Editing file

Feedback

Sample Size Estimation

Objectives:

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.

Click here for the practical

15th lecture

Color Index:

- Boys' Slides
- Girls' Slides
- Doctors Notes
- Golden Notes
- Important
- Extra

MED439

KING SAUD UNIVERSITY

Research Team

RESEARCH TEAM 438

Sample size estimation

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? e.g. the null hypothesis states that there is no difference in the prevalence of smoking between dental college students and medical college students. But suppose the alternative hypothesis states that there is a 5% difference in smoking prevalence between the two colleges. The sample size is basically the number of students needed from each college to show this 5% difference as statistically significant

Power
Opposite to sample size

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?

The power of a study is the ability to see the existence of treatment effect in a study.(1 - β)

Why to calculate sample size?

1. To show that under certain conditions, the hypothesis test has a good chance of showing a desired difference (if it exists).
2. To show to the IRB committee and funding agency that the study has a reasonable chance to obtain a conclusive result.
3. 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. **So what is hypothesis test? In statistics we have 2 statements one is null hypothesis and the other is alternate hypothesis, always there must be a comparison group. Ex. If there's no difference among male and female students then its null and if there's a difference then its alternate hypothesis**
- 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 estimation

Before determining the sample size answer the following:

- What is the primary objective of the study?
- What is the main outcome measure? Is it a continuous (Quantitative and calculates means (e.g. BMI, Hemoglobin levels, etc) or dichotomous (Categorical variables (including prevalence percentages) + Binary options (e.g. dead or survived) outcome?
- How will the data be analyzed to detect a group difference?
- How small a difference is clinically important to detect? The smaller the difference, the larger the sample size needed
- How much variability is in our target population?
- What is the desired α and β (Type I and Type II errors)?
- What is the anticipated drop out and non- response % ?

Where to get this knowledge? The more you read the more you have the knowledge.

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

Estimation of sample size by using:

Formulae (manual calculation) → Tables or nomogram → Softwares

From a statistical point of view studies divided into:

Descriptive study	VS	Hypothesis testing study "comparing"
✓ Sample surveys. ✓ Here you are looking for Precision.(single means\proportion) If no comparisons are involved, sample size calculation should be based on precision		✓ Simple → 2 groups. ✓ More than 2 groups & complex studies. ✓ Here you are looking for Power.(two means\proportions)

Sample size estimation

Type I error	Type II error
Rejecting H_0 when H_0 is true α: type I error rate we use α as 5% ↳ Is called significance level.	Failing to reject H_0 when H_0 is false β : type II error rate we use β as 20%

Power (1 - β): Probability of detecting group difference given the size of the effect(Δ) and the sample size of the trial (N).(Power is the sensitivity in statistical terminology)

↳ Is called statistical power.

example : we are looking for the effect of some new intervention in compare with the existing treatment ,Whether you have enough power to detect the difference between the two treatments. For that you must have enough sample size.

- If the sample size is small then the study may not have power to detect the difference.
- Used only in hypothesis testing study (analytical) because you are comparing.

Sensitivity is the ability of the test to detect the disease

Diagnosis and statistical reasoning

Test result	Disease status	
	present	absent
+ve	True +ve (sensitivity)	False +ve
-ve	False -ve	True -ve (specificity)

- the sensitivity (TP) is equal to the power
- A false positive result is a type I error
- A false negative result is a type II error

Test result	Significance difference	
	Present H0 not true	Absent H0 is true
Reject H0	No error (1-β)	Type I error α
Accept H0	Type II error β	No error (1-α)

Sample size for adequate precision

- In descriptive study, we put summary statistics (mean, proportion). For these statistics we have to provide some sort of reliability or (precision)
- Reliability : is repeatability and reproducibility, how reliable your estimates is.
- By giving "confidence interval". The width of CI indicates the precision
  - → Small sample size → wider CI
  - → Large sample size → closer CI
- The wider the confidence interval → sample statistic is not reliable and it may not give an accurate estimate of the true value of the population parameter.
- To have good precision you must have a close CI, to have close CI you must have good sample size.
- **Example:** if the prevalence of obesity is 10% (take it as a point estimates) and you calculate the CI = ±3 then (7%-13%), then if someone does the same study their prevalence will be between these intervals (hence the closer the interval is, the more accurate and reliable the study is)

Sample size formula for reporting precision:

(SINGLE GROUP STUDIES no comparisons are involved 'descriptive')

- Single mean: (for continuous/quantitative outcome: age, weight, height, blood pressure, body temperature etc)

  $n = Z^2_{\alpha} S^2 / d^2$

  Where : S = sd

- Single proportion; (for categorical/qualitative outcome: classified into one of a number of categories on the basis of characteristics of variable, yes/no..etc)

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

  Where : 1- Z_{α} =1.96 for 95% confidence level (used most of the time)

  2- Z_{α} = 2.58 for 99% confidence level:

- ✓ n = sample size
- ✓ Z = normal distribution
- ✓ α = type I error
- ✓ S = standard deviation (σ) (variability)
- ✓ d = precision
- ✓ P = proportion (prevalence)

So standard deviation for:
1- quantitative data is S
2- categorical data is P(1-P)

Confidence level = (1- α)%
Where α is the level of significance.
So if:
Confidence level is 95% , α will be 5% = 0.05
Then
$Z_{1-0.05/2} = Z_{0.975}$ (check Z table) = 1.96

Three bits of information required to determine the sample size: "in analytical studies"

1- Type I & II errors: (in descriptive studies we only need type I error)

- Researcher fixes probabilities of type I and II errors:

  1- prob (type I error) = prob (reject H_0 when H_0 is true) = α

  Smaller error → greater precision → need more information → need larger sample size. (usually it's fixed to 5%)

  2- prob (type II error) = prob (don't reject H_0 when H_0 is false) = β (usually fixed to 20%)

  3- power = 1 - β

  More power → smaller error → need larger sample size

- Size of the measure of interest to be detected:

  Difference between two or more means, 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 objectives of the study (for example , clinical relevance).

2- Clinical effect size: (effect size, group difference, or clinical effect → different terminologies)

"What is the meaningful difference between the groups"

- It is truly an estimate and often the most challenging aspect of sample size planning.(you need to decide how much the group difference you want to detect)
  - → Large difference - small sample size.
  - → Small difference - large sample size (cost/benefit), so do case control study.

3- variation:

- All statistical tests are based on the following ratio:

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

✓ v = variability

As ↑ n , ↓ $v / \sqrt{n}$, so ↑ test statistic

- This is the relation between the effect size , the difference between the two groups, the variability, and the sample size. (n is the sample size . N is the most significant variable which makes the most changes so sample size is the most important out of all)
  - → As the sample size is small this quantity $v / \sqrt{n}$ will be larger , whatever the difference between the two groups is you may not get the statistical significant result.
  - → When this quantity $v / \sqrt{n}$ is small even if the difference between the two groups is small you will get the significant result.
  - → When the difference between the two groups is large but you have this quantity $v / \sqrt{n}$ large due to small sample size we may not get the statistical significant

Sample size formula for comparing: (TWO-INDEPENDENT GROUP STUDIES, 'analytical,hypothesis testing')

1- Two mean: (continuous outcome)

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

- ✓ S = standard deviation
- ✓ d = difference between two means (contrast with the previous d parameter).
- ✓ α = type I error
- ✓ β = type II error (power = 1-β)

2- Two proportion: (categorical outcome)

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

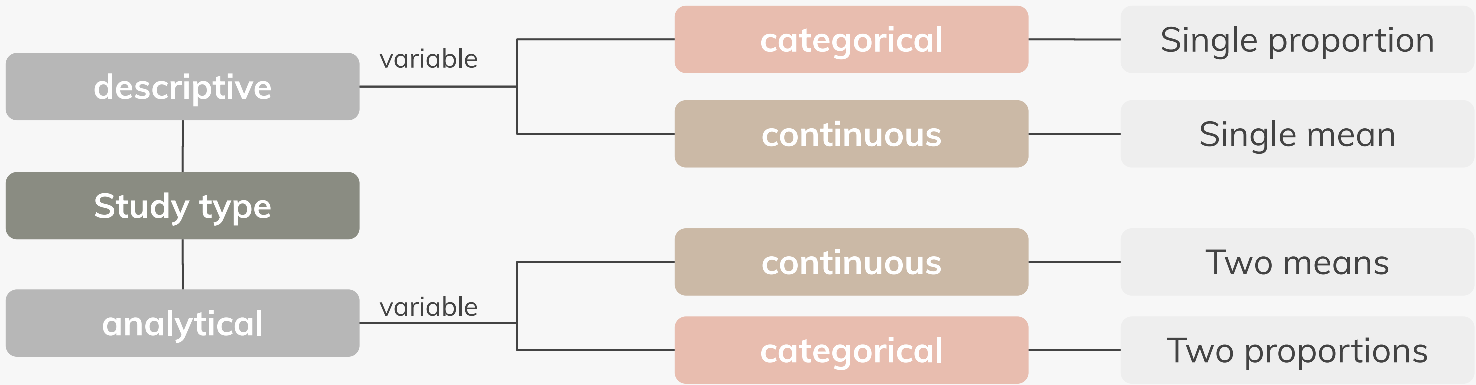
Where :

→ Zα =1.96 for 95% confidence level
→ Zα =2.58 for 99% confidence level
→ Z_β= 0.842 for 80% power
→ Z_β= 1.282 for 90% power

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.

Sample size estimation

Z table   Check this if you didn't understand



Problem 1

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.(no comparison)

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: (single mean)

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

$$n = Z^2_{\alpha\backslash 2} S^2 / d^2$$

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

Add 10% for Non-response or missing data

Problem 2

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. (no comparison)

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: (single proportion)
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

$$n = Z^2_{\alpha\backslash 2}\, p(1-p) / d^2$$

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

Sample size estimation

Problem 3

<u>Does</u> 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
- Group 1: Control group given placebo pills. Expected to have same disease rate as registry (**150 cases per 100,000**) (comparison)
- 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, Women randomized to Vitamin A vs. placebo.
- Planned statistical analysis: Chi-square test to compare two proportions from independent samples $H_0: p_1 = p_2$ vs. $H_A: p_1 \neq p_2$
- Assume 2-sided test with α=0.05 and 80% power

<u>Answer:</u> (two proportions)

p_1 = 150 per 100,000 = .0015 / q1 = 1 - 0.0015 = 0.9985

p_2 = 120 per 100,000 = .0012 (20% rate reduction) / q2 = 1 - 0.0012 = 0.9988

$\Delta = p_1 - p_2 = .0003$

$z_{1-\alpha/2} = 1.96$ $z_{1-\beta}$ = .84

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

$n = (1.96 + 0.84)^2 * ((0.0015*0.9985) + (0.0012*0.9988)) / (0.0015 - 0.0012)^2$

n = 234,882 (Too many to recruit in one year! Because the incidence is very small)

n = 234,878

Problem 4

To observe <u>whether</u> feeding milk to 5 year old children enhances growth (height in cm)

- Group 1: Extra milk diet
- Group 2: Normal milk diet (comparison)
- <u>Assumptions or specifications:</u> (from literature or "Guesstimate")
  - Type-I error (α) =0.05 , $Z_\alpha = 1.96$)
  - Type-II error (β) = 0.20
  - i. Power(1-β) = 0.80 , $Z_\beta = 0.842$)
  - Clinically significant difference (Ω) =0.5 cm.,
  - Measure of variation (SD.,) =2.0 cm.,

<u>Answer:</u> (two means)

$n = 2\, S^2\, (Z_\alpha + Z_\beta)^2 / d^2$

$n = 2\,(2)^2 * (1.96 + 0.842)^2 / 0.5^2 =$ 252.8 (in each group)

Sample size estimation

Problem 5

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
- 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)

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

- Assume 2-sided alternative test with α=0.05 and 90% power

Answer: (two means)

$Z_\alpha = 1.96$ for 95% confidence level; $Z_\beta = 1.28$ for 90% power

$d = \mu_1 - \mu_2 = 10$ mg/dl

$\sigma_1 = \sigma_2 = (50$ mg/dl)

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

$$n = 2 (50)^2 * (1.96 + 1.28)^2 / 10^2 = 525 \text{ per group}$$

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

→ Suppose 10% loss to follow-up expected:
525*10% = 52.5 , adjust n = 525 + 52.5 = 577.5 per group

Problem 6

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: (two means)

$$n = 2S^2 (Z_{1-\alpha\backslash 2} + Z_{1-\beta})^2 / d^2$$

$$= (8^2 + 12^2) * (1.96+1.28)^2 / 3^2$$

= 242.6~243 in each group

$S_p = S1^2 + S2^2 / 2$
$= 8^2 + 12^2 / 2$
$= 104$
$2 S_p = 208$
So
$2 (S1^2 + S2^2 / 2)$
Equals
$(S1^2 + S2^2)$

Lecture Summary

General information

- Sample size calculation is an educated guess & more appropriate for studies involving hypothesis testing.
- **Descriptive studies** use summary statistics (mean, proportion) to measure the **precision or reliability** of the study, it measures the reliability by giving the "confidence interval".
- **Hypothesis testing** includes simple or complex studies that are used to measure the **power**.
- Three bits of information are required to determine the sample size: Type I and II errors & Clinical effect & Variation.
- You can estimate the sample size by 3 ways: Formulae & Sample size tables or Nomogram & Softwares.
- The test used in statistical analysis to compare two proportions from independent sample is **Chi-square test.**
- The test used in statistical analysis to compare two means from independent sample is **two sample t-test.**

Definitions

- **Sample size:** The number of subjects needed to assure that a given probability can detect a statistically significant effect of a given magnitude.
- **Power:** The probability of detecting group difference given the size of the effect and the sample size of the trial.
- **Clinical Effect Size:** meaningful difference between the groups.
- **Type I error (α):** Rejecting H0 when H0 is true.
- **Type II error (β):** Accepting (Failing to reject) H0 when H0 is false.
- **α:** The significance level of the test.
- **1-β:** The statistical power of the test.

Formulas

- $n = Z_\alpha^2 \cdot (1-P) / d^2$ — For single proportion (Qualitative e.g. diabetes)
- $n = Z_\alpha^2 \cdot S^2 / d^2$ — For single mean (Quantitative e.g. weight)
- $n = (Z_\alpha + Z_\beta)^2 \cdot ((p_1q_1) + (p_2q_2)) / (p_1-p_2)^2$ — For comparing two proportions.
- $n = 2S^2 (Z_\alpha + Z_\beta)^2 / d^2$ — For comparing two means.

Relationships

- The wider the confidence interval the more it's not reliable (not precise).
- Smaller error > greater precision > need more information > need larger sample size.
- More power > smaller error > need larger sample size.
- The larger the differences the smaller the sample size.
- The smaller the differences the larger the sample size.
- As the sample size (n) increases the variability ratio (v/n) decreases and the test statistic increases.

Remember to define the 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.

Questions

(1) Type II error is?

A) Accepting H0 when H0 is false.

B) Rejecting H0 when H0 is true.

C) Rejecting H0 when H0 is false.

D) Accepting H0 when H0 is true.

(2) If the population mean is 250, sample mean is 210, standard deviation is 10, and N is 9 what is the Z-score?

A) 5.5

B) 2

C) 4

D) 3

(3) two sample t-test is used for?

A) Single mean.

B) Single proportion.

C) Comparison of two means.

D) Comparison of two proportions.

(4) What happens when the sample size increases?

A) The test statistic decreases.

B) The variability ratio increases.

C) Both of the above.

D) None of the above.

(5) Chi-square test is used for?

A) Single mean.

B) Single proportion.

C) Comparison of two means.

D) Comparison of two proportions.

(6) The 'd' in the following formula: $Z_a^2 \cdot S^2 / d^2$, stands for what?

A) The accuracy of estimate.

B) The proportion to be estimated.

C) The difference between the means.

D) The standard deviation.

Answers:
1: A | 2: C | 3: C | 4:D | 5: D | 6: A

Research Team

RESEARCH TEAM 438

Keep it simple, Keep it focused

Thank you for checking our work!

Leaders:

Shuaa Khdary   Sarah AlQuwayz

Abdulrhman Alsuhaibany

Member:

Rand AlRefaei

Note Taker:

Rand AlRefaei

Contact us:

Research4390@gmail.com