Basic principles in hypothesis testing

An RCT was conducted to assess the effect of intervention-conservative management of persistent postnatal urinary incontinence.

Setting: Community setting

Intervention: Assessment by nurses of urinary incontinence with conservative advice on pelvic floor exercise at five, seven and nine months after delivery.

Control : No visit by nurses

Results

Among 279 women in the intervention group, 167 (59.9%) experienced any form of urinary incontinence compared to 169 (69%) in the control group (245).

Does this result gives enough evidence that conservative management by nurses reduces any form of urinary incontinence?

Data

Group	End point-Urinary incontinence	
	Yes	No
Intervention	167 (59.9%)	112
Control	169 (69%)	76

Statistical terms

- **Null hypothesis**

  Conservative management not effective in reducing any form of urinary incontinence.

- **Alternate hypothesis**

  Conservative management is effective in reducing urinary incontinence

Process of hypothesis testing

		Possible conditions of null hypothesis (Intervention is not effective)	
		True	False
Possible actions on Null hypothesis	Accept	*Correct action*	Type II error
	Reject	Type I error	*Correct action*

Prob. {Type I error} = α = level of significance.

Prob. {Type II error} = β, 1 - β = Power of the test.

Test of significance

This is a statistical procedure by which one can conclude, whether the difference observed between two groups is only due to chance or not.

Analysis

Group	Urinary incontinence		Chi Square	P value
	Yes	No		
Intervention	167 (59.9%)	112	4.33	0.04
Control	169 (69%)	76		

What is p-value?

Probability that difference at least as large as those found in the observed data would have occurred by chance.

Low p Value

High p value

Clinical significance

Types of tests of significance

-Two sided or two tail test

-One sided or one tail test

Types of tests of significance

-Parametric tests

- Non parametric tests

Important parametric tests

Students' t test- paired and unpaired

Analysis of variance

- One way
- Two way
- Repeated measures

Important non parametric tests

Chi-square test

Mann –Whitney

Willcoxon

Non parametric analysis of variance

Kruskal wallis

Freidmann

Chi square test

Association between type of ICU admission & Mortality

		Vital status		Total
		Lived	Died	
Service at ICU admission	Medical	67	26	93
	Surgical	93	14	107
Total		160	40	200

Chi square test

$$\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

Where

O_i = Observed frequency

E_i = Expected frequency

Association between type of ICU admission & Mortality

		Vital status		Total
		Lived	Died	
Service at ICU admission	Medical			93
	Surgical			107
Total		160	40	200

Association between type of ICU admissions & Mortality

		Vital status		Total
		Lived	Died	
Service at ICU admission	Medical	67	26	93
	Surgical	93	14	107
Total		160	40	200

Chi square = 6.88

P value = 0.009

t- test and analysis of variance

Two independent sample t test

		Age		
		N	Mean	S.D
Vital status	Medical	160	55.7	20.43
	Surgical	40	65.1	16.50

Two independent sample t test

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}}$$

Two independent sample t test

		Age			P value
		N	Mean	S.D	
Vital status	Medical	160	55.7	20.43	0.003
	Surgical	40	65.1	16.50	

One way ANOVA

		Age			P value
		N	Mean	S.D	
Level of consciousness at ICU admission	No comma	185	57.0	20.51	0.403
	Deep stupor	5	61.2	11.17	
	Coma	10	65.4	12.50	

Sample size in health science research

Sample size in health science research

- What would be the sample size?
- What is the size of population to be studied to get an estimate of prevalence of cataract in a district?
- How many subjects we would need in each group to have a valid comparison for incidence of diabetes in two groups?

Sample size is determined based on:

1. Main research question, the outcome measure and the statistical procedure
2. Statistical and clinical assumptions
3. Study constraints

Example

A randomized, double blind placebo controlled trial was planned to study the effect of estrogen replacement therapy after Ischemic stroke. Under the assumption that the rate of primary outcome (death or nonfatal stroke) would be reduced from 25% in the placebo group to 15% in the estradiol group, expecting a dropout rate of 10%, to achieve a power of 80% at a two -tailed alpha level of 0.05, what would be the number of subjects recruited in each group? (N Engl J Med 2001; 345(7):1243-9.).

Main research question, the outcome measure and the statistical procedure

Main research question	Statistical Procedure
Does estrogen replacement therapy after ischemic stroke reduce the rate of death or nonfatal stroke?	Difference in two proportions
Does dietary intervention reduce blood pressure?	Difference in two Means
What is the Prevalence of cataract in a district?	Estimation of proportion

Statistical and clinical assumptions

- The sample size required for a one sided test is smaller than that needed for the same precision in a two sided test.
- Less α → more subjects
- Less β → More power (1-β)

  → more subjects

Statistical and clinical significance

- Small difference in two groups is clinically significant → Large sample size.
- Large difference only clinically significant → small sample size.

Sample size for comparing two proportions

प्रज्ञानं ब्रह्म
Manipal
INSPIRED BY LIFE

A randomized, double blind placebo controlled trial was planned to study the effect of estrogen replacement therapy after Ischemic stroke. Under the assumption that the rate of primary outcome (death or nonfatal stroke) would be reduced from 25% in the placebo group to 15% in the estradiol group, expecting a dropout rate of 10%, to achieve a power of 80% at a two -tailed alpha level of 0.05, what would be the number of subjects recruited in each group? (N Engl J Med 2001; 345(7):1243-9.).

Sample size for comparing two proportions

Required information:

i) Test value of the difference between proportions, p_1-p_2

ii) Anticipated values of population proportions

iii) Level of significance, Z_α

iv) Power of the test, Z_β

v) Alternative hypothesis: Two sided test : $p_1 - p_2 \# 0$

One sided test : $p_1 - p_2 < 0$ or $P_1 - p_2 > 0$

$$n = \frac{[Z_\alpha\sqrt{2PQ} + Z_\beta\sqrt{P_1Q_1+P_2Q_2}]^2}{(P_1 - P_2)^2}$$

$P_1 = 0.25, \quad Q_1 = 0.75 \qquad P_2 = 0.15 \qquad Q_2 = 0.85$

$P = (0.25 + 0.15)/2 = 0.20, \quad Q = 0.80, \quad Z\alpha = 1.96 \quad Z\beta = 0.84$

$$n = \frac{[1.96\sqrt{2 \times 0.2 \times 0.8} + 0.84\sqrt{0.25 \times 0.75 + 0.15 \times 0.85}]^2}{(0.10)^2} = 249.7 \approx 250$$

Total number of subjects required for the trial = 2 x 250 = 500

Dropout rate = 10% → Final trial size = 500 X 100/90 =556

Sample size for comparing two means

A randomized clinical trial was planned to study the effect on blood pressure reduction through dietary approaches. The study has two groups. One group receive a control diet and the other test diet. What would be the sample size in order to provide the study with a power of 90% to detect a difference in systolic BP of 2.1 mmHg between two groups at 5% level of significance. The standard deviation of systolic BP measurement is observed to be 6.5 mmHg. (N Engl J med 2001; 344(1) : 3-10).

Sample size for comparing two means

Required information:

Test value of the difference between means, δ.

Level of significance, Z_α

Required power, Z_β

Alternative hypothesis: one sides or two sided

Anticipated standard deviation of the parameter, σ

$$\text{Sample size, } n = \frac{2 (Z_\alpha + Z_\beta)^2 \ \sigma^2}{\delta^2}$$

δ = 2.1 mmHg σ = 6.5 mm Hg

α = 0.05, $Z\alpha$ = 1.96

Power = 0.90, $Z\beta$ = 1.282

$$n = \frac{2 \times (1.96 + 1.282)^2 (6.5)^2}{(2.1)^2} = 201.4 \approx 202$$

Total trial size = 2 X 202 = 404.

Estimating proportion with absolute precision

A postgraduate student aims to estimate the prevalence of physical disability in a district. How many individuals should be included in the study so that the prevalence may be estimated within 5% points of the true value with 95% confidence interval. Anticipated that true rate is unlikely to exceed 15%.

$$n = \frac{Z_{\alpha}^{2} P(1-P)}{d^{2}}$$

P = 15% = 0.15, d = 5% = 0.05, $Z_{0.05}$ = 1.96

$$n = \frac{(1.96)^{2} \times 0.15 \times 0.85}{(0.05)^{2}} = 195.9 \approx 196$$

Estimating proportion with relative precision

In problem 3, if everything else being same except absolute precision is changed to relative precision, (5%) What would be the required sample size ?

$$n = \frac{Z_{\alpha}^{2} \; p(1-p)}{d^{2} \; p^{2}}$$

P = 15% = 0.15, ε = 5% = 0.05, $Z_{0.05}$ = 1.96

$$n = \frac{(1.96)^{2} \times 0.85}{(0.05)^{2} \times 0.15} = 8707.6 \approx 8708$$

Study constraints

- Availability of resources:
  - Finance
  - Material
  - Man power
  - Logistic support
- Time
- Ethical consideration

Points to remember

Check list for sample size determination

What is the primary outcome measure?

What is the statistical test for comparison?

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?

Estimating proportion with Relative precision

A survey is being planned to determine the prevalence of diabetics in a community. With an anticipation that prevalence unlikely to exeed 65/100000, with 95% confidence interval & 10% precision what would be the required sample size?.

प्रज्ञानं ब्रह्म

Manipal

INSPIRED BY LIFE

Thank you

MANIPAL

UNIVERSITY