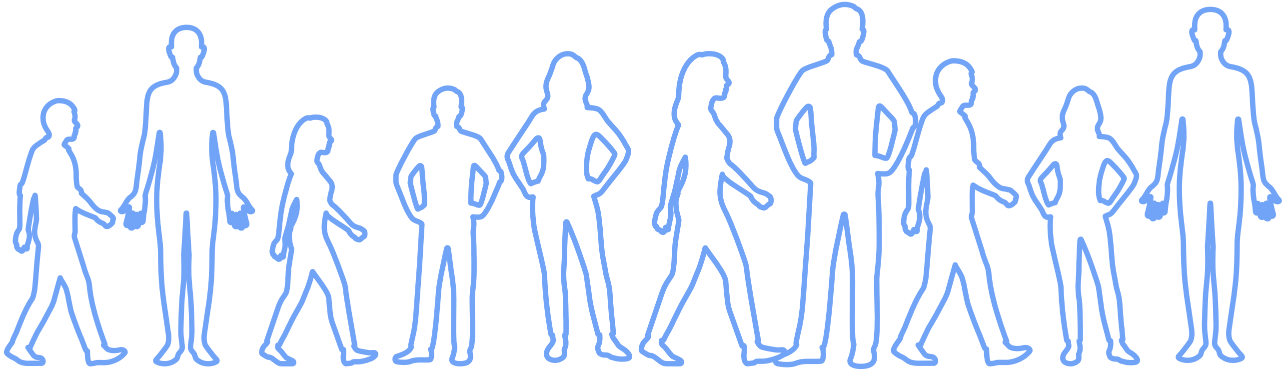
Example: Heights

Example: Heights



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

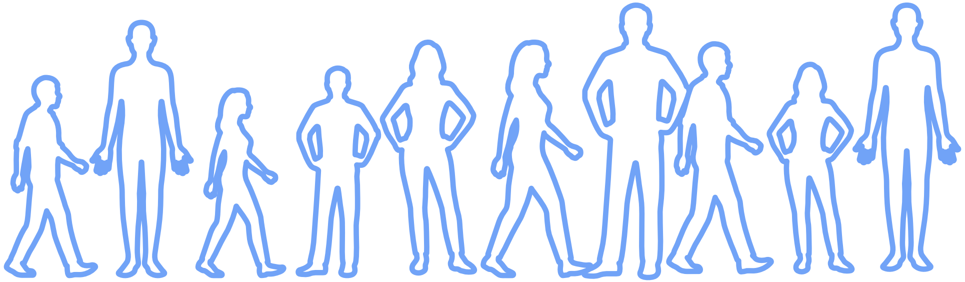
Example: Heights



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$$\bar{x} = 68.442$$

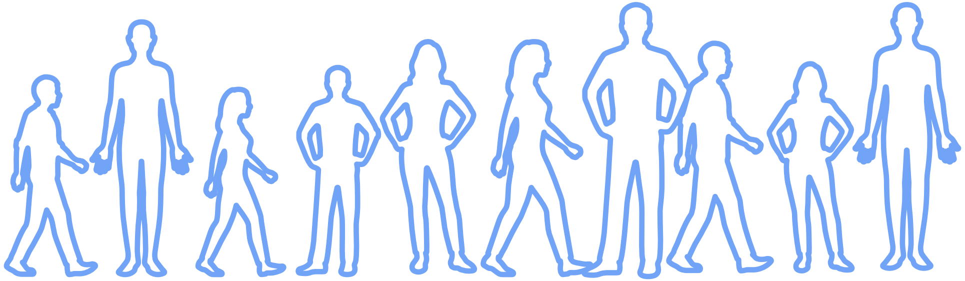
Data Quality



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$$\bar{x} = 68.442$$

Data Quality

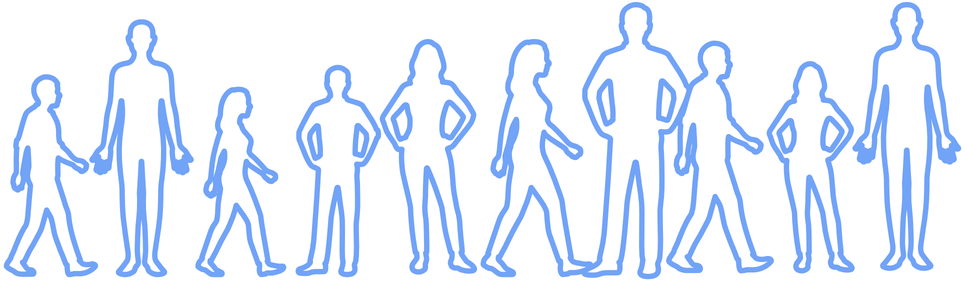


x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$$\bar{x} = 68.442$$

Reliable

Data Quality



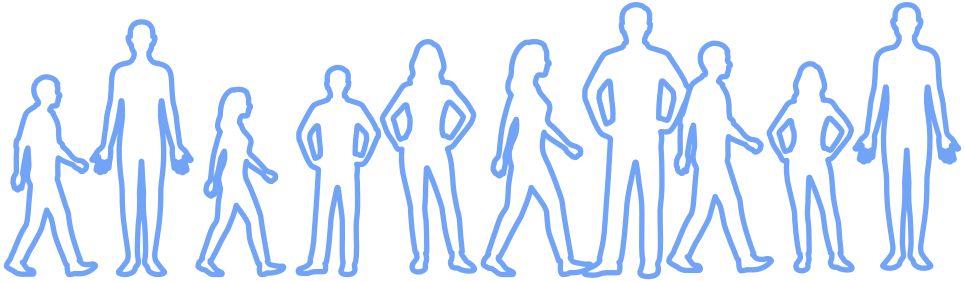
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

Reliable

Representative

Data Quality



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

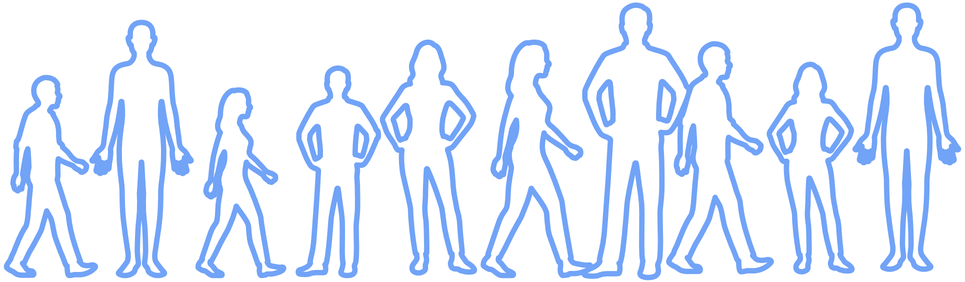
$$\bar{x} = 68.442$$

Reliable

Representative

Randomized

Data Quality



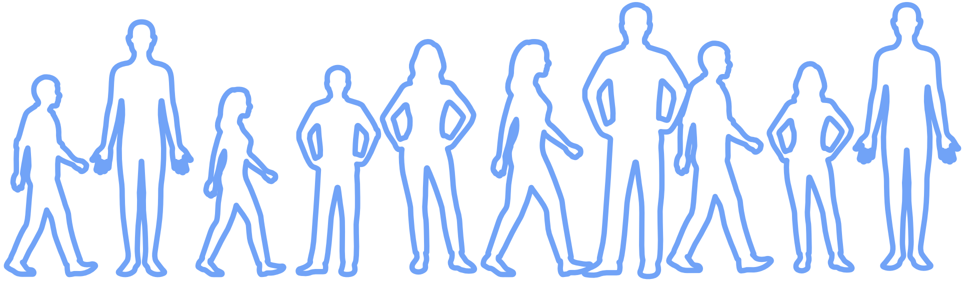
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$$\bar{x} = 68.442$$

Reliable

- Representative
- Randomized
- Sample size

Determining the Hypothesis



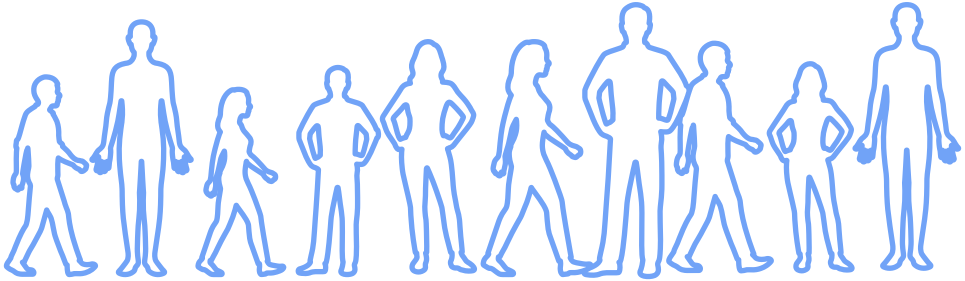
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

vs. $H_1 : \mu > 66.7$

Population

Determining the Hypothesis



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

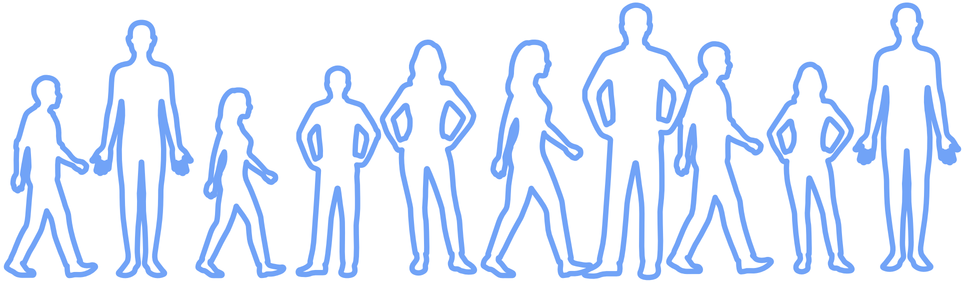
$\bar{x} = 68.442$

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Population

Test Statistic



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

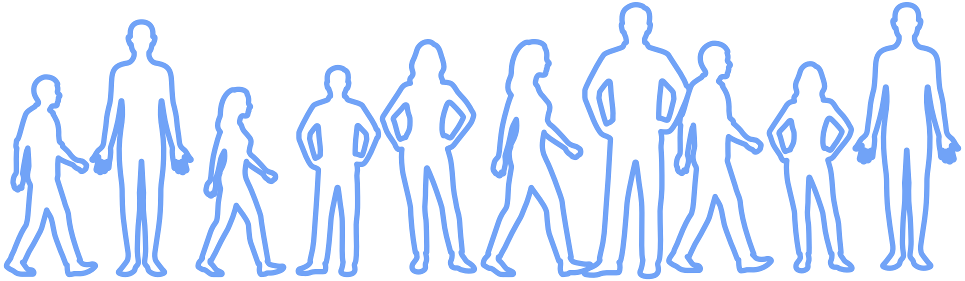
$\bar{x} = 68.442$

Observed statistic

vs. $H_1 : \mu > 66.7$

Test statistic $\longrightarrow \bar{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$

Test Statistic



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

Observed statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

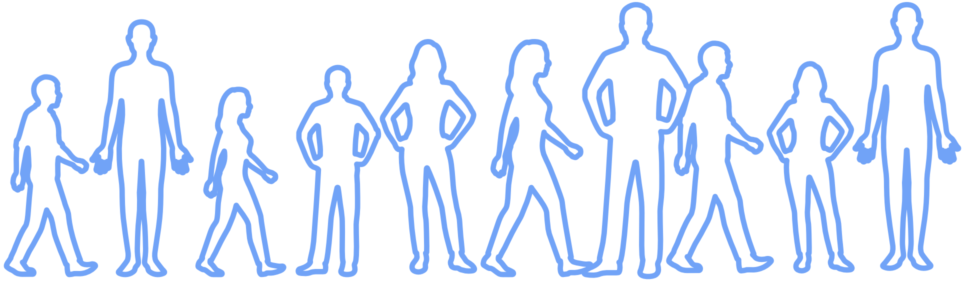
$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Test statistic $\longrightarrow \bar{X} = \dfrac{1}{10} \sum_{i=1}^{10} X_i$

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10} \sum_{i=1}^{10} X_i$$



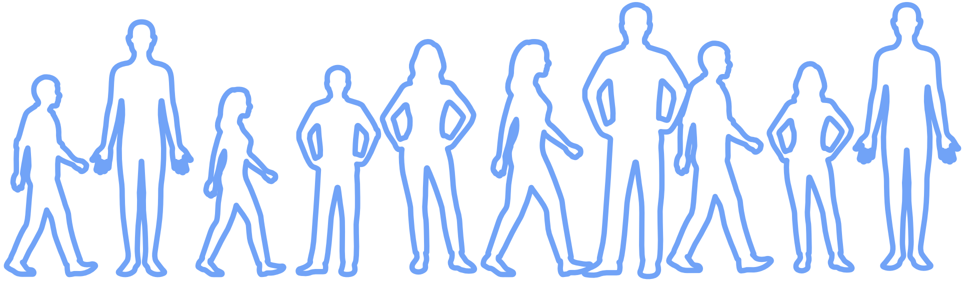
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$$\bar{x} = 68.442$$

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

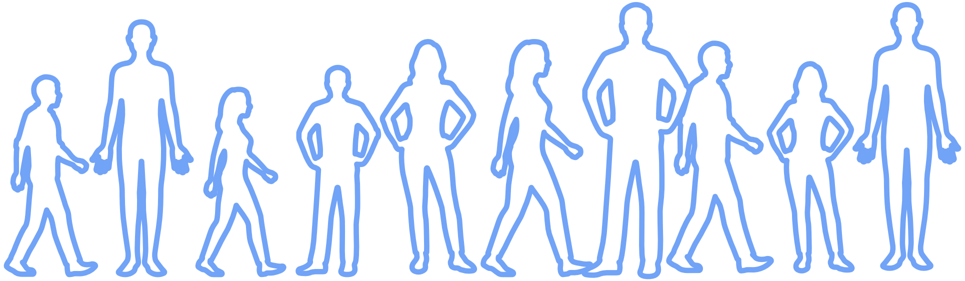
$\bar{x} = 68.442$

Test statistic: $T(\boldsymbol{X})$ $\quad \boldsymbol{X} = (X_1, \ldots, X_n)$

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

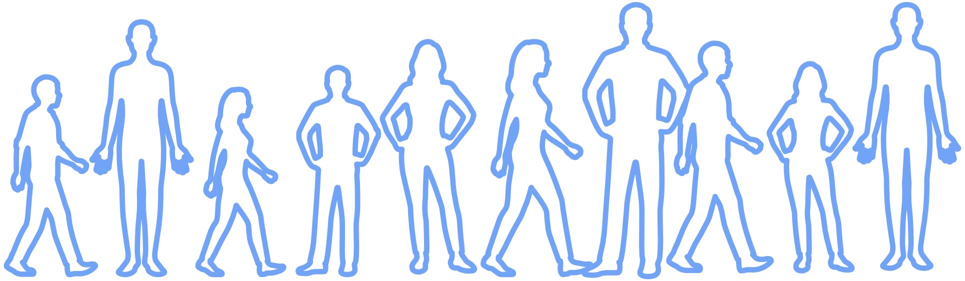
Test statistic: $T(\boldsymbol{X}) \quad \boldsymbol{X} = (X_1, \ldots, X_n)$

Information about the population parameter under study

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

Test statistic: $T(\boldsymbol{X})$ $\quad \boldsymbol{X} = (X_1, \ldots, X_n)$

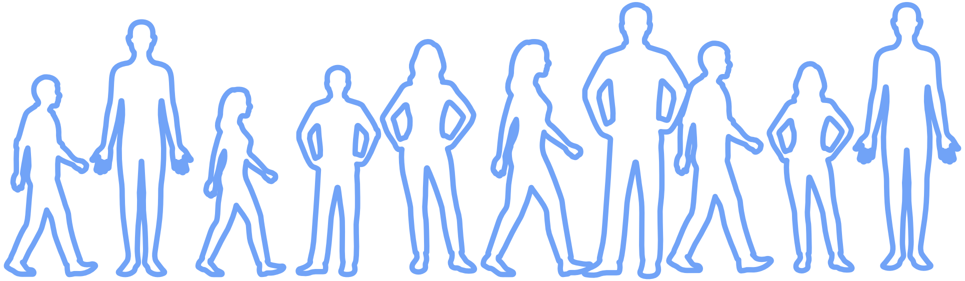
Information about the population parameter under study

$\mu \to \overline{X}$

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

Test statistic: $T(\boldsymbol{X})$ $\quad \boldsymbol{X} = (X_1, \ldots, X_n)$

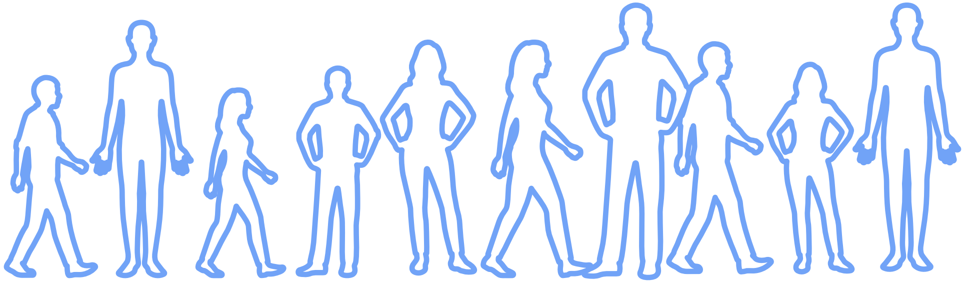
Information about the population parameter under study

$\mu \rightarrow \overline{X}$ $\qquad$ $p \rightarrow \overline{X}$

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

Test statistic: $T(\boldsymbol{X})$ $\quad \boldsymbol{X} = (X_1, \ldots, X_n)$

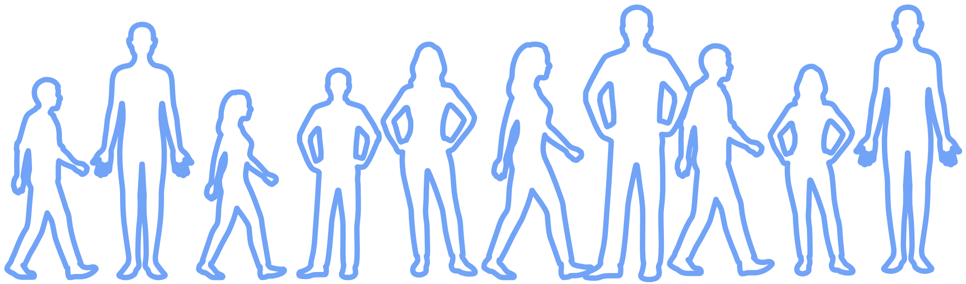
Information about the population parameter under study

$$\mu \rightarrow \overline{X} \qquad p \rightarrow \overline{X} \qquad \sigma^2 \rightarrow S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

Test Statistic

Test statistic

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$$



x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
66.75	70.24	67.19	67.09	63.65	64.64	69.81	69.79	73.52	71.74

$\bar{x} = 68.442$

Test statistic: $T(\boldsymbol{X}) \quad \boldsymbol{X} = (X_1, \ldots, X_n)$

Information about the population parameter under study

$\mu \rightarrow \overline{X}$ $\quad$ $p \rightarrow \overline{X}$ $\quad$ $\sigma^2 \rightarrow S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

Not unique!

Example: Heights

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions

3 sets of hypothesis

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions

3 sets of hypothesis

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions

3 sets of hypothesis

Right-Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

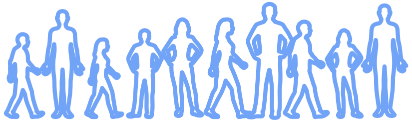
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions

3 sets of hypothesis

Right-Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



→ $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions

3 sets of hypothesis

Right-Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Left Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

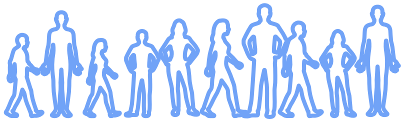
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions | **3 sets of hypothesis**

Right-Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Left Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

 → $H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

3 questions

3 sets of hypothesis

Right-Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Left Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Two-Tailed Test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Example: Heights

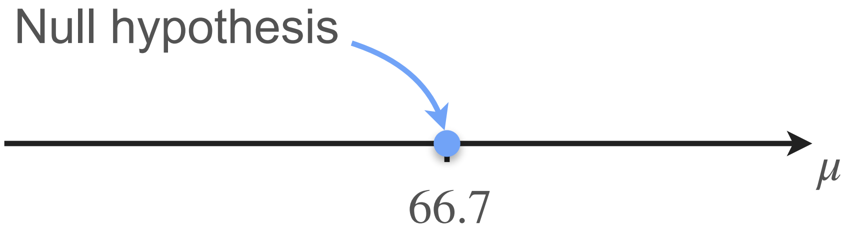
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Right-tailed test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Right-tailed test $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



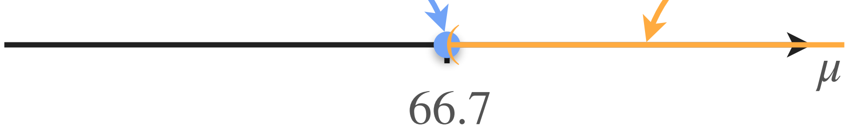
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Right-tailed test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Null hypothesis

Alternative hypothesis



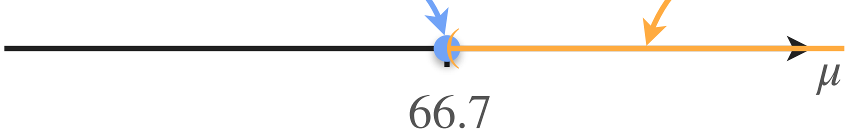
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Right-tailed test → $H_0 : \mu \leq 66.7$ vs. $H_1 : \mu > 66.7$

Null hypothesis

Alternative hypothesis

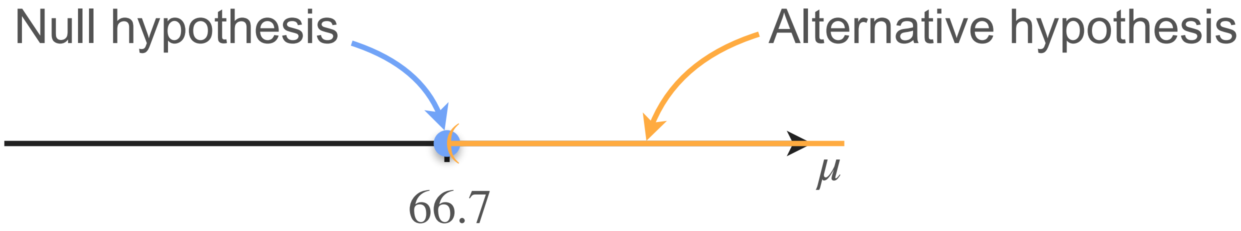


Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$ Test statistic

Right-tailed test $\longrightarrow$ $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



Example: Heights

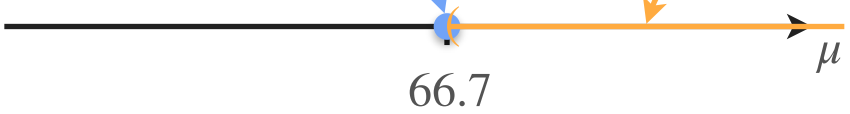
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$ Test statistic

Right-tailed test $\longrightarrow$ $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Null hypothesis

Alternative hypothesis



If $\overline{x} \gg 66.7 \Rightarrow$ Reject H_0

Example: Heights

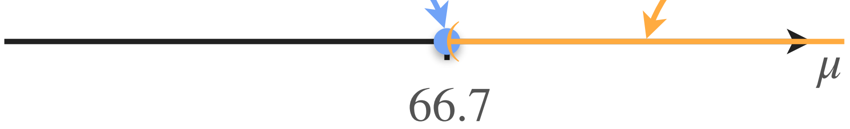
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$ Test statistic

Right-tailed test $\rightarrow$ $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Null hypothesis

Alternative hypothesis



Type I error: Determine $\mu > 66.7$, when population mean did not change

If $\overline{x} \gg 66.7 \Rightarrow$ Reject H_0

Example: Heights

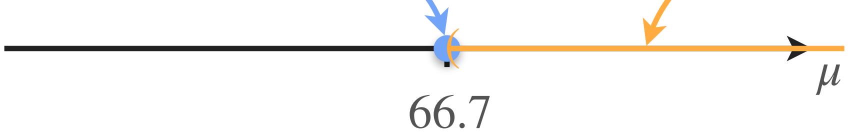
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$ Test statistic

Right-tailed test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Null hypothesis

Alternative hypothesis



If $\bar{x} \gg 66.7 \Rightarrow$ Reject H_0

Type I error: Determine $\mu > 66.7$, when population mean did not change

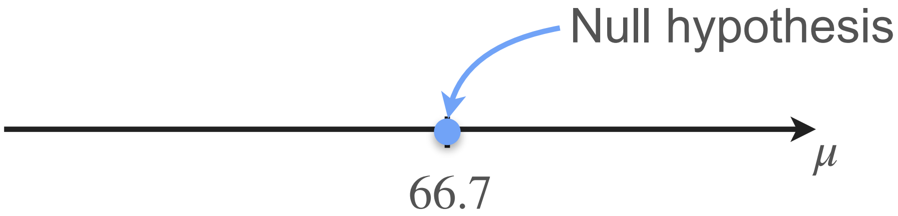
Type II error: Do not reject that $\mu = 66.7$ when in true $\mu > 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Type I error:

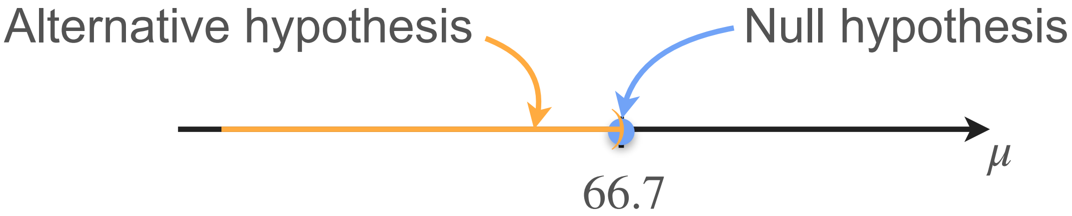
Type II error:

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Left tailed test $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Type I error:

Type II error:

Example: Heights

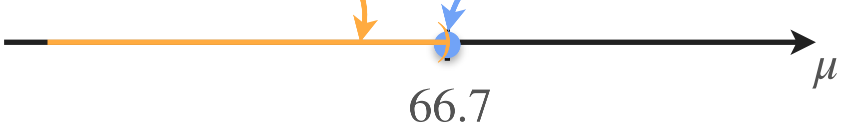
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Left tailed test $H_0 : \mu \geq 66.7$ vs. $H_1 : \mu < 66.7$

Alternative hypothesis

Null hypothesis



Type I error:

Type II error:

Example: Heights

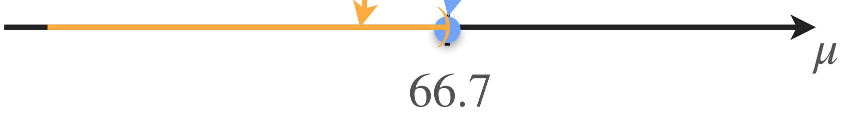
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Left tailed test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Alternative hypothesis

Null hypothesis



If $\overline{x} \ll 66.7 \Rightarrow$ Reject H_0

Type I error:

Type II error:

Example: Heights

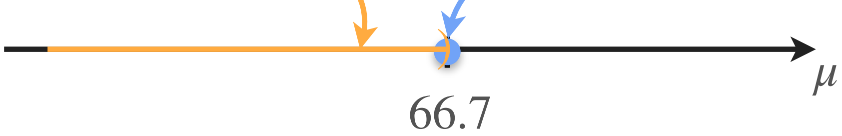
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Left tailed test $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Alternative hypothesis

Null hypothesis



If $\bar{x} \ll 66.7 \Rightarrow$ Reject H_0

Type I error: Determine $\mu < 66.7$, when population mean did not change

Type II error:

Example: Heights

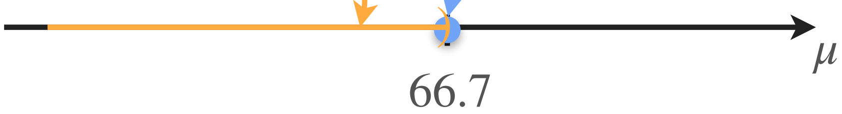
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Left tailed test $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Alternative hypothesis

Null hypothesis



If $\bar{x} \ll 66.7 \Rightarrow$ Reject H_0

Type I error: Determine $\mu < 66.7$, when population mean did not change

Type II error: Don't reject that $\mu = 66.7$ when true $\mu < 66.7$

Example: Heights

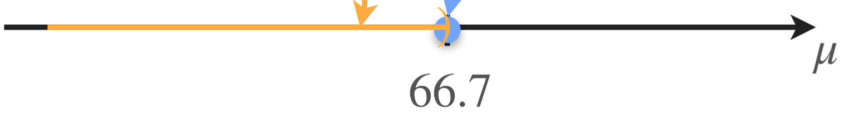
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$ Test statistic

Left tailed test $\longrightarrow$ $H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Alternative hypothesis

Null hypothesis



If $\bar{x} \ll 66.7 \Rightarrow$ Reject H_0

Type I error: Determine $\mu < 66.7$, when population mean did not change

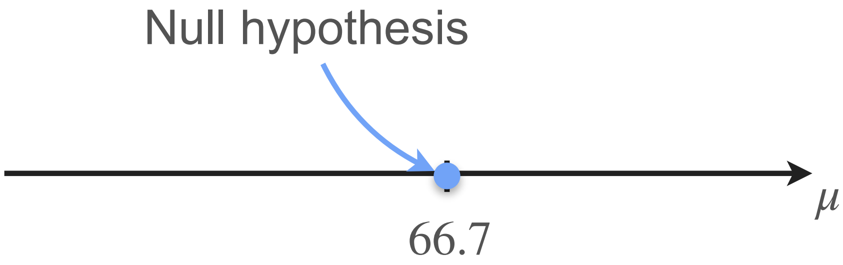
Type II error: Don't reject that $\mu = 66.7$ when true $\mu < 66.7$

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Type I error:

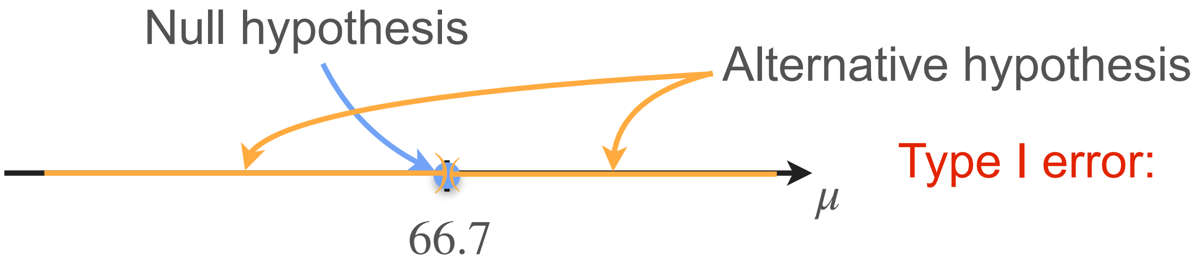
Type II error:

Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Two tailed test $H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Type I error:

Type II error:

Example: Heights

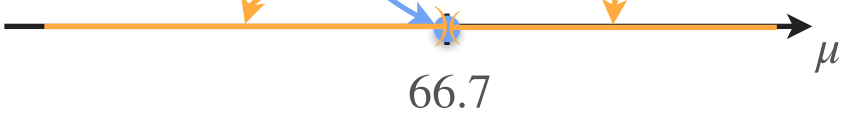
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Two tailed test $H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Null hypothesis

Alternative hypothesis



Type I error:

Type II error:

$$\text{If} \quad \begin{matrix} \bar{x} \gg 66.7 \\ \text{or} \\ \bar{x} \ll 66.7 \end{matrix} \quad \Rightarrow \quad \text{Reject } H_0$$

Example: Heights

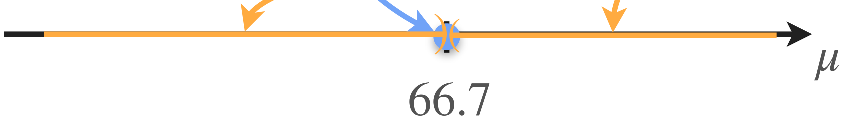
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\bar{X}$

Two tailed test → $H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Null hypothesis

Alternative hypothesis



$$\text{If} \quad \begin{matrix} \bar{x} \gg 66.7 \\ \text{or} \\ \bar{x} \ll 66.7 \end{matrix} \quad \Rightarrow \quad \text{Reject } H_0$$

Type I error: Determine $\mu \neq 66.7$, when population mean did not change

Type II error:

Example: Heights

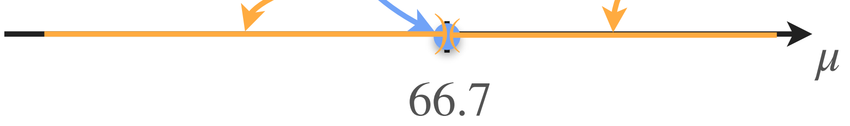
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$

Two tailed test $\quad H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Null hypothesis

Alternative hypothesis



$$\text{If} \quad \begin{array}{c} \bar{x} \gg 66.7 \\ \text{or} \\ \bar{x} \ll 66.7 \end{array} \quad \Rightarrow \quad \text{Reject } H_0$$

Type I error: Determine $\mu \neq 66.7$, when population mean did not change

Type II error: Don't reject that $\mu = 66.7$ when true $\mu \neq 66.7$

Example: Heights

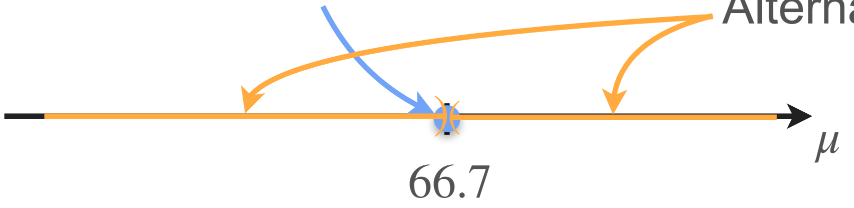
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\overline{X}$ Test statistic

Two tailed test $\rightarrow$ $H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Null hypothesis

Alternative hypothesis



If $\bar{x} \gg 66.7$ or $\bar{x} \ll 66.7$ $\Rightarrow$ Reject H_0

Type I error: Determine $\mu \neq 66.7$, when population mean did not change

Type II error: Don't reject that $\mu = 66.7$ when true $\mu \neq 66.7$

DeepLearning.AI

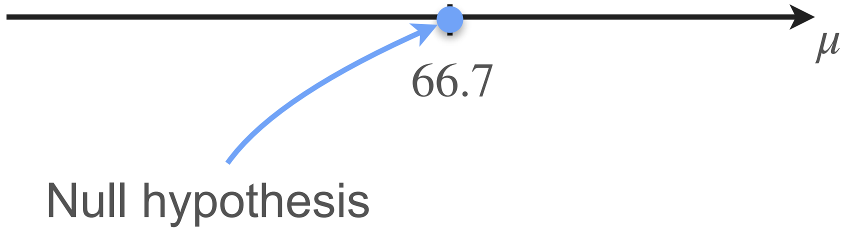
Hypothesis Testing

p-Value

Example: Heights

Example: Heights

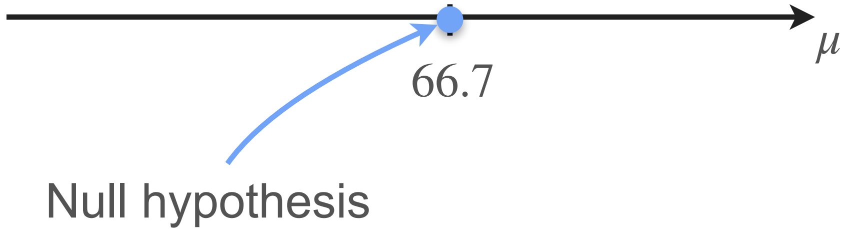
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.



Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$$\sigma = 3$$
$$n = 10$$



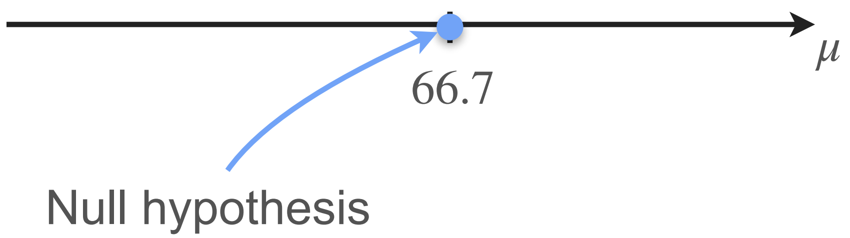
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

If H_0 is true:



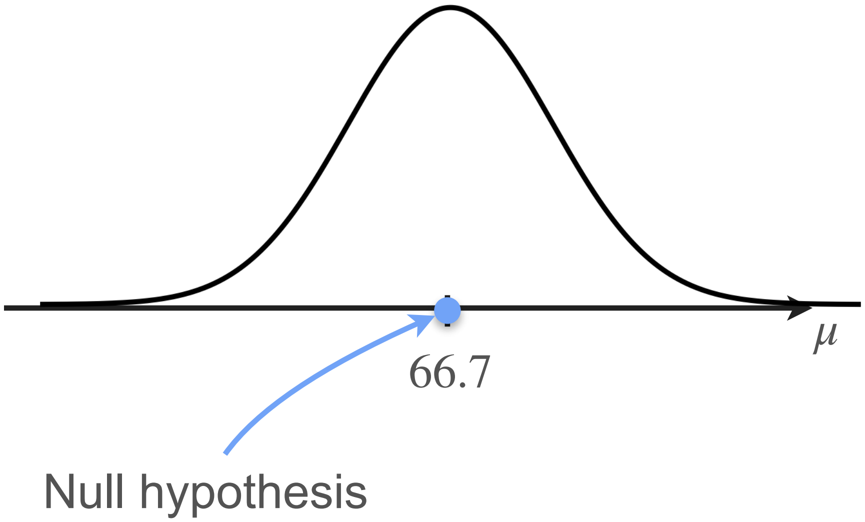
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

If H_0 is true: $\overline{X} \sim \mathcal{N}\left(\quad , \quad \right)$



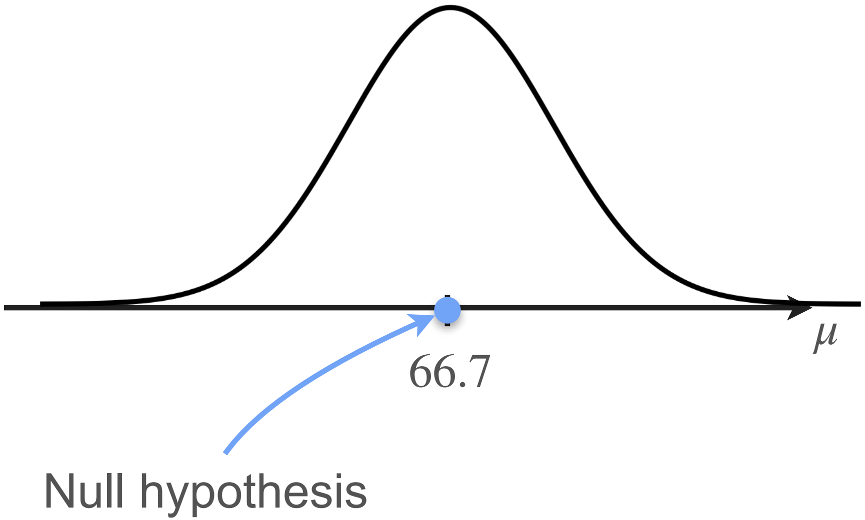
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

If H_0 is true: $\overline{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$



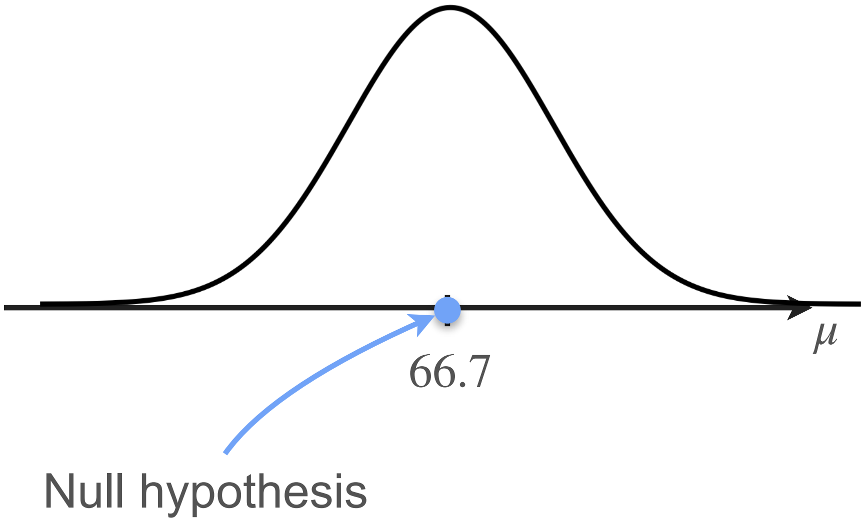
Example: Heights

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

If H_0 is true: $\overline{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$



How likely was your sample if H_0 is true?

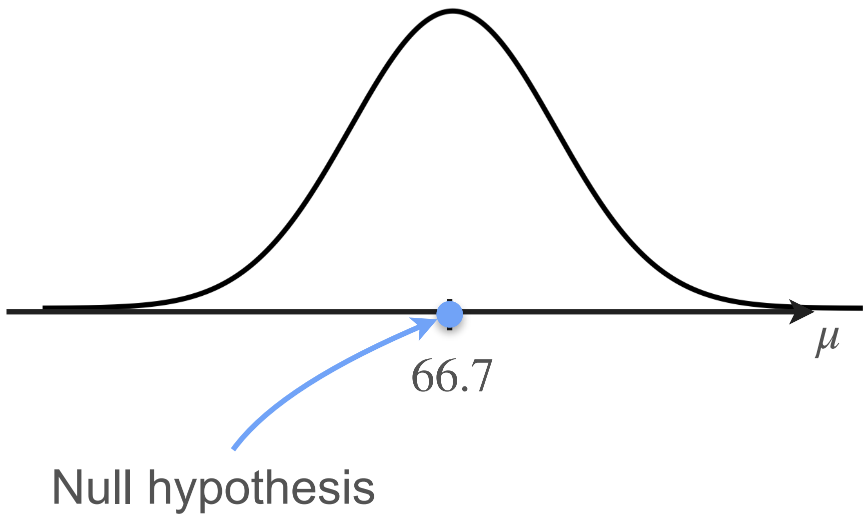
If the answer is very unlikely, then reject H_0

Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$



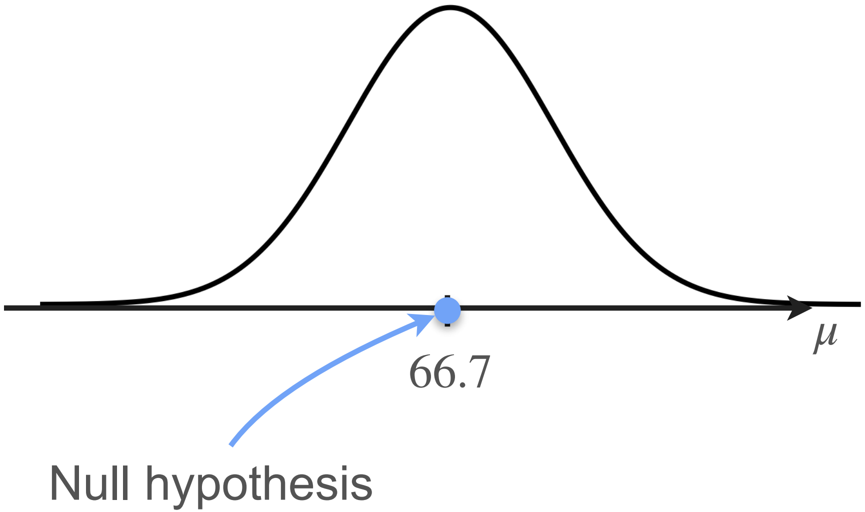
Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



Right-Tailed Test for Gaussian Data (Known σ)

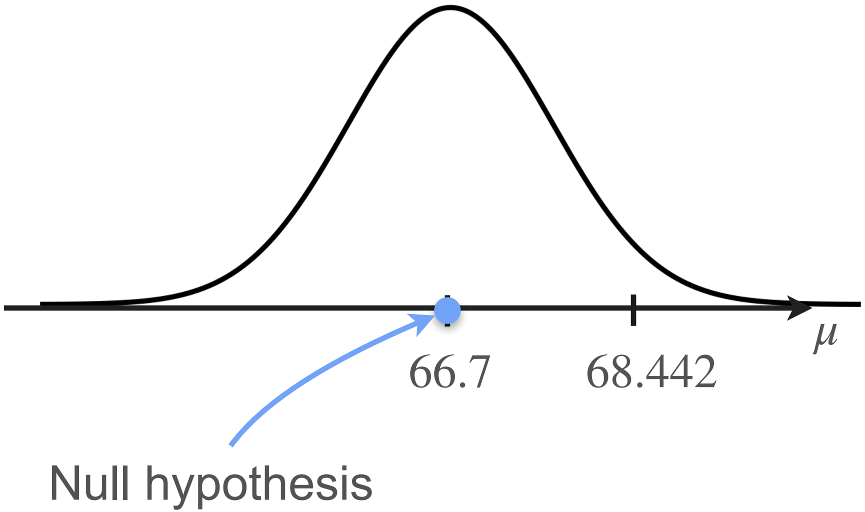
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

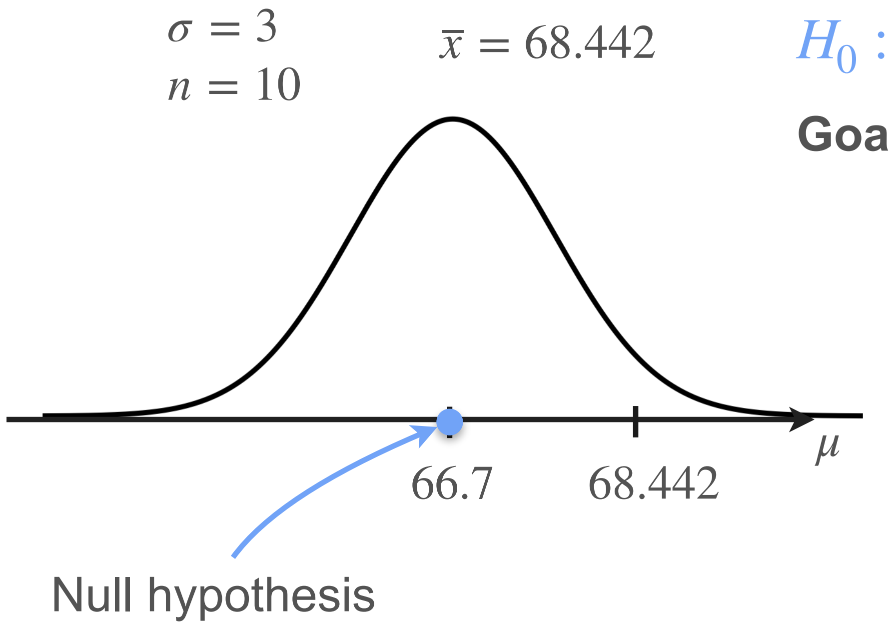


Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$



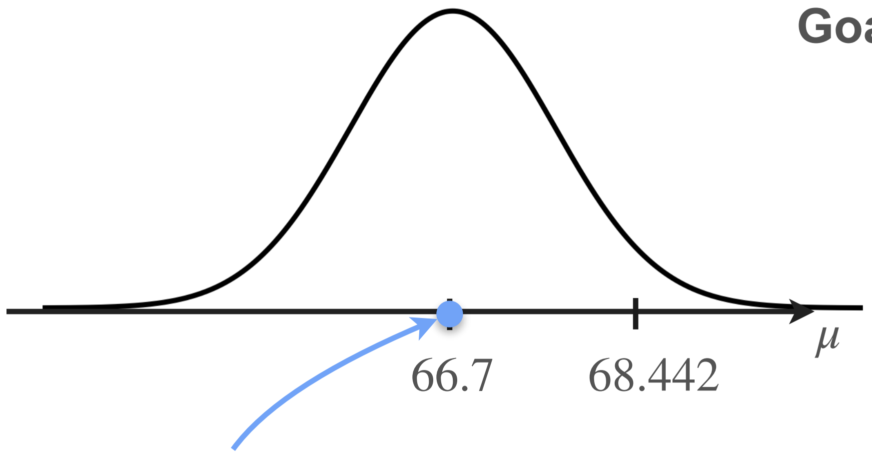
Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



Null hypothesis

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

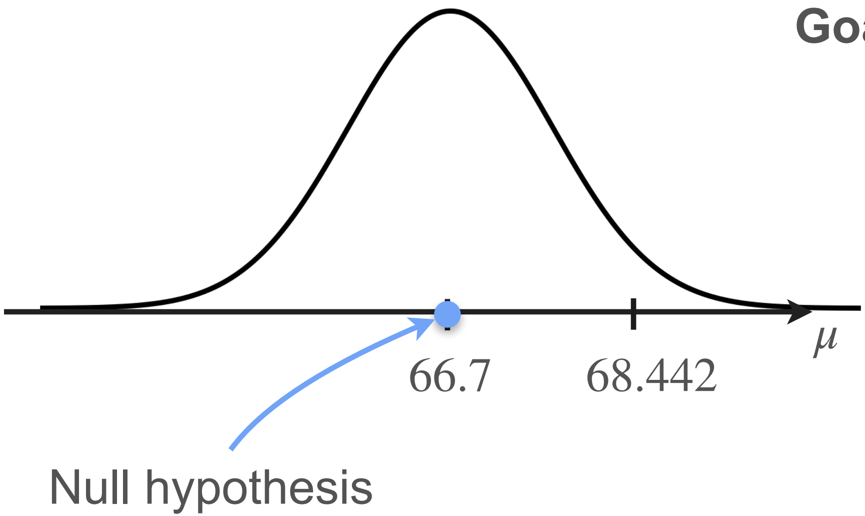
Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right)?$$

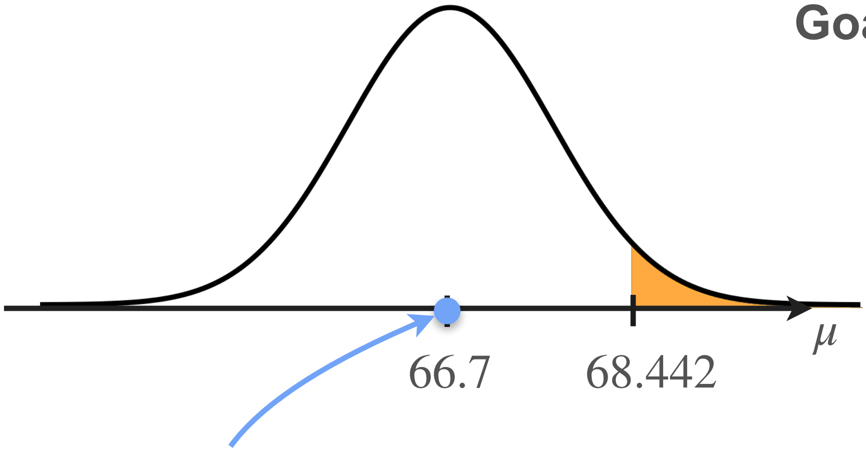
Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



Null hypothesis

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407$$

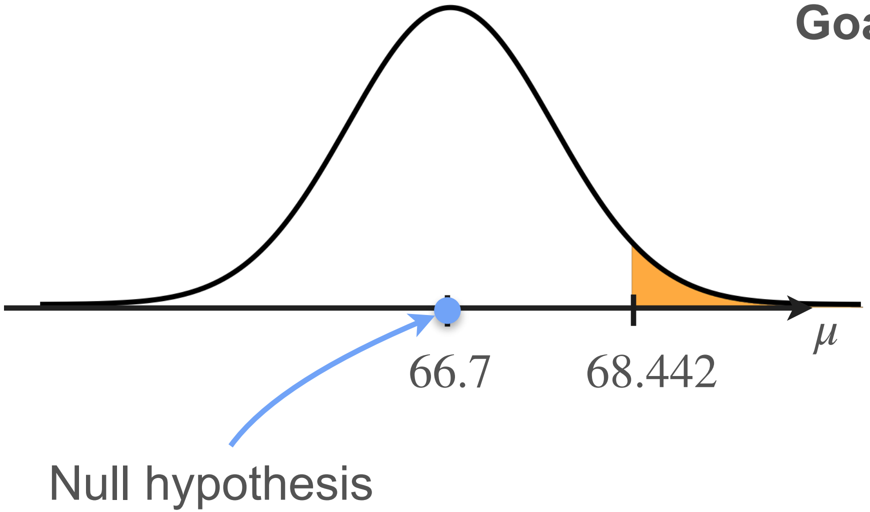
Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

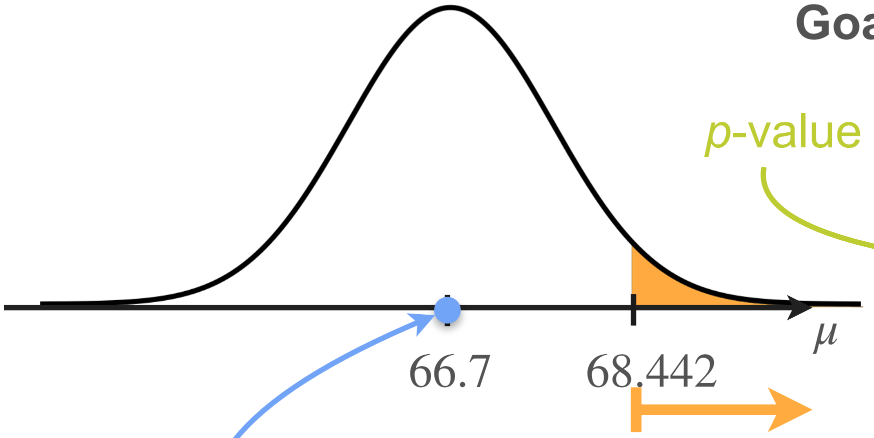
Right-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



Null hypothesis

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

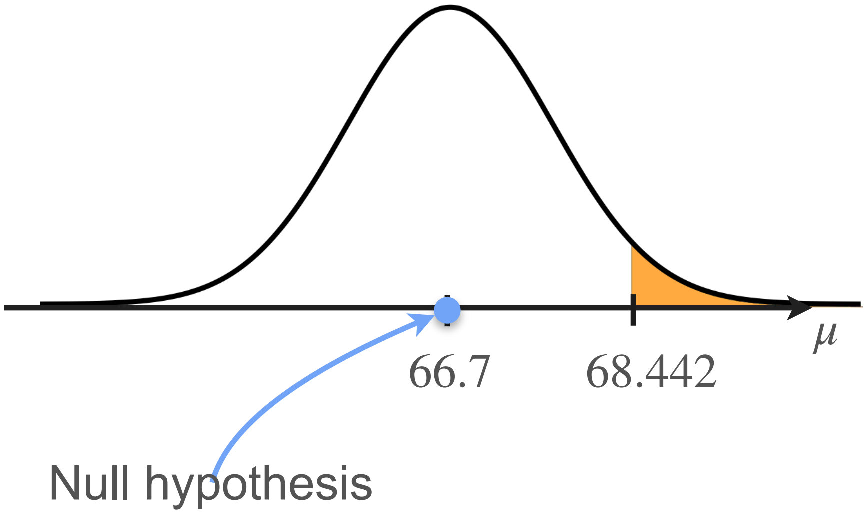
Type I error: Determine $\mu > 66.7$, when population mean did not change

p-value

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

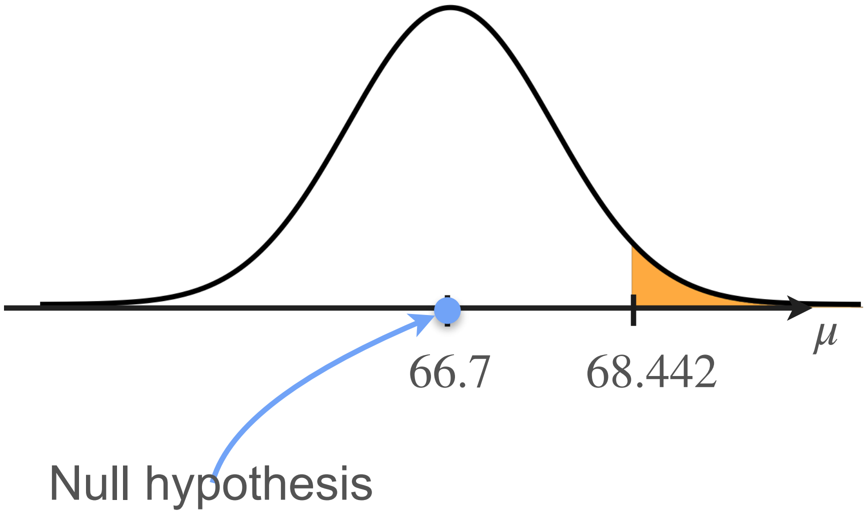
P-Values



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

P-Values

A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed



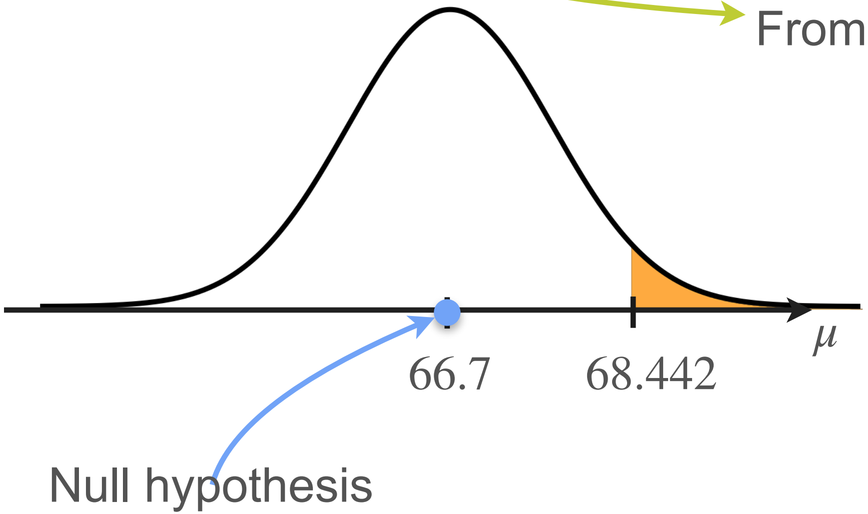
$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

P-Values

A **p-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

From the observed value to the direction of H_1

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

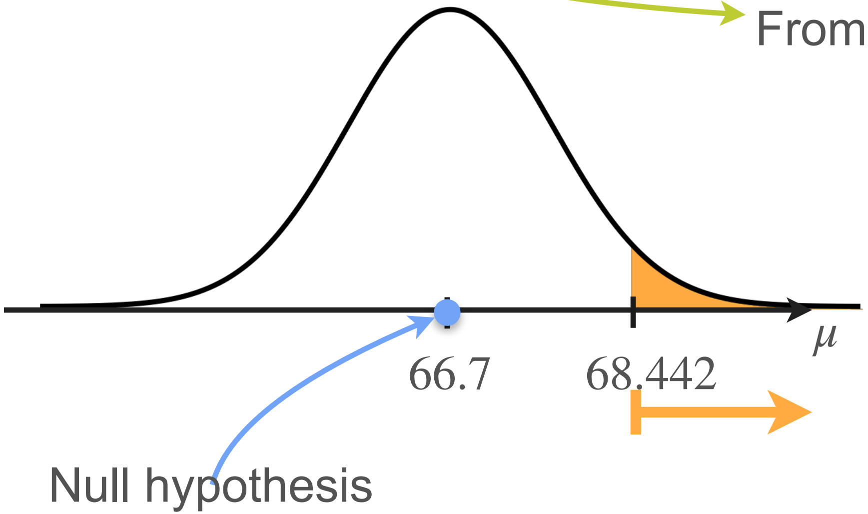


P-Values

A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

From the observed value to the direction of H_1

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



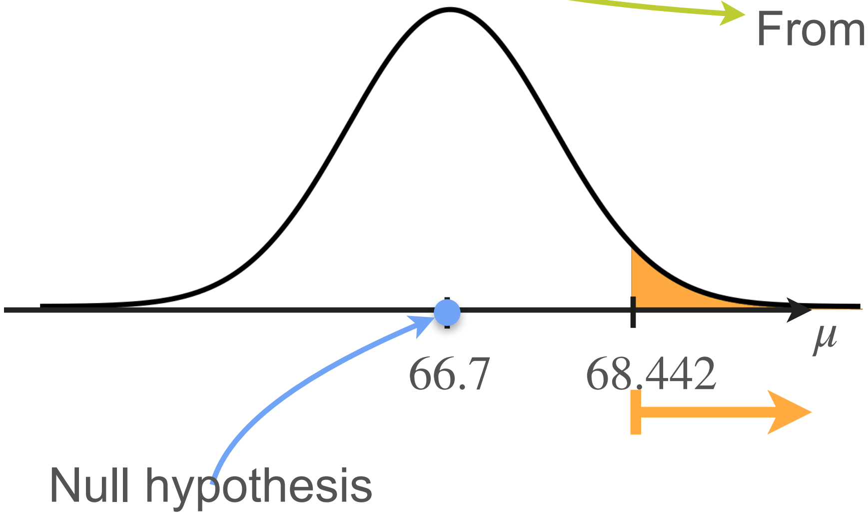
P-Values

A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

From the observed value to the direction of H_1

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

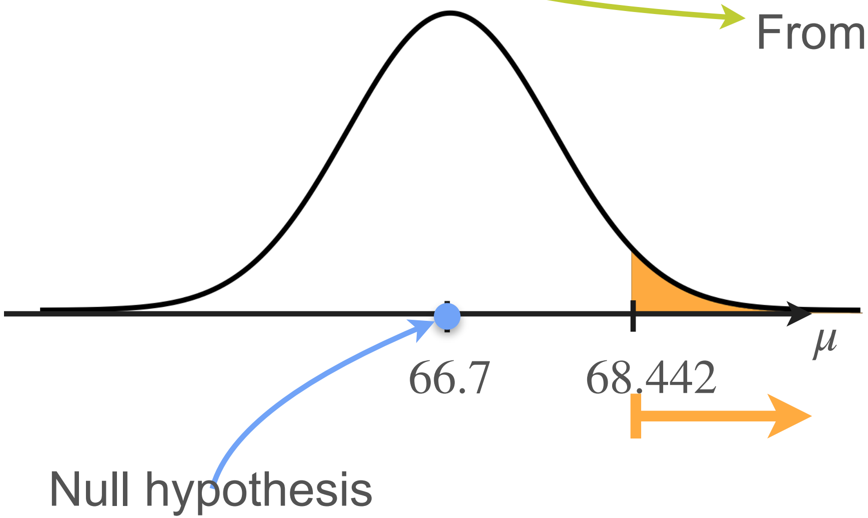
Decision rule:



P-Values

A **p-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

From the observed value to the direction of H_1



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

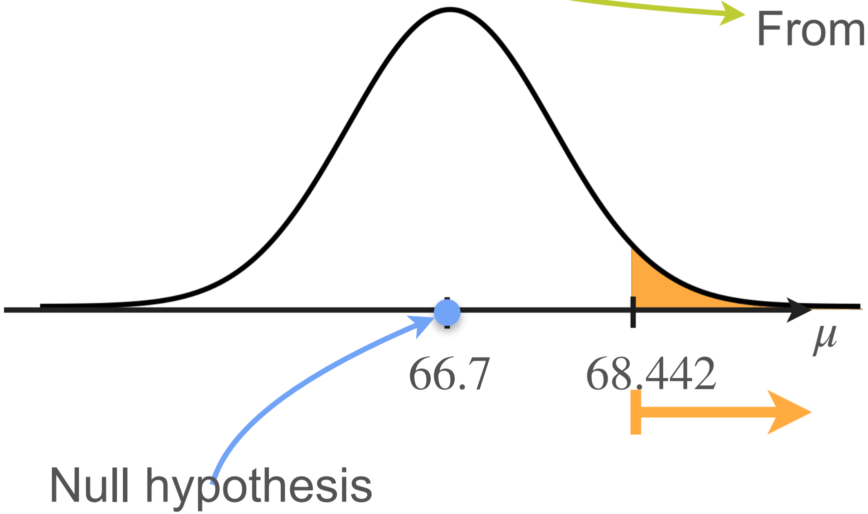
Decision rule:

If *p*-value $< \alpha$ reject H_0 (and accept H_1 as true)

P-Values

A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

From the observed value to the direction of H_1



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Decision rule:

If p-value $< \alpha$ reject H_0 (and accept H_1 as true)

If p-value $> \alpha$ don't reject H_0

p-values

A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

p-values

A **p-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

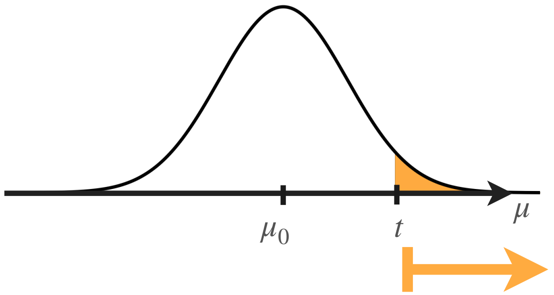
$T(X)$: test statistic t: observed statistic H_0: $\mu = \mu_0$

p-values

A **p-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

$T(X)$: test statistic t: observed statistic H_0: $\mu = \mu_0$

Right-tailed test



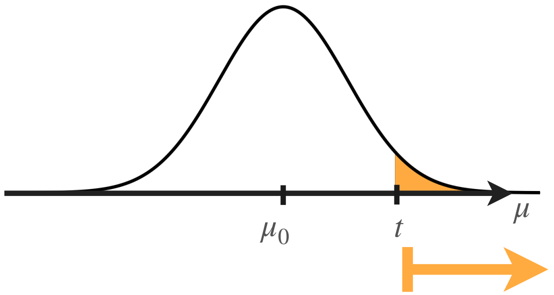
$$\mathbf{P}(T(X) > t \,|\, H_0)$$

p-values

A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

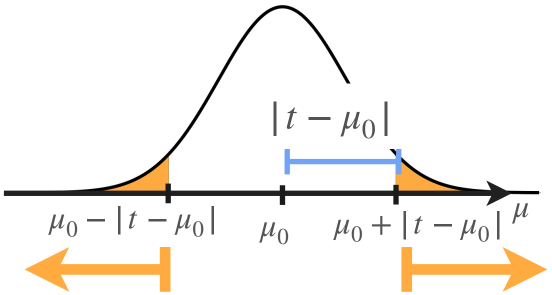
$T(X)$: test statistic $\quad$ t: observed statistic $\quad$ H_0: $\mu = \mu_0$

Right-tailed test



$$\mathbf{P}(T(X) > t \,|\, H_0)$$

Two-tailed test



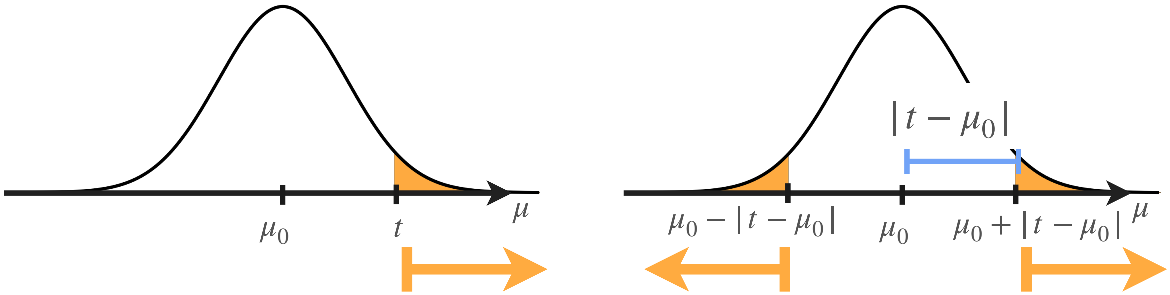
$$\mathbf{P}(|T(X) - \mu_0| > |t - \mu_0| \,|\, H_0)$$

p-values

A **p-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

$T(X)$: test statistic $\quad$ t: observed statistic $\quad$ H_0: $\mu = \mu_0$

Right-tailed test $\quad$ Two-tailed test $\quad$ Left-tailed test



$\mathbf{P}(T(X) > t \,|\, H_0)$ $\quad$ $\mathbf{P}(|T(X) - \mu_0| > |t - \mu_0| \;|\, H_0)$ $\quad$ $\mathbf{P}(T(X) < t \,|\, H_0)$

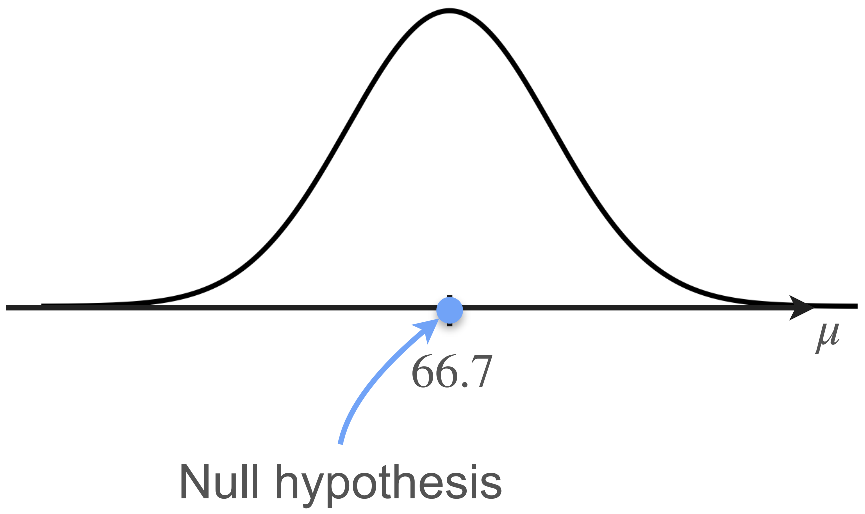
Two-Tailed Test for Gaussian Data (Known σ)

Two-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$



Two-Tailed Test for Gaussian Data (Known σ)

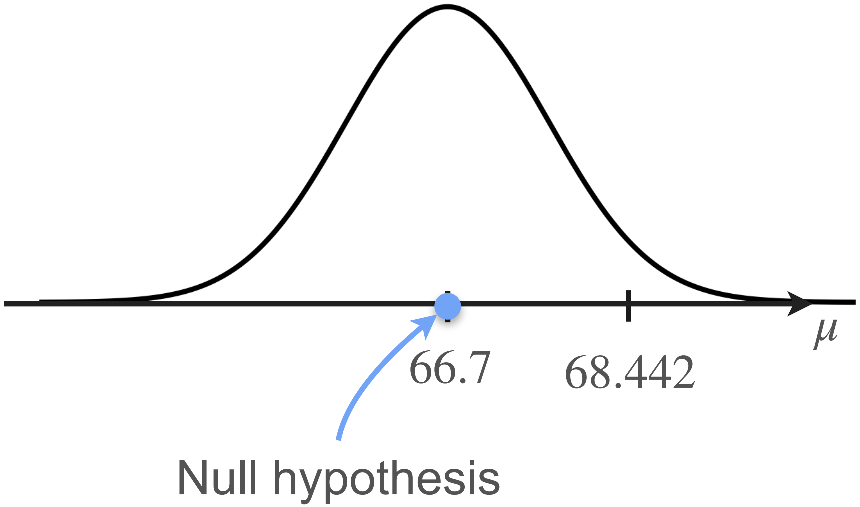
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Two-Tailed Test for Gaussian Data (Known σ)

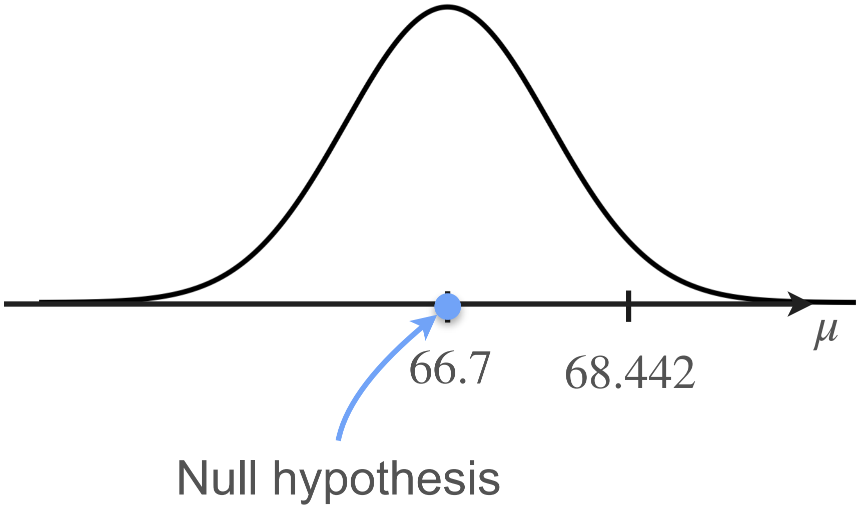
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Type I error: Determine $\mu \neq 66.7$, when population mean did not change

Two-Tailed Test for Gaussian Data (Known σ)

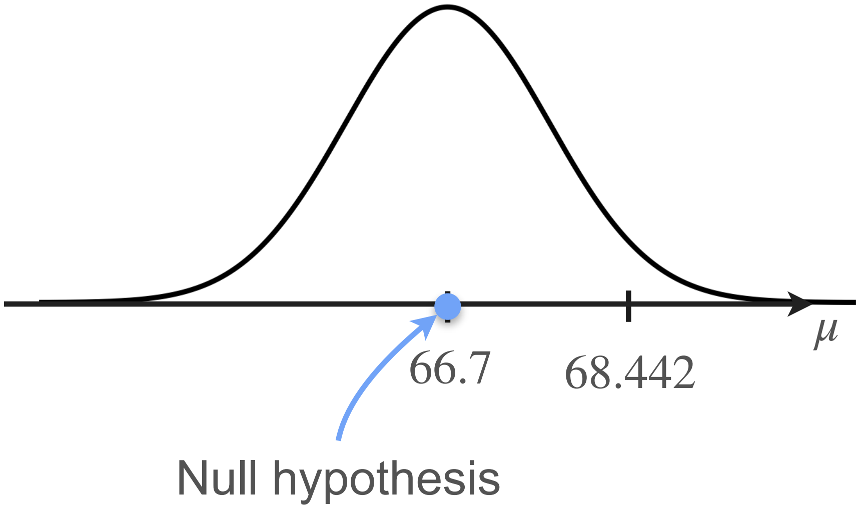
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Type I error: Determine $\mu \neq 66.7$, when population mean did not change

$$\mathbf{P}\left(\left|\bar{X} - 66.7\right| > \left|68.442 - 66.7\right| \,\middle|\, \mu = 66.7\right)?$$

Two-Tailed Test for Gaussian Data (Known σ)

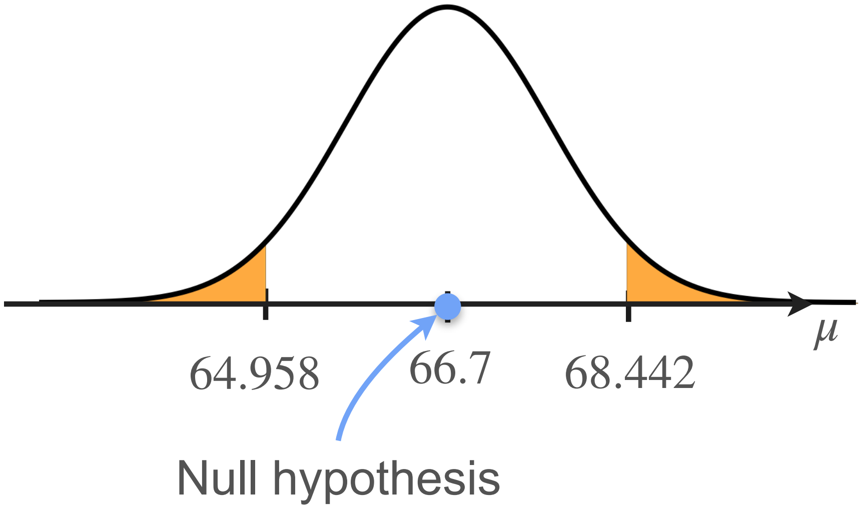
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Type I error: Determine $\mu \neq 66.7$, when population mean did not change

$$\mathbf{P}\left(\left|\bar{X} - 66.7\right| > \left|68.442 - 66.7\right| \;\middle|\; \mu = 66.7\right)$$

$$= 0.082$$

Two-Tailed Test for Gaussian Data (Known σ)

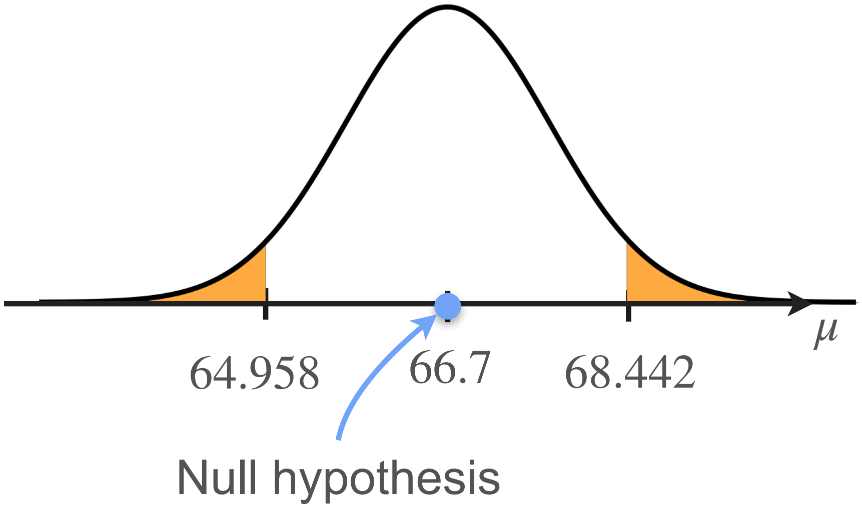
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$



Type I error: Determine $\mu \neq 66.7$, when population mean did not change

$$\mathbf{P}\left(\left| \bar{X} - 66.7 \right| > \left| 68.442 - 66.7 \right| \;\middle|\; \mu = 66.7 \right)$$

$$= 0.082 > \alpha$$

Conclusion: Do not reject H_0

(with a 5% significance level)

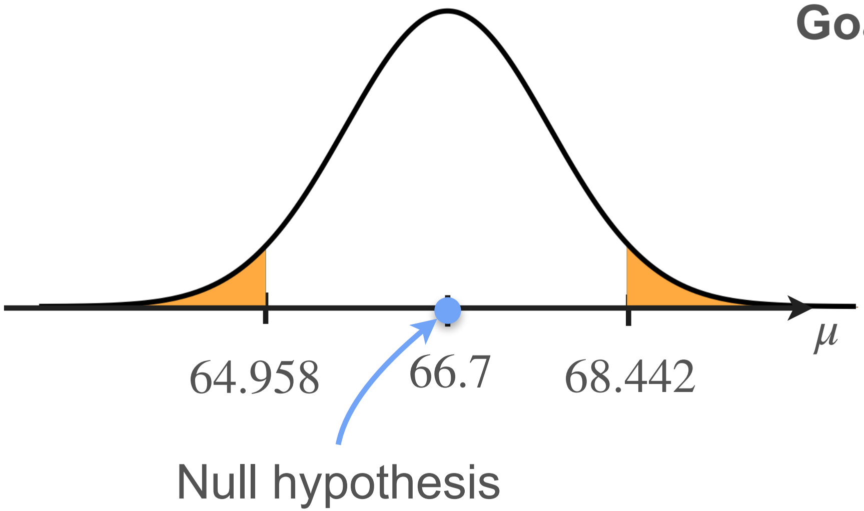
Two-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu \neq 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu \neq 66.7$, when population mean did not change

$$\mathbf{P}\left(\left| \bar{X} - 66.7 \right| > \left| 68.442 - 66.7 \right| \,\middle|\, \mu = 66.7 \right)$$

$$= 0.082 > \alpha$$

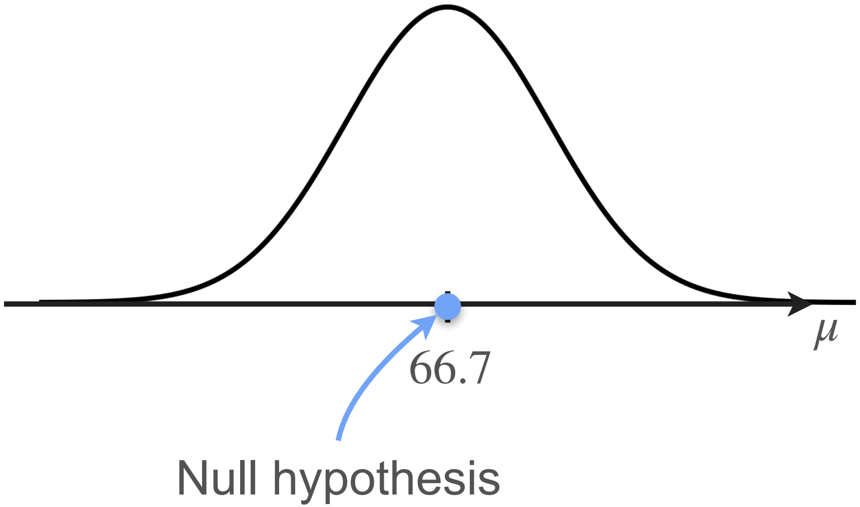
Conclusion: Do not reject H_0
(with a 5% significance level)

Left-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$



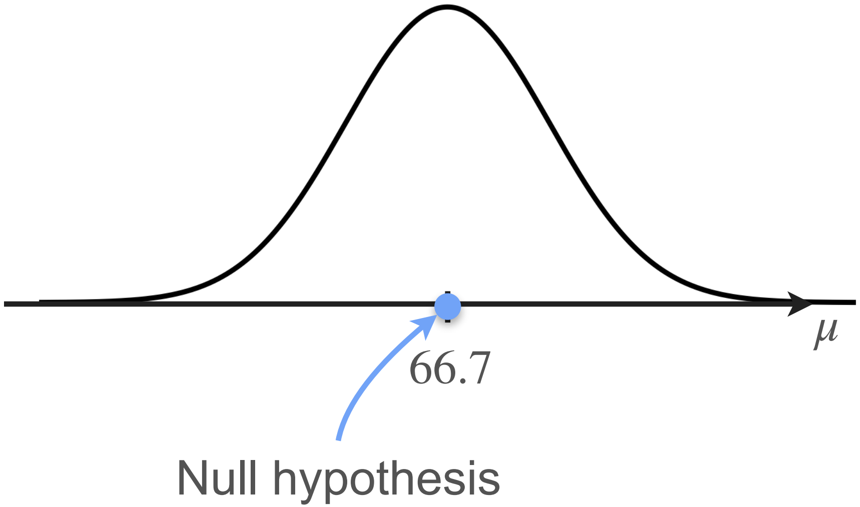
Left-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Left-Tailed Test for Gaussian Data (Known σ)

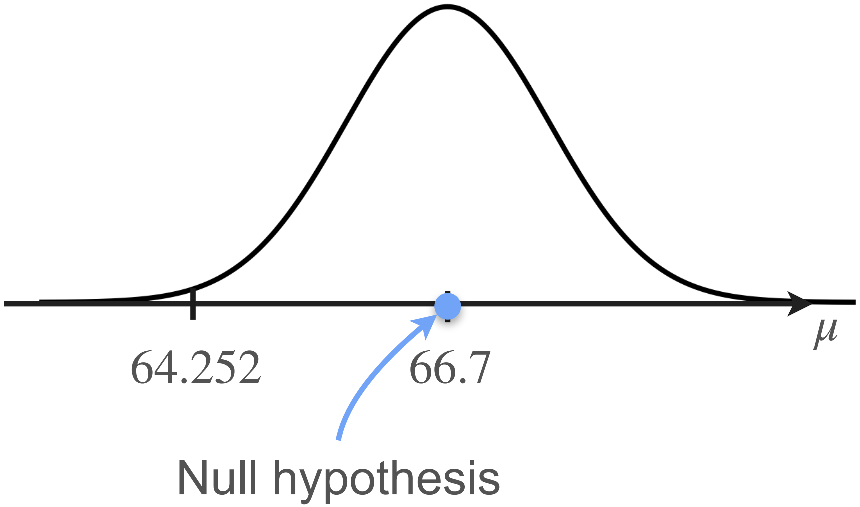
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 64.252$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Left-Tailed Test for Gaussian Data (Known σ)

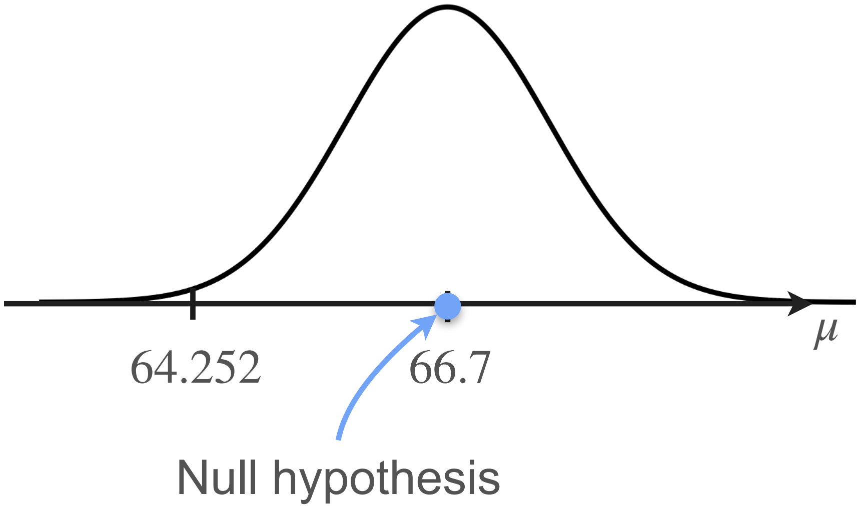
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 64.252$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Type I error: Determine $\mu < 66.7$, when population mean did not change

Left-Tailed Test for Gaussian Data (Known σ)

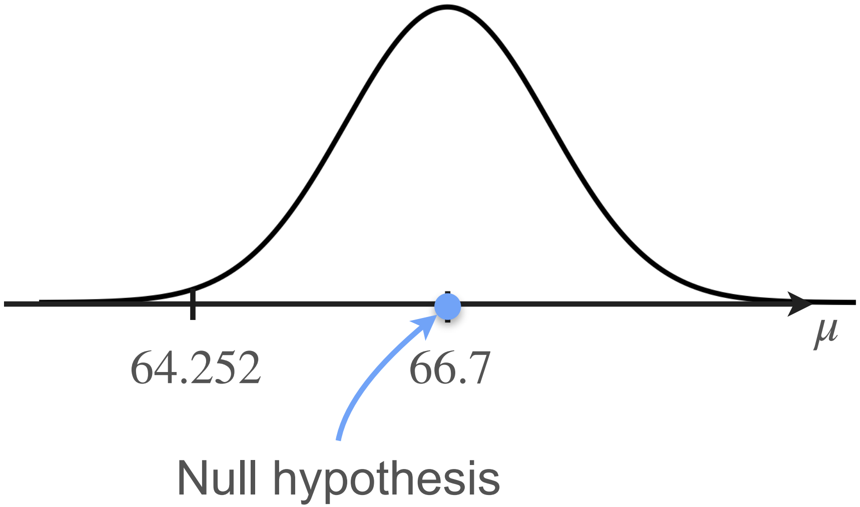
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 64.252$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Type I error: Determine $\mu < 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} < 64.252 \,\middle|\, \mu = 66.7\right)?$$

Left-Tailed Test for Gaussian Data (Known σ)

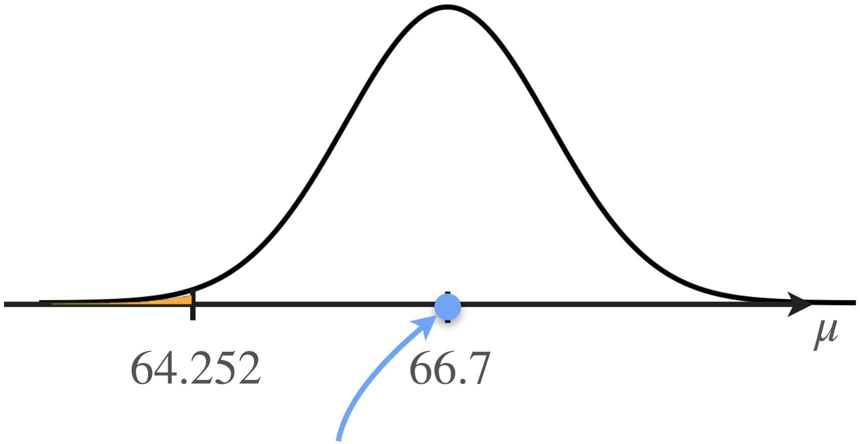
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 64.252$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Type I error: Determine $\mu < 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} < 64.252 \,\middle|\, \mu = 66.7\right)$$

$$= 0.0094$$

Left-Tailed Test for Gaussian Data (Known σ)

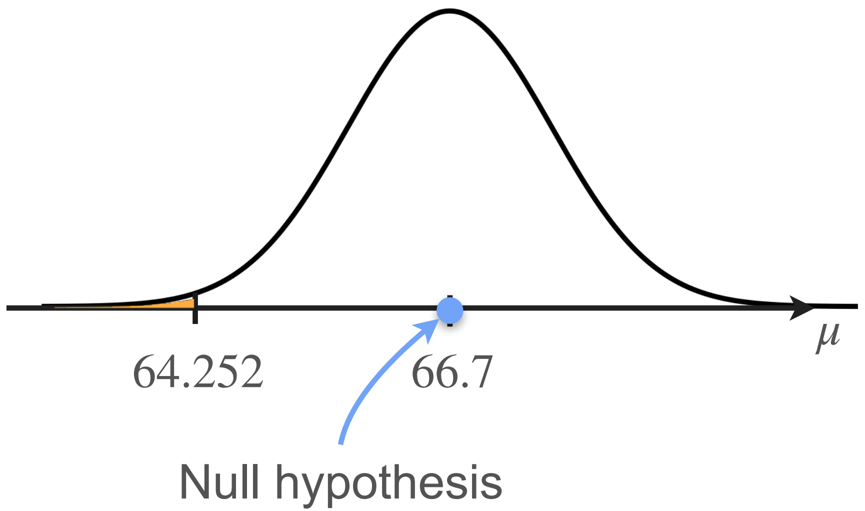
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 64.252$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$



Type I error: Determine $\mu < 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} < 64.252 \,\middle|\, \mu = 66.7\right)$$

$$= 0.0094 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

DeepLearning.AI

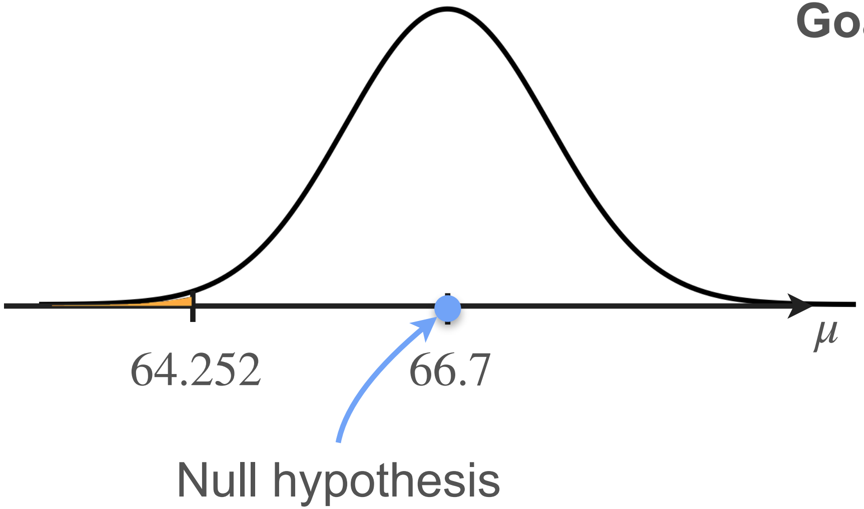
Left-Tailed Test for Gaussian Data (Known σ)

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 64.252$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu < 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu < 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} < 64.252 \,\middle|\, \mu = 66.7\right)$$

$$= 0.0094 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

Tests Using the Z-Statistic

Tests Using the Z-Statistic

So far, you used the statistic $\overline{X}$

Tests Using the Z-Statistic

So far, you used the statistic $\overline{X}$

If H_0 is true: $\overline{X} \sim \mathcal{N}\left(\mu_0, \frac{3^2}{10}\right)$

Tests Using the Z-Statistic

So far, you used the statistic $\overline{X}$

If H_0 is true: $\overline{X} \sim \mathcal{N}\left(\mu_0, \frac{3^2}{10}\right)$

Applying standardization, you can write equivalent tests using the

Z-statistic $Z = \dfrac{\overline{X} - \mu_0}{3/\sqrt{10}}$

Tests Using the Z-Statistic

So far, you used the statistic $\overline{X}$

If H_0 is true: $\overline{X} \sim \mathcal{N}\left(\mu_0, \dfrac{3^2}{10}\right)$

Applying standardization, you can write equivalent tests using the

$$Z\text{-statistic } Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}}$$

If H_0 is true, $Z = \dfrac{\overline{X} - \mu_0}{3/\sqrt{10}} \sim \mathcal{N}(0,1)$

Right-Tailed Test Using the Z Statistic

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

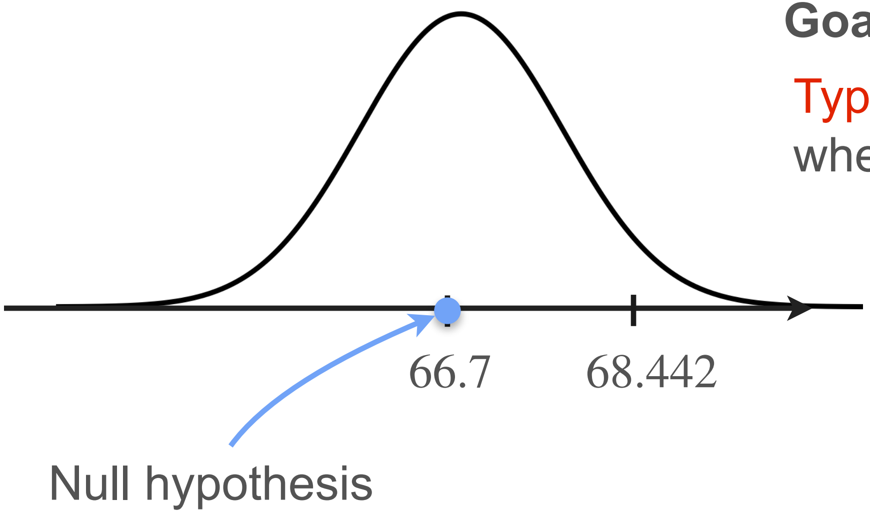
Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

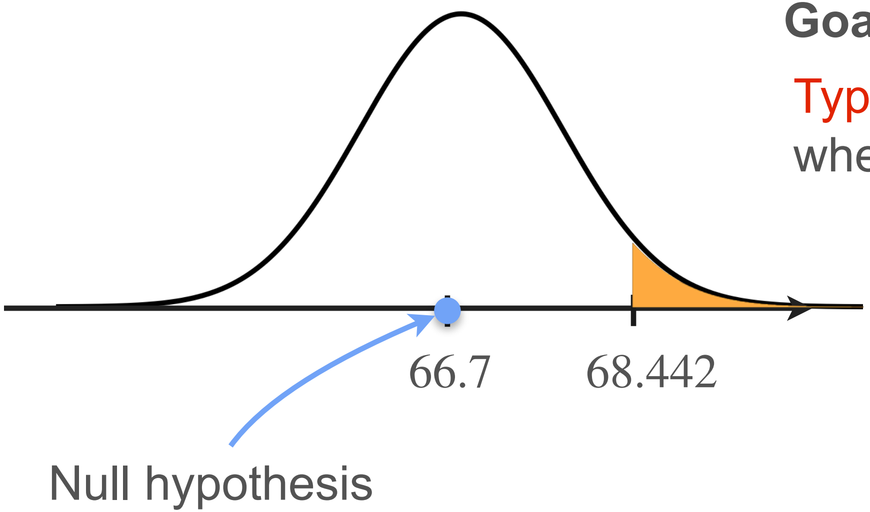
Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

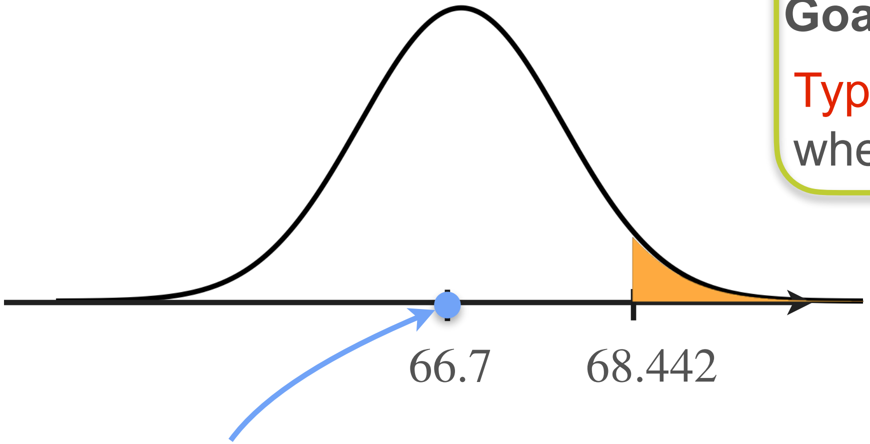
Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$

$n = 10$

$\bar{x} = 68.442$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

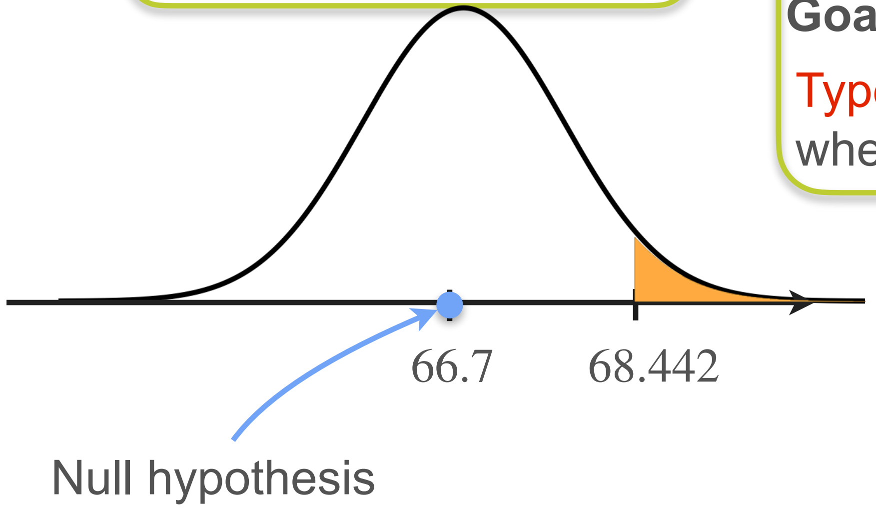
Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$



$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}}$$

$$\mathbf{P}\left(\overline{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}} \rightarrow z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\mathbf{P}\left(\overline{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}} \rightarrow z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\overline{X} > 68.442$$

$$\mathbf{P}\left(\overline{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}} \rightarrow z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\overline{X} - 66.7 > 68.442 - 66.7$$

$$\mathbf{P}\left(\overline{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}} \longrightarrow z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\mathbf{P}\left(\overline{X} > 68.442 \,\middle|\, \mu = 66.7\right) = 0.0407 < \alpha$$

$$\frac{\overline{X} - 66.7}{3/\sqrt{10}} > \frac{68.442 - 66.7}{3/\sqrt{10}}$$

Conclusion: reject H_0

(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\bar{X} - \mu_0}{3/\sqrt{10}} \rightarrow z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\frac{\bar{X} - 66.7}{3/\sqrt{10}} > \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\mathbf{P}\left(\bar{X} > 68.442 \,\middle|\, \mu = 66.7\right)$$
$$= 0.0407 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$Z = \frac{\overline{X} - \mu_0}{3/\sqrt{10}} \rightarrow z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\frac{\overline{X} - 66.7}{3/\sqrt{10}} > \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$$\mathbf{P}\left(\qquad > \qquad \middle| \mu = 66.7 \right)$$

$$= 0.0407 < \alpha$$

Conclusion: reject H_0

(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

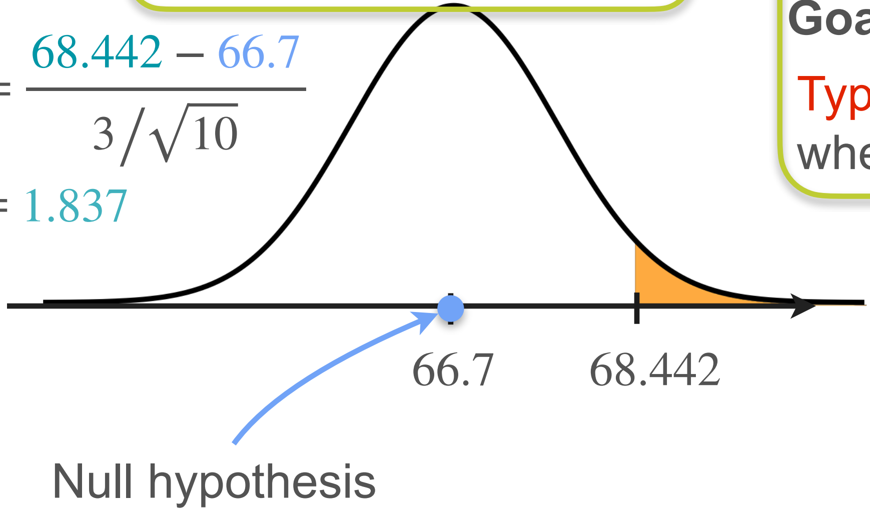
$n = 10$

$$z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change



$$\mathbf{P}\left(\frac{\bar{X} - 66.7}{3/\sqrt{10}} > 1.837 \;\middle|\; \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad \bar{x} = 68.442$

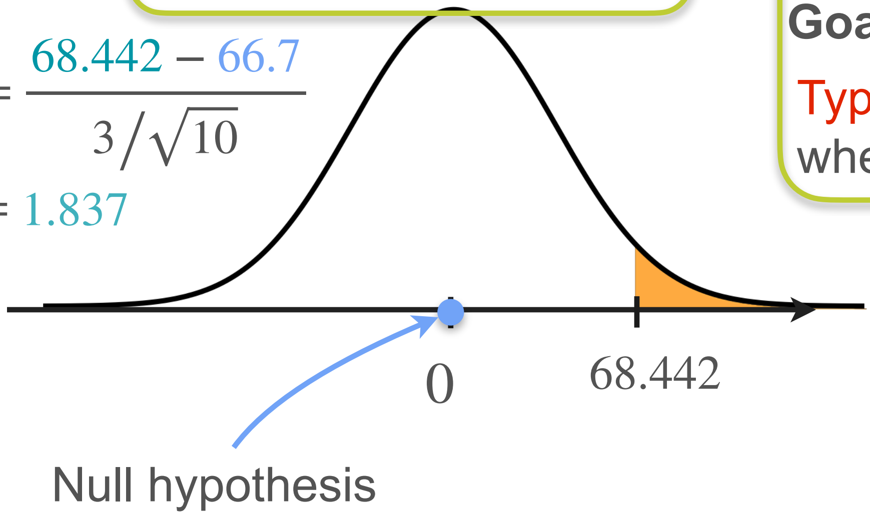
$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$



$$\mathbf{P}\left(\frac{\bar{X} - 66.7}{3/\sqrt{10}} > 1.837 \;\middle|\; \mu = 66.7\right) = 0.0407 < \alpha$$

Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$\sigma = 3$ $\quad$ $\bar{x} = 68.442$

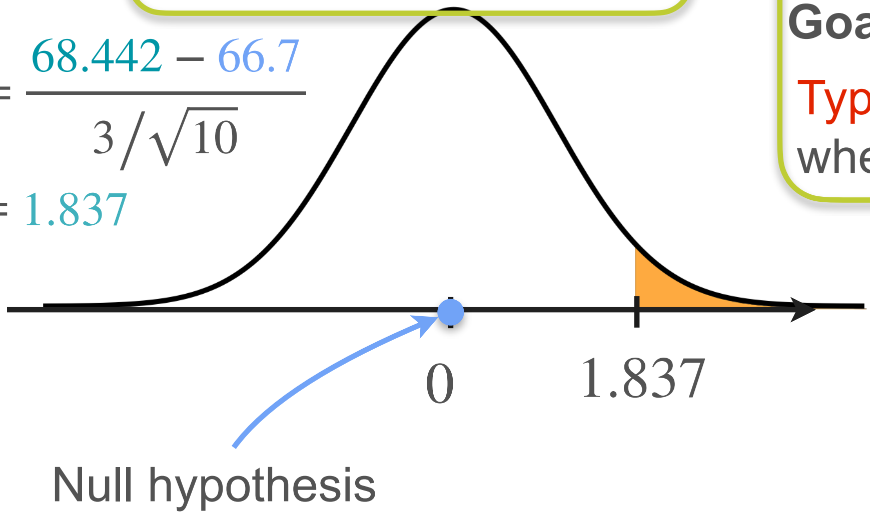
$n = 10$

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

Goal: Type I error probability < $\alpha = 0.05$

Type I error: Determine $\mu > 66.7$, when population mean did not change

$$z = \frac{68.442 - 66.7}{3/\sqrt{10}} = 1.837$$



$$\mathbf{P}\left(\frac{\bar{X} - 66.7}{3/\sqrt{10}} > 1.837 \;\middle|\; \mu = 66.7\right) = 0.0407 < \alpha$$

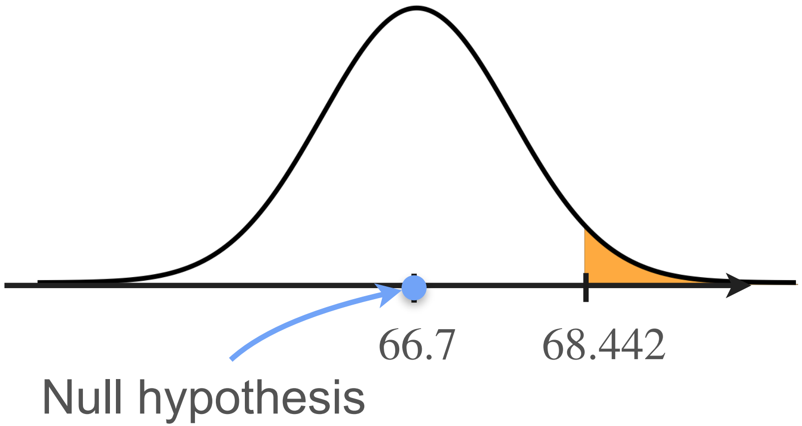
Conclusion: reject H_0
(with a 5% significance level)

Hypothesis Testing

Critical Values

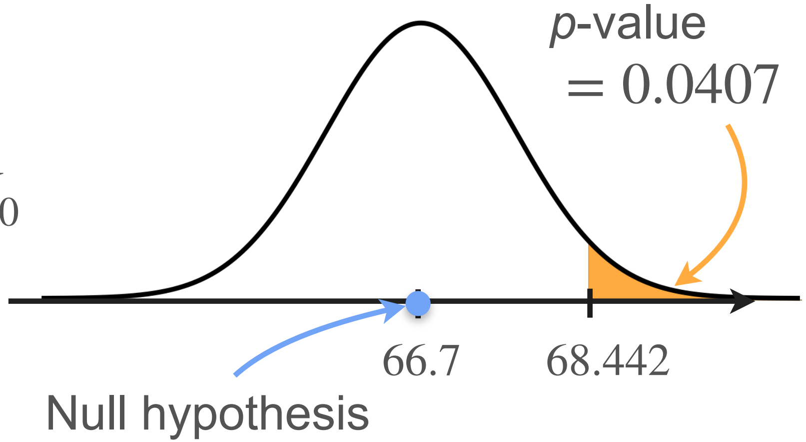
P-Values and Critical Values

P-Values and Critical Values



P-Values and Critical Values

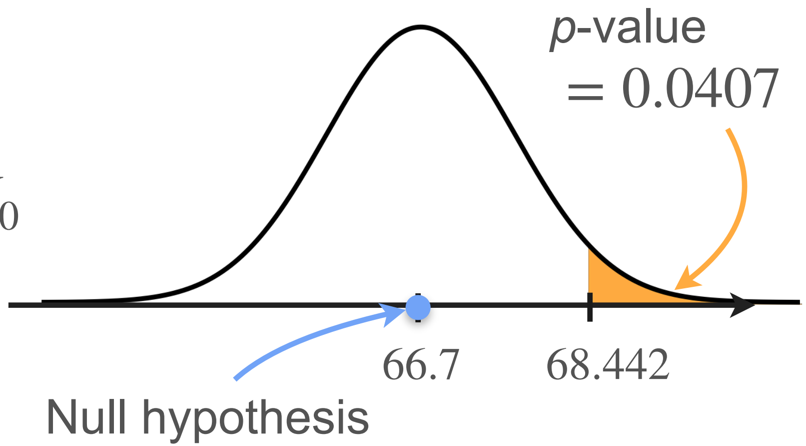
If p-value $< \alpha$ ⇨ Reject H_0



P-Values and Critical Values

If p-value $< \alpha$ ⇨ Reject H_0

What is the least extreme sample you could get that you would still reject H0?

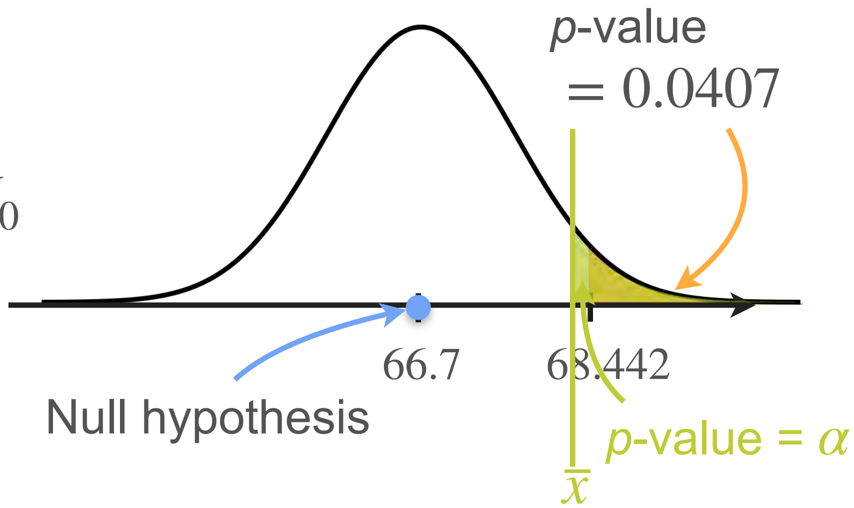


P-Values and Critical Values

If p-value $< \alpha$ $\Rightarrow$ Reject H_0

What is the least extreme sample you could get that you would still reject H0?

Sample that has p-value = α



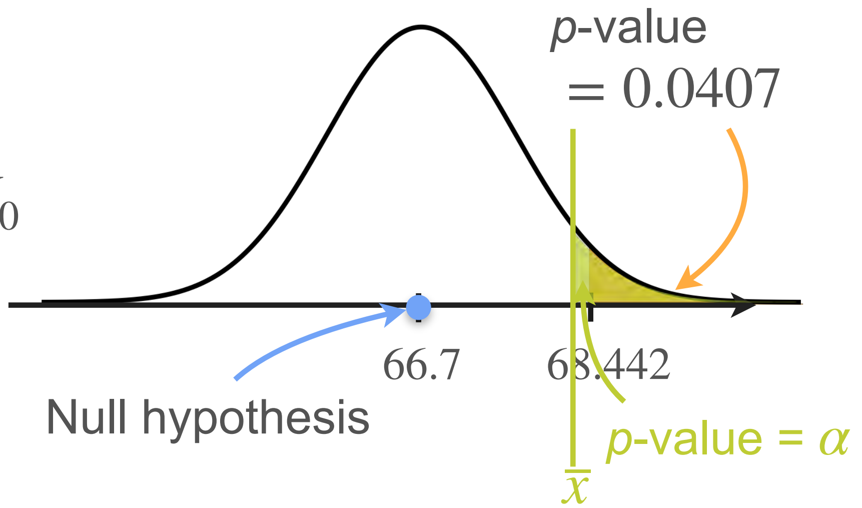
P-Values and Critical Values

If p-value $< \alpha$ ⇨ Reject H_0

What is the least extreme sample you could get that you would still reject H0?

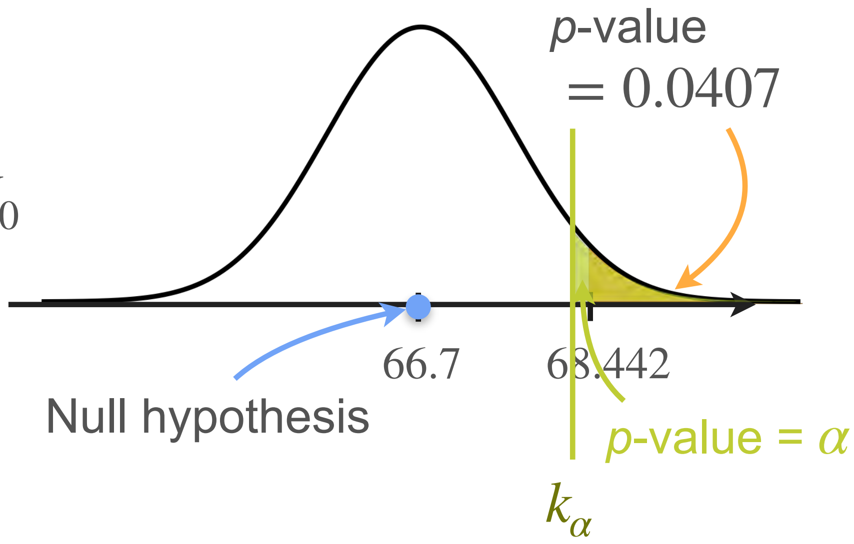
Sample that has p-value = α

Critical values



P-Values and Critical Values

If p-value $< \alpha$ ⇨ Reject H_0



What is the least extreme sample you could get that you would still reject H0?

Sample that has p-value = α

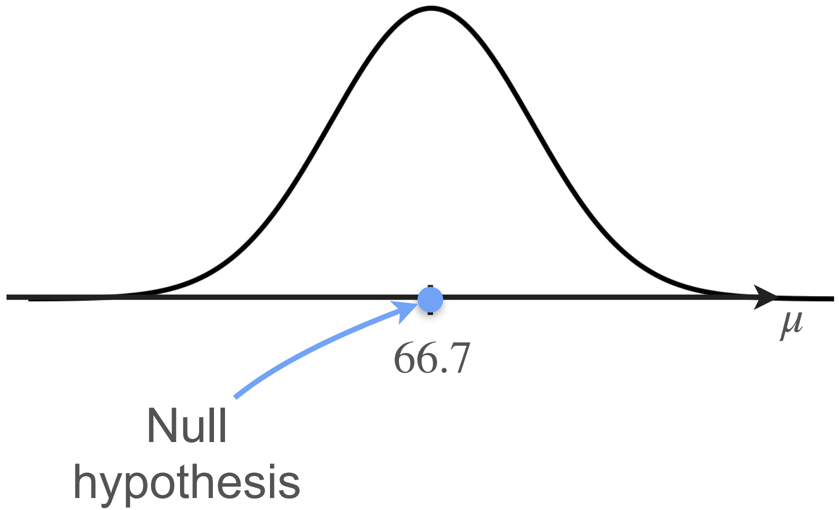
Critical values

Computing Critical Values

Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

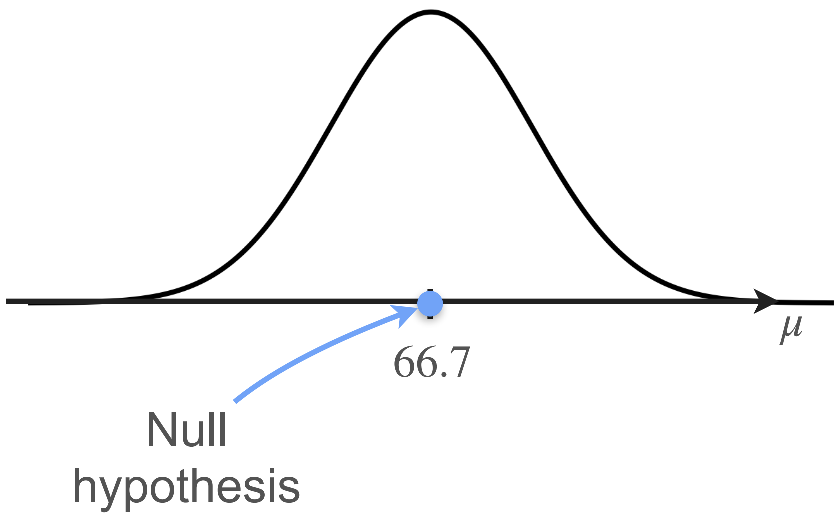
$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$ $\quad n = 10 \quad \sigma = 3$

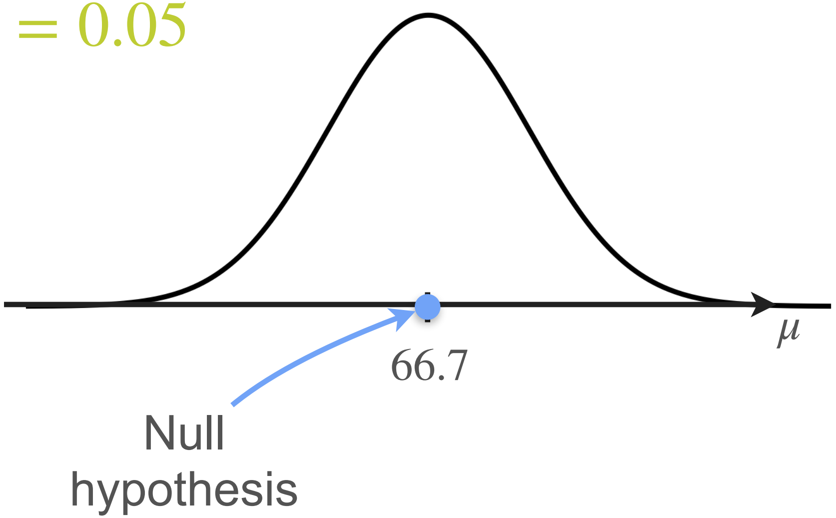


Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$ $n = 10$ $\sigma = 3$

$\alpha = 0.05$



Computing Critical Values

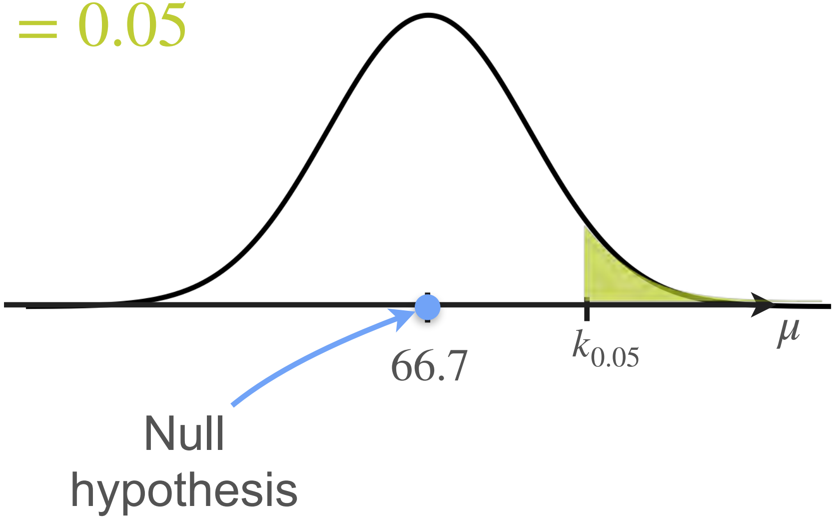
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3$

$\alpha = 0.05$

$$0.05 = \mathbf{P}\left(\overline{X} > k_{0.05} \,\middle|\, \mu = 66.7\right)$$



Computing Critical Values

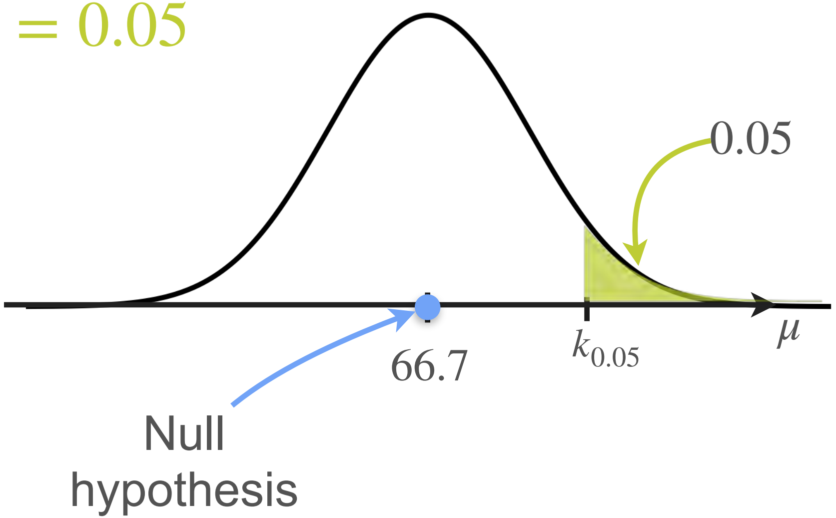
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3$

$\alpha = 0.05$

$$0.05 = \mathbf{P}\left(\overline{X} > k_{0.05} \,\middle|\, \mu = 66.7\right)$$



Computing Critical Values

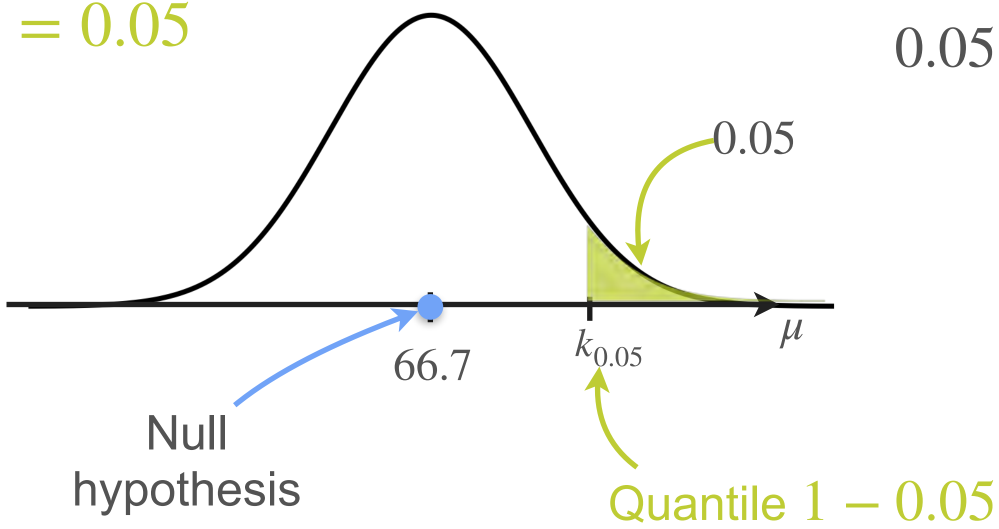
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3$

$\alpha = 0.05$

$$0.05 = \mathbf{P}\left(\bar{X} > k_{0.05} \,\middle|\, \mu = 66.7\right)$$



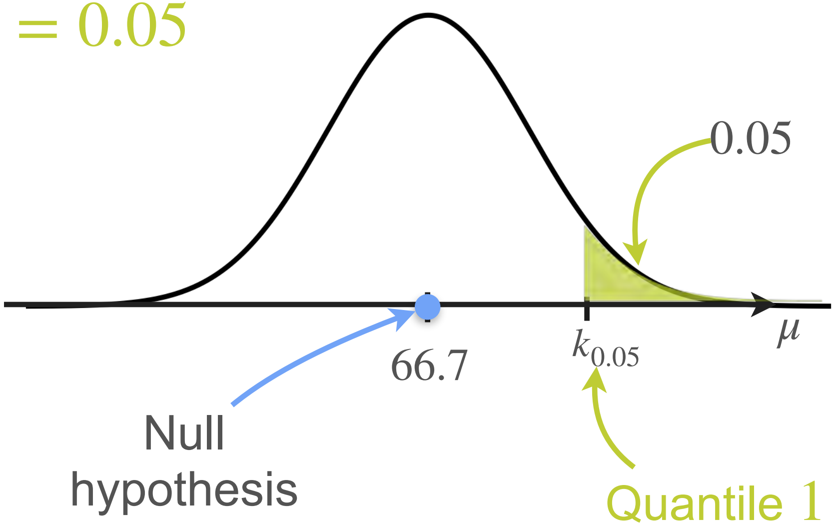
Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3$

$\alpha = 0.05$



$$0.05 = \mathbf{P}\left(\bar{X} > k_{0.05} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

$$k_{0.05} = 68.26$$

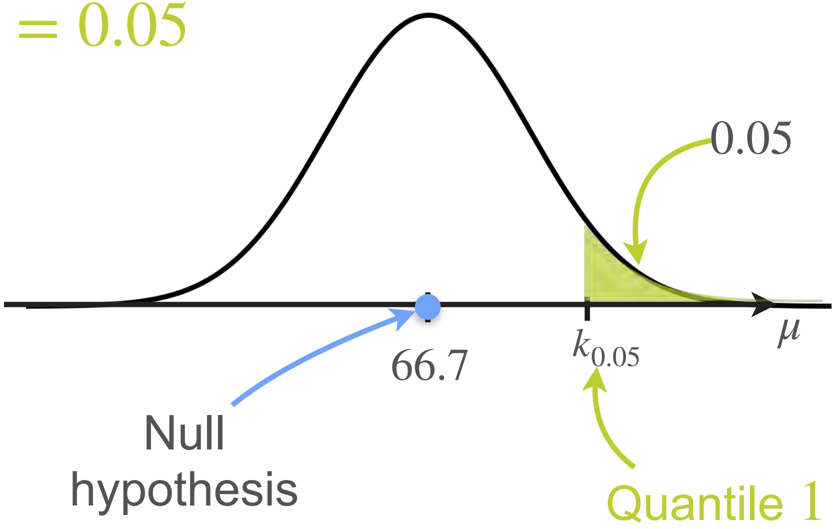
Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3$

$\alpha = 0.05$



$$0.05 = \mathbf{P}\left(\bar{X} > k_{0.05} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

$k_{0.05} = 68.26$

Decision rule: Reject H_0 if $\bar{x} > 68.26$

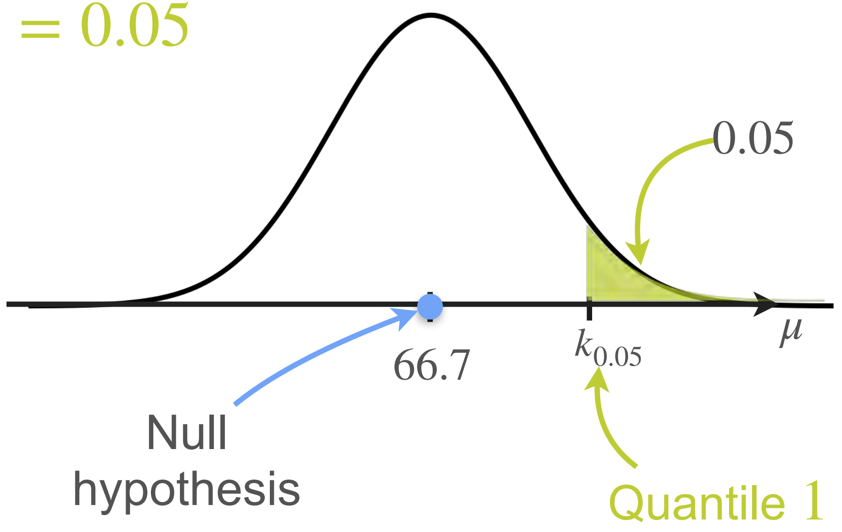
Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Reject H_0

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$\alpha = 0.05$



$n = 10 \qquad \sigma = 3 \quad \bar{x} = 68.442$

$$0.05 = \mathbf{P}\left(\bar{X} > k_{0.05} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7 \ \ \bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

$k_{0.05} = 68.26$

Decision rule: Reject H_0 if $\bar{x} > 68.26$

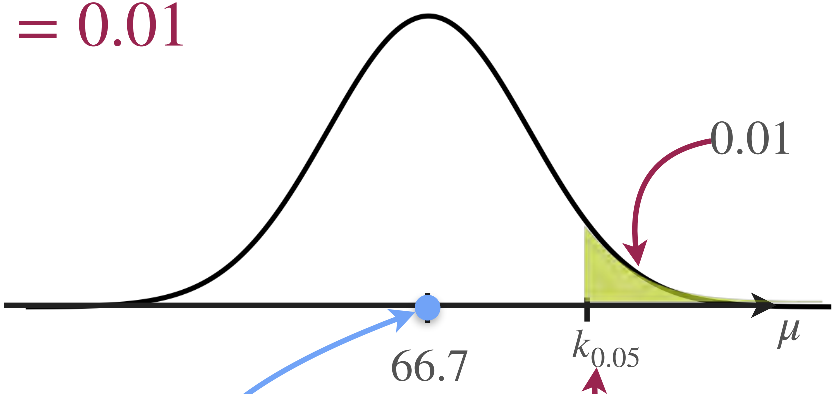
Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Reject H_0

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$\alpha = 0.01$



$n = 10$ $\sigma = 3$ $\bar{x} = 68.442$

$$0.05 = \mathbf{P}\left(\bar{X} > k_{0.01} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

$k_{0.05} = 68.26$

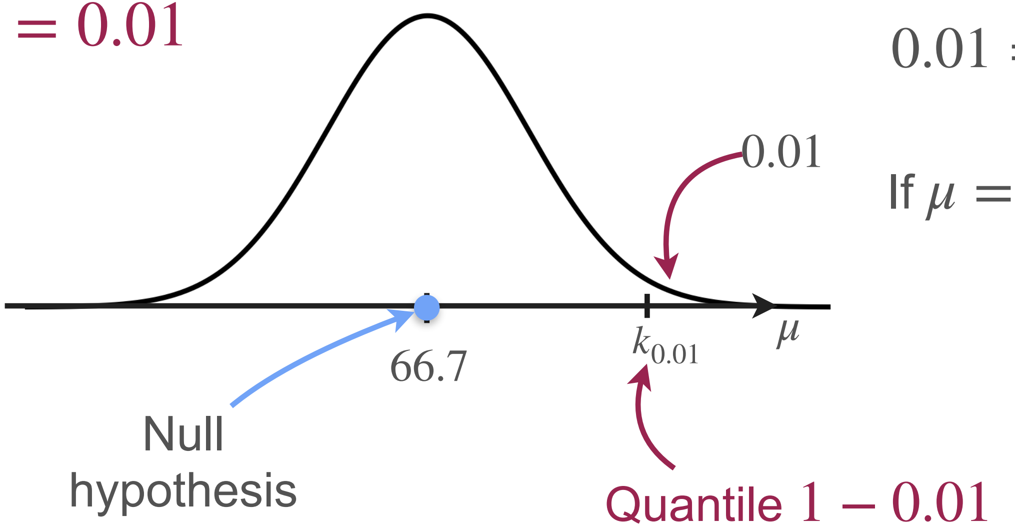
Decision rule: Reject H_0 if $\bar{x} > 68.26$

Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$\alpha = 0.01$



$n = 10 \qquad \sigma = 3 \qquad \bar{x} = 68.442$

$$0.01 = \mathbf{P}\left(\bar{X} > k_{0.01} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

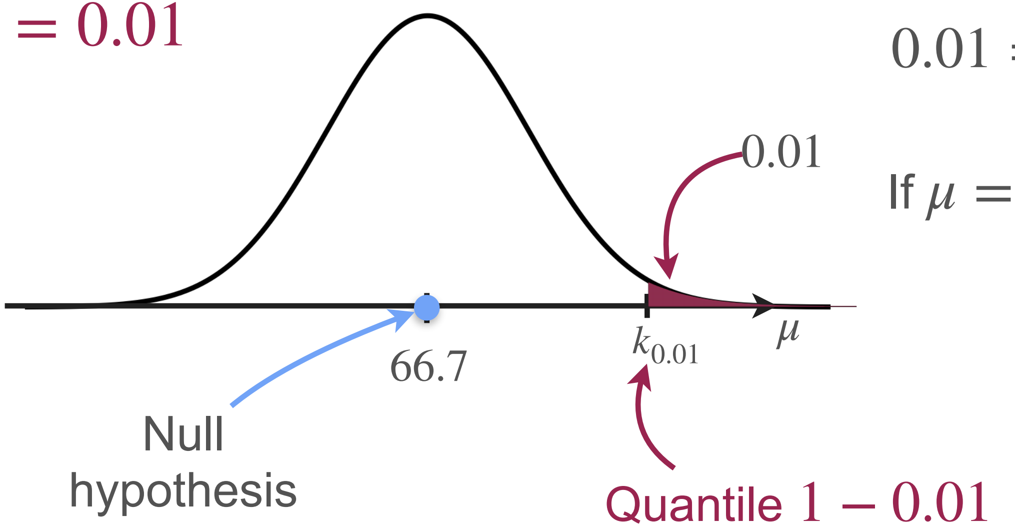
$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3 \qquad \bar{x} = 68.442$

$\alpha = 0.01$

$$0.01 = \mathbf{P}\left(\bar{X} > k_{0.01} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\quad \bar{X} \sim \mathcal{N}\left(66.7, \dfrac{3^2}{10}\right)$



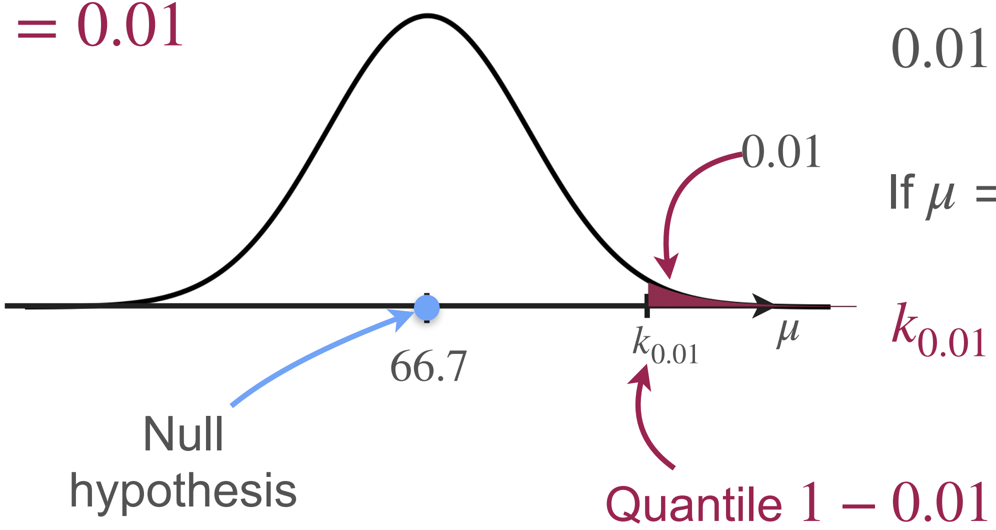
Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3 \quad \bar{x} = 68.442$

$\alpha = 0.01$



$$0.01 = \mathbf{P}\left(\bar{X} > k_{0.01} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

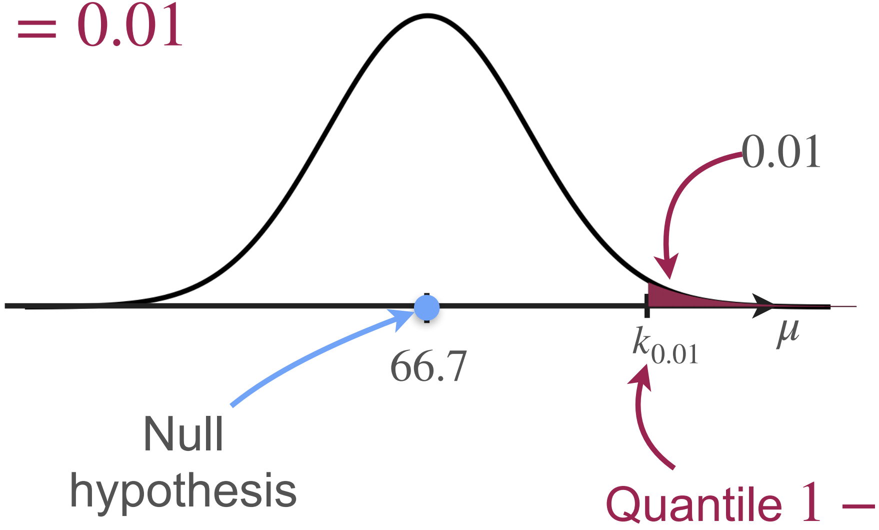
$$k_{0.01} = 68.91$$

Computing Critical Values

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$\alpha = 0.01$



$n = 10$ $\sigma = 3$ $\bar{x} = 68.442$

$$0.01 = \mathbf{P}\left(\bar{X} > k_{0.01} \middle| \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

$$k_{0.01} = 68.91$$

Decision rule: Reject H_0 if $\bar{x} > 68.91$

Computing Critical Values

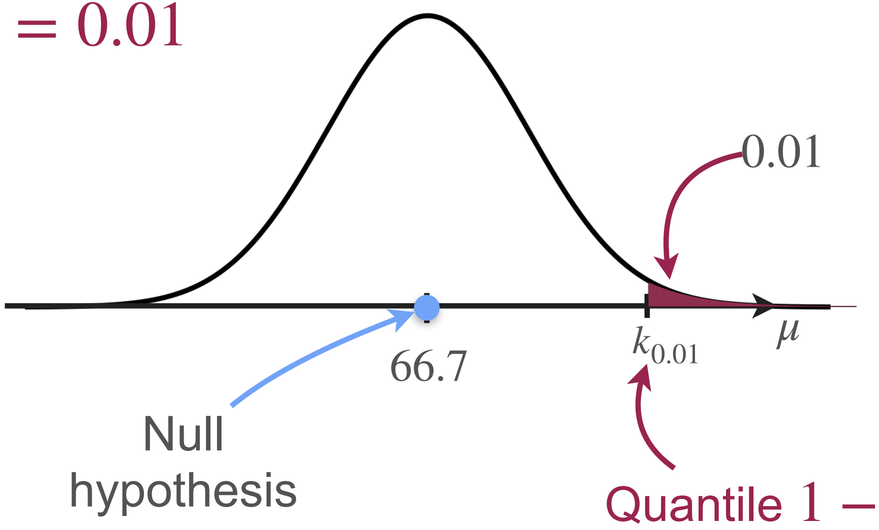
The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

Do not reject H_0

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$

$n = 10 \qquad \sigma = 3 \qquad \bar{x} = 68.442$

$\alpha = 0.01$



$$0.01 = \mathbf{P}\left(\bar{X} > k_{0.01} \,\middle|\, \mu = 66.7\right)$$

If $\mu = 66.7$ $\bar{X} \sim \mathcal{N}\left(66.7, \frac{3^2}{10}\right)$

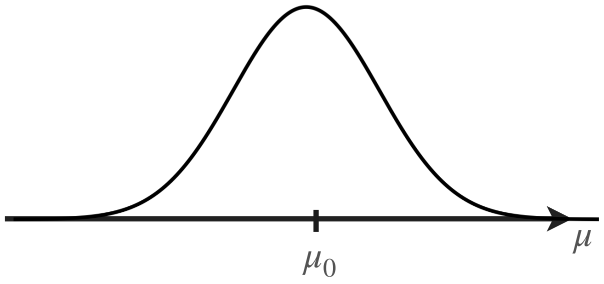
$k_{0.01} = 68.91$

Decision rule: Reject H_0 if $\bar{x} > 68.91$

Critical Values

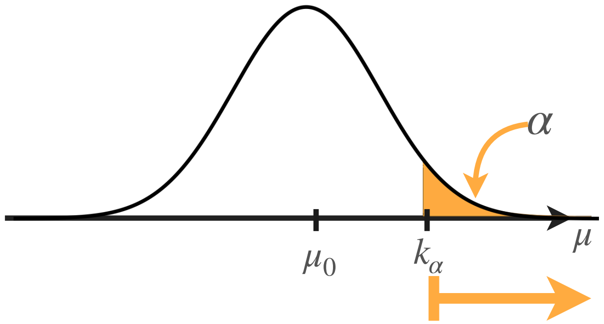
Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$



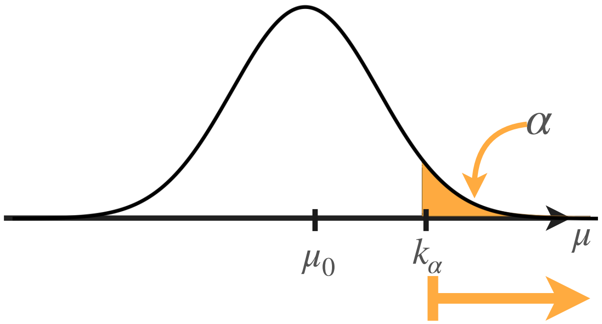
Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$



Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

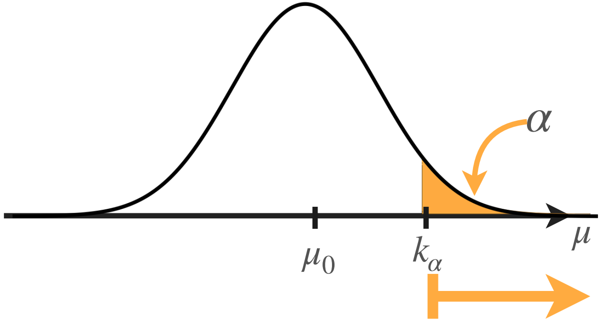


k_α is the quantile $1 - \alpha$

Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$



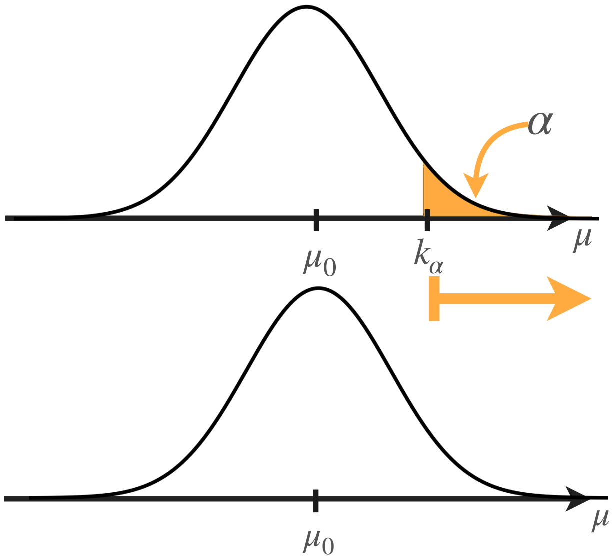
k_α is the quantile $1 - \alpha$

Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$



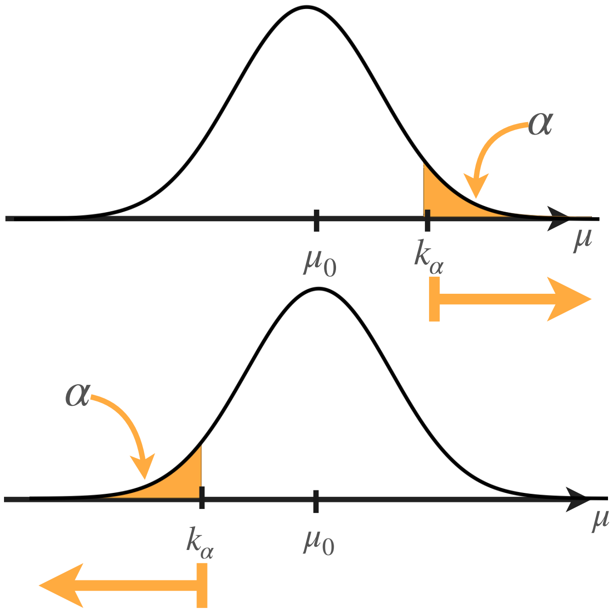
k_α is the quantile $1 - \alpha$

Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$



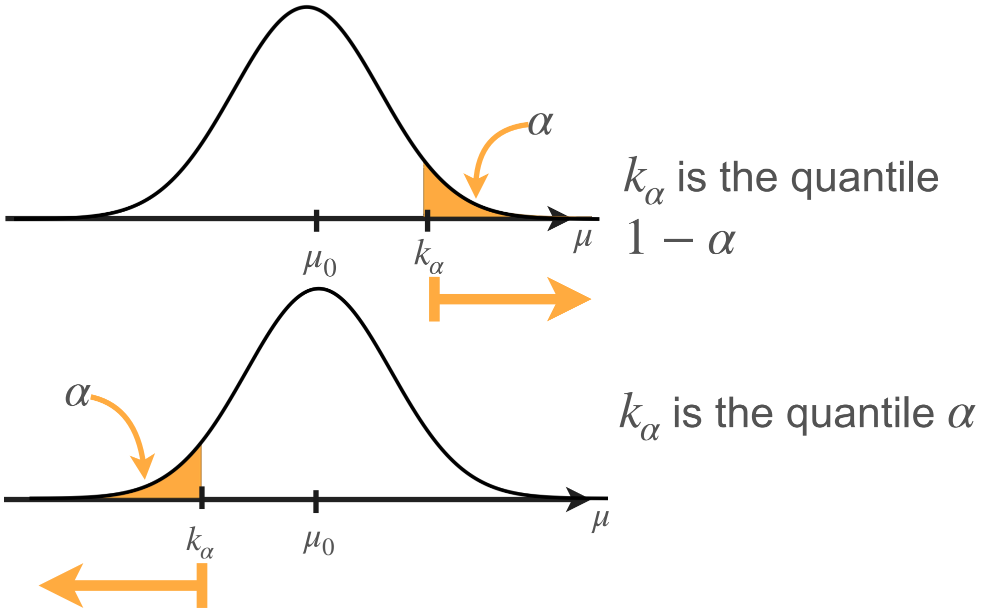
k_α is the quantile $1 - \alpha$

Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$



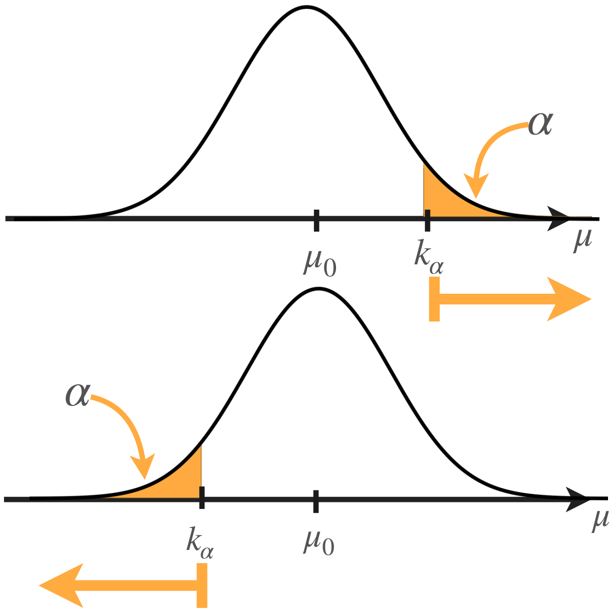
Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

Decision rule: Reject H_0 if $t < k_\alpha$



k_α is the quantile $1 - \alpha$

k_α is the quantile α

Critical Values

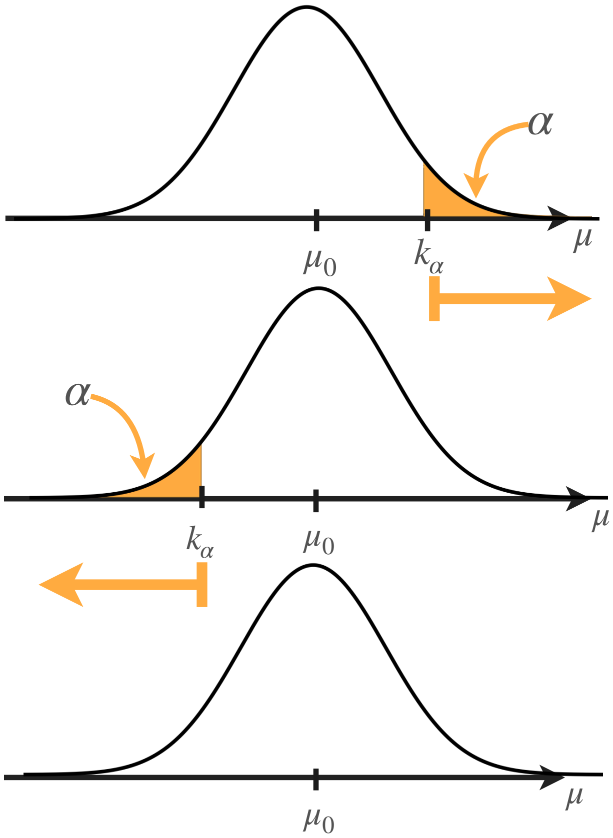
$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

Decision rule: Reject H_0 if $t < k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$



k_α is the quantile $1 - \alpha$

k_α is the quantile α

Critical Values

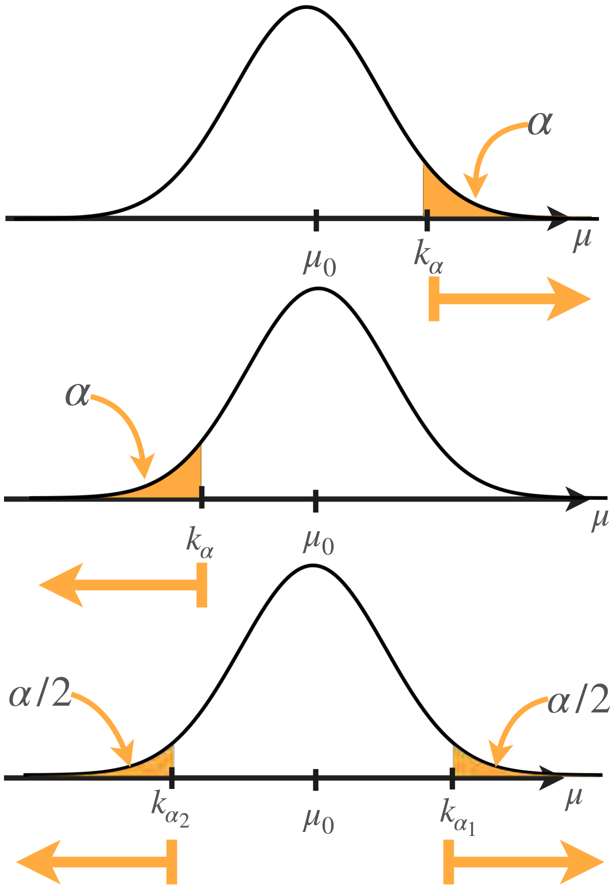
$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

Decision rule: Reject H_0 if $t < k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$



k_α is the quantile $1 - \alpha$

k_α is the quantile α

Critical Values

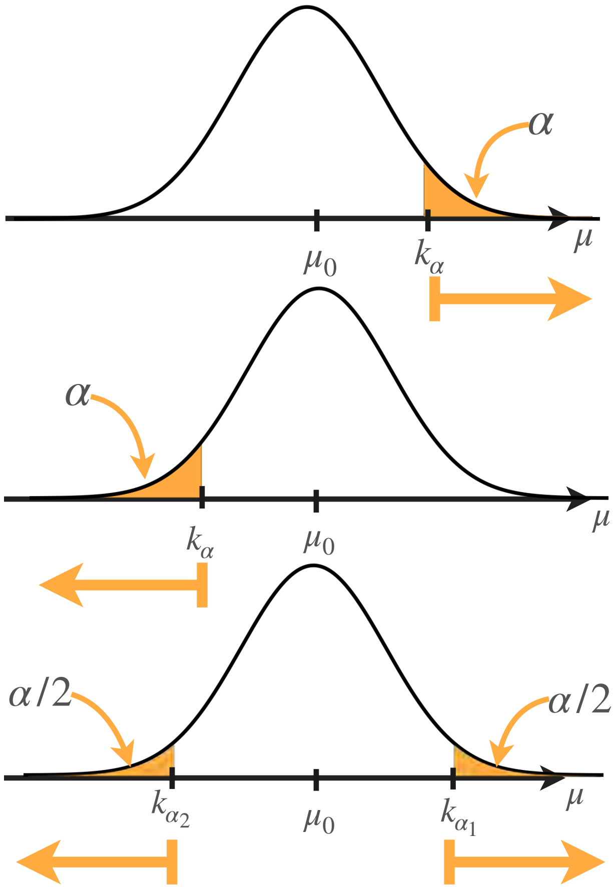
$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

Decision rule: Reject H_0 if $t < k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$



k_α is the quantile $1-\alpha$

k_α is the quantile α

k_{α_1} is the quantile $1-\dfrac{\alpha}{2}$, and k_{α_2} is the quantile $\dfrac{\alpha}{2}$

Critical Values

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu > \mu_0$

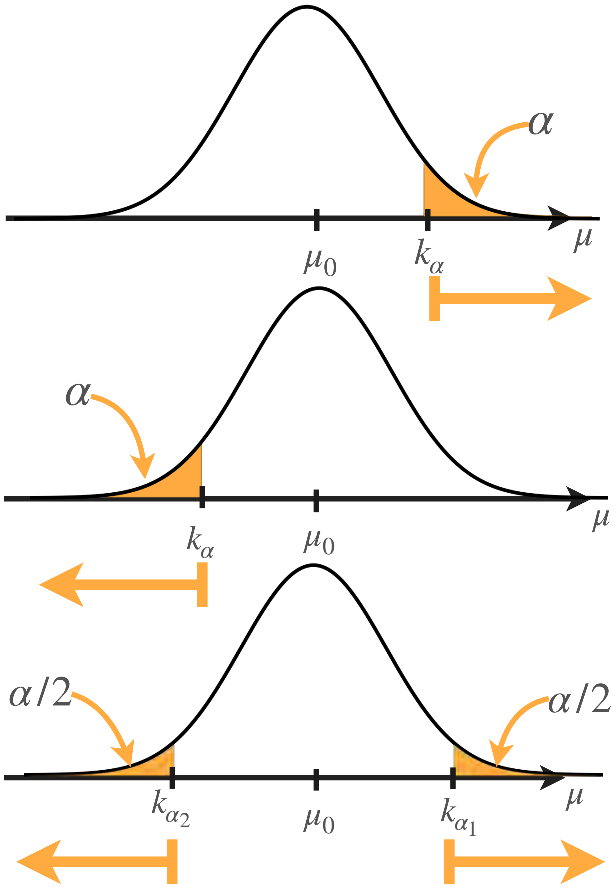
Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu < \mu_0$

Decision rule: Reject H_0 if $t < k_\alpha$

$H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

Decision rule: Reject H_0 if $t > k_{\alpha_1}$ or $t < k_{\alpha_2}$



k_α is the quantile $1 - \alpha$

k_α is the quantile α

k_{α_1} is the quantile $1 - \frac{\alpha}{2}$, and

k_{α_2} is the quantile $\frac{\alpha}{2}$

Critical Values: Concluding Remarks

Critical Values: Concluding Remarks

- You can define the critical value in advance

Critical Values: Concluding Remarks

- You can define the critical value in advance
- For a given sample, using *p*-value and critical value will lead to the same conclusion

Critical Values: Concluding Remarks

- You can define the critical value in advance
- For a given sample, using *p*-value and critical value will lead to the same conclusion
- Defining a test in terms of critical values makes determining Type II error probabilities for the decision rule.

Hypothesis Testing

Power of a test

Type I and Type II Errors

Type I and Type II Errors

The **mean** height for 18 y/o in the US in the 70s was **66.7 in**.

$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$