Unit 3 - Estimation

Confidence Interval Estimation

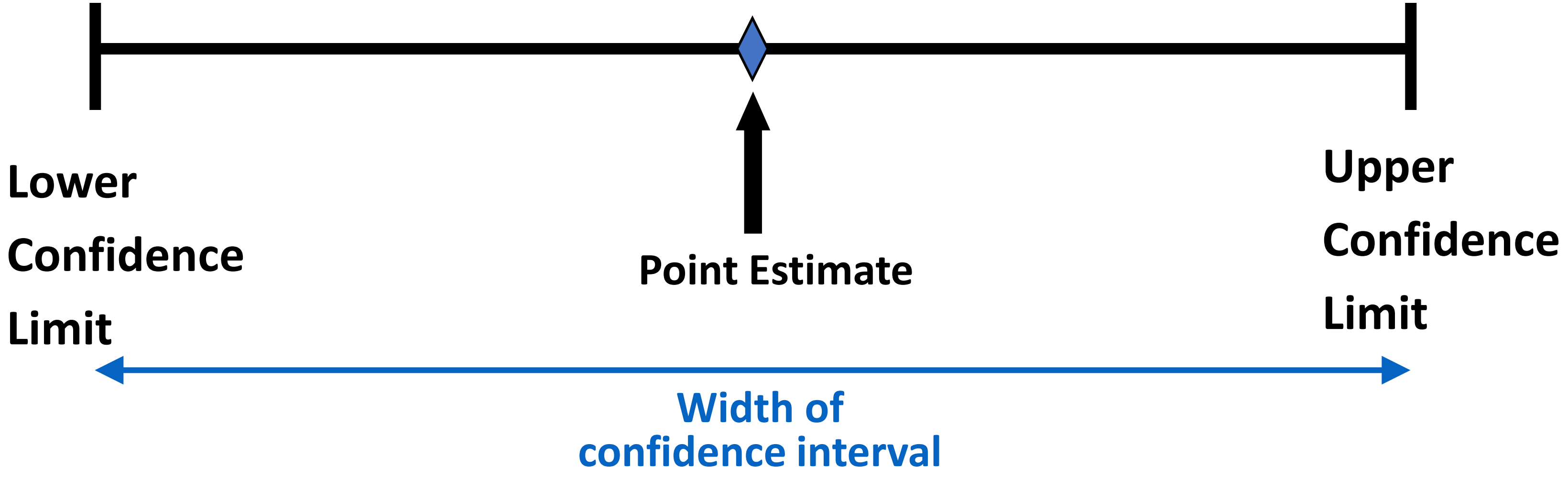
Learning Objectives

In this unit, you learn:

- To construct and interpret confidence interval estimates for the Population Mean, μ
  - when Population Standard Deviation σ is Known
  - when Population Standard Deviation σ is Unknown

Point and Interval Estimates

- A point estimate is a single number,
- a confidence interval provides additional information about variability



Point Estimates

We can estimate a Population Parameter …		with a Sample Statistic (a Point Estimate)
Mean	μ	$\bar{X}$
Proportion	π	p

- How much uncertainty is associated with a point estimate of a population parameter?

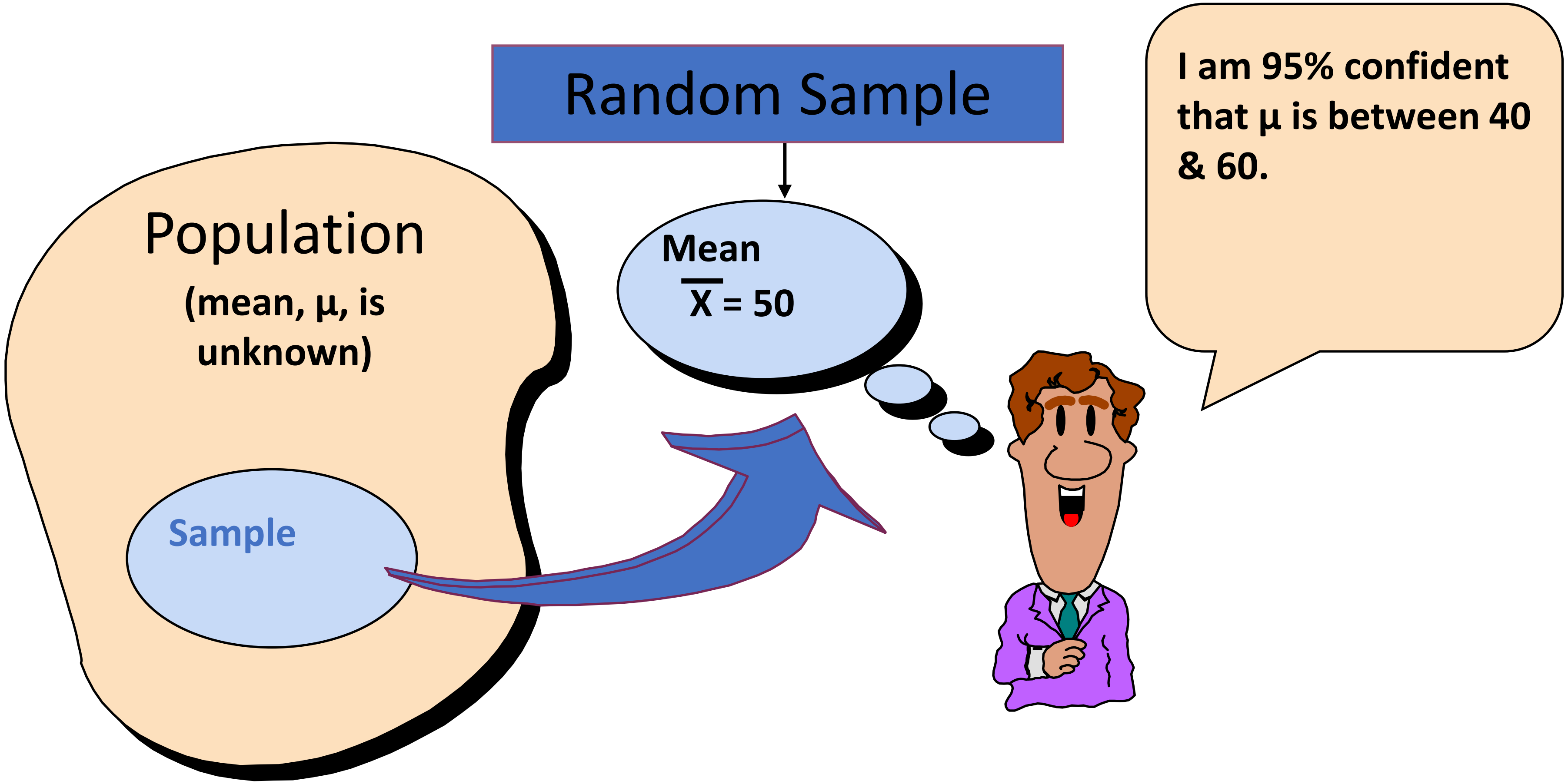
- An interval estimate provides more information about a population characteristic than does a point estimate

- Such interval estimates are called confidence intervals

Confidence Interval Estimate

- An interval gives a range of values:
  - Takes into consideration variation in sample statistics from sample to sample
  - Based on observations from 1 sample
  - Gives information about closeness to unknown population parameters
  - Stated in terms of level of confidence
    - Can never be 100% confident

Estimation Process



General Formula

- The general formula for all confidence intervals is:

Point Estimate ± (Critical Value)(Standard Error)

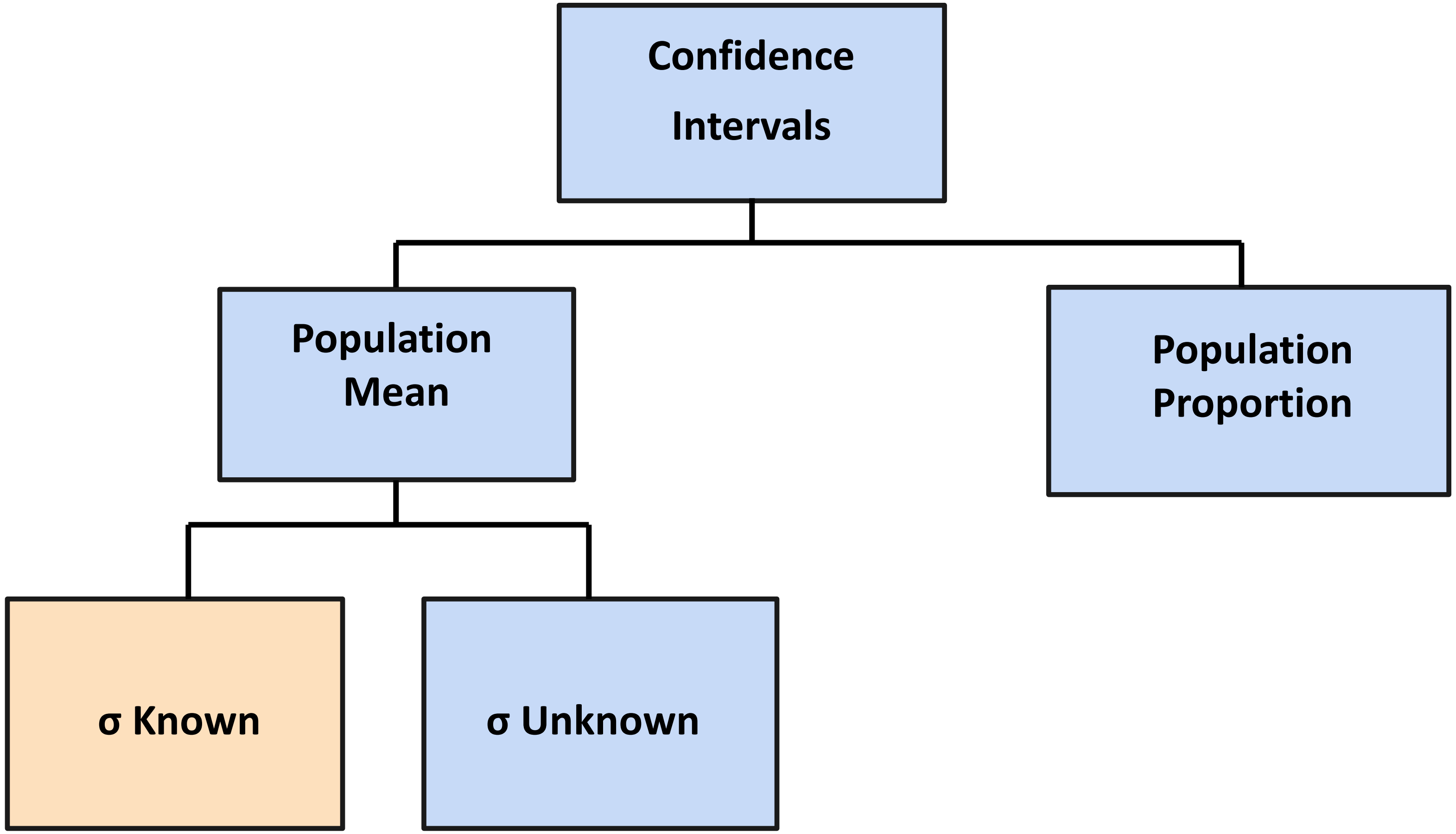
- Confidence Level
  - Confidence for which the interval will contain the unknown population parameter
- A percentage (less than 100%)

Confidence Level, $(1 - \alpha)$

(continued)

- Suppose confidence level = 95%
- Also written $(1 - \alpha) = 0.95$
- A relative frequency interpretation:
  - In the long run, 95% of all the confidence intervals that can be constructed will contain the unknown true parameter
- A specific interval either will contain or will not contain the true parameter
  - No probability involved in a specific interval

Confidence Intervals



Confidence Interval for μ (σ known)

- Assumptions
  - Population standard deviation σ is known
  - Population is normally distributed
  - If population is not normal, use large sample
- Confidence interval estimate:

$$\bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

where $\bar{X}$ is the point estimate

Z is the normal distribution critical value for a probability of $\alpha/2$ in each tail

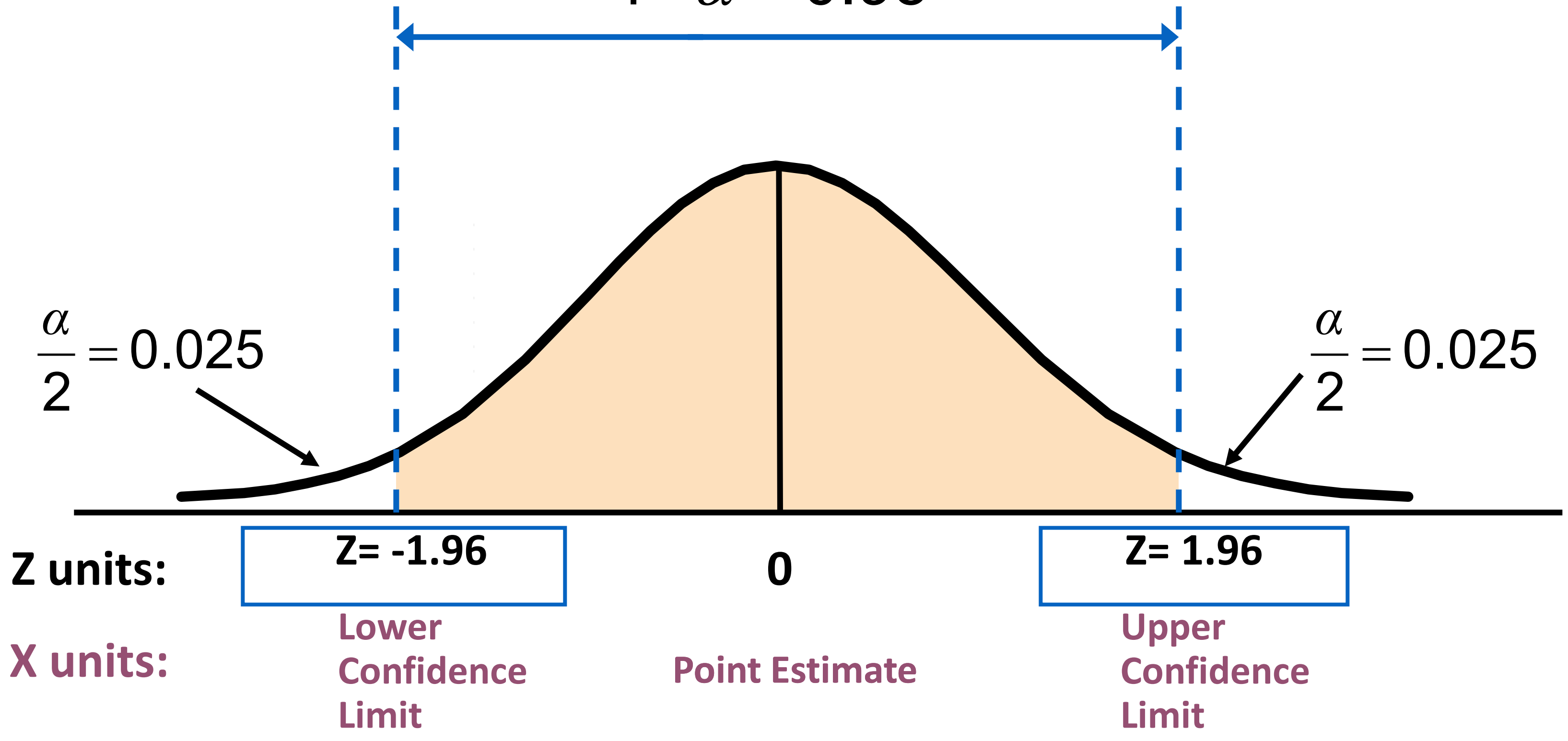
$\sigma/\sqrt{n}$ is the standard error

Finding the Critical Value, Z

$$Z = \pm 1.96$$

- Consider a 95% confidence interval:

$$1 - \alpha = 0.95$$

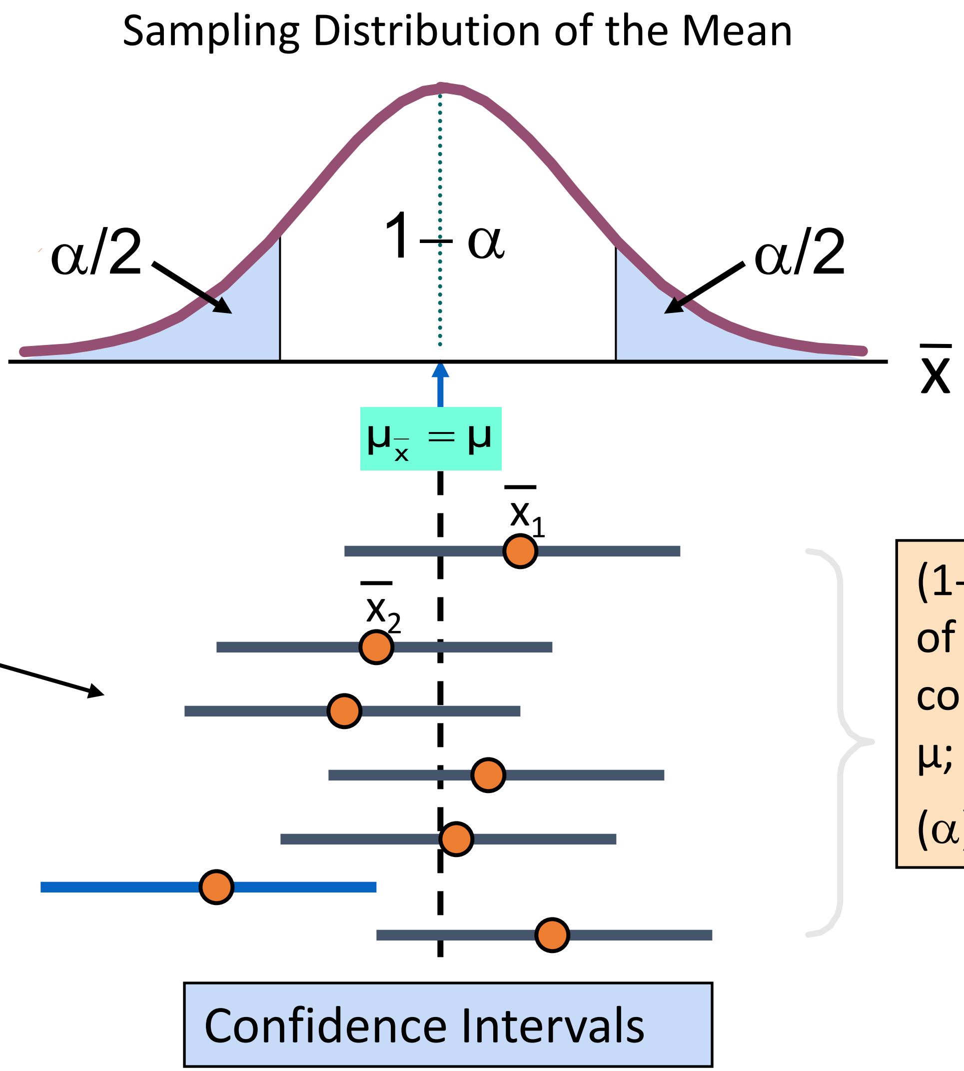


Common Levels of Confidence

- Commonly used confidence levels are 90%, 95%, and 99%

Confidence Level	***Confidence Coefficient,*** $1-\alpha$	***Z value***
80%	**0.80**	**1.28**
90%	**0.90**	**1.645**
95%	**0.95**	**1.96**
98%	**0.98**	**2.33**
99%	**0.99**	**2.58**
99.8%	**0.998**	**3.08**
99.9%	**0.999**	**3.27**

Intervals and Level of Confidence



Intervals extend from

$$\bar{X} + Z\frac{\sigma}{\sqrt{n}}$$

to

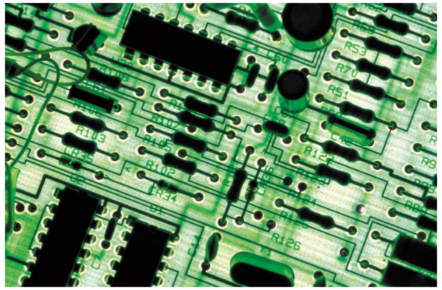
$$\bar{X} - Z\frac{\sigma}{\sqrt{n}}$$

$(1-\alpha)$x100% of intervals constructed contain μ;

(α)x100% do not.

- A sample of 11 circuits from a large normal population has a mean resistance of 2.20 ohms. We know from past testing that the population standard deviation is 0.35 ohms.

- Determine a 95% confidence interval for the true mean resistance of the population.



Confidence Interval Estimations

Example

(continued)

- A sample of 11 circuits from a large normal population has a mean resistance of 2.20 ohms. We know from past testing that the population standard deviation is 0.35 ohms.

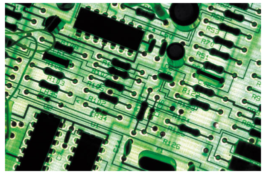
- Solution:

$$\bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

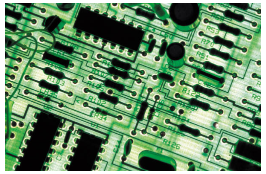
$$= 2.20 \pm 1.96\,(0.35/\sqrt{11})$$

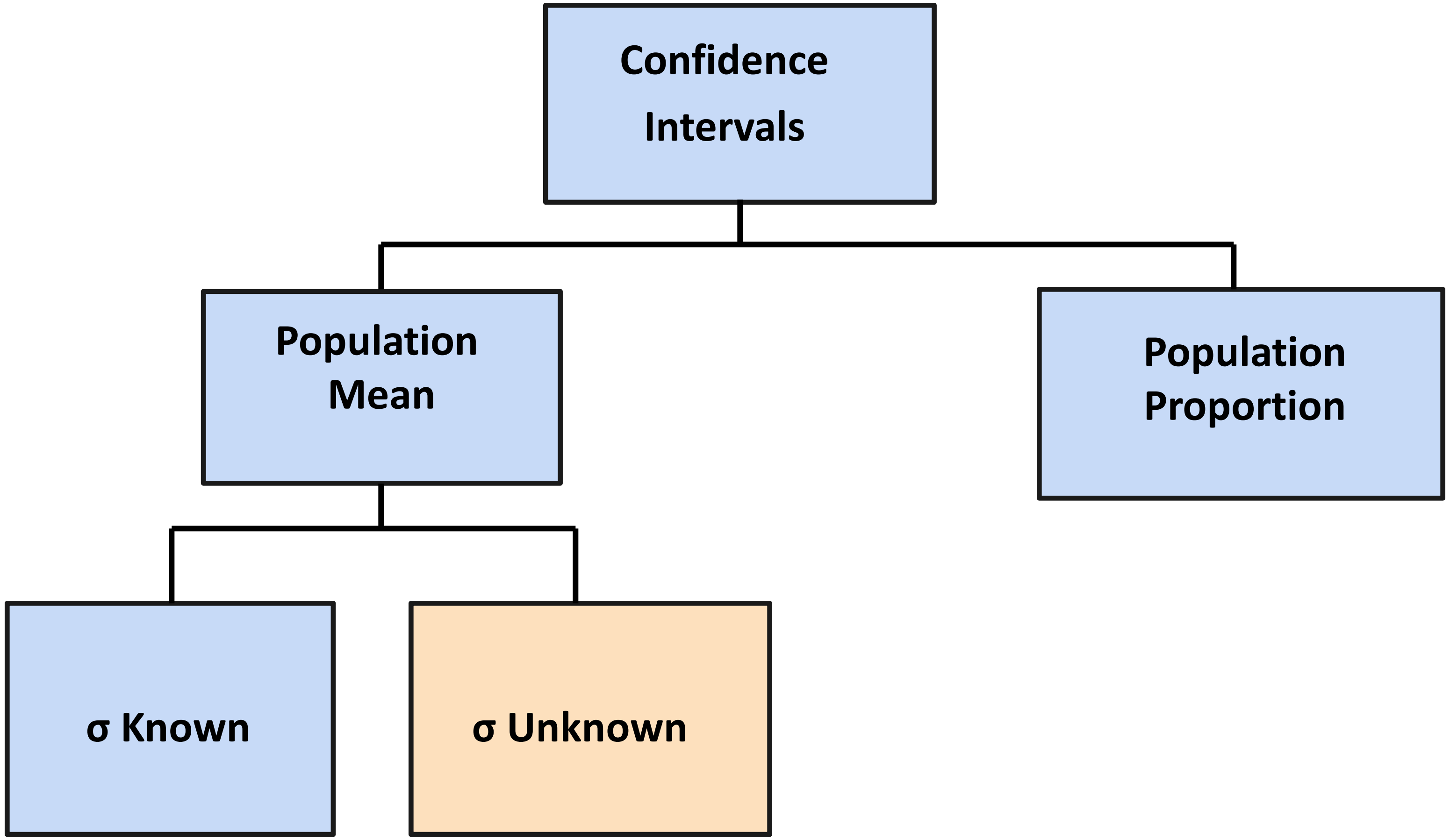
$$= 2.20 \pm 0.2068$$

$$1.9932 \le \mu \le 2.4068$$



- We are 95% confident that the true mean resistance is between 1.9932 and 2.4068 ohms
- Although the true mean may or may not be in this interval, 95% of intervals formed in this manner will contain the true mean




Confidence Intervals
Population Mean
Population Proportion
σ Known
σ Unknown

Confidence Interval for μ (σ unknown)

- If the population standard deviation σ is unknown, we can substitute the sample standard deviation, S
- This introduces extra uncertainty, since S is variable from sample to sample
- So we use the t distribution instead of the normal distribution

Confidence Interval for μ (σ unknown)

(continued)

- Assumptions
  - Population standard deviation is unknown
  - Population is normally distributed
  - If population is not normal, use large sample
- Use Student's t Distribution
- Confidence Interval Estimate:

$$\bar{X} \pm t_{n-1} \frac{S}{\sqrt{n}}$$

(where t is the critical value of the t distribution with n -1 degrees of freedom and an area of α/2 in each tail)

- The t is a family of distributions
- The t value depends on degrees of freedom (d.f.)
  - Number of observations that are free to vary after sample mean has been calculated

$$d.f. = n - 1$$

Degrees of Freedom (df)

Idea: Number of observations that are free to vary after sample mean has been calculated

Example: Suppose the mean of 3 numbers is 8.0

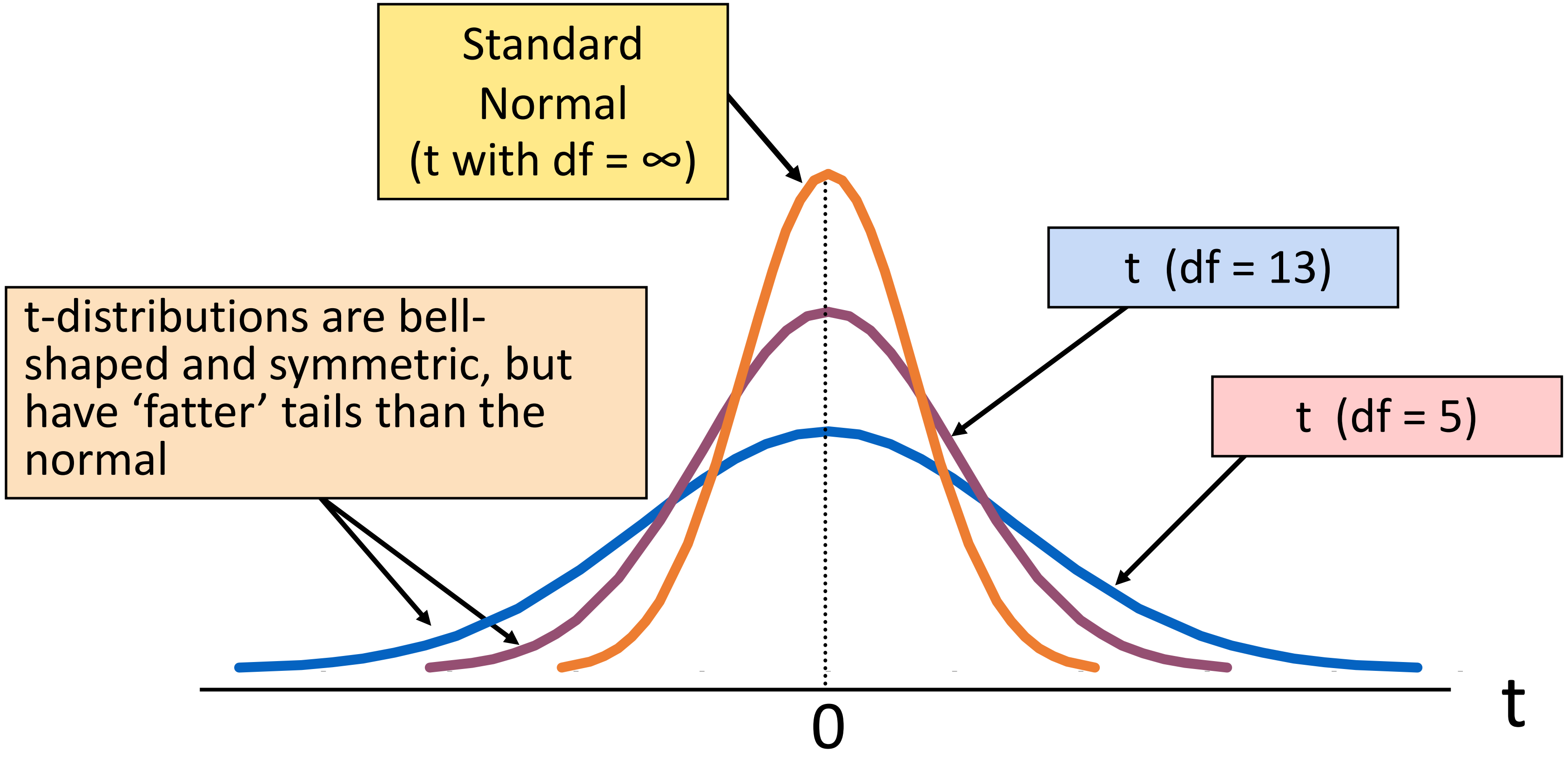
Let $X_1 = 7$
Let $X_2 = 8$
What is X_3?

If the mean of these three values is 8.0, then X_3 must be 9 (i.e., X_3 is not free to vary)

Here, n = 3, so degrees of freedom = n – 1 = 3 – 1 = 2

(2 values can be any numbers, but the third is not free to vary for a given mean)

Note: $t \longrightarrow Z$ as n increases

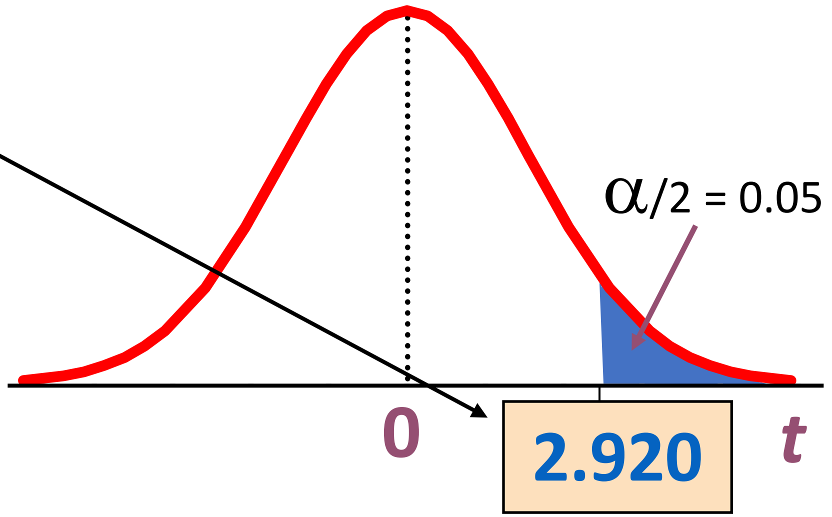


Student's t Table

	Upper Tail Area		
df	.25	.10	**.05**
1	1.000	3.078	6.314
2	0.817	1.886	**2.920**
3	0.765	1.638	2.353

The body of the table contains t values, not probabilities

Let: $n = 3$
$df = n - 1 = 2$
$\alpha = 0.10$
$\alpha/2 = 0.05$



UNIVERSITI TEKNOLOGI MARA

With comparison to the Z value

Confidence Level	t (10 d.f.)	t (20 d.f.)	t (30 d.f.)	Z
0.80	1.372	1.325	1.310	1.28
0.90	1.812	1.725	1.697	1.645
0.95	2.228	2.086	2.042	1.96
0.99	3.169	2.845	2.750	2.58

Note: $t \longrightarrow Z$ as n increases

Example

A random sample of n = 25 has $\bar{X} = 50$ and S = 8. Form a 95% confidence interval for μ

- d.f. = n – 1 = 24, so $t_{\alpha/2,\, n-1} = t_{0.025,24} = 2.0639$

The confidence interval is

$$\bar{X} \pm t_{\alpha/2,\, n-1} \frac{S}{\sqrt{n}} = 50 \pm (2.0639) \frac{8}{\sqrt{25}}$$

$46.698 \le \mu \le 53.302$

Ethical Issues

- A confidence interval estimate (reflecting sampling error) should always be included when reporting a point estimate
- The level of confidence should always be reported
- The sample size should be reported
- An interpretation of the confidence interval estimate should also be provided

Summary

- Introduced the concept of confidence intervals
- Discussed point estimates
- Developed confidence interval estimates
- Created confidence interval estimates for the mean (σ known)
- Determined confidence interval estimates for the mean (σ unknown)
- Addressed confidence interval estimation and ethical issues