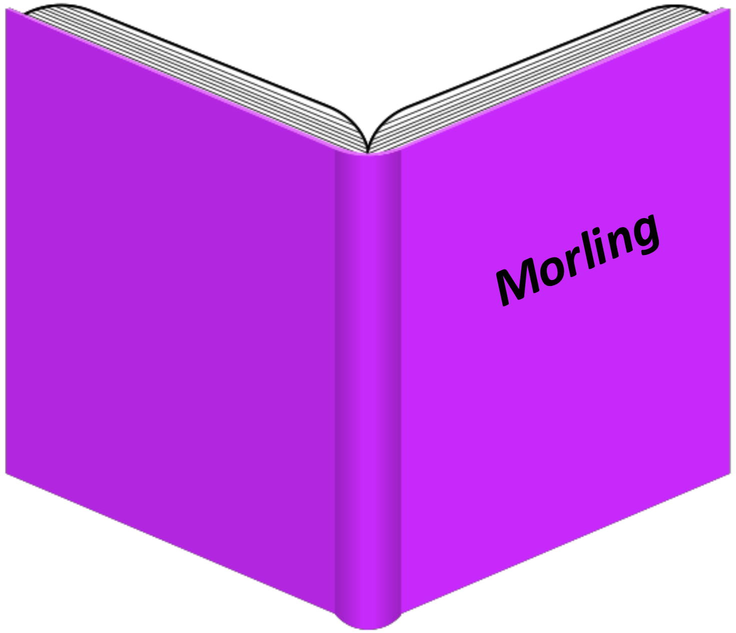
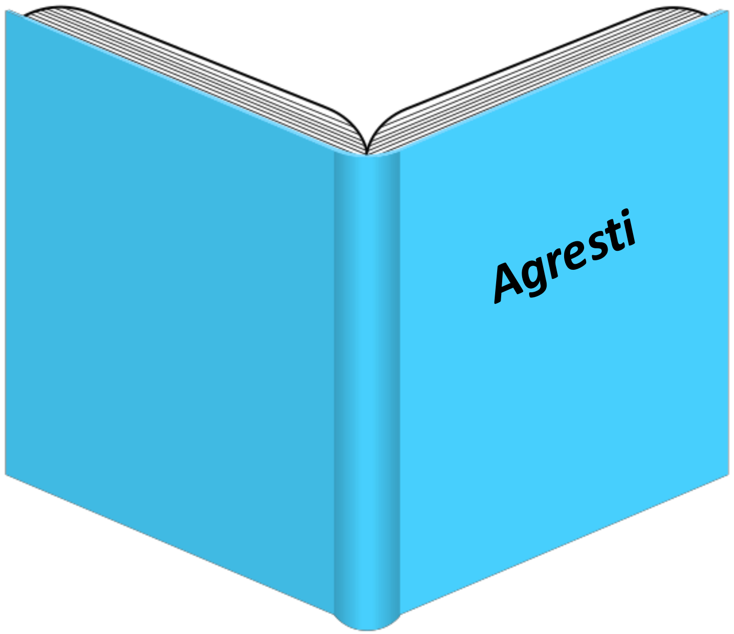


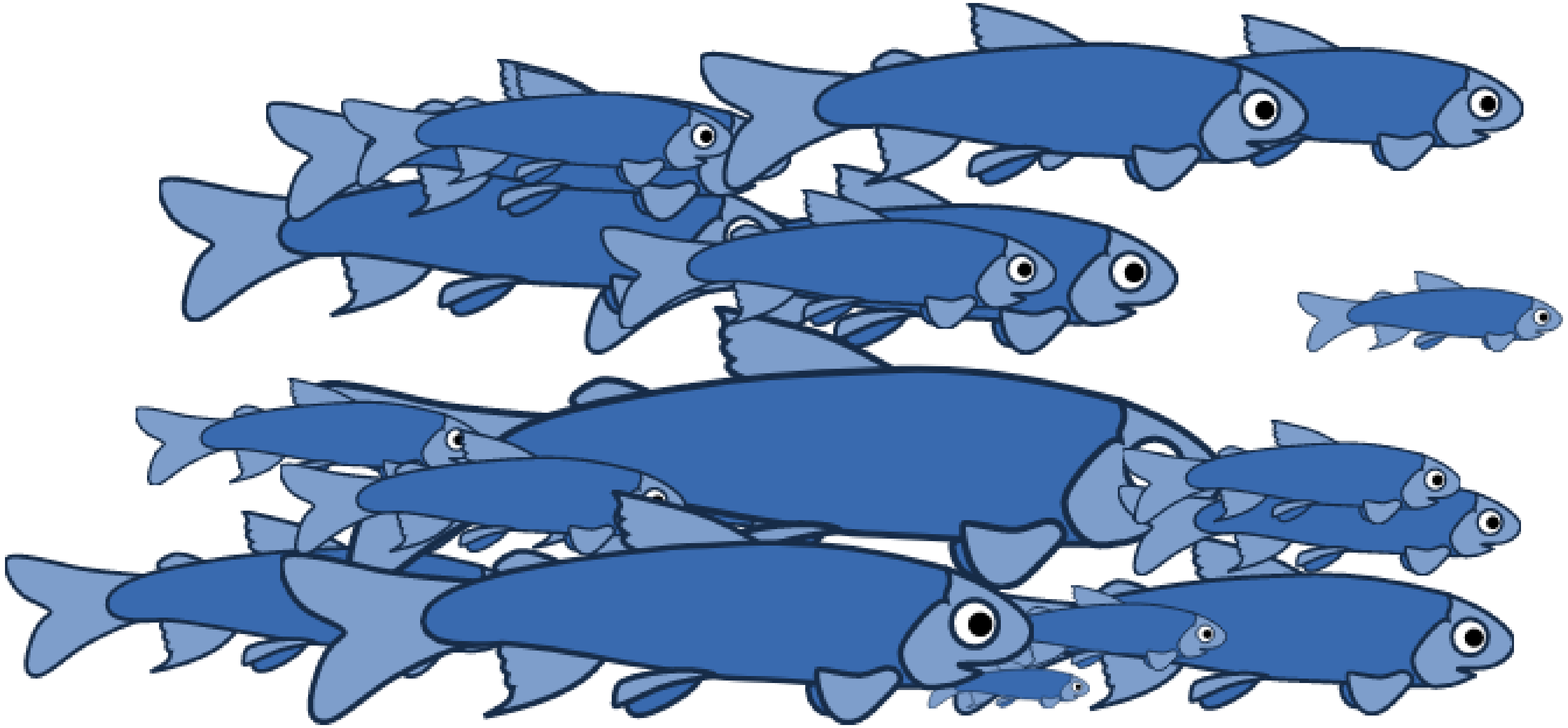


Research Methods and Statistics

Lecture 12: Confidence Intervals II

Johnny van Doorn

Pictures source: pixabay



https://www.zoology.ubc.ca/~whitlock/Kingfisher/SamplingNormal.htm

Overview of Today

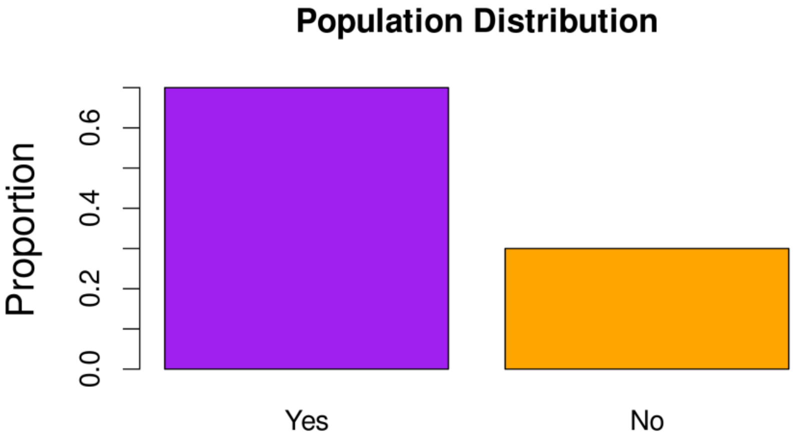
1. **Recap of last Thursday and Tuesday**
2. Confidence interval for the mean
   - *t*-distribution
3. Robustness against extreme values
4. Choosing a sample size
5. Recap
   - Example exam question

Recap

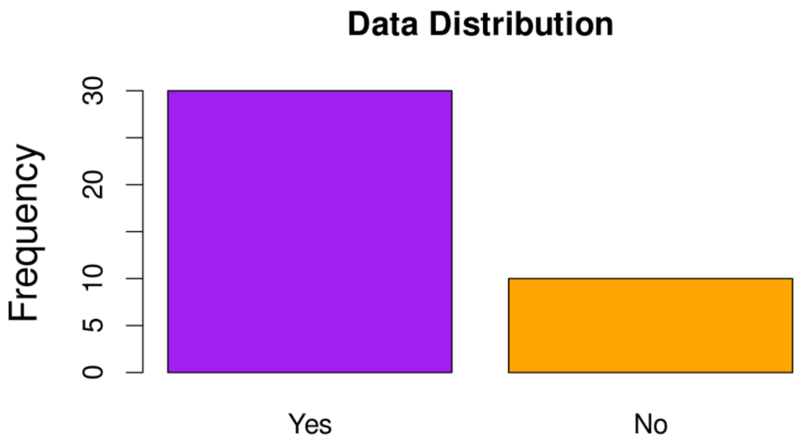
Sample (n=40)

Sampling distribution

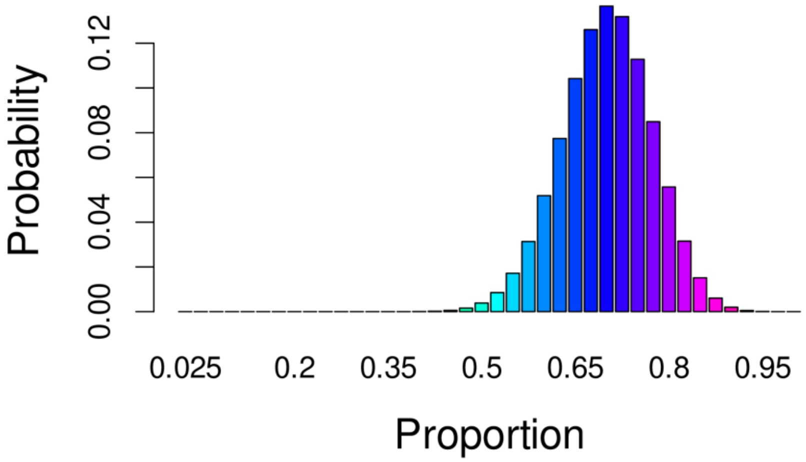
Binomial distribution
n=40, p=0.7



$p = 0.7$



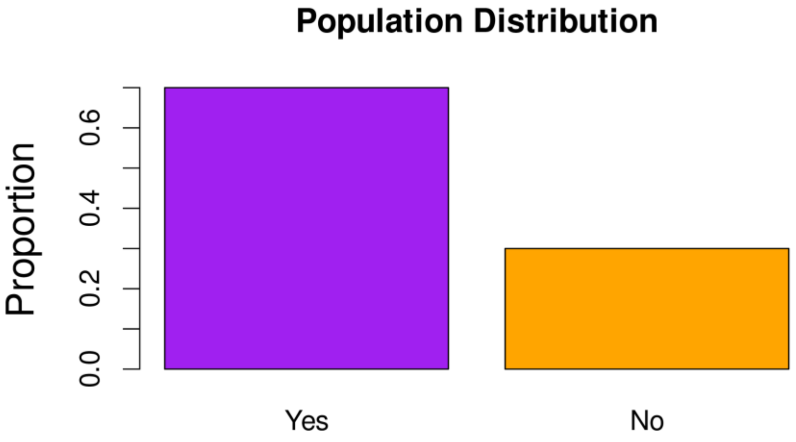
$\hat{p} = 0.75$



Sampling distribution: What values can you expect, if you would repeat an experiment?

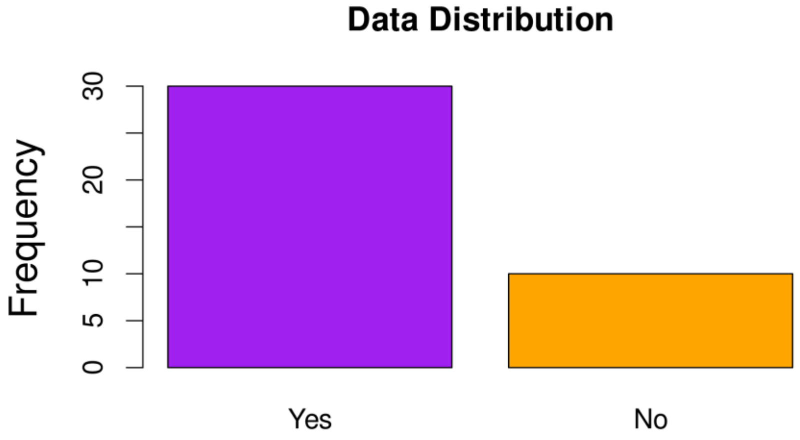
Recap

p = 0.7



Sample (n=40)

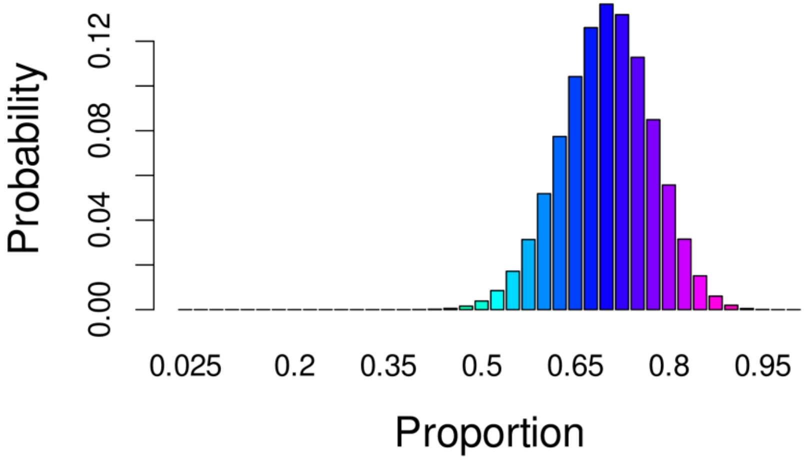
$\hat{p}$ = 0.75



Sampling distribution

Binomial distribution

n=40, p=0.7



Behavior of the **variable**

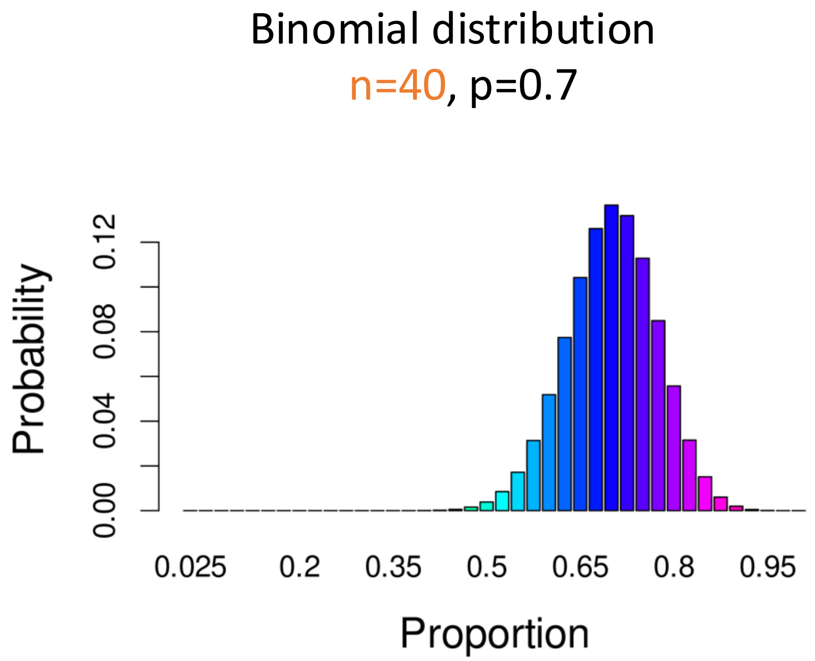
Behavior of the **statistic**

Recap

- The sampling distribution of the statistic tells us something about the *certainty* of our observation

- For instance, what could be likely outcomes if we repeat the experiment?

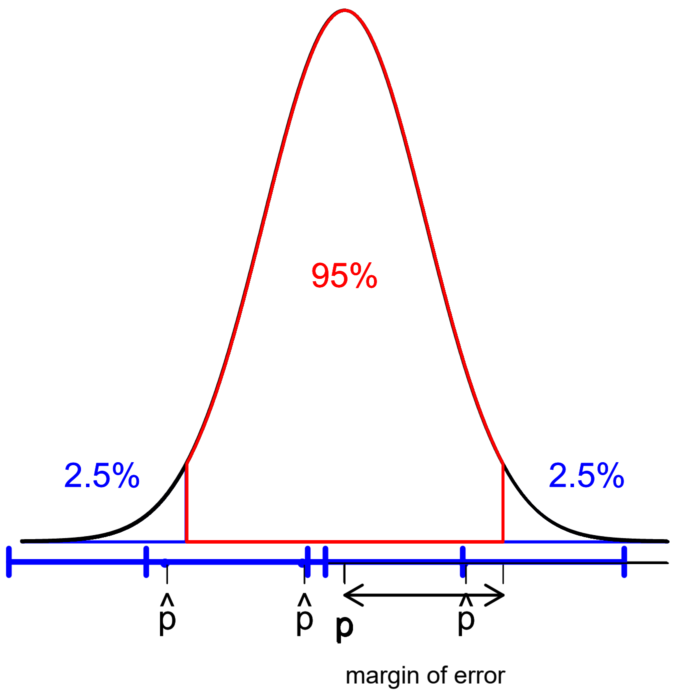
Sampling distribution



Behavior of the **statistic**

Recap

Sampling distribution



$\hat{\mu} = \hat{p}$

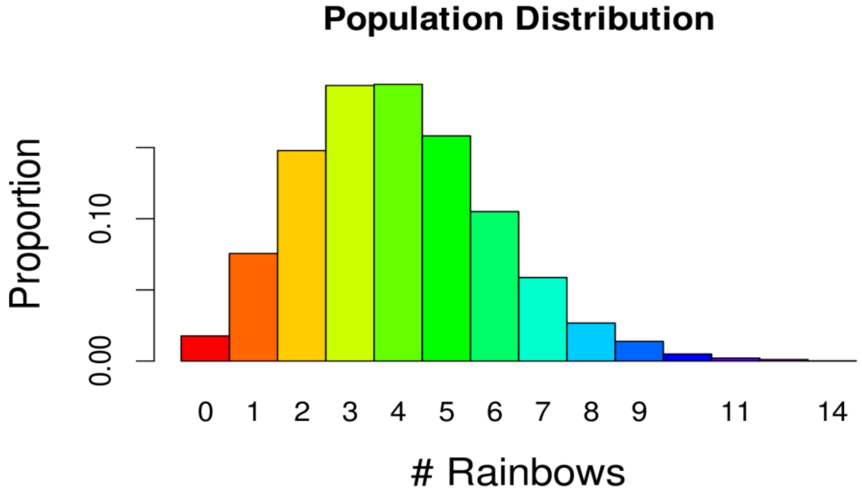
$\hat{\sigma} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

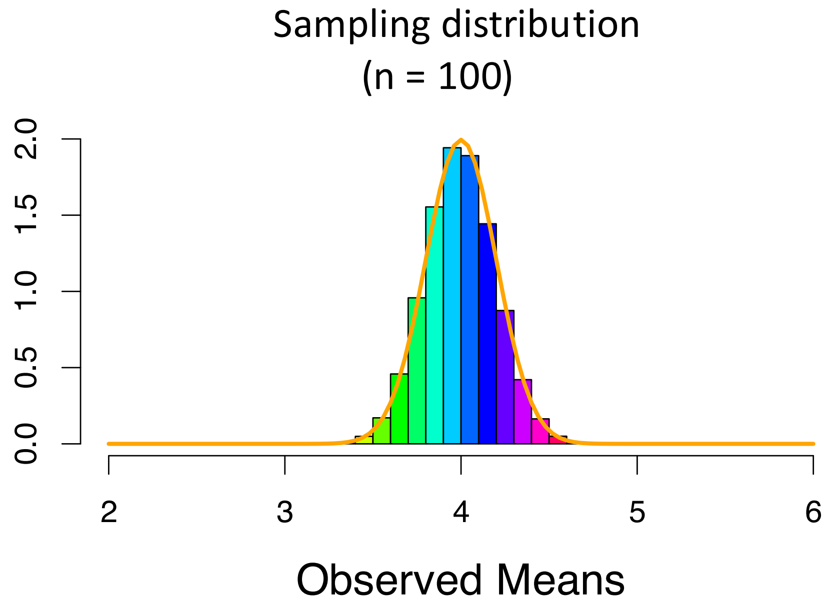
$\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Confidence interval for a proportion: The 95% Confidence Internal for a proportion means that – *in the long run* – the TRUE p is in the interval around the ESTIMATED $\hat{p}$ in 95% of the cases

Recap





- The sampling distribution of x̄:
  1. **IF** we have a population distribution with parameter μ
  2. **IF** we draw a random sample of large enough *n* (because of CLT)
  3. **THEN** the mean of that sample (x̄) is normally distributed with
     - mean μ
     - standard deviation $\frac{\sigma}{\sqrt{n}}$
     - How to get these values in practice?

Recap

1. We are trying to estimate a point (i.e., population mean) based on a sample and its statistics (i.e., sample mean/sd)
2. The sampling distribution tells us how reliable our estimate is (i.e., the sd of the sampling distribution = the standard error)
3. Because we do not know the mean and sd of the sampling distribution, we approximate this using the sample statistics

Today: How to get these values in practice?

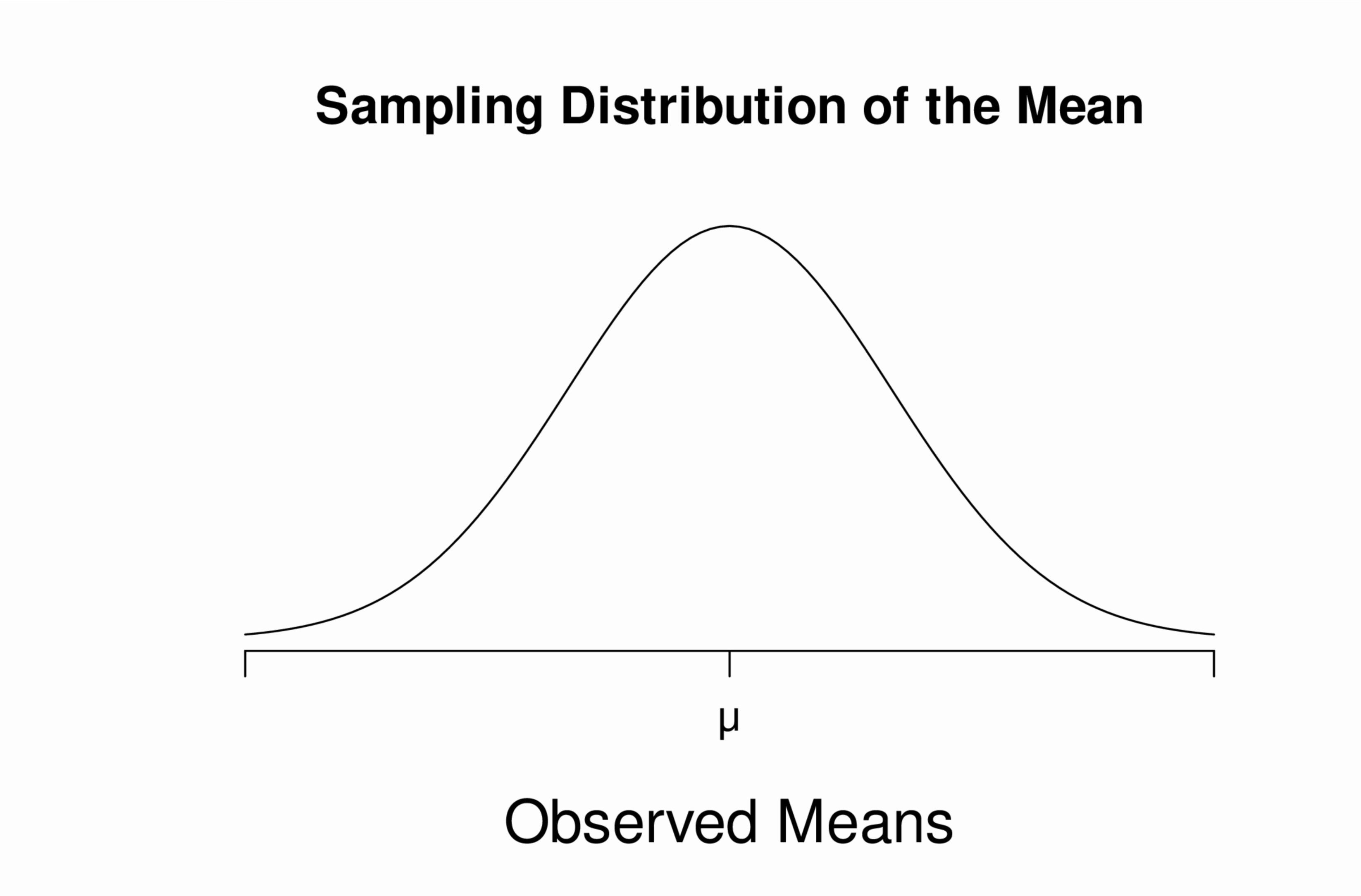
Confidence interval for the **mean**

In other words, we want to know the mean and standard deviation of the sampling distribution of the observed statistic

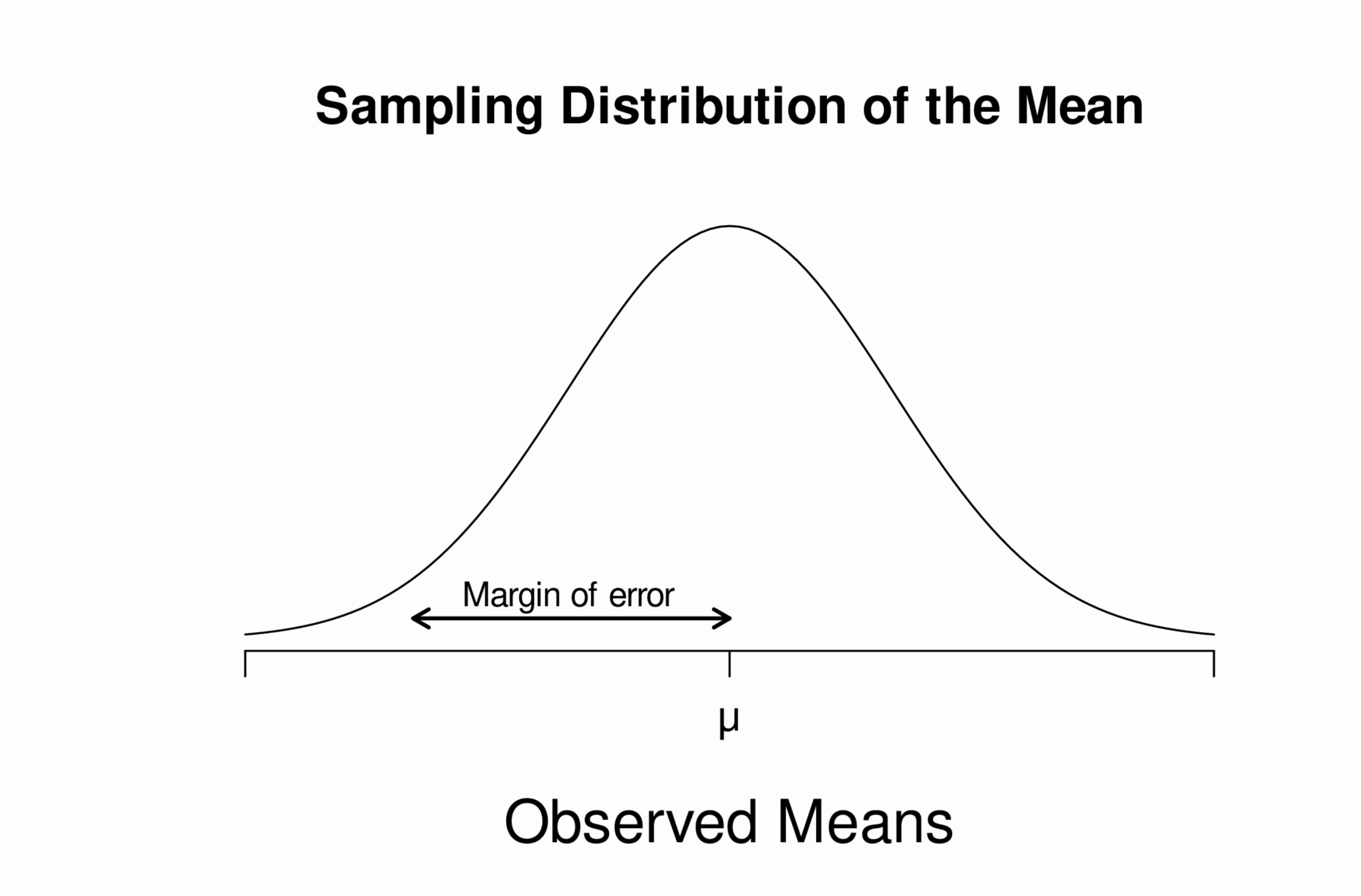
Overview of Today

1. Recap of last Thursday and Tuesday
2. **Confidence interval for the mean**
    - *t*-distribution
3. Robustness against extreme values
4. Choosing a sample size
5. Recap
    - Example exam question

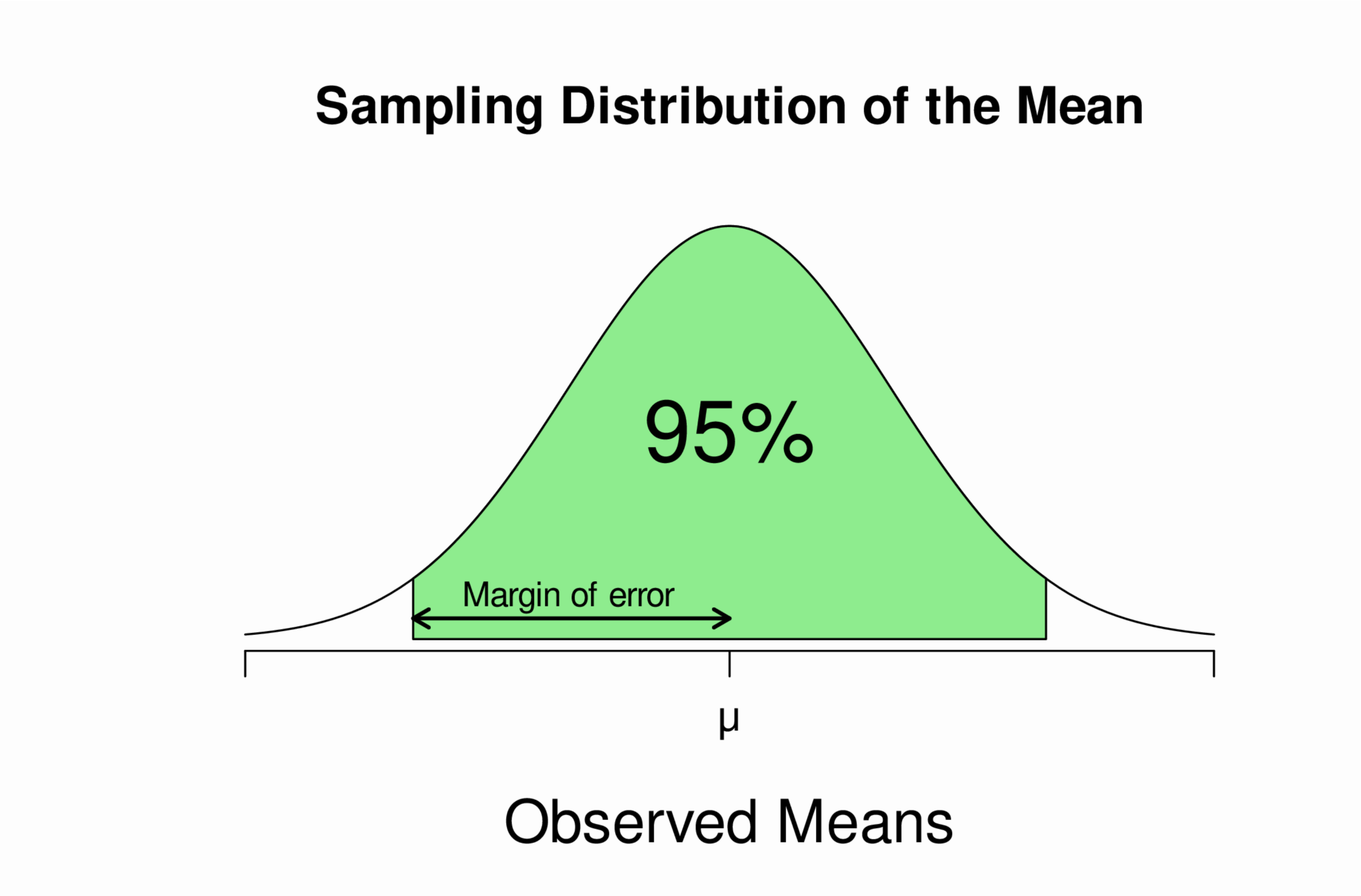
- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error*?



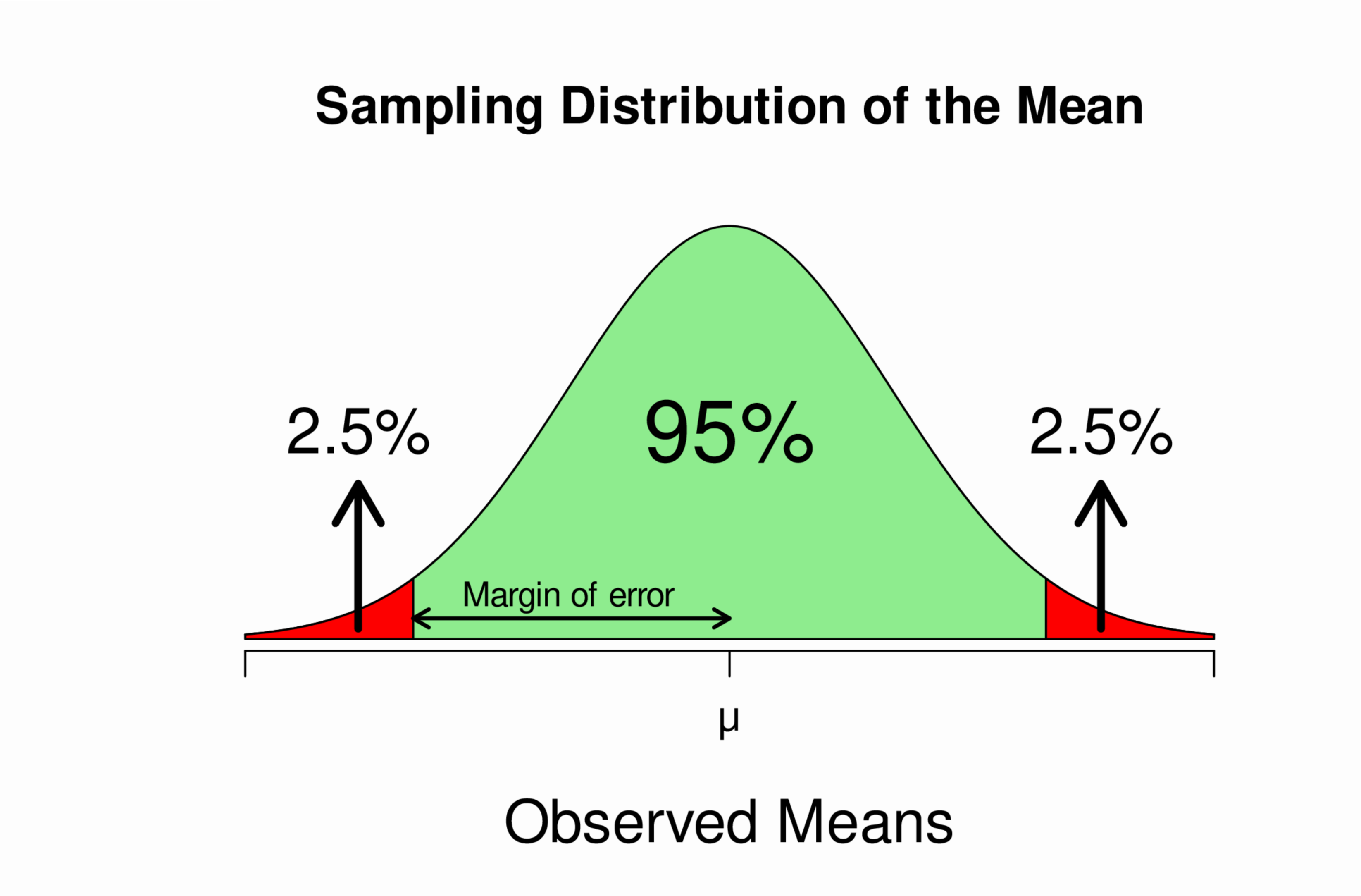
- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error*?



- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error*?



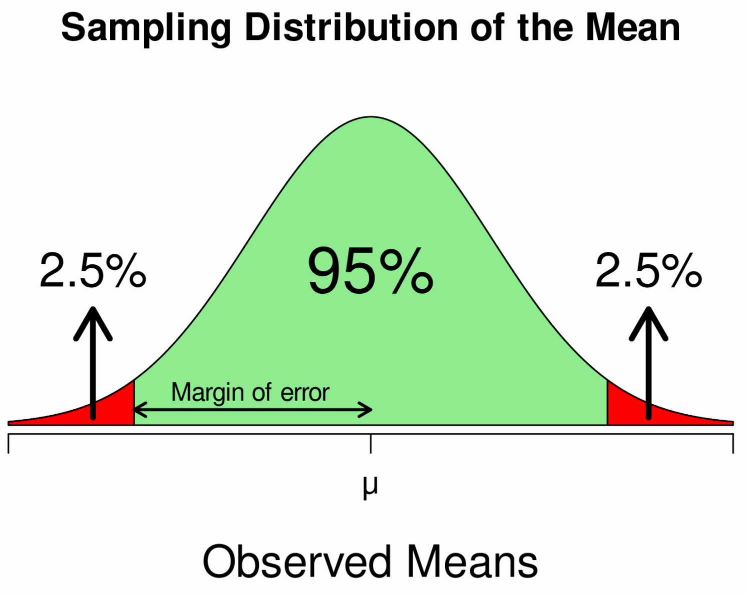
- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error*?



- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error*?

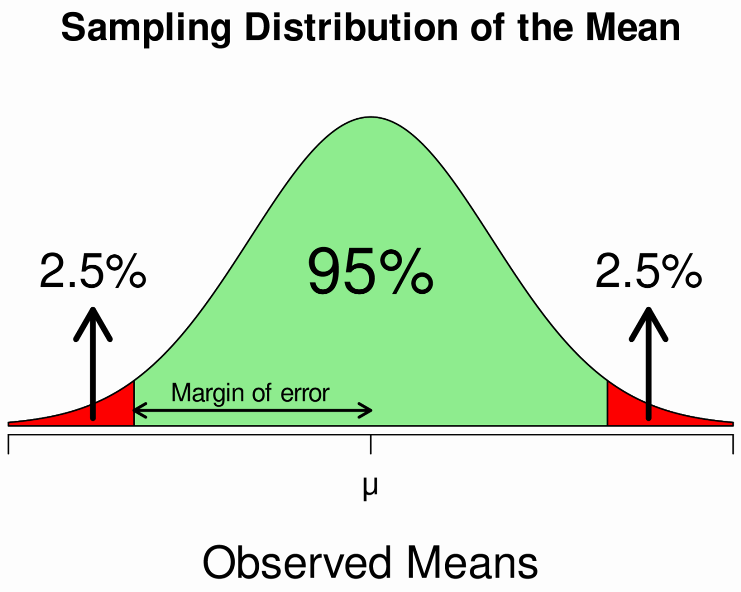
95%

Mostly what we discussed last week

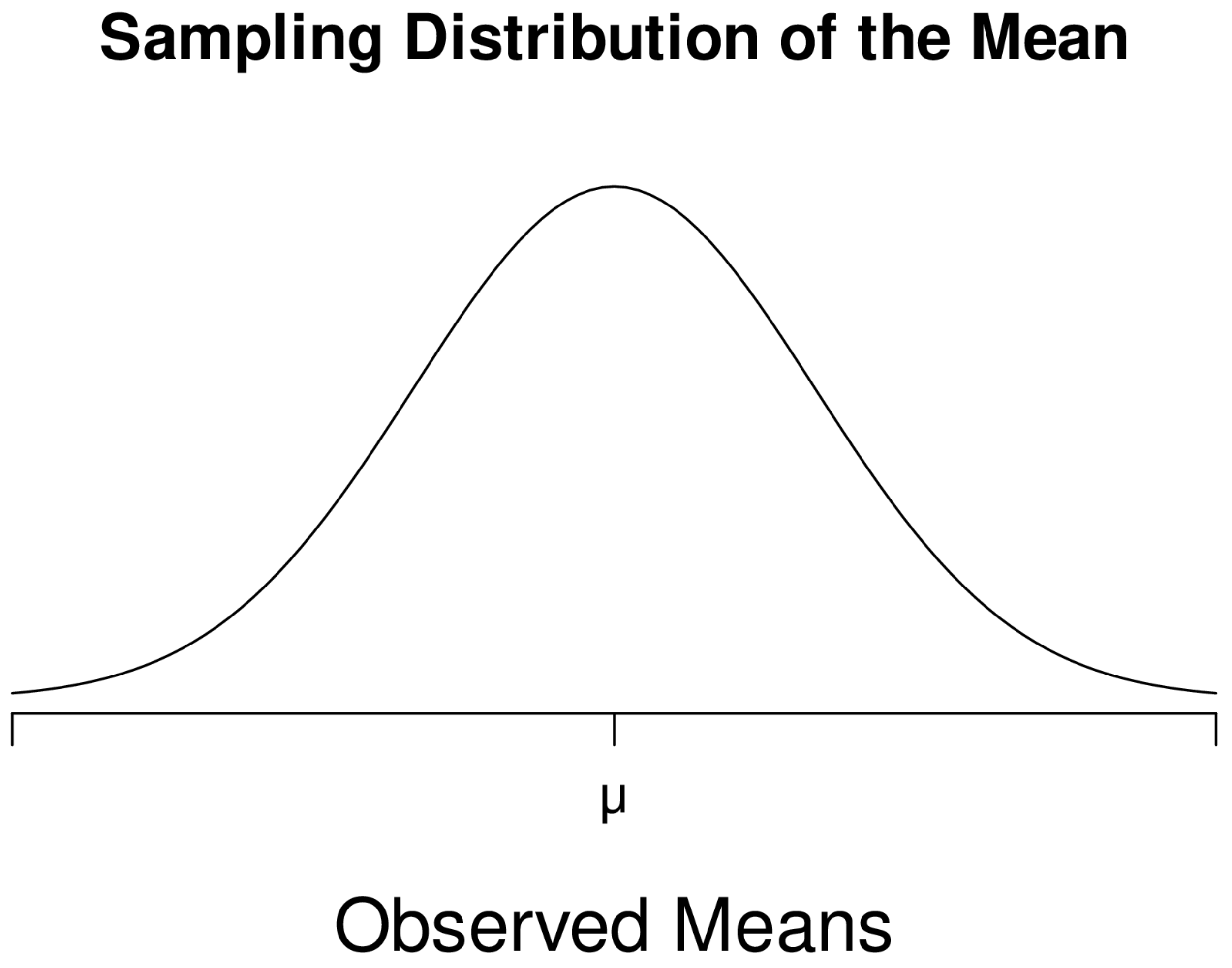


- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error* ?
  - 95%

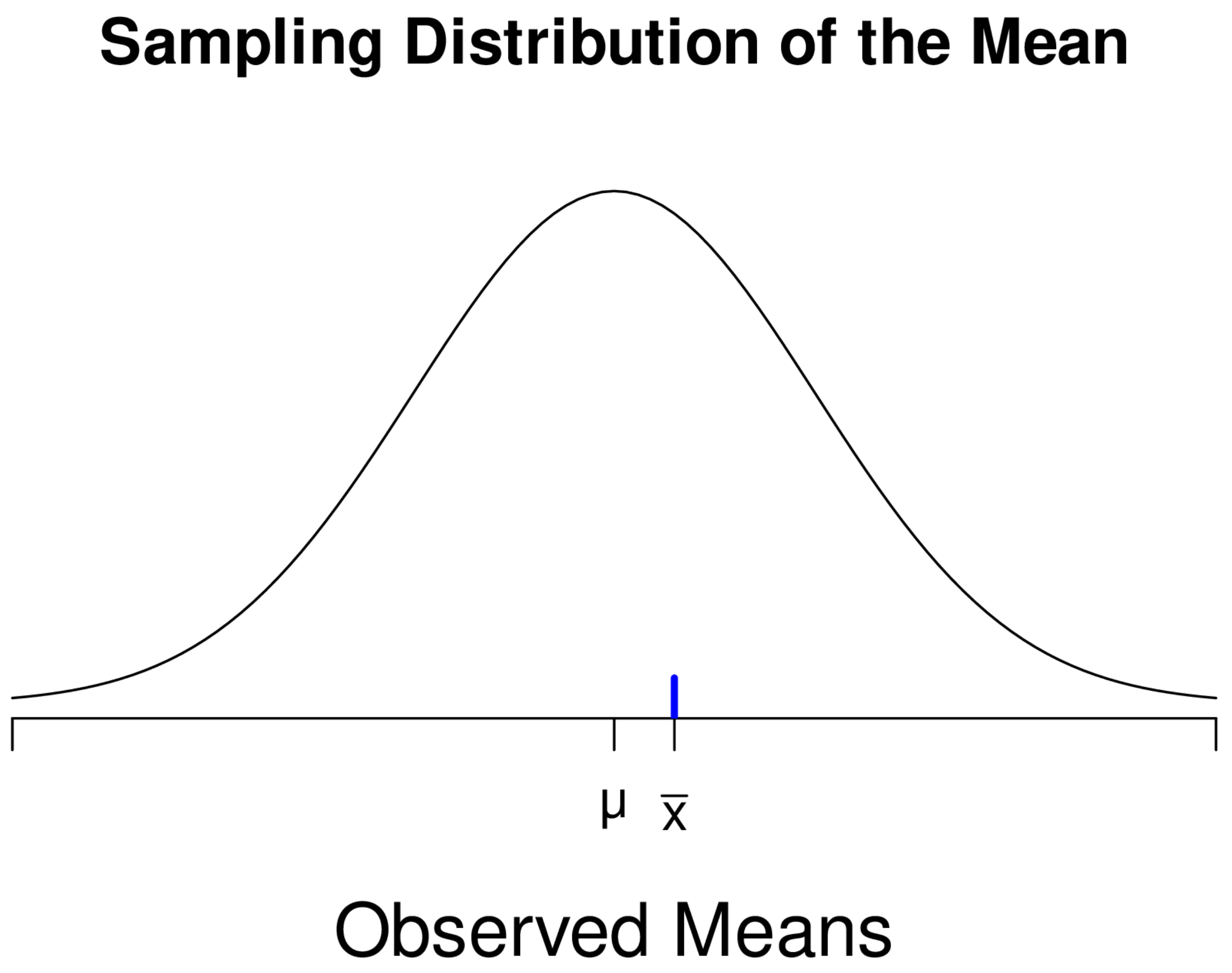
- Question 2: If we draw keep drawing samples and <u>compute</u> $\bar{X}$ - *margin of error* and $\bar{X}$ + *margin of error* each time, what percentage of intervals contain μ?



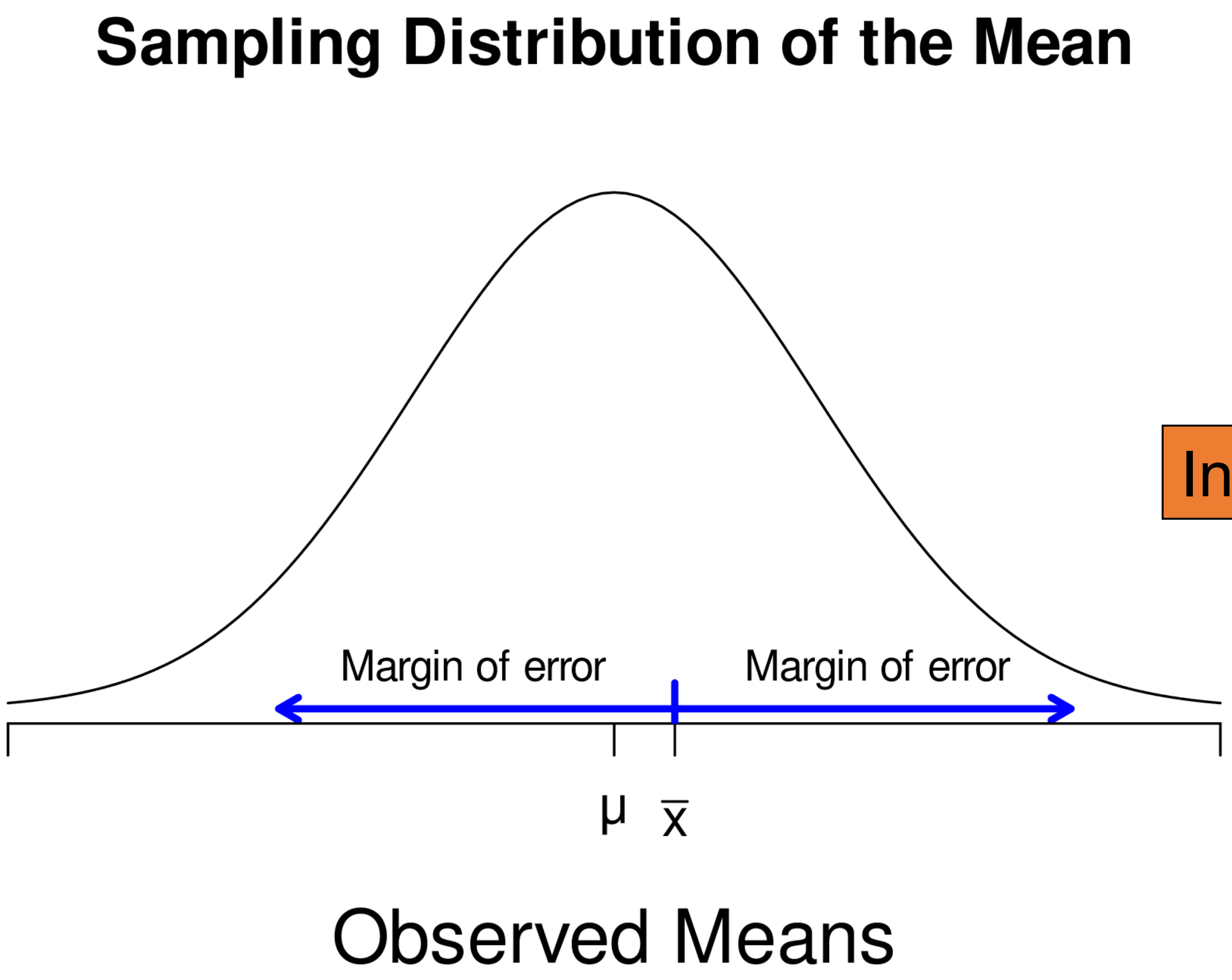
Confidence interval for the mean



Confidence interval for the mean

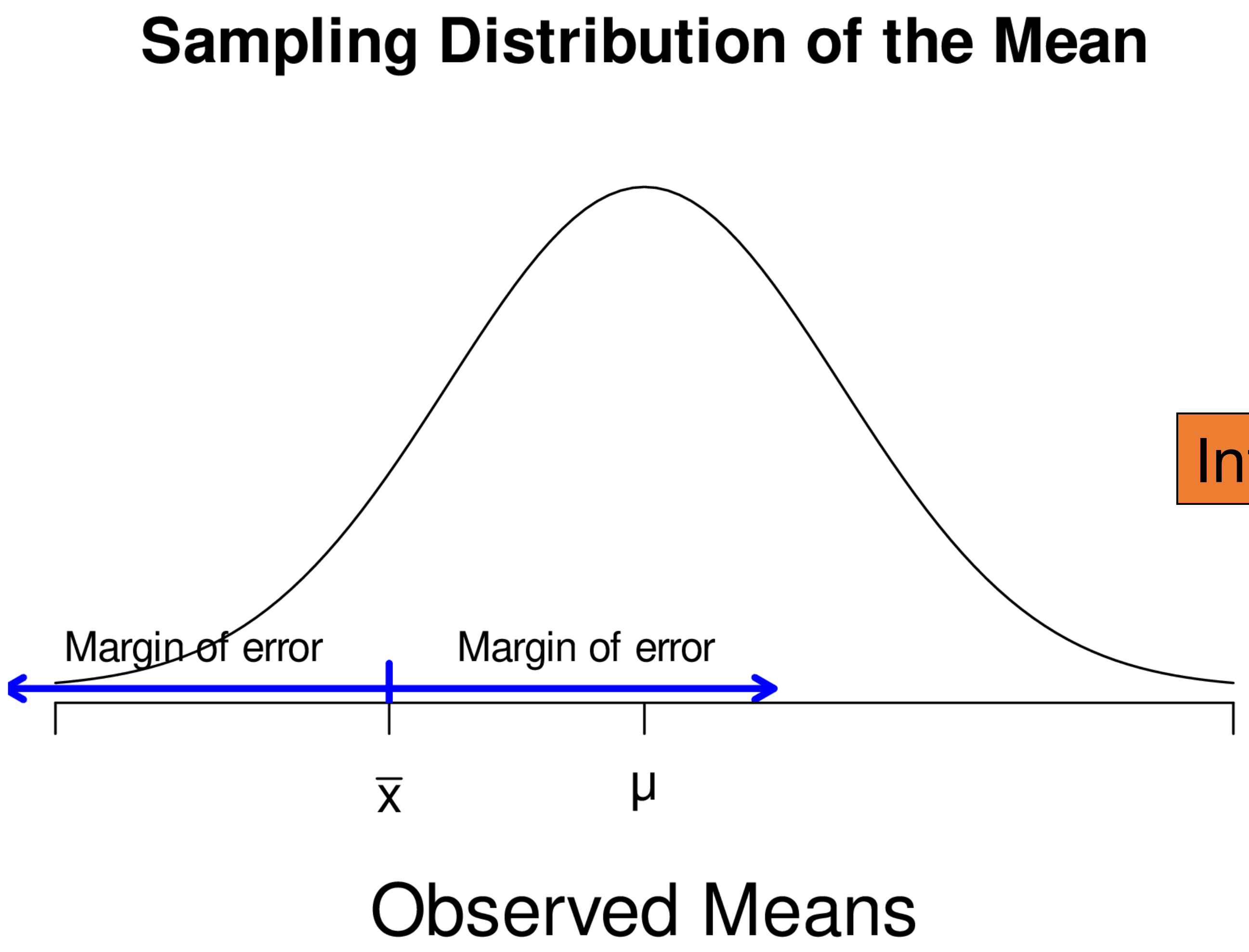


Confidence interval for the mean



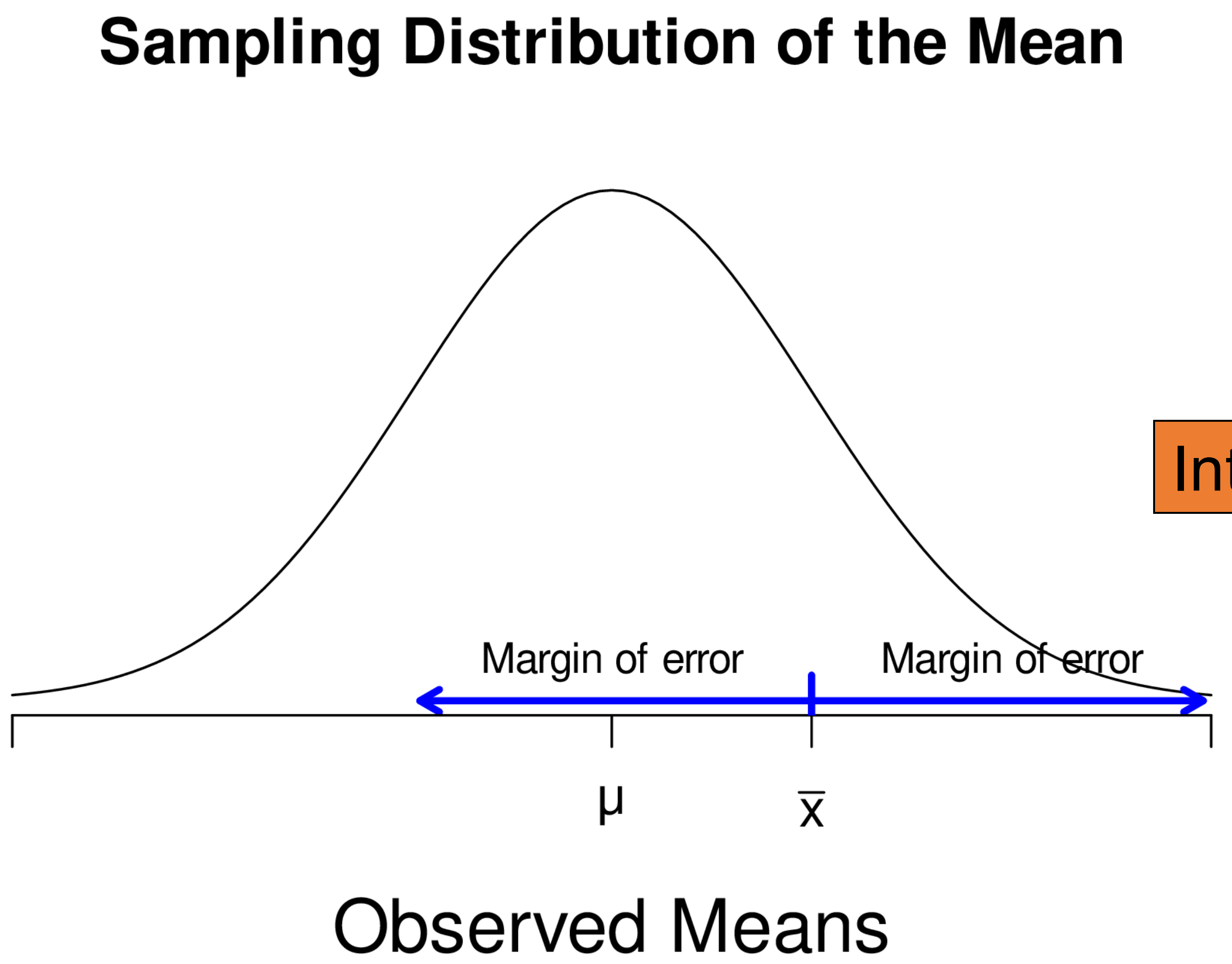
Interval contains μ

Confidence interval for the mean



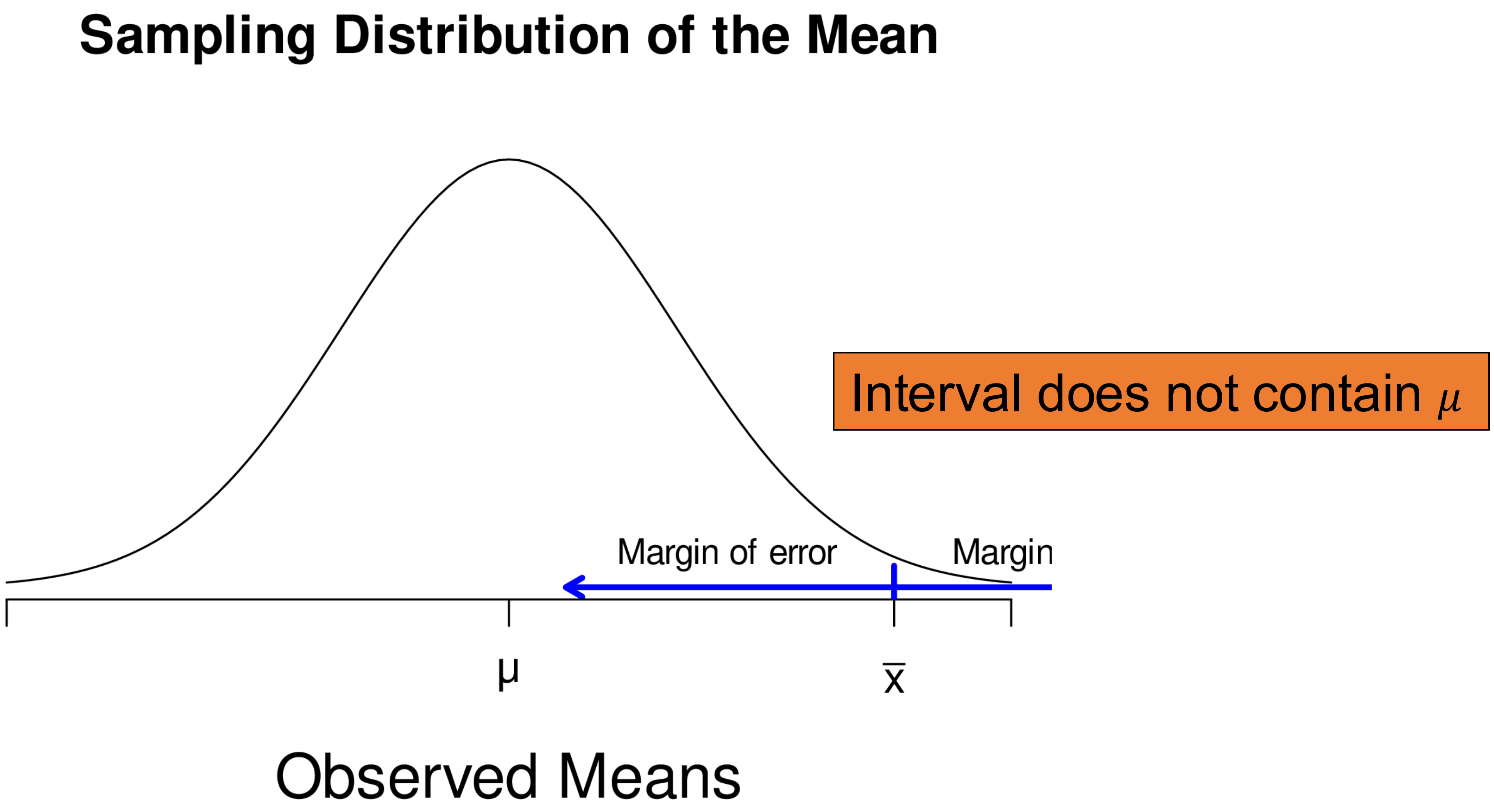
Interval contains μ

Confidence interval for the mean



Interval contains μ

Confidence interval for the mean



Questions

- Question 1: What is the probability that a sample mean $\bar{X}$ is between μ - *margin of error* and μ + *margin of error* ?

  95%

Reasoning starting with μ

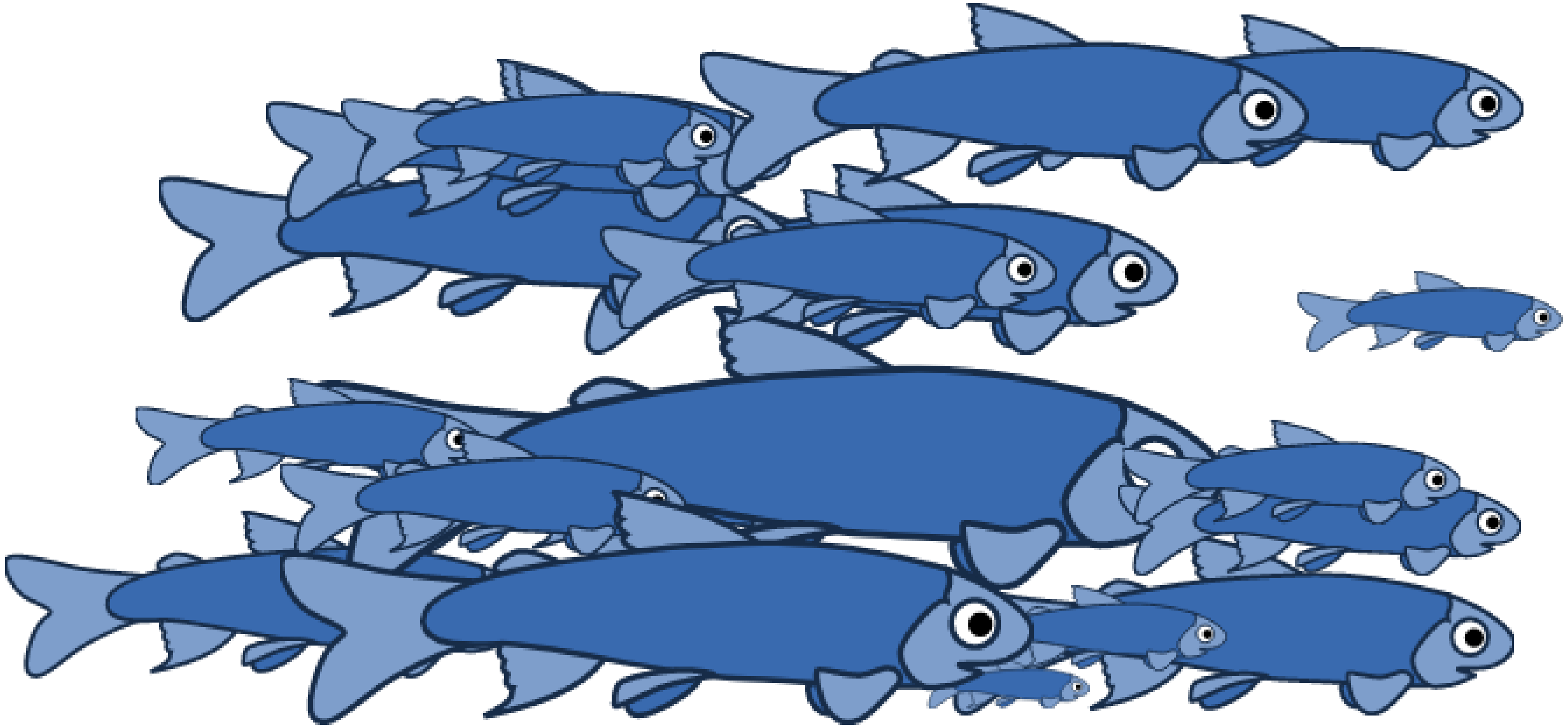
- Question 2: If we draw keep drawing samples and <u>compute</u> $\bar{X}$ - *margin of error* and $\bar{X}$ + *margin of error* each time, what percentage of intervals contain μ?

  95%

Reasoning starting with $\bar{x}$

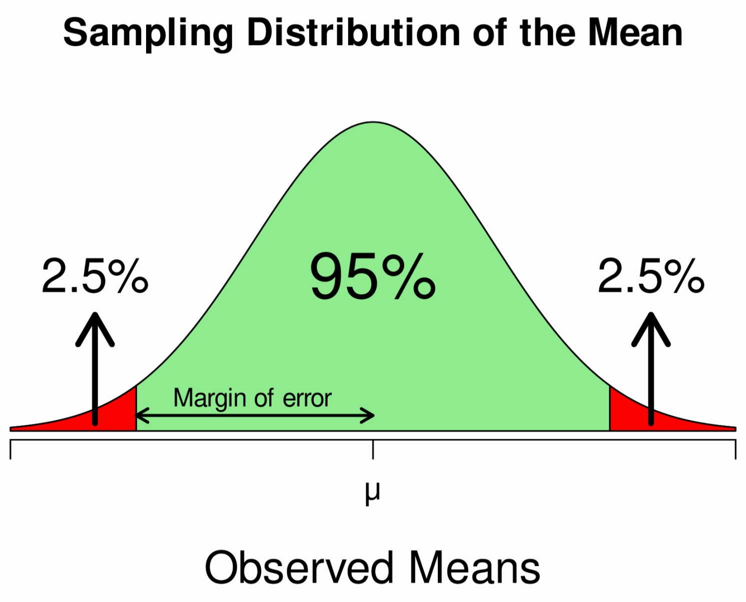
↓

Typically the case



https://www.zoology.ubc.ca/~whitlock/Kingfisher/CIMean.htm

How do we choose margin of error?



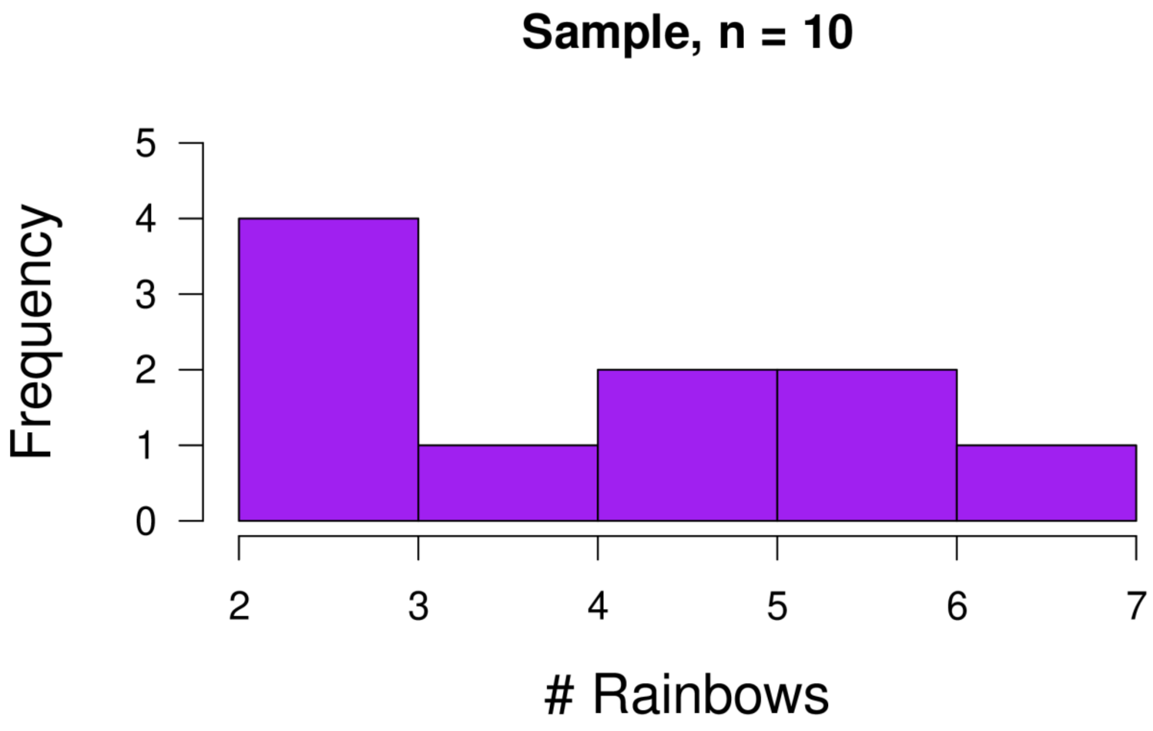
- 1.96 standard deviations above the mean μ
  - Because of 95% probability
- Standard deviation of sampling distribution is $\dfrac{\sigma}{\sqrt{n}}$
- We need to know σ
- Solution: estimate σ with *s*
  - Is this good enough?

Standard error: The *estimated* standard deviation of a sampling distribution.

Let's try it out

- The 95% CI interval based on <u>one</u> sample is:

$$\bar{x} - 1.96 * \frac{s}{\sqrt{n}} \text{ and } \bar{x} + 1.96 * \frac{s}{\sqrt{n}}$$



$$\bar{x} = 4.4$$

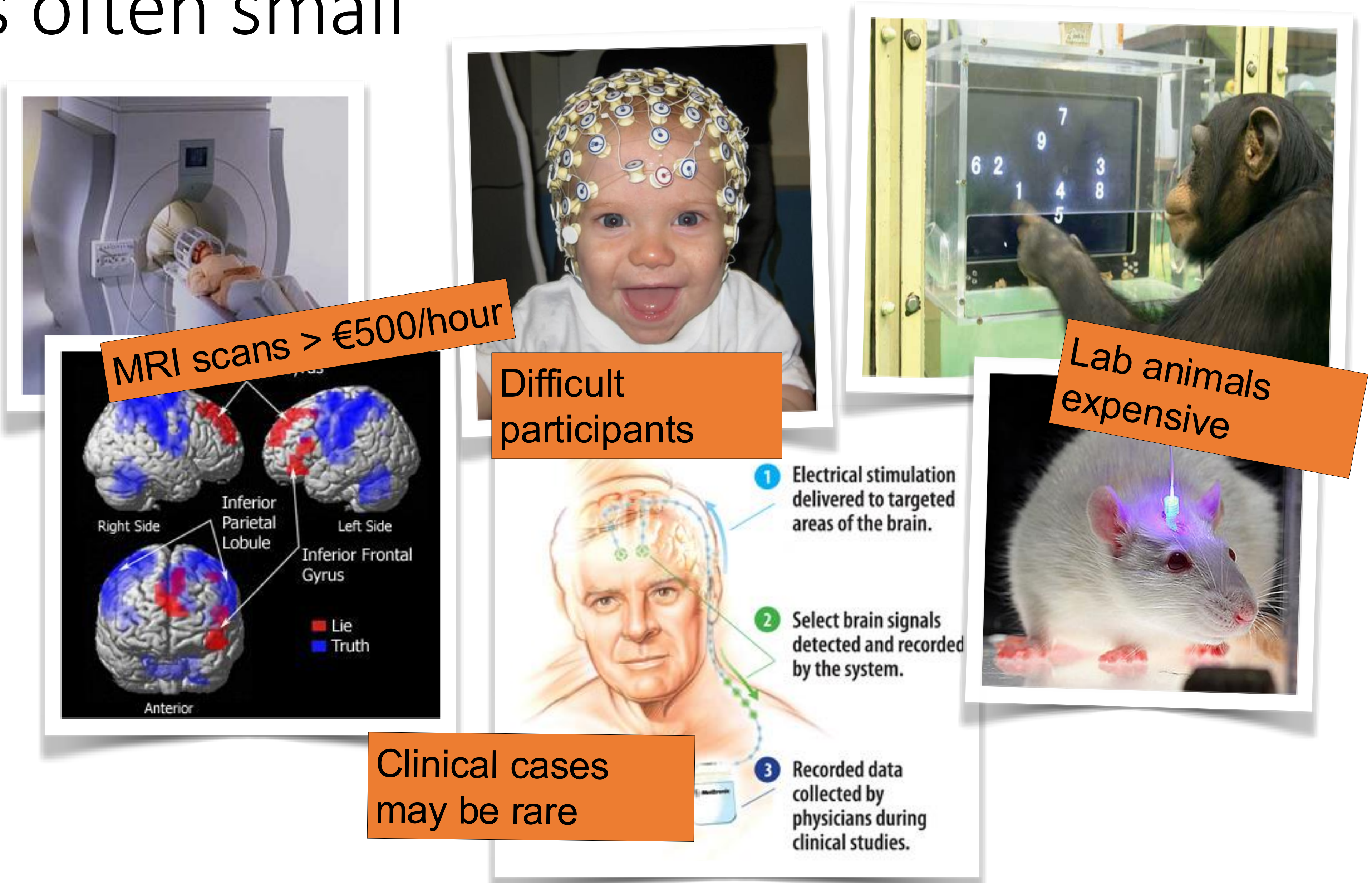
$$s = 1.65$$

$$\bar{x} - 1.96 \frac{s}{\sqrt{n}} = 3.38$$

$$\bar{x} + 1.96 \frac{s}{\sqrt{n}} = 5.42$$

- However, *s* is a noisy estimate of σ, especially for small *n*

n is often small



Let’s try it out

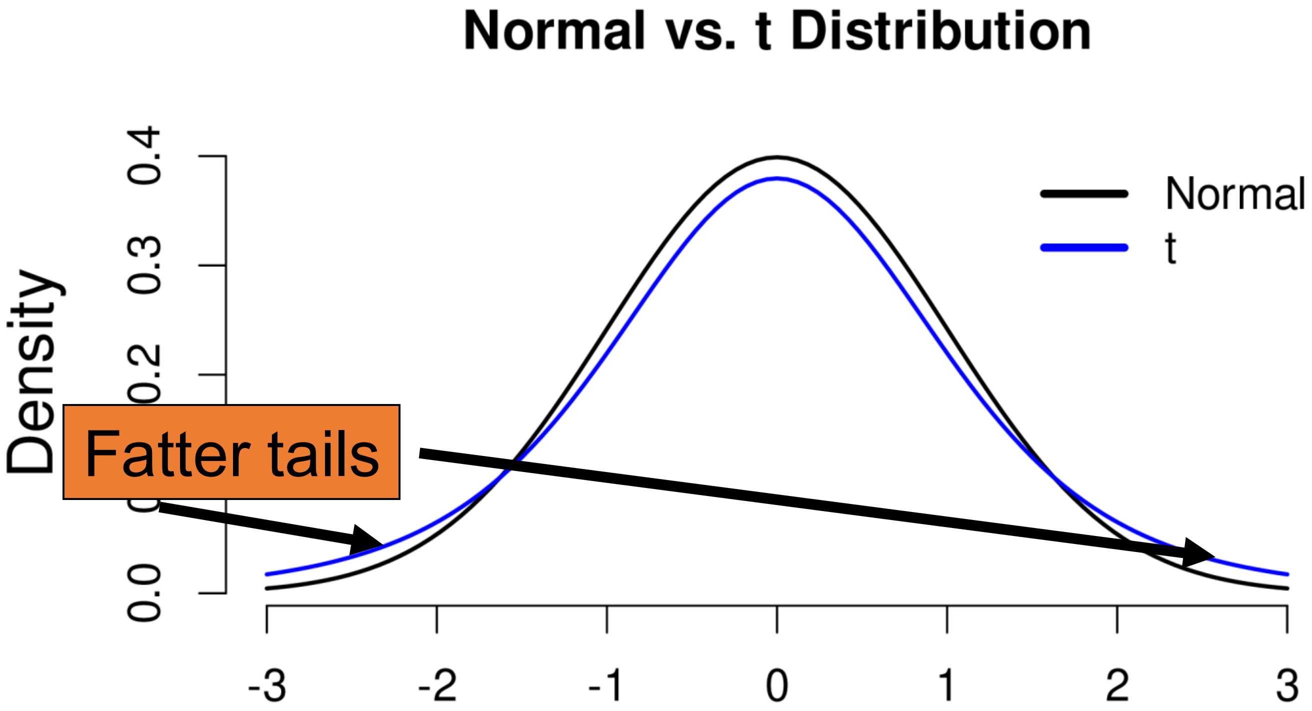
- However, s is a noisy estimate of σ, especially for small n

In order to make sure we do not underestimate σ (and therefore the standard error) and grow “overconfident”, we can use a different distribution than the normal distribution:
The *t*-distribution!

The *t*-distribution

- Fatter tails lead to ***larger*** confidence intervals

GUINNESS



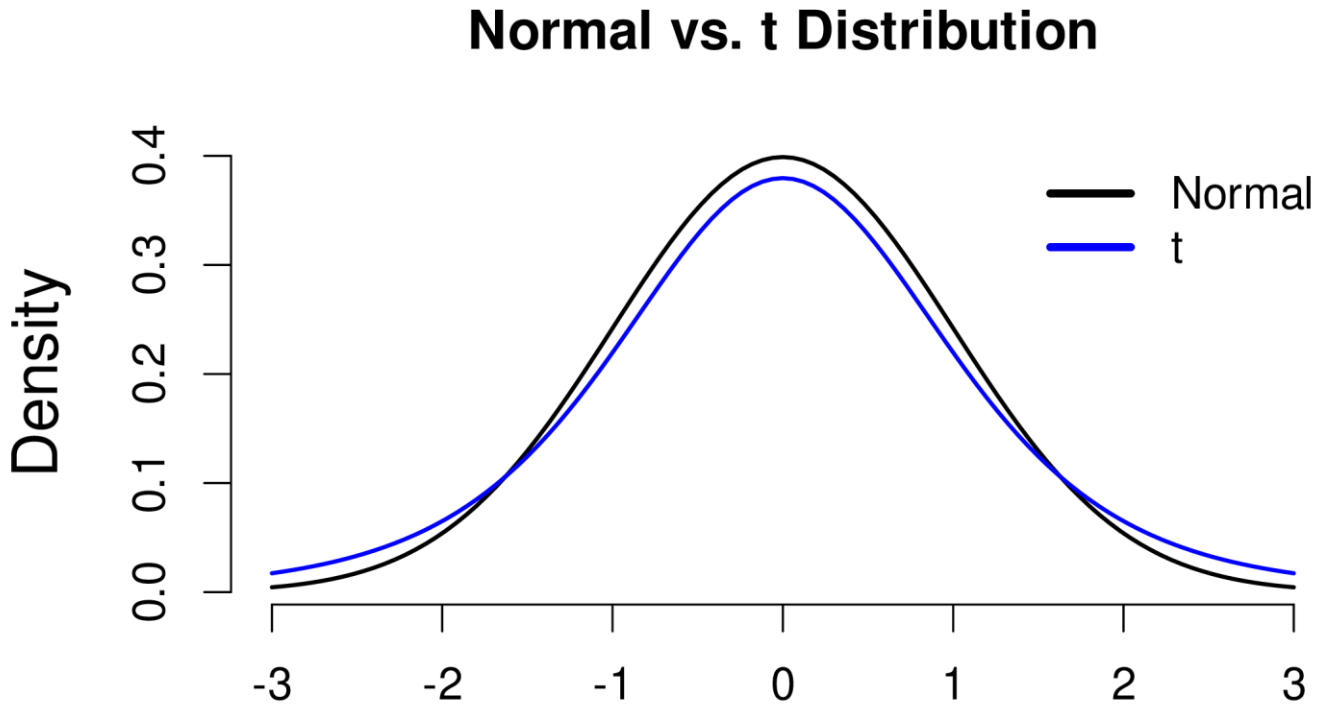
Source: https://en.wikipedia.org/wiki/William_Sealy_Gosset

The *t*-distribution approximates the normal distribution

- Black: Normal
- Blue: *t*

n = 6
***df* = n-1 = 5**

df: degrees of freedom



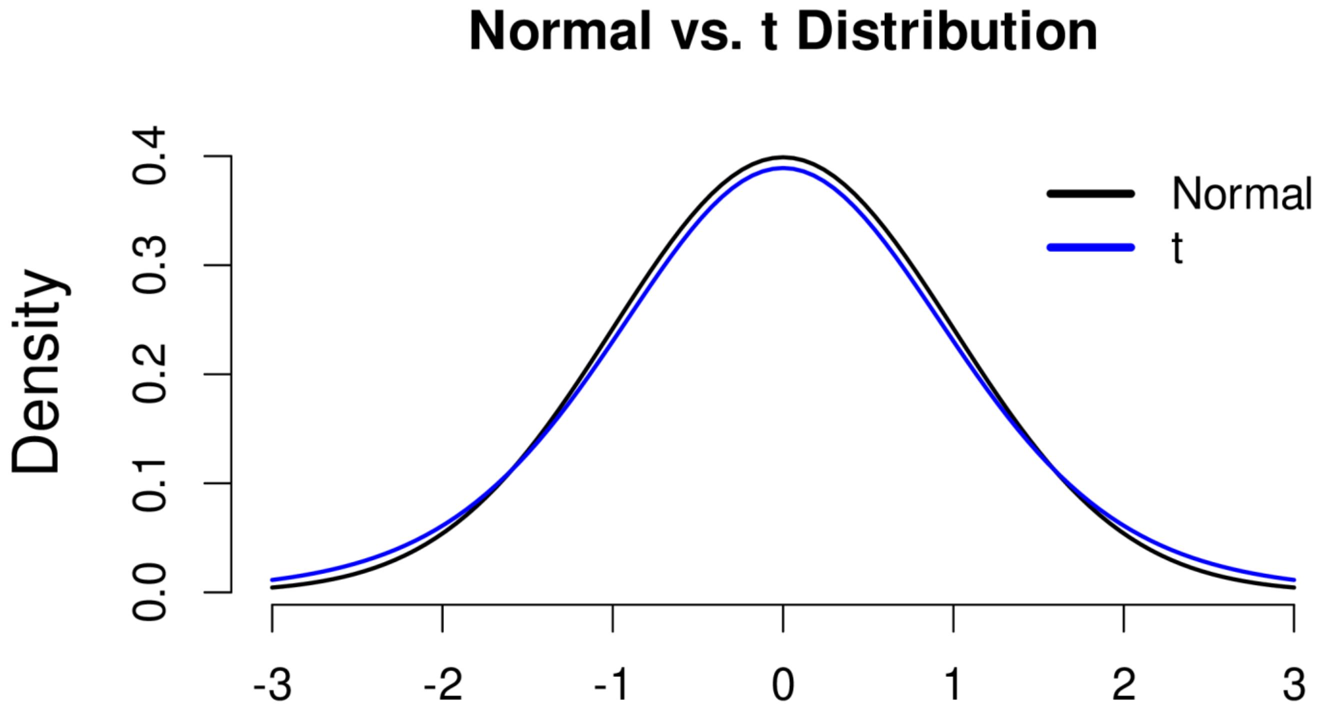
The *t*-distribution approximates the normal distribution

- Black: Normal
- Blue: *t*

n = 11

***df* = n-1 = 10**

df: degrees of freedom



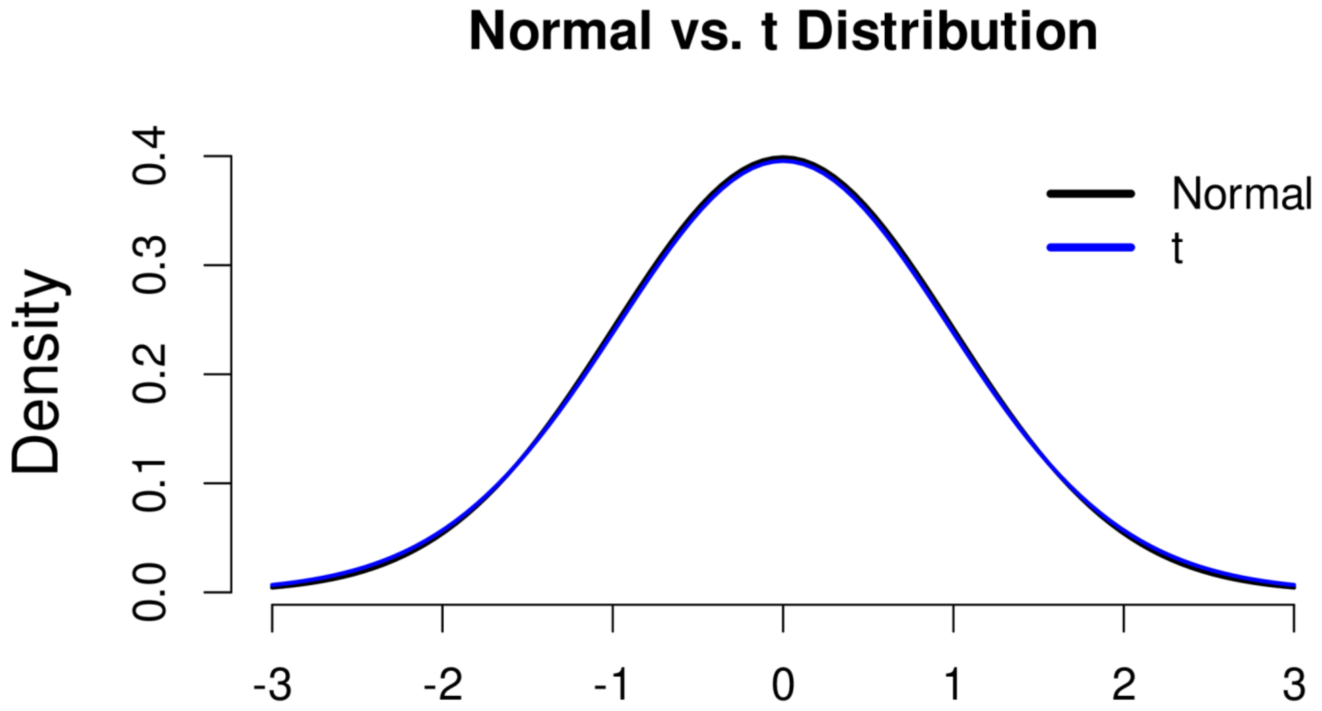
The *t*-distribution approximates the normal distribution

- Black: Normal
- Blue: *t*

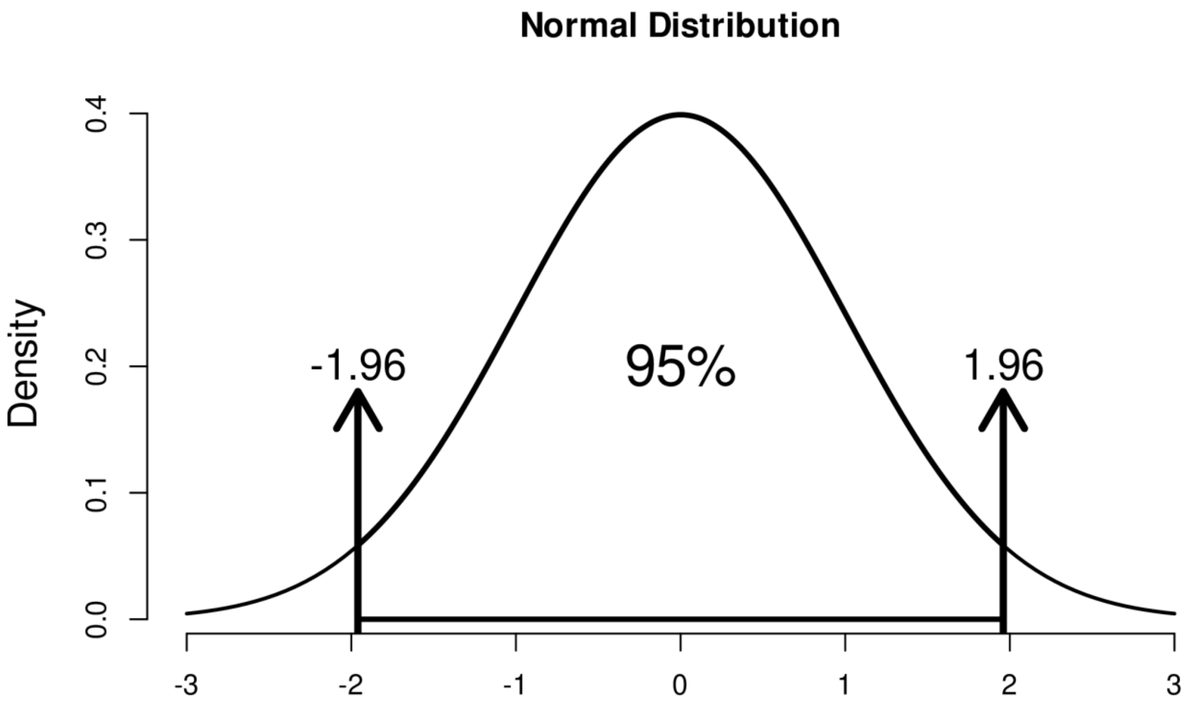
n = 31

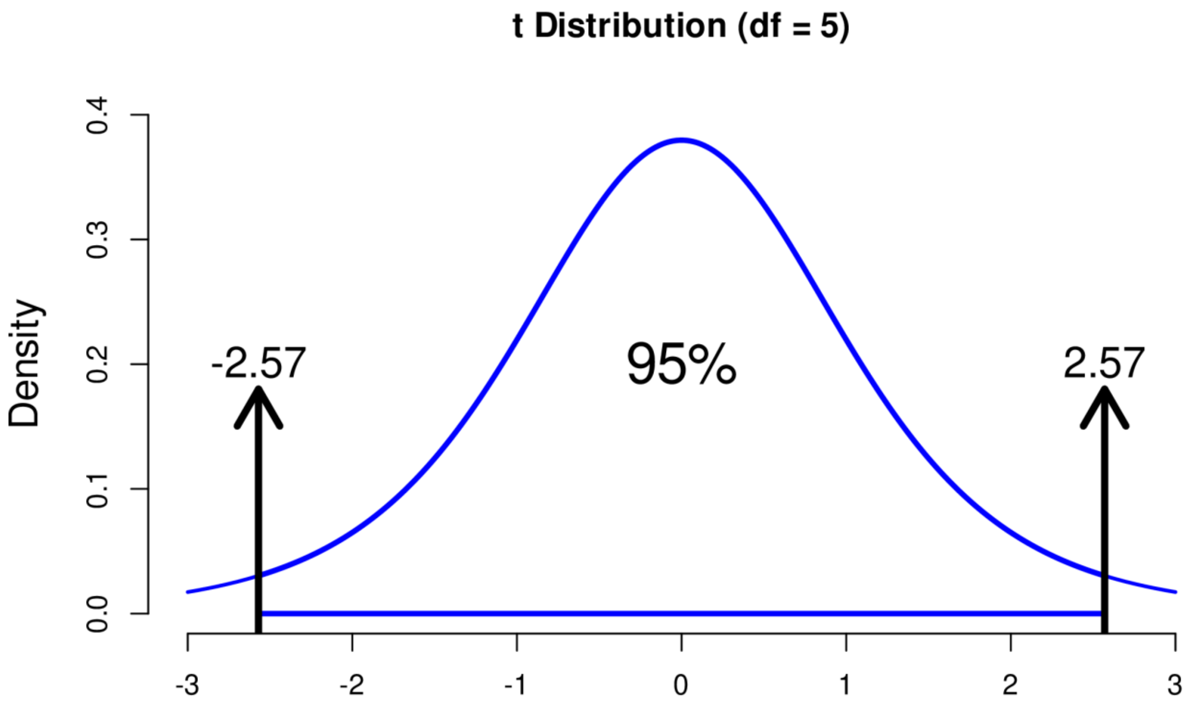
***df* = n-1 = 30**

df: degrees of freedom

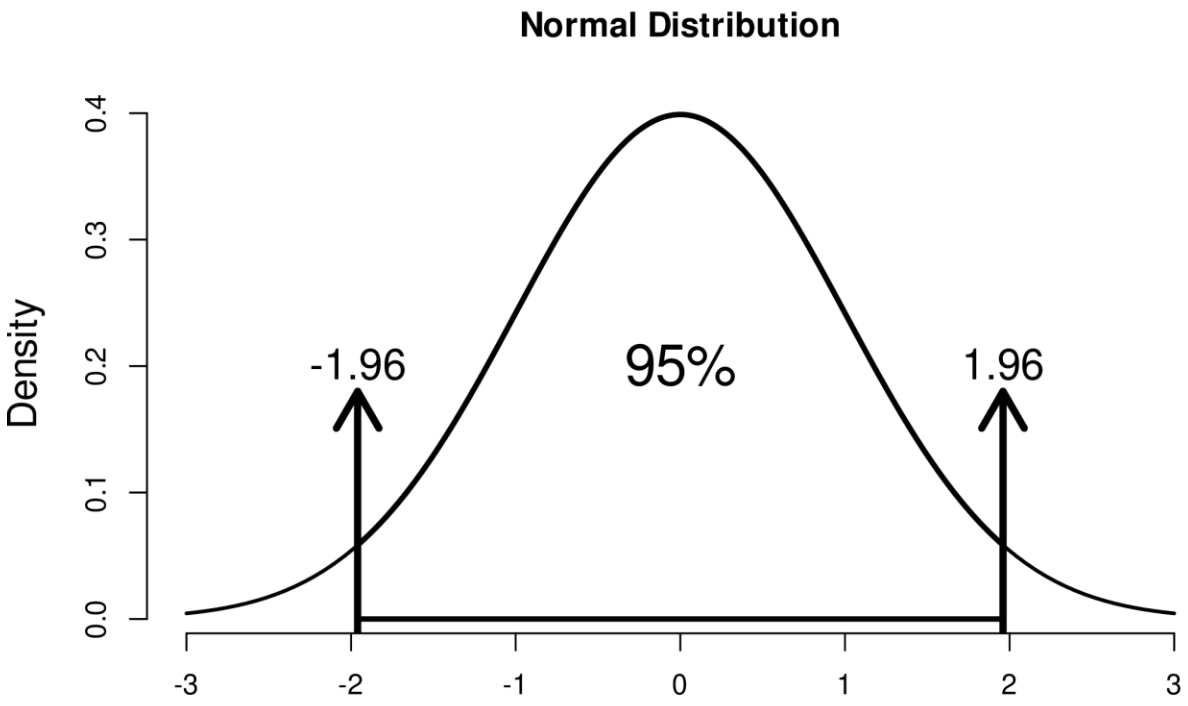


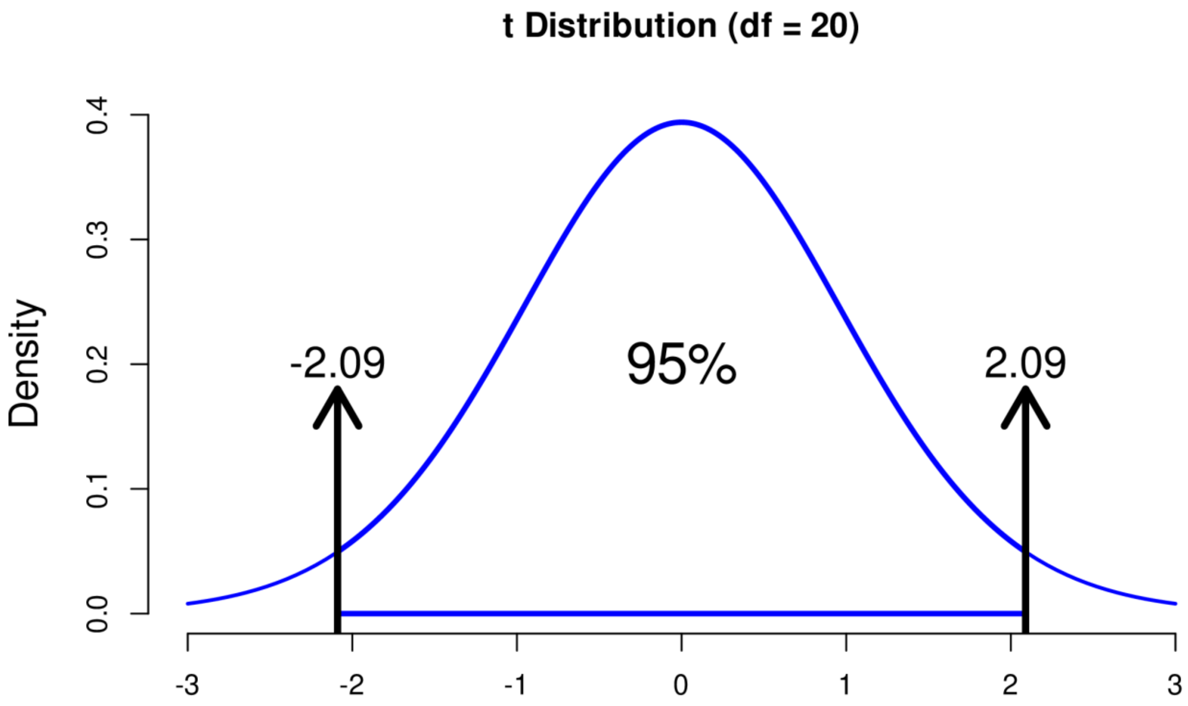
Fatter tails lead to *larger* confidence intervals



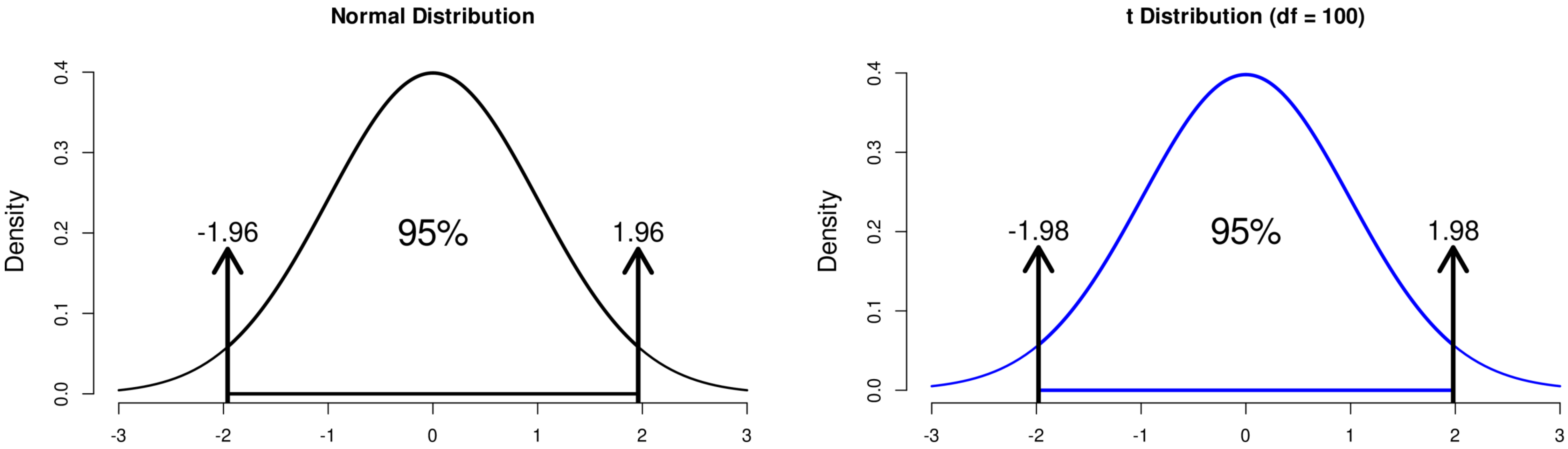


Fatter tails lead to *larger* confidence intervals





Fatter tails lead to *larger* confidence intervals



Confidence interval for the mean

Confidence interval for the mean: A 95% confidence interval for the population mean μ is

$$\bar{x} \pm margin\ of\ error$$

$$\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}}$$

where $t_{.025}$ depends on the sample size *n*.

Example: Estimating the Average Sleep

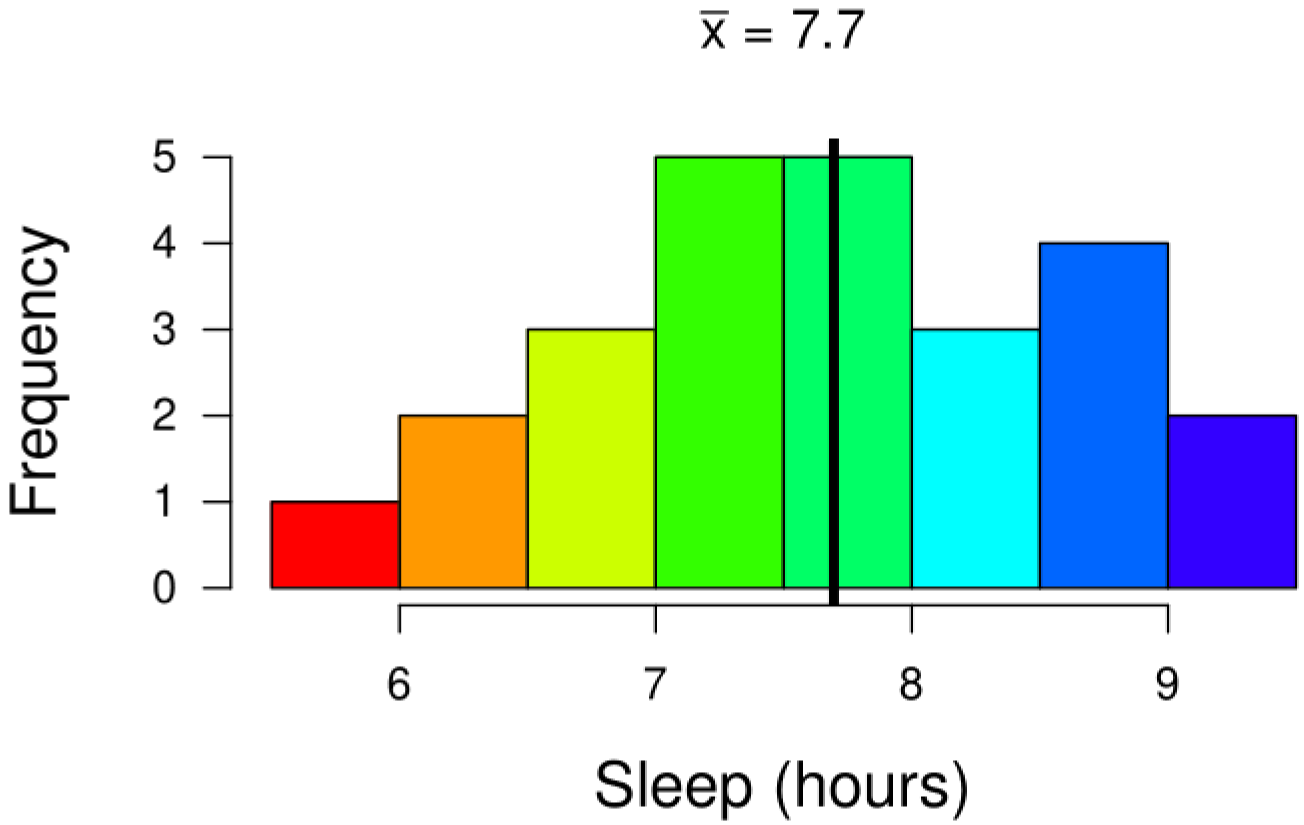
- The article claims a lack of sleep among students
- How can we estimate the average sleep time (in hours) if we take a sample of 25 students?
- How can we compare this to the recommended length of 8 hours?

Hershner SD, Chervin RD. Causes and consequences of sleepiness among college students. *Nature and Science of Sleep*. 2014;6:73-84. Published 2014 Jun 23. doi:10.2147/NSS.S62907

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4075951/pdf/nss-6-073.pdf

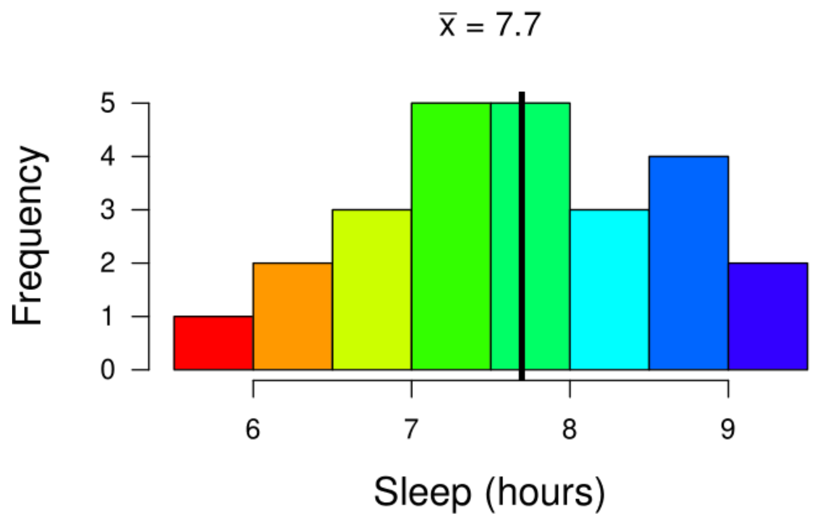
Example: Sleepers

- $n = 25$
- $\bar{x} = 7.7$
- $s = 0.98$



- Question is: What is the population mean of this group?

Example: Sleepers

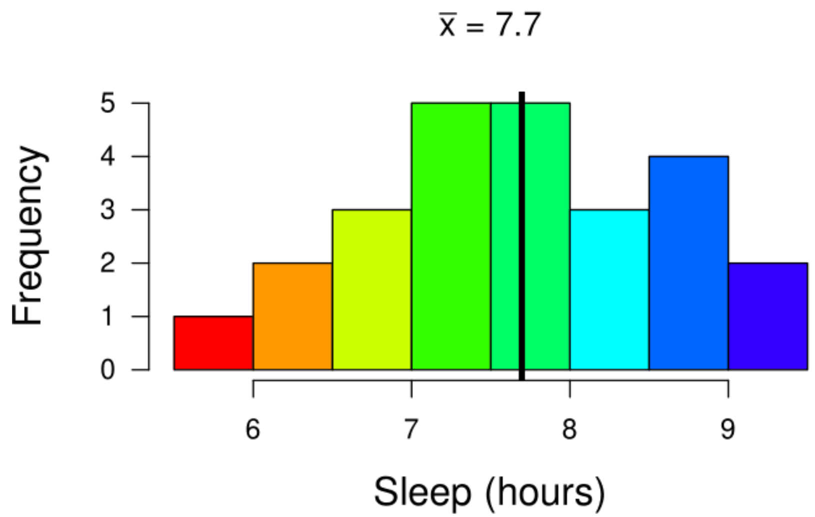


- n = 25
- $\bar{x}$ = 7.7
- s = 0.98

- Estimated population mean:
  - $\bar{x}$ = 7.7
- 95% CI:
  - Since we do not know σ, we have to estimate it, and use the *t*-distribution:

$$\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}}$$

Example: Sleepers

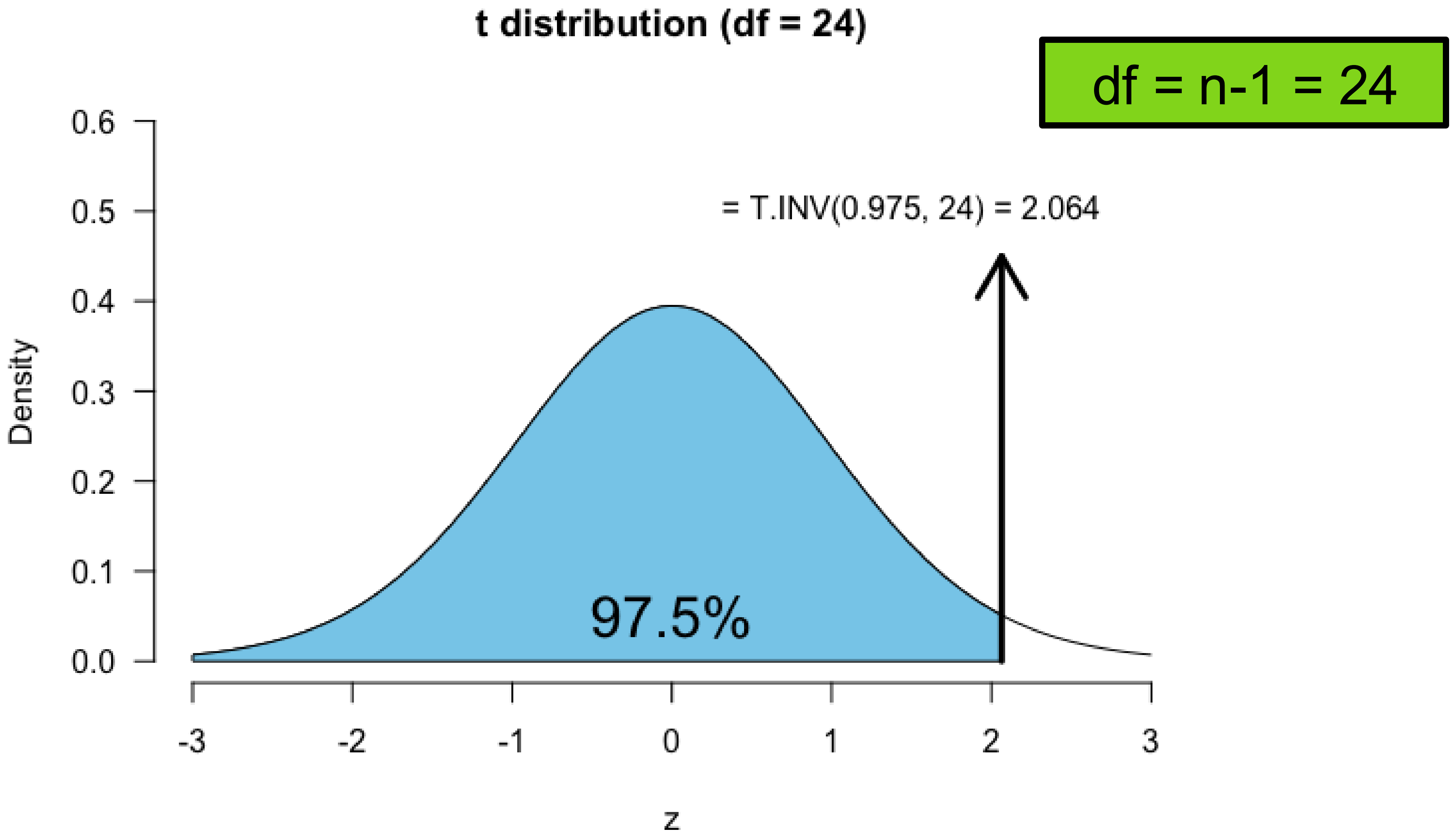


- n = 25
- $\bar{x}$ = 7.7
- s = 0.98

- Estimated population mean:
  - $\bar{x}$ = 7.7
- 95% CI:
  - Since we do not know σ, we have to estimate it, and use the *t*-distribution:

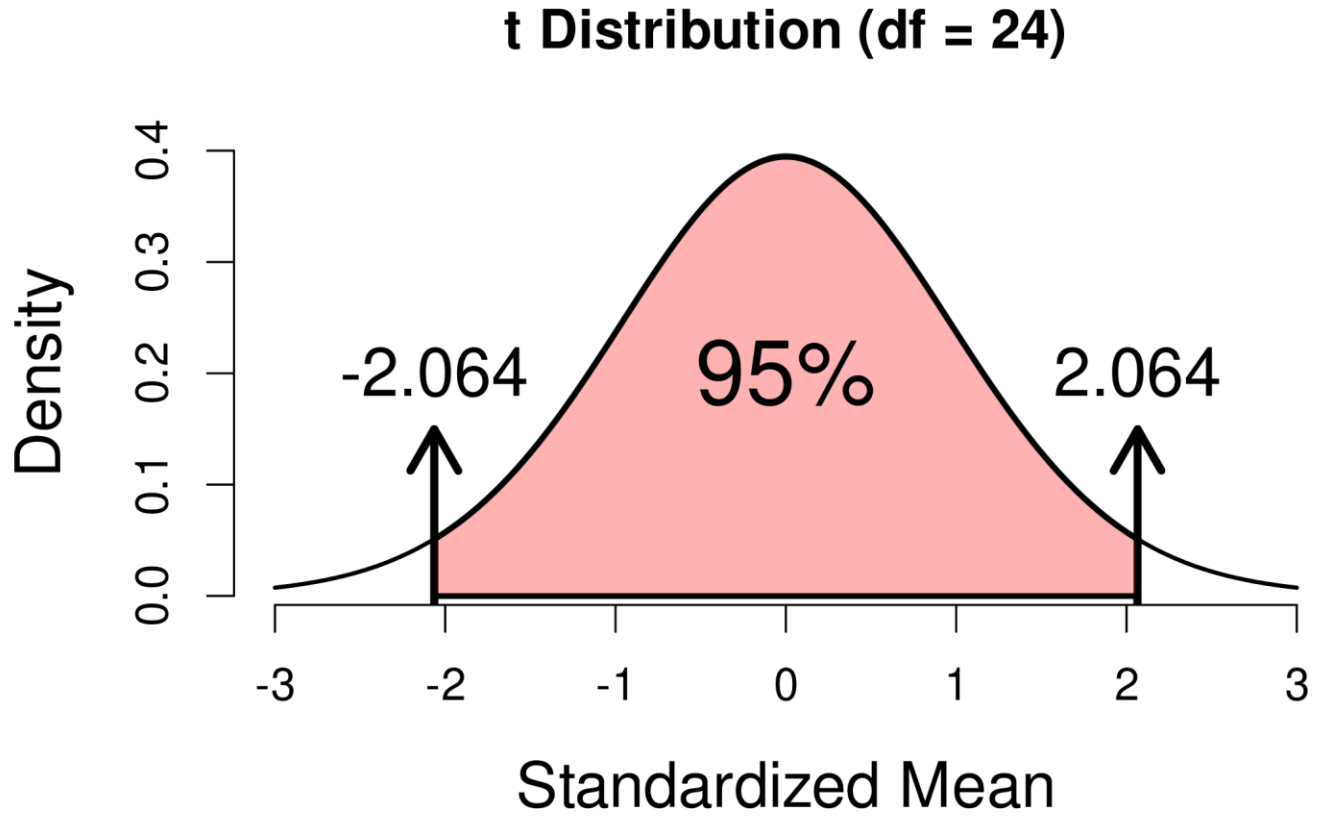
$$\bar{x} \pm t_{.025} \frac{0.98}{\sqrt{25}}$$

Look up $t_{.025}$ in Excel

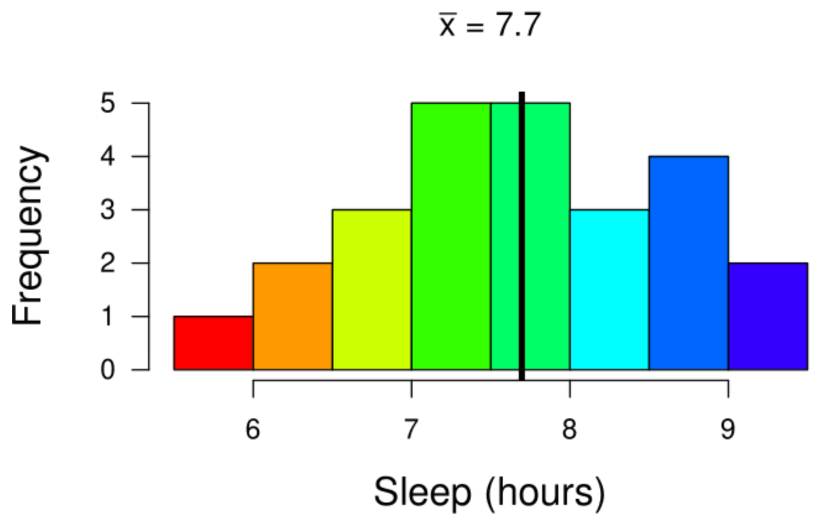


Look up $t_{.025}$ in Excel

df = n-1 = 24



Example: Sleepers



- $n = 25$
- $\bar{x} = 7.7$
- $s = 0.98$

- Estimated population mean:
  - $\bar{x} = 7.7$
- 95% CI:
  - Since we do not know σ, we have to estimate it, and use *t*-distribution:

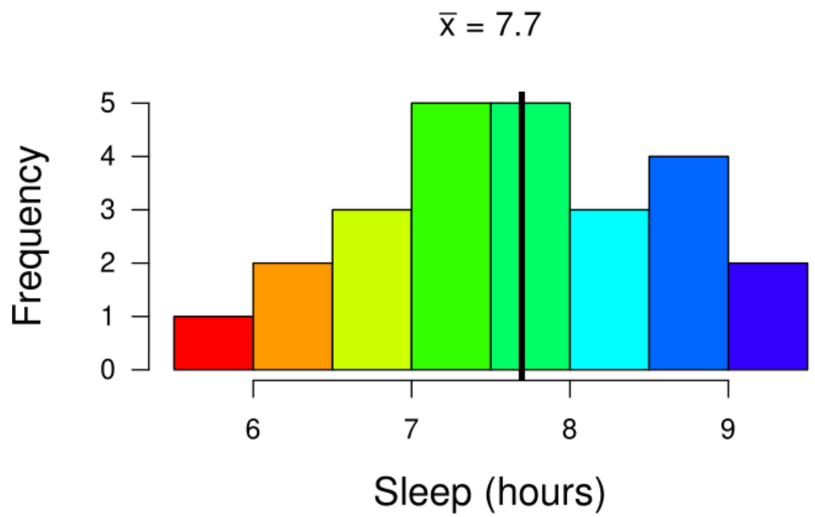
$$\bar{x} \pm 2.064 \frac{0.98}{\sqrt{25}}$$

7.7 - 2.064 * (0.98 / 5) = 7.3

7.7 + 2.064 * (0.98 / 5) = 8.1

**95% CI:
From 7.3 to 8.1**

Example: Sleepers



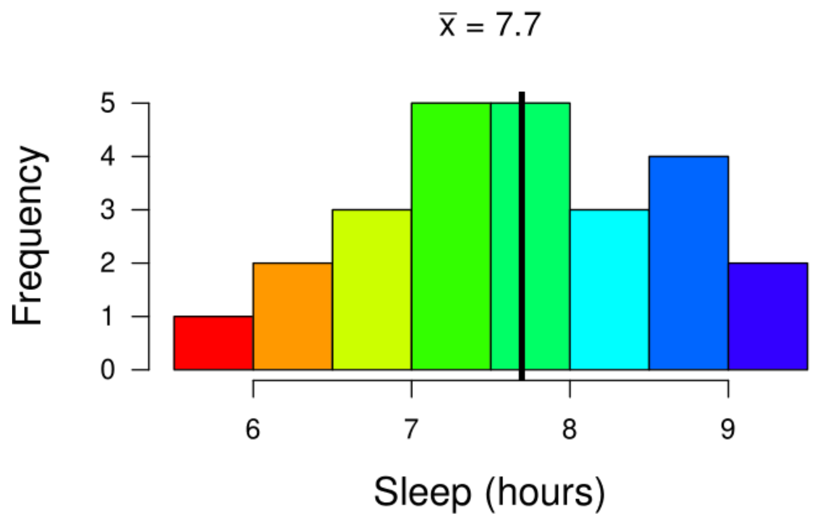
- n = 25
- $\bar{x}$ = 7.7
- s = 0.98

- Estimated population mean:
  - $\bar{x}$ = 7.7
- 95% CI:
  - Since we do not know σ, we have to estimate it, and use *t*-distribution:

We do not have evidence that these 25 participants are not getting enough sleep, since 8 hours falls inside the interval

95% CI:
From 7.3 to 8.1

Example: Sleepers



- n = 25
- $\bar{x}$ = 7.7
- s = 0.98

- Estimated population mean:
  - $\bar{x}$ = 7.7
- 95% CI:
  - Since we do not know σ, we have to estimate it, and use *t*-distribution:

If we were to repeat this experiment, each time construct a 95% CI, then 95% of those CI's will contain the population mean μ

95% CI:
From 7.3 to 8.1

Intermezzo:

On the interpretation of Confidence Intervals

95% CI:
From 7.3 to 8.1

“If we repeat this experiment, each time construct a 95% CI, in 95% of the cases the interval contains μ“

?

“There is a 95% probability that μ is between 7.3 and 8.1“

If this...

… then **not** necessarily this.

95% CI:
From 7.3 to 8.1

Intermezzo:
On the interpretation of Confidence Intervals

“If we repeat this experiment, each time construct a 95% CI, in 95% of the cases the interval contains μ“

“There is a 95% probability that μ is between 7.3 and 8.1“

My procedure for estimating sleep:

- I roll a 20-sided die
  - If 1 (5% of cases) my CI will be [-∞,0]
  - If 20 (95% of cases) my CI will be [0, ∞]

Will result in 95% of the CI’s containing μ, but is kind of ridiculous

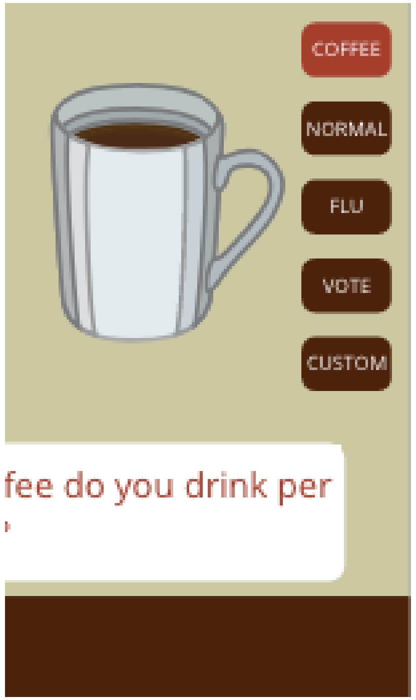
CI’s are about long-term probability: a single CI will have either 0% or 100% chance to contain μ

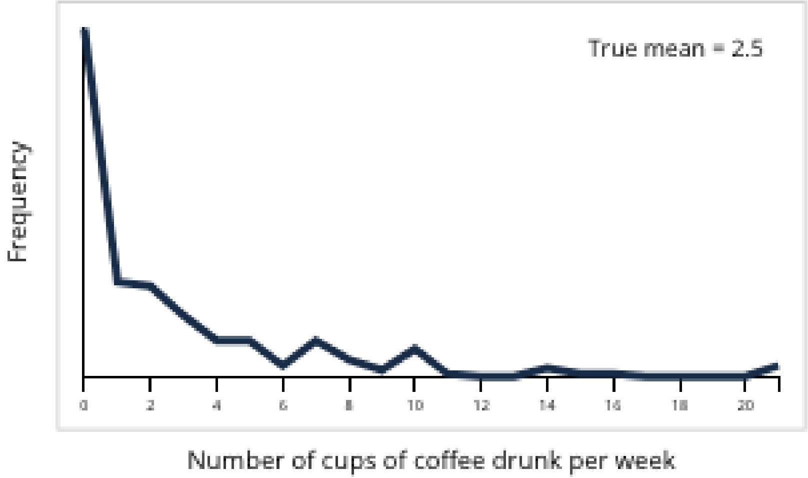
Overview of Today

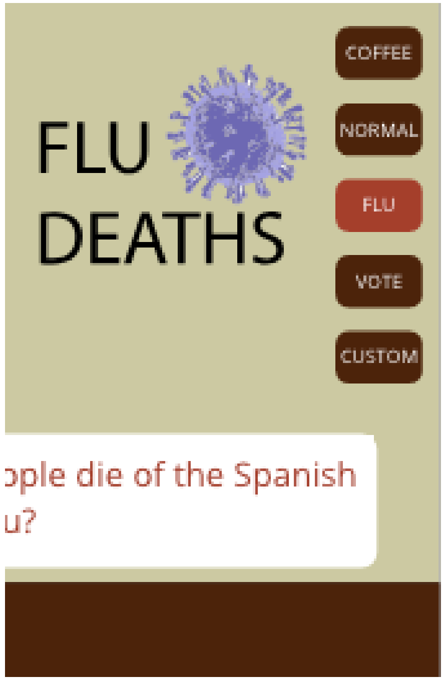
1. Recap of last Thursday and Tuesday
2. Confidence interval for the mean
   - *t*-distribution
3. **Robustness against extreme values**
4. Choosing a sample size
5. Recap
   - Example exam question

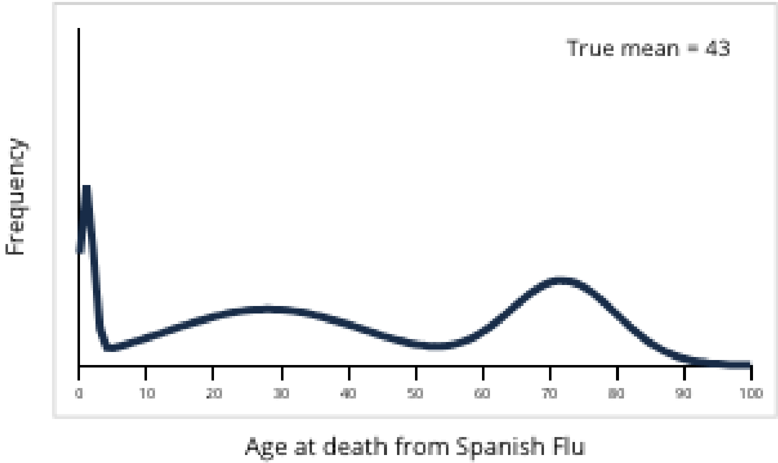
Robustness against extreme values

- Often, the population is not normally distributed
- If sample is large, then no problem (CLT)
- If sample is small, then a high probability of an outlier







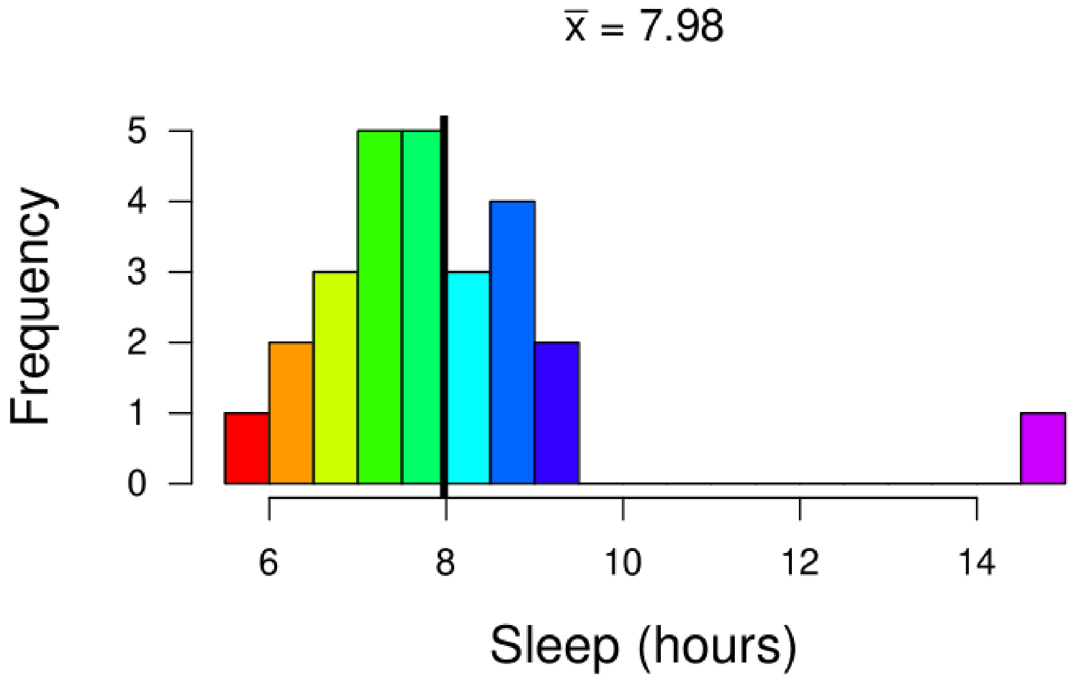


http://www.zoology.ubc.ca/~whitlock/Kingfisher/CLT.htm

Robustness against extreme values?

- Example: Previous sleepers + 1 extreme observation
- Outlier?
  - Use interquartile range
  - Q3 + 1.5 * IQR = 10.19
  - Yes

- Effect on confidence interval?

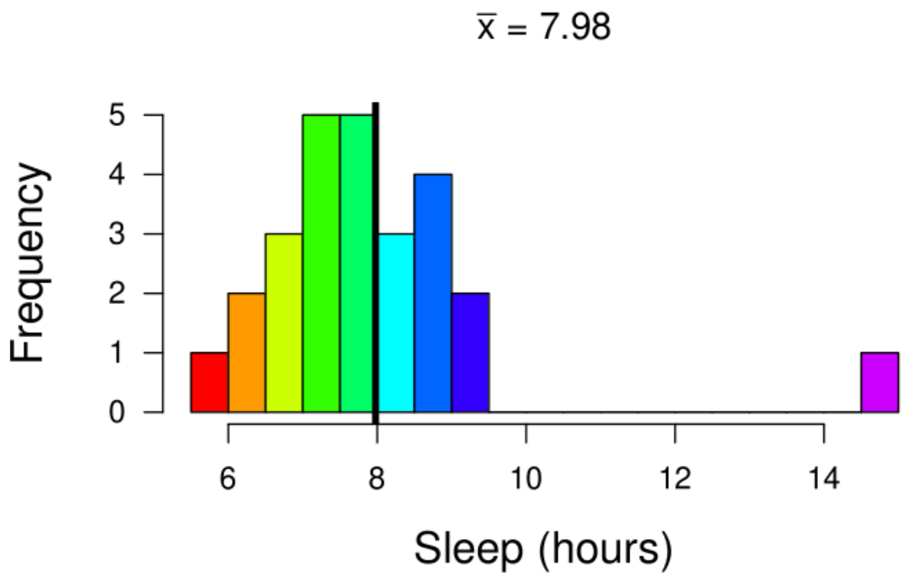


Robustness against extreme values?

- n = 26 (was 25)
- $\bar{X}$ = 7.98 (was 7.7)
- S = 1.73 (was 0.98)

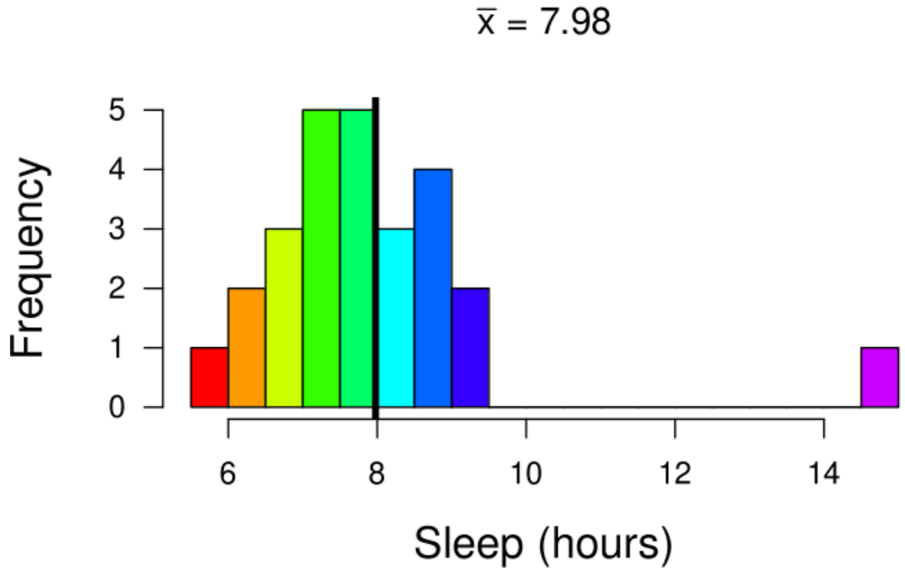
- $t_{.025}$ *also changes*, because we added one person to the sample (df = 25)

22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450



Robustness against extreme values?

- n = 26 (was 25)
- $\bar{X}$ = 7.98 (was 7.7)
- S = 1.73 (was 0.98)

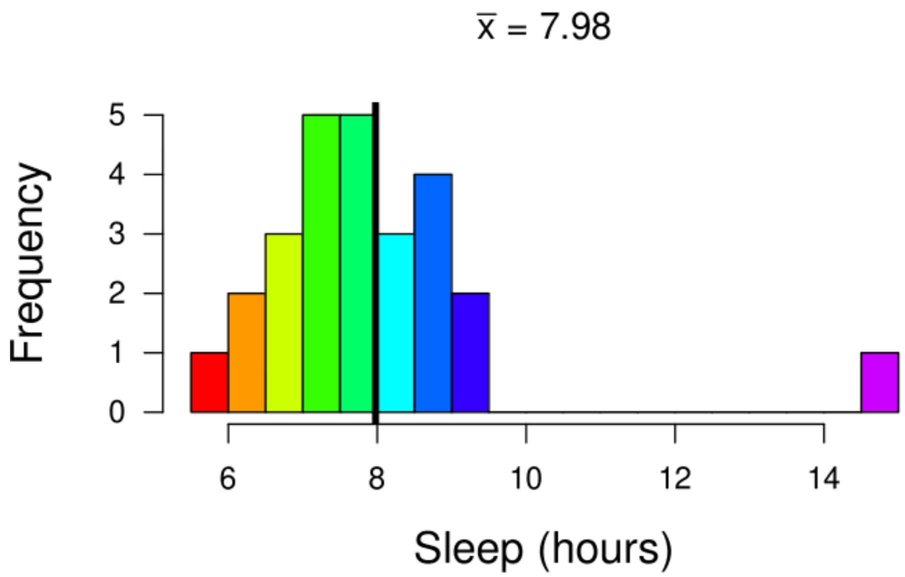


- CI:

$$\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}}$$

Robustness against extreme values?

- n = 26 (was 25)
- $\bar{X}$ = 7.98 (was 7.7)
- S = 1.73 (was 0.98)



- CI:

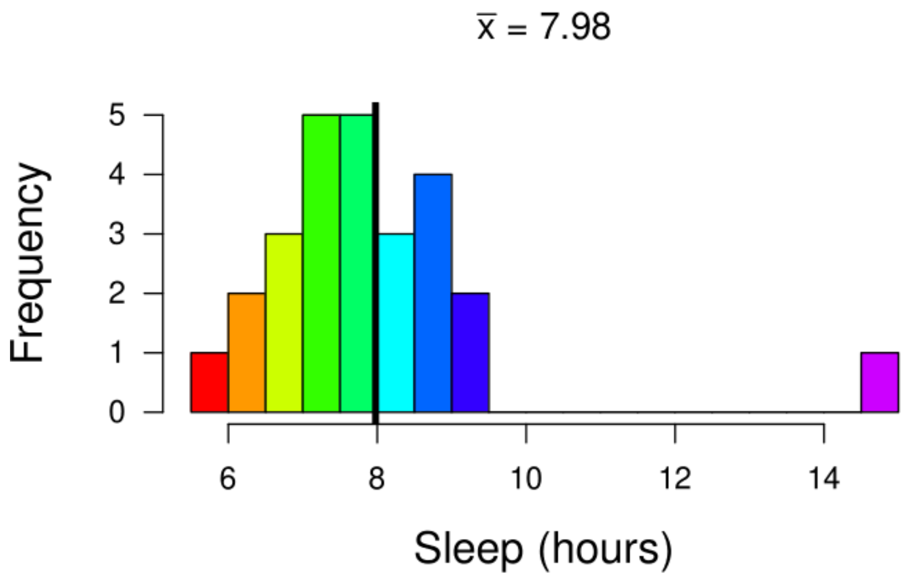
$$\bar{x} \pm t_{.025}\frac{s}{\sqrt{n}} = 7.98 \pm 2.060\frac{1.73}{\sqrt{26}}$$

OLD 95%CI:
From 7.3 to 8.1

NEW 95%CI:
From 7.28 to 8.68

Robustness against extreme values?

- n = 26 (was 25)
- X̄ = 7.98 (was 7.7)
- S = 1.73 (was 0.98)



- CI:

$$\bar{x} \pm t_{.025}\frac{s}{\sqrt{n}} = 7.98 \pm 2.060\frac{1.73}{\sqrt{26}}$$

OLD 95%CI:
From 7.3 to 8.1

NEW 95%CI:
From 7.28 to 8.68

Wider on *BOTH* sides
Still, 95% confidence

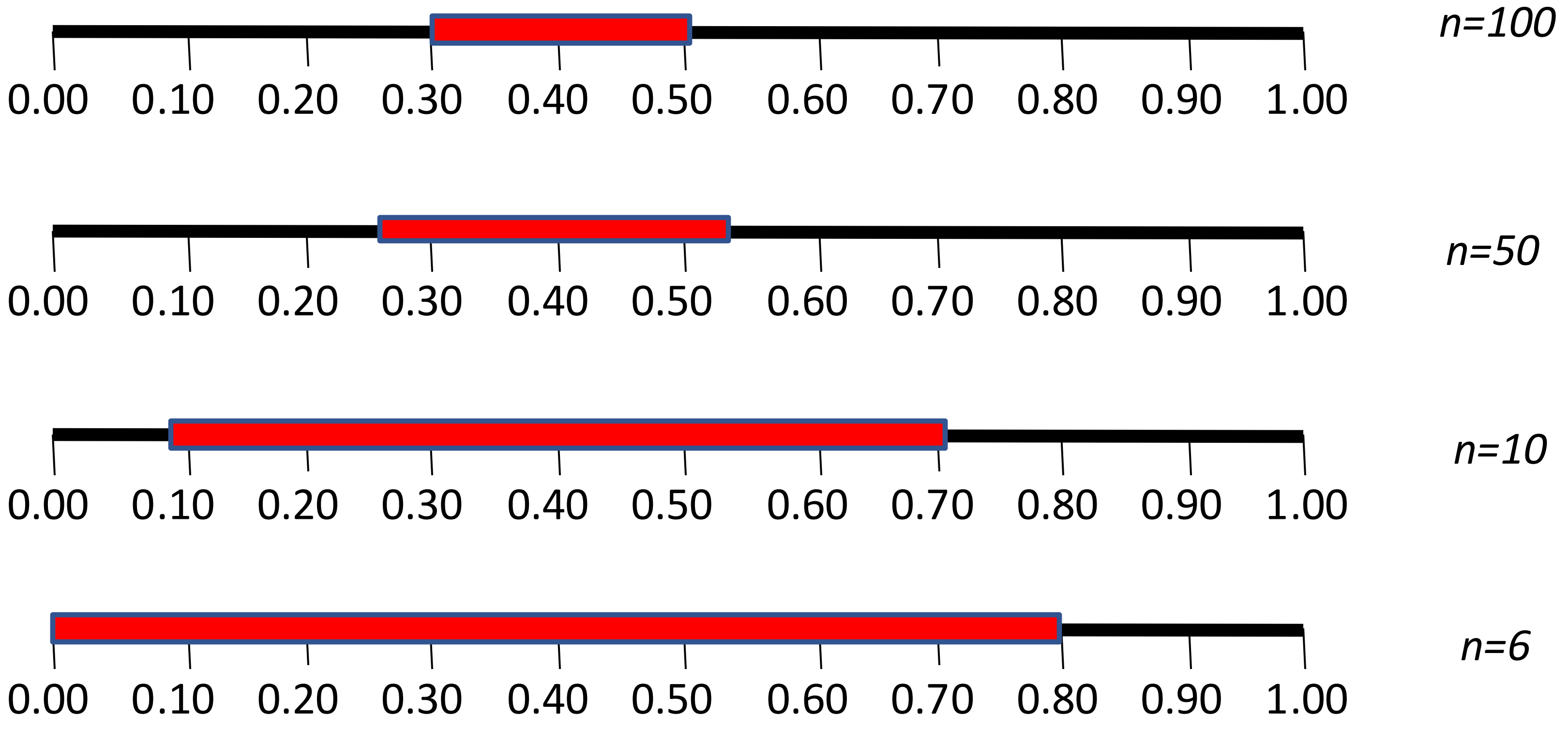
Overview of Today

1. Recap of last Thursday and Tuesday
2. Confidence interval for the mean
   - *t*-distribution
3. Robustness against extreme values
4. **Choosing a sample size**
5. Recap
   - Example exam question

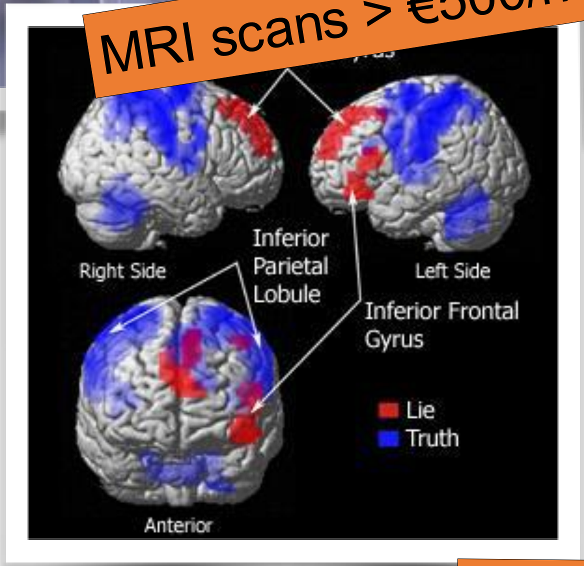
Riet: What affects confidence intervals?

Formula for confidence interval:
$\hat{p} \pm z(se)$ which is $\hat{p} \pm z\sqrt{\hat{p}(1-\hat{p})/\boldsymbol{n}}$

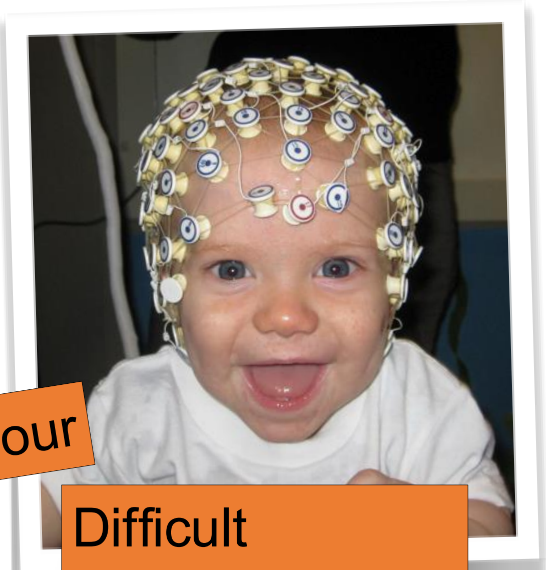
2) The sample size, n (*if **n** goes up, se goes down, making the interval smaller*)



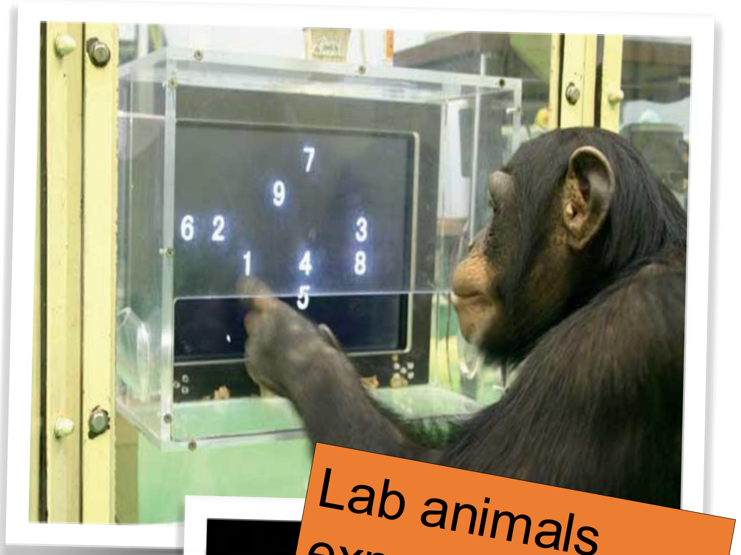
n is often small



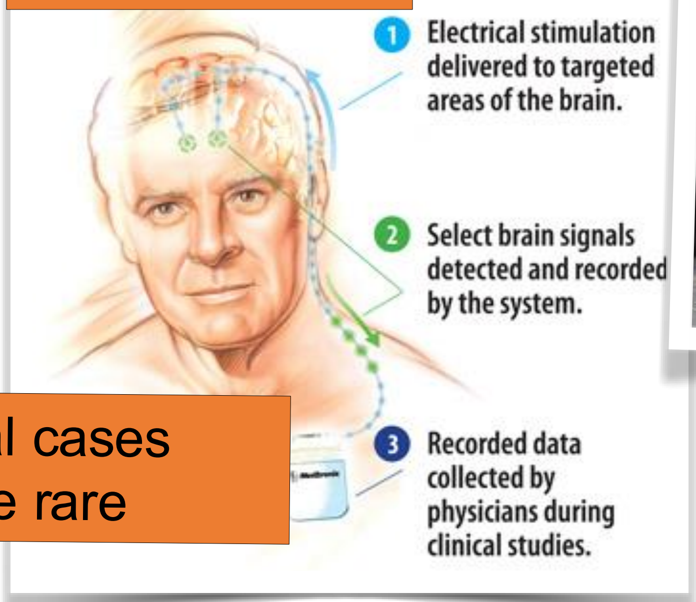
MRI scans > €500/hour



Difficult participants



Lab animals expensive



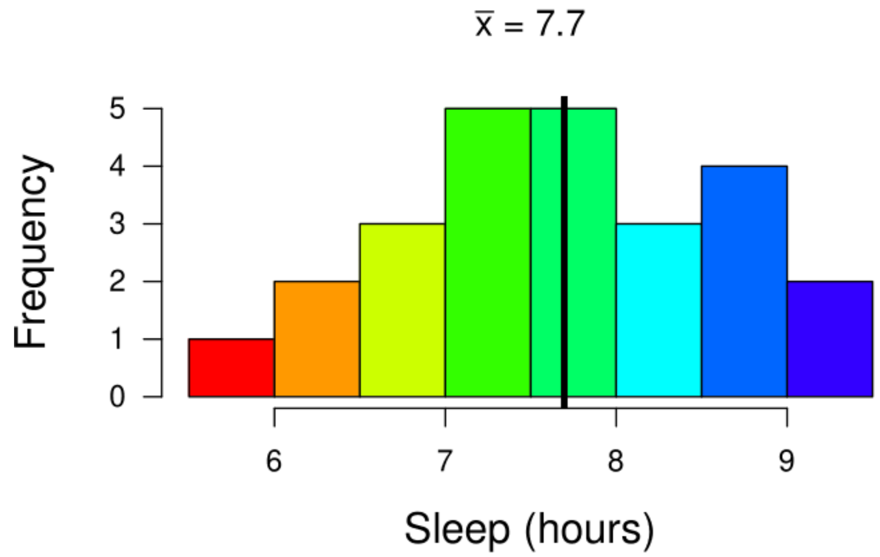
Clinical cases may be rare

Example: Sleepers

- $n = 25$
- $\bar{x} = 7.7$
- $s = 0.98$

- Estimated population mean:
  - $\bar{X} = 7.7$
- 95% CI:
  - From 7.3 to 8.1

- We do not have evidence for an average lack of sleep
  - Because of no effect?
  - Because of small sample?



How can we determine what sample size to use, in order to detect a difference of 0.3?

Example: Sleepers

- We could make an educated guess for a better sample size using $\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}}$
- We have to assume a z-distribution again (because *t* depends on the sample size)
- We have to specify the margin of error *m* to be detected
  - We want to be able to detect an effect of 8 – 7.7 = 0.3
- We have to guess the standard error/standard deviation
  - In our case, we can use the result of a previous experiment!
  - S = 0.98

$$n = \frac{\sigma^2 z^2}{m^2}$$

Example: Sleepers

- In our case, we can use the result of a previous experiment!
  - $\bar{X}$ = 7.7, so maybe appropriate *m* = 8 – 7.7 = 0.3
  - S = 0.98

- If we want 95% certainty:

$$n = \frac{\sigma^2 z^2}{m^2} = \frac{0.98^2 \times 1.96^2}{0.3^2} \approx 41$$

Example: Sleepers

- In our case, we can use the result of a previous experiment!
  - $\bar{X} = 7.7$, so maybe appropriate $m = 8 - 7.7 = 0.3$
  - $S = 0.98$
- If we want 99% certainty:

$$n = \frac{\sigma^2 z^2}{m^2} = \frac{0.98^2 \times 2.58^2}{0.3^2} \approx 71$$

Example: Sleepers

- In our case, we can use the result of a previous experiment!
  - $\bar{X}$ = 7.7, so maybe appropriate *m* = 8 – 7.7 = 0.3
  - S = 0.98

- If we want a smaller margin of error m (in other words, detect a smaller effect), for instance 0.1:

$$n = \frac{\sigma^2 z^2}{m^2} = \frac{0.98^2 \times 1.96^2}{0.1^2} \approx 369$$

Choosing a sample size

The minimal sample size to obtain a confidence interval for μ *with margin of error m is*

$$n = \frac{\sigma^2 z^2}{m^2}$$

The minimal sample size to obtain a confidence interval for p with margin of error m is

$$n = \frac{\hat{p}(1-\hat{p})z^2}{m^2}$$

Overview of Today

1. Recap of last Thursday and Tuesday
2. Confidence interval for the mean
    - *t*-distribution
3. Robustness against extreme values
4. Choosing a sample size
5. **Recap**
    - Example exam question

Confidence interval for the mean

Confidence interval for the mean: A 95% confidence interval for the population mean μ is

$$\bar{x} \pm margin\ of\ error$$

$$\bar{x} \pm t_{.025} \frac{s}{\sqrt{n}}$$

where $t_{.025}$ depends on the sample size *n*.

What Affects the Confidence Interval?

- Confidence level → different critical t/z-value
  - Balancing between confidence and accuracy
  - What would a 100% CI look like?
- Standard deviation → different standard error
- Sample size → different standard error
- Mean → different centre

Recap 1/2

1. We are trying to estimate a point (i.e., population mean) based on a sample and its statistics (i.e., sample mean/sd)
2. The sampling distribution tells us how reliable our estimate is (i.e., the sd of the sampling distribution = the standard error)
3. Because we do not know the mean and sd of the sampling distribution, we approximate this using the sample statistics and use the t-distribution

Recap 2/2

1. Based on the confidence interval, we can make claims about the population and *quantify our uncertainty*

2. We can calculate what sample size we need to detect effects: smaller effects (e.g., differences between groups) require higher sample size!

A 95% confidence intervals means that in the long run,
95% of your confidence intervals will include the true parameter value

Overview of Today

1. Recap of last Thursday and Tuesday
2. Confidence interval for the mean
   - *t*-distribution
3. Robustness against extreme values
4. Choosing a sample size
5. Recap
   - **Example exam question**

Example Exam question

- An educational psychologist studies school absence of children in the Netherlands. She has a sample of 41 children. The children in the sample are on average 19.8 days per year absent. The sample standard deviation is 4.7.
- What is the 99% confidence interval for the estimate of the mean number of sick days in the population of school-going children?

  a) (10.59, 29.01)

  b) (18.32, 21.28)

  c) (17.82, 21.78)

Solution

- Because we are computing 99% CI for the *mean*, we use $\bar{x} \pm t_{.005} \frac{s}{\sqrt{n}}$
  - *n* = 41
  - $\bar{x}$ = 19.8
  - *s* = 4.7
  - $t_{.005} = 2.704$ (because we are looking at a *t*-distribution with 40 df, for a 99% CI)

Solution

- Because we are computing 99% CI for the *mean*, we use $\bar{x} \pm t_{.005} \frac{s}{\sqrt{n}}$
  - $n = 41$
  - $\bar{x} = 19.8$
  - $s = 4.7$
  - $t_{.005} = 2.704$ (because we are looking at a *t*-distribution with 40 df, for a 99% CI)

  - $\bar{x} - t_{.005} \frac{s}{\sqrt{n}} = 19.8 - 2.704 \frac{4.7}{\sqrt{41}} = 17.82$
  - $\bar{x} + t_{.005} \frac{s}{\sqrt{n}} = 19.8 + 2.704 \frac{4.7}{\sqrt{41}} = 21.78$

Example Exam question

- An educational psychologist studies school absence of children in the Netherlands She has a sample of 41 children. The children in the sample are on average 19.8 days per year absent. The sample standard deviation is 4.7.
- What is the 99% confidence interval for the estimate of the mean number of sick days in the population of school-going children?

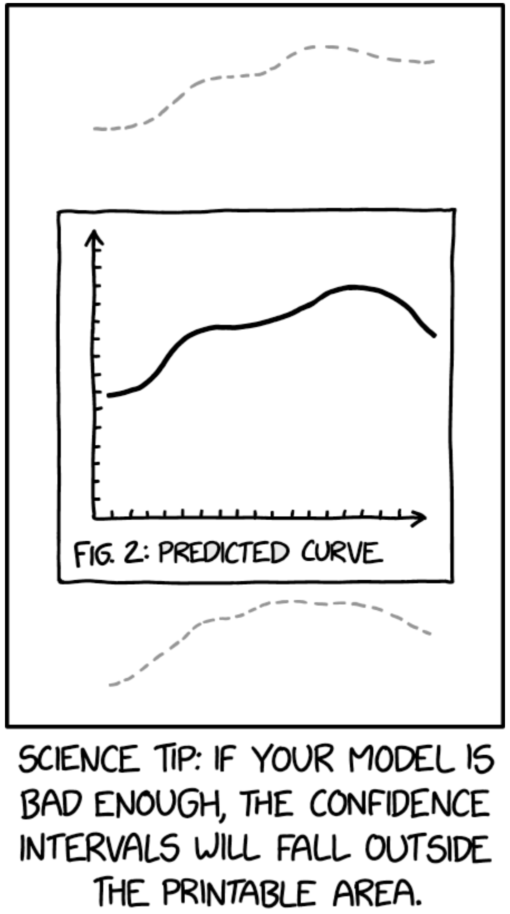
~~a) (10.59, 29.01)~~

~~b) (18.32, 21.28)~~ → Solution assuming 95% CI!

c) (17.82, 21.78)

Questions?

Thank you for your attention



https://www.xkcd.com/2311/

Bonus Video

Hans Rosling on estimating the global future population (Dutch intro, but captions available), it's from 2015 so getting pretty old – still really like the enthusiasm of the guy

https://www.youtube.com/watch?v=fPtfx0C-34o&feature=youtu.be

For more further understanding of confidence intervals, Khan academy statistics has a great little video about the CI of a proportion: https://www.youtube.com/watch?v=hlM7zdf7zwU

Excel commands for calculations:

Regular numeric operations:
= 2.3 + 3 - 7 * SQRT(10) / 4^2

For distributions → surface area:
= NORM.DIST(2, 0, 1, TRUE)
= T.DIST(2, 14, TRUE)

You start with inputting the observed test statistic. Then you give the df (for X^2 and t distributions), or the mean and sd (for the normal/z distribution). Then you add TRUE to get the area under the curve, to the LEFT of the test statistic.

For distributions-→ critical value:
= NORM.INV(0.025, 0, 1)
= T.INV(0.025, 20)

You start with inputting the desired probability. For instance, 0.025 gives you the value of the distribution, where to the left of that value, 2.5% of the distribution is situated. This is useful when constructing the 95% CI.
Then you give the df (for t distributions), or the mean and sd (for the normal/z distribution).

→ Paste these into any cell and press enter!

Highlighted exercises from the book

- 8.36
- 8.68
- 8.76

→ try yourself first, then check next slides for answers

8.36

n = 590
sample mean = 2.56
sample SD = 0.84

A. Point estimate of population mean = point estimate = $\bar{x}$ = 2.56
B. Standard error of the sample mean:

   SE = s / sqrt(n) = 0.84 / sqrt(590) = 0.0346
C. Interpretation of 95% CI (2.49, 2.62):

   If we repeat this experiment over and over, 95% of these confidence intervals contain the true population mean; we are 95% confident that the true mean lies in this interval of [2.49, 2.62]
D. Is μ = 2 plausible? Explain:

   No — 2 is outside the 95% CI (2.49, 2.62), so it is not plausible at the 5% level

Given:

8.68

- $n = 825$
- $x = 618$
- $\hat{p} = 618 / 825 = 0.7491$
- Check normal-approximation conditions:
  - $n\hat{p} = 825(0.7491) \approx 618$, greater than 15
  - $n(1-\hat{p}) \approx 207$, greater than 15, so both OK

a. Standard error:

- $SE = \text{sqrt}[\hat{p}(1 - \hat{p}) / n]$
  $= \text{sqrt}[0.7491 \times (1 - 0.7491) / 825]$
  $= 0.0151$

b. Margin of error (99% confidence, $z = 2.576$):

- $ME = z \times SE$
  $= 2.576 \times 0.0151$
  $= 0.0389$

c. 99% confidence interval:

- $\hat{p} \pm ME = 0.7491 \pm 0.0389$
  $= (0.7102, 0.7880)$

8.76

The proportion to be estimated is 0.5, where the error margin is 0.05 and the desired confidence is 95% (so a z-score of 1.96 needs to be used). We can plug these values into the formula for computing the required sample size:

$(0.5^2 * 1.96^2) / 0.05^2 = 384$