ABC OF STATISTICS

Statistics in Medical Research

Dinesh Kumar

This article is 2nd in series of articles on 'Statistics in Medical Research' trying to explore the subject in a simplistic manner without ignoring the fundamental framework.

Knowledge of statistical principles alone has a limited utility since its full application requires additional expertise of epidemiological principles especially for those researchers who wish to pursue research independently. Therefore, a brief introduction about various epidemiological designs, formulation of research questions and objectives, measurement of variables, questionnaire formulation, issues related to internal and external validity (generalizability) and sample size considerations shall be an integral part of the present series which predominantly focus upon the statistical principles. Few more commonly used terms need clarification before we proceed with first step i.e., presentation of data.

Parametric and Non-Parametric Statistic

Parametric Statistics : Some assumptions are to be met before a particular test of significance can be applied to a set of data. Strictly speaking, the investigator needs to ensure that these conditions are met. Sample measurements drawn from normally distributed population of measurements in a random manner, measurement of these on a continuous scale and issues related to variance (variability of measurements in a dataset) are prerequisites for the application of tests of significance belonging to parametric statistics. Unless these assumptions are met, the validity of the conclusions drawn is questionable. Commonly employed parametric tests are Student 't' test (paired and unpaired), F test for analysis of variance, correlation and regression analyses etc. Some of these tests require advanced mathematical skills and are time consuming if calculations are manual.

Non-Parametric Statistics or Distribution Free Methods : Doesn't require all assumptions to be met as in case of classical parametric statistical technique. The nonparametric tests are suitable alternative particularly when the data is in the form of ranks or counts. Also, these could be gainfully used when sample size is a problem. Chi- squared test (pronounced as Ky) is the most commonly employed nonparametric test in practice. Wilcoxan Rank Sum, Mann Whitney U or median test , Kruskal-Wallis 1-way and Friedman 2-way analysis of variance are some of the nonparametric equivalent to 't' test and F test for analysis of variance. These tests can be performed manually as they don't involve extensive calculations.

Statistic and Parameter : The measurements made on a set of data for describing its characteristics e.g., mean and standard deviation are used to draw conclusions about the population values. When measurements are computed from a sample they are termed as *statistics* whereas the population measurements are known as *parameters*.

Statistics	Parameter
$\overline{X}$ or $\overline{Y}$ (pronounced X bar or Y bar)	μ (mu)
S or SD (standard deviation)	σ (sigma)

Example : A researcher was interested in determining the average blood sugar in healthy adults residing in an urban area. Employing probability- sampling (random) methods, the researcher selected 100 healthy adults. They were subjected to blood sugar measurements according to a standardized protocol. In this example, the average blood sugar computed from a sample of 100 healthy adults is referred as *statistic* and the *parameter* of interest is the average blood sugar of all healthy adults residing in the urban area. In reality, we rarely know about the value of the population parameters since we are rarely

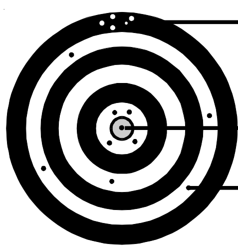
From the Postgraduate Department of Preventive & Social Medicine, Govt. Medical College, Jammu 180001 (J&K) India.
Correspondence to : Dr. Dinesh Kumar, Statistical Editor, JK Science, Journal of Medical Education and Research.

able to study the whole populations. That is why we often select a representative portion of the population (sample) and compute the measures of average and dispersion.

Thus, a sample statistic is measured to estimate a population parameter

Validity and Reliability The accuracy of a measurement is defined in terms of validity and reliability. Let us try to understand the difference between the two with the help of an example. Consider three shooters who are asked to hit the bull's eye. Each one of them is allowed to make five attempts. The results of the attempts made by them are displayed in Fig. below.

Which one of them is Valid and Reliable?



Did you guess it right ?

Yes, it is the "First Shooter" because he is hitting the place he is supposed to and in a manner, which is very consistent.

The "Second Shooter" is not valid though he is very reliable as he is hitting nearly the same place all the time.

The "Third Shooter" is neither valid nor reliable.

A formal definition of Validity and Reliability might read like this:

Validity, Measurement : An expression of the extent to which a test/measurement measures what it purports to measure. As illustrated above, it is considered equivalent to a marksman's capacity to hit the bull's eye. A measure may be of different validity in different groups or populations e.g., cough and sputum production may have a low validity for the diagnosis of chronic bronchitis in areas having high prevalence of pulmonary tuberculosis or presence of fever may not be a valid measure of malaria in areas with low transmission of malaria. Different forms of validity like face validity; content validity; construct and criterion validity; predictive validity shall be discussed in the present series as and when felt necessary to ease understanding.

Validity, Study : The degree to which the inferences drawn from the study, especially generalizations extending beyond the study sample are warranted when account is taken of the study methods employed, the representativeness of the study sample and the nature of the population from which it is drawn. The two key subsets of study validity are *Internal and External Validity*. While the internal validity reflects on the quality (unbiasednesss) of measurements on the subjects under study, the external validity deals with the larger issue of applicability of the conclusions drawn to the population of interest.

It is therefore clear that for a study to be externally valid, it has to internally valid first.

Reliability or Reproducibility on the other hand refers to the ability of the test to produce consistent results when independently repeated and interpreted under nearly identical circumstances. It is measured by performing two or more independent measurements of the same attribute and comparing the findings e.g., two or more interviewers seeking the same information from same respondents, measurement of blood pressure at different times in the same subjects made by the same investigators. Lack of reliability may arise from divergences between observers or instruments of measurements or instability of the attribute being measured.

So, which of the two needs to be fulfilled : Validity or reliability, ideally both. But it is clear that if the measure is not reliable, this must reduce its validity. Therefore, validity should probably be defined assuming that the measure is reliable.

Suggested reading

1. Last JM. A dictionary of epidemiology, 3rd Ed. New York, Oxford University Press, 1995.
2. Bradford Hill A. A short Textbook of medical statistics, 11th ed. Hodder and Stoughton, London, 1984.
3. Knapp RG. Basic Statistics for Nurses, 2nd ed. New York, John Wiley & Sons, 1985.
4. Lwanga SK, Tye Cho-Yook. Ayeni O. Teaching health statistics, Lesson and seminar outlines, 2nd ed. Geneva, World Health Organization, 1999.
5. Abramson JH. Survey Methods in Community Medicine. An Introduction to epidemiological and evaluative studies, 3rd ed. Churchill Livingstone, New York, 1984.