International Journal of Preventive Medicine

Which Statistical Hypothesis Test Should I Apply? A Simple Guide for Beginners

Inaamul Haq, Aanisa Nazir[1]

Department of Community Medicine, Government Medical College, Srinagar, Jammu and Kashmir, India, [1]Center for DNA Fingerprinting and Diagnostics, Laboratory of Bacterial Genetics, Hyderabad, Telangana, India

Correspondence to:
Dr. Inaamul Haq, Department of Community Medicine, Government Medical College, Srinagar - 190 010, Jammu and Kashmir, India.
E-mail: haqinaam@yahoo.co.in

INTRODUCTION

Researchers in the field of medical and allied sciences (MAS) are generally mathematically faint at heart. The evolution from a student to a researcher is abrupt. Even though statistical methods are taught to an MAS student at the postgraduate level in most Indian universities, practical application of the same is mostly lacking. Once these students are thrown into the realm of research, they fail to apply even the basic methods in analytical statistics. A researcher works hard and honestly to collect good data but may report the wrong findings because an incorrect statistical test was used for data analysis. With this study, we intend to present a simpler way of answering the common question, "Which statistical hypothesis test should I apply?"

THE BUILDING BLOCKS

There are three essential aspects which we must understand before we can start analyzing any data.

The type of variable used

"Any aspect of an individual that is measured, like blood pressure, or recorded, like age or sex, is called a variable."[1] A "variable" may take different values in different individuals (or animals, objects, organisms, and populations) or in an individual at different times. Common examples include age, sex, weight, height, caste, religion, income, education, colony count, bacterial strain, antibiotic sensitivity, bacterial motility, bacterial morphology, and bacterial growth rate. A researcher must be very clear about the variable(s) being recorded or measured in the research. For the purpose of this study, we shall define four types of variables: "dichotomous," "polychotomous," "ordinal score," and "scale."

Dichotomous

Dichotomous variables have only two categories or levels. Variables with values such as "yes" or "no," "present" or "absent," "test" or "control" are dichotomous. Gender is a dichotomous variable with two values - "male" and "female." Disease status (present/absent), exposure status (present/absent), residence (urban/rural), antibiotic sensitivity (sensitive/resistant), type of bacterial strain (wild type/mutant), and motility (motile/immotile) are examples of dichotomous variables.

Polychotomous: Polychotomous variables are categorical variables with more than two categories. Examples include caste, religion, blood group, political affiliation, type of organism, socioeconomic status, severity of disease or symptom, and Likert scale. The categories of a polychotomous variable may or may not have an inherent order.

Ordinal score

Sometimes, polychotomous variables have more than just few ordered categories based on a ranking or scoring system. A typical example is the commonly used scale for measuring pain - the visual analog scale.[2] It has at least 10 ordered categories. Other examples include Clinical Global Impression[3] and Yale–Brown Obsessive Compulsive Scale.[4] For this study, we shall consider all polychotomous variables with seven or more ordered categories as "ordinal scores."

Scale

These are quantitative variables. The characteristic that the variable indicates can be counted or measured. The measured ones usually have some unit of measurement (years, min, kg, cm, g, etc.). Examples include weight, height, age, clinical biochemistry parameters, pulse rate, bacterial doubling time, and bacterial colony size (in mm).

Table 1: Choosing the statistical test, between-subjects design, and no repeated measures

Variable 1	Variable 2			
	Dichotomous	**Polychotomous**	**Ordinal score**	**Scale**
Dichotomous	Chi-square test	Chi-square test	Mann-Whitney U-test	Unpaired *t*-test
Polychotomous	Chi-square test	Chi-square test	Kruskal-Wallis test	ANOVA
Ordinal score	Mann-Whitney U-test	Kruskal-Wallis test	Spearman correlation	Spearman correlation
Scale	Unpaired *t*-test	ANOVA	Spearman correlation	Pearson correlation

Sometimes, the way a variable has been measured may be different than how it is being recorded or analyzed. Age is usually measured as a scale variable but may be later grouped at the time of analysis into "age groups." In such instances, variables such as age will be treated as if they were polychotomous.

Identification of the question to be analyzed

A researcher should be very clear about the question for which an answer is being sought. An MAS researcher usually wants to (a) explore possible associations between variables (find out relationship between variables), (b) determine differences between two or more groups or treatments, or (c) predict an outcome given one or more exposure variables. To apply the correct statistical test, a researcher must be able to frame his/her question clearly. "Is there an association (or relationship) between gender and exam performance?", "Is there a relationship between type of bacterial strain and sensitivity to an antibiotic?", "Is there a difference in hospital stay between patients treated with operation 'A' versus operation 'B'?", "Is there a difference in the pain score of patients before and after giving an analgesic?" These are some questions which an MAS researcher might be interested in finding an answer to.

Identifying variables in the question

Most of the questions framed by an MAS researcher relate two variables. The trick to choose the correct statistical test is to identify the two variables in the question. The question "Is there an association between gender and exam performance?" has two variables: (1) Gender, a dichotomous variable and (2) exam performance, a scale variable (when measured as the marks obtained in an exam. The question "Is there a difference in hospital stay between patients treated with operation 'A' versus operation 'B'?" has two variables: (1) Hospital stay, a scale variable and (2) type of operation, a dichotomous variable with two values - operation "A" and operation "B."

CHOOSING THE CORRECT STATISTICAL TEST

Once the researcher is ready with the building blocks (the question to be answered, the two variables, and their type), it becomes easy to choose the correct statistical test. Tables 1 and 2 provide an easy guide.

In a between-subjects design, measurements are made on two or more different groups of "cases" (individuals,

Table 2: Choosing a statistical test, within-subjects design, and repeated measures

Variable of interest	Two repeated measures (before and after studies)	>2 repeated measures
Dichotomous	McNemar Chi-square test	Cochran *Q*-test*
Polychotomous	NA	NA
Ordinal	Wilcoxon signed rank test	Friedman test*
Scale	Paired *t*-test	Repeated-measures ANOVA

*Rarely used tests because such situations are rare in medical and allied sciences research. NA=Not available

animals, objects, organisms, etc.). In a within-subjects design, measurements are made on the same cases at two (before and after treatment) or more (baseline, at time t_1, and at time t_2) different times or through two or more different techniques (e.g. ELISA and Western Blot).

DISCUSSION

This write-up has been written for beginners in research. This is not an all-encompassing text in statistical methods. We have omitted tests used for predictions because we think those to be beyond the scope of a beginner in research. Mentioning about regression analysis would be akin to opening the Pandora's box.

The tests mentioned in this write-up need to be used with caution. Most of these tests are based on certain assumptions which a beginner may find difficult to understand and test. In general, when we have a large dataset (more than 30 or 40 "cases"),[1,5] most of the test assumptions are met and the tests can be used effectively. Problems might arise with small datasets (<30 "cases"). Hence, caution should be exercised in "small dataset" situations, and an expert advice should be sought.

With the advent of computers and advances in the field of machine computing and internet, researchers can access free online sources for performing most of the tests mentioned in Tables 1 and 2. Open Source Epidemiologic Statistics for Public Health is one such source available at http://www.openepi.com/Menu/OE_Menu.htm. Other useful sites are VassarStats (http://vassarstats.net/) and GraphPad QuickCalcs (http://www.graphpad.com/quickcalcs/).

REFERENCES

1. Kirkwood BR, Sterne JA. Essential Medical Statistics. 2nd ed. Oxford: Wiley-Blackwell; 2003. p. 501.
2. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990;13:227-36.
3. Busner J, Targum SD. The clinical global impressions scale: Applying a research tool in clinical practice. Psychiatry (Edgmont) 2007;4:28-37.
4. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, *et al.* The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Arch Gen Psychiatry 1989;46:1006-11.
5. Daniel WW. Biostatistics: A Foundation for Analysis in the Health Sciences. 9th ed. New York: John Wiley & Sons; 2009. p. 783.

Access this article online	
Quick Response Code:	**Website:** www.ijpvmjournal.net/www.ijpm.ir
	DOI: 10.4103/2008-7802.183653