Concepts of Power Test (α), Sample Size and sampling

Nandkishore Singh[1],Neha Nehra[2],Shivam Dixit[3],JagMohan Dhakare[4]
1. Statistician-cum-Assistant Professor,Dept.of Community Medicine SHKM GMC Nalhar Mewat.
2.Ex-M.Ed.Student GFIE. Faridabad,
3. Statistician-cum-Tutor, Dept. of Community Medicine, ESIC, Faridabad.
4.Statistician-cum-Assistant Professor,Dept.of Community Medicine,Santosh University Ghaziyabad.

Abstract

The concepts of hypothesis testing, trade-offs between Type I and Type II error, and the use of power in choosing an appropriate sample size based on power when designing an experiment are routinely included in many surveys and research. However, many researchers do not fully grasp the importance of these ideas and are unable to implement them in any meaningful way at the conclusion of the research. This paper presents methods, questions, tables to help understand the role of power in hypothesis testing and allow obtaining numerical values without having to perform any calculations.

Introduction

The choice of test, identification of the correct calculations, and interpretation of a critical value or a p-value are some of the complications that make for a difficult topic. Conceptually need to use hypothesis testing is based on the idea of proof by contradiction. We assume that the null hypothesis is true and look for evidence in the data to either support or reject the null hypothesis. If we reject the null hypothesis, we are concluding that the alternative hypothesis or research hypothesis is more plausible. A fundamental concept of the method is the notion that we will continue to assume that the null hypothesis is true until there is overwhelming evidence against the coverage of the notion of power is understandable, it does represent a significant level in statistical use. we commonly see experimenters completely ignore this aspect of hypothesis testing in the design phase. First, may use a sample size considerably larger than needed, wasting valuable resources. More commonly, use a sample size that is too small, so that only an unreasonably large effect could be detected. For a researcher, deciding what size of difference between treatments groups will give a meaningful result is crucial to the success of the study. This assessment should typically be performed before any data is collected, in order that the results of the data do not cloud his or her judgment, and to avoid wasting resources. If do not have an adequate grasp of power.

Sample Size and Sampling.

Sample Size for Estimation of Proportion

$$n = \frac{Z^2 pq(DEFF)}{d^2}$$

Estimates for the expected proportion (p)

- Normal deviate at desired level of confidence (z)
- Level of absolute precision (d)
- Design effect (DEFF)

Anticipated Proportion (p)

- *Problem 1*: If the proportion is expected in the range, select the values closer to 50%.
  - e.g if the proportion is thought to be between 15% & 30%, use 30% for the sample size calculation.
- *Problem 2*: If the proportion is not known, 0.5 (50%) is used because it produces the largest sample size (for given value of z and d).

Z-Values

		Two sided test	**One sided test**
		$Z_{1-\alpha/2}$	$Z_{1-\alpha}$
Significance Level	**0.01**	**2.576**	**2.326**
	0.05	**1.960**	**1.645**
	0.10	**1.645**	**1.28**
			$Z_{1-\beta}$
Power	**0.80**		**0.842**
	0.90		**1.282**
	0.95		**1.645**
	0.99		**2.326**

Anticipated Proportion %	**Absolute precision** ±5%	**Relative precision** as 10 % of proportion
20	$\frac{(1.96)^2(20)(80)}{5^2} = 246$	$\frac{(1.96)^2(20)(80)}{2^2} = 1536$
40	$\frac{(1.96)^2(40)(60)}{5^2} = 369$	$\frac{(1.96)^2(40)(60)}{4^2} = 576$
50	$\frac{(1.96)^2(50)(50)}{5^2} = 384$	$\frac{(1.96)^2(50)(50)}{5^2} = 384$
70	$\frac{(1.96)^2(70)(30)}{5^2} = 323$	$\frac{(1.96)^2(70)(30)}{7^2} = 164$
90	$\frac{(1.96)^2(90)(10)}{5^2} = 138$	$\frac{(1.96)^2(90)(10)}{9^2} = 43$

Formula for sample size when prevalence is given $N = \frac{Z^2 P Q}{e^2}$

Formulae for sample size for two groups.

$$n = \frac{[Z_{1-\alpha}\sqrt{\{2P(1-P)\}} + Z_{1-\beta}\sqrt{\{P_1(1-P_1)+P_2(1-P_2)\}}]^2}{(P_1-P_2)^2}$$

Sample Size in Absence of Literature: There are three options:

- Use best guessed observations & calculate
- Do preliminary study & use its observations
- Do only preliminary study

• Probability sampling: – Simple Random Sampling – Stratified random sampling – Systematic sampling – Cluster sampling – Probability Proportion to size (PPS) – Multistage sampling	• Non-probability sampling – Convenience sampling – Snowball sampling – Quota sampling – Focus Group discussion

Variations in Sampling Process
Rural areas with many small hamlets or homesteads
Urban areas:

- Neighborhoods, census zones
- Where is the center?
- Where is the periphery?

– How to account for apartment buildings?
– In all cases, need to establish unambiguous methods; eliminate personal choice of research workers.

Testing of Hypothesis and Power of test.

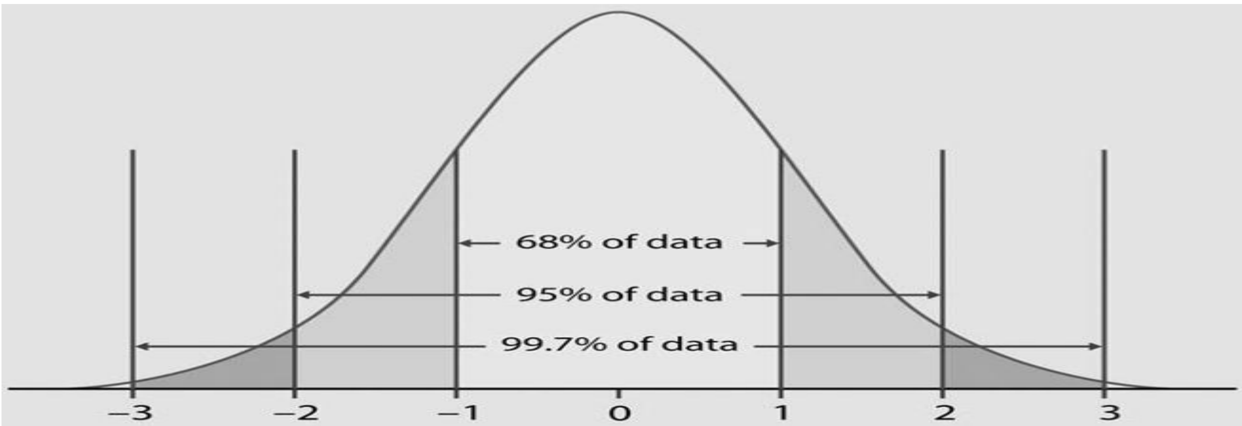
The purpose of the hypothesis test is to decide between two explanations: 1. The difference between the sample and the population can be explained by sampling error (there does not appear to be a treatment effect) 2. The difference between the sample and the population is too large to be explained by sampling error (there does appear to be a treatment effect).

Null Hypothesis (denoted H 0): is the statement being tested in a test of hypothesis. Alternative Hypothesis (H 1): is what is believe to be true if the null hypothesis is false.

Null Hypothesis: H0 Must contain condition of equality =, >, or <
Test the Null Hypothesis directly Reject H 0 or fail to reject H 0
Alternative Hypothesis: H1 □ Must be true if H0 is false =, <, >
'opposite' of Null
Example: $H0 : \mu = 30$ versus $H1 : \mu > 30$

Decision taken by the investigator	**Existing Reality** **Group A=Group B**	**Group A # Group B**
Group A # Group B	**P[Type-I Error]** **(Level of Significance)** **-**	**P[Correct Decision]** **(Power of the study)**
Group A=Group B	**P[Correct Decision]** **(Level of Confidence)**	**Type-II Error** **(1 – Level of significance)**

Step-1: State the hypotheses and select an α level. The null hypothesis, H0, always states that the treatment has no effect (no change, no difference). According to the null hypothesis, the population mean after treatment is the same is it was before treatment. The α level establishes a criterion, or "cut-off", for making a decision about the null hypothesis. The alpha level also determines the risk of a Type I error.



The α-level: **α = 0.05 = ±1.96,** **α = 0.01 = ±2.58,** **α = 0.001 = ±3.30**

Step-2: Locate the critical region. The critical region consists of outcomes that are very unlikely to occur if the null hypothesis is true. That is, the critical region is defined by sample means that are almost impossible to obtain if the treatment has no effect. The phrase "almost impossible" means that these samples have a probability (p) that is less than the alpha level.

Step-3: Compute the test statistic. The test statistic (z-score) forms a ratio comparing the obtained difference between the sample mean and the hypothesized population mean versus the amount of difference we would expect without any treatment effect (the standard error).

Step-4:A large value for the test statistic shows that the obtained mean difference is more than would be expected if there is no treatment effect. If it is large enough to be in the critical region, we conclude that the difference is significant or that the treatment has a significant effect. In this case we reject the null hypothesis. If the mean difference is relatively small, then the test statistic will have a low value. In this case, we conclude that the evidence from the sample is not sufficient, and the decision is fail to reject the null hypothesis.

Conclusions.

Concepts of power, sample size and hypothesis testing in recent years have improved. On end-of-term evaluations The "Power "concepts have also been used extensively with individual consulting to show them the impact of different sample sizes for experiments. By removing the computational burden, it is easy to discuss power with sample size. Several formula have been substantially redesigned, some with an increase in sample size, and some abandoned completely, once the power for the experiment is quantified.

This believes that the concepts will be an easily accessible tool for researcher to incorporate into their research, lectures and assignments to gain a better working understanding of the concepts and implications of power and sample size for data collection.

References:

1. Indrayan A. Medical Biostatistics ,2rd ed.
2. Austin Bradford hill, Principles of Medical Statistics, Ninth Edition, Published by Lancet Ltd.
3. Anthony N.Glaser,High Yield Biostatistics,ed.-2000.
4. B. Annadurai,A textbook of Biostatistics, New age International Publishers ed.2007.