

## **Basic Statistical Concepts in Environmental Science: An Introduction**

**Bagsmrita Bhagawati**

(Department of Statistics, B.H.College, Howly Assam, India)

---

**Abstract:** *In his forward march towards civilization man was accompanied both by environment and Statistics. Since early days man was found to be knowingly accustomed with environment and unknowingly played with Statistics. Thus both Statistics and environment have shared a long history of mutual reciprocation. In modern times these two subjects are independently able to attract academic attention of the scholars throughout the world. But still one may observe an organic relationship between these two disciplines. The present work has made a modest attempt to discuss some basic statistical concepts used in Environmental Science, thereby trying to establish link between the aforementioned subjects.*

**Keywords:** *environment, Environmental Science, organic relationship, Statistics*

---

### **I. Introduction**

In his forward march towards civilization man was accompanied both by environment and Statistics. Since early days man was found to be knowingly accustomed with environment and unknowingly played with Statistics. Thus both Statistics and environment have shared a long history of mutual reciprocation. In modern times these two subjects are independently able to attract academic attention of the scholars throughout the world. But still one may observe an organic relationship between these two disciplines. The present work has made a modest attempt to discuss some basic statistical concepts used in Environmental Science, thereby trying to establish link between the aforementioned subjects

### **II. Basic Concept Of Environment**

Environment means the surrounding external conditions influencing growth of people, animals or plants; living or working conditions etc. Environmental science is an integrated multidisciplinary approach that studies environment and solutions of environmental problems. In present scenario, environment has become a world agenda and hence increased the scope and importance of the study of environmental science.

### **III. Meaning Of Statistics**

The word 'Statistics' originates from the Latin word 'status', Italian word 'statista' and the German word 'statistik', which means political state. This was actually considered as the by product of the administrative activity of the State. In ancient period information were collected only to get an idea about the manpower and wealth of the nation by the rulers. The concept of modern statistics came during the mid seventeenth century with the introduction of theory of probability and games of chance. Thus statistics now-a days deal with collection, presentation, analysis and interpretation of numerical data. The Oxford English etymological dictionary defines Statistics as 'first applied to the political science concerned with the facts of a state or community XVIII; all derived immediately from German *statistisch*'. The subject statistics, hence deals with exploration, summarisation and inferences about the state of a nation, the state of the health of the people, the state of the environment etc. [1].

### **IV. Application Of Different Statistical Tools**

In modern times statistical theory and methods have been applied to various fields of life. One such important discipline is studying environment i.e. in collecting and planning data, in designing surveys and experiments etc. for an environmental statistics. [2, 3, 4]. In this paper we try to study some basic statistical concepts relevant to environmental study.

#### **4.1 Data Analysis**

Data analysis is divided into two sections: descriptive statistics and inferential statistics.

##### **4.1.1 Descriptive statistics**

Descriptive statistics is the initial stage of data analysis where exploration, visualisation and summarisation of data are done. In this section we study about the definition of population and random sample. Different types of data viz. quantitative or qualitative, discrete or continuous are helpful for studying the

features of the data distribution, patterns and associations. The frequency tables, bar charts, pie diagrams, histograms etc. represent the data distribution: position, spread and shape efficiently. This kind of statistical technique are useful for interpreting the information contained in the data and hence for drawing conclusion. Further, different measures of central tendency viz. mean, median, quartiles, percentiles etc. were calculated for analysing environmental quantitative data. It is also useful to study dispersion measures such as range to measure variability in small samples. The sample variance and standard deviation are found to be suitable where the sample mean is considered as the ideal measure of central tendency. One of the important measure of relative dispersion is coefficient of variation, and it is useful for comparison variability of data with different units. Skewness and kurtosis characterise the shape of the sample distribution. The concepts of association and correlation demonstrate the relationships between variables and are useful tool for clear understanding of the linear and non linear relationships.

**Some examples where descriptive statistics are used have been given below:**

**Example 1:** This example is discussed by Scheaffer *et al.*(1996). A questionnaire was distributed among the students seeking their personal information such as age, gender, weight or height, or their opinion about a critical environmental problem or an environmental factor which they consider more important where they live, etc. Students can construct the two way table and summarise the results using the descriptive statistics. Also they can study the existing relationship between gender and the chosen critical environmental problem using the descriptive analytical tool.

**Example 2:** The Technical University of Valencia work with local groups and individuals around the coast of Europe for protection and sustainable use of coastal resources and informing public participation in environmental planning and management. Different descriptive statistics are applied to collect and gather information.

**Example 3:** Further atmospheric variables viz. pollutants and climatological parameters are measured in Valencia regularly. These are descriptively analyzed by studying the relationship between ozone and nitrogen oxides and meteorological parameters such as solar radiation, wind speed and temperature. Several graphical and numerical tools are used to interpret the results.

#### 4.1.2 Inferential Statistics

In inferential statistics, concept of probability is important for studying the uncertainties in environment. For example, whether it will rain or not tomorrow can be best inferred by using probability. Several theoretical probability distribution viz. Bernoulli distribution, binomial distribution, Poisson distribution etc. are useful for modelling probability distribution of real environmental data. For example decisions such as coin tossing, rain/no rain, yes/no etc. are explained by Bernoulli variable since their outcomes are binary. Also if we are interested in counting the number of times flood occurs in Dhemaji district of Assam, India out of a total number of floods occurred. The probability distribution of such variable is given by Binomial distribution since we are to count the number of times flood occur ( $X$ ), a Bernoulli event, with probability  $p$  out of a total i.e. out of  $n$  trials. Also, if we do not know the total number of occurrence of floods, but know the mean of the occurrence of flood then the distribution is modelled by Poisson distribution. Statistical tools such as estimation i.e. estimating population parameters from sample values, hypothesis testing etc. plays a vital role in analysing environmental data. Some of the frequently used test statistic in atmospheric and environmental science is  $Z$  test,  $t$  test,  $F$  test etc. Another statistical approach is time series analysis, which study environmental quantities with respect to time. For example, monthly/yearly mean temperature, rainfall, humidity etc is best studied by time series.

## V. Conclusion

The statistical concepts act as a valuable tool for monitoring environmental systems. In agricultural activities viz. timing of cropping and harvesting, timing of chemical applications, type of crops planted and irrigation scheduling etc. required the knowledge of environmental statistics. The forestry activities of a country viz. extraction of timber, afforestation, reforestation projects etc need statistical information. Thus application of statistics is important for environmental science for effective and innovative monitoring of environmental variables over the time.

## Reference

- [1]. D.B. Stephenson, Statistical concepts in environmental science, retrieved from <http://www.met.rdg.ac.uk/cag/courses/Stats/> 2003.
- [2]. J. B. Garfield, How students learn statistics, *International Statistical Review*, 63, 1995, 25-34.
- [3]. R. L. Scheaffer, M. Gnanadesikan, A. Watkins and J. Witmer, *Activity-Based Statistic*, New York, Springer-Verlag, 1996.
- [4]. D. S. Moore, New pedagogy and new content: The case of statistics, *International Statistical Review*, 65, 1997, 123-137.