Confidence Interval

Instructor, Nero Chan Zhen Yu

Forward School

Principal idea of confidence interval

- From an estimate made by measuring a sample (eg. Sample mean) AND knowing the error in that value.
- State an interval in which the population value is likely to be

How to define "likely"? >> level of confidence

Eg. 90%, 95%, …

Forward School

Confidence Interval Idea

Best Estimate ± Margin of Error

(or statistic)

"a few" x "average distances"

Reflects how confident we want to be

Reflects sampling variability in the statistic

Confidence Interval Idea

Best Estimate ± Margin of Error

(or statistic)

“a few” x “average distances”

Reflects how confident we want to be

Reflects sampling variability in the statistic

Forward School

One Sample CI for μ

- Take the sample of size n from a Normal population with unknown mean μ and known standard deviation σ.
- A level C confidence interval for μ is:

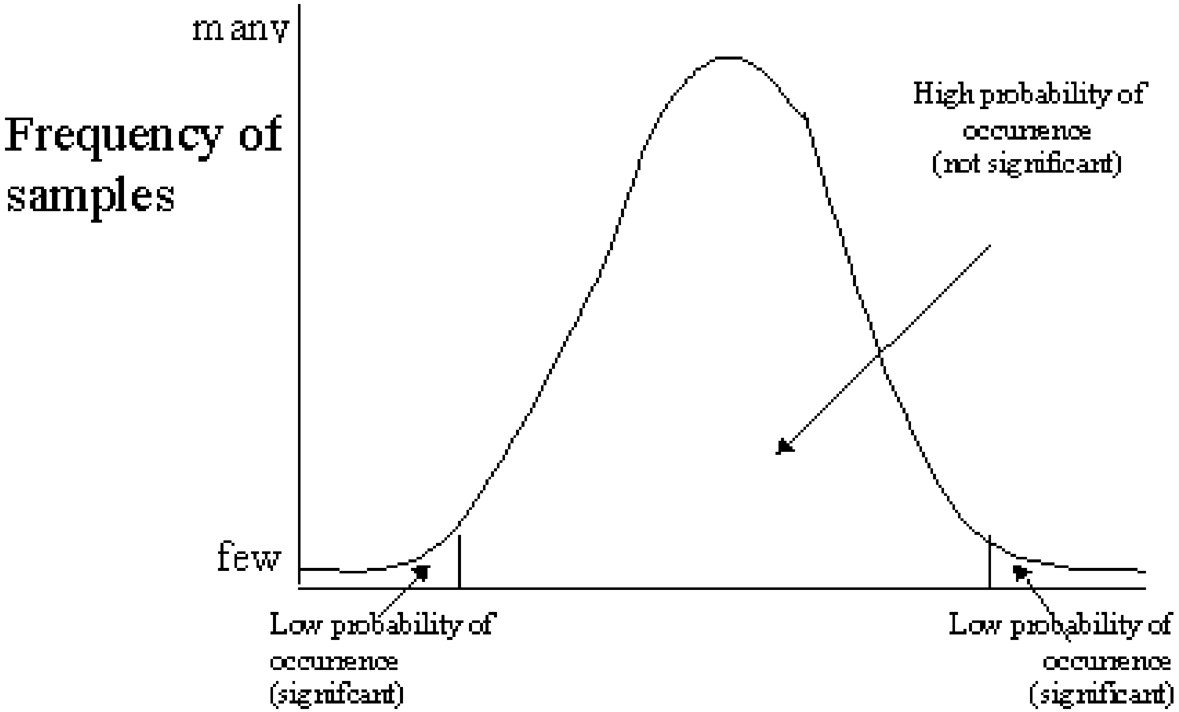
Sample mean ± margin of error

Sample mean ± multiplier × SD ($\bar{x}$)

$$\boxed{\bar{x} \pm \left(z \times \frac{\sigma}{\sqrt{n}}\right)}$$

Forward School

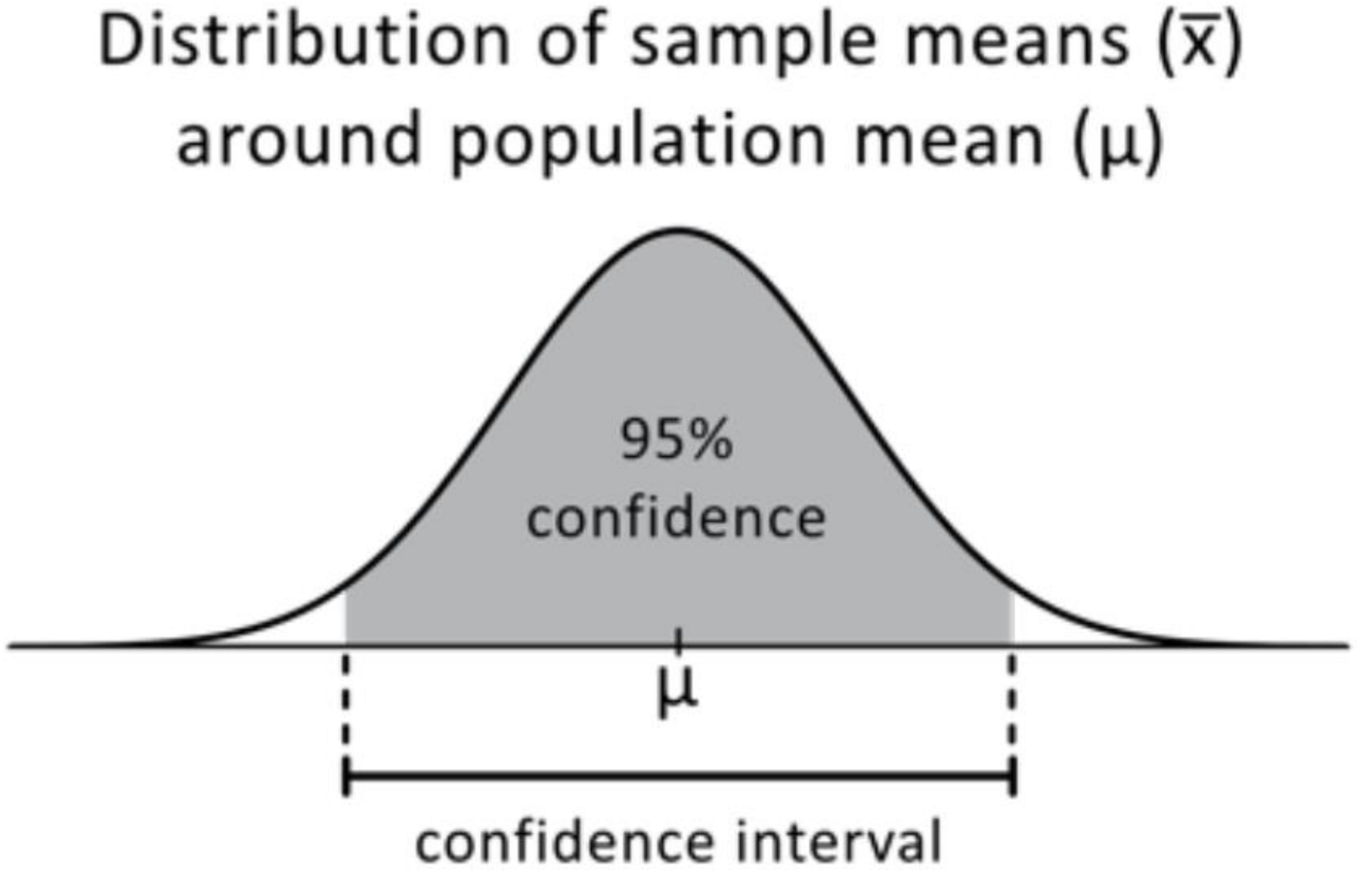
Sampling Variability of the Estimate (Statistic)



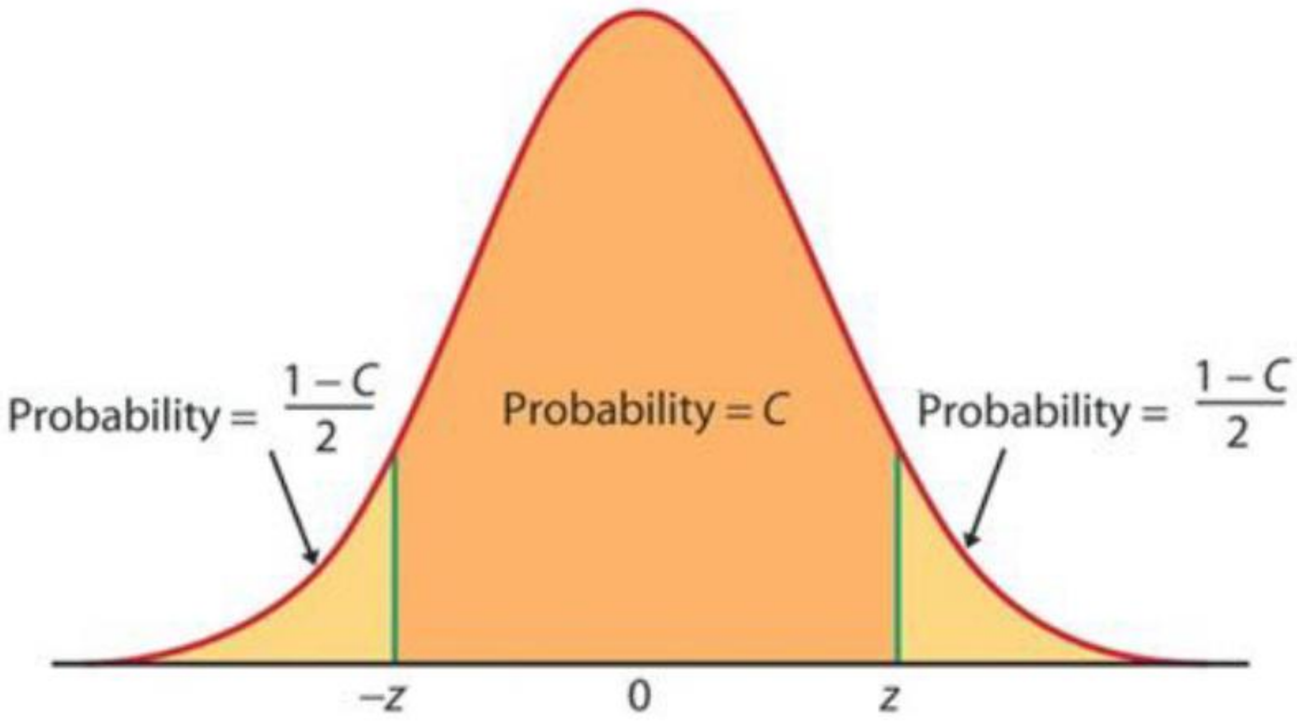
- Standard Deviation or Standard Error of the Statistic

Forward School

Confidence Interval Idea



A Level C confidence interval



z = percentile value,1.645 (for 90% CI), 1.96 (for 95% CI), ... with scipy:

if 90% CI then from

```
scipy.stats import norm
z=norm.ppf(0.95)
```

if 95% CI then

```
z=norm.ppf(0.975)
```

Forward School

Interpreting the level of confidence

The level of confidence is …

- In the long run, the % of all confidence intervals computed in this way will capture the population value.

Eg. 95% CI means in the long run, 95% of all confidence intervals computed in this way will capture the population value.

Forward School

Sample Poll

What proportion of parents report they use a car seat for all travel with their toddler?

Population – parents with a toddler

Parameter of Interest – A proportion

Construct a 95% Confidence Interval for the population proportion of parents reporting they use car seat for all travel with their toddler

Forward School

Sample Poll

A sample of 659 parents with toddler was taken and asked if they used car seat for all travel with their toddler.

540 parents responded **'YES'** to this question.

95% Confidence Interval Calculations

Best Estimate ± Margin of Error

$\hat{p} \pm$ **Margin of Error**

n = 659 **Sample size**

x = 540 **Number responsed 'YES"**

$\hat{p}$ = x/n = 540/659 = **0.85**

Forward School

95% Confidence Interval Calculations

Best Estimate ± Margin of Error

$\hat{p} \pm$ Margin of Error

$\hat{p} \pm$ "a few". estimated se $(\hat{p})$

$$\hat{p} \pm 1.96 \,.\, \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.85 \pm 1.96 \,.\, \sqrt{\frac{0.85(1-0.85)}{659}}$$

0.85 ± 0.0273

(0.8227, 0.8773)

Forward School

Interpreting Confidence Interval

Confidence Interval for ___parameter___?

We make a confidence interval for a ___parameter___?

parameter or statistic

Forward School

Interpreting Confidence Interval

“range of reasonable values for our parameter”

We estimate, with 95% confidence, the population proportion of parents with toddler who report they use a car seat for all travel with their toddler is somewhere between 82.27% - 87.73%

or

Based on our sample of 659 parents with toodlers, with 95% confidence, we estimate between 82.3% and 87.7% of all such parents report they use car seat for all travel with their toodler

Forward School

Interpreting Confidence Interval

“range of reasonable values for our parameter”

We estimate, with 95% confidence, the population proportion of parents with toddler who report they use a car seat for all travel with their toddler is somewhere between 82.27% - 87.73%

or

Based on our sample of 659 parents with toodlers, with 95% confidence, we estimate between 82.3% and 87.7% of all such parents report they use car seat for all travel with their toodler

Forward School

Interpreting Confidence Interval

“range of reasonable values for our parameter”

We estimate, with 95% confidence, the population proportion of parents with toddler who report they use a car seat for all travel with their toddler is somewhere between 82.27% - 87.73%

or

Based on our sample of 659 parents with toodlers, with 95% confidence, we estimate between 82.3% and 87.7% of all such parents report they use car seat for all travel with their toodler

Forward School

Interpreting Confidence Interval

Car Seats for toddlers example

Does our confidence interval of (0.8227 - 0.8773) contain ***sample proportion*** of parents with toddlers who report they use car seat for all travel with toddler

‘YES’ it most certainly does.. our interval is centered at that sample proportion of 0.85 or 85%

Forward School

Interpreting Confidence Interval

Car Seats for toddlers example

Does our confidence interval of (0.8227 - 0.8773) contain ***population proportion*** of parents with toddlers who report they use car seat for all travel with toddler

Based on the limited data that we got,
WE DON'T KNOW

Wrong understanding of Confidence Level

95% chance or **probability** that the population proportion is in this computed interval of (0.8227, 0.8773)

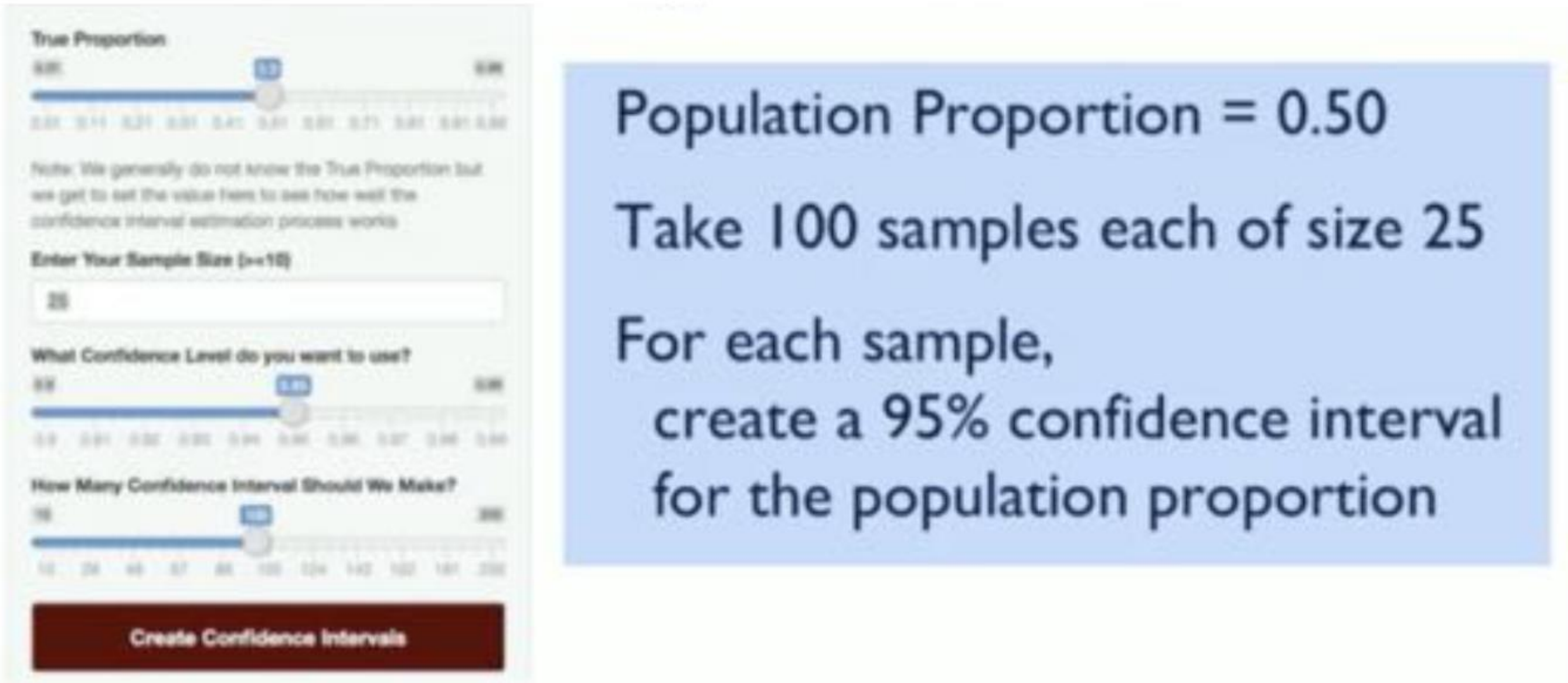
Forward School

Correct understanding of Confidence Level

95% confidence level refers to **our confidence in the statistical procedures** that was used to compute this interval

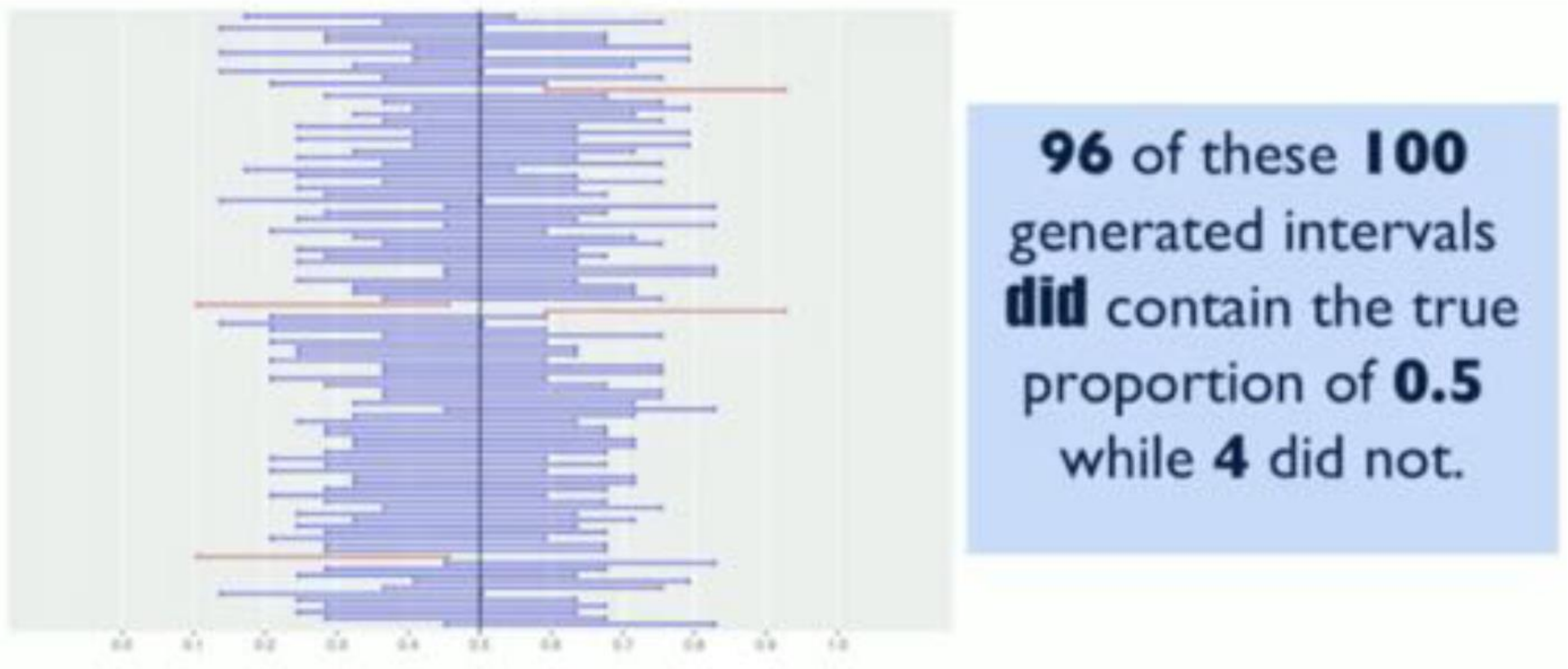
Forward School

Understanding Confidence Level



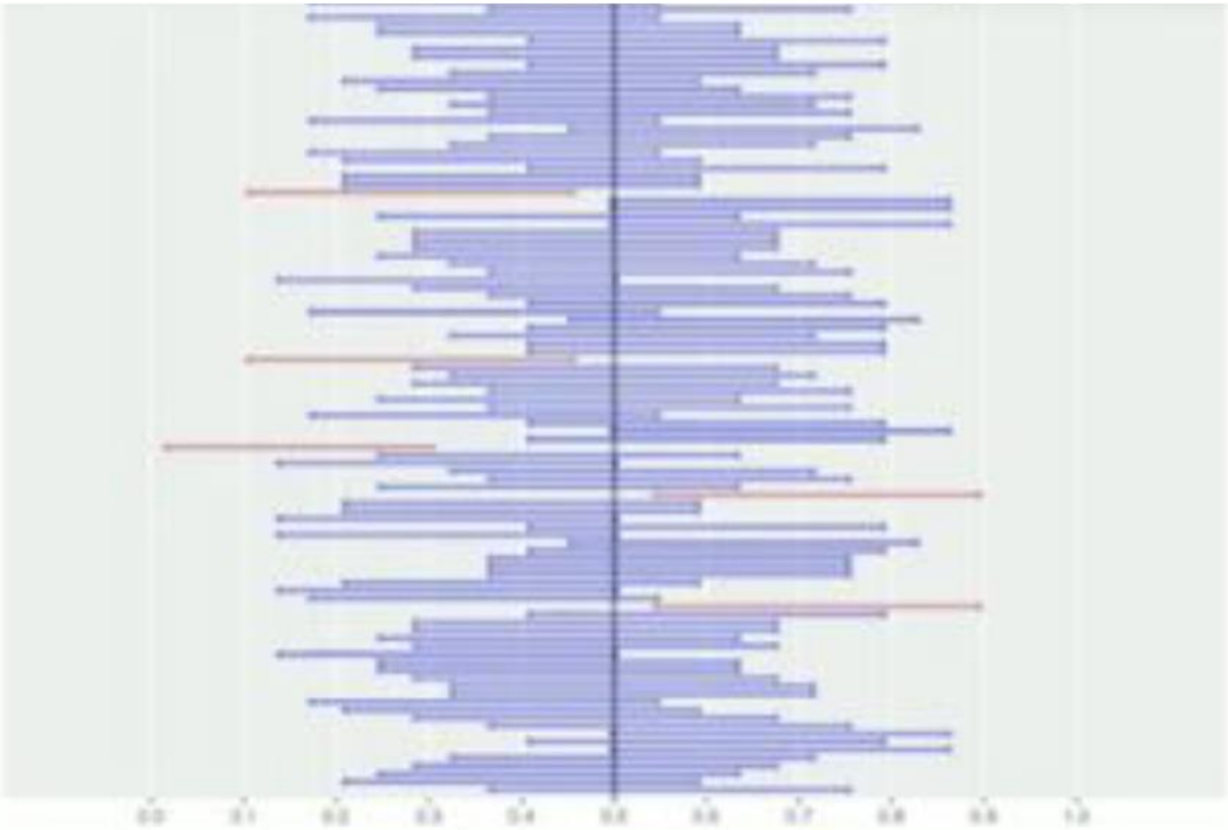
Forward School

Understanding Confidence Level



Forward School

Understanding Confidence Level



95 of these **100** generated intervals **did** contain the true proportion of **0.5** while **5** did not.

Forward School

Understanding Confidence Level

Forward School

Different Z Multipliers

90%	95%	98%	99%
1.645	1.96	2.326	2.576

Best Estimate ± Margin of Error

Best Estimate ± "a few" (estimated) standard errors

More confident → Larger Multiplier → Wider Interval

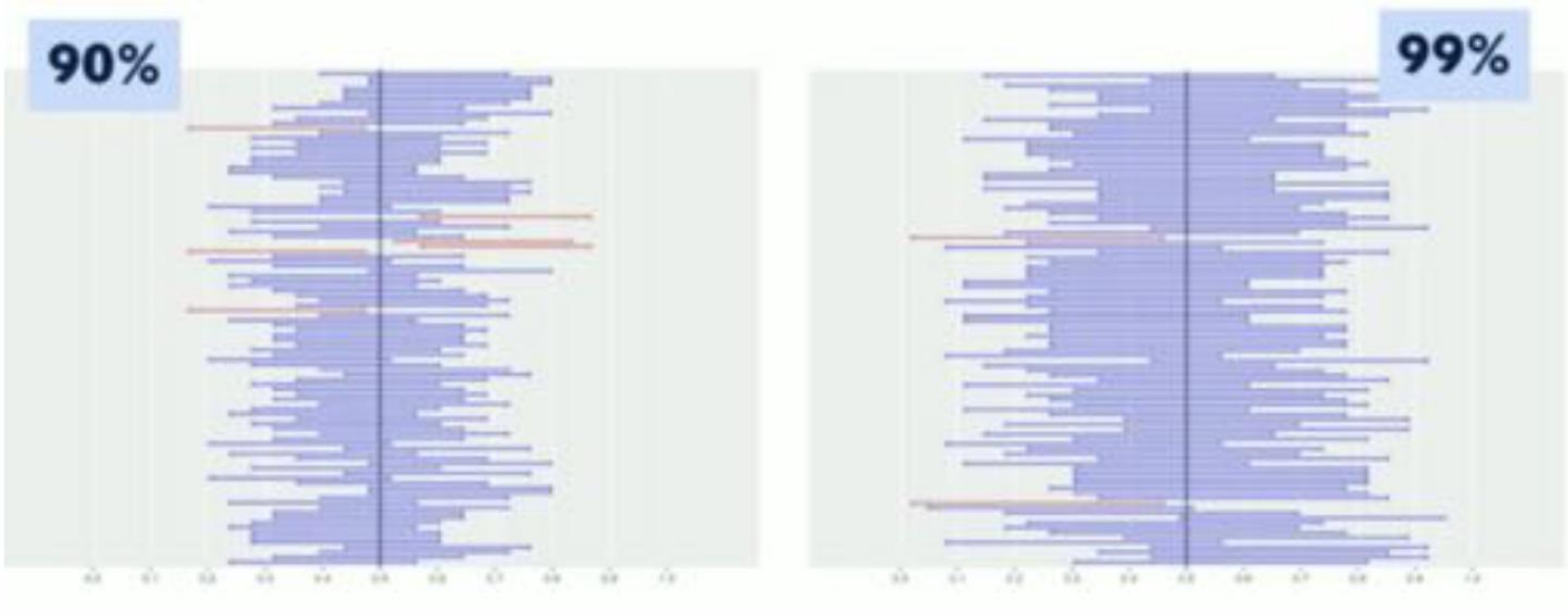
Forward School

Sample Poll Example

In a sample of 659 parents with a toddler, 540 (or **85%**) stated they **use a car seat** for all travel with their toddler.

90% CI:	95% CI:	99% CI:
0.85 ± 0.0229	0.85 ± 0.0273	0.85 ± 0.0358
82.7% to 87.3%	82.3% to 87.7%	81.4% to 88.6%

Changing Confidence Level



Forward School

In short…

- Confidence Intervals are used to give an *interval* estimate for our parameter of interest.
- Center of Confidence of Interval = our best estimate
- Margin of Error (MoE) = “a few” estimated standard of errors

Forward School

Demo

Confidence Interval Estimation

Demo Site

- Seeing Theory – A Visual Introduction to Probability and Statistics
- https://seeing-theory.brown.edu/frequentist-inference/index.html

Example 1

Stream health is measured in part from the dissolved oxygen content (DOC). A water authority collected a litre of water from each of 45 locations along a stream measured the DOC in each specimen. The mean content was found to be 4.62 mg/L.

Dissolved oxygen is known in this area (from long-term data) to be distributed Normally with σ = 0.92 mg/L.

a) Construct a 95% confidence interval for the average DOC in this particular stream at this time.

b) Is this strong evidence that the stream has a mean DOC less than 5 mg/L?

Forward School

Example 1

$\bar{x} = 4.62$, $\sigma = 0.92$, $n = 45$

a) Construct a 95% confidence interval for the average DOC in this particular stream at this time.

95% CI for μ:

$$\bar{x} \pm \left(z^* \times \frac{\sigma}{\sqrt{n}}\right) = 4.62 \pm \left(1.96 \times \frac{0.92}{\sqrt{45}}\right) = (4.35, 4.89)$$

Or using scipy: import numpy as np

```
(4.62 + np.array([+1, -1])*norm.ppf(0.975)*(0.92)/np.sqrt(45))
```

b) Is this strong evidence that the stream has a mean DOC less than 5 mg/L?

Yes, since the value of the CI is below 5mg/L

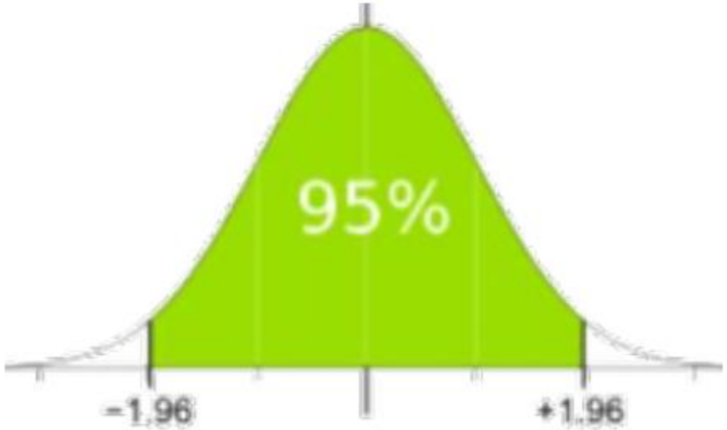
Example 2

Della wants to make a one-sample *z* interval to estimate what proportion of her community members favor a tax increase for more local school funding. She wants her margin of error to be no more than ±2% at the 95% confidence interval.

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

Forward School

Example 2



$\hat{p} \pm$ **Margin of Error**

$$\hat{p} \pm z^{*}.\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\hat{p} \pm \boxed{\mathbf{1.96}\,.\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \quad \leq 2\%$$

$$1.96.\sqrt{\frac{0.5.0.5}{n}} \leq 2$$

$$n \geq 49^2$$

$$n \geq 2401$$

Minimum sample size is 2401

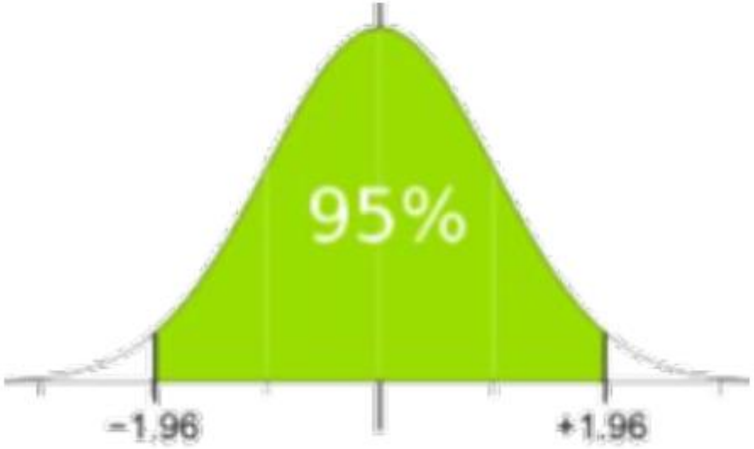
Example 3

A sample of 659 parents with toddler was taken and asked if they used car seat for all travel with their toddler. 540 parents responded **'YES'** to this question.

With confidence interval of 95%, calculate the Margin of Error.

Forward School

Example 3



$\hat{p} \pm$ **Margin of Error**

$$\hat{p} \pm z^{*} . \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\hat{p} \pm \boxed{\mathbf{1.96} \,.\, \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

$$\pm\, 1.96 . \sqrt{\frac{0.819\,(1-0.819)}{659}}$$

± 2.94

Margin of Error is 2.94

Forward School

Two Proportions

Research Question

What is the difference in population proportions of parents reporting that their children age 6-18 have had some swimming lessons between white and black children?

Populations - All parents of white children age 6-18 and all parents of black children age 6-18

Parameter of Interest - Different in Population Proportions (p_1-p_2)

We will let 1 = White and 2 = Black

Forward School

Survey Results

- A sample of 247 parents of black children age 6-18 was taken with 91 saying that their child has had some swimming lessons
- A sample of 988 parents of white children age 6-18 was taken with 543 saying that their child has had some swimming lessons

Difference in Proportion Confidence Interval

Best Estimate ± Margin of Error

$$\hat{p}_1 - \hat{p}_2 \pm \text{Margin of Error}$$

$$\hat{p}_1 - \hat{p}_2 \pm \text{"a few"}.\text{se}(\hat{p}_1 - \hat{p}_2)$$

$$\hat{p}_1 - \hat{p}_2 \pm 1.96.\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Forward School

Best Estimate of the Parameter

$\hat{p}_1 = \frac{543}{988} = 0.55$

$\hat{p}_2 = \frac{91}{247} = 0.37$

$\hat{p}_1 - \hat{p}_2 = 0.55 - 0.37 = 0.18$

1 = White and 2 = Black

$\hat{p}_1 - \hat{p}_2 \pm 1.96 . \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

$0.18 \pm 1.96 . 0.0345$

0.18 ± 0.0677

$(0.1123, 0.2477)$

Forward School

Interpreting Confidence Interval

“range of reasonable vavlues for our parameter”

With 95% confidence, the population proportion of parents with white children who have taken swimming lessons is 11.23% to 24.77% higher than the population proportions of parents with black children who have taken swimming lessons.

Forward School

Intervals for Differences

Is there a difference between 2 parameters?

If parameters are equal $\longrightarrow$ difference is 0

If parameters are unequal $\longrightarrow$ difference is not 0

Look for 0 in the range of reasonable values

Forward School

Assumptions

- We need to assume that we have two independent random samples
- We also need large enough sample sizes to assume that the distribution of our estimate is normal. That we need

$n_1\hat{p}_1, n_1(1-\hat{p}_1), n_2\hat{p}_2 \text{and } n_2(1-\hat{p}_2)$ to be at least 10.

In other words, we need at least 10 YES's and 10 NO's for each sample

One Sample CI for μ

If σ is unknown, then we use *s* to estimate σ and student t distribution is used instead of normal distribution

A level C confidence interval for μ is:

$$\boxed{\bar{x} \pm \left(t_{df} \times \frac{s}{\sqrt{n}}\right)}$$

df: degree of freedom

df = *n*–1, for the cases of one sample or match pair (termed as paired interval: e.g., before and after, follow-up and baseline)

Forward School

Estimating a Mean Difference for Paired Data

Paired Values

- Want to treat the two set of values simultaneously
- Other ways paired data arises:
  - Measurement collected on same individual

Forward School

Paired Values

- Want to treat the two set of values simultaneously
- Other ways paired data arises:
  - Measurement collected on same individual
  - Measurement collected on matched individual



Paired Values

- Want to treat the two set of values simultaneously
- Other ways paired data arises:
  - Measurement collected on same individual
  - Measurement collected on matched individual
- Variable: Difference of measurements within pairs

Research Question

What is the average difference between the older twin's and younger twin's self reported education?

Construct a 95% confidence interval for the mean difference of reported education for a set of identical twins

Population - All identical twins

Parameter of Interest - Population mean difference of self reported education level μ_d

Difference: Older - Younger

Forward School

Difference Calculation

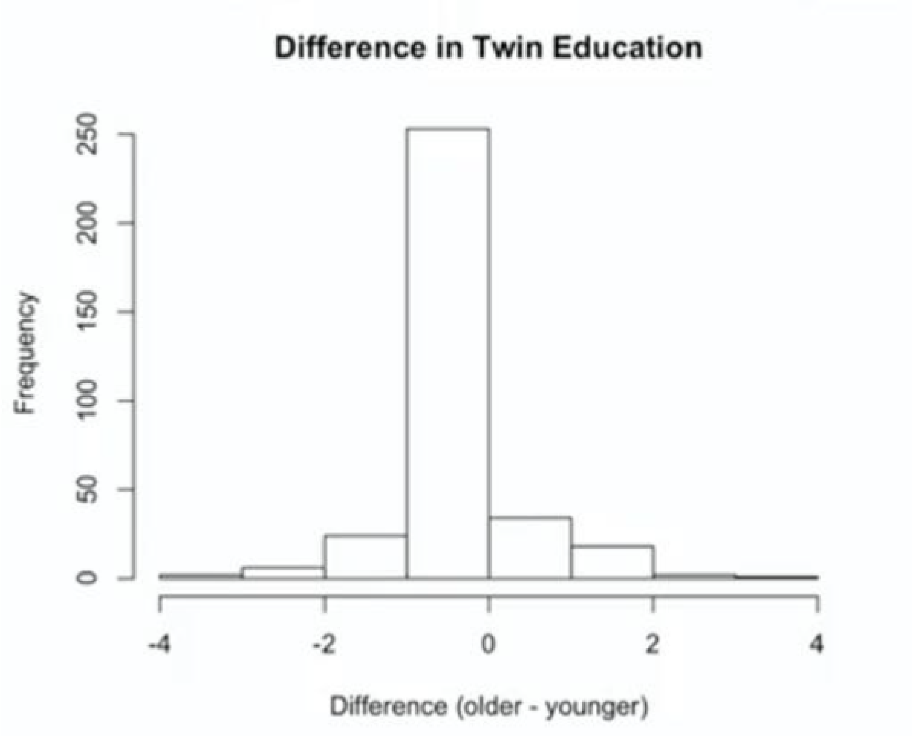
Difference = older twin - younger twin

Older twin education	Younger twin education	Difference (older - younger)
16	16	0
18	16	2
12	12	0
14	14	0
13	15	-2

Forward School

Difference Summary

Difference = older twin - younger twin



n = 340 observations

minimum = - 3.5 years
maximum = 4 years
72.1% had a difference of 0 years

mean = 0.0838 years
Standard Deviation = 0.7627 years

Forward School

95% Confidence Interval Calculations

Best Estimate ± Margin of Error

Sample mean difference ± "a few" · estimated standard error

$$\bar{x}_d \pm t^* \left(\frac{s_d}{\sqrt{n}}\right)$$

t^* multiplier comes from a t-distribution with $n - 1$ degrees of freedom

95% confidence

$n = 25 \rightarrow t^* = 2.064$

$n = 1000 \rightarrow t^* = 1.962$

Forward School

Mean Difference Confidence Interval

mean = 0.0838 years
Standard Deviation = 0.7627 years

n = 340 observations →t*=1.967

$$\bar{x}_d \pm t^*.(\frac{s_d}{\sqrt{n}})$$

$$0.084 \pm 1.967.(\frac{0.76}{\sqrt{340}})$$

$$0.084 \pm 1.967.(0.04)$$

$$0.084 \pm 0.814$$

$$(0.0025, 0.1652)$$ years

Forward School

Intrepreting Confidence Interval

“range of reasonable values for our parameters”

With 95% confidence, the population mean difference of the older twin’s less than younger twin’s self reported education is estimated to be between 0.0025 years and 0.6152 years.

Is there a difference between education levels of the older and younger twins?

Interval for differences

- Is there a mean difference between the education levels of twins?
- If education levels are generally equal $\longrightarrow$ mean difference is **0**
- If education levels are unequal $\longrightarrow$ mean difference is not **0**
- Hence, look for **0** in the range of reasonable values

Forward School

Assumptions

- We need to assume that we have a **random sample of identical twins**
- **Population of difference is normal** (for a **large sample size** can help to bypass this assumption)

Forward School

Estimating a Difference in Population Means with Confidence (for Independent Groups)

Forward School

Research Question

Considering Mexican-American Adults (age 18-29) living in the United States, do males and females differ significantly in mean Body Mass Index (BMI)?

Population - Mexican-American Adults (age 18-29) living in the United States

Parameter of Interest (μ_1-μ_2) - Body Mass Index or BMI (kg/m^2)

Forward School

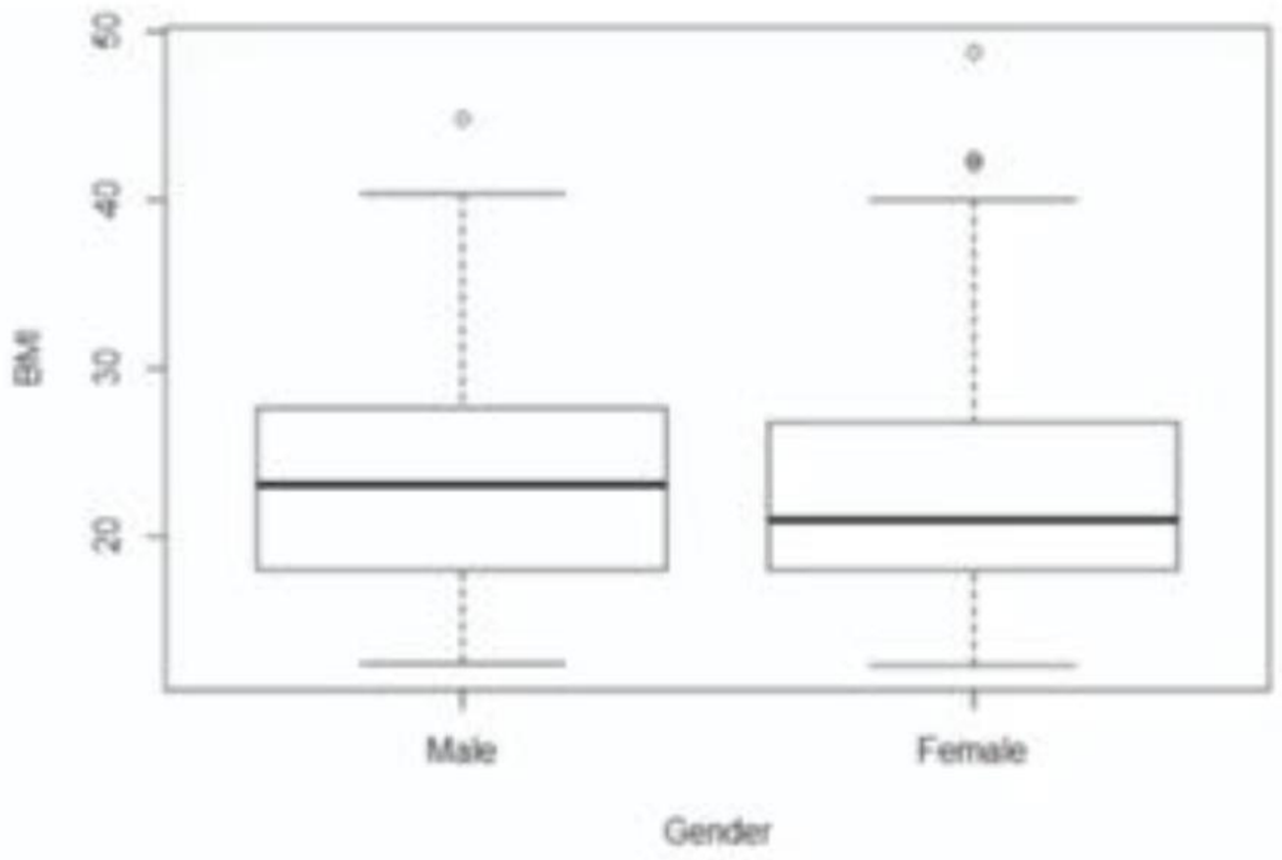
NHANES Data

Gender	BMI	Race	Age 18-29
1	19.9	1	1
2	17.0	1	1
2	26.7	1	1
1	25.6	1	1
...	...	...	...

The data was filtered to include only Mexican-American adults that were between the ages of 18 and 29.

BMI Variable Summary

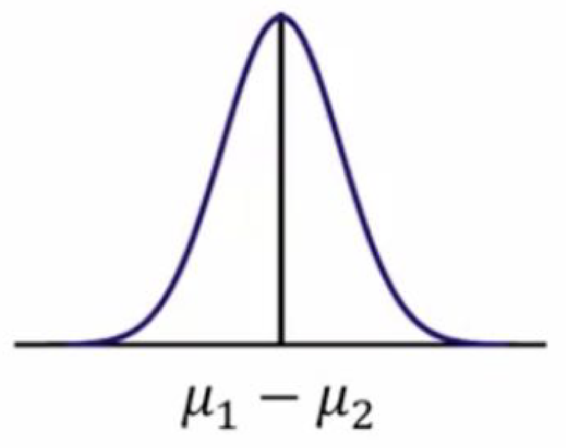
	Male	Female
Mean	23.57	22.83
St. Dev.	6.24	6.43
Min	12.5	12.4
Max	44.9	48.8
n	258	239



Forward School

Sampling Distribution of the Difference in Two (Independent) Sample Means

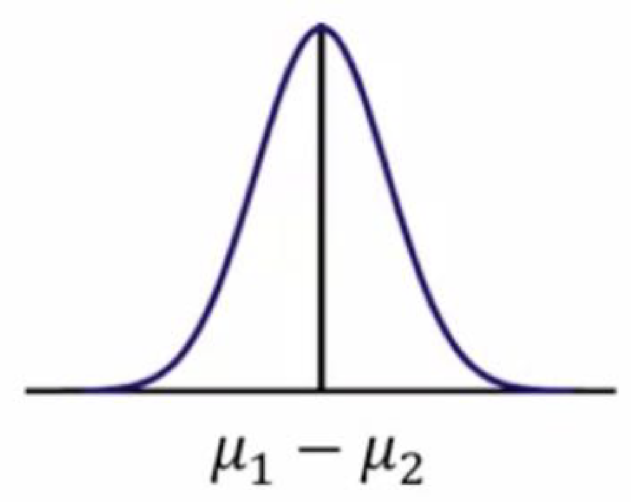
If models for both populations of responses are approximately normal (or sample sizes are both 'large' enough), distribution of the difference in sample means is (approximately) normal.



All possible values of <u>difference in sample means</u>

Forward School

Sampling Distribution of the Difference in Two (Independent) Sample Means



$$\text{Standard Error} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\text{Estimated Standard Error} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

All possible values of <u>difference in sample means</u>

Forward School

Confidence Interval Basics

Best Estimate ± Margin of Error

Best Estimate = Unbiased Point Estimate

Margin of Error = “a few” estimated Standard Errors

“ a few” = multiplier from appropriate distribution based on desired confidence level and sample design

95% confidence level ⟺ 0.05 significance

Forward School

Confidence Intervals Approaches

Pooled Approach

The variance of the two populations are assumed to be equal $(\sigma_1^2 - \sigma_2^2)$

Unpooled Approach

The assumption of equal variances is dropped

Unpooled Confidence Intervals Calculations

Best Estimate ± Margin of Error

Difference in sample means ± "a few" . estimated standard errors

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The degree of freedom (df) for t* is the approximated values of taking the smaller of n_1-1 and n_2 -1 (i.e df = min(n_1-1, n_2-1))

Forward School

Unpooled Confidence Intervals Calculations

Best Estimate ± Margin of Error

Difference in sample means ± "a few" . estimated standard errors

$$(\bar{x}_1 - \bar{x}_2) \pm \_?\_ \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

t* multiplier comes from a t-distribution with n_1+n_2-2 degrees of freedom. With the assumption of population variances are equal.

Forward School

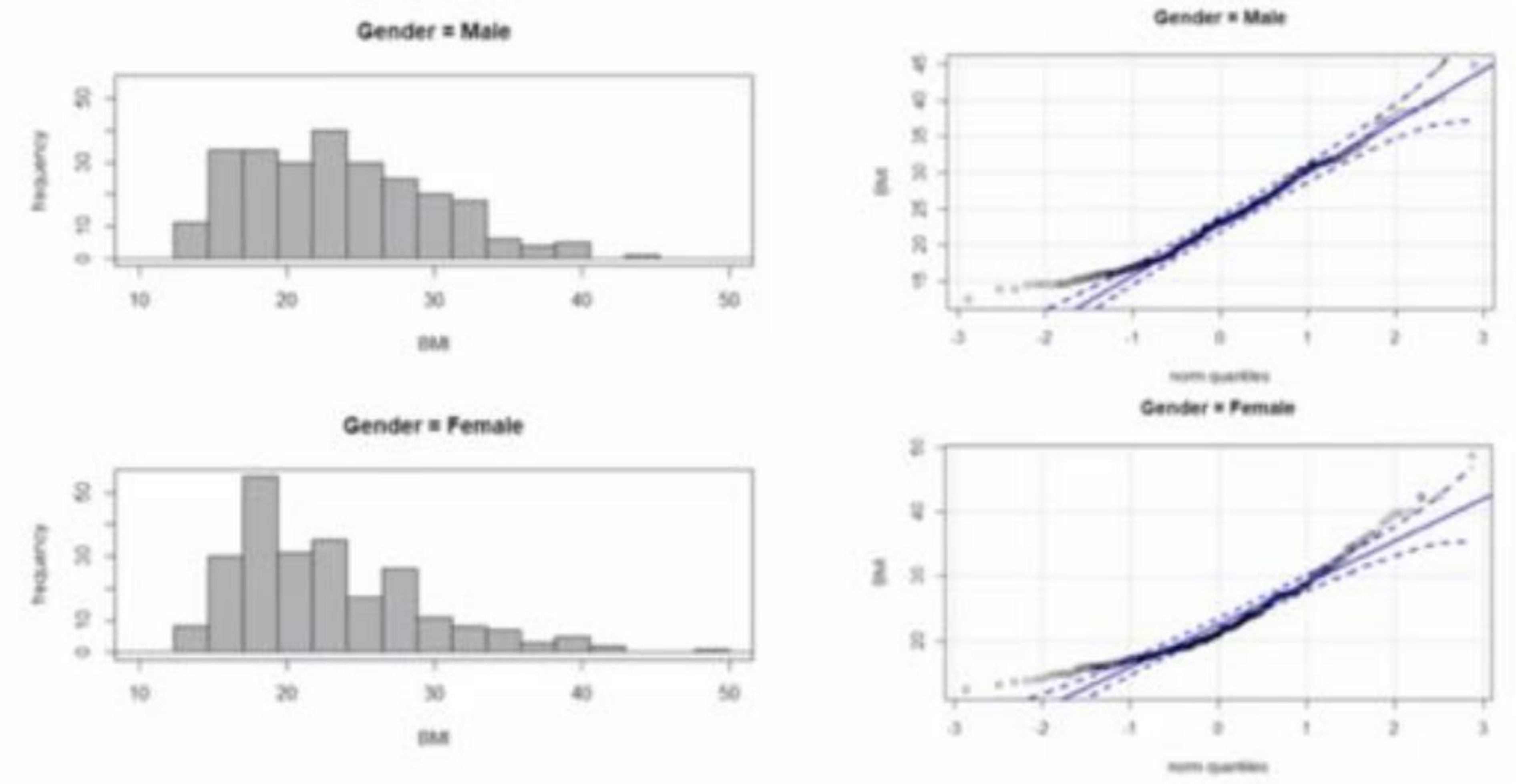
95% Confidence Interval Example

Considering Mexican-American adults (ages 18-29) living in the United States, do males and females differ significantly in Body Mass Index (BMI)?

Normality Assumption: models for both populations of responses are approximately normal (or sample sizes are both ‘large’ enough)

Forward School

95% Confidence Interval Example



Forward School

95% Confidence Interval Example

Considering Mexican-American adults (ages 18-29) living in the United States, do males and females differ significantly in Body Mass Index (BMI)?

Normality Assumption: models for both populations of responses are approximately normal (or sample sizes are both ‘large’ enough)

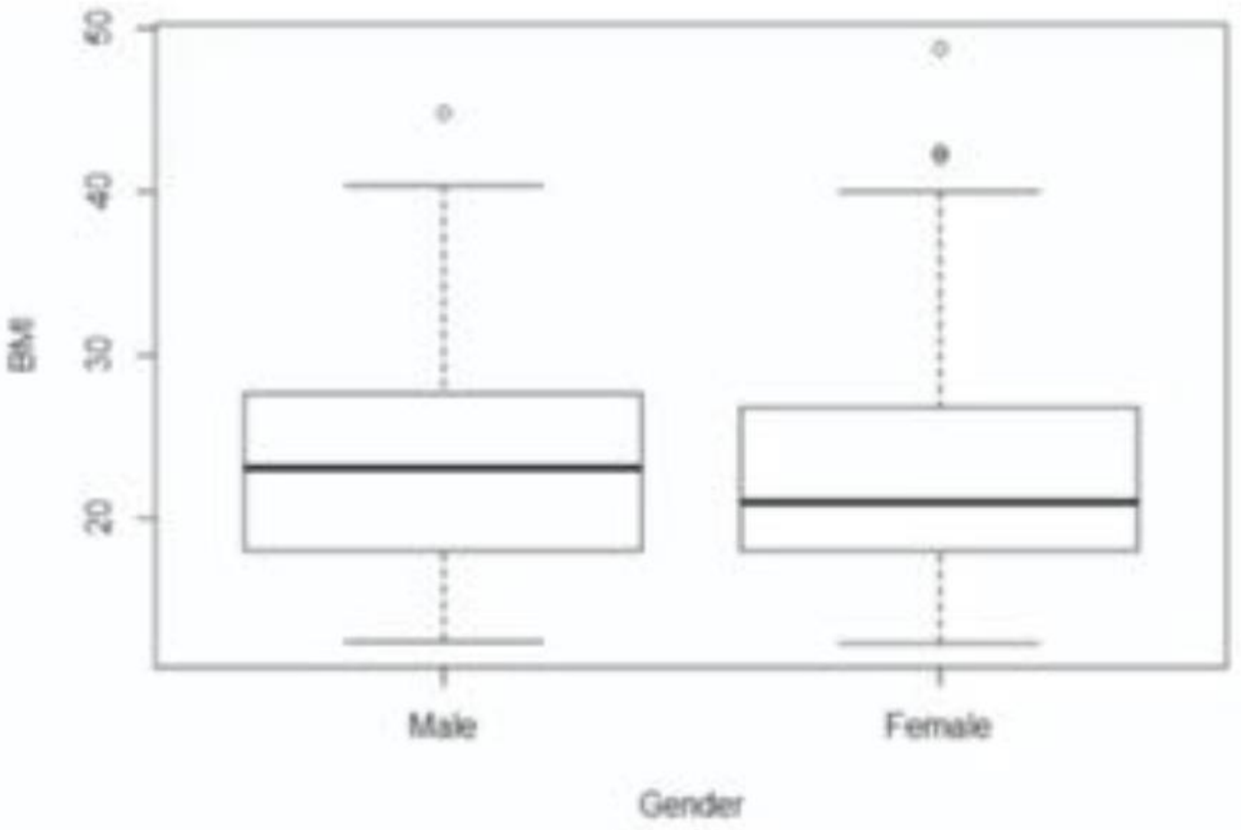
Both distributions have a slight-to-moderate right skew, but with large sample sizes, let us apply Central Limit Theorem (CLT) and continue

Variance Assumption: Enough evidence to assume equal variances between two populations - hence use **“pooled”** approach

Forward School

95% Confidence Interval Example

	Male	Female
Mean	23.57	22.83
St. Dev.	6.24	6.43
Min	12.5	12.4
Max	44.9	48.8
n	258	239



Forward School

95% Confidence Interval Example

	Male	Female
Mean	23.57	22.83
St. Dev.	6.24	6.43
Min	12.5	12.4
Max	44.9	48.8
n	258	239

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Here, t* multiplier of 1.98 will be used

Forward School

95% Confidence Interval Example

	Male	Female
Mean	23.57	22.83
St. Dev.	6.24	6.43
Min	12.5	12.4
Max	44.9	48.8
n	258	239

$$(23.57 - 22.83) \pm 1.98\sqrt{\frac{(258-1)6.24^2(239-1)6.43^2}{258+239-2}}\ \sqrt{\frac{1}{258}+\frac{1}{239}}$$

$$0.74 \pm 1.98\ (6.33)\ (0.0898)$$

$$(-0.385, 1.865)kg/m^2$$

Forward School

Interpreting Confidence Interval

$$(-0.385, 1.865)kg/m^2$$

“range of reasonable values for our parameter”

With 95% confidence, the difference in mean Body Mass Index (BMI) between males and females for all Mexican-American adults (ages 18-29) in the US is **estimated** to be between -0.385 kg/m^2 and 1.865 kg/m^2.

Forward School

Interpreting Confidence Interval

What does "with 95% confidence" mean?

If this procedure were repeated over and over, each time producing a 95% confidence interval estimate,

we would **expect 95% of those resulting intervals, to contain the difference in population mean in BMI.**

Forward School

Summary

- Confidence Intervals are used to give an interval estimate for our parameter of interest - **difference in population means**
- Center of the Confidence Interval is our best estimate **~ difference in sample means**
- Margin of Errors is “a few” (estimated) standard errors **~ for two means we use t* multipliers (pooled vs unpooled)**
- Assumptions for CI’s for Difference in Population Means
  - **Data are two simple random samples, independent**
  - **both populations of responses are normal (else large n helps)**
- Know how to interpret the **interval** and **level**

Forward School

The student *t* Distribution

- The *t* density curve is similar in shape to the standard
- Normal curve: symmetric about 0 and bell-shaped
- The spread of the *t* distributions is a bit greater than that of the standard Normal curve (i.e., the *t* curve is slightly "fatter" at the tail)
- As *n* increases, *s* is a more accurate estimation of σ. So *t* distribution changes in width with sample size as reflected in the degrees of freedom
- When *n* is large $s \approx \sigma$ and the t-distribution approach normal distribution:

Forward School

Example

A Cola maker wants to test a new recipes for loss of sweetness during storage. 10 trained tasters rate the sweetness before and after storage. The average difference (before – after) in the sweetness is 1.02 with a standard deviation 1.196

a) Construct a 95% confidence interval for the average difference of sweetness.

b) Are these data provide good evidence that the cola lost sweetness during storage?

Example

$\bar{x} = 1.02$, $s = 1.196$, $n = 10$

a) Construct a 95% confidence interval for the average difference of sweetness.

95% CI for μ: $\bar{x} \pm (t_{df} \times \frac{s}{\sqrt{n}})$

Using scipy: from scipy.stats import t

(1.02+np.array([+1,-1])* t.ppf(0.975,9)*1.196/np.sqrt(10)) = (0.16, 1.88)

b) Are these data provide good evidence that the cola lost sweetness during storage?

Yes, since the values in the CI are positive (more than 0)

Forward School

CI for two independent samples

Comparing:

- the responses to two treatments or
- the characteristics of two populations

There is

- a separate sample from each treatment or each population.
- individual measurements are not matched, and
- the samples can be of differing sizes or same size

- We are interested in the difference between the means (μ_a-μ_b)

Forward School

CI for μ_a-μ_b

Two cases:

1) Equal variances / common population variance / constant variance / pooled variance

2) Unequal variances / unpooled variance

We can conduct a test to check which is the correct case. However we will not discuss it here.

Forward School

CI for μ_a-μ_b

- Equal variance / common population variance / constant variance / pooled variance

$$s_p^2 = \frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{n_a + n_b - 2}$$

CI for μ_1-μ_2: $(\bar{x}_a - \bar{x}_b) \pm t_{df}(s_p)\sqrt{\frac{1}{n_a} + \frac{1}{n_b}}$

$$df = n_a + n_b - 2$$

Forward School

Example

- Twelve randomly selected mature citrus trees of one variety have a mean height of 13.8 feet with a standard deviation of 1.2 feet, and fifteen randomly selected mature citrus trees of another variety have a mean height of 12.9 feet with a standard deviation of 1.5 feet.
- Assuming that the random samples were selected from normal populations with common variance. Construct a 90% confidence interval for the difference between the true average heights of the two kinds of citrus trees.

$$n_a = 12, \quad \bar{x}_a = 13.8, \quad s_a = 1.2$$
$$n_b = 15, \quad \bar{x}_b = 12.9, \quad s_b = 1.5$$

Forward School

Example

$n_a = 12, \quad \bar{x}_a = 13.8, \quad s_a = 1.2$

$n_b = 15, \quad \bar{x}_b = 12.9, \quad s_b = 1.5$

$$s_p^2 = \frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{n_a + n_b - 2} = \frac{(11)(1.2)^2 + (14)(1.5)^2}{25}$$

90% CI for μ_a-μ_b :

$$(\bar{x}_a - \bar{x}_b) \pm t_{n_a + n_b - 2}(s_p)\sqrt{\frac{1}{n_a} + \frac{1}{n_b}}$$

$$(13.8 - 12.9) \pm t_{df=25}(s_p)\sqrt{\frac{1}{12} + \frac{1}{15}}$$

Forward School

Example

R output:

$s_p^2 = 1.8936$

90% CI for μ_a-μ_b :

```
(13.8-12.9)+np.array([+1,-1])*t.ppf(0.95,25)*np.sqrt(1.8936)*np.sqrt(1/12+1/15)
```

= (-0.01, 1.81)

Forward School

Quiz