

Inferential Statistics

Dr. Wan Nor Arifin

Unit of Biostatistics and Research Methodology, Universiti Sains Malaysia.



Outlines

- Estimation
- Hypothesis Testing

Overview

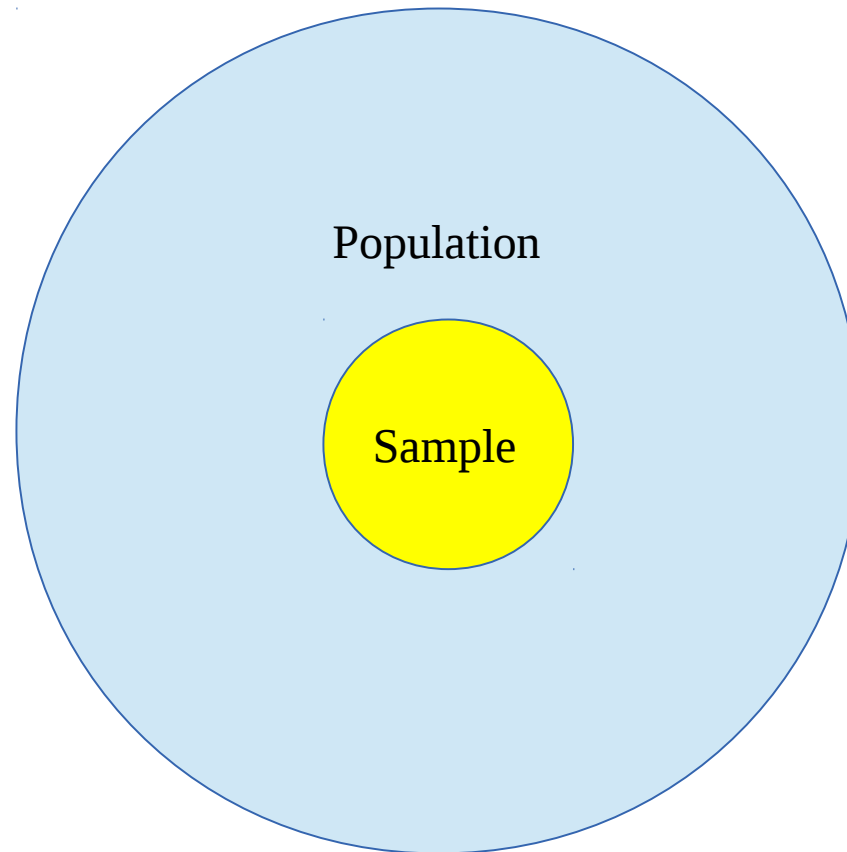
- Statistics?
- Population vs sample?
- Inference?

Overview

- Statistics is a field of study dealing with (Daniel, 1995):
 1. Collection, organization, summarization and analysis of data.
 2. Making inference/conclusion about population data from sample data.

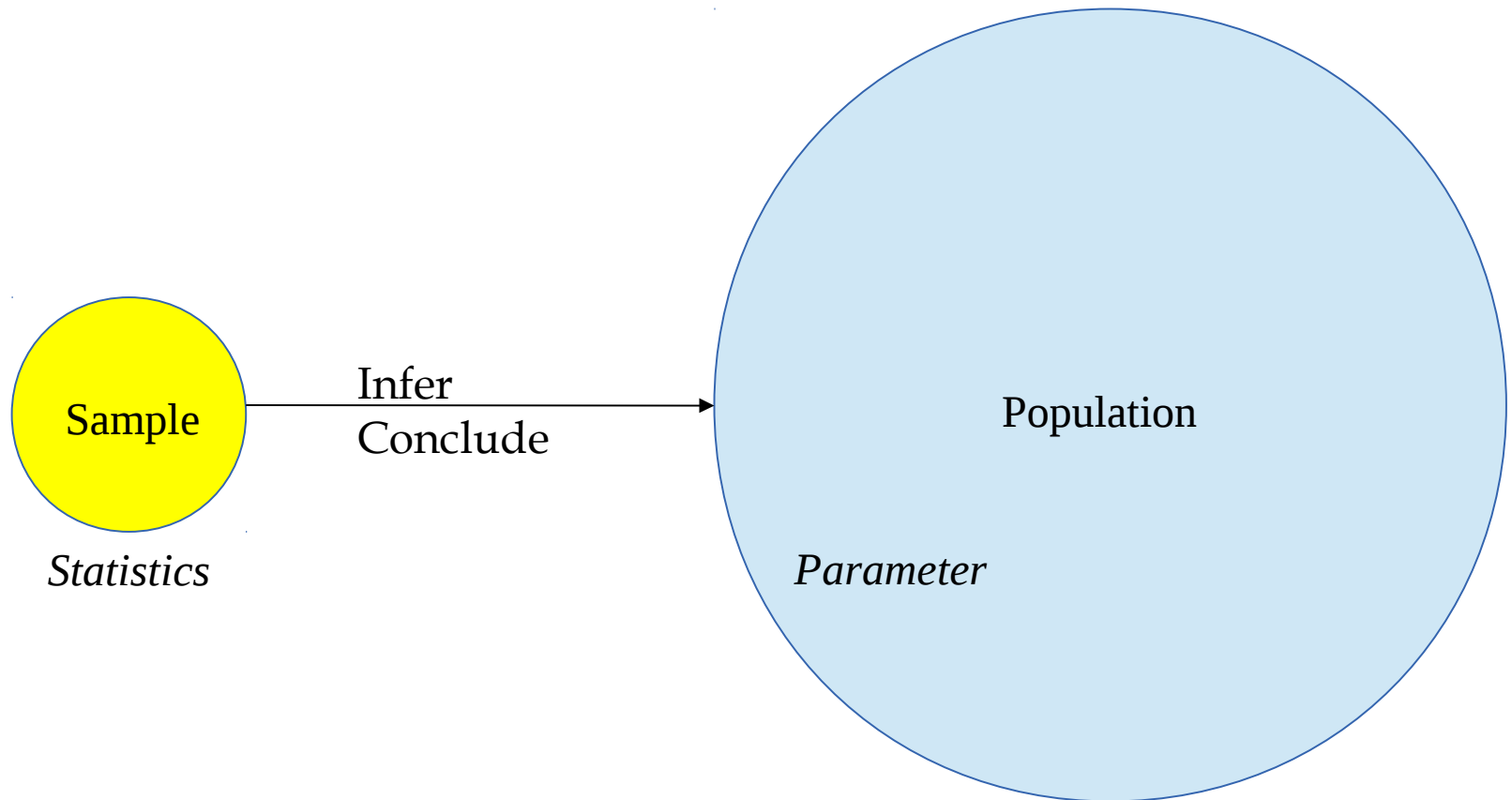
Overview

- Population vs sample

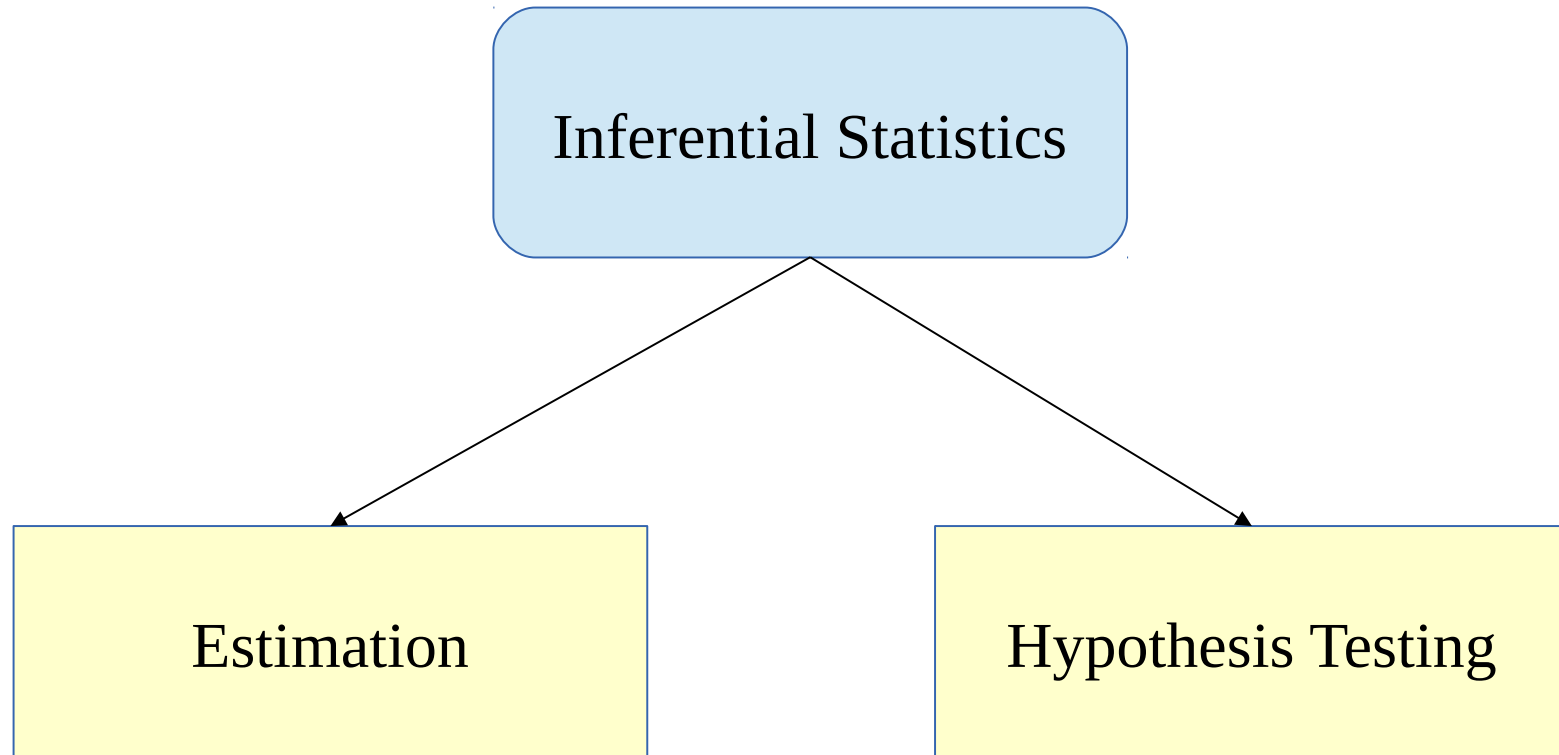


Overview

- Inference:



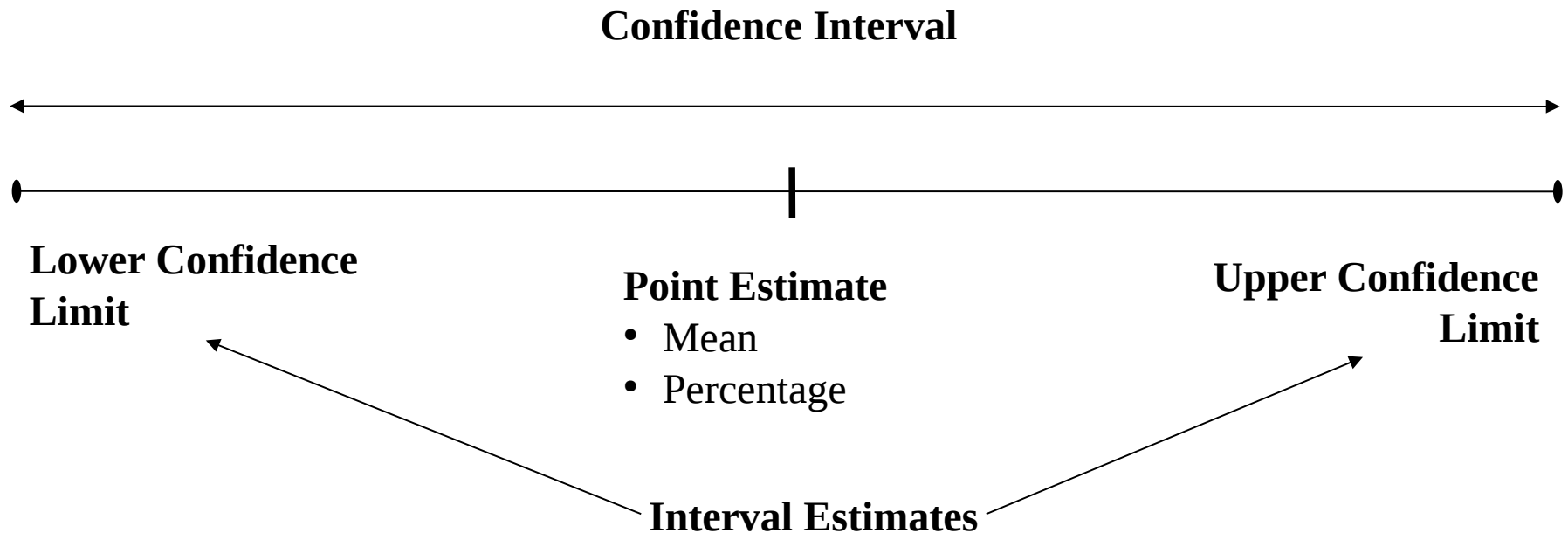
Overview



1. Estimation

1. Estimation

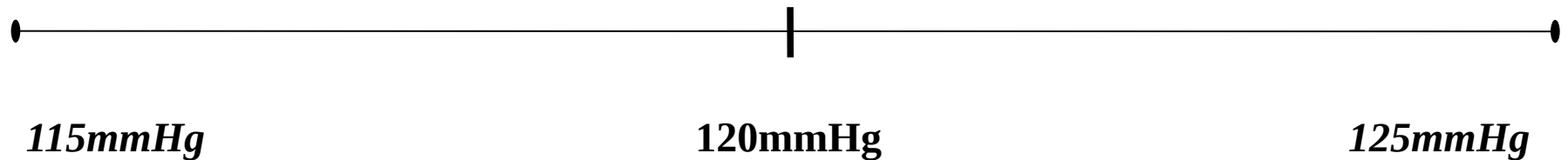
- Usually for One Sample → One Population
- Estimate *parameter* by



1. Estimation

Mean SBP for Normal population

95% Confidence Interval



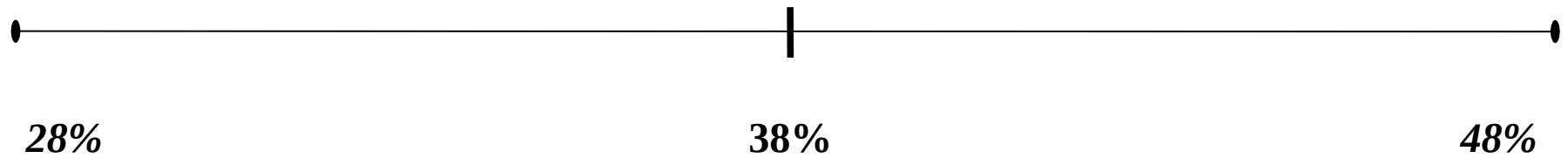
Interpretation: Based on a *sample* of 30 subjects, I am 95% sure that mean SBP of normal *population* is between 115mmHg to 125mmHg. The sample mean is 120mmHg.

Reporting: 120mmHg (95% CI: 115mmHg, 125mmHg)

1. Estimation

Percentage of Obesity among University Students' population

95% Confidence Interval



Interpretation: Based on a *sample* of 100 subjects, I am 95% sure that percentage of obesity of university students' *population* is between 28% to 48%. The sample percentage is 38%.

Reporting: 38% (95% CI: 28%, 48%)

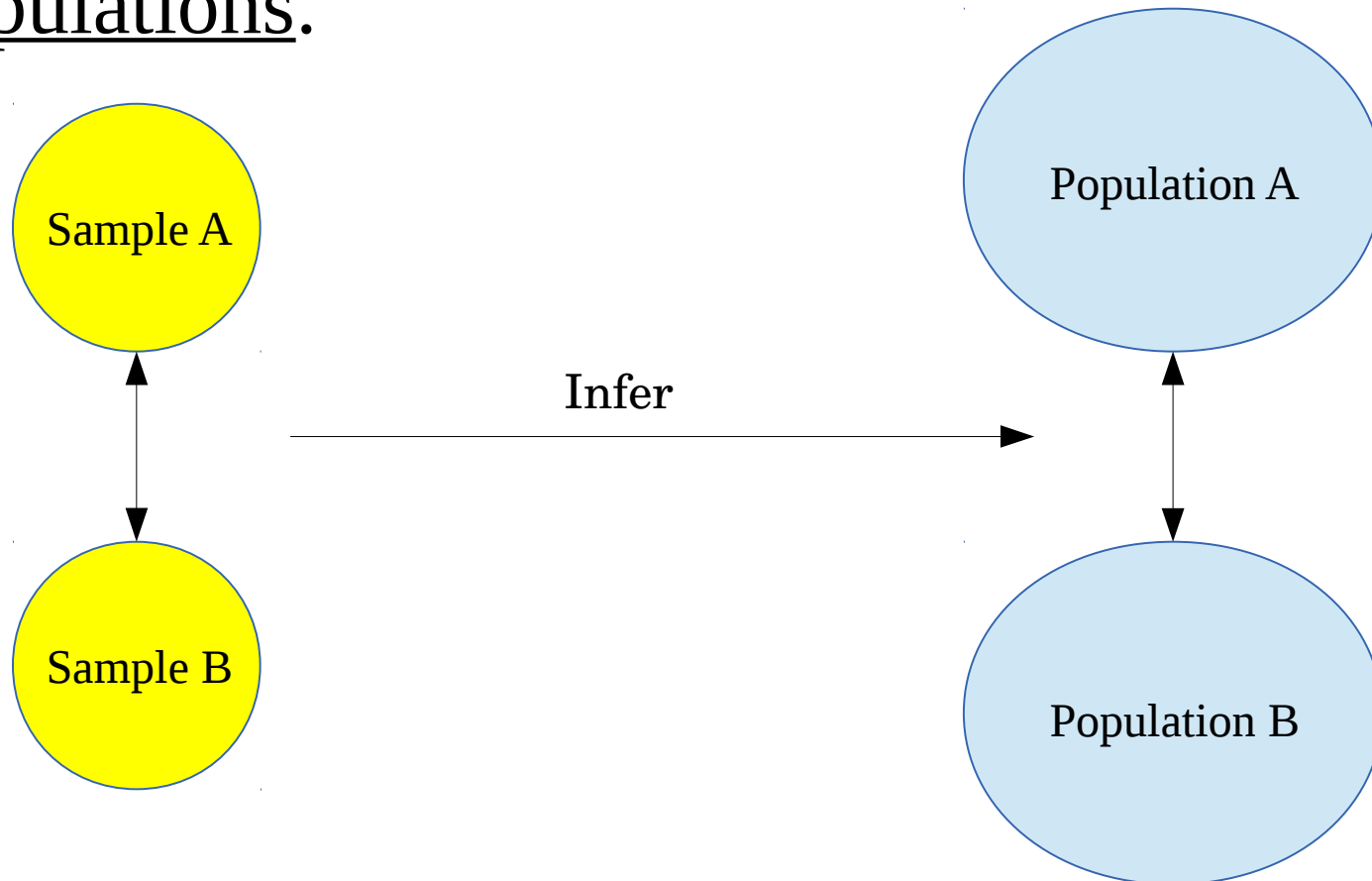
1. Estimation

- **Interval estimates** values depend on ***Confidence level*** (90%, 95%, 99%), ***sample size*** and ***standard deviation*** → Precision.
- Calculation*? Given in SPSS output. It is important to know the interpretation.

2. Hypothesis Testing

2. Hypothesis Testing

- Usually for comparison of samples → comparison of populations.



2. Hypothesis Testing

- Stated in form of **Statistical Hypothesis** → Can be tested with statistical test.

Alternative Hypothesis:

Population A is different from Population B

Null Hypothesis:

Population A is similar to Population B

2. Hypothesis Testing

- **P-value** – Probability that the difference is merely by chance → Calculated from statistical test.
- Set acceptable level so called “chance” → **Significance level, α (0.05, 0.01, 0.001)**

Alternative Hypothesis:

$$P\text{-value} \leq \mathbf{0.05}$$

Null Hypothesis:

$$P\text{-value} > \mathbf{0.05}$$

2. Hypothesis Testing

Alternative Hypothesis:
Population A is different from
Population B

Null Hypothesis:
Population A is similar to
Population B

Statistical Test 

Alternative Hypothesis:
P-value \leq **0.05**

Null Hypothesis:
P-value $>$ **0.05**

2. Hypothesis Testing

Comparing **Mean SBP of Medical Students' population** vs **Lecturers' population**

Alternative Hypothesis:
Mean SBP of MS population
is different from L population

Null Hypothesis:
No difference in Mean SBP
between the populations

Statistical Test

Alternative Hypothesis:
P-value \leq **0.05**

Null Hypothesis:
P-value $>$ **0.05**

Independent t-test

2. Hypothesis Testing

Comparing **Obesity % of Medical Students' population** vs **Lecturers' population**

Alternative Hypothesis:
Obesity % among MS
population is different from L
population

Null Hypothesis:
No difference in Obesity %
between the populations

Statistical Test

Alternative Hypothesis:
P-value \leq **0.05**

Null Hypothesis:
P-value $>$ **0.05**

Chi-squared test

Outcomes

- ✓ Understand basic concept of confidence interval.
- ✓ Able to interpret confidence interval.
- ✓ Understand basic concept of hypothesis testing.
- ✓ Able to interpret P-value.
- ✓ Understand concept of significance level.

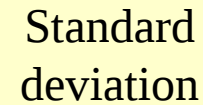
Reference

Daniel, W. W. (1995). *Biostatistics: A foundation for analysis in the health sciences* (6th ed.). USA: John Wiley & Sons.

* 1. Estimation

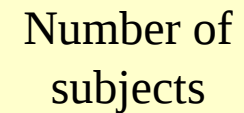
- Interval estimates (mean):

Standard deviation



$$\text{Lower confidence limit} = \text{Point Estimate} - \text{Reliability Coefficient} \times \frac{SD}{\sqrt{n}}$$

Number of subjects



$$\text{Upper confidence limit} = \text{Point Estimate} + \text{Reliability Coefficient} \times \frac{SD}{\sqrt{n}}$$

* 1. Estimation

- Interval estimates (proportion):

p = proportion

$$\text{Lower confidence limit} = \text{Point Estimate} - \text{Reliability Coefficient} \times \sqrt{\frac{p(1-p)}{n}}$$

Number of subjects

$$\text{Upper confidence limit} = \text{Point Estimate} + \text{Reliability Coefficient} \times \sqrt{\frac{p(1-p)}{n}}$$

* 1. Estimation

- Reliability Coefficient:

Confidence level	Reliability coefficient
90%	1.65
95%	1.96
99%	2.56