ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 10, March 2018

Various Statistical Methods Used In Computer Application for Data Analysis

Arnica Jain
Assistant Professor in Economics, S.G.G.J. Girls College, Raikot, Punjab, India

Abstract - This article includes various statistical methods used in the field of computer application. This article discusses some general methods for data interpretation like concepts of central tendency, standard deviation, dispersion, correlation, regression, hypothesis testing.

Keywords- Statistical methods, Frequency, Distribution, skewed, correlation, regression.

1. INTRODUCTION

The word 'statistics' conveys a variety of meaning to people. Statistics is an imposing form of mathematics whereas to others it is tables , charts and figures. In this information age , data is no longer scarce - its overpowering. The key is to sift through the overwhelming volume of data available to organisations and businesses and correctly interpret its implications. But to sort through all this information, we just need the right statistical methods or tools. Before discussing about the statistical methods lets discuss about the meaning of statistical methods.

2. What is Statistical Methods?

According to W.E. Deming, "The statistical methods is more than an array of techniques. The statistical methods is a mode of thought, it is sharpened thinking, it is power."

The fast speed of computers have substantially changed the field of statistics. The use of computer software has brought about a technological revolution. Statistical methods deal with the collection, classification, presentation, comparison and interpretation of statistical data. By their use, useful deductions and inferences are drawn. Following are the main statistical methods:

2.1 Mean, Median and Mode

The arithmetic mean, more commonly known as 'the average' is the sum total of all the values in a series divided by the number of items in the series. The mean is useful in determing the overall trend of a data set or providing a rapid snapshot of your data. Another advantage of the mean is that it is very easy to understand and quick to calculate.

Median is a middle value in a series. It divide the series into two equal parts. one part is having upper values and other lower values. Median is just the 50th percentile value below which 50 percent of the values in the sample fall.

The mode or modal value is that value in a series of observation which occurs with the greatest frequency i.e. the mode is often said to be that value which occurs most frequent.

2.2 Relationship between Mean, Median and Mode

The relationship between mean, median and mode to each other can provide some important information about the relative shape of the data distribution. If the mean, median and mode are approximately equal to each other, the distribution can be assumed to be symmetrical, if mean>median>mode, the distribution will be skewed to right. If mean<median<mode, the distribution will be skewed to left.

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 10, March 2018

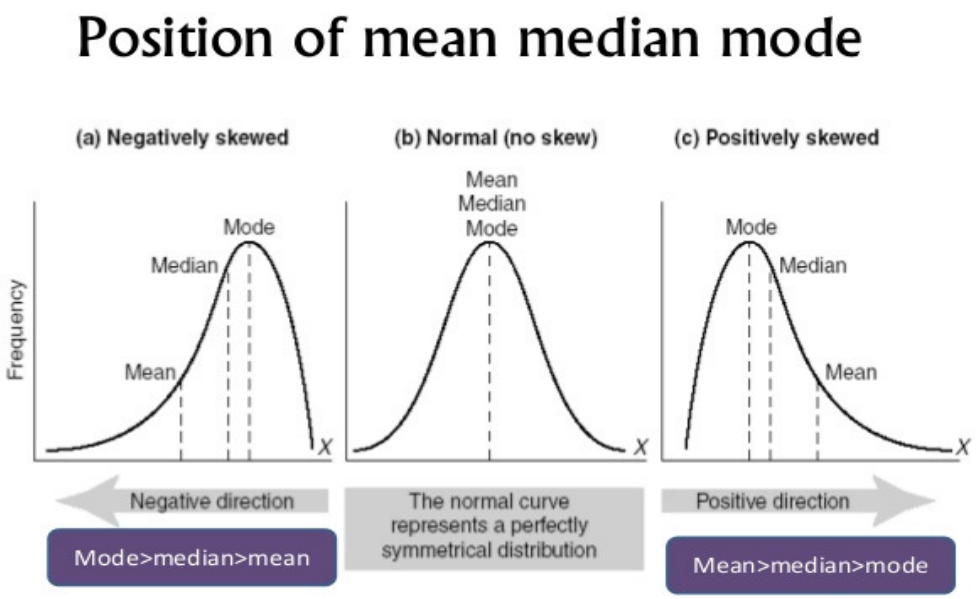

Fig. 1

2.3 Geometric and Harmonic Mean

Geometric mean is defined as the Nth root of the product of N items or values. If there are two items, we take the square root; if there are three items, then cube root, and so on. Of course with large n this can be difficult to calculate then we will take the logarithmic transformation of the geometric mean.

Harmonic mean is based on the reciprocals of numbers averaged. It is defined as the reciprocal of the individual observation.

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 10, March 2018

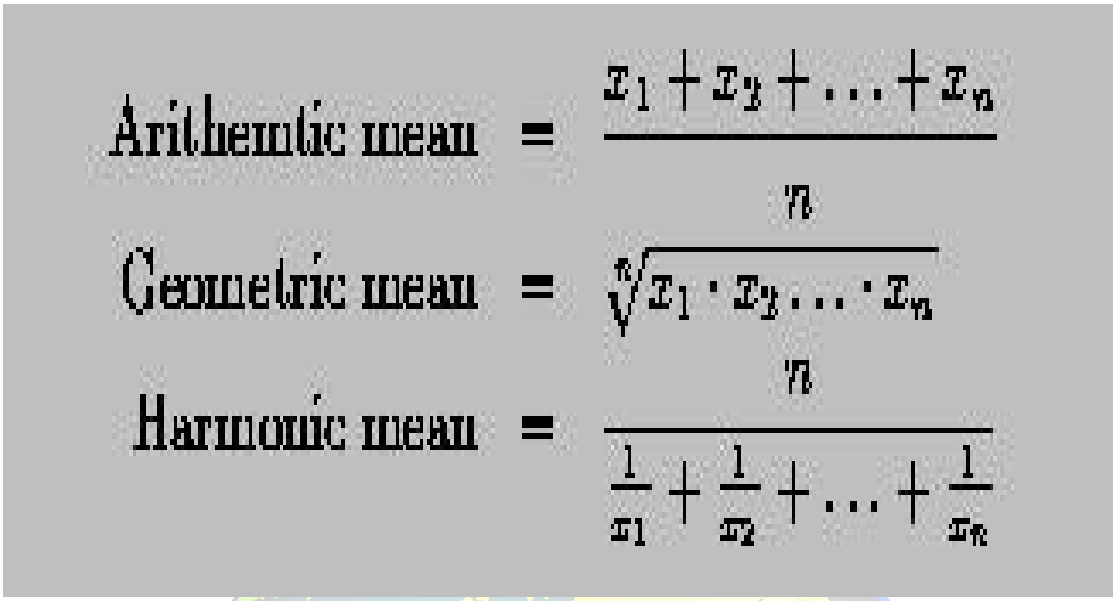

Fig. 2.

The range of a sample(set of data) is simply the maximum possible difference in the data i.e. the difference between the maximum and the minimum value. A more exact term for it is 'range width' and is usually denoted by the letter R and w. the two values (maximum and minimum) are called the "range limits"

2.5 Standard deviation

The standard deviation, often represented with the Greek letter sigma, is the measure of a spread of a data around the mean. A high standard deviation signifies that the data is spread more widely from the mean, where a low standard deviation signals that more data align with the mean. In a portfolio of data analysis methods, the standard deviation is useful for quickly determining dispersion of data points.

2.6 Correlation

Correlation is are measure of the degree of relationship between two variables. Two variables are said to be correlated if change in one variable result in some change in other variable. Correlation is said to exist when two variable vary together directly or indirectly. If they do not vary together, then correlation is said to be absent. So we study the nature and extent of relationship between the variables.

According to Croxton and Cowden, "When the relationship is of a quantitative nature, the appropriate statistical tool for discovering and measuring the relationship and expressing it in a brief formula is known as correlation."

2.7 Regression

The dictionary meaning of term regression is to turn back or revert. Regression is a statistical tool with the help of which one can estimate the unknown values of one variable from known values of another variable. Out of these two variables, one is treated as independent and other as dependent variable. A line which represents the average relationship between two variables is called regression lines. The regression line also designates whether those relationship are strong and weak. Regression is commonly taught in high school or college. Statistics courses with applications for science or business in determining trends over time.

2.8 Hypothesis Testing

Also commonly called testing, hypothesis testing assesses if a certain premise is actually true for your data or population. In data analysis and statistics, you consider the result of a hypothesis test statistically significant if the result couldn't have happened by random chance. Hypothesis tests are used in

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 10, March 2018

everything from science and research to business and economics.

References

[1]https://en.wikibooks.org/wiki/Statistics/Summary/Averages/mean

[2]https://en.wikipedia.org/wiki/Statistical_hypothesis_testing

[3]https://www.mathalino.com/reviewer/derivation-of-formulas/relationship-between-arithmetic-mean-harmonic-mean-and-geometric-mean

Statistical Methods by S.P. Gupta

Statistical Methods By C S Aggarwal

Biography

Arnica Jain
(M.Sc (Hons.School),B.Ed,B.A.)
97818-68294
Assistant Prof. At Swami Ganga Giri Janta Girls College, Raikot(LDH.)