Journal of Islamabad Medical & Dental College (JIMDC);2013;2(2):103-103

Biostatistics and Data Types

Muhammad Afzal and Farwa Rizwi
Department of Community Medicine, Islamabad Medical and Dental College, Islamabad
(Bahria University, Islamabad)

Biostatistics is the science which deals with development and application of the most appropriate methods for the:

- Collection of data.
- Presentation of the collected data.
- Analysis and interpretation of the results.
- Making decisions on the basis of such analysis

Therefore, when different statistical methods are applied in biological, medical and public health data they constitute the discipline of biostatistics.

Different characteristics that take different values in different persons, places and things, or the characteristics and properties, we wish to observe by members of a group (sample) which differ from one another are called "Variables" in statistics. All the information regarding all the variables in the study is called data. There are two main types of data:

1. Qualitative variable: It is a variable or characteristic which cannot be measured in quantitative form but can only be identified by name or categories, for instance place of birth, ethnic group, type of drug, stages of breast cancer (I, II, III, or IV), degree of pain (minimal, moderate, severe or unbearable).

2. Quantitative variable: A quantitative variable is one that can be measured and expressed numerically and it can be either discrete or continuous. The values of a discrete variable are usually whole numbers, such as the number of episodes of diarrhea or number of children. A continuous variable is a measurement on a continuous scale. Examples include weight, height, blood pressure, age, etc.

Although the types of variables could be broadly divided into categorical (qualitative) and numerical (quantitative), it has been a common practice to see four basic types of data (scales of measurement)

i. Nominal data represent categories or names. There is no implied order to the categories of nominal data. In these types of data, individuals are simply placed in the proper category or group, and the number in each category is counted. Each item must fit into exactly one category.

ii. Ordinal Data have order among the response classifications (categories). The spaces or intervals between the categories are not necessarily equal. For example: strongly agree, agree, no opinion, disagree, strongly disagree. In this situation, we only know that the data are ordered.

iii. Interval Data: In interval data the intervals between values are the same. For example, in the Fahrenheit temperature scale, the difference between 70 degrees and 71 degrees is the same as the difference between 32 and 33 degrees. But the scale is not a RATIO Scale. 40 degrees Fahrenheit is not twice as much as 20 degrees Fahrenheit.

iv. Ratio Data: The data values in ratio data do have meaningful ratios, for example, age is a ratio data, and someone who is 40 is twice as old as someone who is 20.

Both interval and ratio data involve measurement. Most data analysis techniques that apply to ratio data also apply to interval data. Therefore, in most practical aspects, these types of data (interval and ratio) are grouped under metric data.

In some other instances, these type of data are also known as numerical discrete and numerical continuous.

Numerical discrete

Numerical discrete data occur when the observations are integers that correspond with a count of some sort. Some common examples are:

- the number of bacteria colonies on a plate,
- the number of cells within a prescribed area upon microscopic examination,
- the number of heart beats within a specified time interval,
- a mother's history of number of births (parity) and pregnancies (gravidity), etc.

Numerical continuous:

The scale with the greatest degree of quantification is a numerical continuous scale. Each observation theoretically falls somewhere along a continuum. One is not restricted, in principle, to particular values such as the integers of the discrete scale. The restricting factor is the degree of accuracy of the measuring instrument. Most clinical measurements, such as blood pressure, serum cholesterol level, height, weight, age etc. are on a numerical continuous scale.

References

1. Text book of Community Medicine by Park J E. Latest Edition.
2. Text book of Community Medicine. 6th Ed. by Ilyas Ansari.
3. Text book of Community Medicine by Maxie Rozani. Latest Edition.
4. Medical Statistics. 2nd Ed. by R. Turkwood.
5. Rizvi F. The Nature and purpose of research. International Journal of Pathology; 2011; 9(1): 35-39.