Chapter 8.1

Estimating μ When σ is Known

Learning Objectives

At the end of this lecture, the student should be able to:

- State the formula for the margin of error
- Explain how to choose z_c when making a confidence interval.
- State one of the assumptions behind x that needs to be met in order to make confidence intervals
- Interpret what a confidence interval means
- Demonstrate how to calculate sample size

Introduction

- Review of key points from Chapter 7
- Explanation of confidence intervals
- Demonstration of how to calculate confidence intervals
- Demonstration of how to calculate sample size

From the US National Archives and Records Administration

Review of Chapter 7

I know you are sick of Chapter 7, but please humor me

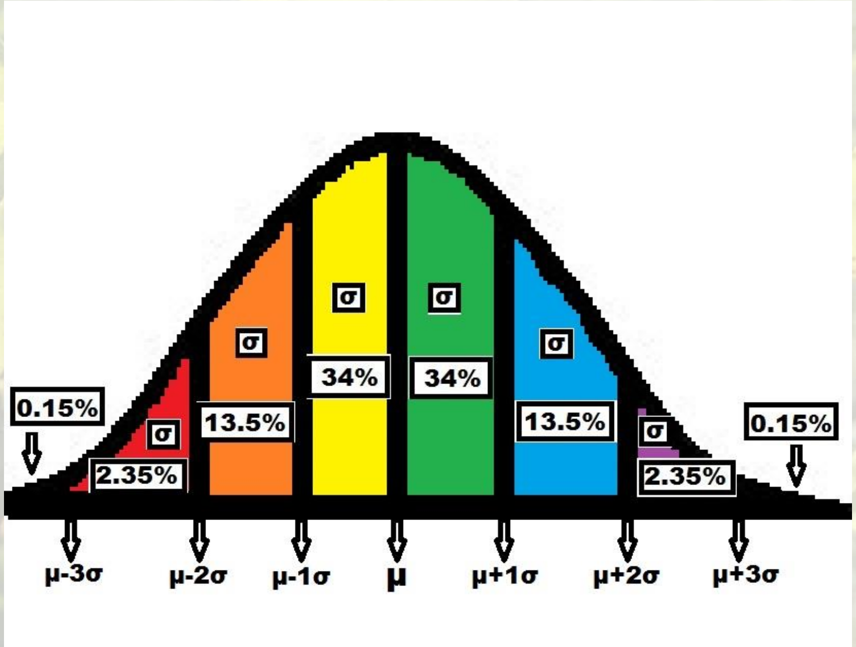
Types of Inferences

1. <u>Estimation:</u> we estimate the value of a population *parameter* using a *sample*

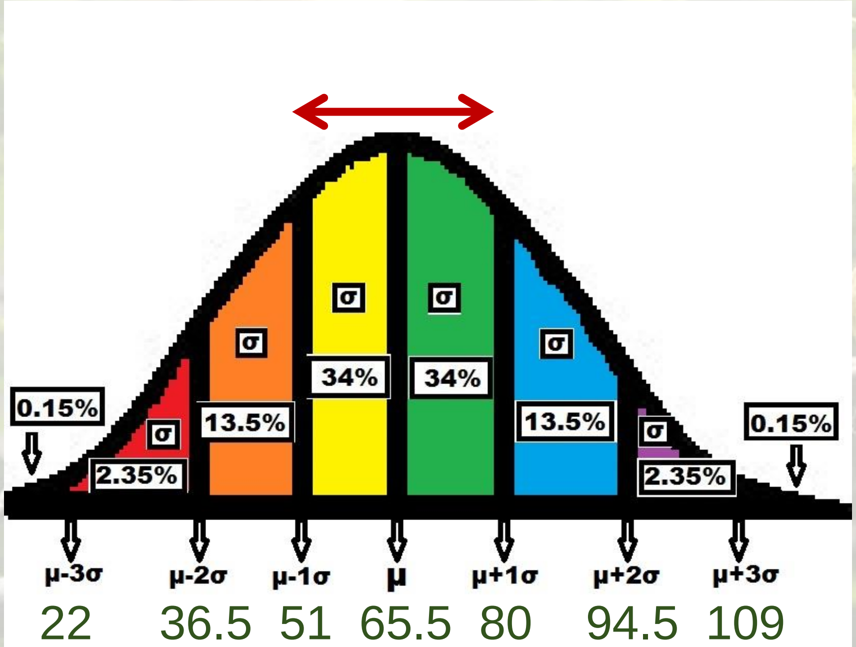
We will practice this in Chapter 8

One Point from 7.1

- In 7.1, we learned about the Empirical Rule.
- Most questions focused on the probability of selecting an x below or above a cutpoint.
- However, there was this one question about the *middle 68%*



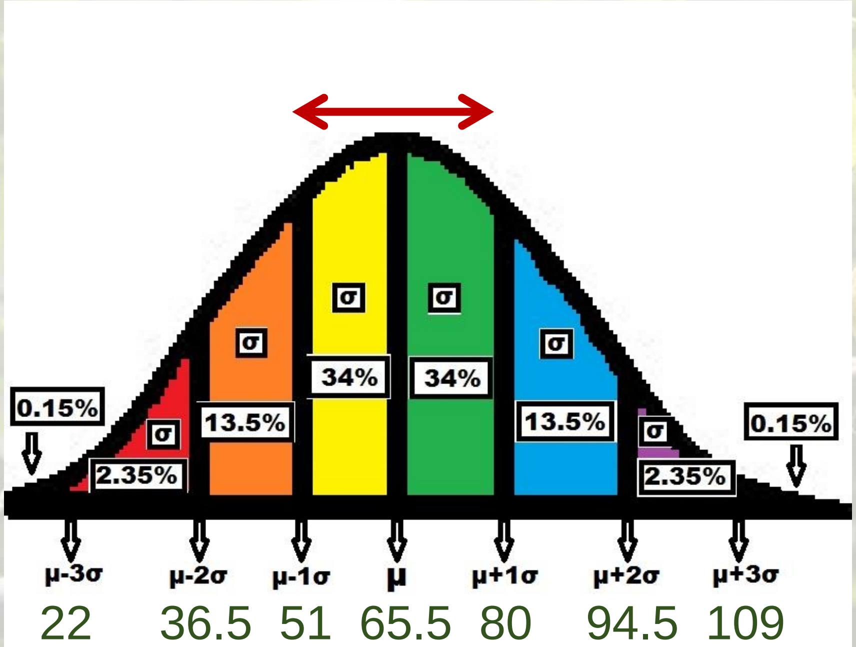
Slide from Chapter 7.1: Empirical Rule



Question: What are the cutpoints for the middle 68% of the data?

Answer: Middle 68% means 34% above mean (80) and 34% below mean (51).

Slide from Chapter 7.1: Empirical Rule



Question: What are the cutpoints for the middle 68% of the data?

Answer: Middle 68% means 34% above mean (80) and 34% below mean (51).

Please notice: This means that the probability of selecting an x between 51 and 80 was 68%.

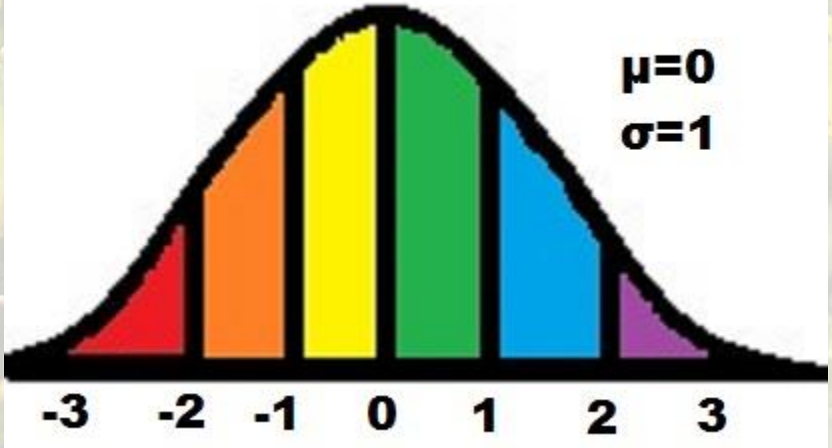
One Point from Chapters 7.2-7.3

$\mu = 65.5$
$\sigma = 14.5$

- In 7.2-7.3, we learned about finding x's not on Empirical Rule cutpoints by calculating z and looking up probabilities in the z table.
- Most questions focused on the probability of selecting an x below or above a cutpoint.
- However, there was this one question about the *middle 20%*

$$z = \frac{x - \mu}{\sigma}$$

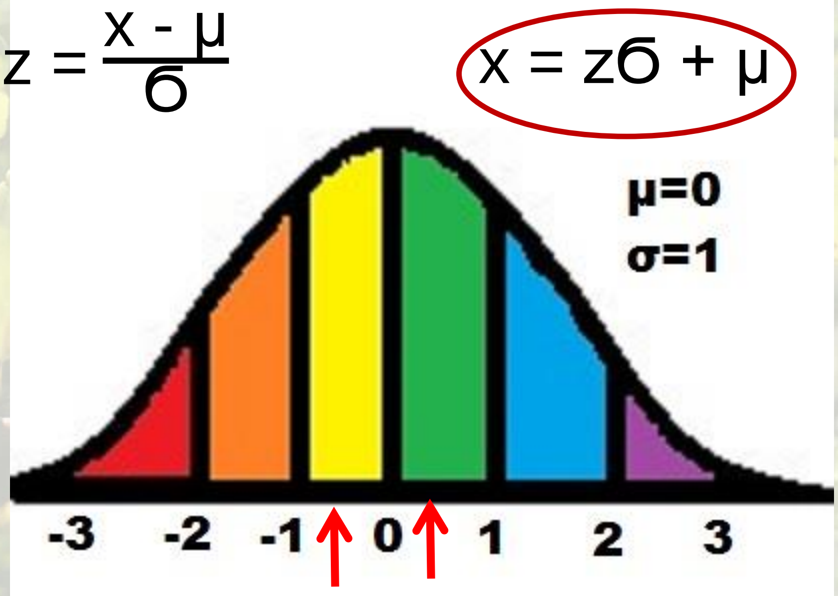
$$x = z\sigma + \mu$$



Slide from Chapters 7.2-7.3

$\mu = 65.5$
$\sigma = 14.5$

- What scores mark the middle 20% of the data?
- Strategy is to find the z-score for (1-0.2000)/2 = 0.4000
- For p = 0.4013, z = -0.25
- Also, z = 0.25 is on the positive side.
- x for the left side:
  - x = (-0.25*14.5) + 65.5 = 61.9
- x for the right side:
  - x = (0.25*14.5) + 65.5 = 69.1
- 61.9 and 69.1 mark the middle 20% of the data.



p = 0.4013	p = 0.4013
z = -0.25	z = 0.25
x = 61.9	x = 69.1

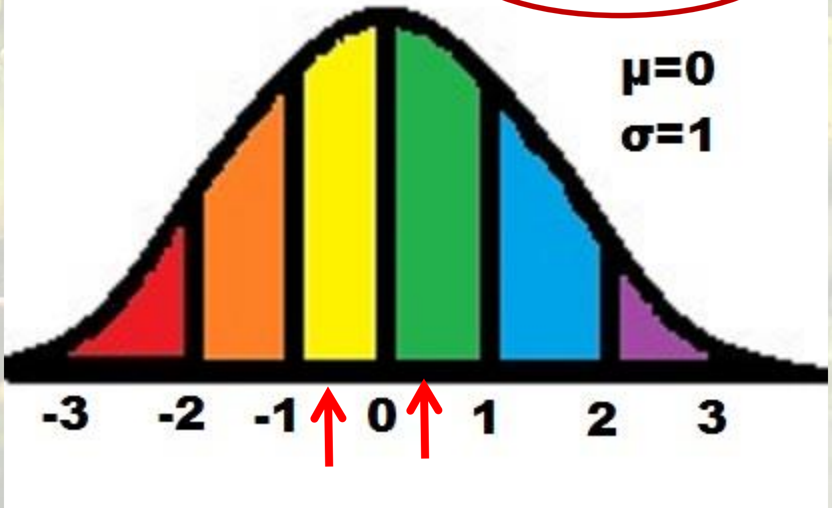
Slide from Chapters 7.2-7.3

μ = 65.5
σ = 14.5

- What scores mark the middle 20% of the data?
- Strategy is to find the z-score for (1-0.2000)/2 = 0.4000
- For p = 0.4013, z = -0.25
- Also, z = 0.25 is on the positive side.
- x for the left side:
  - x = (-0.25*14.5) + 65.5 = 61.9
- x for the right side:
  - x = (0.25*14.5) + 65.5 = 69.1
- 61.9 and 69.1 mark the middle 20% of the data.

$$z = \frac{x - \mu}{\sigma}$$

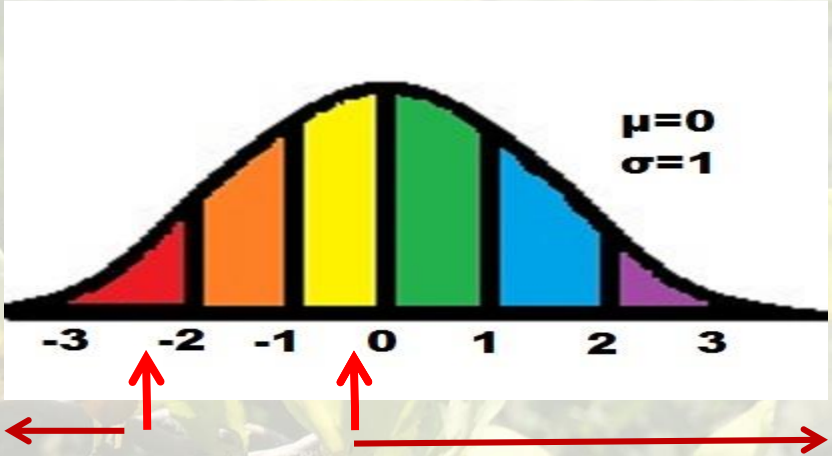
$$x = z\sigma + \mu$$



Please notice: This means that the probability of selecting an x between 61.9 and 69.1 was 20%.

Slide from Chapter 7.4-7.5

- Assume the 100-student class is a population.
- Now I have to pick an n
    - Let's pick 36.
- *Question:* What is the probability of me selecting a sample of 36 students with an x-bar between 60 and 65? *1 – 0.0113 – 0.5832 = 0.4055*



z1 = (60-65.5)/2.4 = -2.28

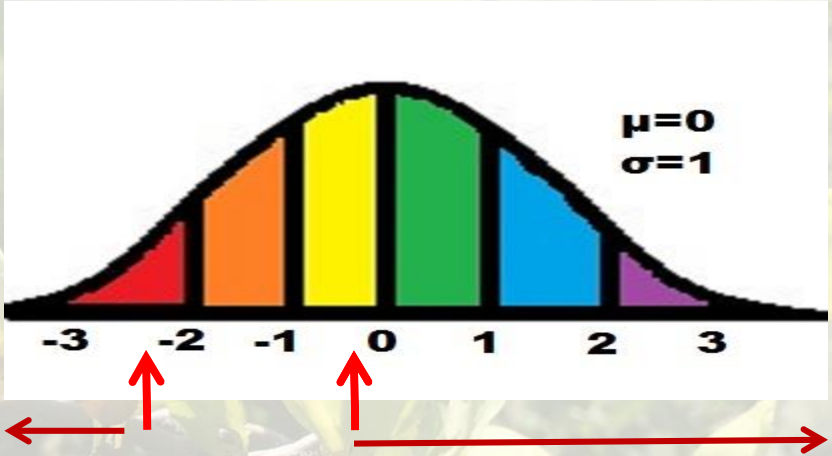
p1 = 0.0113

z2 = (65-65.5)/2.4 = -0.21

p2 = 0.5832

Slide from Chapter 7.4-7.5

- Assume the 100-student class is a population.
- Now I have to pick an n
  - Let's pick 36.
- *Question:* What is the probability of me selecting a sample of 36 students with an x-bar between 60 and 65? *1 – 0.0113 – 0.5832 = 0.4055*



z1 = (60-65.5)/2.4 = -2.28

p1 = 0.0113

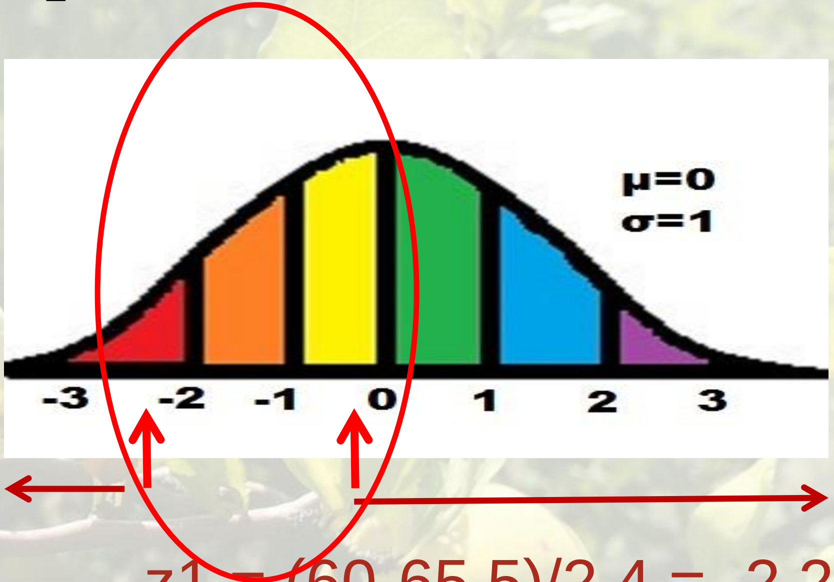
z2 = (65-65.5)/2.4 = -0.21

p2 = 0.5832

The probability of selecting a sample of 36 so that the x-bar is between 60 and 65 is 41%.

Slide from Chapter 7.4-7.5

- Unbalanced question
- What if you wanted to find the x-bar lower and upper limits of the middle 90%, or 95%, or 99% of the distribution?
- Remember the CLT…



z1 = (60-65.5)/2.4 = -2.28

p1 = 0.0113

z2 = (65-65.5)/2.4 = -0.21

p2 = 0.5832

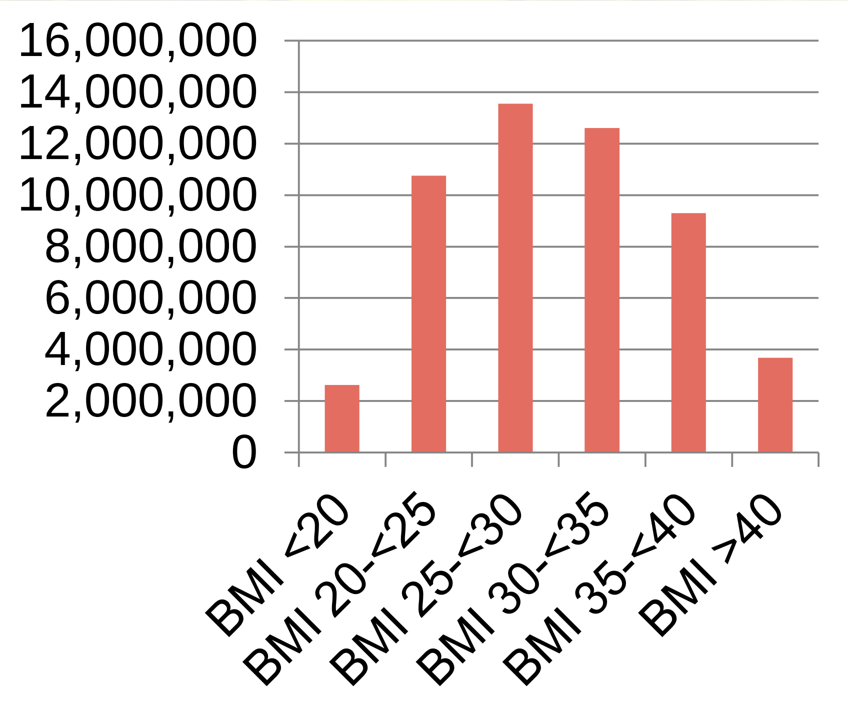
The probability of selecting a sample of 36 so that the x-bar is between 60 and 65 is 41%.

Confidence Intervals and the CLT

More Theory

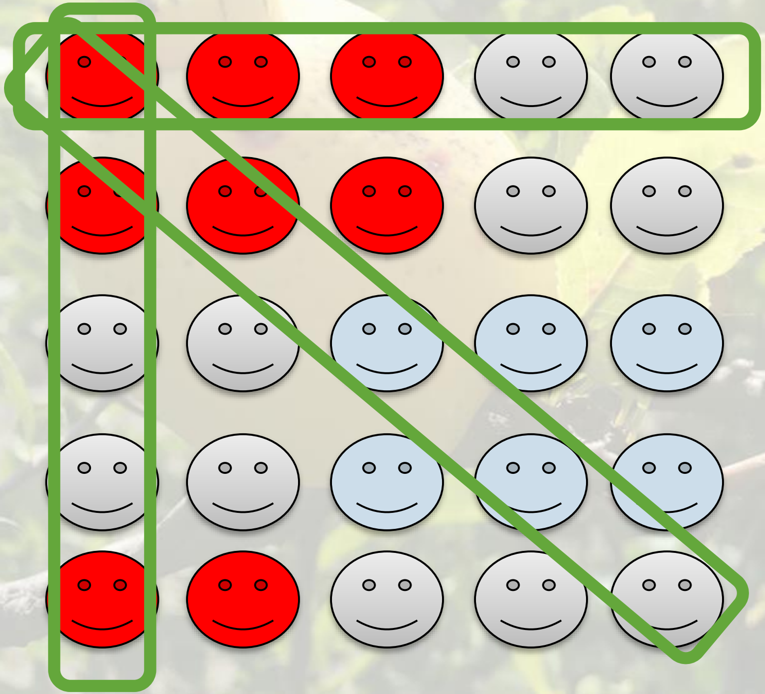
Remember the Histogram of X-bars?

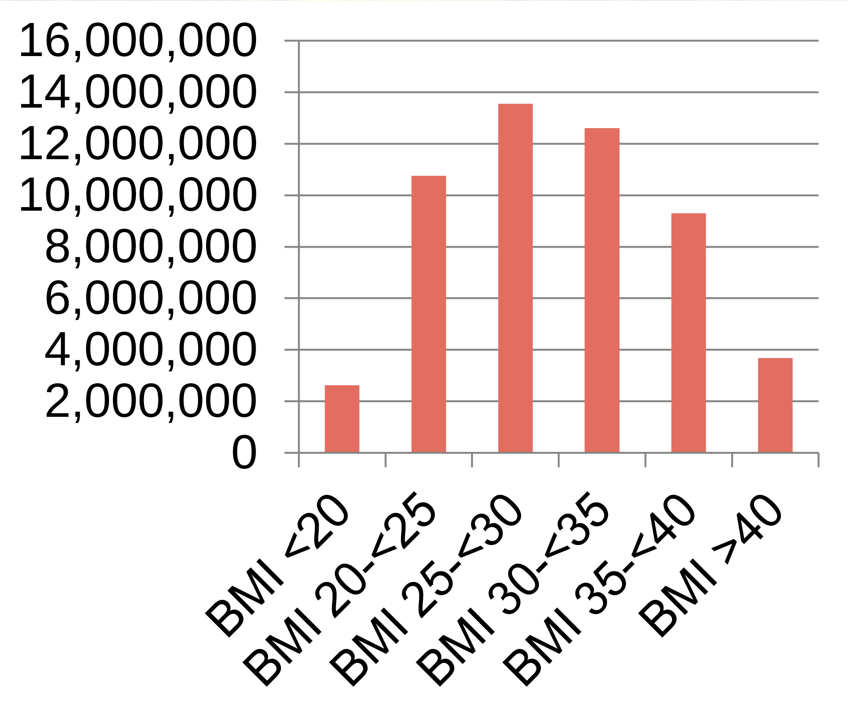
- Each individual frequency count represents an x-bar from a sample.
- In real life, we usually only get one sample
  - Where will its x-bar be?
  - How far will the x-bar be from μ?



Remember the Histogram of X-bars?

x-bar = 23



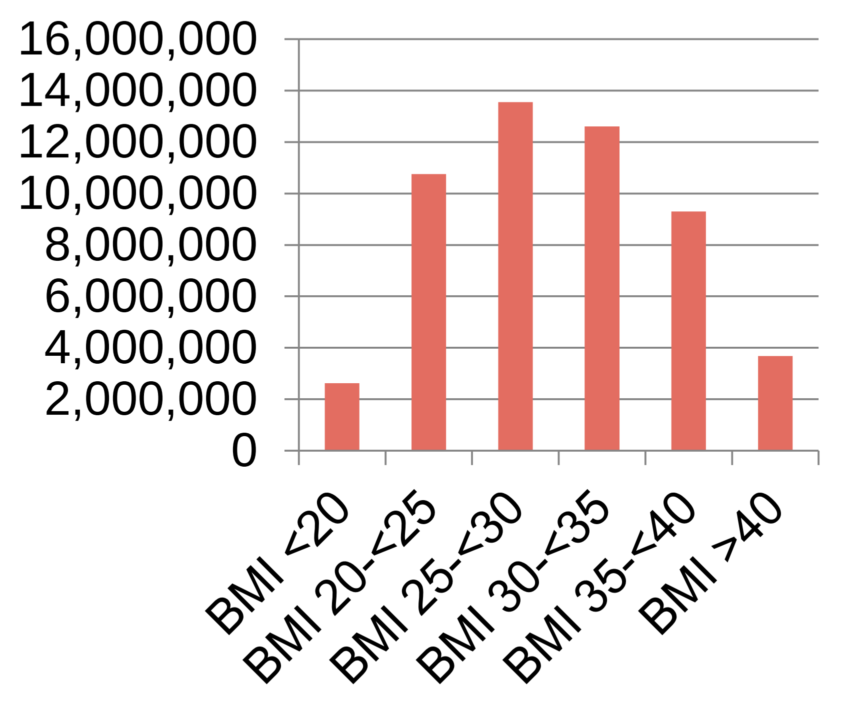


x-bar = 21

x-bar = 25

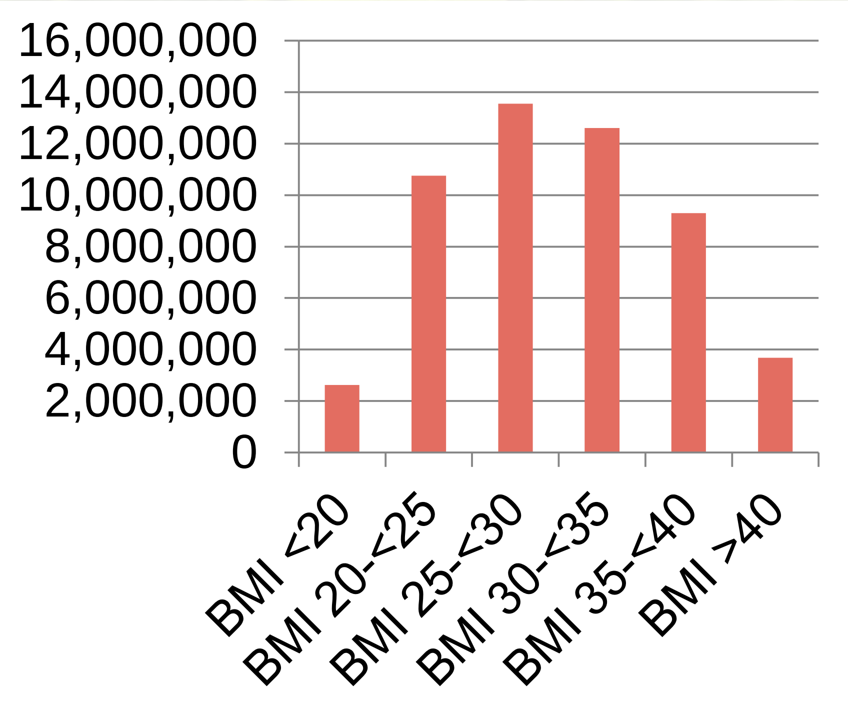
Strategy

- Don't get too excited about x-bar due to sampling error
- You want to create a range, or interval, around x-bar, with a lower limit and an upper limit.
- Within those limits are ***MOST*** of the x-bars you could have gotten from other samples.
- …but not ***ALL*** – probability!



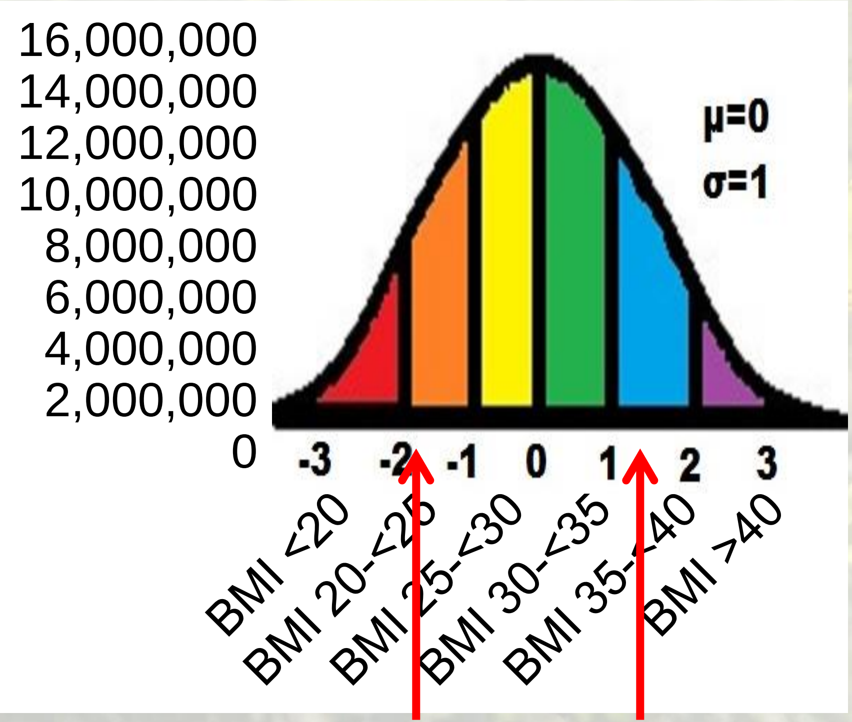
Strategy

- Don't get too excited about x-bar due to sampling error
- You want to create a range, or interval, around x-bar, with a lower limit and an upper limit.
- Within those limits are ***MOST*** of the x-bars you could have gotten from other samples.
- …but not ***ALL*** – probability!



Strategy

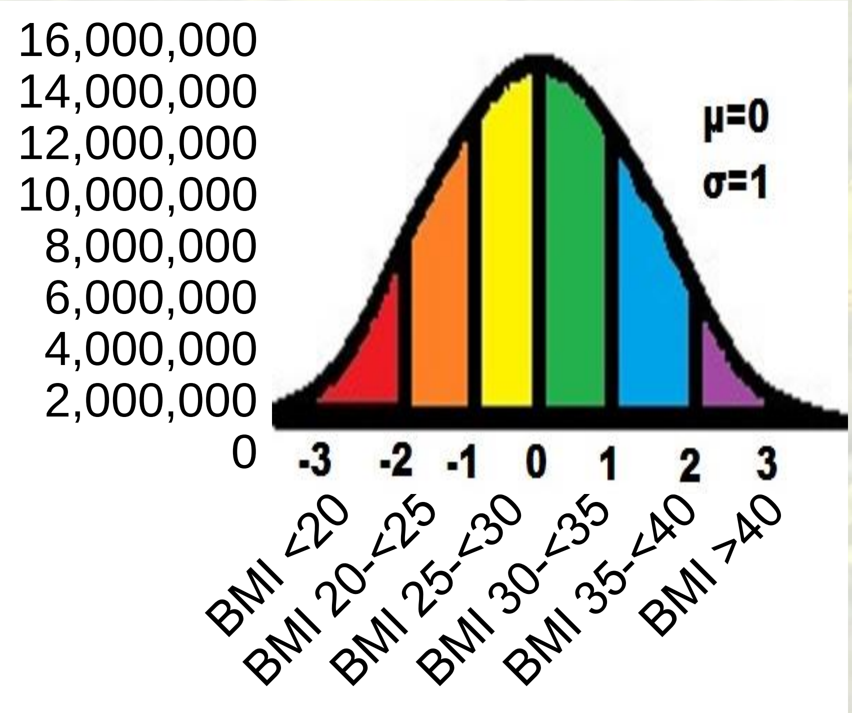
- Don't get too excited about x-bar due to sampling error
- You want to create a range, or interval, around x-bar, with a lower limit and an upper limit.
- Within those limits are ***MOST*** of the x-bars you could have gotten from other samples.
- …but not ***ALL*** – probability!



Middle 95%

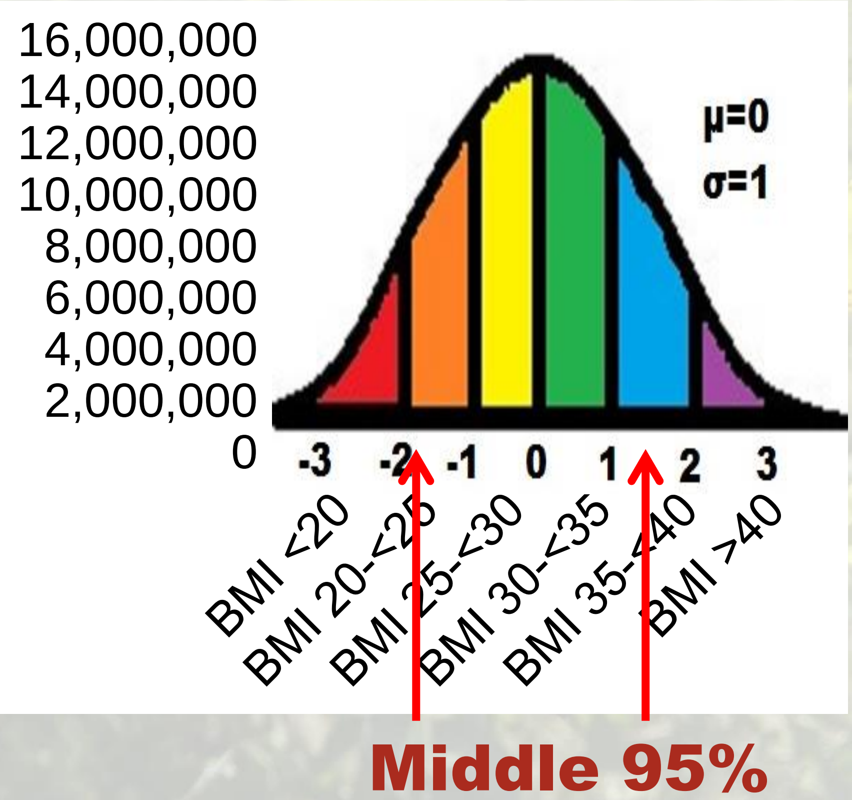
Example

- Imagine you took 100 samples (of any size) from this population. Then you'd get 100 x-bars.



Example

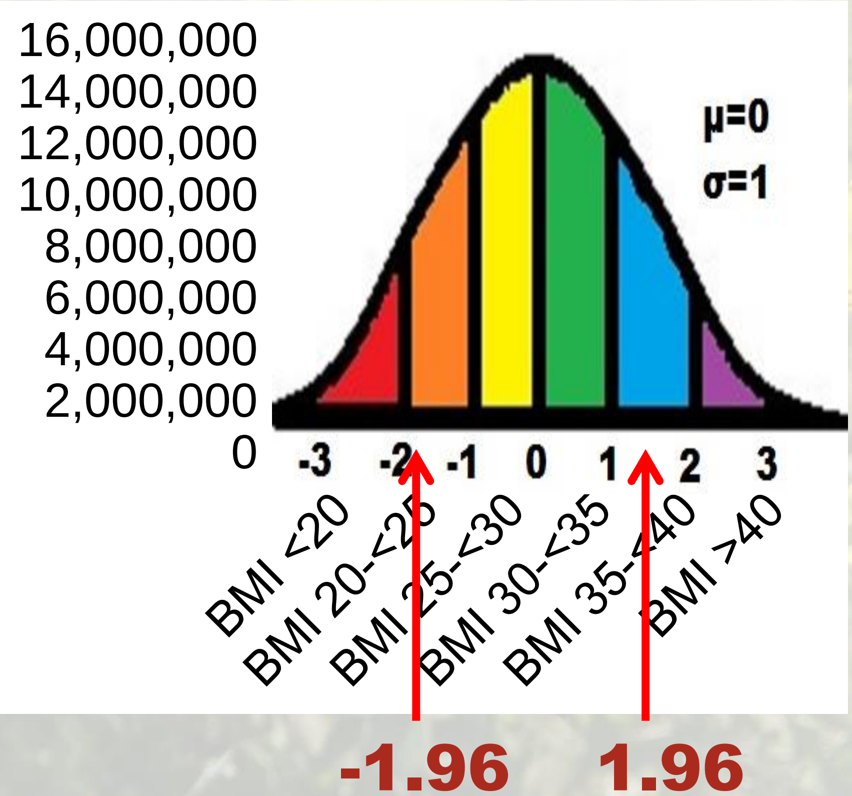
- Imagine you took 100 samples (of any size) from this population. Then you'd get 100 x-bars.
- If you figured out the upper and lower limits for the **middle 95% of sampling distribution**, you'd find that **95 out of 100 of your x-bars were within those limits.**



Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58



Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

These 3 things are all the same thing:

- Critical value z
- "z sub c"
- z_c

The z you pick determines

- The width of the interval around x-bar
- The confidence level
- Note: the more confident you want to be, the bigger the z_c

Confidence Interval: *"I am 95% confident μ is between the lower limit and the upper limit"*

Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

A confidence interval for μ is an interval:

- *Computed from sample data*
- *where "level of confidence" (abbreviated c) refers to the probability of having the resulting interval contain the actual value of μ*

Another way of saying it:

"Level of confidence" is

- *the proportion of confidence intervals*
- *calculated from a random sample of size n samples*
- *that actually contains μ*

Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

Don't get too excited about x-bar, because it is ***a point estimate***, meaning a one-time estimate, of a population parameter. There can be others!

Remember ***sampling error***? When using x-bar as a point estimate for μ, the ***margin of error (ME, or E)*** is the absolute difference between the x-bar you got and μ.

Example from students:

If x-bar = 60, E = 5.5

If x-bar = 90, E = 24.5

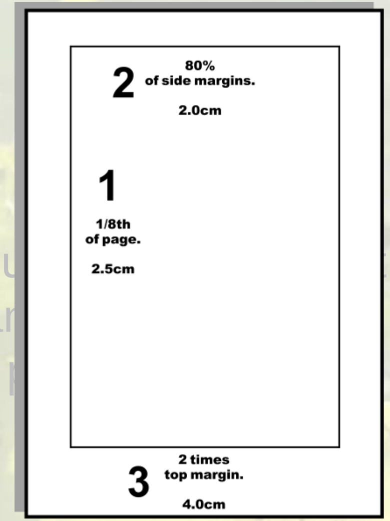
μ = 65.5
σ = 14.5

Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

Don't get too excited abou
is ***a point estimate***, mear
estimate, of a population
can be others!



Remember ***sampling error***? When using x-bar as a point estimate for μ, the ***margin of error (ME, or E)*** is the absolute difference between the x-bar you got and μ.

Example from students:

If x-bar = 60, E = 5.5

If x-bar = 90, E = 24.5

μ = 65.5
σ = 14.5

Image by Jazzmanian in the public domain

Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

Assumptions about x to consider before making confidence intervals

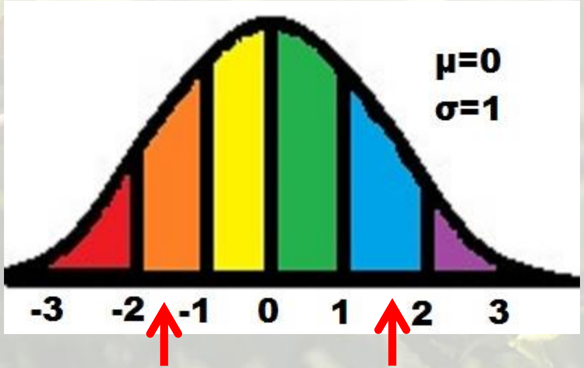
1. Simple random sample of size n has been drawn from a population of x values
2. The value of σ is known
3. *If x itself has a normal distribution*, then we know that x-bar will no matter how big our sample is.
4. *If we aren't sure about x's distribution*, we should get a sample of at least n=30.
5. *If we know that x's distribution is very skewed or definitely not normal*, shoot for n=50 or n=100.

Explanation

- If you want the middle 95%, the z is always the same: 1.96 (on either side of μ).

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

For a confidence level (c), the ***critical value z_c*** is the z score such that the area under the curve is between $-z_c$ and z_c.



-1.96 **1.96**

How to Calculate a Confidence Interval

How to do something with all this theory

Steps to Calculating a Confidence Interval When σ is Known

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

Steps to Calculating a Confidence Interval When σ is Known

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

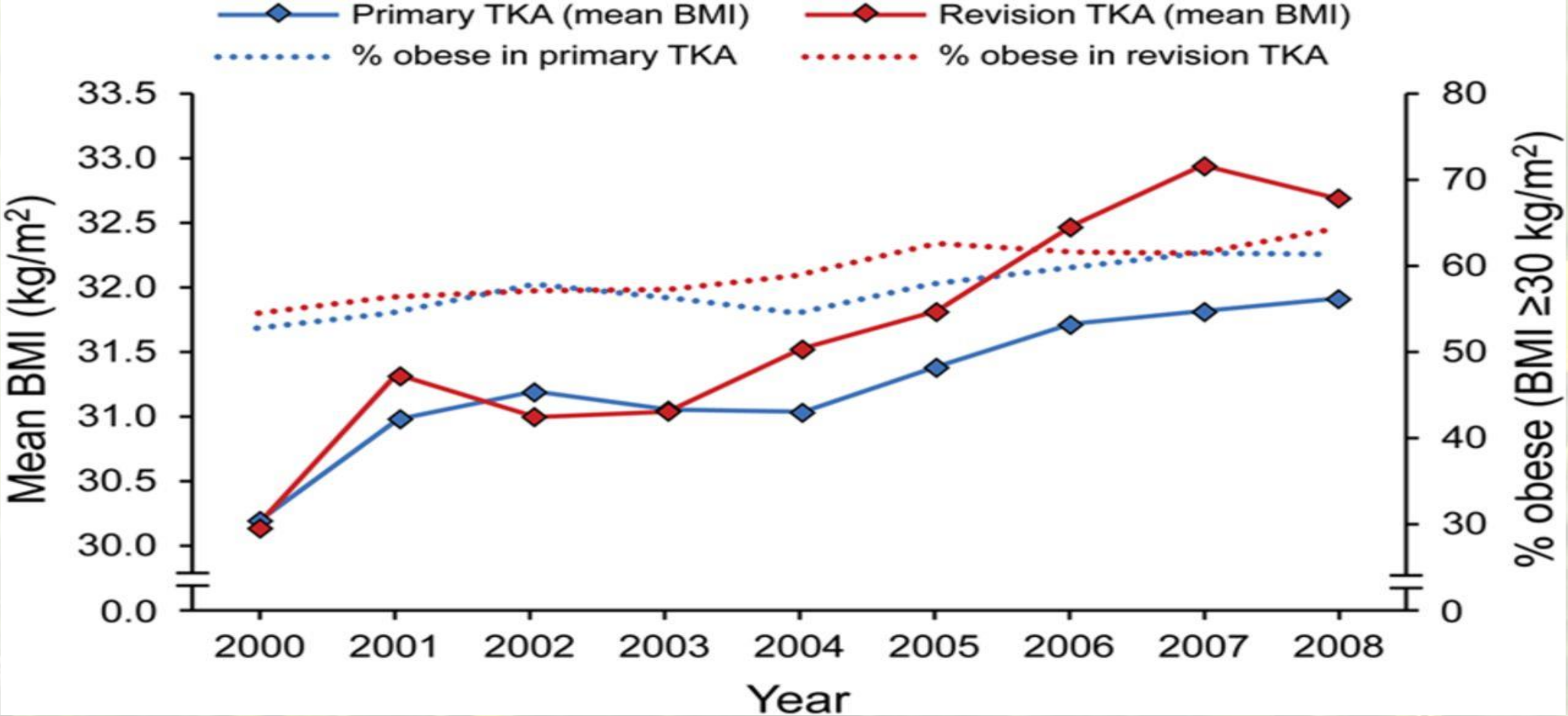
$$E = z_c \frac{\sigma}{\sqrt{n}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

Why would you know the σ, but not know the μ?

- The μ can (and often should) change, but the variation (σ) tends to stay the same.
- Imagine a public health intervention to lower the μ blood pressure. The σ blood pressure would likely stay the same over time.

A line graph showing trends in mean BMI and percentage of obese patients (BMI of ≥30 kg/m2) who underwent total knee arthroplasty (TKA) from 2000 to 2008.



Hilal Maradit Kremers et al. J Bone Joint Surg Am 2014;96:718-724

Steps to Calculating a Confidence Interval When σ is Known

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

$$E = z_c \boxed{\frac{\sigma}{\sqrt{n}}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

Standard Error Fans!

Please note that the formula for E can also be written like this:

$$E = z_c * SE$$

On a quiz, keep track of your SE and reuse it if questions are about different levels of confidence.

Steps to Calculating a Confidence Interval When σ is Known

$\sigma = 14.5$

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

New class of students!

What is their μ?

1. They got an x-bar of 63, we have old σ we can use, and our n=100 again.

Steps to Calculating a Confidence Interval When σ is Known

σ = 14.5

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula: $$E = z_c \frac{\sigma}{\sqrt{n}}$$
4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

New class of students!

What is their μ?

2. Picking 99% for c, so 2.58 as z_c.

Level of Confidence c	Critical value z_c
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

Steps to Calculating a Confidence Interval When σ is Known

σ = 14.5

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

New class of students!

What is their μ?

3. Calculating E.

x-bar = 63
σ = 14.5
n = 100
c = 99%
z_c = 2.58

2.58 * (14.5/√100) = 3.7 = E

Steps to Calculating a Confidence Interval When σ is Known

σ = 14.5

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

New class of students!

What is their μ?

4. Subtract E from x-bar for lower limit

x-bar = 63
σ = 14.5
n = 100
c = 99%
z_c = 2.58
E = 3.7

63 – 3.7 = 59.3

Steps to Calculating a Confidence Interval When σ is Known

σ = 14.5

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:
$$E = z_c \frac{\sigma}{\sqrt{n}}$$
4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

New class of students!

What is their μ?

5. Add E to x-bar for upper limit

$63 + 3.7 = 66.7$

x-bar = 63
σ = 14.5
n = 100
c = 99%
z_c = 2.58
E = 3.7
LL = 59.3

Steps to Calculating a Confidence Interval When σ is Known

$\sigma = 14.5$

1. Make sure you have your x-bar, σ, and n available
2. Pick c, so then you have z_c
3. Calculate the E using this formula:

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

4. Subtract the E from x-bar to get the lower limit for the confidence interval
5. Add the E to the x-bar to get the upper limit for the confidence interval.

New class of students!

What is their μ?

"I am 99% confident the μ of the new class of students is between 59.3 and 66.7".

x-bar = 63
$\sigma = 14.5$
n = 100
c = 99%
$z_c = 2.58$
E = 3.7
LL = 59.3
UL = 66.7

How to Calculate Sample Size

Rearrangement of the equation

Steps to Calculating Sample Size When σ is Known

1. Make sure you have your σ available.
2. Pick c, so then you have z_c
3. Pick how big of an E you want
4. Calculate the n (sample size) using this formula:

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

Level of Confidence c	Critical value z_c
0.80, or 80%	1.28
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

Steps to Calculating Sample Size When σ is Known

σ = 14.5

1. Make sure you have your σ available.
2. Pick c, so then you have z_c
3. Pick how big of an E you want
4. Calculate the n (sample size) using this formula:

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

How many students do I need if:

- I want a 95% confidence interval, and
- I want my E = 5

Steps to Calculating Sample Size When σ is Known

$\sigma = 14.5$

1. Make sure you have your σ available.
2. Pick c, so then you have z_c
3. Pick how big of an E you want
4. Calculate the n (sample size) using this formula:

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

How many students do I need if:

- I want a 95% confidence interval, and
- I want my E = 5

$\sigma = 14.5$
$c = 95\%$
$z_c = 1.96$
$E = 5$

Steps to Calculating Sample Size When σ is Known

$\sigma = 14.5$

1. Make sure you have your σ available.
2. Pick c, so then you have z_c
3. Pick how big of an E you want
4. Calculate the n (sample size) using this formula:

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

How many students do I need if:

- I want a 95% confidence interval, and
- I want my E = 5

$\sigma = 14.5$
$c = 95\%$
$z_c = 1.96$
$E = 5$

$[(1.96*14.5)/5]^2 = 32.3$

Need to always round up due to not being able to have 0.3 of a person.

Answer: n=33

Conclusion

- Reviewed some slides from Chapter 7 that relate to the subject of confidence intervals
- Described z_c, level of confidence, point estimate
- Demonstrated how to construct a confidence interval when the σ is known
- Demonstrated how to calculate sample size.

Image by Shmuel Shahar in the public domain