

Copyright Notice

These slides are distributed under the Creative Commons License.

[DeepLearning.ai](#) makes these slides available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite [DeepLearning.AI](#) as the source of the slides.

For the rest of the details of the license, see <https://creativecommons.org/licenses/by-sa/2.0/legalcode>



DeepLearning.AI

Math for Machine Learning

Probability and Statistics - Week 4

W4 Lesson 1

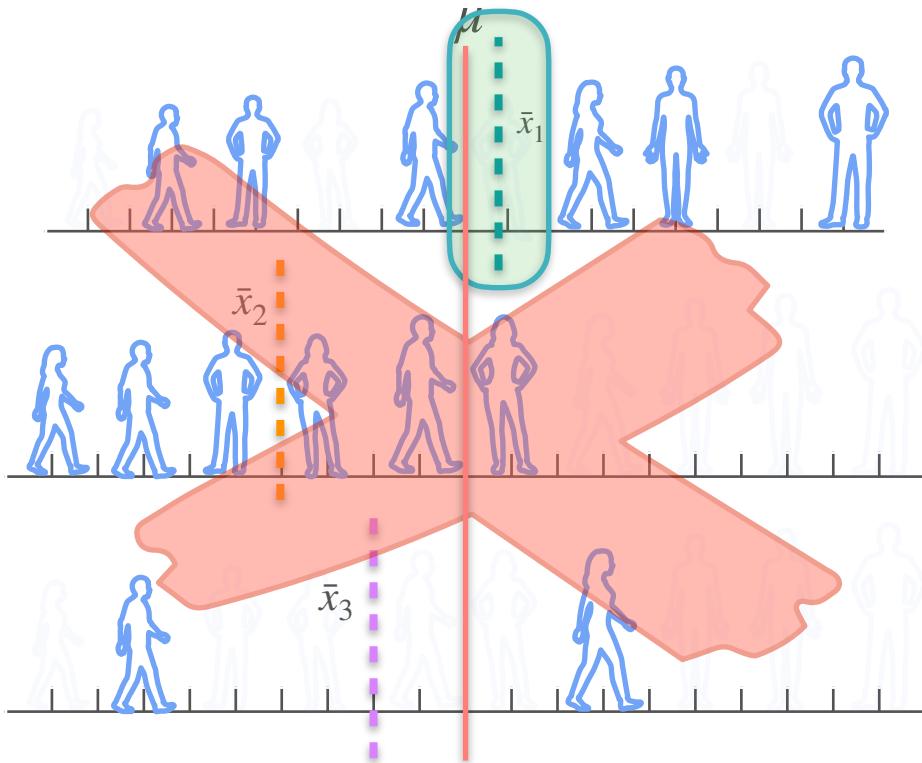


DeepLearning.AI

Confidence Interval

**Confidence Interval
(Known Standard Deviation)**

Confidence Interval - Intuition



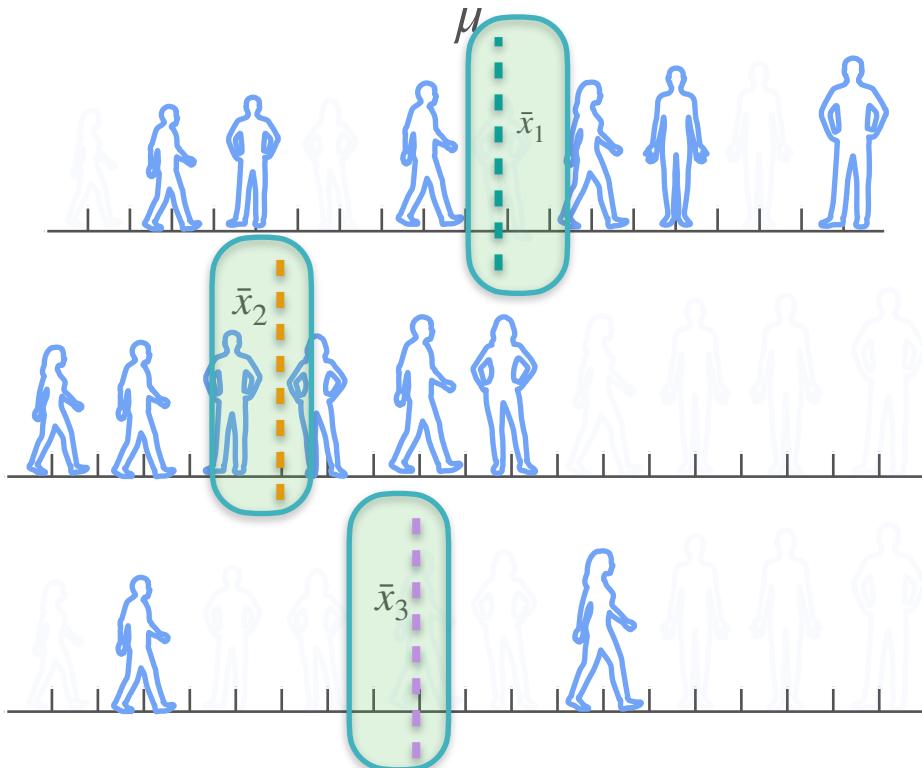
Statistopia

10,000 people

μ
(mean height of the population)

Can you use these sample means with
some degree of certainty?

Confidence Interval - Intuition



Can you use these sample means with some degree of certainty?

Confidence Interval

lower limit $< \mu <$ upper limit

Confidence Interval - Intuition

$n = 1$



Confidence level



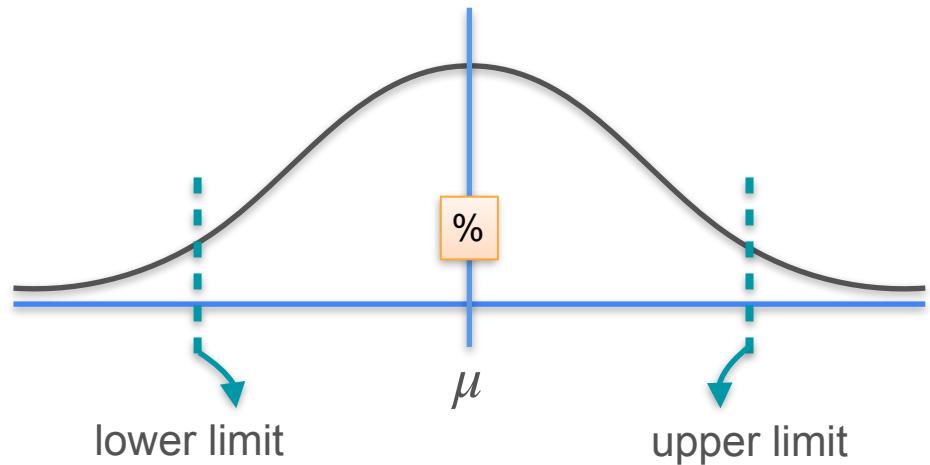
\bar{x}

$1 - \alpha$

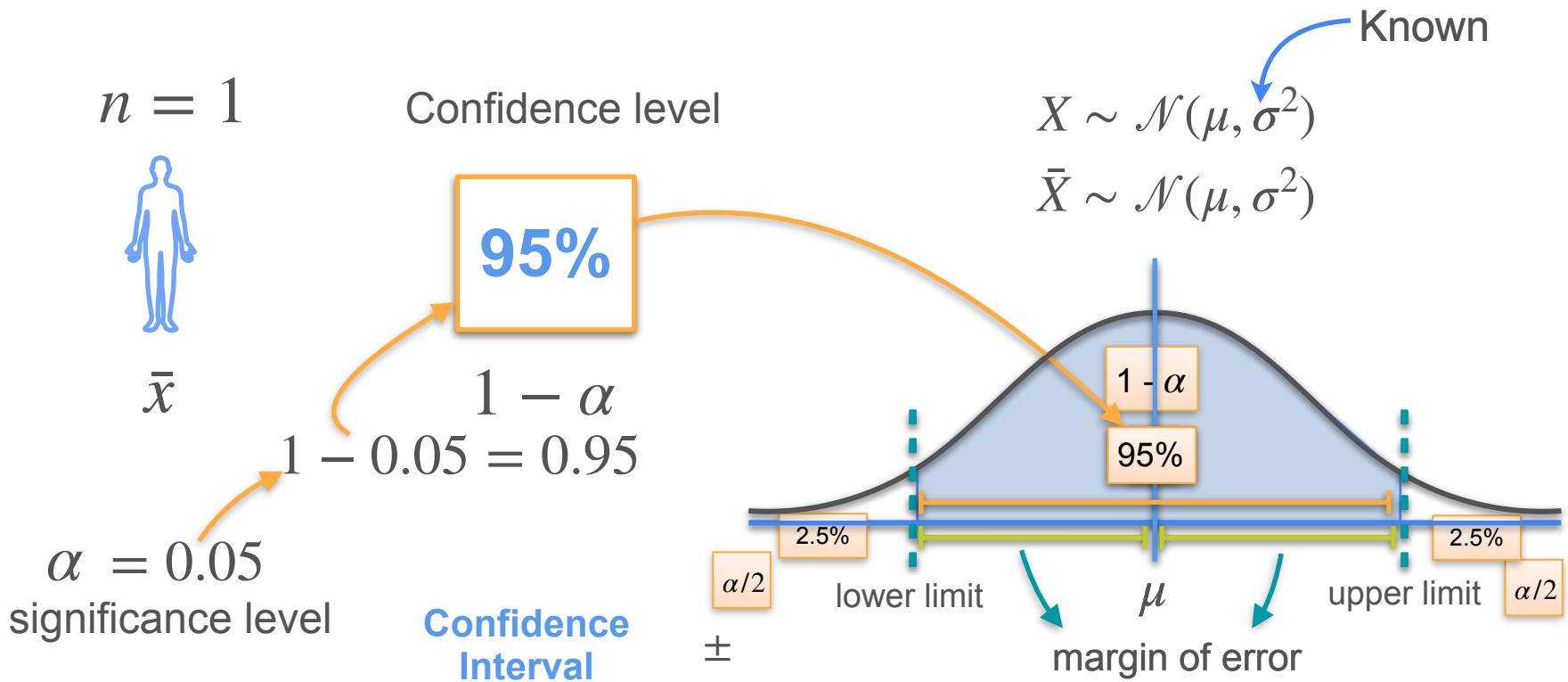
α
significance level

Known

$$X \sim \mathcal{N}(\mu, \sigma^2)$$
$$\bar{X} \sim \mathcal{N}(\mu, \sigma^2)$$



Confidence Interval - Intuition



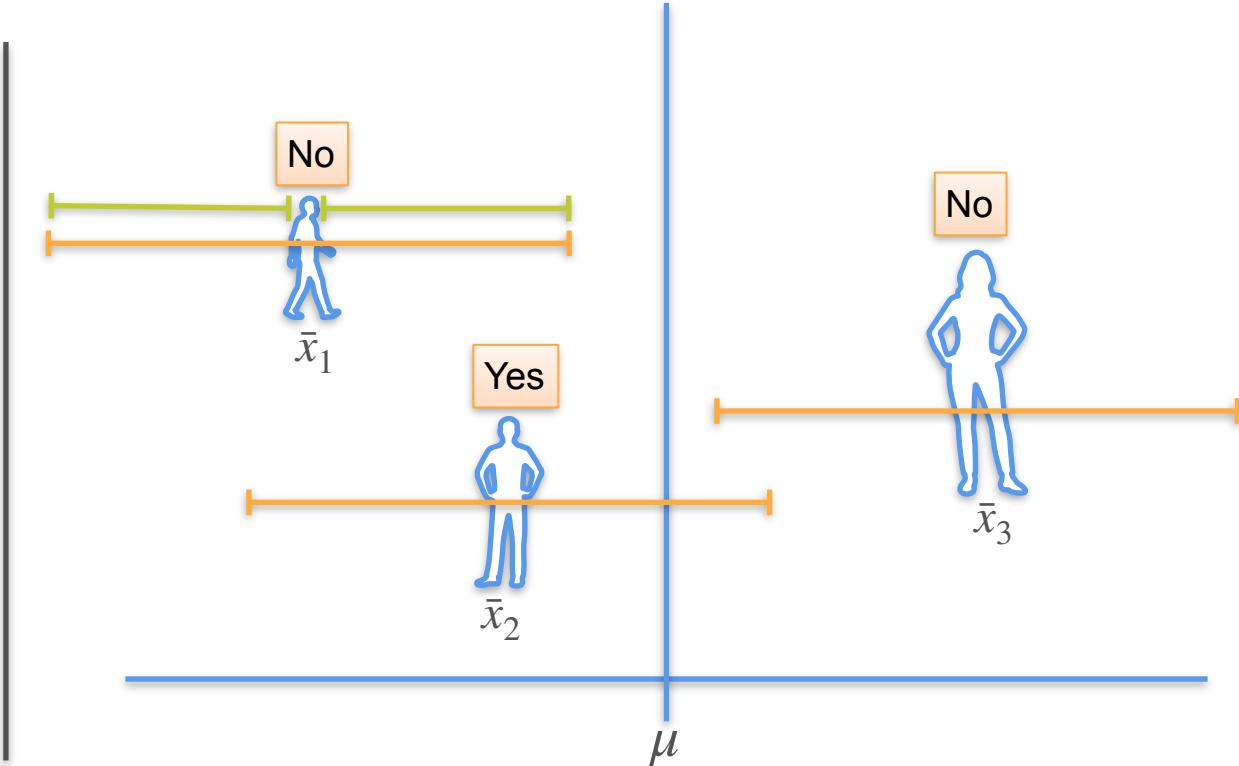
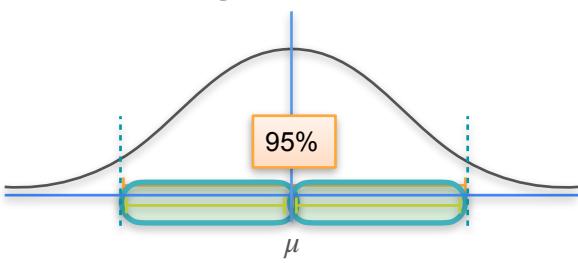
Confidence Interval - Intuition

$n = 1$

Known σ

95%

Margin of error



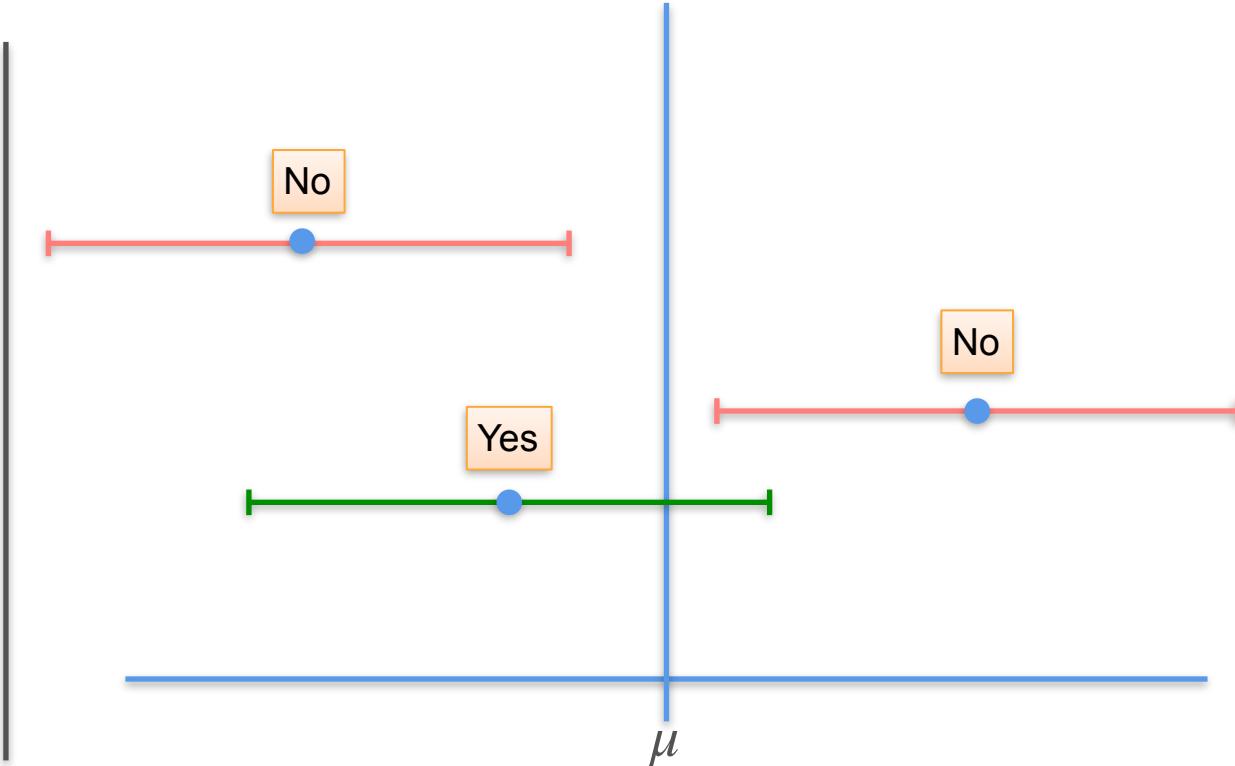
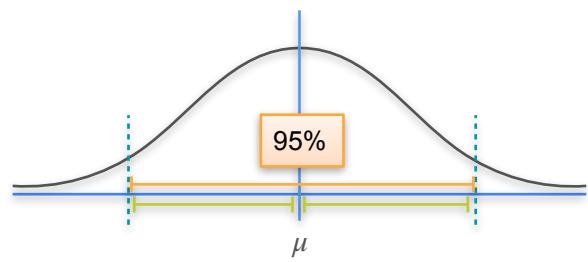
Confidence Interval - Intuition

$n = 1$

Known σ

95%

Margin of error



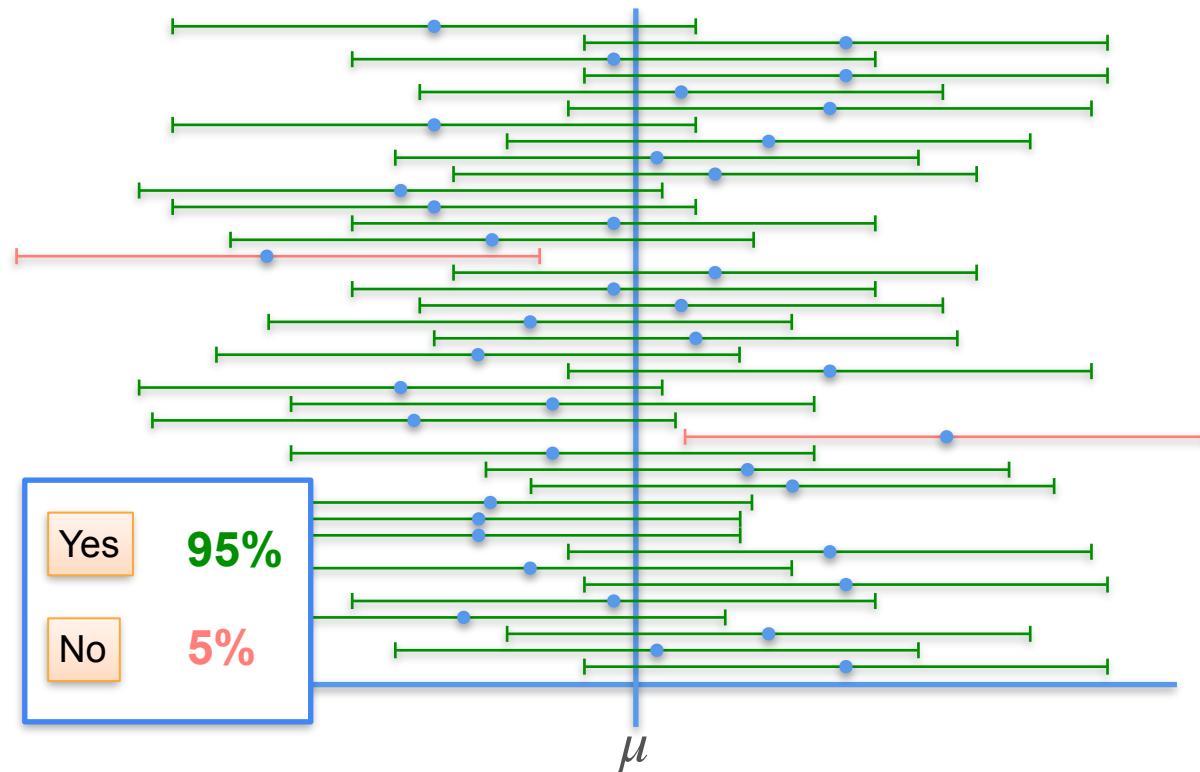
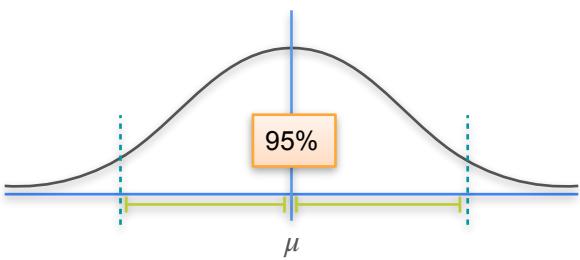
Confidence Interval - Intuition

$n = 1$

Known σ

95%

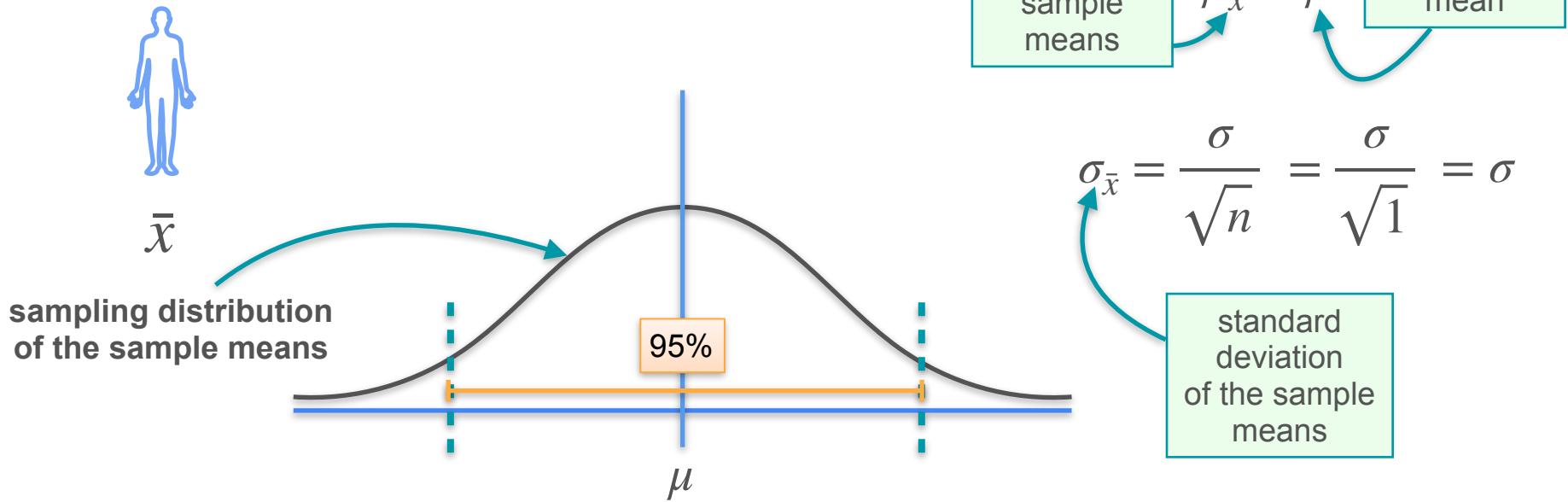
Margin of error



Confidence Interval - Intuition

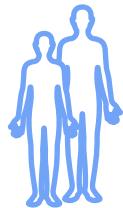
$$n = 1$$

$$\bar{X} \sim \mathcal{N}(\mu, \sigma^2)$$



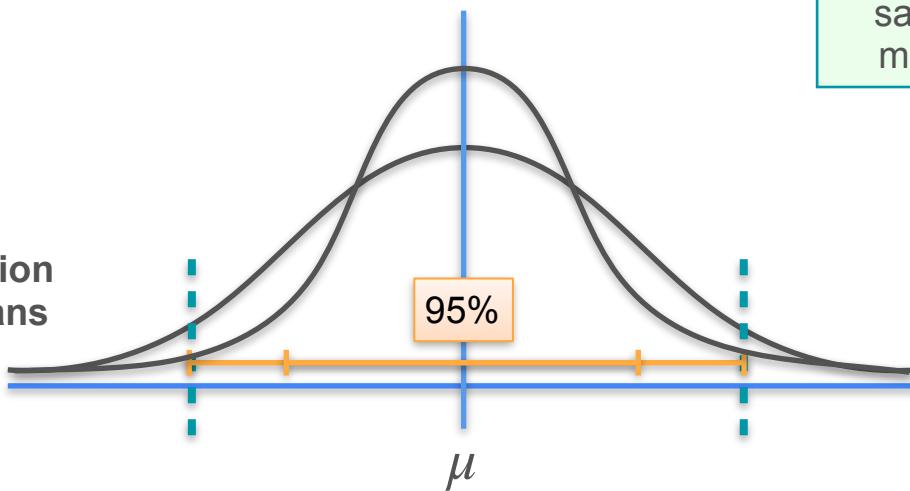
Confidence Interval - Intuition

$$n = 2$$



$$\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{2}\right)$$

sampling distribution
of the sample means



Known population standard deviation (σ)

mean of the
sample
means

$$\mu_{\bar{x}} = \mu$$

population
mean

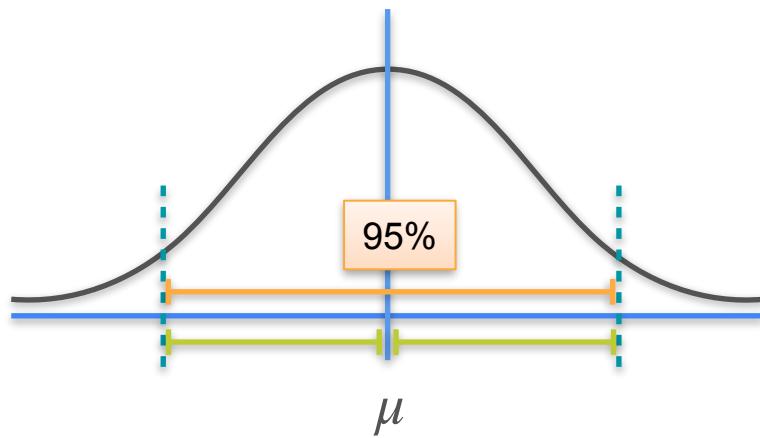
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{2}}$$

standard
deviation
of the sample
means

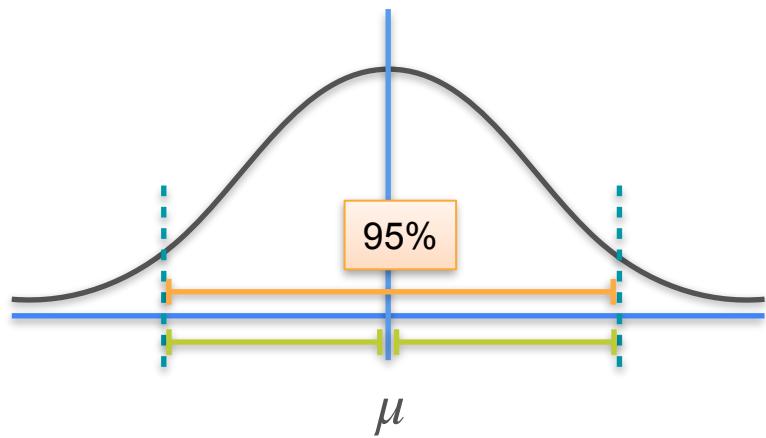
Confidence Interval - Intuition

95%

$$n = 1$$
$$\mathcal{N}\left(\mu, \frac{\sigma^2}{1}\right)$$



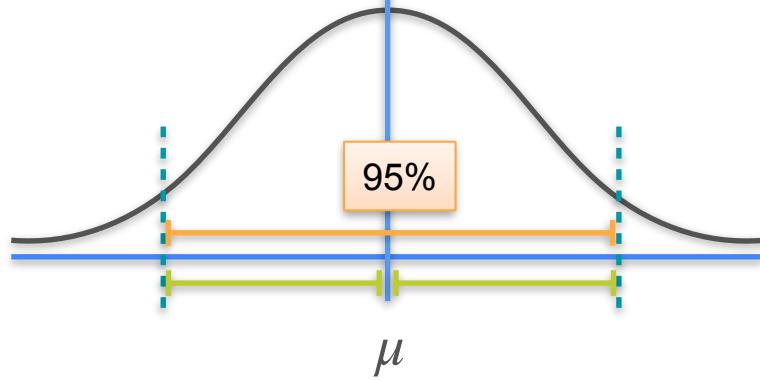
$n = 2$



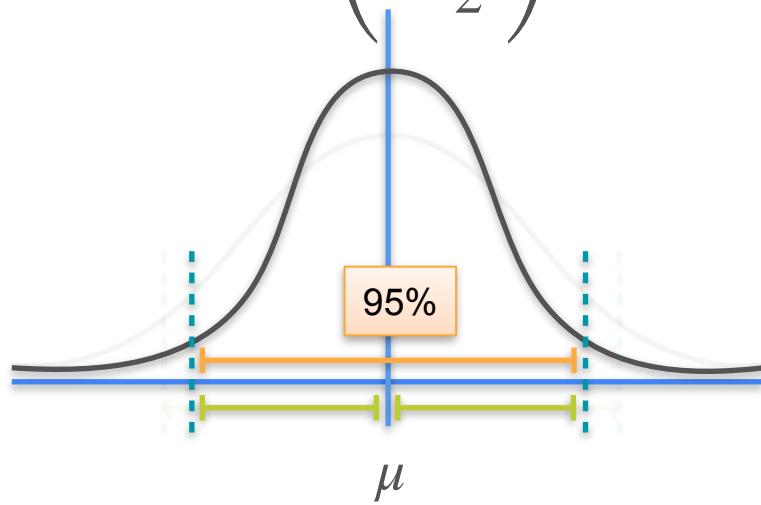
Confidence Interval - Intuition

95%

$$n = 1$$
$$\mathcal{N}\left(\mu, \frac{\sigma^2}{1}\right)$$



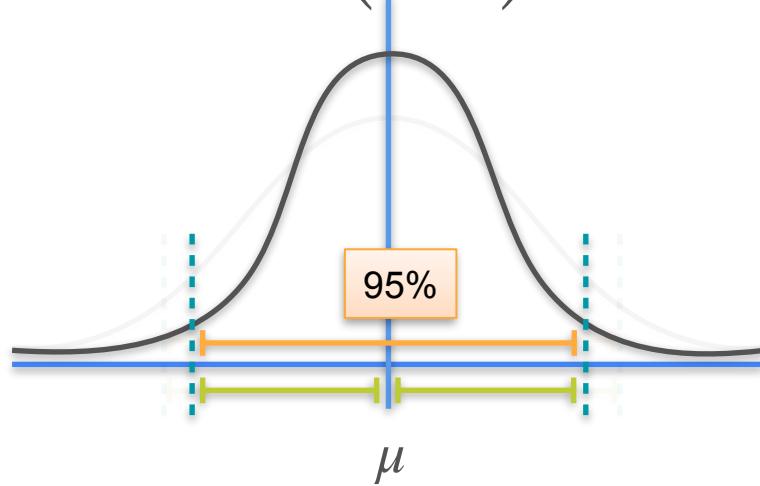
$$n = 2$$
$$\mathcal{N}\left(\mu, \frac{\sigma^2}{2}\right)$$



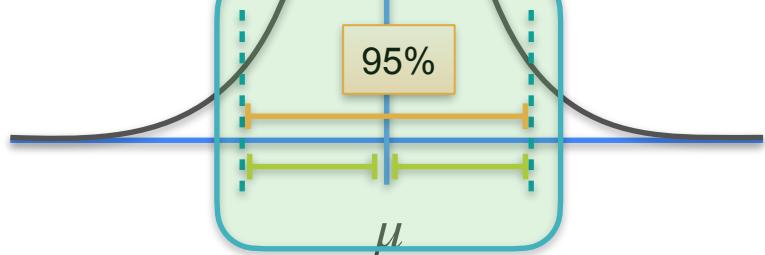
Confidence Interval - Intuition

95%

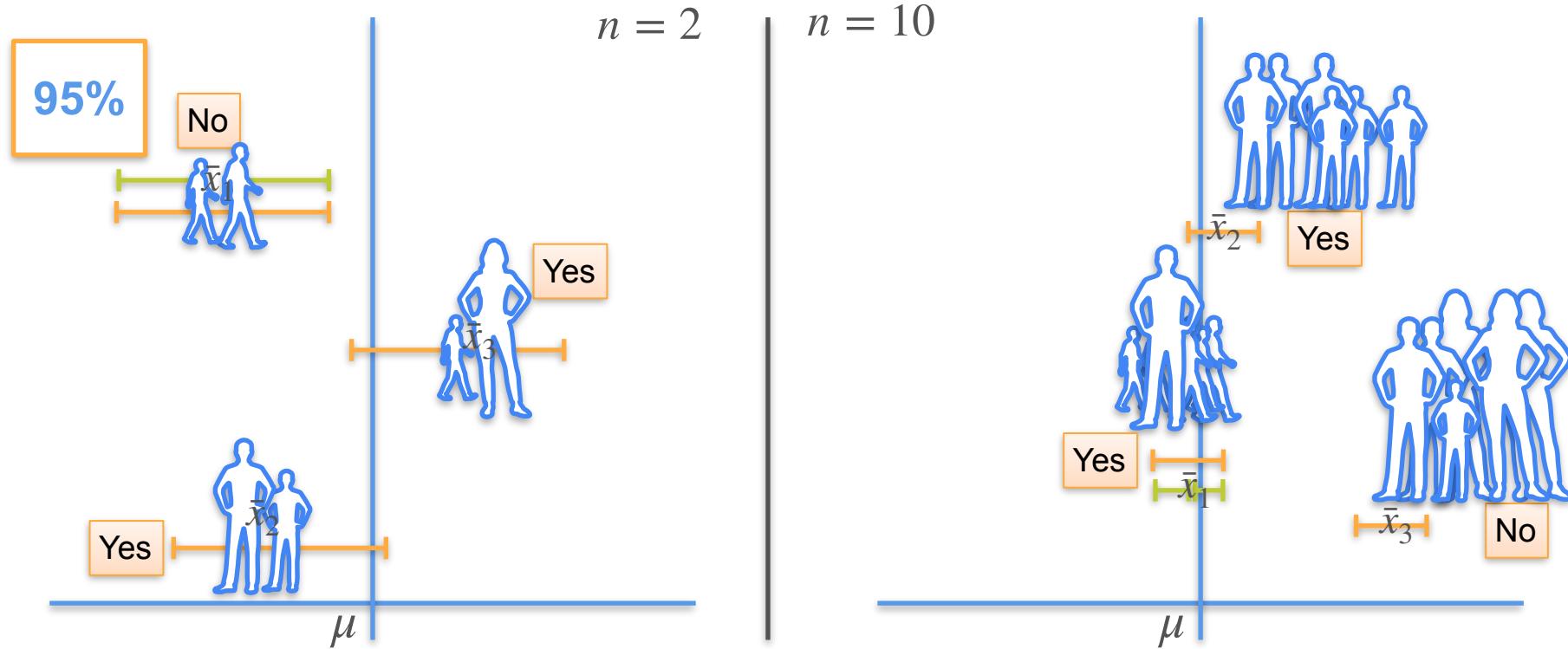
$$n = 2$$
$$\mathcal{N}\left(\mu, \frac{\sigma^2}{2}\right)$$



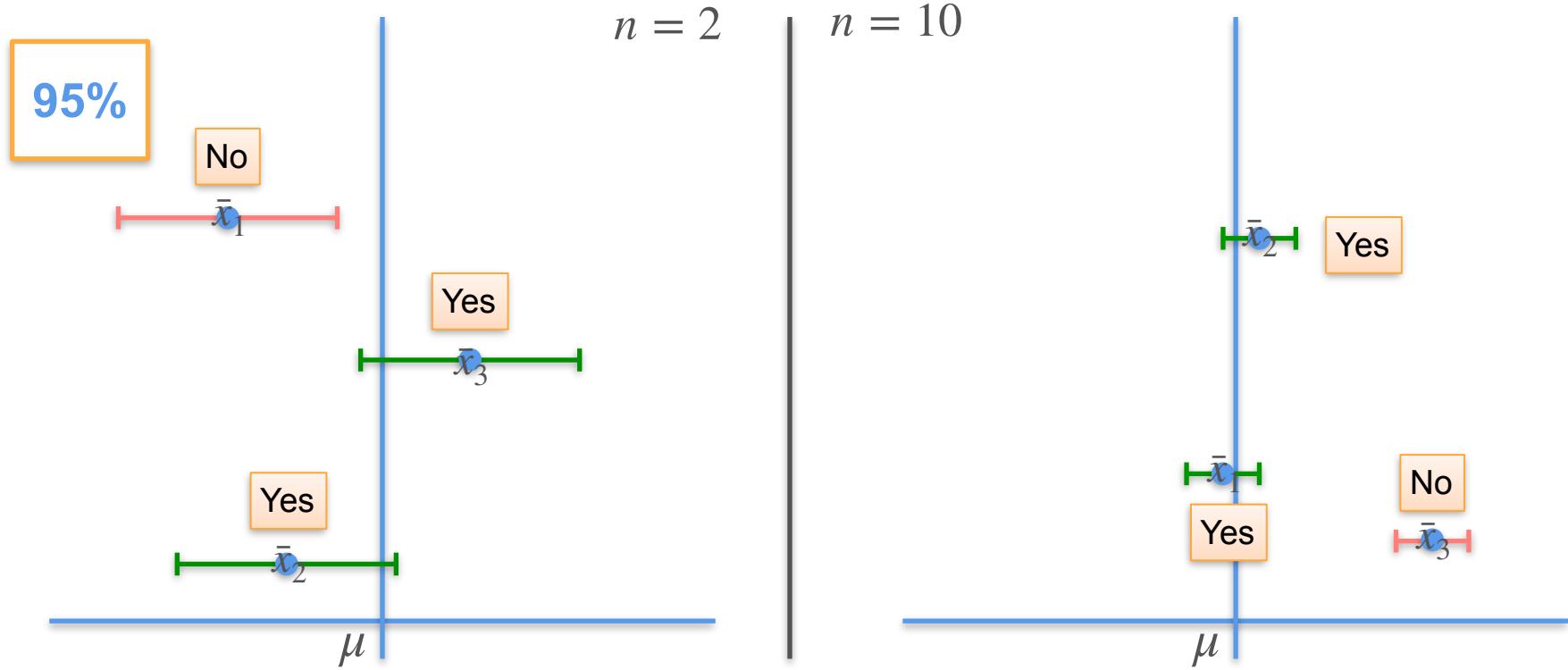
$$n = 10$$
$$\mathcal{N}\left(\mu, \frac{\sigma^2}{10}\right)$$



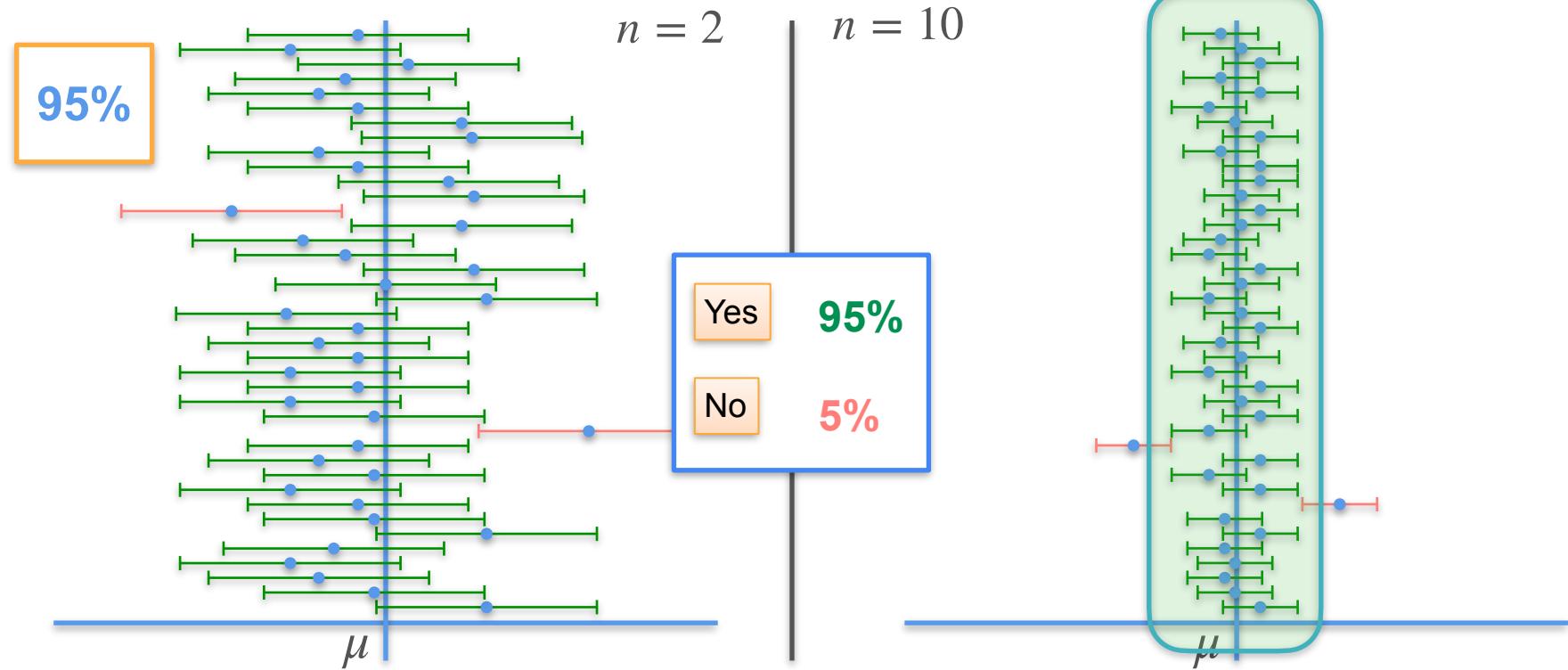
Confidence Interval - Intuition



Confidence Interval - Intuition

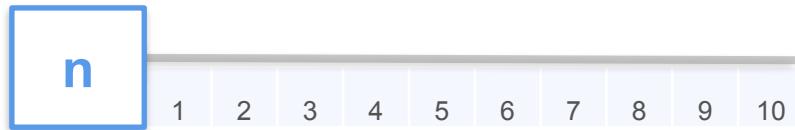


Confidence Interval - Intuition

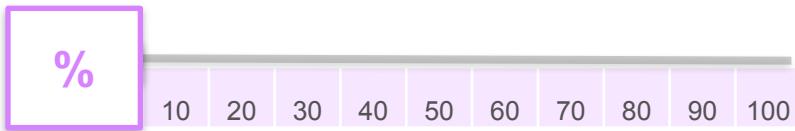


Effect of the Sample Size

sample size

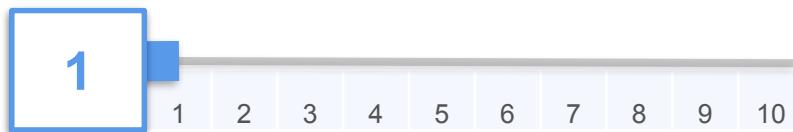


Confidence level

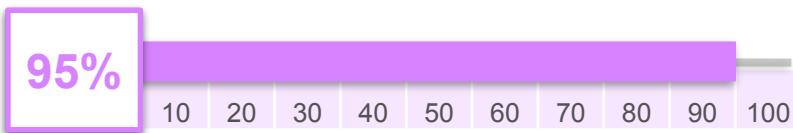


Effect of the Sample Size

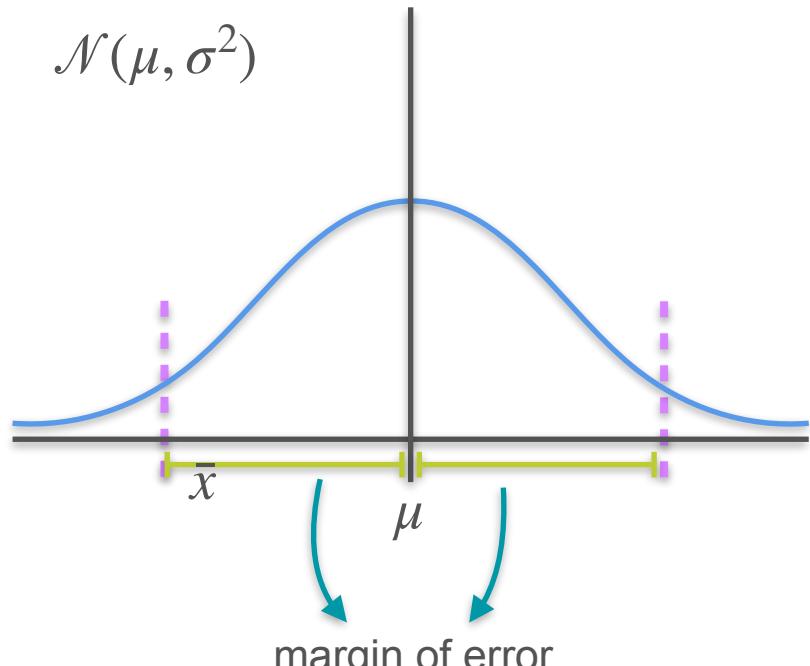
sample size



Confidence level

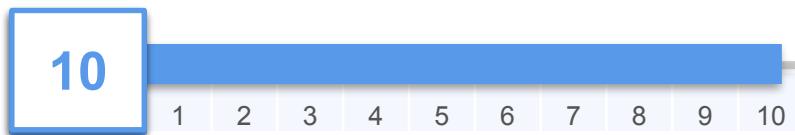


$$\mathcal{N}(\mu, \sigma^2)$$

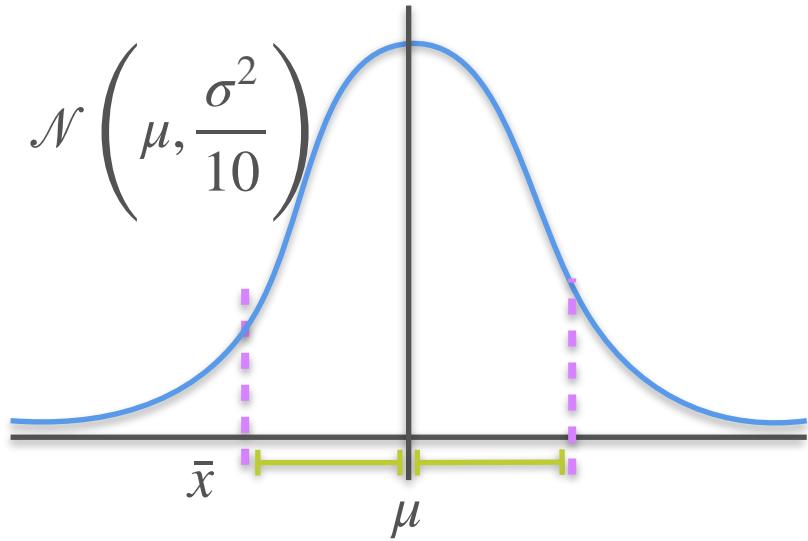
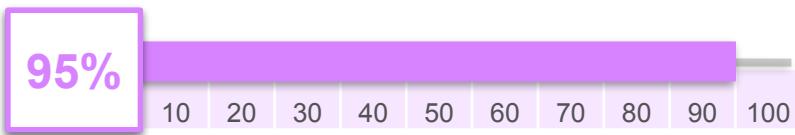


Effect of the Sample Size

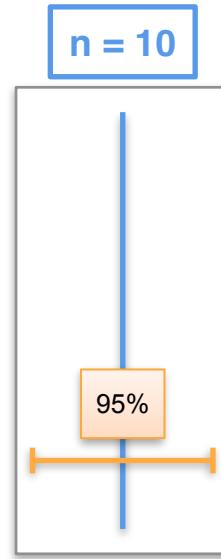
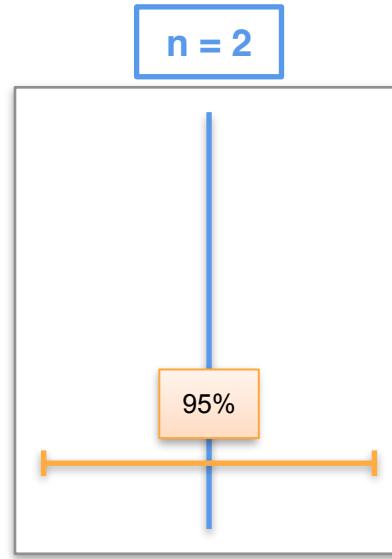
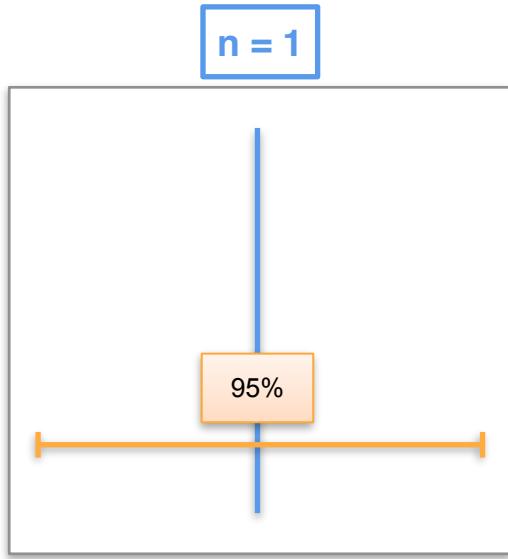
sample size



Confidence level



Effect of the Sample Size

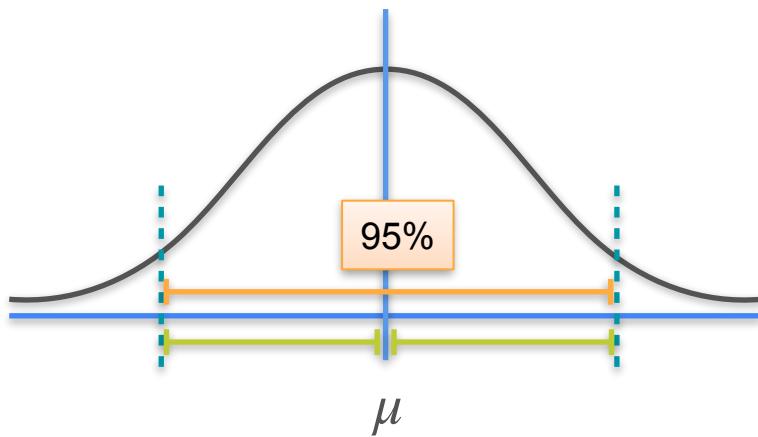


As n increases, the confidence interval shrinks

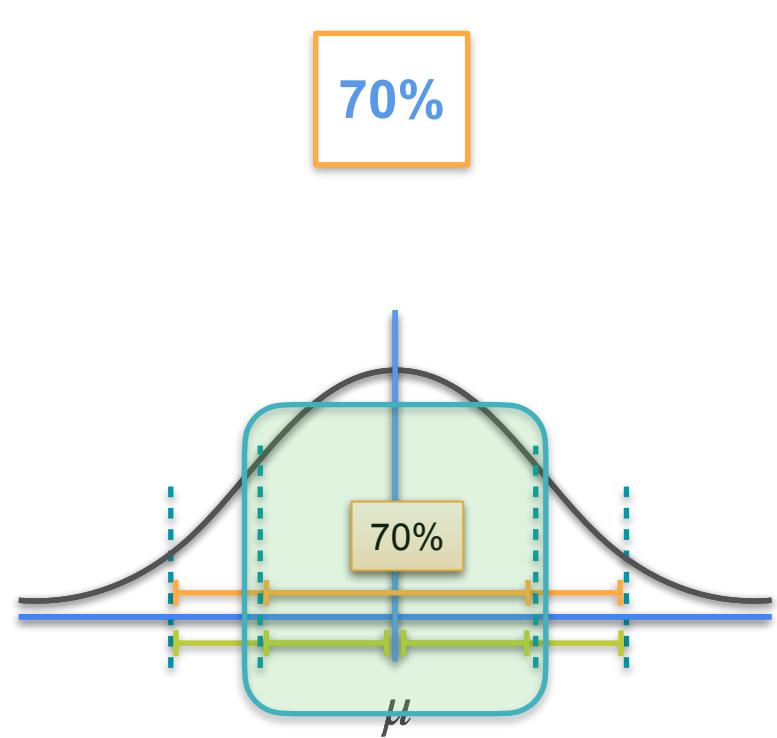
Effect of the Confidence Level

$$n = 1$$

$$\mathcal{N}(\mu, \sigma^2)$$

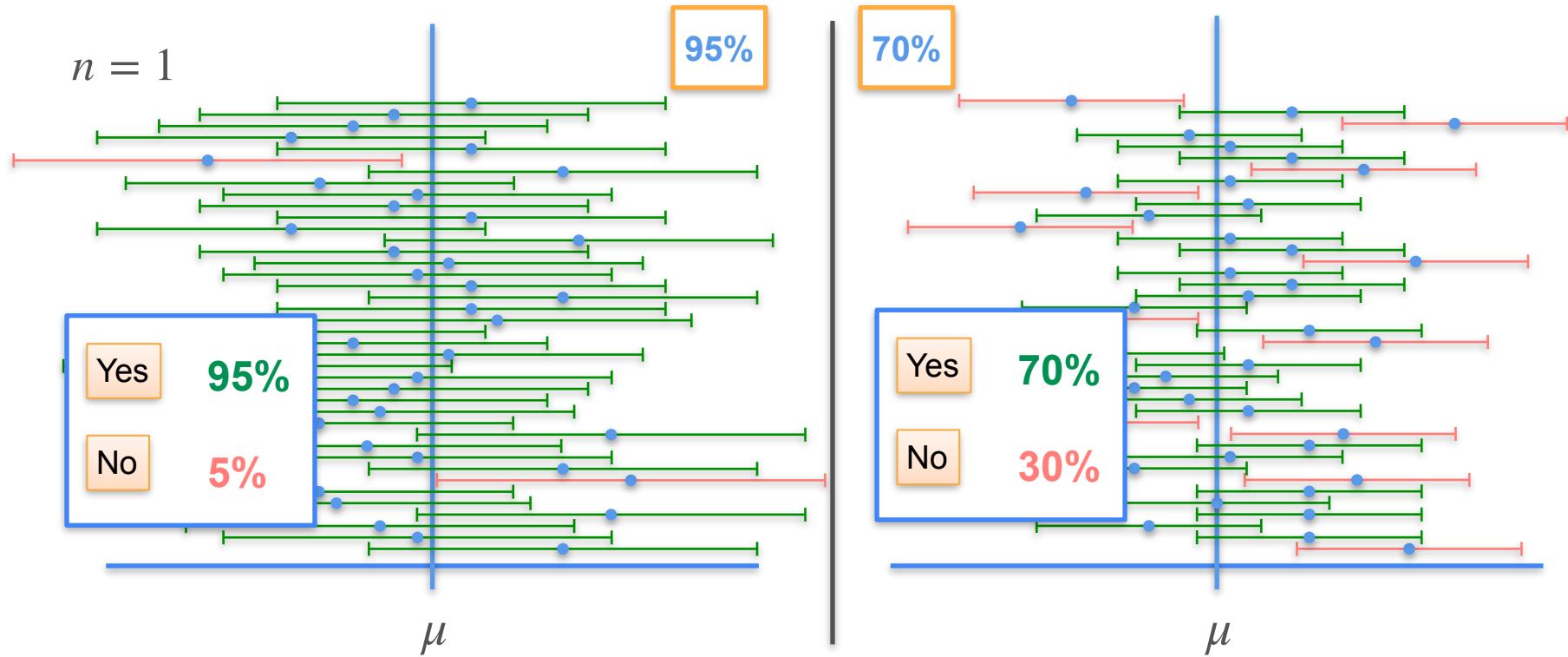


95%



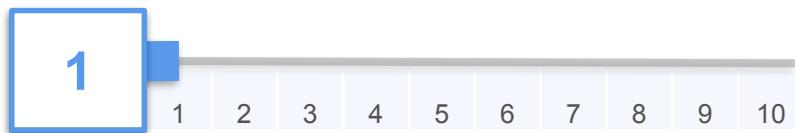
70%

Effect of the Confidence Level

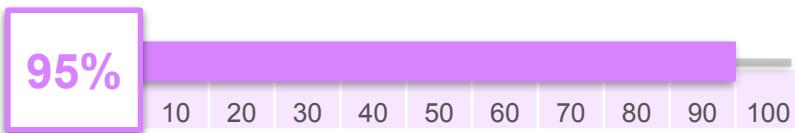


Effect of the Confidence Level

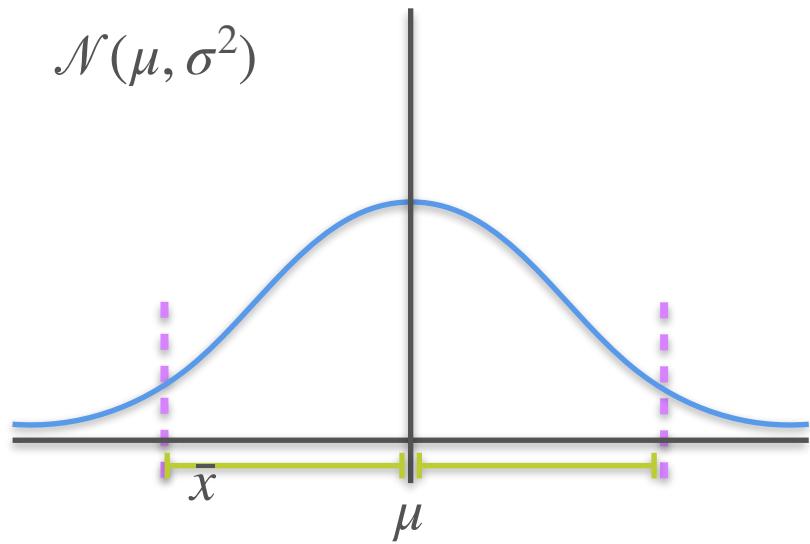
sample size



Confidence level

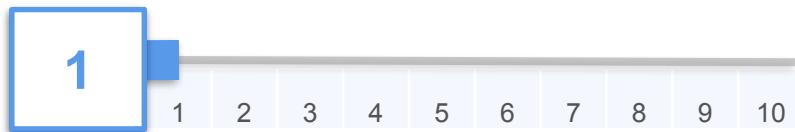


$$\mathcal{N}(\mu, \sigma^2)$$

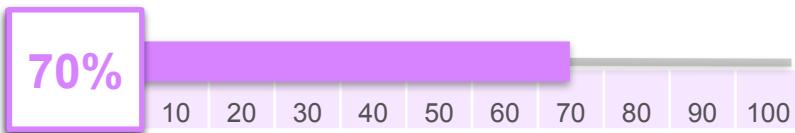


Effect of the Confidence Level

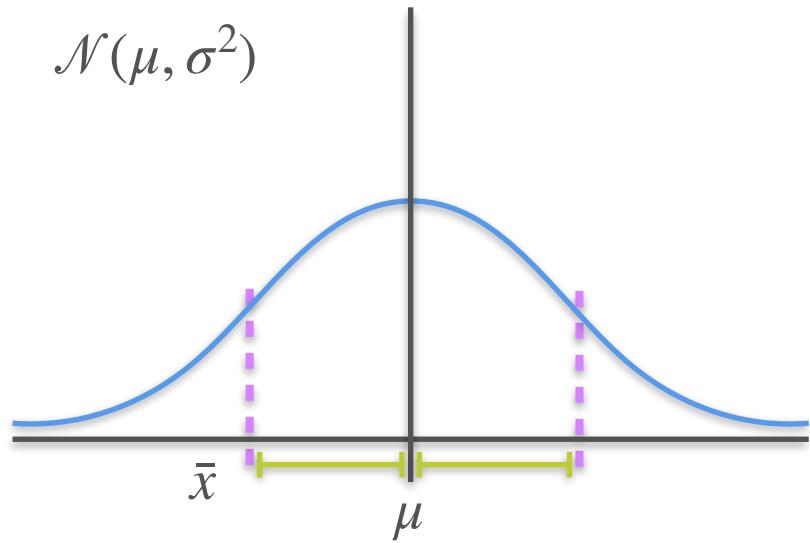
sample size



Confidence level



$$\mathcal{N}(\mu, \sigma^2)$$

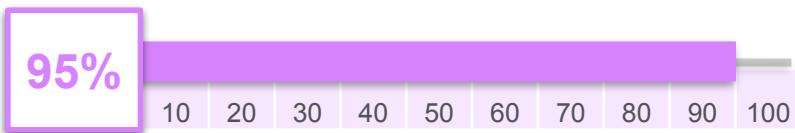


Effect of the Confidence Level

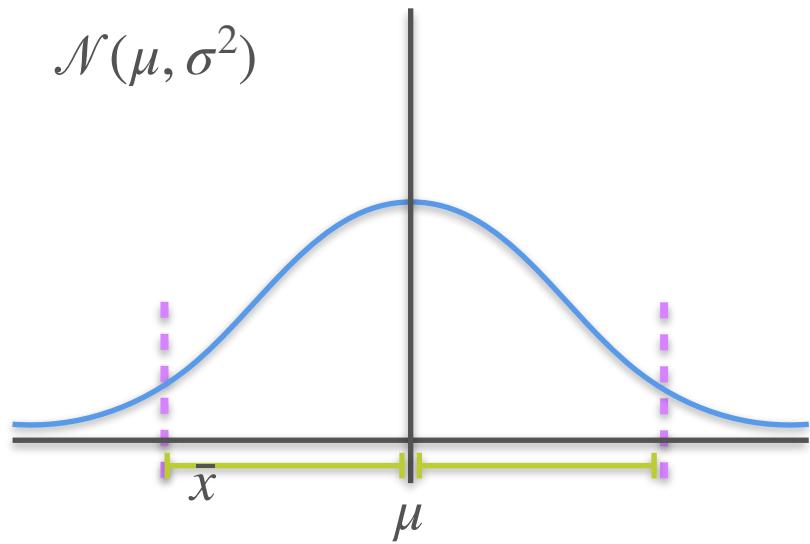
sample size



Confidence level



$$\mathcal{N}(\mu, \sigma^2)$$



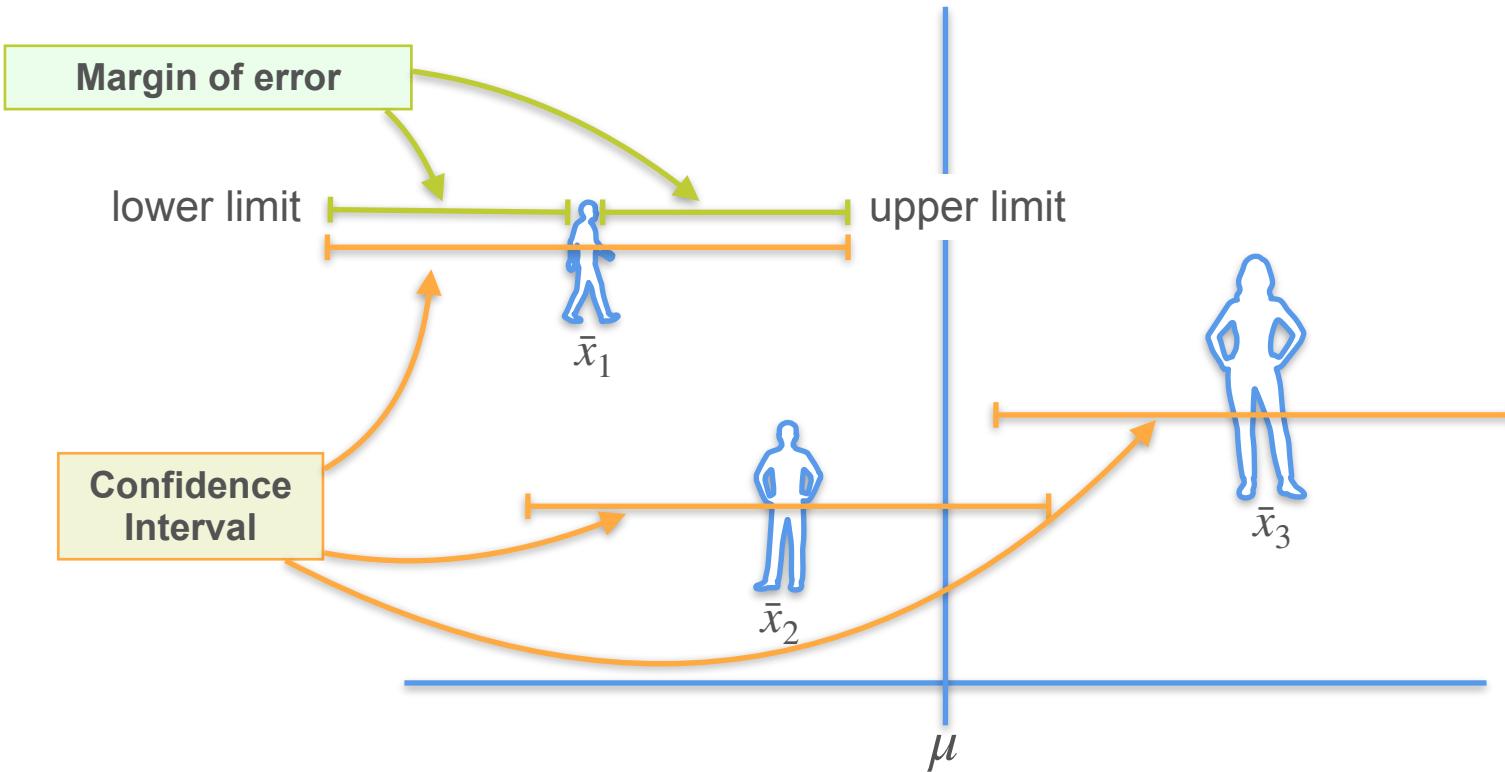


DeepLearning.AI

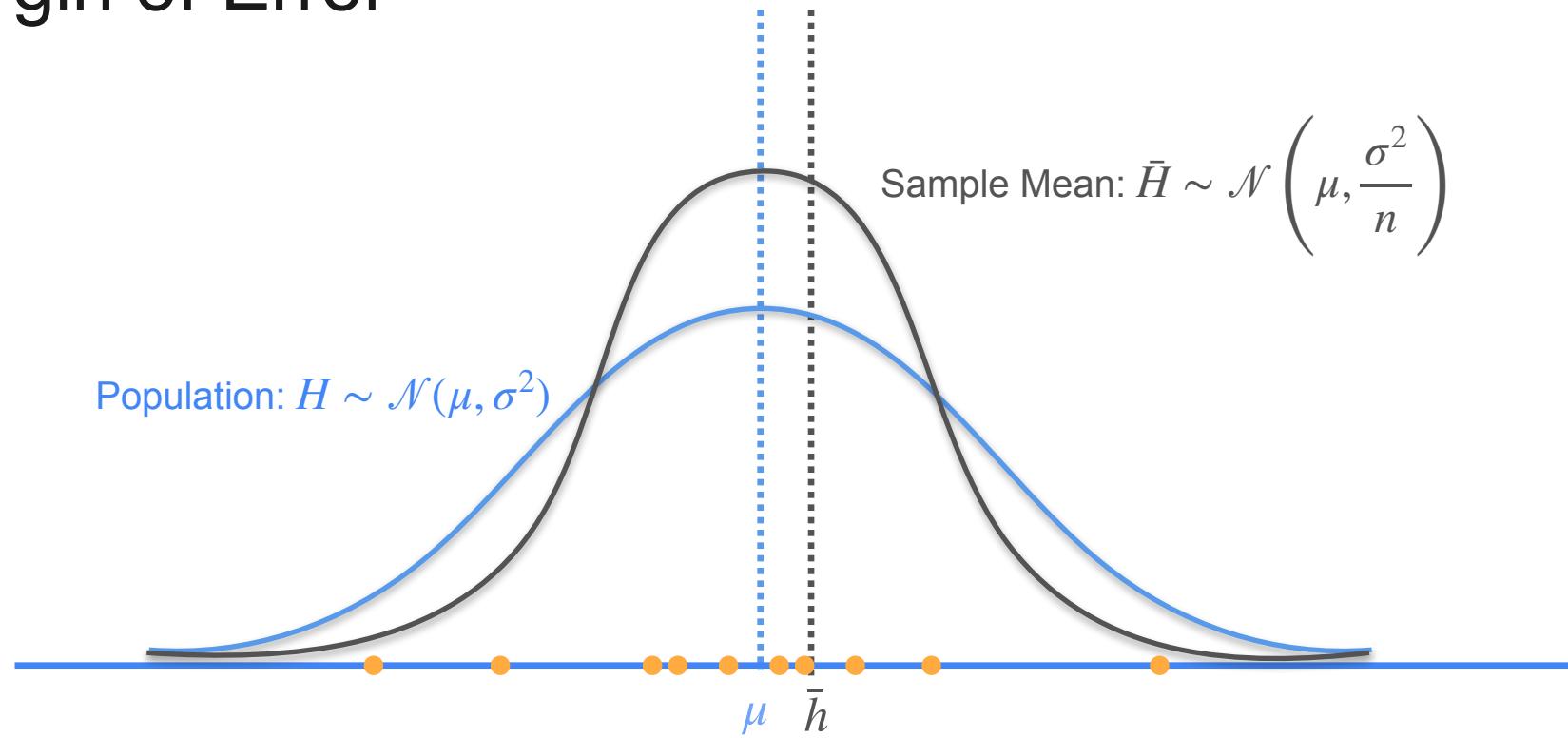
Confidence Interval

Margin of Error

Margin of Error - Introduction

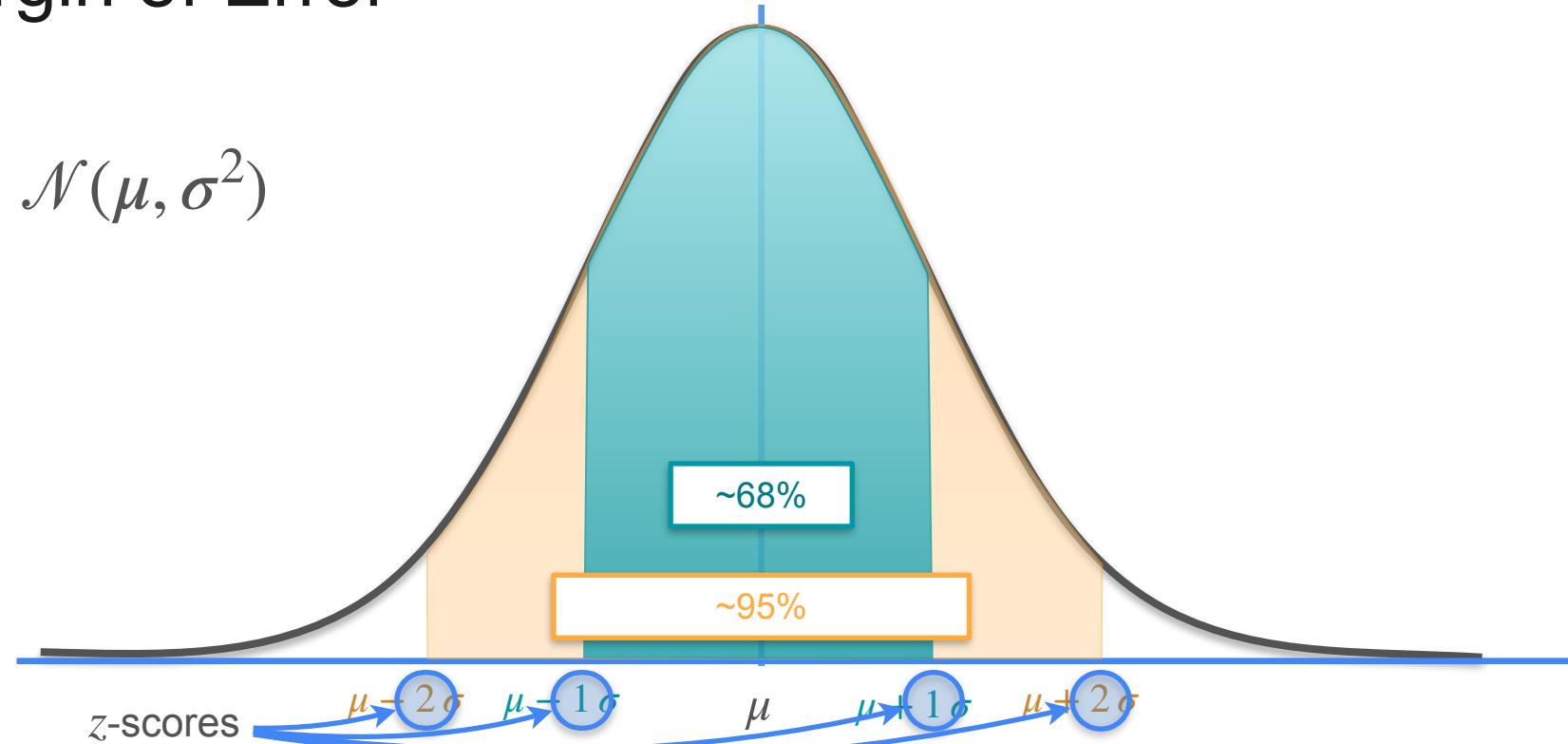


Margin of Error



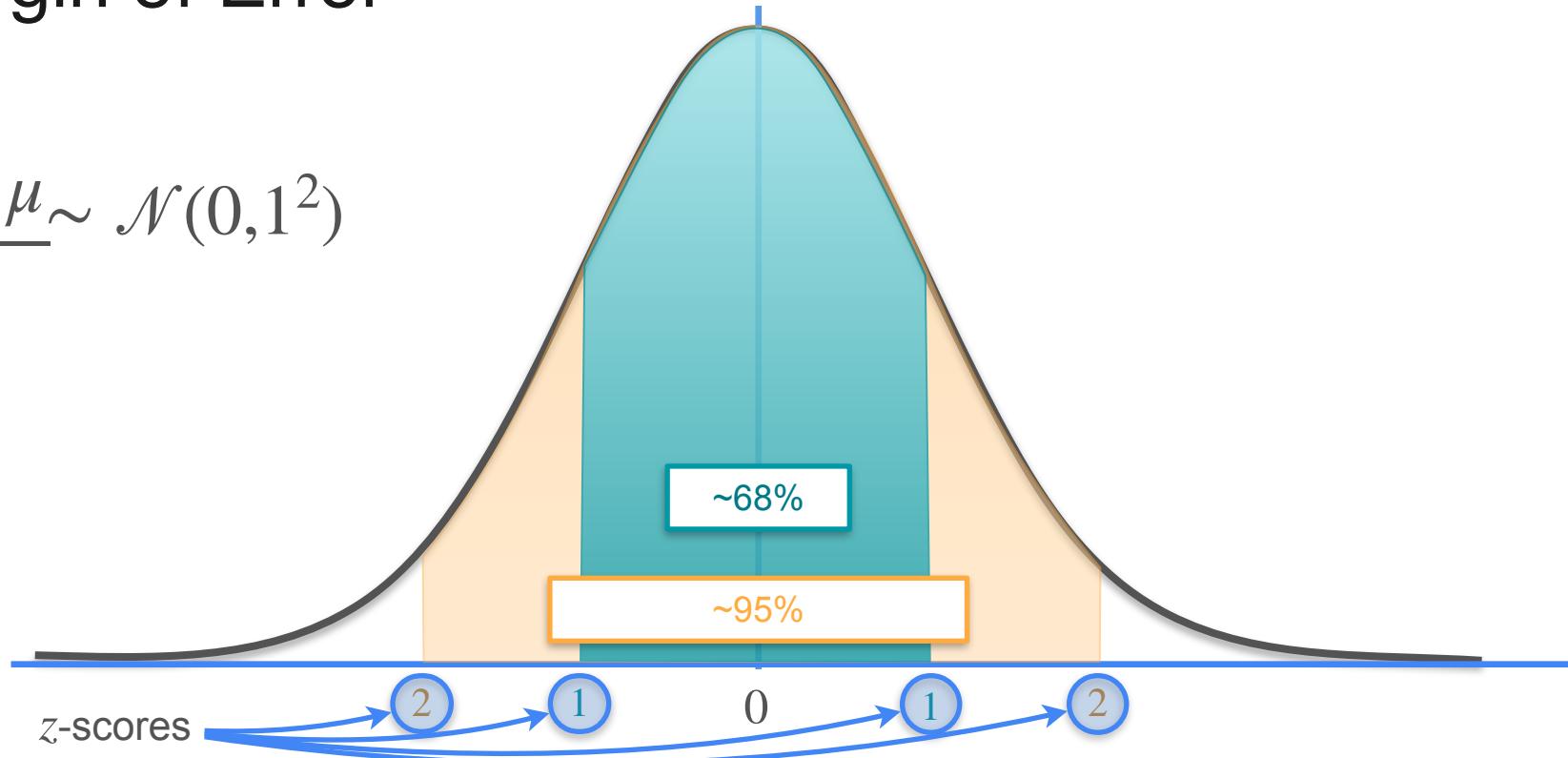
Margin of Error

$$X \sim \mathcal{N}(\mu, \sigma^2)$$



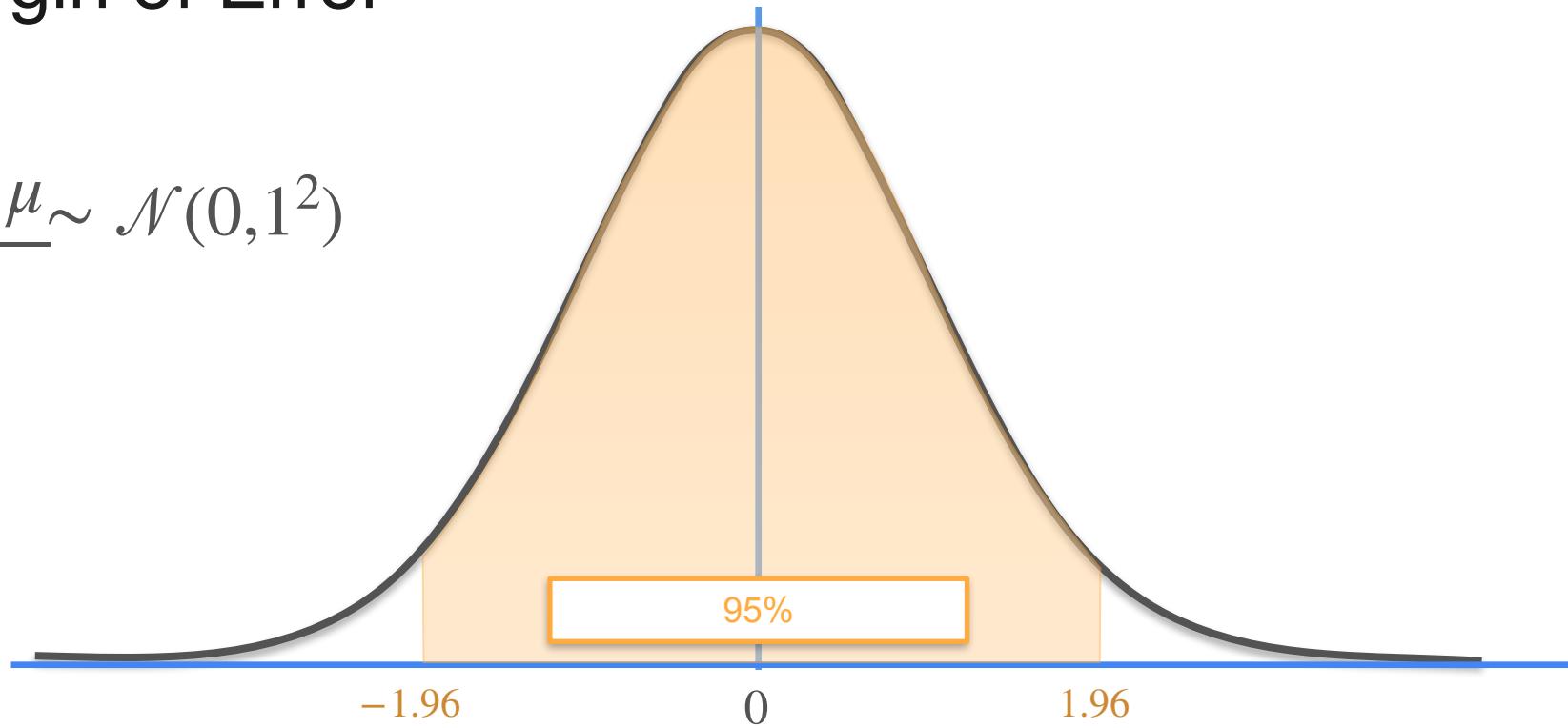
Margin of Error

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1^2)$$



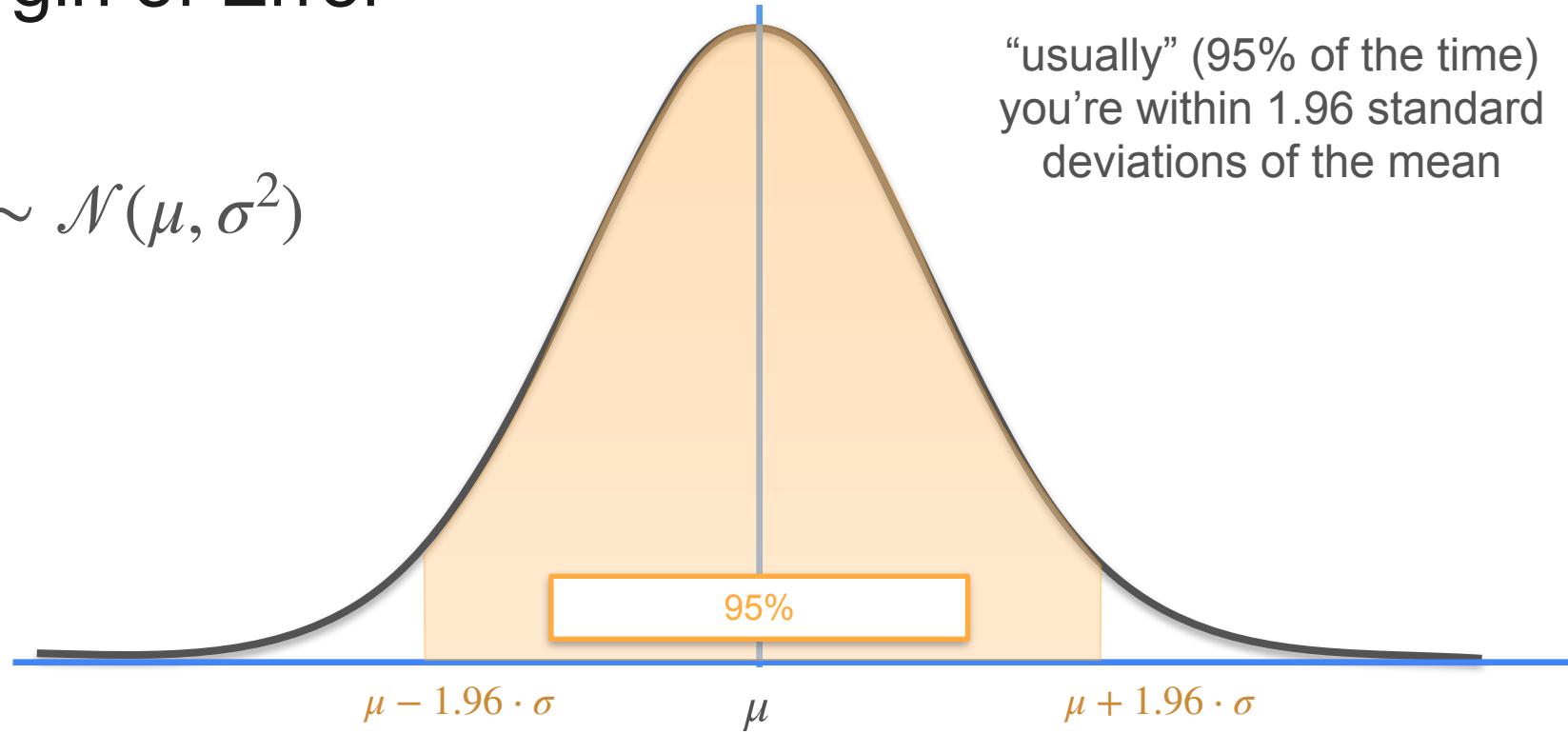
Margin of Error

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1^2)$$



Margin of Error

$$X \sim \mathcal{N}(\mu, \sigma^2)$$



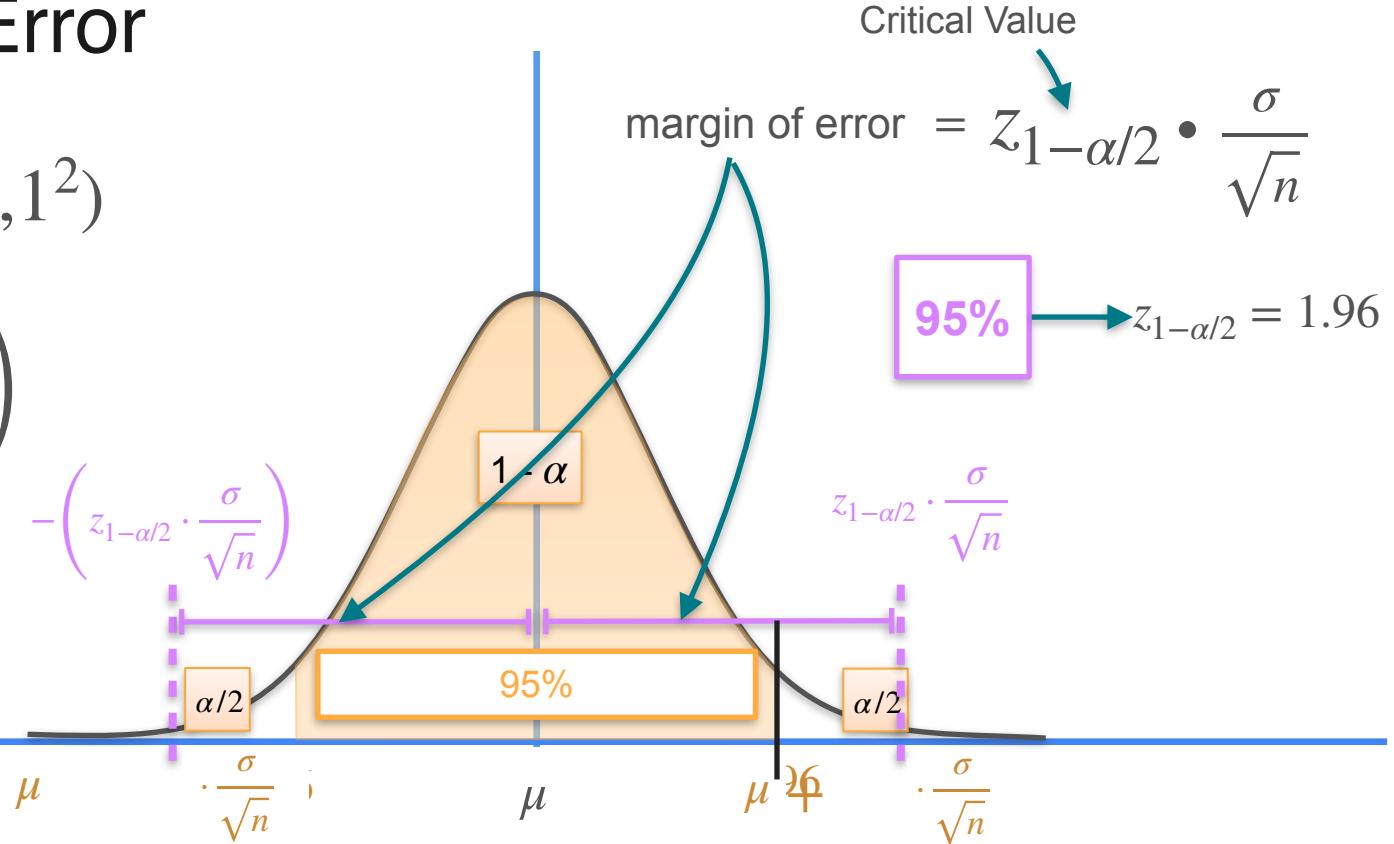
Margin of Error

$$\bar{X} - \mu \sim \mathcal{N}(0, 1^2)$$

$$\bar{X} \sim \mathcal{N}\left(\mu_{\bar{x}}, \frac{\sigma^2}{n}\right)$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Standard error





DeepLearning.AI

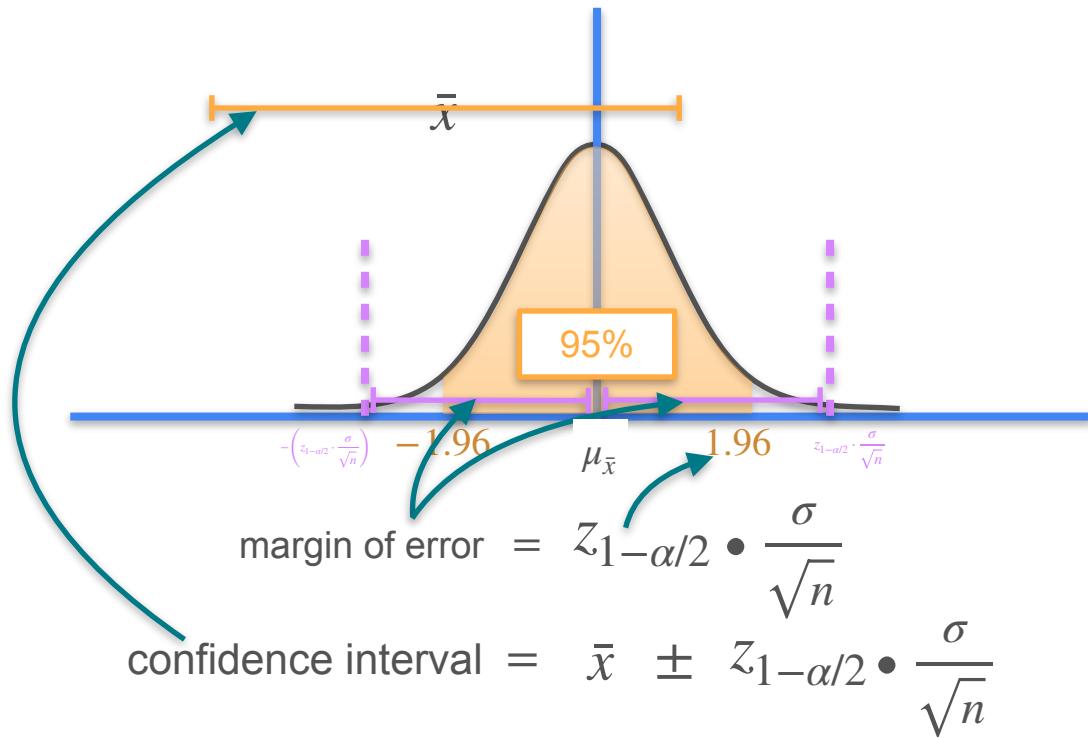
Confidence Interval

**Confidence Interval -
Calculation Steps**

Confidence Interval - Calculation Steps

STEPS:

- Find the sample mean
- Define a desired confidence level ($1 - \alpha$)
- Get the critical value ($z_{1-\alpha/2}$)
- Find the standard error ($\frac{\sigma}{\sqrt{n}}$)
- Find the margin of error
- Add/subtract the margin of error to the sample mean



Confidence Interval - Calculation Steps

STEPS:

- Find the sample mean
- Define a desired confidence level ($1 - \alpha$)
- Get the critical value ($z_{1-\alpha/2}$)
- Find the standard error ($\frac{\sigma}{\sqrt{n}}$)
- Find the margin of error
- Add/subtract the margin of error to the sample mean

$$\text{confidence interval} = \bar{x} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Assumptions

- Simple random sample
- Sample size > 30 or population is approximately normal



DeepLearning.AI

Confidence Interval

**Confidence Interval -
Example**

Confidence Interval - Example

Statistopia

6,000 adults

Random Selection



$$\bar{x} = 170\text{cm}$$

$$\sigma = 25\text{cm}$$

95%

$$\rightarrow z_{1-\alpha/2} = 1.96$$

Calculate a 95% confidence interval for the average height of adults on Statistopia.

Confidence Interval - Example

Random Selection



49

$\sigma = 25cm$

95%

$$z_{1-\alpha/2} = 1.96$$

Confidence Interval

$170cm \pm \text{margin of error}$

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$= 1.96 \cdot \frac{25}{\sqrt{49}}$$

$$= 1.96 \cdot \frac{25}{7}$$

$$= 7$$

Confidence Interval - Example

Random Selection

2500



$$\sigma = 10\text{cm}$$

95%

$$z_{1-\alpha/2} = 1.96$$

Confidence Interval

$$170\text{cm} \pm \text{margin of error}$$

$$\text{margin of error} = 7$$

Confidence Interval

$$170\text{cm} - 7 = 163\text{cm}$$

$$170\text{cm} + 7 = 177\text{cm}$$

$$163\text{cm} < \mu < 177\text{cm}$$



DeepLearning.AI

Confidence Interval

Calculating Sample Size

Calculating Sample Size

6,000 adults

Random Selection

49

Margin of error: 7cm

$$\bar{x} \pm 7\text{cm}$$

$$163\text{cm} < \mu < 177\text{cm}$$



95% → $z_{1-\alpha/2} = 1.96$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Calculating Sample Size

6,000 adults

Random Selection

Margin of error: 7cm

$$\bar{x} \pm 7\text{cm}$$



95%

$$\rightarrow z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm}$$

$$\sigma$$

What is the smallest sample size to obtain the
desired margin of error?

Margin of error: 3 cm

$$\bar{x} \pm 3\text{cm}$$

Calculating Sample Size

6,000 adults

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$3 = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$3 \geq z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$3 = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$3 \geq z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$3 \geq z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$3 \geq 1.96 \times \frac{25}{\sqrt{n}}$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$3 \geq z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\frac{3}{1.96} \geq \frac{25}{\sqrt{n}}$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$3 \geq z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\frac{3}{1.96 \times 25} \geq \frac{1}{\sqrt{n}}$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$3 \geq z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\frac{1.96 \times 25}{3} \leq \frac{\sqrt{n}}{1}$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%

$$z_{1-\alpha/2} = 1.96$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$\frac{1.96 \times 25}{3} \leq \frac{\sqrt{n}}{1}$$

$$\left(\frac{1.96 \times 25}{3} \right)^2 \leq n$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%  $z_{1-\alpha/2} = 1.96$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

Margin of error: 3 cm

$$\frac{1.96 \times 25}{3} \leq \frac{\sqrt{n}}{1}$$

$$n \geq \left(\frac{1.96 \times 25}{3} \right)^2$$

Calculating Sample Size

6,000 adults

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

95%  $z_{1-\alpha/2} = 1.96$

$$n \geq \left(\frac{1.96 \times 25}{3} \right)^2$$

$$\bar{x} = 170\text{cm} \quad \sigma = 25\text{cm}$$

$$n \geq 266.78 \approx 267$$

Margin of error: 3 cm

Calculating Sample Size

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$n \geq \left(\frac{1.96 \times 25}{3} \right)^2$$

$$n \geq \left(\frac{z_{\alpha/2} \cdot \sigma}{MOE} \right)^2$$



DeepLearning.AI

Confidence Interval

**Difference Between
Confidence and Probability**

Difference Between Confidence and Probability

95%
Confidence
Level



The confidence interval contains the true population parameter approximately 95% of the time.

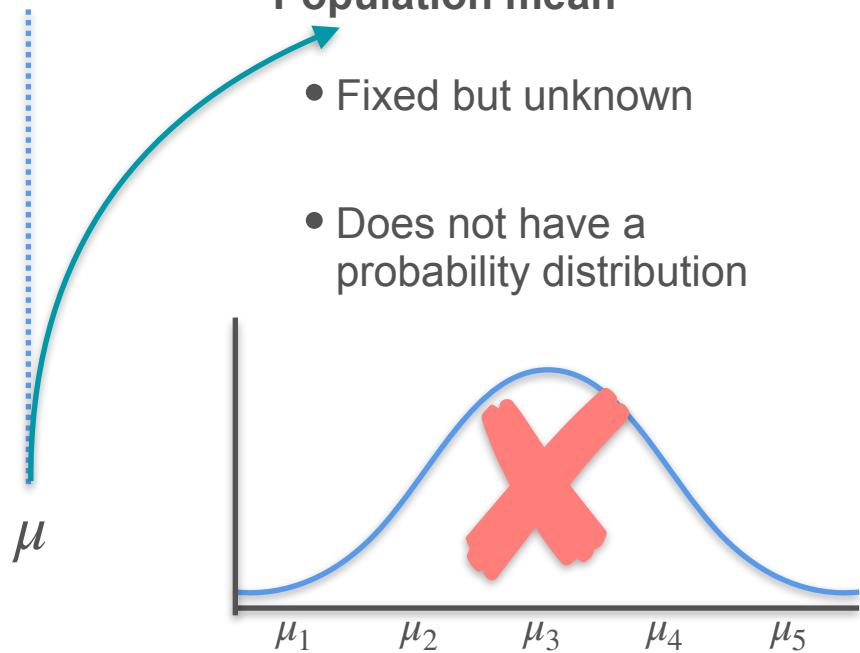


There's a 95% probability that the population parameter falls within the confidence interval.



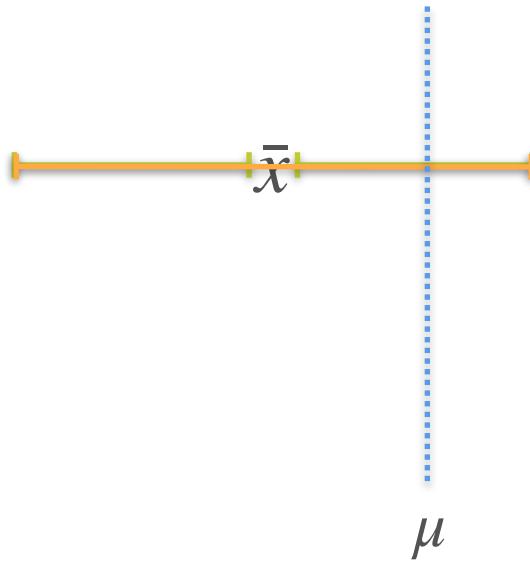
Difference Between Confidence and Probability

95%
Confidence
Level



Difference Between Confidence and Probability

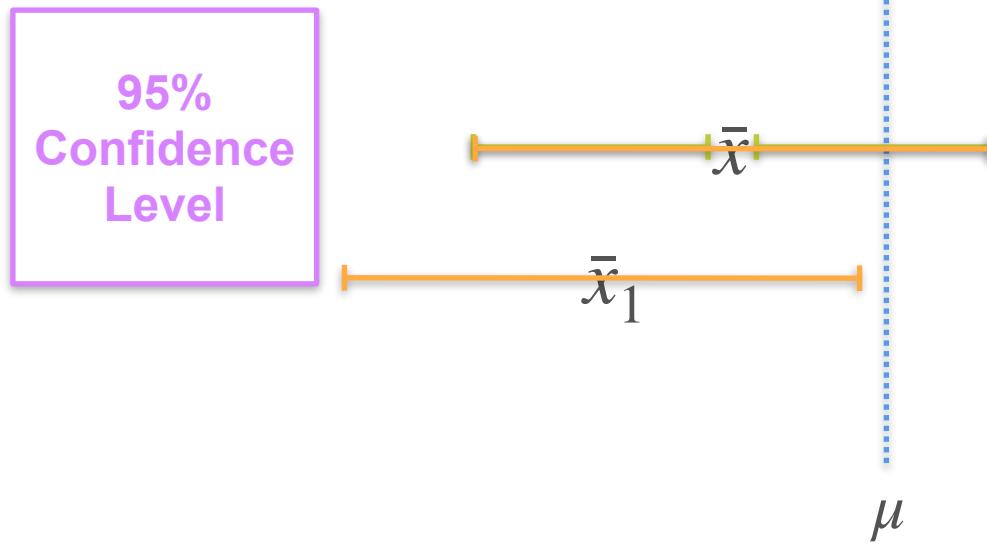
95%
Confidence
Level



Population mean

- Fixed but unknown
- Does not have a probability distribution
- In the interval

Difference Between Confidence and Probability

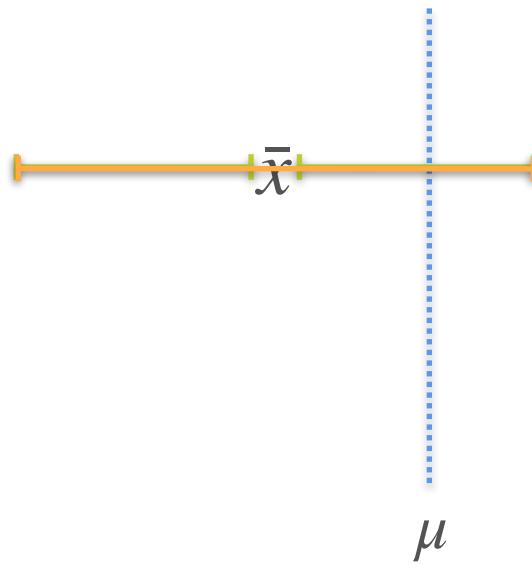


Population mean

- Fixed but unknown
- Does not have a probability distribution
- In the interval or not

Difference Between Confidence and Probability

95%
Confidence
Level



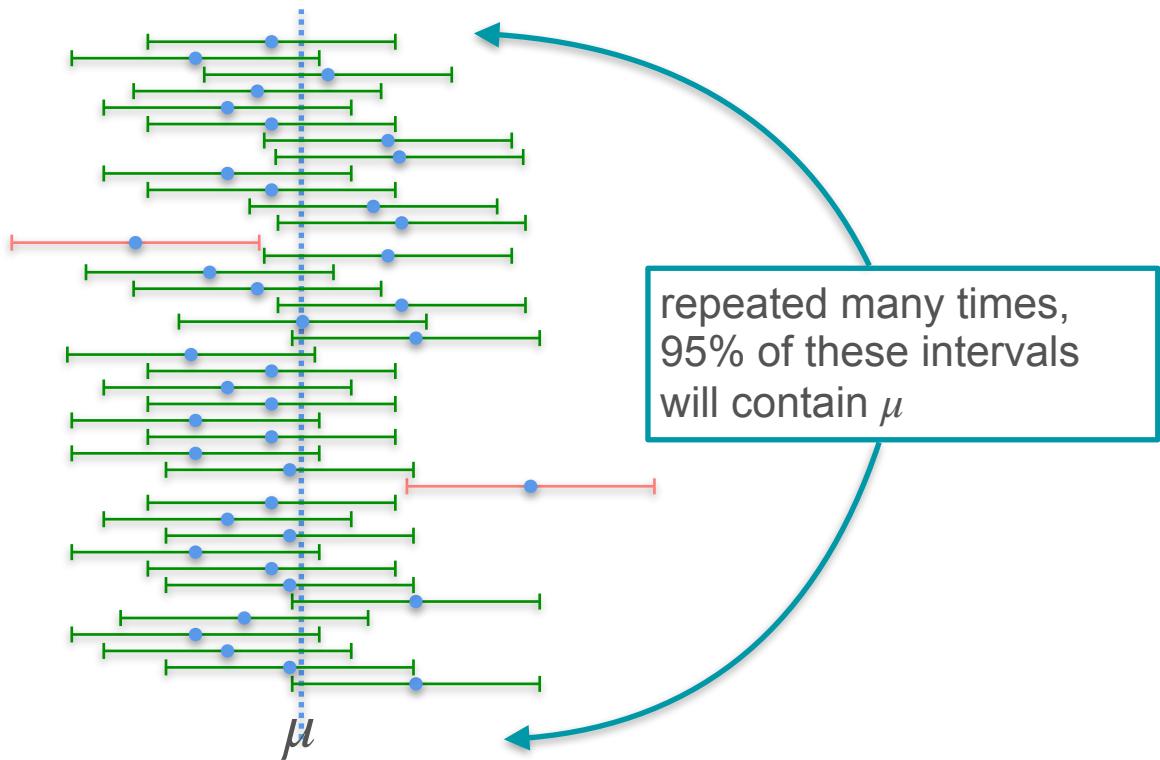
Population mean

- Fixed but unknown
- Does not have a probability distribution
- In the interval or not
- Does not fall within a specific interval 95% of the time

Difference Between Confidence and Probability

95%
Confidence
Level

success rate for constructing
the confidence interval

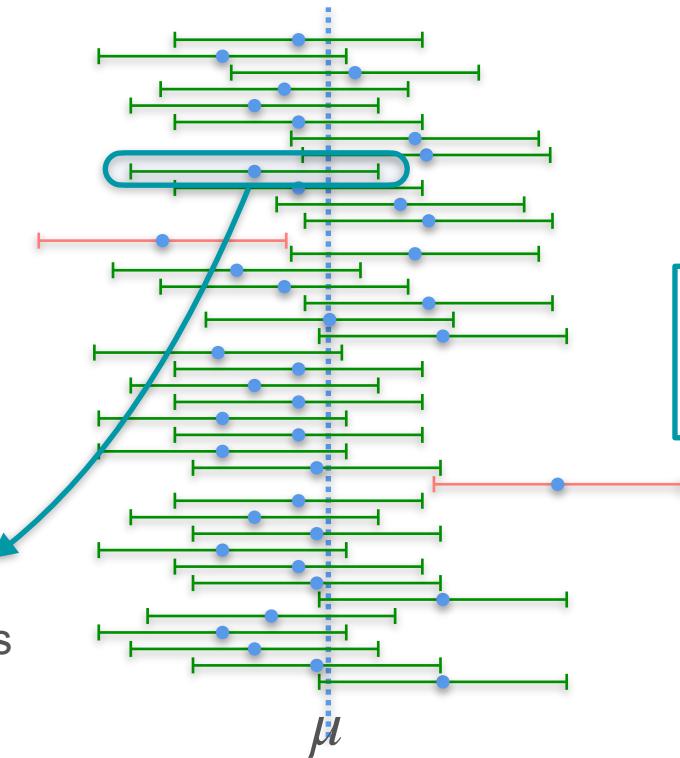


Difference Between Confidence and Probability

95%
Confidence
Level

success rate for constructing
the confidence interval

not the probability that
one specific intervals contains
the population mean



repeated many times,
95% of these intervals
will contain μ



DeepLearning.AI

Confidence Interval

Unknown Standard Deviation

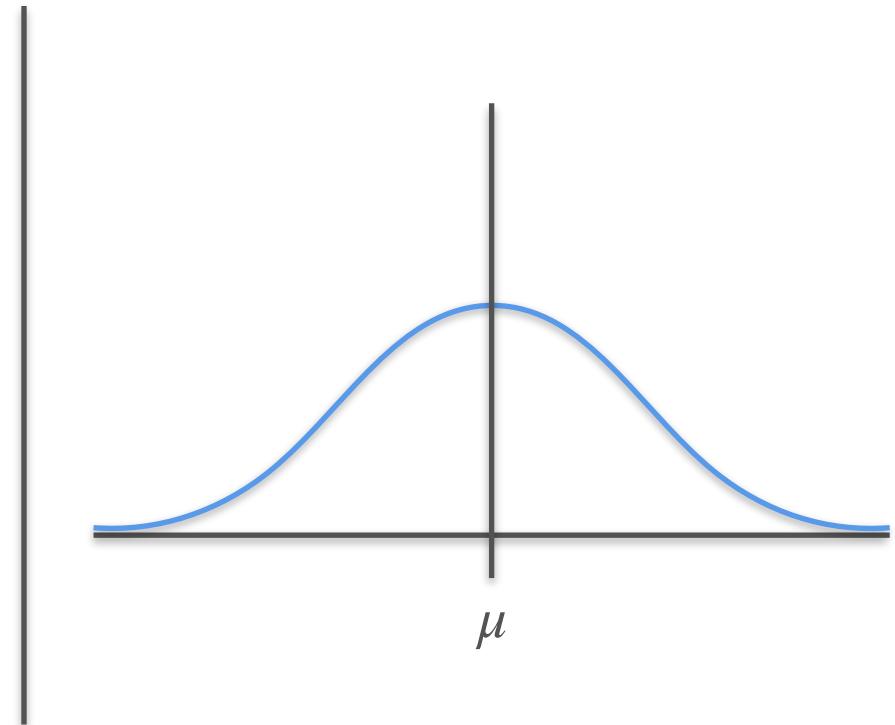
Confidence Interval - t Distribution

known σ

$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

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

normal
distribution



Confidence Interval - t Distribution

known σ ? $\rightarrow s$

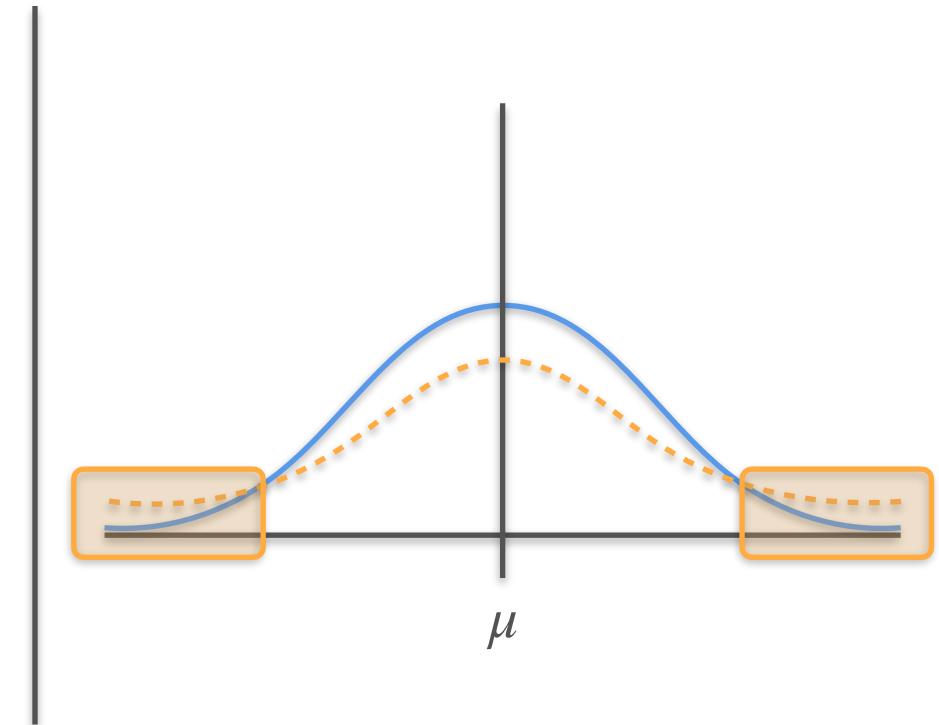
$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{\text{dotted arrow}} \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

normal distribution

not a normal distribution

student's t distribution



Confidence Interval - t Distribution

known σ

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

normal distribution

$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$



unknown σ

replace with s

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

student's t distribution

$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$



Confidence Interval - t Distribution

known σ

$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$



unknown σ

$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$



$$\bar{x} \pm t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$



Confidence Interval - t Distribution

known σ

$$\bar{x} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

unknown σ

$$\bar{x} \pm t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

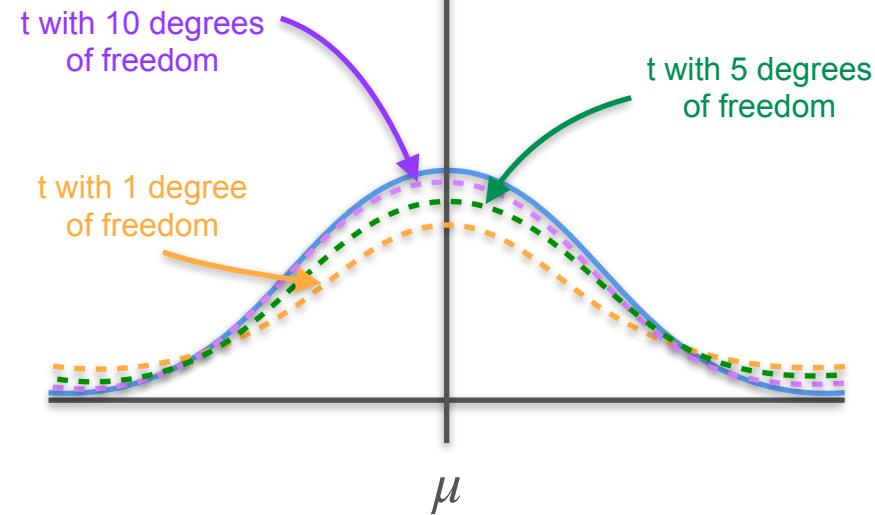
Confidence Interval - t Distribution

unknown σ

$$\bar{x} \pm t_{1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

degrees of freedom

$$n - 1$$





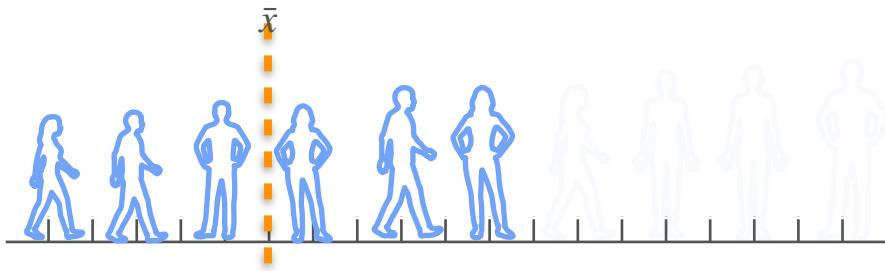
DeepLearning.AI

Confidence Interval

Confidence Intervals for Proportion

Confidence Interval for Proportions

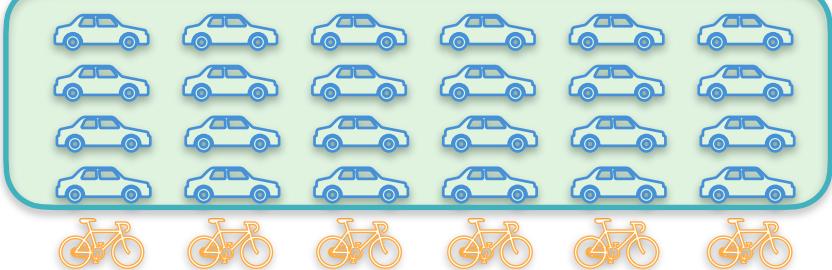
Confidence Interval for Means



$$\text{confidence interval} = \bar{x} \pm \text{margin of error}$$

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Confidence Interval for Proportions



$$x = 24$$

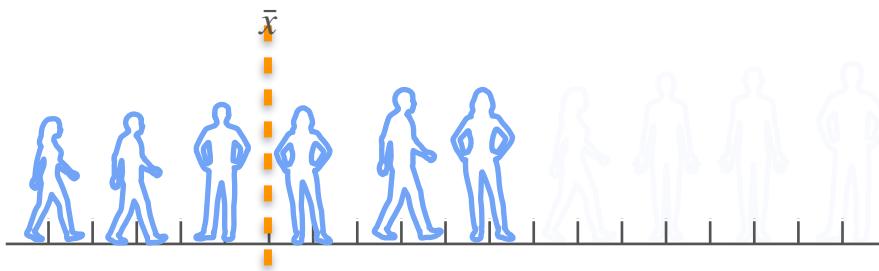
$$n = 30$$

$$\hat{p} = \frac{x}{n} = \frac{24}{30} = 80\%$$

How do you calculate a 95% confidence interval for this sample proportion?

Confidence Interval for Proportions

Confidence Interval for Means



$$\text{confidence interval} = \bar{x} \pm \text{margin of error}$$

$$\text{margin of error} = z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Confidence Interval for Proportions

$$\hat{p} = \frac{x}{n} = \frac{24}{30} = 80\%$$

$$\text{confidence interval} = \hat{p} \pm \text{margin of error}$$

$$\text{margin of error} = z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

standard error

Confidence Interval for Proportions

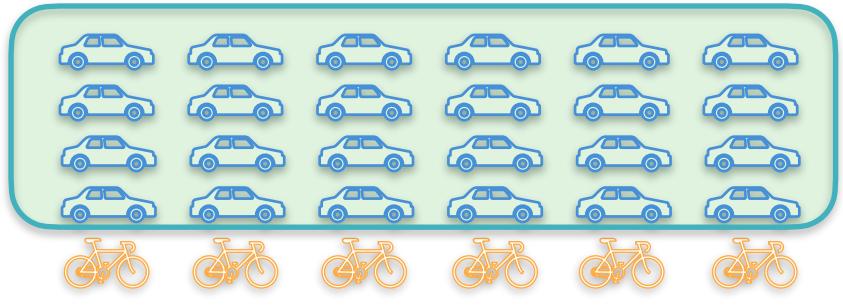
Confidence Interval for Proportions

confidence interval = $\hat{p} \pm$ margin of error

$$\text{margin of error} = z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

$$\text{margin of error} = 1.96 \cdot \sqrt{\frac{0.8(1 - 0.8)}{30}}$$

$$\text{margin of error} = 0.14$$



$$n = 30 \quad \hat{p} = 80\% = 0.8$$

Calculate a 95% confidence interval for this sample proportion

95% → $z_{1-\alpha/2} = 1.96$

Confidence Interval for Proportions

Confidence Interval for Proportions

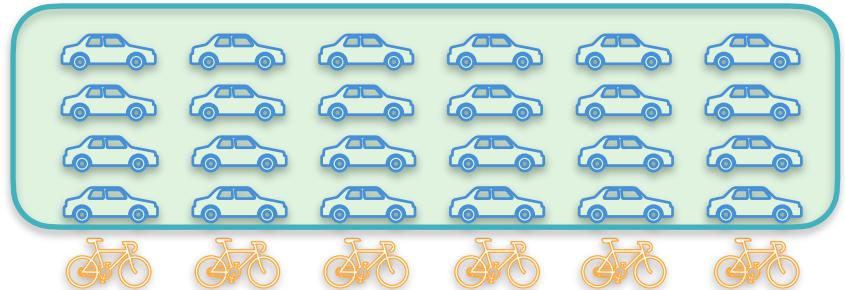
confidence interval = $\hat{p} \pm$ margin of error

margin of error = 0.14

confidence interval = 0.8 ± 0.14

$$0.66 < p < 0.94$$

$$66\% < p < 94\%$$



$$n = 30 \quad \hat{p} = 80\% = 0.8$$

Calculate a 95% confidence interval for this sample proportion

95%

$$\rightarrow z_{1-\alpha/2} = 1.96$$

W4 Lesson 2

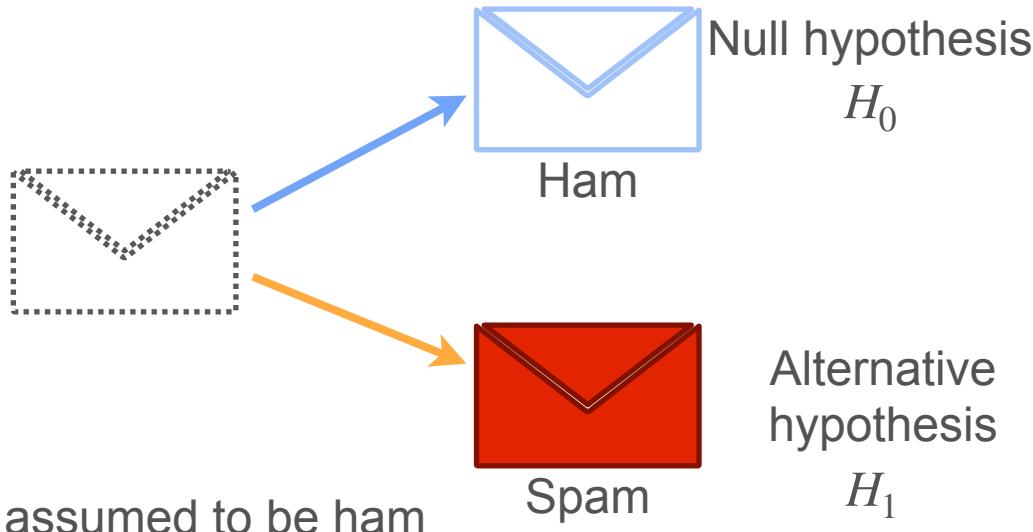


DeepLearning.AI

Hypothesis Testing

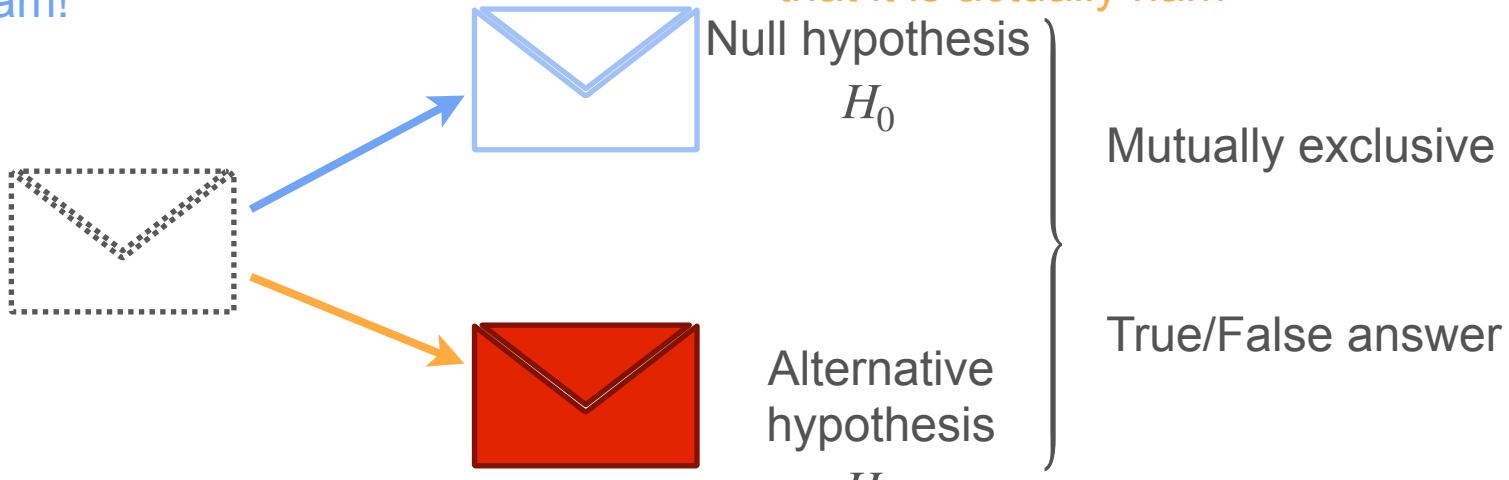
Defining Hypotheses

Motivation



Motivation

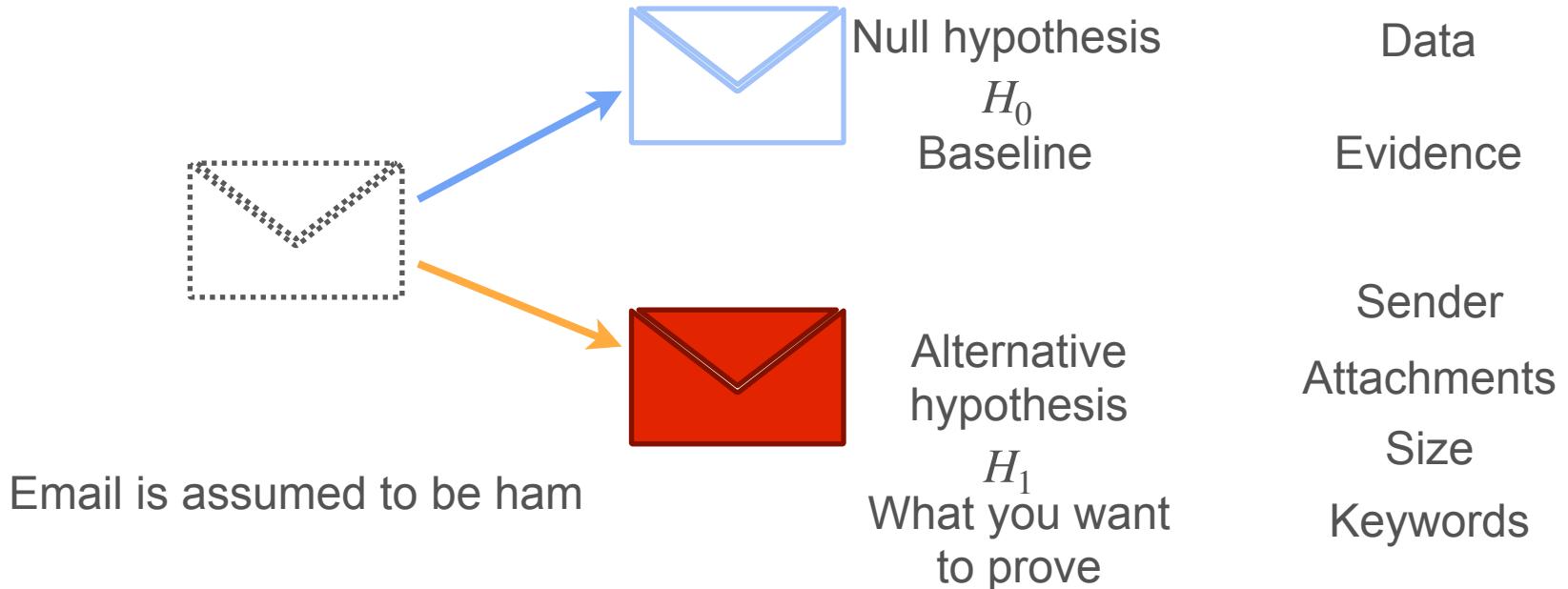
When rejecting that the email is not spam, you are accepting that the email is spam!



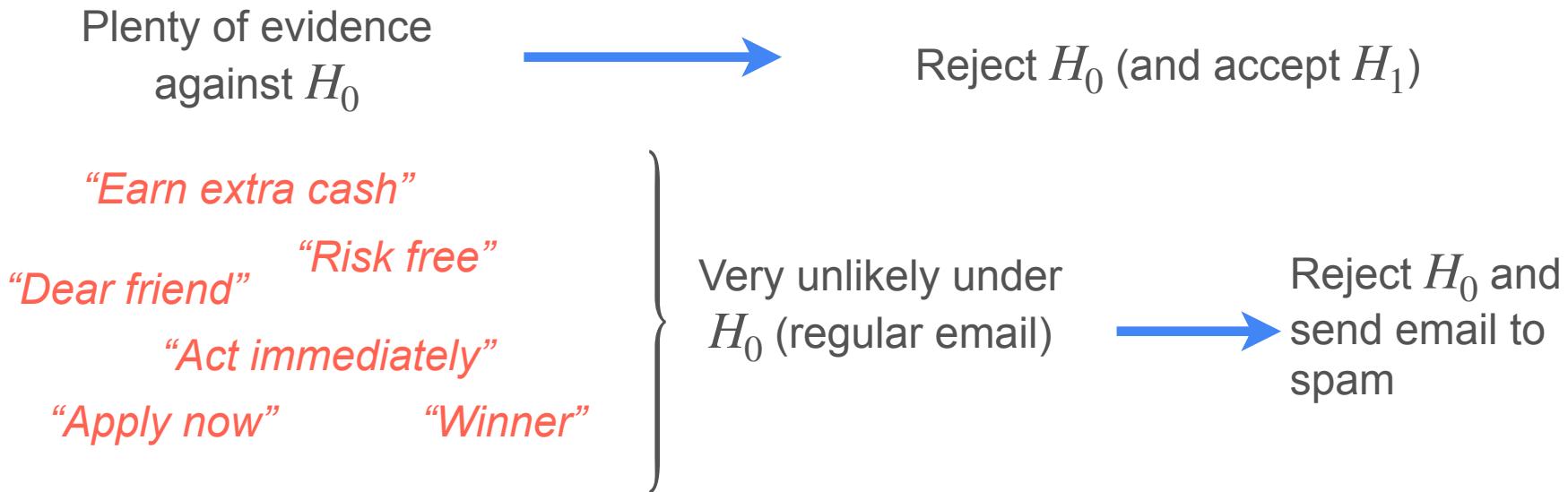
Email is assumed to be ham

Motivation

Not labeling the email spam, doesn't mean the email is ham!



How To Determine the Result of the Test





DeepLearning.AI

Hypothesis Testing

Type I and Type II errors

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

| 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

Sending a regular email to spam is worse than sending a spam email to the regular inbox.

Type I error



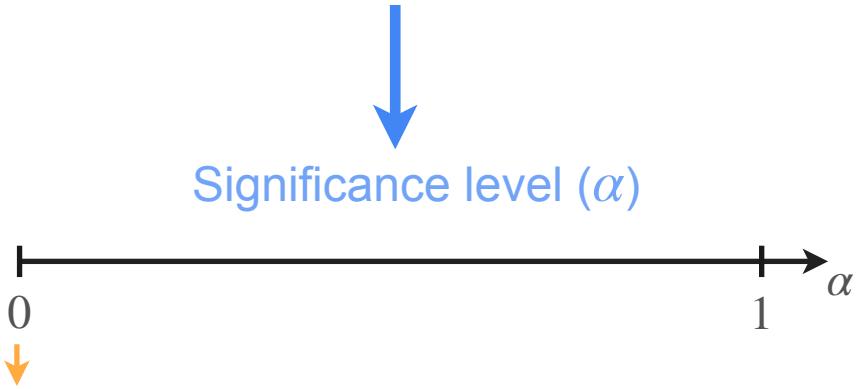
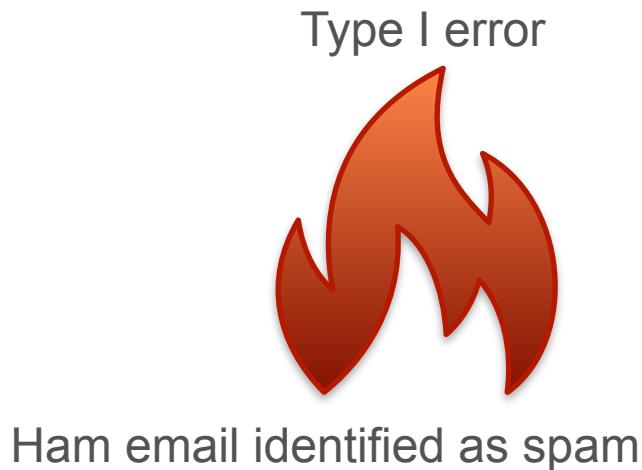
Type II error



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

Significance Level

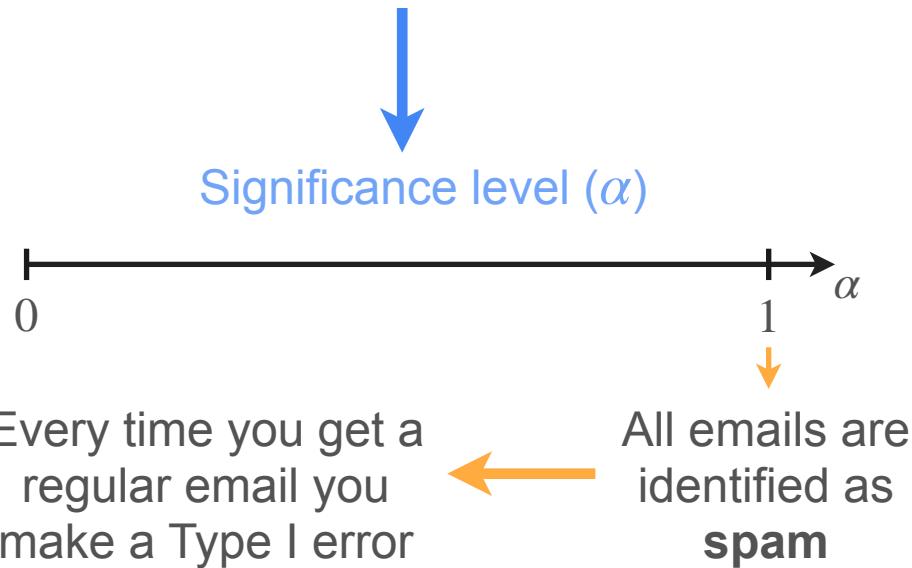
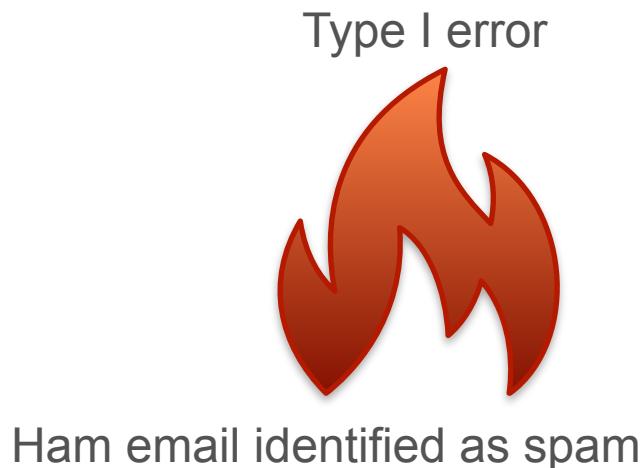
What is the greatest probability of type I error you are willing to tolerate?



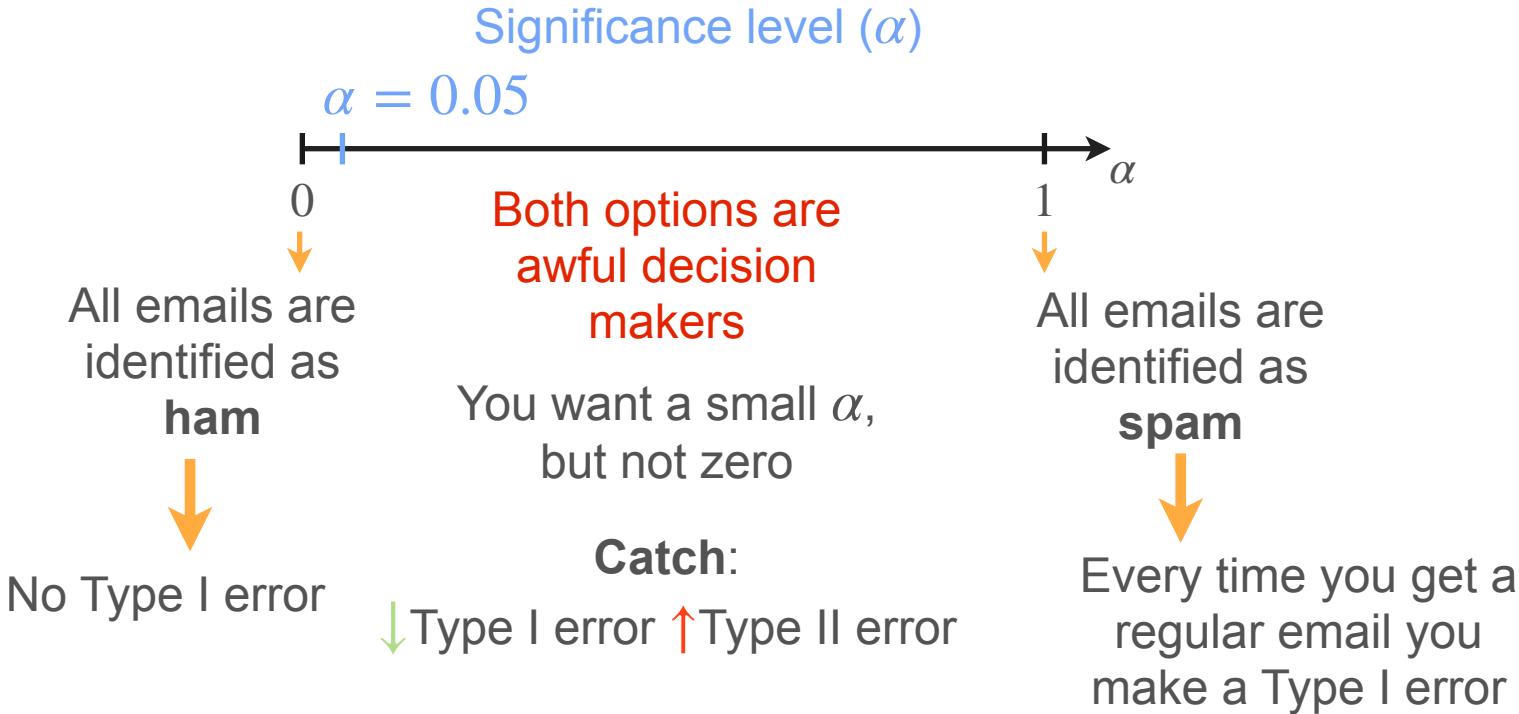
All emails are identified as **not spam** \longrightarrow No Type I error

Significance Level

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



Significance Level



Significance Level

Type I error



Ham email identified as spam

$$\begin{aligned}\alpha &= \max \mathbf{P}(\text{Type I error}) \\ &= \max \mathbf{P}(\text{Reject } H_0 | H_0)\end{aligned}$$

The value of α is your criteria for designing your test

Given sample, α will determine if you reject H_0 or not

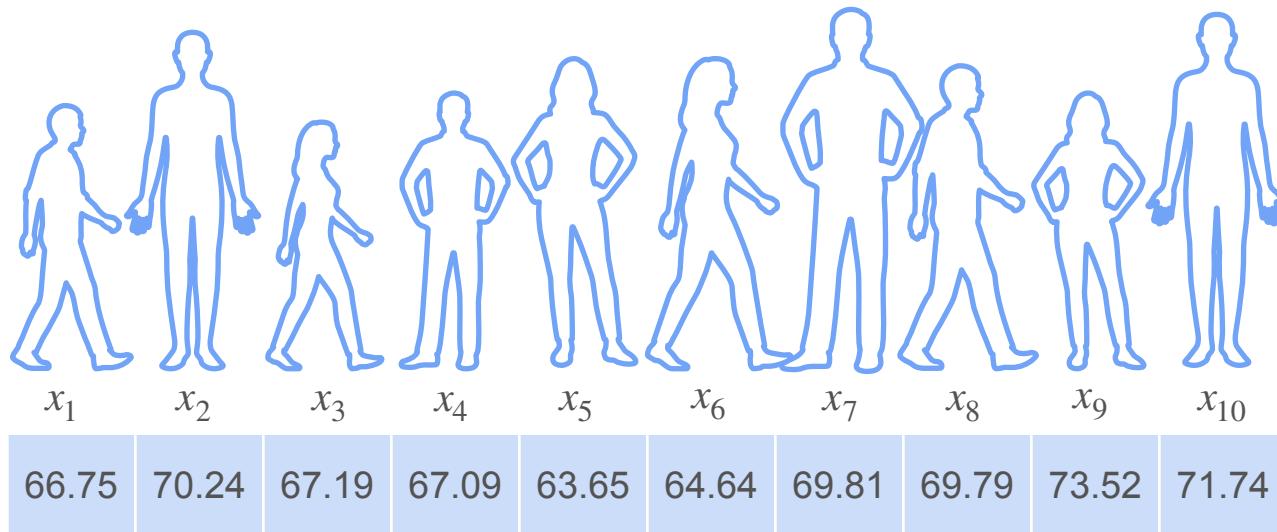


DeepLearning.AI

Hypothesis Testing

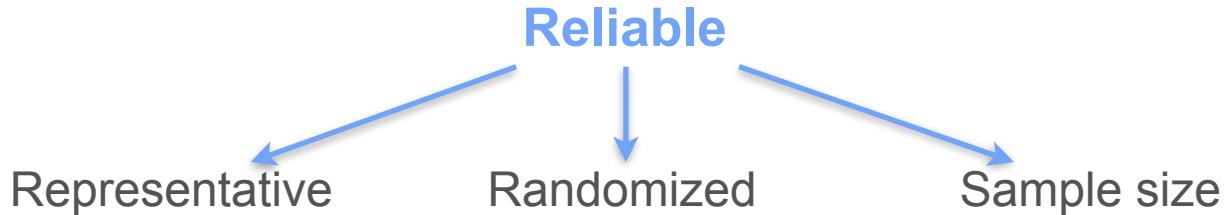
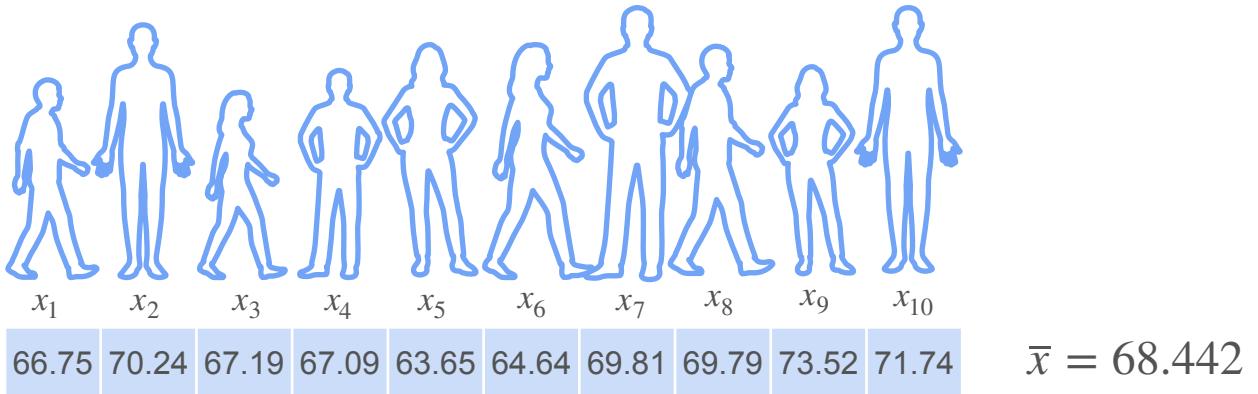
**Right-Tailed, Left-Tailed and
Two-Tailed tests**

Example: Heights

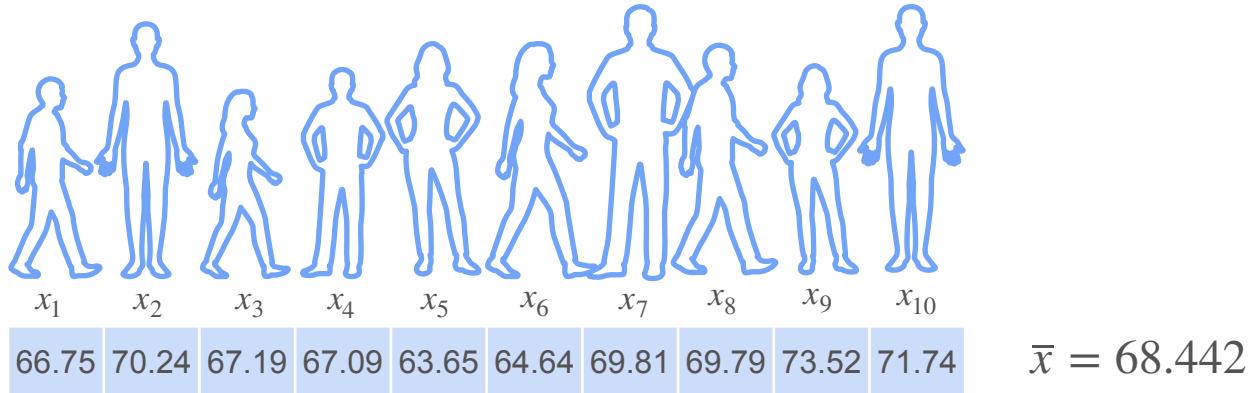


$$\bar{x} = 68.442$$

Data Quality



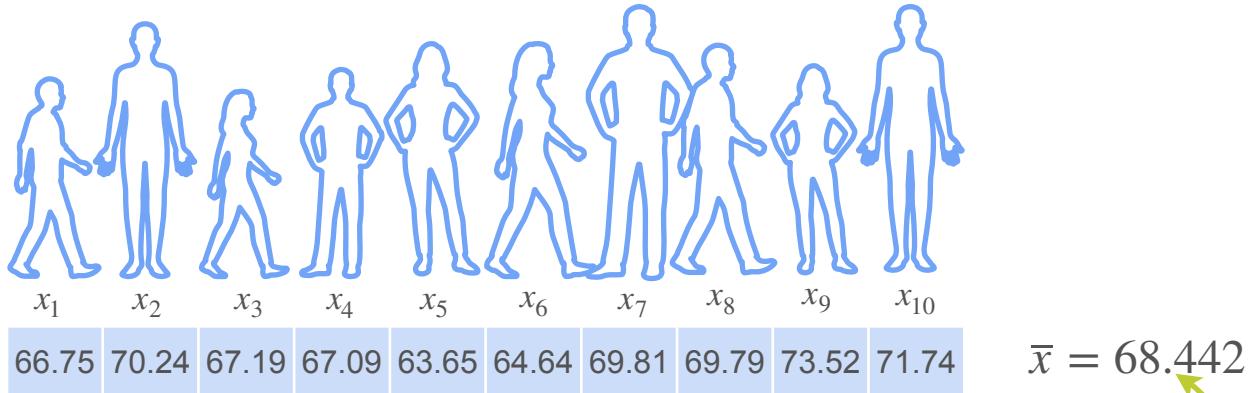
Determining the Hypothesis



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

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

Test Statistic



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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

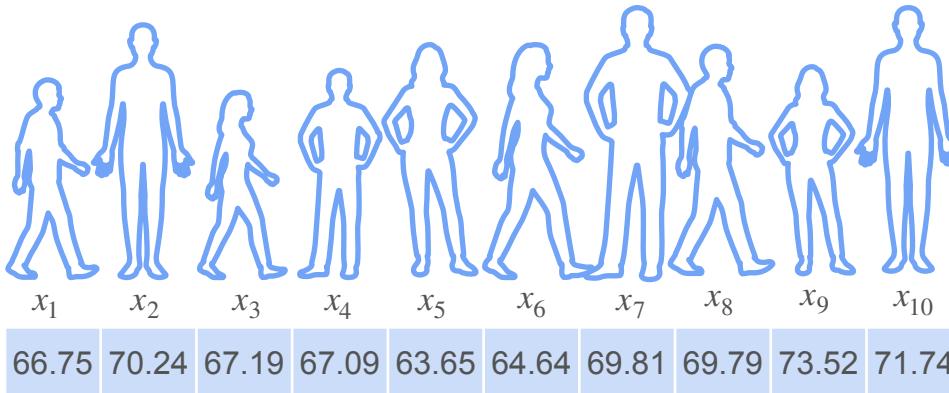
Observed statistic

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

Test Statistic

Test statistic

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



$$\bar{x} = 68.442$$

Test statistic: $T(X)$ $X = (X_1, \dots, X_n)$

Information about the population parameter under study

$$\mu \rightarrow \bar{X}$$

$$p \rightarrow \bar{X}$$

Not unique!

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

Example: Heights

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

3 questions



Right-Tailed Test

3 sets of hypothesis

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$



Left Tailed Test

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu < 66.7$$



Two-Tailed Test

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu \neq 66.7$$

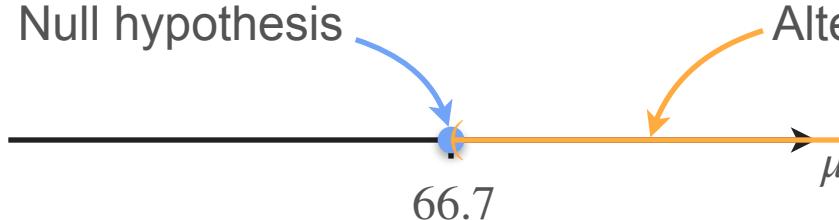
Example: Heights

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

\bar{X} Test statistic

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

Null hypothesis



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

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

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

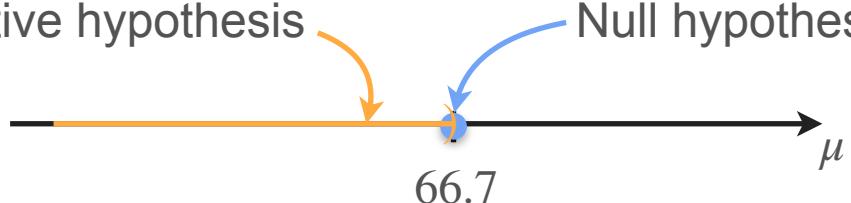
Example: Heights

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

\bar{X} Test statistic

Left tailed test $\longrightarrow H_0 : \mu \geq 66.7$ 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

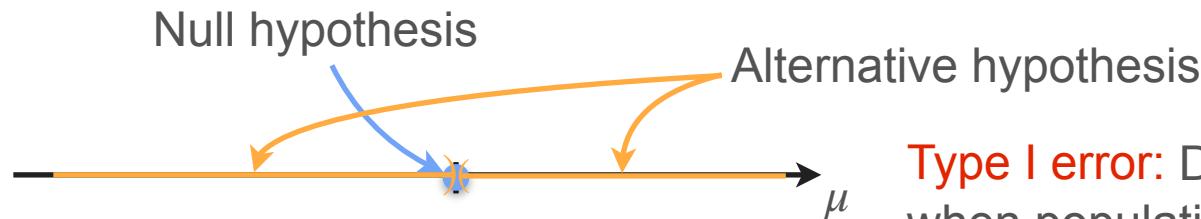
Type II error: Don't reject that $\mu = 66.7$ when true $\mu < 66.7$

Example: Heights

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

\bar{X} Test statistic

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



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

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

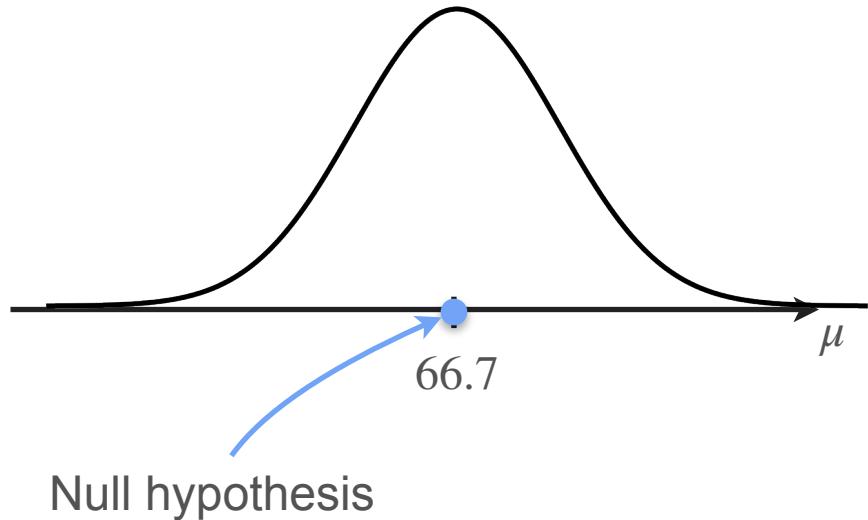
p -Value

Example: Heights

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

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

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



How likely was your sample if H_0 is true?

If the answer is very unlikely, then reject H_0

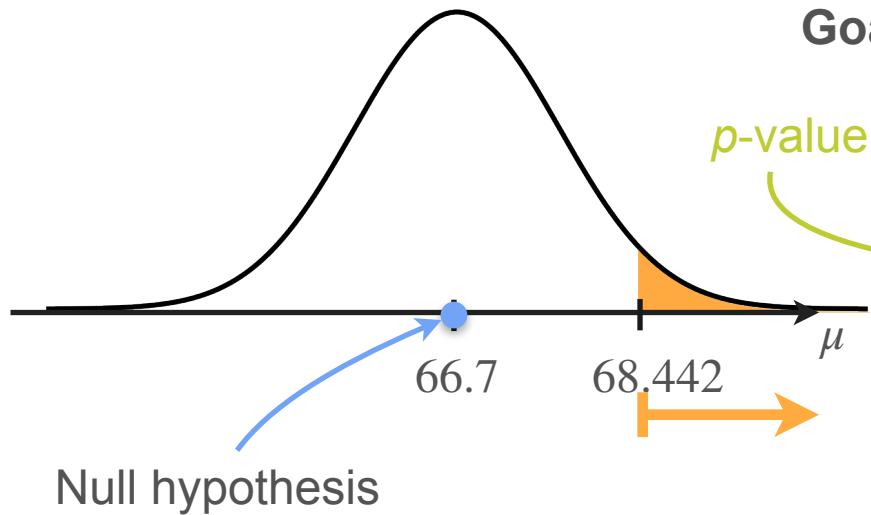
Right-Tailed Test for Gaussian Data (Known σ)

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

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

$$\bar{x} = 68.442$$

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$



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

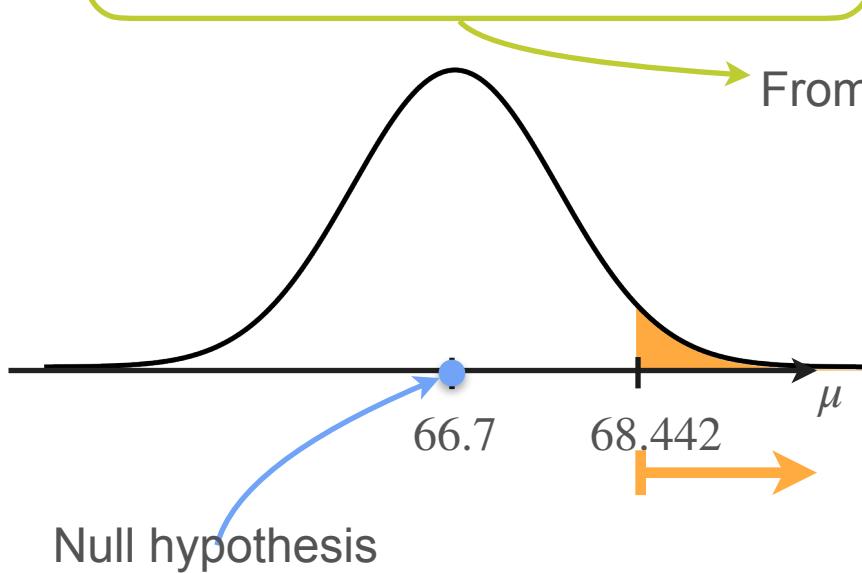
Type I error: Determine $\mu > 66.7$, when population mean did not change

$$\begin{aligned}P(\bar{X} > 68.442 \mid \mu = 66.7) \\ = 0.0332 < \alpha\end{aligned}$$

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

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

Decision rule:

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

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

p-values

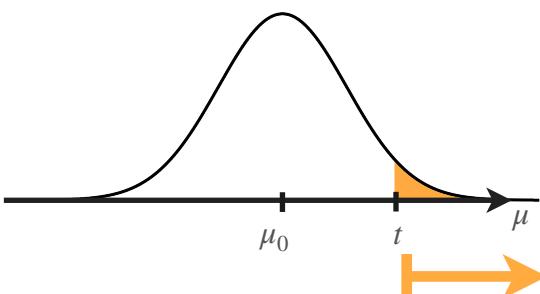
A ***p*-value** is the probability, assuming H_0 is true, that the test statistic takes on a value as extreme as or more extreme than the value observed

$T(X)$: test statistic

t : observed statistic

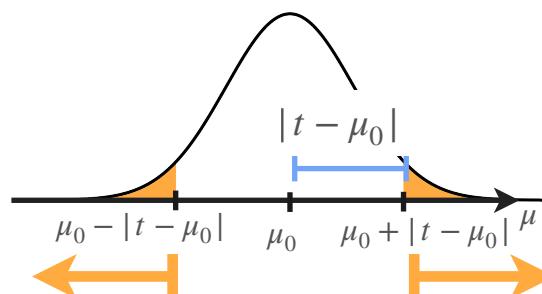
$H_0: \mu = \mu_0$

Right-tailed test



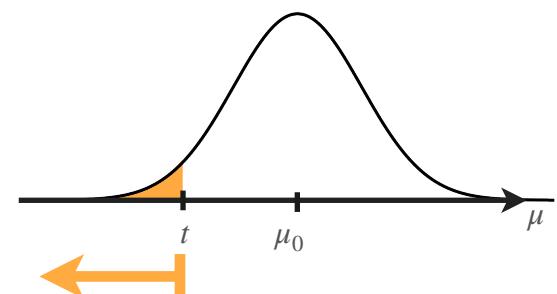
$$\mathbf{P}(T(X) > t | H_0)$$

Two-tailed test



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

Left-tailed test



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

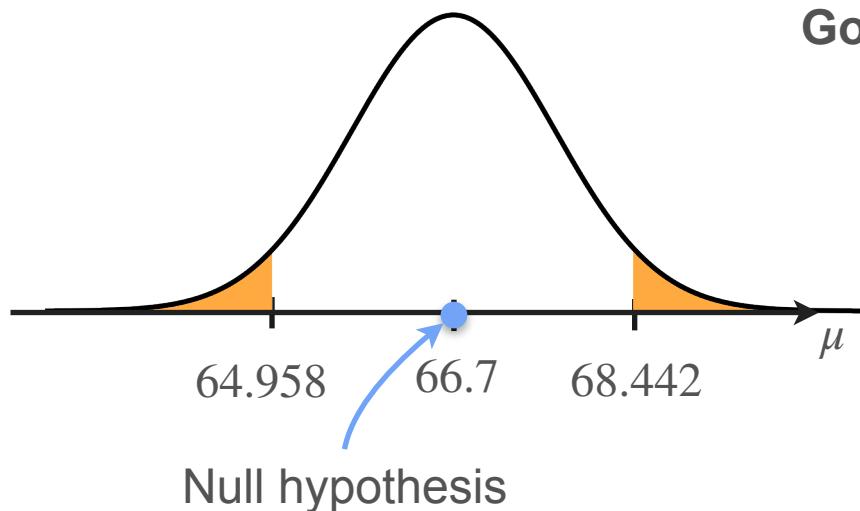
Two-Tailed Test for Gaussian Data (Known σ)

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

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

$$\bar{x} = 68.442$$

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu \neq 66.7$$



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

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

$$\begin{aligned}P\left(\left|\bar{X} - 66.7\right| > |68.442 - 66.7| \mid \mu = 66.7\right) \\ = 0.0663 > \alpha\end{aligned}$$

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

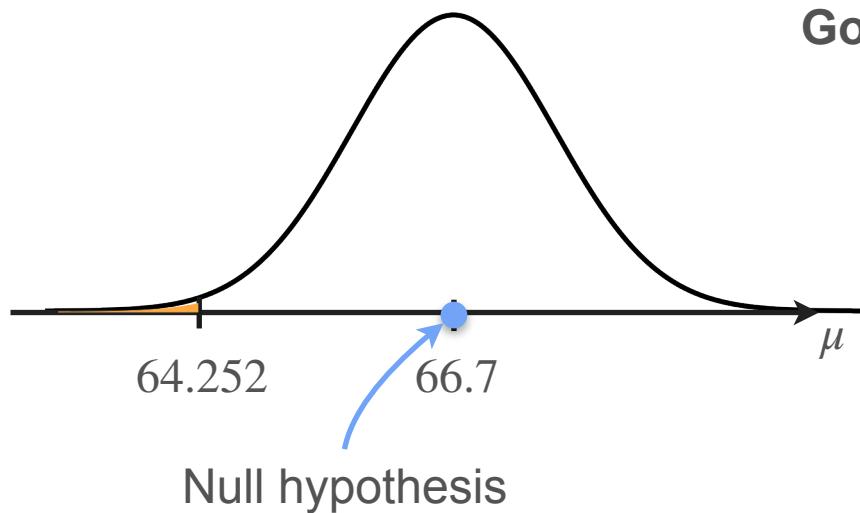
Left-Tailed Test for Gaussian Data (Known σ)

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

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

$$\bar{x} = 64.252$$

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu < 66.7$$



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

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

$$\begin{aligned}P(\bar{X} < 64.252 \mid \mu = 66.7) &: \\ &= 0.0049 < \alpha\end{aligned}$$

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

Tests Using the Z-Statistic

So far, you used the statistic \bar{X}

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

Applying standardization, you can write equivalent tests using the

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

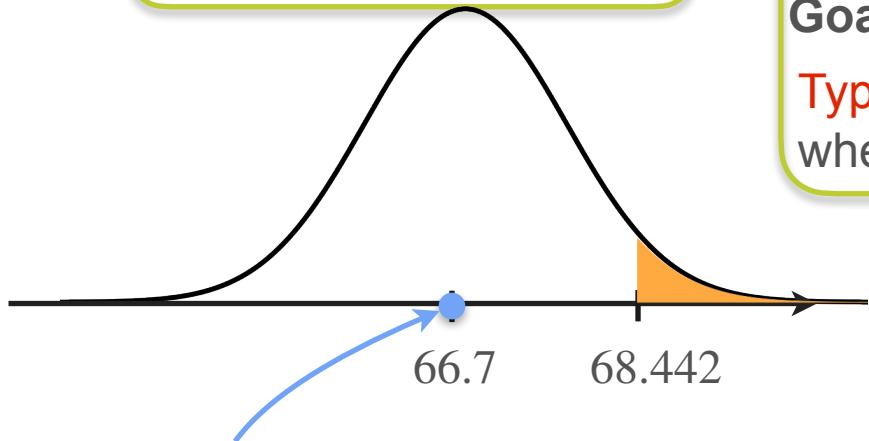
$$\text{If } H_0 \text{ is true, } Z = \frac{\bar{X} - \mu_0}{3/\sqrt{10}} \sim \mathcal{N}(0,1)$$

Right-Tailed Test Using the Z Statistic

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

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

$$\bar{x} = 68.442$$



$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

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

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

$$\begin{aligned}P(\bar{X} > 68.442 \mid \mu = 66.7) \\ = 0.0332 < \alpha\end{aligned}$$

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.

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

$$\bar{x} = 68.442$$

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

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

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

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

$$\begin{aligned}P(\bar{X} > 68.442 \mid \mu = 66.7) \\ = 0.0332 < \alpha\end{aligned}$$

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

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.

$$\begin{aligned}\sigma &= 3 \\ n &= 10\end{aligned}$$

$$\bar{x} = 68.442$$

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

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

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

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

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

$$P\left(\frac{\bar{X} - 66.7}{3/\sqrt{10}} > 1.837 \mid \mu = 66.7 \right) = 0.0332 < \alpha$$

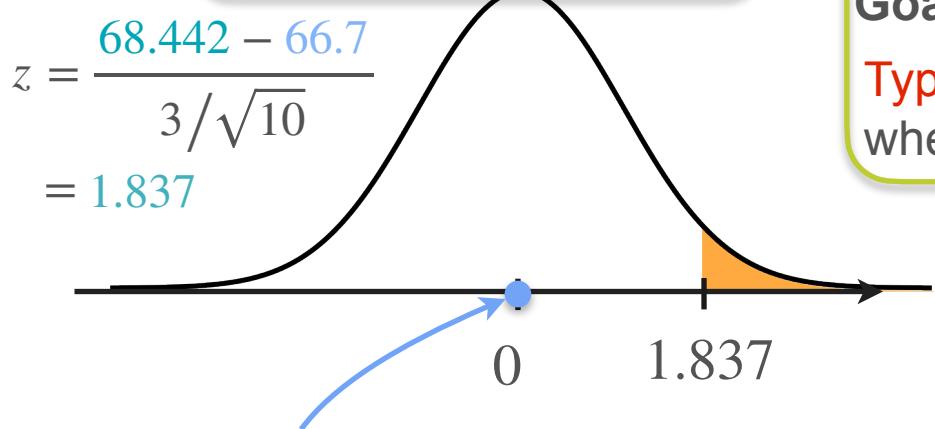
Conclusion: reject H_0
(with a 5% significance level)

Right-Tailed Test Using the Z Statistic

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

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

$$\bar{x} = 68.442$$



$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

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

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

$$P\left(\frac{\bar{X} - 66.7}{3/\sqrt{10}} > 1.837 \mid \mu = 66.7\right)$$

Conclusion: reject H_0 if $0.0332 < \alpha$
(with a 5% significance level)



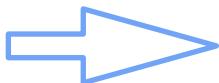
DeepLearning.AI

Hypothesis Testing

Critical Values

P-Values and Critical Values

If $p\text{-value} < \alpha$

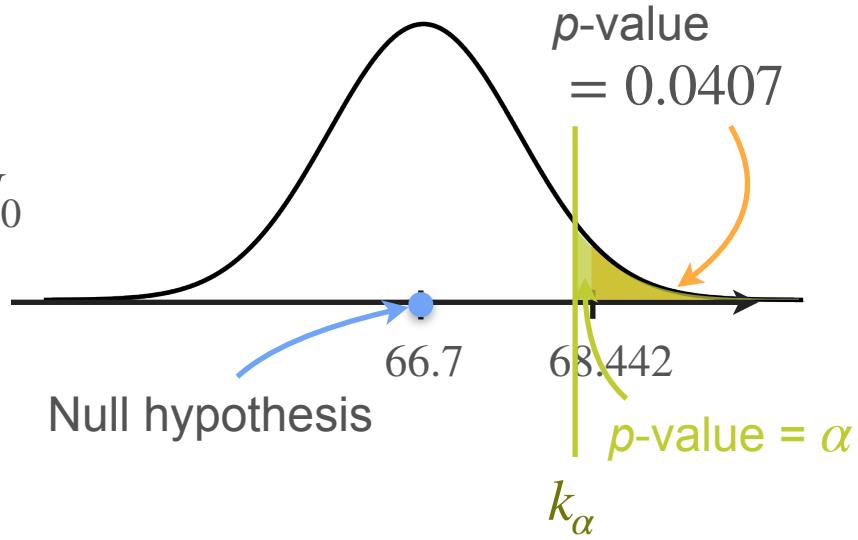


Reject H_0

What is the least extreme sample you could get that you would still reject H_0 ?

Sample that has $p\text{-value} = \alpha$

Critical values



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. $H_1 : \mu > 66.7$

$\alpha = 0.05$

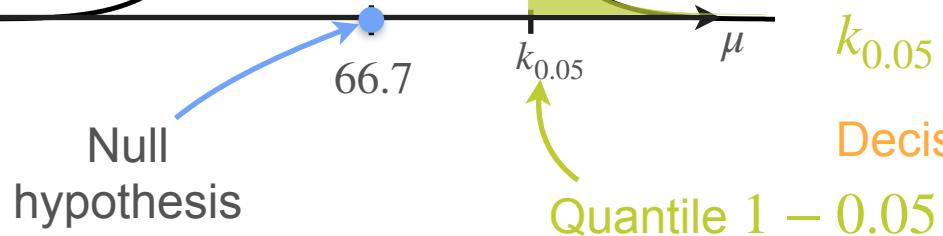
$n = 10 \quad \sigma = 3 \quad \bar{x} = 68.442$

$$0.05 = P(\bar{X} > k_{0.05} \mid \mu = 66.7)$$

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**

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

$$\alpha = 0.01$$

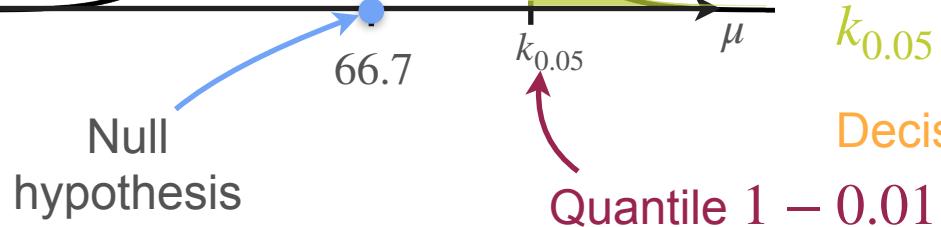
$$n = 10 \quad \sigma = 3 \quad \bar{x} = 68.442$$

$$0.05 = P(\bar{X} > k_{0.01} \mid \mu = 66.7)$$

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.**

Do not reject H_0

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

$\alpha = 0.01$

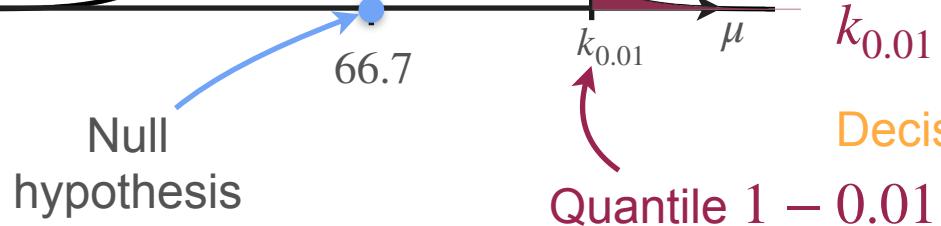
$$n = 10 \quad \sigma = 3 \quad \bar{x} = 68.442$$

$$0.01 = P(\bar{X} > k_{0.01} \mid \mu = 66.7)$$

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

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

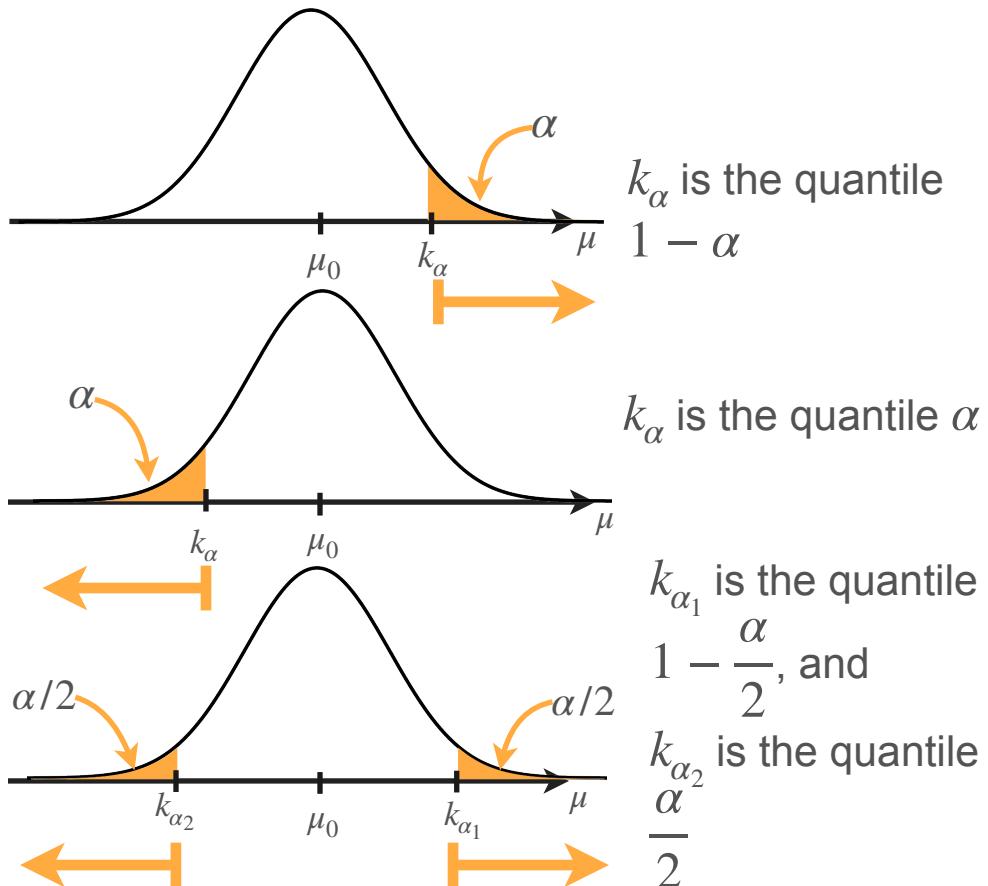
Decision rule: Reject H_0 if $t > k_\alpha$

$H_0 : \mu = \mu_0$ 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
 $t < k_{\alpha_2}$



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.



DeepLearning.AI

Hypothesis Testing

Power of a Test

Type I and Type II Errors

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

| Decision | Reality | |
|--|--------------------------------|---------------------------------|
| | H_0 True ($\mu = 66.7$) | H_0 False ($\mu > 66.7$) |
| Reject H_0 (Decide $\mu > 66.7$) | Type I error | Correct |
| Don't reject H_0 (Decide $\mu = 66.7$) | Correct | Type II error |

Finding the Type II Error Probabilities

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

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

For $\alpha = 0.05$: $k_\alpha = 68.26$

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

What is the **Type II error probability** if the true value is $\mu = 70$?

$$\mathbf{P}(\text{Do not reject } H_0 | \mu = 70) \longrightarrow \mathbf{P}(\bar{X} < 68.26 | \mu = 70)$$

Finding the Type II Error Probabilities

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

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

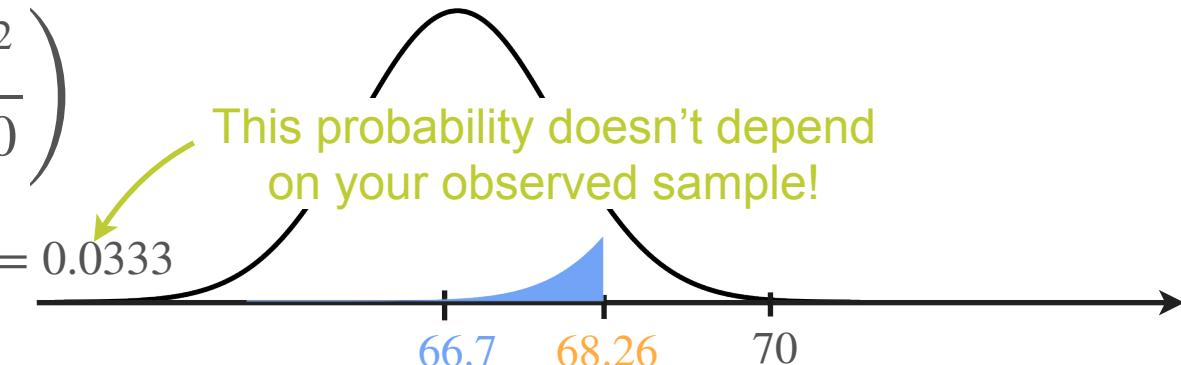
For $\alpha = 0.05$: $k_\alpha = 68.2604$ Decision rule: Reject H_0 if $\bar{X} > 68.26$

What is the **Type II error probability** if the true value is $\mu = 70$?

$$\text{If } \mu = 70 \quad \bar{X} \sim \mathcal{N}\left(70, \frac{3^2}{10}\right)$$

$$\beta = P(\bar{X} < 68.26 | \mu = 70) = 0.0333$$

This probability doesn't depend on your observed sample!



Power of the Test

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$

| Decision | Reality | | Power of the test $P(\text{Reject } H_0 \mu \in H_1)$ |
|--|--------------------------------|---------------------------------|--|
| | H_0 True ($\mu = 66.7$) | H_0 False ($\mu > 66.7$) | |
| Reject H_0 (Decide $\mu > 66.7$) | Type I error | Correct | |
| Don't reject H_0 (Decide $\mu = 66.7$) | Correct | Type II error | |

Power of the Test

$$\left. \begin{array}{l} \text{Type II error: } \mathbf{P} \left(\text{Do not reject } H_0 \mid \mu \in H_1 \right) \\ \text{Power of the test: } \mathbf{P} \left(\text{Reject } H_0 \mid \mu \in H_1 \right) \end{array} \right\} \begin{array}{l} \beta \\ 1 - \beta \end{array}$$

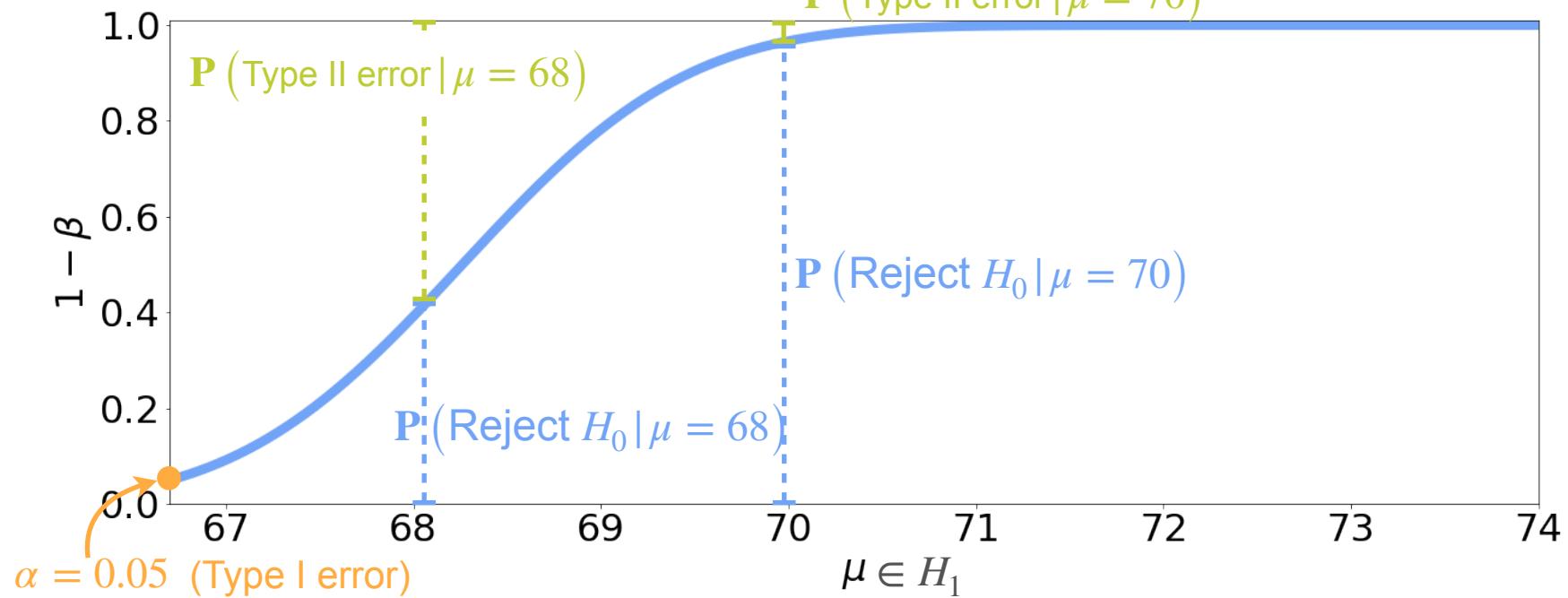
Complementary probabilities

Power of the test = 1 – Type II error probability

$$= 1 - \mathbf{P} \left(\text{Do not reject } H_0 \mid \mu \in H_1 \right)$$

Power of the Test

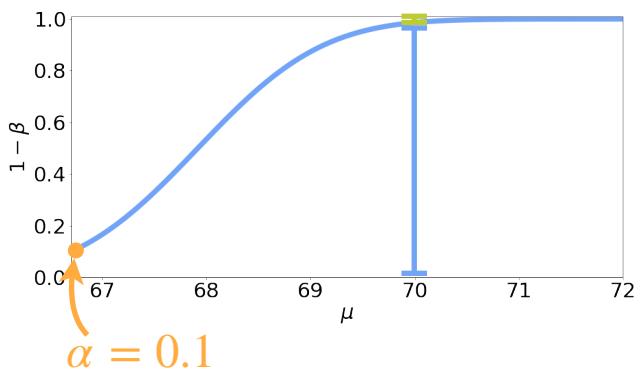
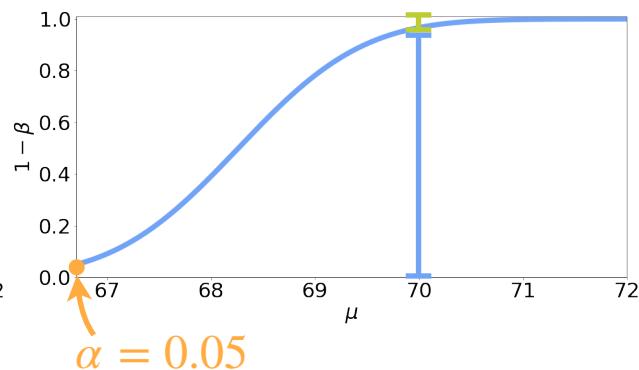
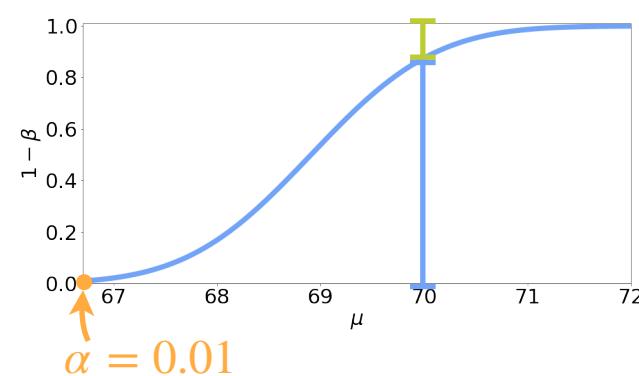
$H_0 : \mu = 66.7$ vs. $H_1 : \mu > 66.7$



Power of the Test

$$\mu = 70$$

Power

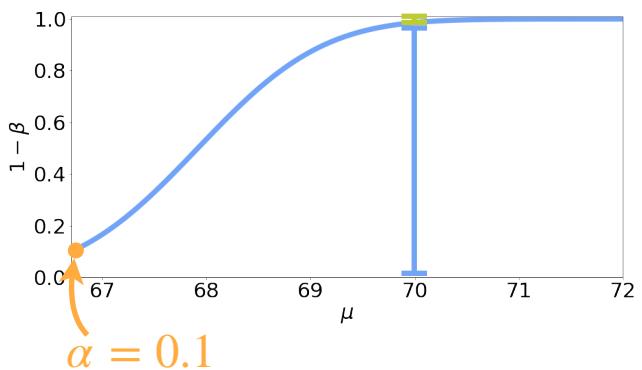
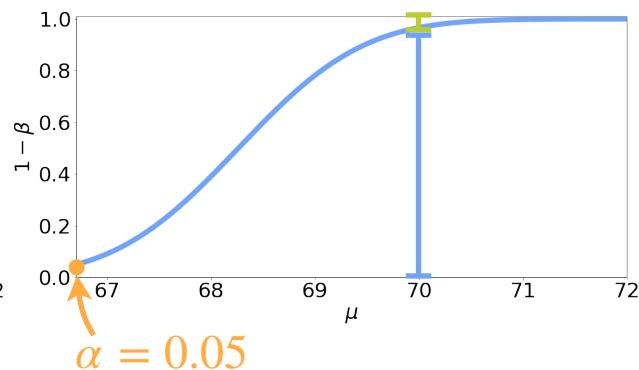
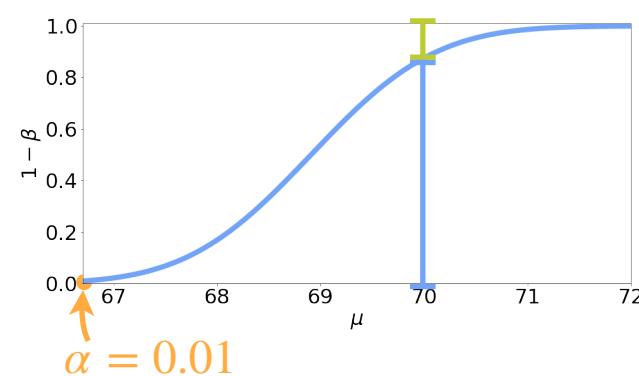


Type I error

Power of the Test

$$\mu = 70$$

Type II error



Type I error



DeepLearning.AI

Hypothesis Testing

Interpreting results

Steps for Performing Hypothesis Testing

1. State your hypotheses.

- Null hypothesis: the baseline $\rightarrow H_0 : \mu = 66.7$
- Alternative hypothesis: the statement you want to prove $\rightarrow H_1 : \mu > 66.7$

2. Design your test

- Decide the test statistic to work with. $\rightarrow \bar{X}$
- Decide the significance level $\rightarrow \alpha = 0.05$

3. Compute the observed statistic (based on your sample) $\rightarrow \bar{x} = 68.442$

4. Reach a conclusion:

- If the p -value is less than the significance level reject H_0
 $\rightarrow P(\bar{X} > 68.442 | \mu = 66.7) > ? 0.05$

Important Remarks - Interpreting Tests

- Type I error:
 - Reject null hypothesis when it is true \rightarrow Reject H_0 when $\mu = 66.7$
- Type II error:
 - Do not reject null hypothesis, when it was false \rightarrow Do not reject H_0 when $\mu > 66.7$
- Significance level (α):
 - It is the maximum probability of incurring in a Type I error
- Errors:
 - $\downarrow \alpha \uparrow \beta$

Important Remarks - Interpreting Tests

- *p*-values:

- If $\mathbf{P}(\text{Reject } H_0 | H_0) < \alpha \longrightarrow \text{Reject } H_0 \text{ and then accept } H_1$
- The *p*-value represents the probability of H_0 being true.

Important Remarks - Interpreting Tests

- *p*-values:

- If $\mathbf{P}(\text{Reject } H_0 | H_0) < \alpha \longrightarrow \text{Reject } H_0 \text{ and then accept } H_1$
- A small *p*-value indicates that the probability of seeing the observed data by chance is small

Important Remarks - Interpreting Tests

- *p*-values:

- If $\mathbf{P}(\text{Reject } H_0 | H_0) < \alpha \longrightarrow \text{Reject } H_0 \text{ and then accept } H_1$
- A small *p*-value indicates that the probability of seeing the observed data by chance is small

- Test conclusions

- Reject $H_0 \rightarrow H_1$ true
- Do not reject $H_0 \rightarrow H_0$ true



You can only say that there is not enough evidence

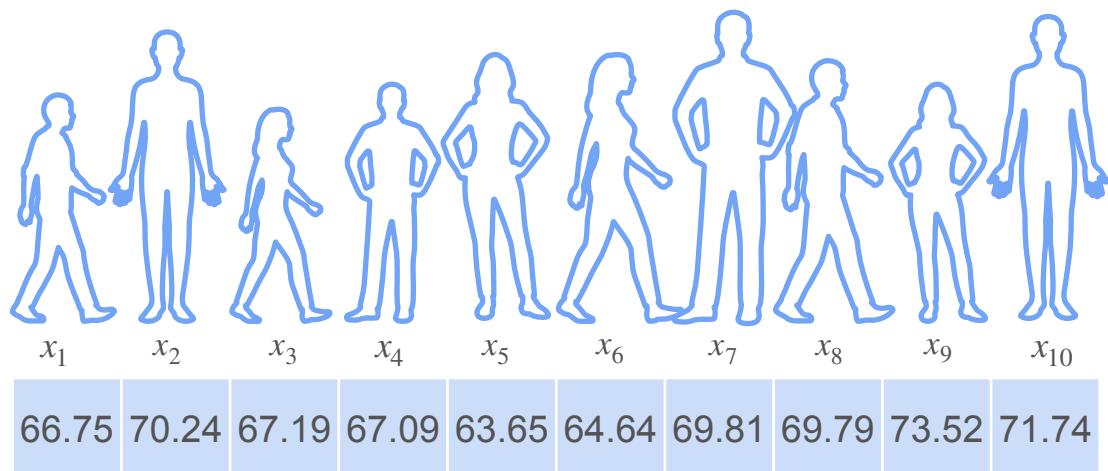


DeepLearning.AI

Hypothesis Testing

t-Distribution

t -Distribution: Motivation



$$X_i \stackrel{i.i.d}{\sim} \mathcal{N}(\mu, \sigma^2)$$

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{10}\right)$$

This is fine if you know μ and σ

What if σ is unknown?

t -Distribution: Motivation

$$X_i \stackrel{i.i.d}{\sim} \mathcal{N}(\mu, \sigma^2)$$

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{10}\right)$$

$$S = \sqrt{\frac{1}{10-1} \sum_{i=1}^{10} (X_i - \bar{X})^2}$$

If μ, σ are known

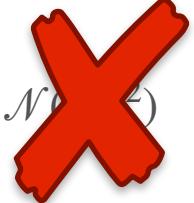
$$\frac{\bar{X} - \mu}{\sigma / \sqrt{10}} \sim \mathcal{N}(0, 1^2) \text{ (Standardization)}$$

Z statistic

What if σ is unknown?

Replace σ with its estimate

$$\frac{\bar{X} - \mu}{S / \sqrt{10}} \sim \mathcal{N}(0, 1^2)$$



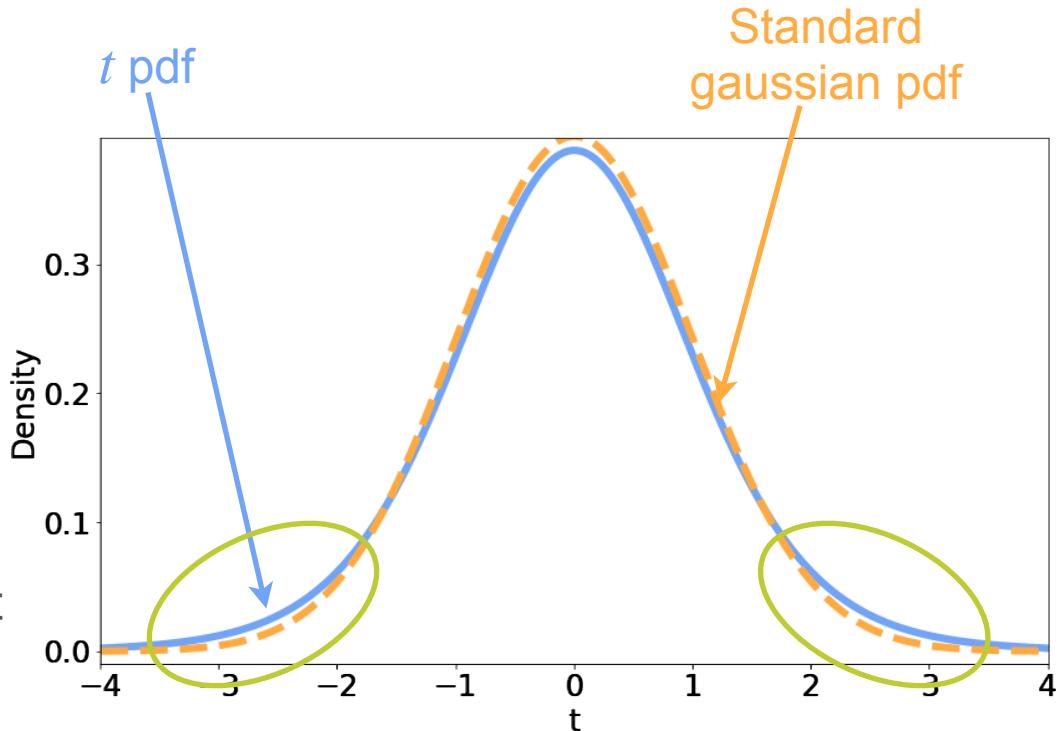
t -Distribution

$\frac{\bar{X} - \mu}{S/\sqrt{10}}$ follows a t distribution

What does it look like?

Still bell-shaped

It has heavier tails that account for the uncertainty introduced with the std estimation



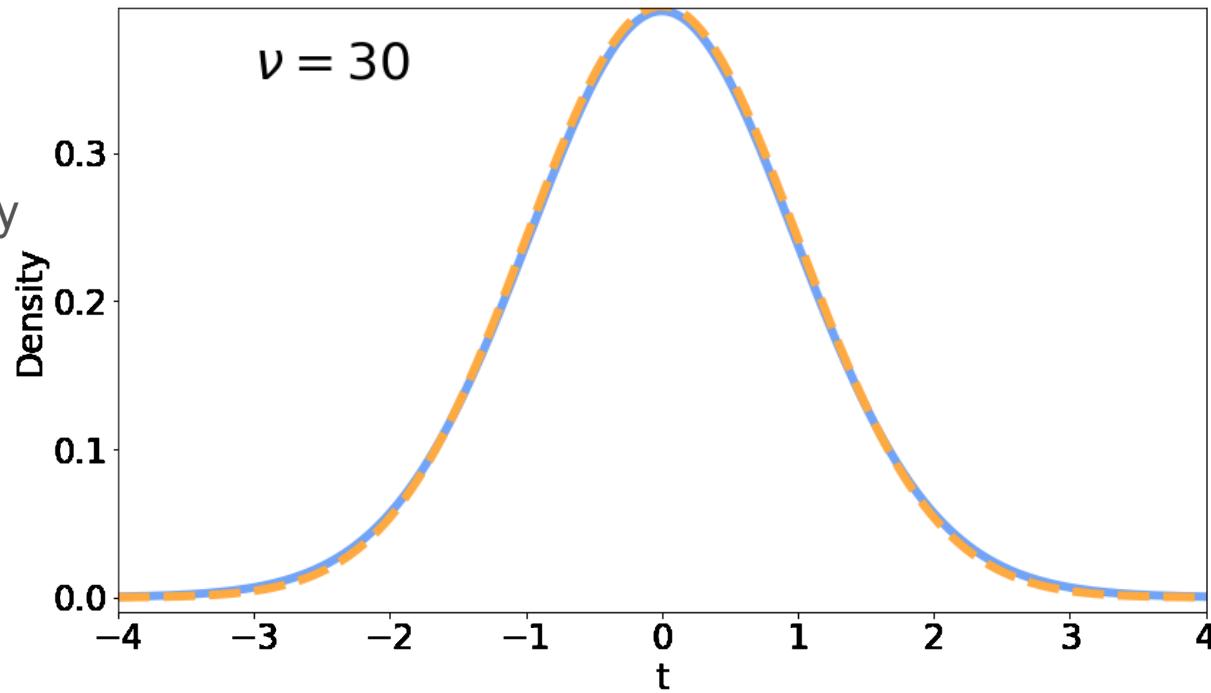
t -Distribution

Parameters:

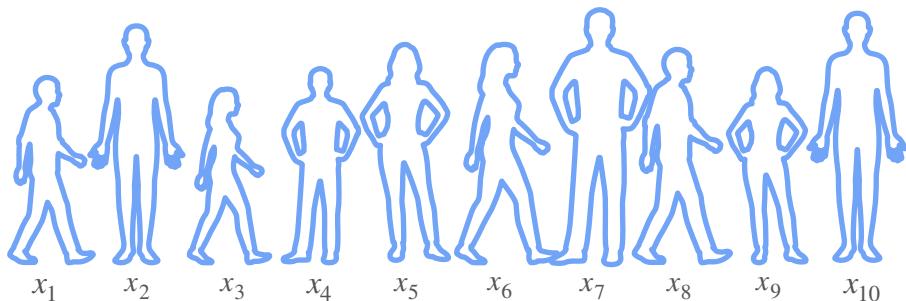
- Degrees of freedom (ν)

Controls how heavy
the tails are

$$X \sim t_{\nu}$$



t -Distribution and T -Statistic



$$n = 10$$
$$T = \frac{\bar{X} - \mu}{S/\sqrt{10}} \sim t_9$$

$\nu = 10 - 1$

Degrees of freedom (ν) = sample size - 1
= $(n - 1)$

As n increases, this looks more like a $\mathcal{N}(0, 1^2)$

T -statistic is used when

- The population has a Gaussian distribution
- But you don't know the variance

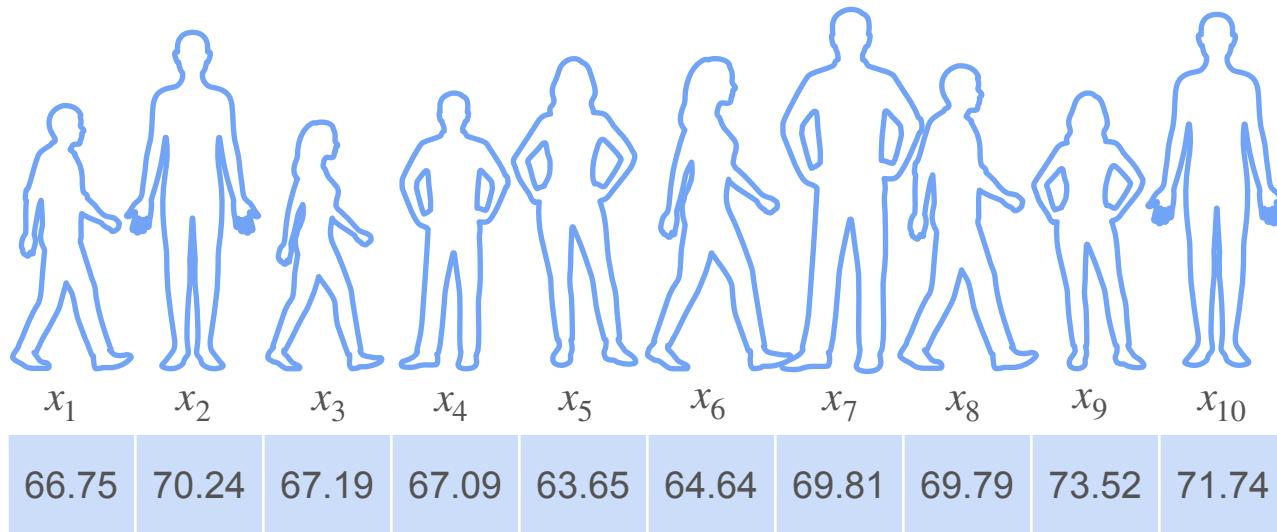


DeepLearning.AI

Hypothesis Testing

t-Tests

Example: Heights

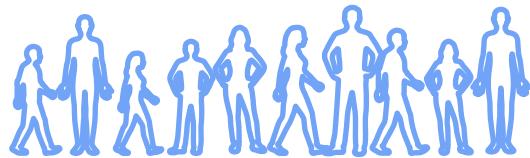


$$\bar{x} = 68.442$$

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7$$



$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu < 66.7$$



$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu \neq 66.7$$

Example: Heights

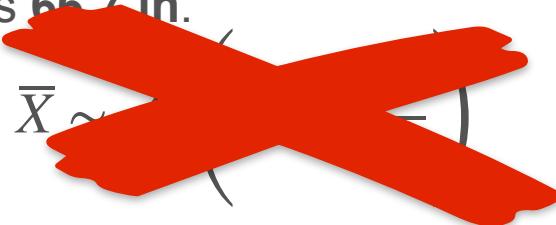
The **mean** height for 18 y/o in the US in the 70s was 66.7 in.



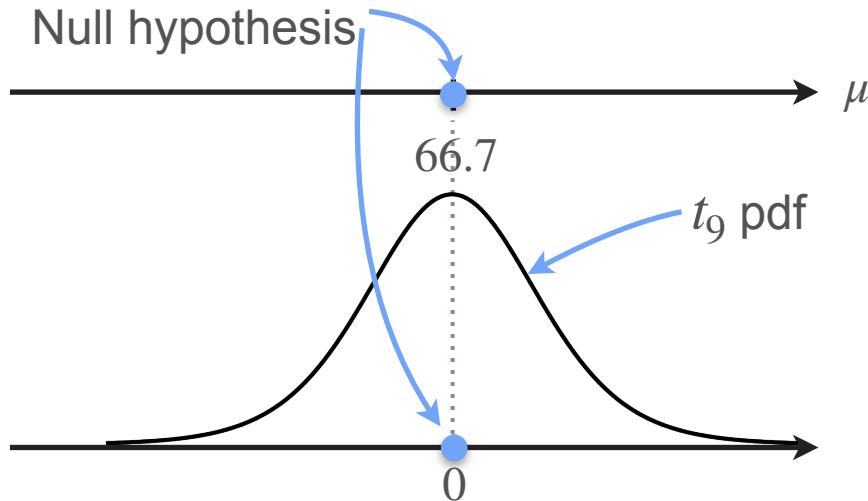
~~$n = 10$~~

$$H_0 : \mu = 66.7$$

If H_0 is true: $\bar{X} \sim$



Null hypothesis



$$\text{If } H_0 \text{ is true: } T = \frac{\bar{X} - 66.7}{S/\sqrt{10}} \sim t_9$$

Right-Tailed Test for Gaussian Data (Unknown σ)

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu > 66.7 \quad n = 10$$

$$\bar{x} = 68.442$$

$$s = 3.113$$

$$t = \frac{68.442 - 66.7}{3.113/\sqrt{10}} = 1.770$$

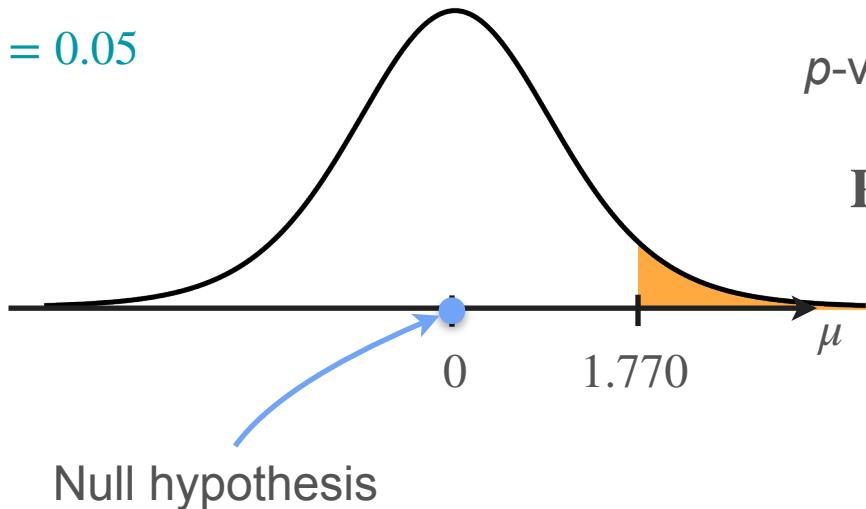
$$\alpha = 0.05$$

p-value:

$$P\left(\frac{\bar{X} - 66.7}{S/\sqrt{10}} > 1.770 \mid \mu = 66.7\right)$$

$$= 0.0552 > \alpha$$

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



Two-Tailed Test for Gaussian Data (Unknown σ)

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu \neq 66.7$$

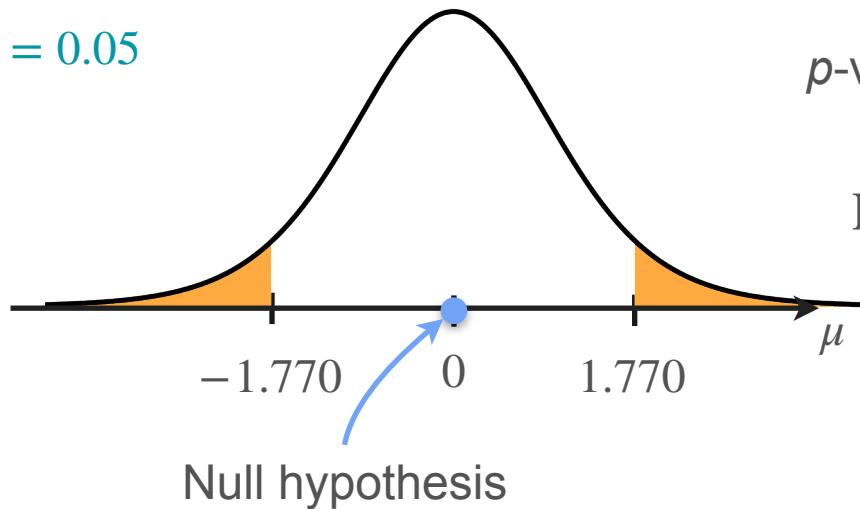
$$n = 10$$

$$\bar{x} = 68.442$$

$$s = 3.113$$

$$t = \frac{68.442 - 66.7}{3.113/\sqrt{10}} = 1.770$$

$$\alpha = 0.05$$



p-value:

$$P\left(\left|\frac{\bar{X} - 66.7}{S/\sqrt{10}}\right| > |1.770| \mid \mu = 66.7\right) = 0.1105 > \alpha$$

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

Left-Tailed Test for Gaussian Data (Unknown σ)

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

$$H_0 : \mu = 66.7 \text{ vs. } H_1 : \mu < 66.7$$

$$n = 10$$

$$\bar{x} = 64.252$$

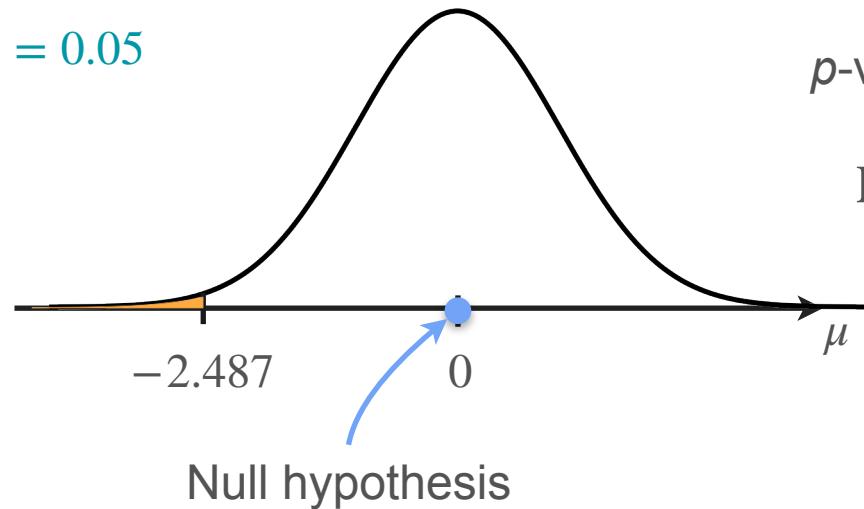
$$s = 3.113$$

$$t = \frac{64.252 - 66.7}{3.113/\sqrt{10}} = -2.487$$

$$\alpha = 0.05$$

p-value:

$$P\left(\frac{\bar{X} - 66.7}{S/\sqrt{10}} < -2.487 \mid \mu = 66.7\right)$$



$$= 0.0173 < \alpha$$

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

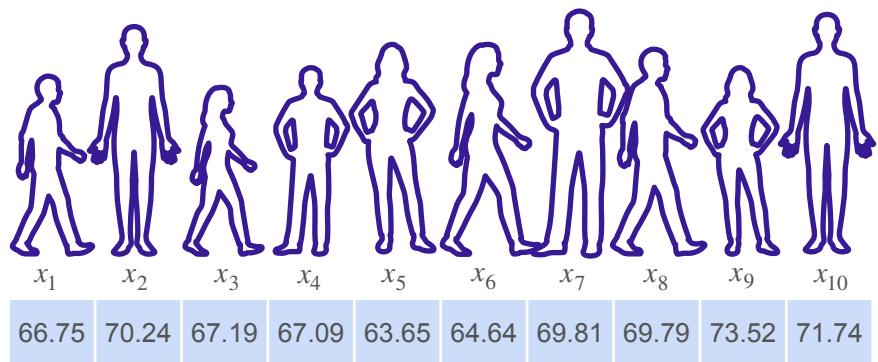


DeepLearning.AI

Hypothesis Testing

Two sample t-test

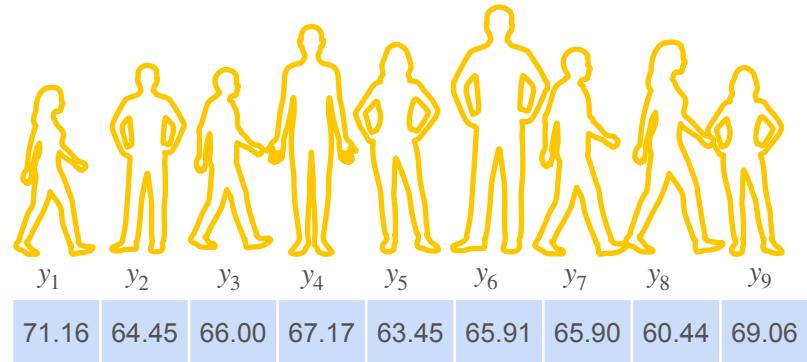
Independent Two-Sample t -Test



$$n_X = 10 \quad \bar{x} = 68.442$$
$$s_X = 3.113$$

Height of 18 y/o in the US

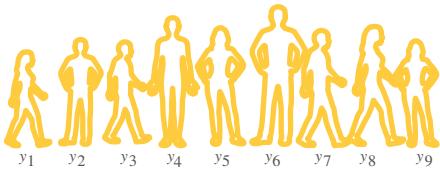
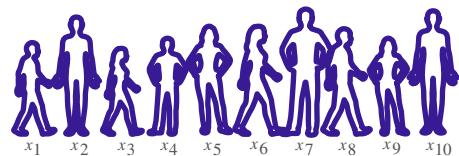
$$\mu_{US} \neq \mu_{Ar}$$



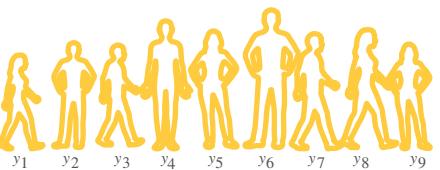
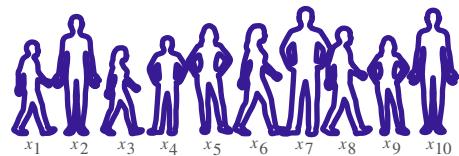
$$n_Y = 9 \quad \bar{y} = 65.949$$
$$s_Y = 3.106$$

Height of 18 y/o in Argentina

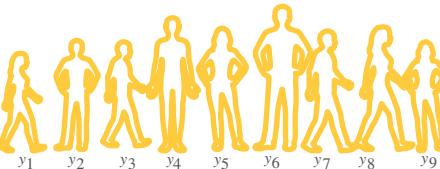
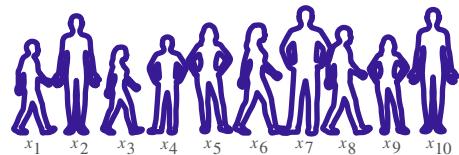
Independent Two-Sample t -Test: Hypothesis



$$H_0 : \mu_{US} - \mu_{Ar} = 0 \text{ vs. } H_1 : \mu_{US} - \mu_{Ar} > 0$$



$$H_0 : \mu_{US} - \mu_{Ar} = 0 \text{ vs. } H_1 : \mu_{US} - \mu_{Ar} < 0$$



$$H_0 : \mu_{US} - \mu_{Ar} = 0 \text{ vs. } H_1 : \mu_{US} - \mu_{Ar} \neq 0$$

Independent Two-Sample t -Test: Assumptions

- All people in the sample from the two groups are different
- Each person in both samples are independent
- Populations are normally distributed

$$X \sim \mathcal{N}(\mu_{US}, \sigma_{US}^2)$$

$$Y \sim \mathcal{N}(\mu_{Arg}, \sigma_{Arg}^2)$$

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

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\text{---} , \text{---} + \text{---} \right)$$

$$\bar{Y} = \frac{1}{9} \sum_{i=1}^{9} Y_i$$

Independent Two-Sample t -Test: Assumptions

- All people in the sample from the two groups are different
- Each person in both samples are independent
- Populations are normally distributed

$$X \sim \mathcal{N}(\mu_{US}, \sigma_{US})$$

$$Y \sim \mathcal{N}(\mu_{Arg}, \sigma_{Arg})$$

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

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_{US} - \mu_{Arg}, \sqrt{\frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}}\right)$$

$$\bar{Y} = \frac{1}{9} \sum_{i=1}^9 Y_i$$

Independent Two-Sample t -Test: Statistic

$$X \sim \mathcal{N}(\mu_{US}, \sigma_{US}) \quad Y \sim \mathcal{N}(\mu_{Arg}, \sigma_{Arg})$$

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_{US} - \mu_{Arg}, \frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}\right) \rightarrow \frac{(\bar{X} - \bar{Y}) - (\mu_{US} - \mu_{Arg})}{\sqrt{\frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}}} \sim \mathcal{N}(0, 1^2)$$

You don't know σ_{US} , σ_{Arg}

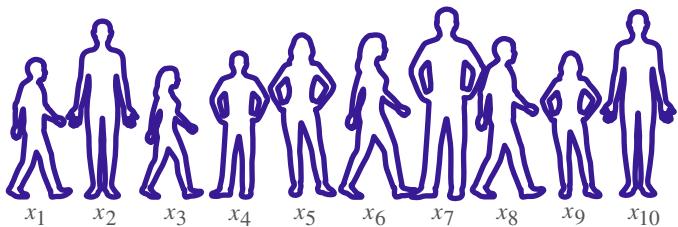


Replace it with the sample standard deviation

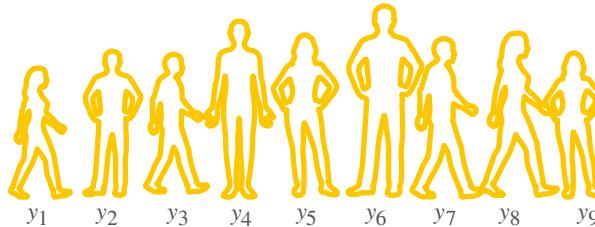
$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_{US} - \mu_{Arg})}{\sqrt{\frac{s_X^2}{10} + \frac{s_Y^2}{9}}} \sim t_{\nu}$$

Degrees of freedom = $\frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{s_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{s_Y^2}{n_Y}\right)^2}{n_Y - 1}}$

Independent Two-Sample t -Test



$$n_X = 10 \quad \bar{x} = 68.442$$
$$s_X = 3.113$$



$$n_Y = 9 \quad \bar{y} = 65.949$$
$$s_Y = 3.106$$

Degrees of freedom = $\frac{\overbrace{\left(\frac{16.8}{\left(\frac{-1}{2} + \frac{-1}{2} \right)^2} \right)^2}}{\left(\frac{-1}{2} \right)^2 + \left(\frac{-1}{2} \right)^2}$

Independent Two-Sample t -Test: Statistic

$$X \sim \mathcal{N}(\mu_{US}, \sigma_{US})$$

$$Y \sim \mathcal{N}(\mu_{Arg}, \sigma_{Arg})$$

$$\bar{X} - \bar{Y} \sim \mathcal{N}\left(\mu_{US} - \mu_{Arg}, \sqrt{\frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}}\right) \rightarrow \frac{(\bar{X} - \bar{Y}) - (\mu_{US} - \mu_{Arg})}{\sqrt{\frac{\sigma_{US}^2}{10} + \frac{\sigma_{Arg}^2}{9}}} \sim \mathcal{N}(0,1)$$

You don't know σ_{US} , σ_{Arg}



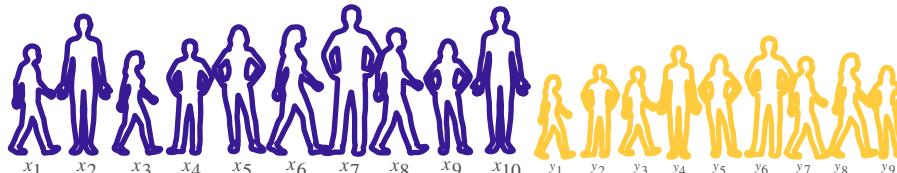
Replace it with the sample standard deviation

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_{US} - \mu_{Arg})}{\sqrt{\frac{s_X^2}{10} + \frac{s_Y^2}{9}}} \sim t_{16.8}$$

Degrees of freedom

$$\text{Degrees of freedom} = \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{s_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{s_Y^2}{n_Y}\right)^2}{n_Y - 1}}$$

Independent Two-Sample t -Test: Right Tailed Test



$$\bar{x} = 68.442$$

$$s_X = 3.113$$

$$n_X = 10$$

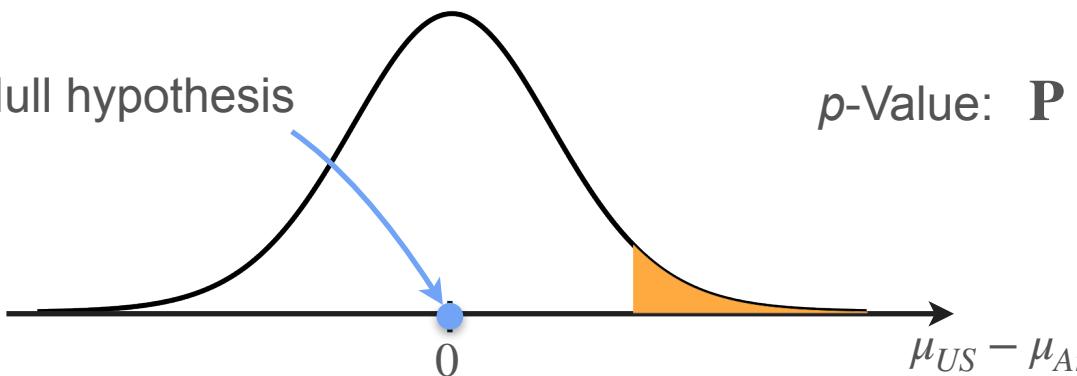
$$\bar{y} = 65.949$$

$$s_Y = 3.106$$

$$n_Y = 9$$

$$t = 1.7450$$

Null hypothesis



$$H_0 : \mu_{US} - \mu_{Ar} = 0 \text{ vs. } H_1 : \mu_{US} - \mu_{Ar} > 0$$

$$\alpha = 0.05$$

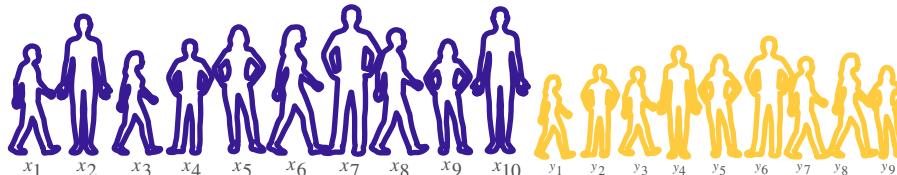
$$\text{If } H_0 \text{ is true: } T = \frac{(\bar{X} - \bar{Y}) - 0}{\sqrt{\frac{s_X^2}{10} + \frac{s_Y^2}{9}}} \sim t_{16.8}$$

$$p\text{-Value: } P(T > |\mu_{US} - \mu_{Ar}|)$$

$$= 0.0495 < 0.05$$

\Rightarrow Reject H_0 (and accept H_1)
(with a 5% significance level)

Independent Two-Sample t -Test: Two Tailed Test



$$\bar{x} = 68.442$$

$$s_x = 3.113$$

$$n_X = 10$$

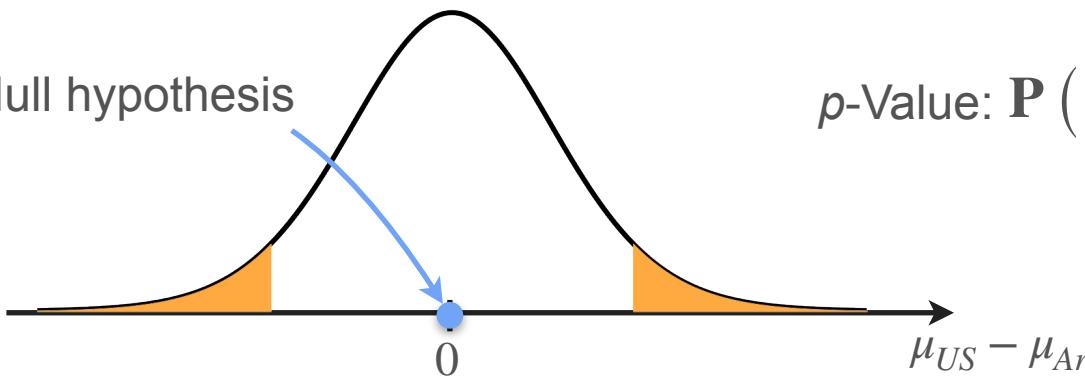
$$\bar{y} = 65.949$$

$$s_Y = 3.106$$

$$n_Y = 9$$

$$t = 1.7450$$

Null hypothesis



p -Value: $\mathbf{P}(|T| > |t| \mid \mu_{US} - \mu_{Ar} = 0)$

$$= 0.0991 > 0.05$$

\Rightarrow Do not reject H_0
(with a 5% significance level)

$$H_0 : \mu_{US} - \mu_{Ar} = 0 \text{ vs. } H_1 : \mu_{US} - \mu_{Ar} \neq 0$$

$$\alpha = 0.05$$

$$\text{If } H_0 \text{ is true: } T = \frac{(\bar{X} - \bar{Y}) - 0}{\sqrt{\frac{s_x^2}{10} + \frac{s_y^2}{9}}} \sim t_{16.8}$$

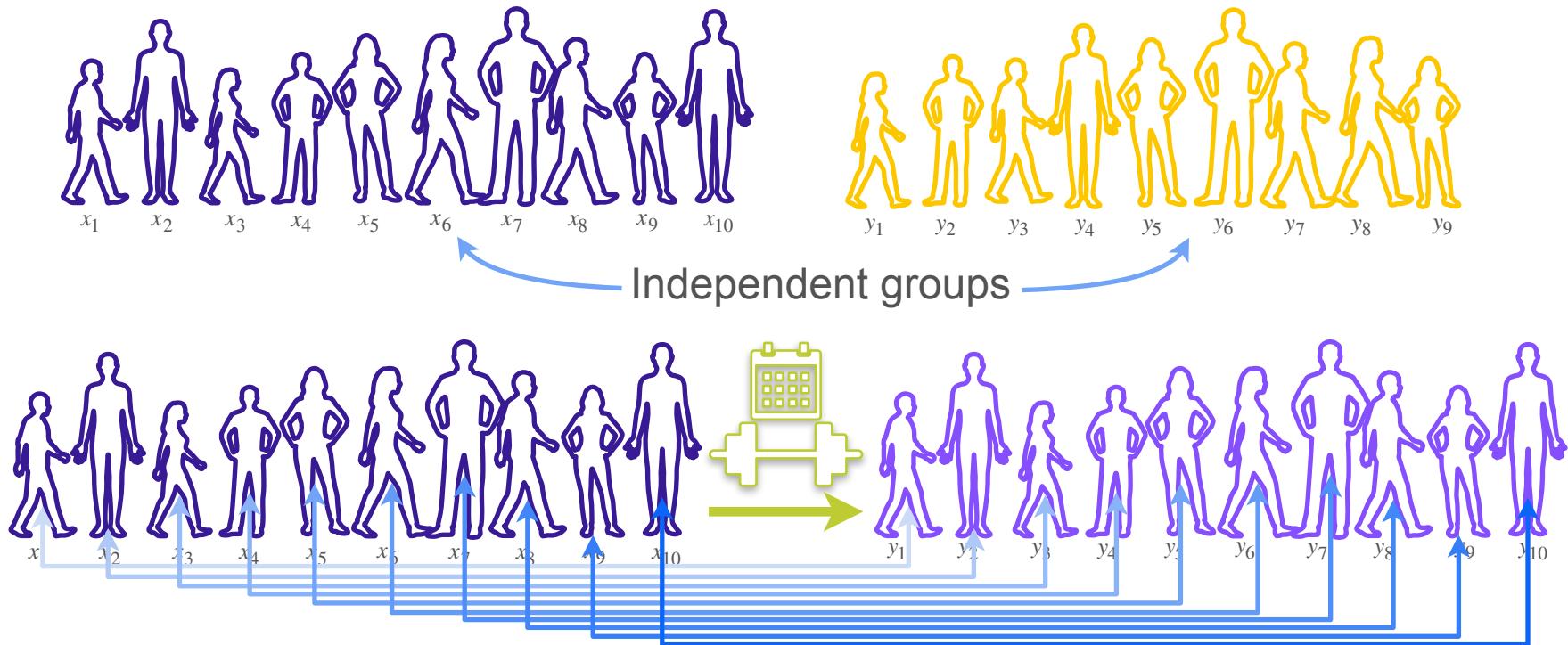


DeepLearning.AI

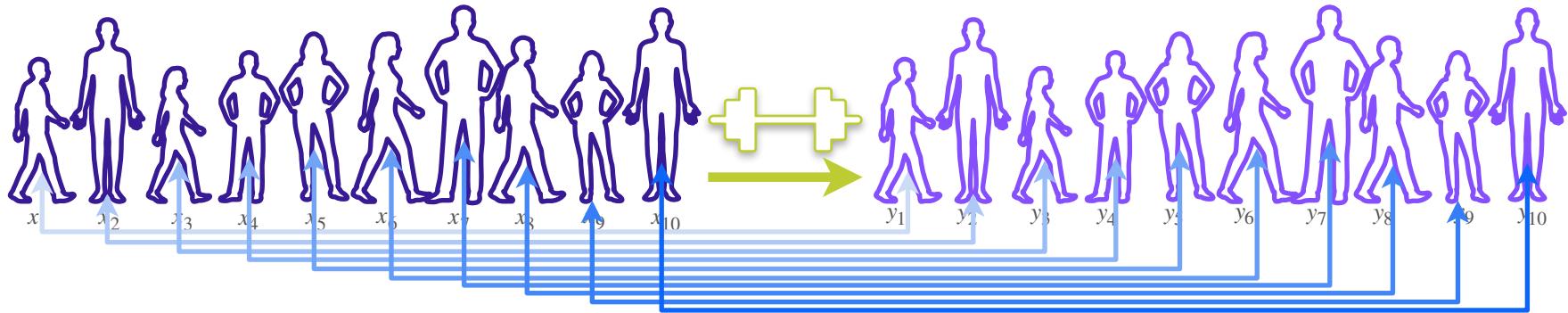
Hypothesis Testing

Paired t-test

Paired t -Test and Two-Sample t -Test



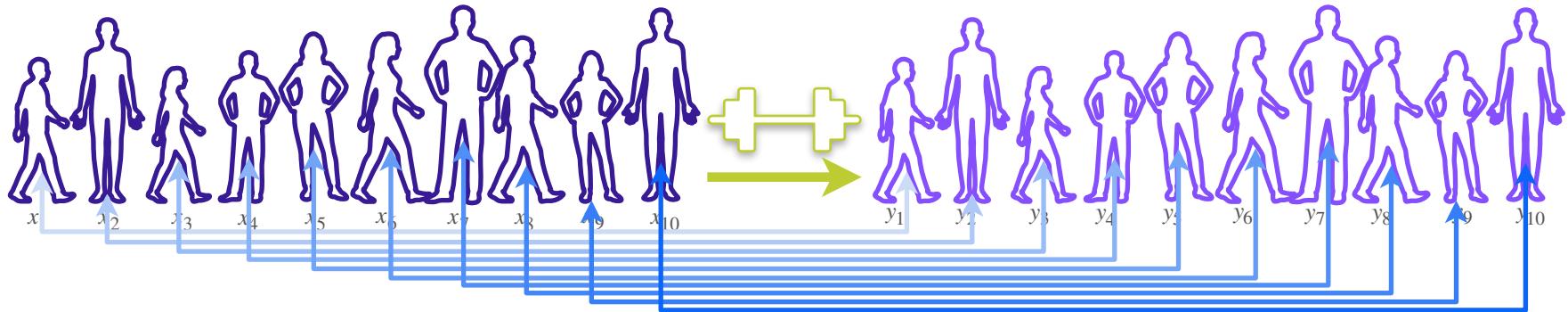
Paired t -Test: Statistic



Now you're interested in the difference between pair of samples

$$\frac{(X_1 - Y_1) + (X_2 - Y_2) + (X_3 - Y_3) + (X_4 - Y_4) + (X_5 - Y_5) + (X_6 - Y_6) + (X_7 - Y_7) + (X_8 - Y_8) + (X_9 - Y_9) + (X_{10} - Y_{10})}{10}$$
$$\frac{D_1 + D_2 + D_3 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9 + D_{10}}{10}$$

Paired t -Test: Statistic



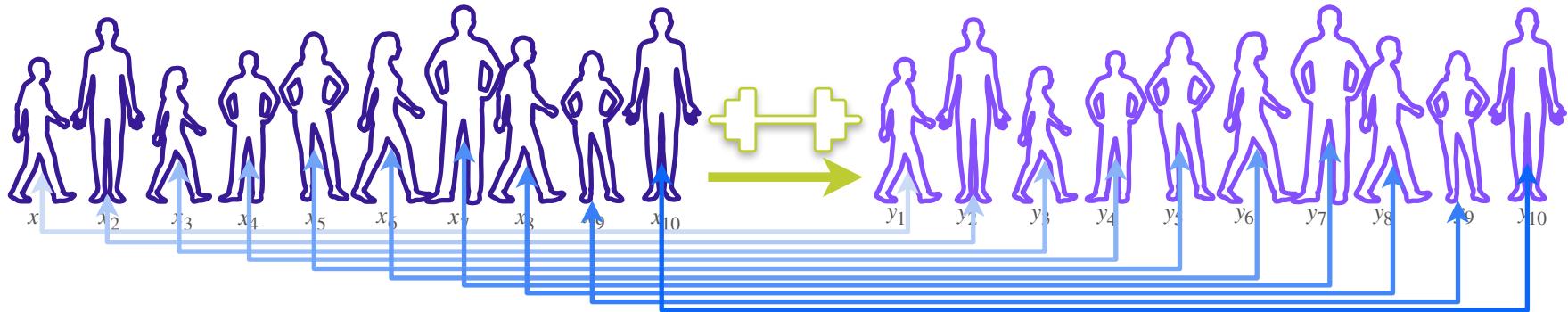
Now you're interested in the difference between pair of samples

$$(X_1 - Y_1) + (X_2 - Y_2) + (X_3 - Y_3) + (X_4 - Y_4) + (X_5 - Y_5) + (X_6 - Y_6) + (X_7 - Y_7) + (X_8 - Y_8) + (X_9 - Y_9) + (X_{10} - Y_{10})$$

10

$$\bar{D} = \frac{D_1 + D_2 + D_3 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9 + D_{10}}{10}$$

Paired t -Test: Statistic



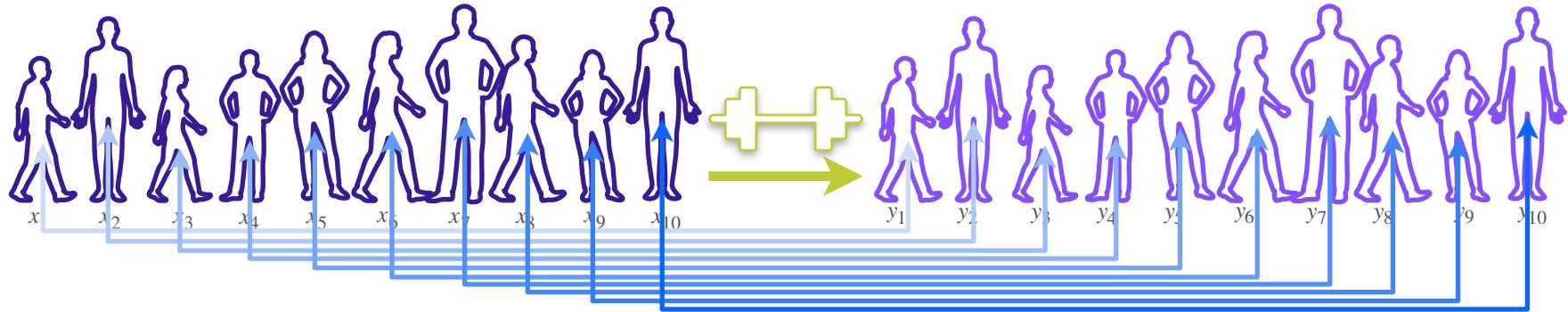
Now you're interested in the difference between pair of samples

$$\bar{D} = \frac{D_1 + D_2 + D_3 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9 + D_{10}}{10}$$

If X_i, Y_i are gaussian $\Rightarrow D_i = X_i - Y_i$ is gaussian.

$$D_i \stackrel{i.i.d}{\sim} \mathcal{N}(\mu_D, \sigma_D^2)$$

Paired t -Test: Statistic



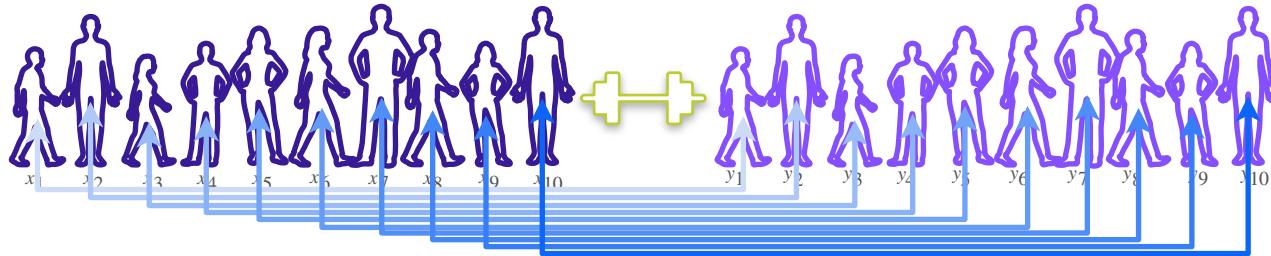
Now you're interested in the difference between pair of samples

$$\bar{D} = \frac{D_1 + D_2 + D_3 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9 + D_{10}}{10}$$

If X_i, Y_i are gaussian $\Rightarrow D_i = X_i - Y_i$ is gaussian.

$$D_i \stackrel{i.i.d}{\sim} \mathcal{N}(\mu_D, \sigma_D^2)$$

Paired t -Test: Statistic



Now you're interested in the difference between pair of samples

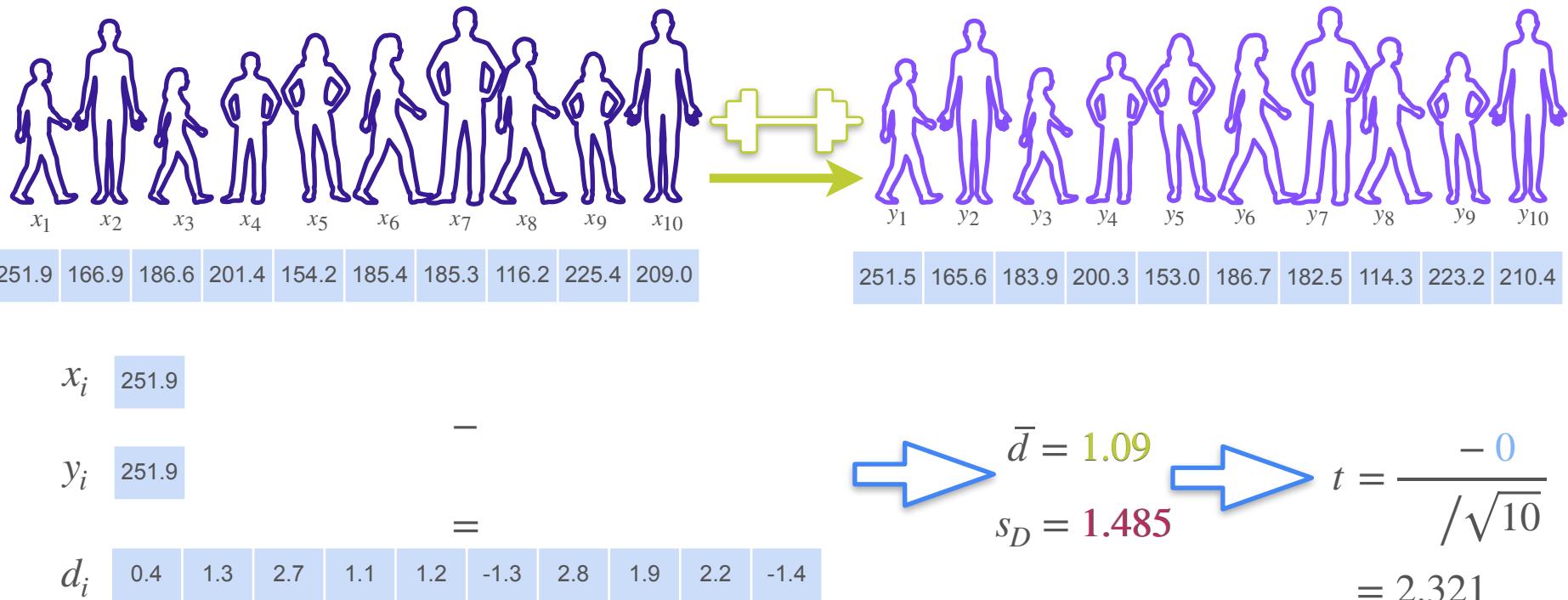
$$\bar{D} = \frac{D_1 + D_2 + D_3 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9 + D_{10}}{10} \quad D_i = X_i - Y_i$$
$$D_i \stackrel{i.i.d}{\sim} \mathcal{N}(\mu_D, \sigma_D^2)$$

$$\frac{\bar{D} - \mu_D}{\sigma_D / \sqrt{10}} \sim \mathcal{N}(0, 1^2)$$

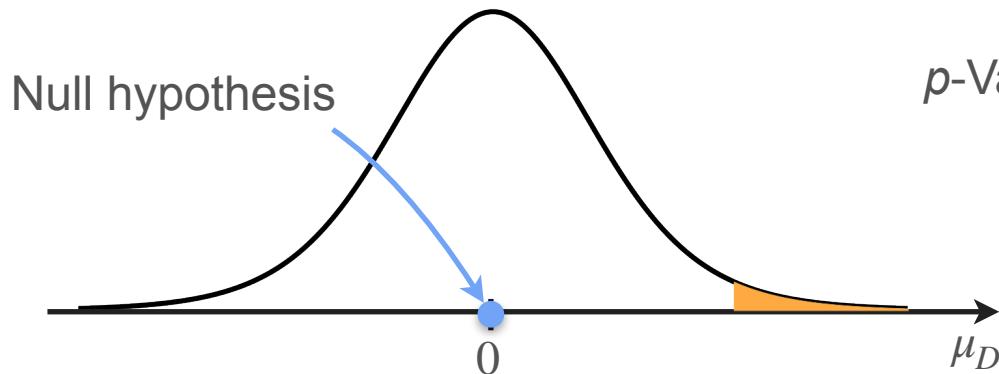
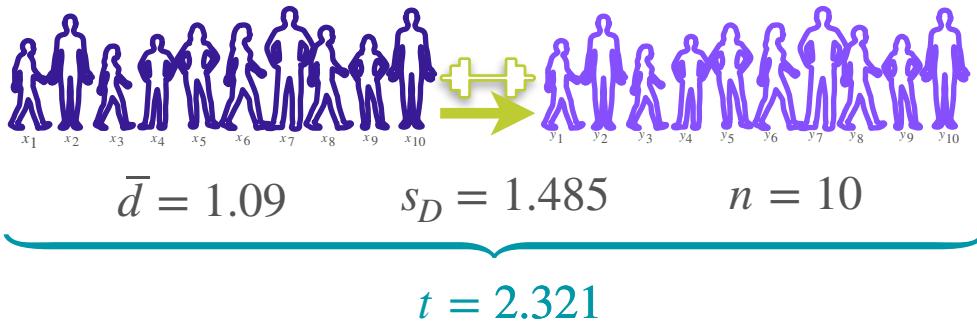
But σ_D is unknown $\Rightarrow \sigma_D \rightarrow S_D = \sqrt{\frac{\sum_{i=1}^{10} (D_i - \bar{D})^2}{10 - 1}}$ $\Rightarrow T = \frac{\bar{D} - 0}{S_D / \sqrt{10}} \sim t_{10-1}$

$H_0 : \mu_D = 0$ Test statistic

Paired t -Test: Observations



Paired Two-Sample t -Test: Right Tailed Test



$$H_0 : \mu_D = 0 \text{ vs. } H_1 : \mu_D > 0$$

$$\alpha = 0.05$$

$$\text{If } H_0 \text{ is true: } T = \frac{\bar{D} - 0}{S_D / \sqrt{10}} \sim t_{10-1}$$

$$p\text{-Value: } P(T > \quad \quad \quad | \mu_D = 0)$$

$$= 0.0227 < 0.05 \\ \Rightarrow \text{Reject } H_0 \text{ (and accept } H_1 \text{)} \\ (\text{with a 5\% significance level})$$

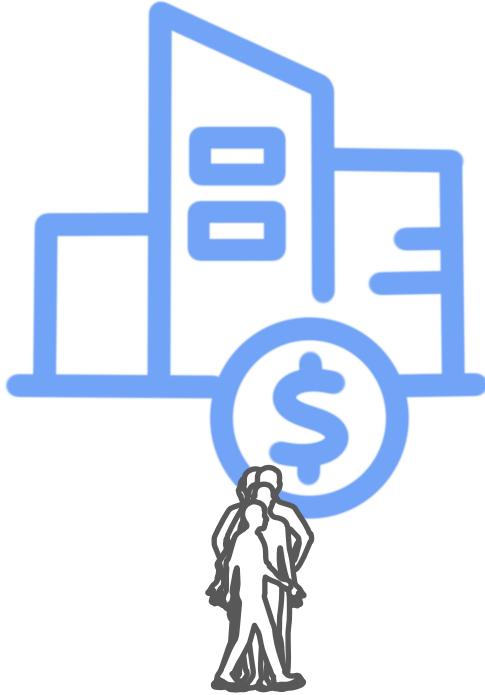


DeepLearning.AI

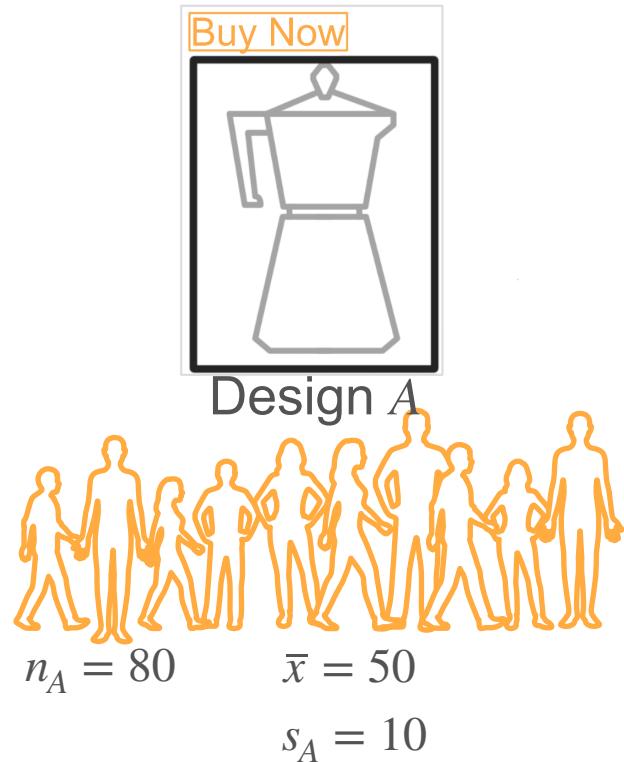
Hypothesis Testing

ML Application: A/B testing

A/B Testing: Purchase Amount



A/B Testing: Purchase Amount



A/B Testing: Purchase Amount



$$n_A = 80$$

$$\bar{x} = 50$$

$$s_A = 10$$

$$X \sim \mathcal{N}(\mu_A, \sigma_A^2)$$



$$n_B = 20$$

$$\bar{y} = 55$$

$$s_B = 15$$

$$Y \sim \mathcal{N}(\mu_B, \sigma_B^2)$$

$$H_0 : \mu_A - \mu_B = 0 \text{ vs. } H_1 : \mu_A - \mu_B < 0$$

$$\alpha = 0.05$$

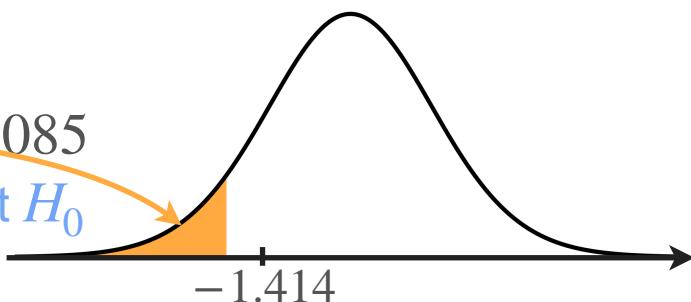
$$\text{If } H_0 \text{ is true: } T = \frac{(\bar{X} - \bar{Y}) - 0}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} \sim t_{23.38}$$

$$t = \frac{(-) - 0}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

$$-1.414$$

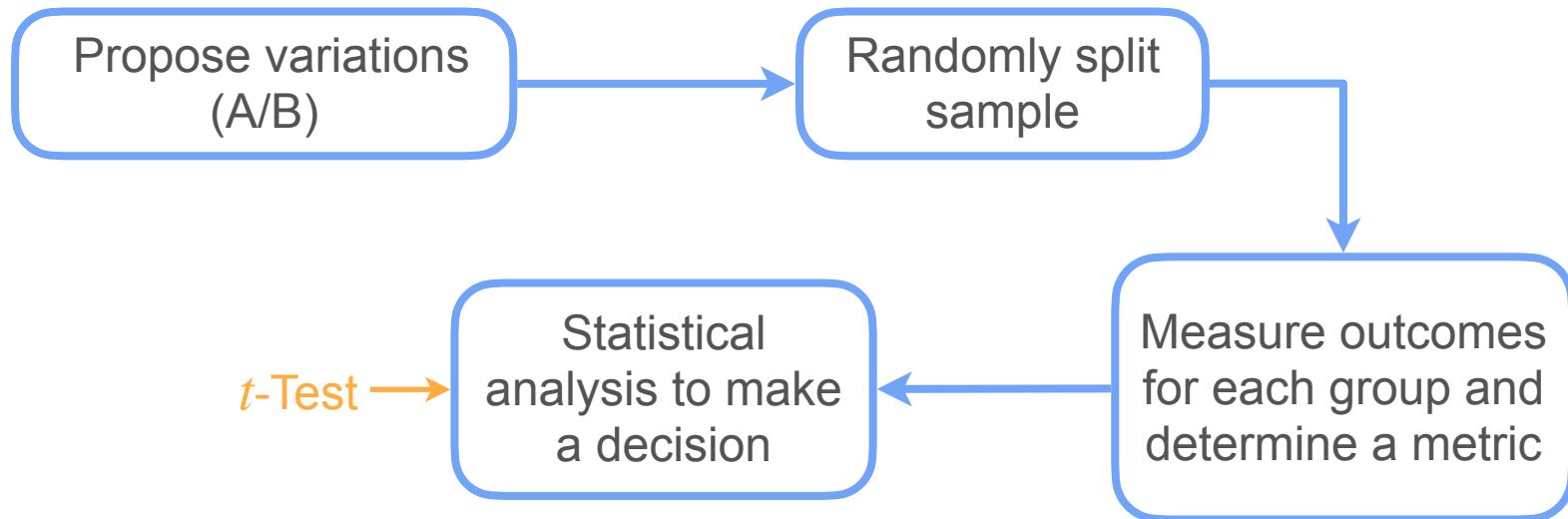
$$p\text{-Value: } 0.085$$

Don't reject H_0

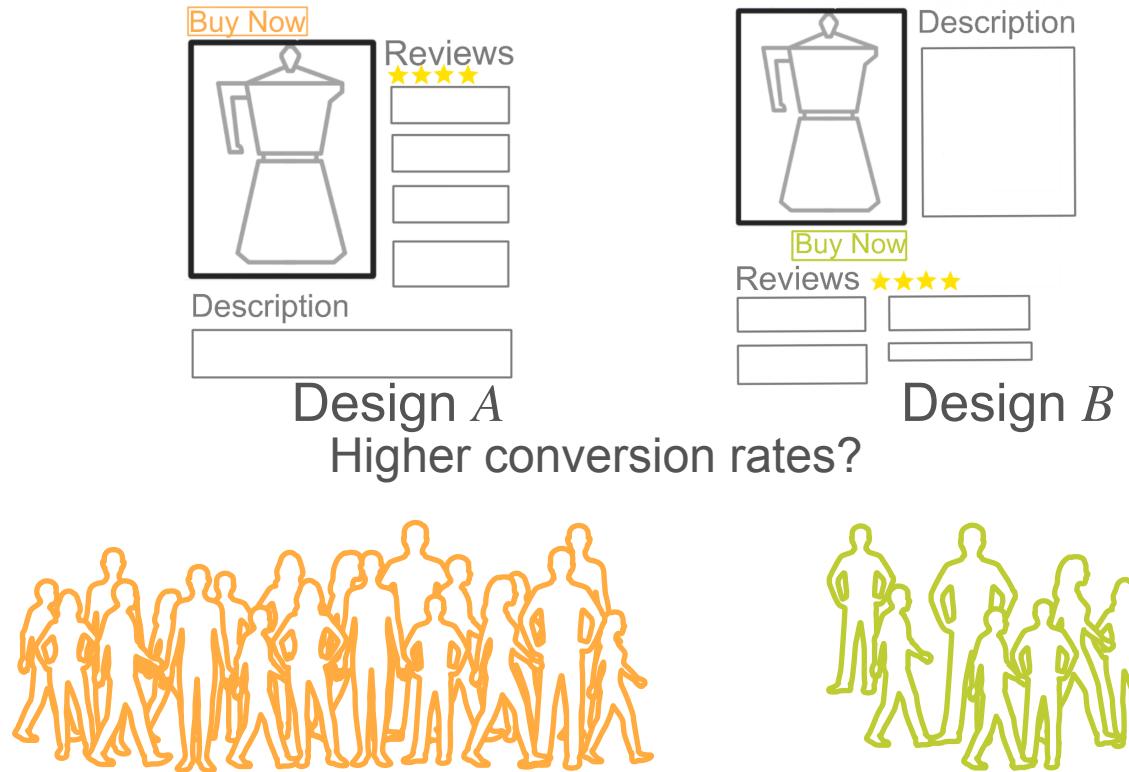


A/B Testing and t -Tests

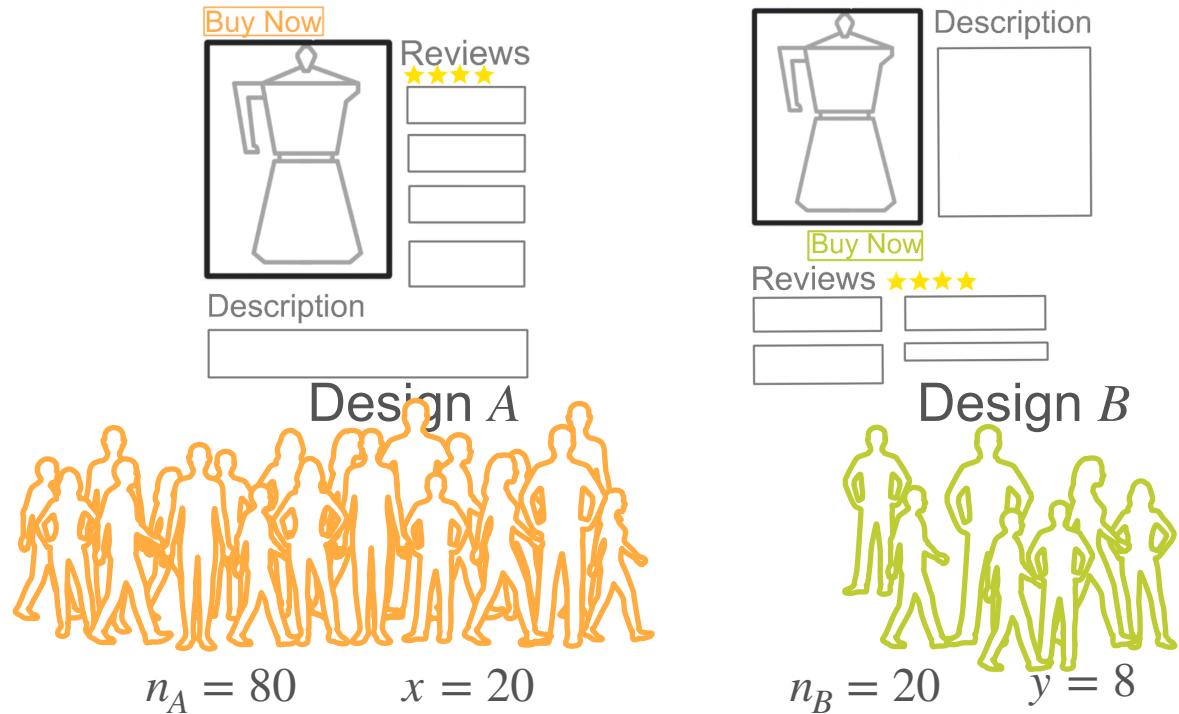
A/B testing is a methodology for comparing two variations (A/B)



A/B Testing: Conversion Rates



A/B Testing: Conversion Rates



A/B Testing: Conversion Rates



$$n_A = 80$$

$$x = 20$$



$$n_B = 20$$

$$y = 8$$

$$X \sim \text{Binomial}(n_A, p_A) \quad Y \sim \text{Binomial}(n_B, p_B)$$

$$H_0 : p_A - p_B = 0 \text{ vs. } H_1 : p_A - p_B < 0$$

p_A = Conversion rate from Design A

p_B = Conversion rate from Design B

$$\alpha = 0.05$$

A/B Testing: Conversion Rates

Statistic?

Law of large numbers

$$\frac{X}{n_A} \rightarrow p_A$$



$$\frac{X}{n_A} \sim \mathcal{N}\left(p_A, \frac{p_A(1-p_A)}{n_A}\right)$$

$$\frac{Y}{n_B} \rightarrow p_B$$

C.L.T.



$$\frac{Y}{n_B} \sim \mathcal{N}\left(p_B, \frac{p_B(1-p_B)}{n_B}\right)$$

A/B Testing: Conversion Rates

Statistic?

$$\left. \begin{array}{l} \frac{X}{n_A} \stackrel{a}{\sim} \mathcal{N} \left(p_A, \frac{p_A(1-p_A)}{n_A} \right) \\ \frac{Y}{n_B} \stackrel{a}{\sim} \mathcal{N} \left(p_B, \frac{p_B(1-p_B)}{n_B} \right) \end{array} \right\} \quad \frac{X}{n_A} - \frac{Y}{n_B} \rightarrow p_A - p_B \quad \frac{\frac{X}{n_A} - \frac{Y}{n_B}}{\sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}} \stackrel{a}{\sim} \mathcal{N} (0, 1^2)$$

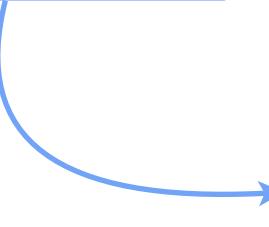
$\frac{X}{n_A} - \frac{Y}{n_B} \stackrel{a}{\sim} \mathcal{N} \left(p_A - p_B, \frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B} \right)$

↔

A/B Testing: Conversion Rates

If H_0 is true $\Rightarrow p_A = p_B = p$

$$\frac{\left(\frac{X}{n_A} - \frac{Y}{n_B}\right) - (p - p)}{\sqrt{\frac{p(1-p)}{n_A} + \frac{p(1-p)}{n_B}}} \stackrel{a}{\sim} \mathcal{N}(0, 1^2)$$



$$= p(1-p) \left(\frac{1}{n_A} + \frac{1}{n_B} \right) = p(1-p)(n_A + n_B) \frac{1}{n_A n_B}$$

A/B Testing: Conversion Rates

If H_0 is true $\Rightarrow p_A = p_B = p$

$$\frac{\left(\frac{X}{n_A} - \frac{Y}{n_B}\right) - (p - p)}{\sqrt{\frac{p(1-p)}{n_A} + \frac{p(1-p)}{n_B}}} \stackrel{a}{\sim} \mathcal{N}(0, 1^2) \longrightarrow \frac{\left(\frac{X}{n_A} - \frac{Y}{n_B}\right) - 0}{\sqrt{(n_A + n_B)p(1-p)}} \sqrt{n_A n_B} \stackrel{a}{\sim} \mathcal{N}(0, 1^2)$$

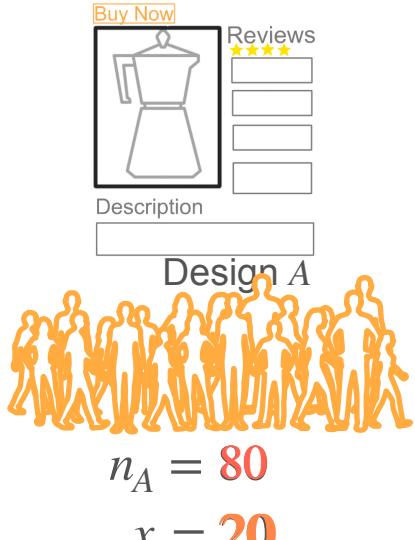
But you don't know p

Replace it by estimation! $\hat{p} = \frac{X + Y}{n_A + n_B}$

Test statistic

$$\frac{\left(\frac{X}{n_A} - \frac{Y}{n_B}\right) - 0}{\sqrt{(X + Y)\left(1 - \frac{X + Y}{n_A + n_B}\right)}} \sqrt{n_A n_B} \stackrel{a}{\sim} \mathcal{N}(0, 1^2)$$

A/B Testing: Conversion Rates



$$X \sim \text{Binomial}(n_A, p_A) \quad Y \sim \text{Binomial}(n_B, p_B)$$

$$H_0 : p_A - p_B = 0 \text{ vs. } H_1 : p_A - p_B < 0$$
$$\alpha = 0.05 \quad \text{If } H_0 \text{ is true } \Rightarrow p_A = p_B = p$$

$$Z = \frac{\left(\frac{X}{n_A} - \frac{Y}{n_B} \right) - 0}{\sqrt{(X+Y)\left(1 - \frac{X+Y}{n_A+n_B}\right)}} \sqrt{n_A n_B} \sim \mathcal{N}(0, 1^2)$$

$$z = \frac{\left(\frac{x}{n_A} - \frac{y}{n_B} \right) - 0}{\sqrt{\left(\frac{x}{n_A} + \frac{y}{n_B} \right) \left(1 - \frac{\frac{x}{n_A} + \frac{y}{n_B}}{n_A + n_B} \right)}} \sqrt{n_A n_B}$$

$$z = -1.336$$

A/B Testing: Conversion Rates



$$n_A = 80$$

$$x = 20$$

$$X \sim \text{Binomial}(n_A, p_A) \quad Y \sim \text{Binomial}(n_B, p_B)$$



$$n_B = 20$$

$$y = 8$$

$$H_0 : p_A - p_B = 0 \text{ vs. } H_1 : p_A - p_B < 0$$

$$\alpha = 0.05 \quad \text{If } H_0 \text{ is true} \Rightarrow p_A = p_B = p$$

$$Z = \frac{\left(\frac{X}{n_A} - \frac{Y}{n_B} \right) - 0}{\sqrt{(X+Y)\left(1 - \frac{X+Y}{n_A+n_B}\right)}} \sqrt{n_A n_B} \sim \mathcal{N}(0, 1^2)$$

$$z = -1.336$$

$$p\text{-value} = 0.091$$

Do not reject
 H_0

-1.336



DeepLearning.AI

Hypothesis Testing

Conclusion