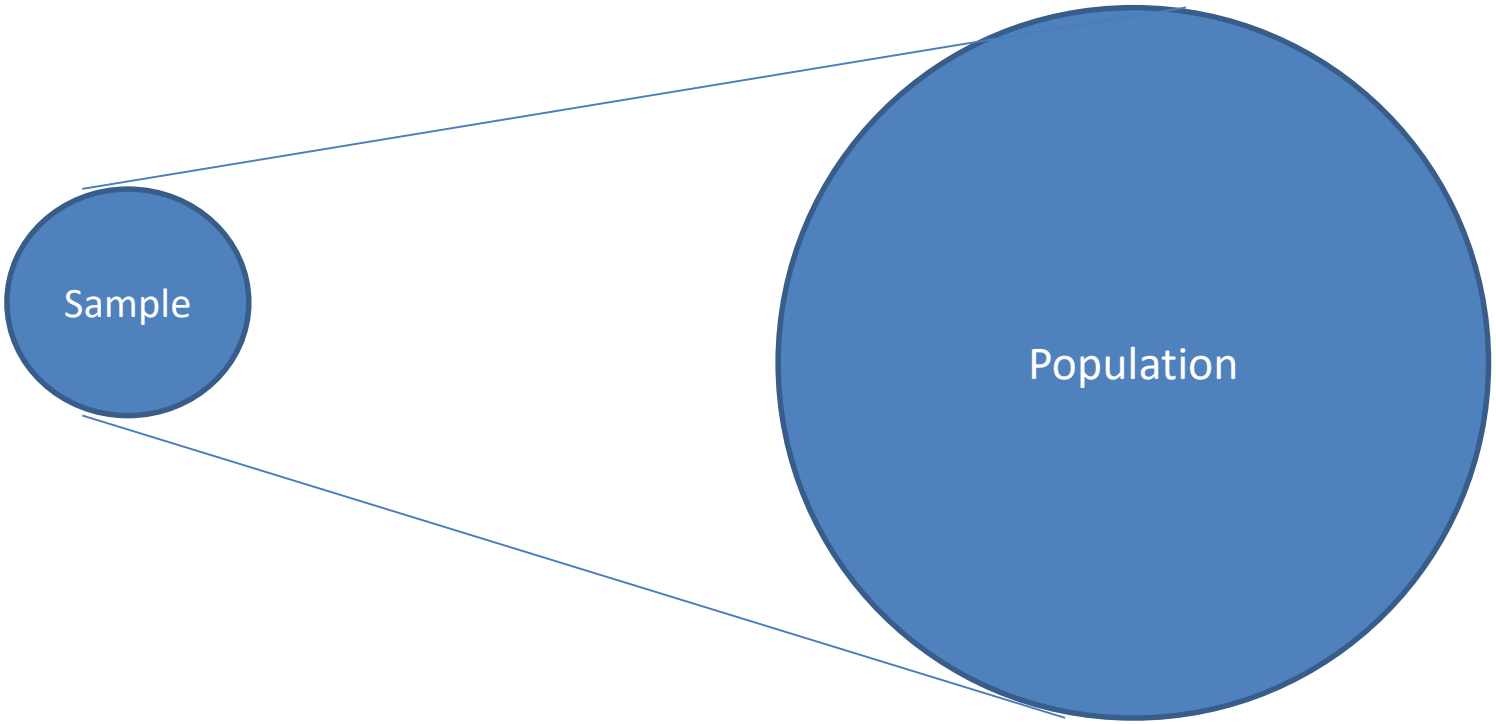
Sample Size Calculation in Biomedical Research

Dr.G.Ravi Prabhu

Prof & HOD

Community Medicine, SVMC, Tirupati

Sample and Population



Sample characteristics

- Adequate Size
- Representative

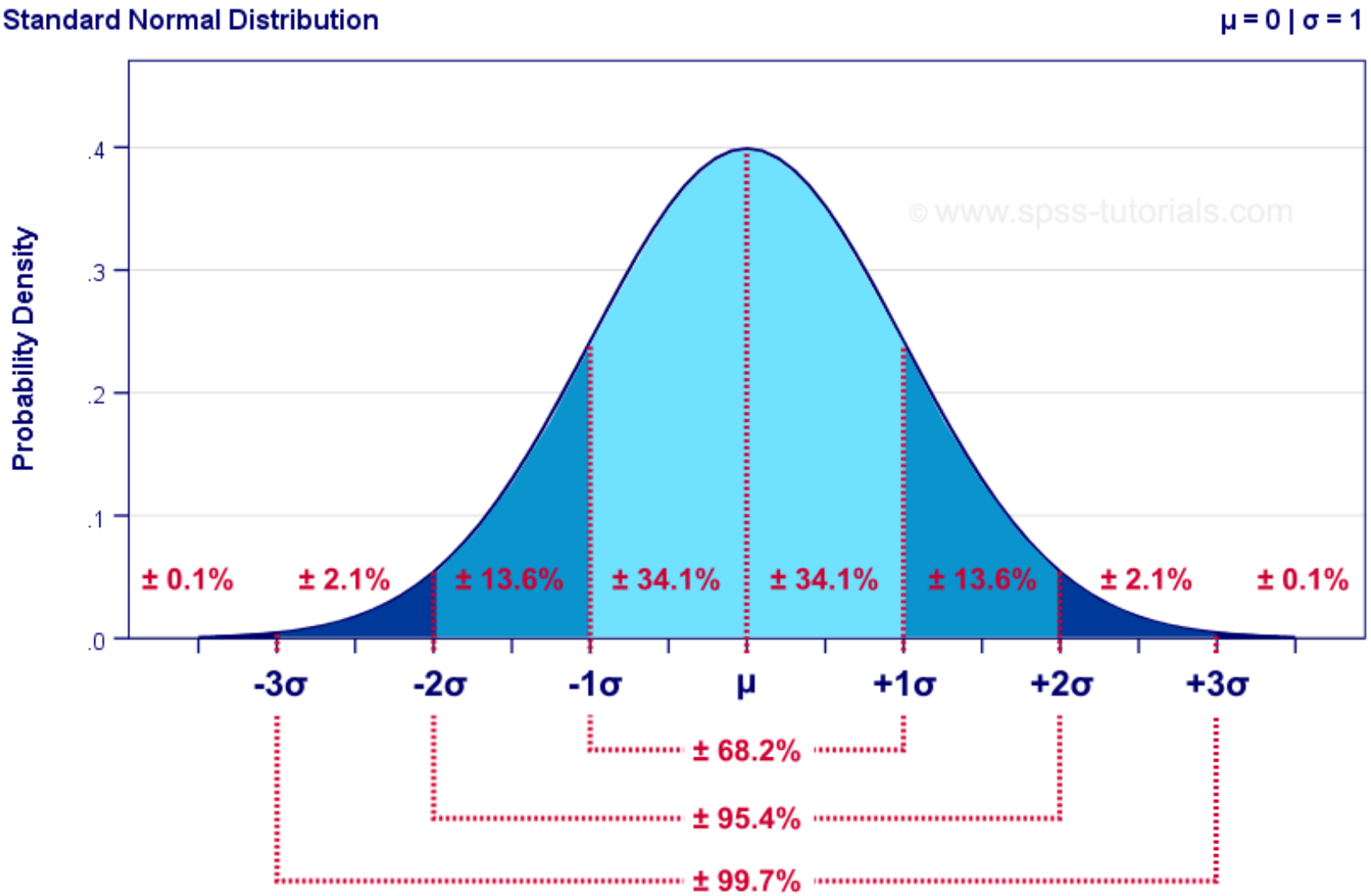
Sample Size

- Ideal
- Small: Cannot give valid conclusions
- Large: Wastage of Resources

Basis for sample size

Normal Distribution

Normal Distribution



Which will show Normal Distribution?

- All biological variables
- Several sample means
- Several sample proportions
- Differences between sample means
- Differences between sample proportions

Formulae

- No Single Formula
- Separate formulae for Descriptive, Cross Sectional, Case Control, Cohort, Experimental study
- Separate for Mean (continuous variable) and Proportion (Categorical variable)

Basically

- 4 Formulae
- Corresponding to 4 tests of Statistical Significance based on Normal Distribution

Basic 4 Formulae

TEST	Sample Size for
SE of Mean	Single Mean
SE of Proportion	Single Proportion
SE of Diff 2 means	2 Means
SE of Diff 2 Propo.	2 Proportions

Single Mean: Example

- Mean SBP of Male adults 20-60 years in Nellore district.
- Estimated: 10,00,000

Question

- How do you estimate Population Mean?

Answer

- Population Mean = Sample Mean ± Z_{α} . $S_{\bar{X}}$

Standard error of Mean

- $S_{\bar{X}} = \frac{S}{\sqrt{N}}$
- Difference between Popu Mean and Sample Mean is not more than 1.96 times $S_{\bar{X}}$

  or Z_{α} times $S_{\bar{X}}$
- This difference is called as E or Precision

Sample Size

- E = (Z_α) SE

  E = $(Z_\alpha).\ \frac{S}{\sqrt{N}}$

By Squaring both sides

E 2 = $\frac{(Z_\alpha)^2 X\ S^2}{N}$.

N = $\frac{(Z_\alpha)^2 X\ S^2}{E^2}$

Single Proportion: Example

- Proportion of Hypertension among Male adults 20-60 years in Nellore district.
- Estimated: 10,00,000

Question

- How do you estimate Population proportion?

Answer

- Population Prop. = Sample Prop ± Z_{α} . S_p

Standard error of Proportion

- $S_p = \sqrt{\frac{PQ}{N}}$
- Difference between Popu Proportion and Sample Proportion is not more than 1.96 times S_p

  or Z_α times S_p
- This difference is called as E or Precision

Sample Size

- E = (Z_α) SE

$$E = (Z_\alpha).\sqrt{\frac{PQ}{N}}$$

By Squaring both sides

$$E\ 2 = \frac{(Z_\alpha)^2 X\ PQ}{N}.$$

$$N = \frac{(Z_\alpha)^2 X\ PQ}{E^2}$$

Note

- For Estimation of Population mean and Population proportion, only Alpha error is used in the Calculation of sample size
- When it is hypothesis testing of difference between two means or proportions, alpha and beta errors are used in the calculation of sample size

Alpha error, Beta Error

Court Decision	Person is really	
	Guilty	Not Guilty
Guilty	Correct	Alpha error
Not Guilty	Beta Error	Correct

New Treatment Evaluation

Study decision	New Treatment is really	
	Effective	Not effective
Effective	Correct	Alpha error
Not Effective	Beta Error	Correct

Note

- Alpha error is more serious than beta error
- (1-β) = Power of the test

Z alpha value

- For 0.05 alpha error: 1.96
- For 0.01 alpha error: 2.58

Z beta value

Beta error	Z value	Power (%)
0.20	0.84	80
0.15	1.04	85
0.10	1.29	90
0.05	1.64	95

Two means: Example

- Wish to compare the Mean SBP of Male and Female Adults 20-60 years in Nellore District.
- Estimated: 10,00,000 and 10,00,000

Question

- How do you determine if the two Sample Means are same?

Answer

- Second Sample Mean

= First Sample mean $\pm Z_{\alpha} \cdot S\left(\overline{X}_1 - \overline{X}_2\right)$

Standard error of Diff between 2 means

$$= \sqrt{\frac{S_1^{\,2}}{n1} + \frac{S_2^{\,2}}{n_2}}$$

- Difference between two means is not more than 1.96 times $S\,(\overline{X}_1 - \bar{X}_2)$

  or $Z_{(\alpha+\beta)}$ times $S\,(\overline{X}_1 - \bar{X}_2)$
- This difference is called as E or Precision

Sample Size

- $E = (Z_{\alpha})\ SE$

$$E = (Z_{\alpha+\beta}).\sqrt{\frac{S_1^{\ 2}}{n1} + \frac{S_2^{\ 2}}{n_2}}$$

By Squaring both sides and assuming

$S_1=S_2=S$ and $N_1=N_2=N$

$$E^{\ 2} = \frac{(Z_{\alpha+\beta})^2 X\ \ 2\ S^2}{N}$$

$$N = \frac{(Z_{\alpha+\beta})^2 X\ \ 2\ S^2}{E^2}$$

Two proportions: Example

- Wish to compare the Proportion of Hypertension in Male and Female Adults 20-60 years in Nellore District.
- Estimated: 10,00,000 and 10,00,000

Question

- How do you determine if the two Sample proportions are same?

Answer

- Second Sample Proportion

= First Sample mean $\pm Z_{\alpha} . S\left(_{p_1 - p_2}\right)$

Standard error of Diff between 2 proportions

$$= \sqrt{\frac{p_1 q_1}{n1} + \frac{p_2 q_2}{n_2}}$$

Difference between two proportions is not more than 1.96 times $S\ (_{p_1 - p_2})$

or $Z_{(\alpha+\beta)}$ times $S\ (_{p_1 - p_2})$

- This difference is called as E or Precision

Sample Size

- $E = (Z_{\alpha+\beta})\ SE$

$$E = (Z_{\alpha+\beta}).\sqrt{\frac{p_1 q_1}{n1} + \frac{p_2 q_2}{n_2}}$$

By Squaring both sides and assuming

$p_1=p_2=P$ and $q_1=q_2=Q$ and $n_1=n_2=N$

$$E^2 = \frac{(Z_{\alpha+\beta})^2 X\ 2\ PQ}{N}$$

$$N = \frac{(Z_{\alpha+\beta})^2 X\ 2\ PQ}{E^2}$$

Case control study: Binary Exposure

- Similar to Experimental study two proportions

- N = $\frac{(r+1)}{r} \cdot \frac{Z_{(\alpha+\beta)}{}^2 \cdot (PQ)}{(p_2-p_1)^2}$

  where r is number of controls to cases

Additionally: P = $\frac{p_1+p_2}{2}$

$p_2 = \frac{p_{1.OR}}{1+ p_{1.(OR-1)}}$

Case control study: Continuous Exposure

- Similar to Experimental study two means
- N = $\frac{(r+1)}{r} \cdot \frac{Z_{(\alpha+\beta)}{}^2 . S^2}{E^2}$

  where r is number of controls to cases

Prospective Cohort Study

- Similar to Case Control study Binary Exposure.

- N = $\frac{(r+1)}{r} \cdot \frac{Z_{(\alpha+\beta)}{}^{2} \cdot (PQ)}{(p_2-p_1)^2}$

  where r is number of controls to cases

Additionally: P = $\frac{p_1+p_2}{2}$

$p_2 = p_1 \times RR$

Screening Test

- If sensitivity is known.
- Formula similar to Single Proportion
- N = $\frac{(Z_\alpha)^2\ Sen.FN}{E^2 Pre.}$

Screening Test

- If Specificity is known.
- Formula similar to Single Proportion
- $N = \frac{(Z_\alpha)^2 \, Spe.FP}{E^2.(1-Pre.)}$


THANK
YOU