How many study subjects are required ? (Estimation of Sample size)

By

Dr.Shaik Shaffi Ahamed
Professor
Dept. of Family & Community Medicine
College of Medicine
King Saud University

Objectives of this session: Students able to

- (1) know the importance of sample size in a research project.
- (2) understand the simple mathematics & assumptions involved in the sample size calculations.
- (3) apply sample size methods appropriately in their research projects.

Why to calculate sample size?

- To show that under certain conditions, the hypothesis test has a good chance of showing a desired difference (if it exists)
- To show to the IRB committee and funding agency that the study has a reasonable chance to obtain a conclusive result
- To show that the necessary resources (human, monetary, time) will be minimized and well utilized

What do I need to know to calculate sample size?

- Most Important: sample size calculation is an educated guess
- It is more appropriate for studies involving hypothesis testing
- There is no magic involved; only statistical and mathematical logic and some algebra
- Researchers need to know something about what they are measuring and how it varies in the population of interest

- SAMPLE SIZE:

How many subjects are needed to assure a given probability of detecting a statistically significant effect of a given magnitude if one truly exists?

- POWER:

If a limited pool of subjects is available, what is the likelihood of finding a statistically significant effect of a given magnitude if one truly exists?

Before We Can Determine Sample Size We Need To Answer The Following:

1. What is the primary objective of the study?
2. What is the main outcome measure?

   Is it a continuous or dichotomous outcome?
3. How will the data be analyzed to detect a group difference?
4. How small a difference is clinically important to detect?

5. How much variability is in our traget population?

6. What is the desired α and β?

7. What is the anticipated drop out and non-response % ?

Where do we get this knowledge?

- Previous published studies
- Pilot studies
- If information is lacking, there is no good way to calculate the sample size

- Type I error: Rejecting H_0 when H_0 is true
- α: The type I error rate.
- Type II error: Failing to reject H_0 when H_0 is false
- β: The type II error rate
- Power $(1 - \beta)$: Probability of detecting group difference given the size of the effect (Δ) and the sample size of the trial (N)

Diagnosis and statistical reasoning

	Disease status: Present	Disease status: Absent
Test result		
+ve	True +ve (sensitivity)	False +ve
-ve	False –ve	True -ve (Specificity)

Test result	Significance Difference is: Present (Ho *not* true)	Significance Difference is: Absent (Ho is true)
Reject Ho	No error $1-\beta$	Type I err. α
Accept Ho	Type II err. β	No error $1-\alpha$

α : significance level

$1-\beta$: power

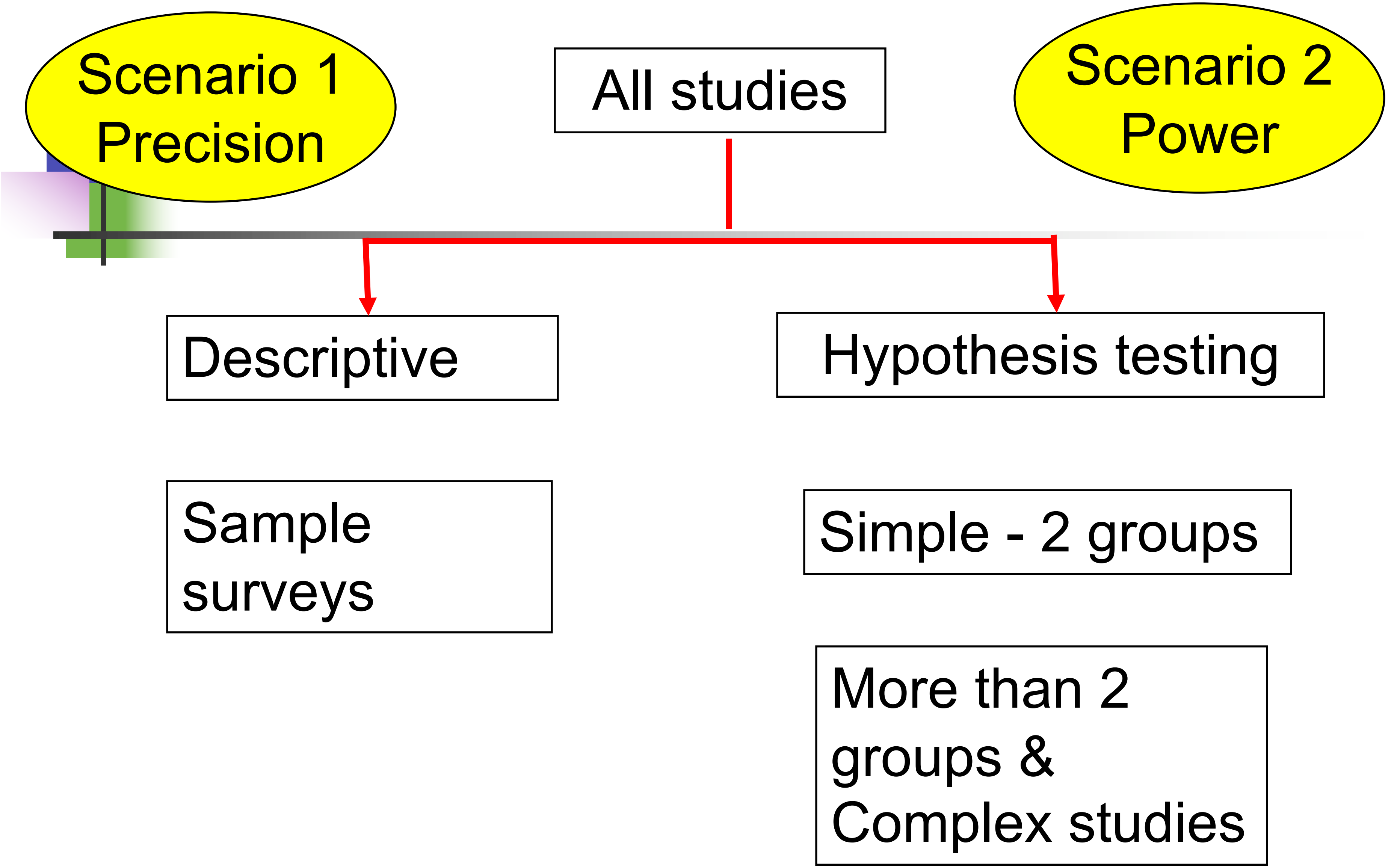
Estimation of Sample Size by Three ways:

By using

(1) Formulae (manual calculations)

(2) Sample size tables or Nomogram

(3) Softwares


Scenario 1
Precision
All studies
Scenario 2
Power
Descriptive
Hypothesis testing
Sample surveys
Simple - 2 groups
More than 2 groups &
Complex studies

SAMPLE SIZE FOR ADEQUATE PRECISION

- In a descriptive study,
- Summary statistics (mean, proportion)
- Reliability (or) precision
- By giving “confidence interval”
- Wider the C.I – sample statistic is not reliable and it may not give an accurate estimate of the true value of the population parameter

Sample size formulae for reporting precision

For single mean : $n = Z^2_\alpha S^2 / d^2$

where S=sd (σ)

For a single proportion : $n = Z^2_\alpha P(1-P)/d^2$

Where , $Z\alpha$ =1.96 for 95% confidence level

$Z\alpha$ = 2.58 for 99% confidence level

Problem 1 (Single mean)

A study is to be performed to determine a certain parameter(BMI) in a community. From a previous study a sd of 46 was obtained.

If a sample error of up to 4 is to be accepted. How many subjects should be included in this study at 99% level of confidence?

Answer

$n = (Z_{\alpha/2})^2\ \sigma^2 / d^2$

σ: standard deviation = 46

d: the accuracy of estimate (how close to the true mean)= given sample error =4

$Z_{\alpha/2}$: A Normal deviate reflects the type I error. For 99% the critical value =2.58

$$n = \frac{2.58^2 \ \text{x} \ 46^2}{4^2} = 880.3 \sim 881$$

Problem 2 (Single proportion)

It was desired to estimate proportion of anemic children in a certain preparatory school. In a similar study at another school a proportion of 30 % was detected.

Compute the minimal sample size required at a confidence limit of 95% and accepting a difference of up to 4% of the true population.

Answer

$n = (Z_{\alpha/2})^2\, p(1\text{-}p) \,/\, d^2$

p: proportion to be estimated = 30% (0.30)

d: the accuracy of estimate (how close to the true proportion) = 4% (0.04)

$Z_{\alpha/2}$: A Normal deviate reflects the type I error

For 95% the critical value =1.96

$$n = \frac{1.96^2 \text{ x } 0.3(1-0.3)}{(0.04)^2} = 504.21 \sim 505$$

Scenario 2

Three bits of information required to determine the sample size

Type I & II errors

Clinical effect

Variation

Quantities related to the research question (defined by the researcher)

- α = Probability of rejecting H_0 when H_0 is true
- α is called significance level of the test
- β = Probability of not rejecting H_0 when H_0 is false
- **$1-\beta$** is called statistical power of the test

- Researcher fixes probabilities of type I and II errors
  - Prob (type I error) = Prob (reject H_0 when H_0 is true) = α
    - Smaller error $\Rightarrow$ greater precision $\Rightarrow$ need more information $\Rightarrow$ need larger sample size
  - Prob (type II error) = Prob (don't reject H_0 when H_0 is false) = β
  - Power =1- β
    - More power $\Rightarrow$ smaller error $\Rightarrow$ need larger sample size

Quantities related to the research question (defined by the researcher)

- Size of the measure of interest to be detected
  - Difference between two or more means
  - Difference between two or more proportions
  - Odds ratio, Relative risk, Correlation, Regression coefficients
  - Change in R^2, etc

- The magnitude of these values depend on the research question and objective of the study (for example, clinical relevance)

Clinical Effect Size

"What is a meaningful difference between the groups"

- It is truly an estimate and often the most challenging aspect of sample size planning
- Large difference – small sample size
- Small differences – large sample size

- Cost/benefit

Variation

Variability

All statistical tests are based on the following ratio:

$$\text{Test Statistic} = \frac{\text{Difference between parameters}}{v / \sqrt{n}}$$

As n ↑ v/√n ↓ Test statistic ↑

Sample size formulae for comparing

two means : $n = 2 S^2 (Z_\alpha + Z_\beta)^2 / d^2$

where S=sd; d= difference

two proportions :

$$n = \frac{(Z_\alpha + Z_\beta)^2 ((p_1 q_1) + (p_2 q_2))}{(p_1 - p_2)^2}, \text{ where } q_1 = (1 - p_1), q_2 = (1 - p_2)$$

Z_α= 1.96 for 95% confidence level

Z_α = 2.58 for 99% confidence level ;

Z_β= 0.842 for 80% power

Z_β= 1.282 for 90% power

Example 1: *Does the consumption of large doses of vitamin A in tablet form prevent breast cancer?*

- Suppose we know from our tumor-registry data that incidence rate of breast cancer over a 1-year period for women aged 45 – 49 is 150 cases per 100,000
- Women randomized to Vitamin A vs. placebo

Example 1 continued

- *Group 1*: Control group given placebo pills. Expected to have same disease rate as registry (**150 cases per 100,000**)

- *Group 2*: Intervention group given vitamin A tablets. Expected to have 20% reduction in risk (**120 cases per 100,000**)

- Want to compare incidence of breast cancer over 1-year

- *Planned statistical analysis*: Chi-square test to compare two proportions from independent samples

$$H_0: p_1 = p_2 \quad \text{vs.} \quad H_A: p_1 \neq p_2$$

Example 1: *Does ingestion of large doses of vitamin A prevent breast cancer?*

- **Test H_0: $p_1 = p_2$ vs. H_A $p_1 \neq p_2$**
- **Assume 2-sided test with α=0.05 and 80% power**

$$n = \frac{(Z_\alpha + Z_\beta)^2((p_1 q_1) + (p_2 q_2))}{(p_1 - p_2)^2}, \text{ where } \quad q_1 = (1 - p_1), q_2 = (1 - p_2)$$

- **p_1 = 150 per 100,000 = .0015**
- **p_2 = 120 per 100,000 = .0012 (20% rate reduction)**
- **$\Delta = p_1 - p_2$ = .0003**
- **$z_{1-\alpha/2}$ = 1.96 $z_{1-\beta}$ = .84**
- **n per group = 234,882**
- **Too many to recruit in one year!**

Example 2: *Does a special diet help to reduce cholesterol levels?*

- Suppose an investigator wishes to determine sample size to detect a 10 mg/dl difference in cholesterol level in a diet intervention group compared to a control (no diet) group

- Subjects with baseline total cholesterol of at least 300 mg/dl randomized

Example 2 continued

- *Group 1*: A six week diet intervention
- *Group 2*: No changes in diet

- Investigator wants to compare total cholesterol at the end of the six week study

- *Planned statistical analysis*: two sample t-test (for independent samples)(comparison of two means)

$$H_0: \mu_1 = \mu_2 \quad \text{vs.} \quad H_A: \mu_1 \neq \mu_2$$

Sample Size Formula

To Compare Two Means From Independent Samples: H_0: $\mu_1 = \mu_2$

1. α level
2. β level (1 – power)
3. Expected population difference ($\Delta = |\mu_1 - \mu_2|$)
4. Expected population standard deviation (σ_1 , σ_2)

Continuous Outcome
(2 Independent Samples)

- Test H_0: $\mu_1 = \mu_2$ vs. H_A: $\mu_1 \neq \mu_2$
- Two-sided alternative
- Assume outcome normally distributed with:

$$n_{per/group} = \frac{(2S^2)(z_\alpha + z_\beta)^2}{d^2}$$

S= standard deviation; d=difference between two means;

$Z\alpha$= 1.96 for 95% confidence level; Z_β= 1.28 for 90% power

Example 2: *Does a special diet help to reduce cholesterol levels?*

- Test $H_0: \mu_1=\mu_2$ vs. $H_A : \mu_1 \neq \mu_2$
- Assume 2-sided test with $\alpha=0.05$ and 90% power

- $d = \mu_1 - \mu_2 = 10$ mg/dl
- $\sigma_1 = \sigma_2 = (50$ mg/dl$)$
- $z_\alpha = 1.96 \quad z_\beta = 1.28$

- n per group = 525

- Suppose 10% loss to follow-up expected, adjust n = 525 / 0.9 = 584 per group

Problem (comparison of two means)

- A study is to be done to determine effect of 2 drugs (A and B) on blood glucose level. From previous studies using those drugs, Sd of BGL of 8 and 12 g/dl were obtained respectively.
- A significant level of 95% and a power of 90% is required to detect a mean difference between the two groups of 3 g/dl. How many subjects should be include in each group?

Answer

$$n = \frac{(SD1 + SD2)^2}{\Delta^2} * f(\alpha,\beta)$$

$$n = \frac{(8^2 + 12^2)x10.5}{3^2} = 242.6 \sim 243$$

in each group

Comparison of two means

- Objective:

  To observe whether feeding milk to 5 year old children enhances growth.

Groups:

Extra milk diet

Normal milk diet

Outcome:

Height (in cms.)

Assumptions or specifications:

Type-I error (α) =0.05

Type-II error (β) = 0.20

i.e., Power($1-\beta$) = 0.80

Clinically significant difference (Δ) =0.5 cm.,

Measure of variation (SD.,) =2.0 cm.,

(from literature or “Guesstimate”)

Using the appropriate formula:

$$N = \frac{2(SD)^2}{\Delta^2} f(\alpha,\beta)$$

$$= \frac{2(2)^2}{(0.5)^2} 7.9$$

= 252.8 (in each group)

The following steps constitute a pragmatic approach to decision taking on Sample size:

(1) Remember that there is no stock answer.

(2) Initiate early discussion among research team members.

(3) Use correct assumptions – consider various possibilities.

(4) Consider other factors also– eg., availability of cases, cost, time.

(5) Make a balanced choice

(6) Ask if this number gives you a reasonable prospect of coming to useful conclusion.

(7) If yes, proceed if no, reformulate your problem for study.

Summary

- Define research question well
- Consider study design, type of response variable, and type of data analysis
- Decide on the type of difference or change you want to detect (make sure it answers your research question)
- Choose α and β
- Use appropriate equation for sample size calculation or sample size tables/ nomogram or software.

Any Q's