Sample size Estimation

Select an area of interest: Select a topic

Collect and evaluate existing information

Research question: Set aim and objectives

Set a relevant study design

Define your study subjects

Develop the tool for data collection

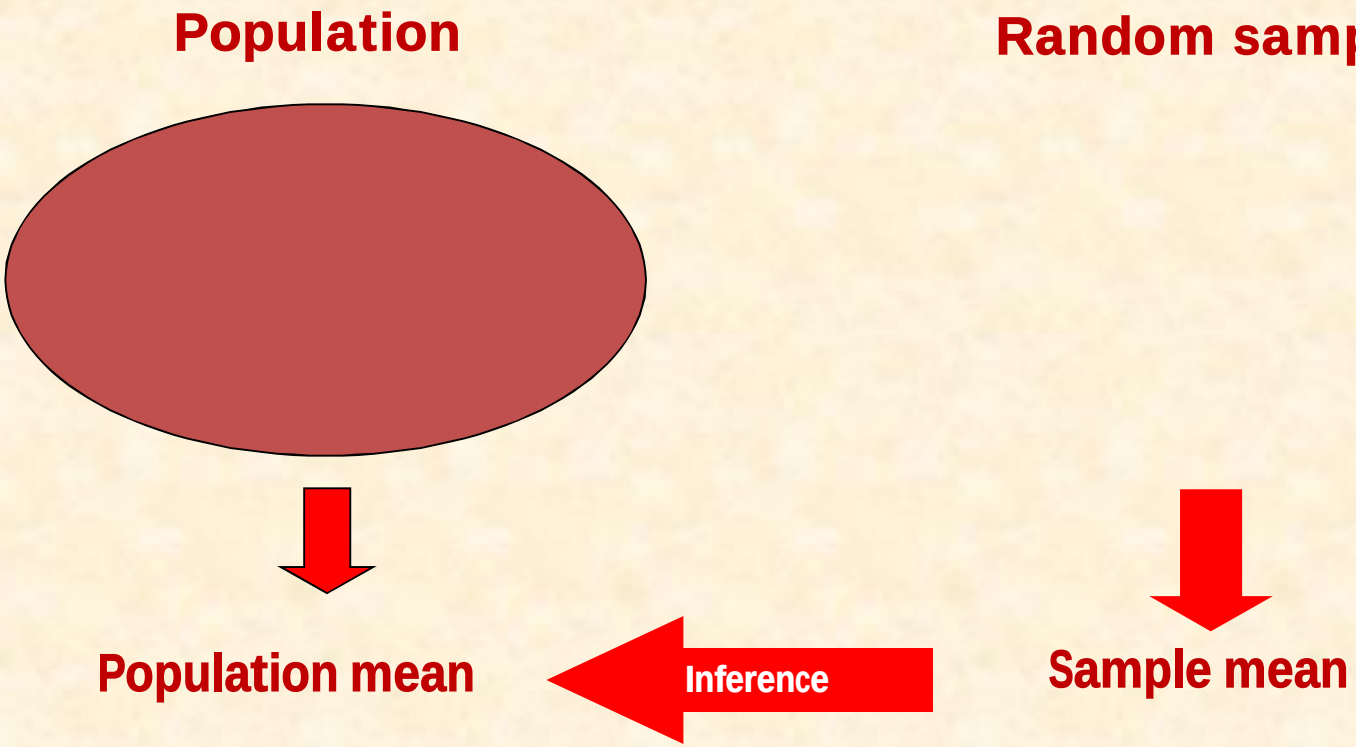
Collect the data

Data analysis

Write the report

Dissemination as publication

Population *vs.* sample



Census enumeration:

When an investigation is carried out on the entire population.

Sample enumeration:

When an investigation is carried out on the representative part of that population.

Large sample size

Advantages

- Estimated Characteristic of interest approaches that of population
- Provide better estimates
- More confidence level
- Smaller test errors

Disadvantages

- Expensive
- Loss of time
- Waste of manpower
- More support facilities
- Ethical issues

Small sample size

Advantages

- Less expensive
- Less time consuming
- Man power and support facilities requirement are less

Disadvantages

- Unable to detect clinically important results
- Efforts may go waste
- Test errors will be more

Main research question, the outcome measure and the statistical procedure

Main Research question	Statistical Procedure
• How common is osteoporosis among post-menopausal women?	Estimation of proportion
• What is the average haemoglobin level among Mysuru population?	Estimation of mean
• Difference in knowledge regarding rabies among people residing in two villages	Difference in two proportions
• Is treatment A better than treatment B in blood pressure reduction?	Difference in two Means

Desired Precision Level (d)

- **It is the range in which the true population value is estimated to be.**
- **Extent of error you will allow in the estimate**

Eg:

65% couples adopted a particular type of family planning method with 5% precision – 55-65% have adopted that particular family planning method

Estimation

Formulae

&

Problems

Scenario 1

From a pilot study it was reported that among headache patients 28% had vascular headache. It was decided to have 95% C.I and 10% variability. How many patients are necessary to conduct the study.

Outcome : Proportion (p) is the parameter of our study

$$n = \frac{Z_{\alpha}^{2} * p * q}{d^{2}}$$

Where

Z = Standardized Normal deviate (Z value)

p = Proportion or Prevalence of interest.

q = 1- p

d = Clinically expected variation.

α = Level of significance

P = 28%

q = 72%

Z_α= 1.96 for α at 0.05

d = 10% of 28% = 2.8

$$n = \frac{(1.96)^2 * 28 * 72}{(10)^2} = 77$$

Scenario 2

In a Health Survey of schoolchildren it is found that the mean hemoglobin level of 55 boys is 10.2/100ml with a standard deviation of 2.1. Consider the precision as 0.8.

Outcome: mean is the parameter of our study

$$n = \frac{Z_{\alpha/2}^{2} * S^{2}}{d^{2}}$$

Z = Standardized Normal deviate (Z value)

S = Sample standard deviation

d = Clinically expected variation

α = Level of Significance

Mean = 10.2

SD= 2.1

Z_{α}= 1.96 for α at 0.05

d = 0.8

$$n = \frac{(1.96)^2 * 2.1^2}{(0.8)^2} = 26$$

Scenario 3

	Mortality		
	Dead	**Alive**	**Total**
High Dose	**4 (16.7)**	**20**	**24**
Low Dose	**1 (4.2)**	**23**	**24**
Total	**5 (10.4)**	**43**	**48**

Outcome: Two proportions is the parameter of our study

$$n = \frac{(Z_{\alpha} + Z_{\beta})^2 \; p * q * 2}{d^2}$$

Where

Z_{α} = Z value for α level Z_{β} = Z value for β level

p = average percentage between two groups

q = 100 - p

d = Clinically meaningful difference between two groups

P = (16.7+4.2)/2=10.4%

q = 89.6%

Z_α= 1.96 for α at 0.05 Z_β = 1.282 for β at 0.10

$$n = \frac{(1.96+1.282)^2 * 10.4 * 89.6 * 2}{(12.5)^2} = 125$$ **in each arm**

Scenario 4

Duration of ICU study (in hours):

	Low Dose	High Dose
Mean	61.7	33.2
SD	93.8	89.1
n	23	20

Outcome: Comparison of two means is the parameter of our study

$$n = \frac{(Z_{\alpha} + Z_{\beta})^2 * S^2 * 2}{d^2}$$

Where

Z_{α} = Z value for α error

Z_{β} = Z value for β error

S = Common standard deviation between two groups

d = Clinically meaningful difference

$$\alpha = 0.05 \qquad \beta = 0.10$$

$$\mathrm{n} = \frac{(1.96 + 1.282)^2 * 2 * 90^2}{28.5^2} = 210$$

Scenario 5

From pilot study, it was observed that Before and After treatment of hypertensive patients with Propranolol the average of blood pressure estimated to be 120 and 80, respectively with clinically significant difference of 40mmHg, if SD = 38, calculate sample size.

Outcome: Paired observations

$$n = 2 + \frac{(Z_{\alpha/2} + Z_{\beta})^2 \times S^2}{d^2}$$

S: Common Standard deviation

D : Clinically meaningful difference

$$n = 2 + \frac{(Z_{\alpha/2} + Z_{\beta})^2 \times S^2}{d^2}$$

$$n = 2 + 7.58\left(\frac{38}{40}\right)^2$$

9 patients in each group

Check list for sample size determination

- ✓ What is the primary outcome measure?
- ✓ What difference in outcome measure is clinically significant?
- ✓ What would be the level of significance or level of confidence?
- ✓ What is the power?
- ✓ Which are the other constraints?

Thank you