

THE TEACHING OF STATISTICS AT THE TERTIARY LEVEL IN GHANA
THE SCIENCE STATISTICS: AN INTRODUCTORY LECTURE
TO THE FIRST YEAR

Nsowah-Nuamah Nicholas Nicodemus Nana, University of Ghana, Ghana

Statistics at the tertiary level may be read in Ghana at the Polytechnics for the High National Diploma or at the University for the Diploma or the Degree. Students get sceptical about the subject especially when they gain the impression they have from school is that Statistics is a branch of Mathematics. During my first lectures in Statistical Methods I usually ask the question “why they have chosen to read Statistics?” After discussing the importance of Statistics I explain with examples why Statistics is not a branch of Mathematics, something students would always want to know. This paper seeks to discuss my approach which aims at convincing Students that Statistics is a separate Science which uses the knowledge in Mathematics and which will create a strong interest for the subject.

BE AWARE OF STATISTICS

History of Statistics

The word Statistics, as far as is known, was first used by a German Professor , Gottfried Achenwall about 1770. It was then defined as “the Science that teaches us what is the political arrangement of all the modern states of the new world”. As it can be seen, the word was originally referred to collections of facts (not necessarily numerical) about the state, or the people who composed the State. The facts were then given in a verbal form.

In 1834 the Royal Statistical Society was founded, and in its first journal, issued in 1838-9, defined statistics as “the ascertaining and bringing together of those facts which are calculated to illustrate the condition and prospects of society”, and that “the statisticians commonly prefers to employ figures and tabular exhibitions”. Thus Statistics which started as verbal descriptions was now mostly concerned with numerical facts. Further changes of meanings took place and the word came to stand for the numbers themselves. Thus figures relating to economics are known as “economic statistics”, to education as educational statistics, to industries as “industrial statistics”, to agriculture as “agricultural statistics”, and so on.

Statistics now generally has a dual meaning. It may be used in a singular or a plural sense. In a plural sense, the term “statistics” carries only the nebulous and too often distasteful connotation of figures. It is used to indicate a group of numbers or collections of numerical data. Here statistics are the figures themselves, suitably classified and tabulated together with any secondary statistics such as percentages or averages derived from them.

The term “statistics” in a singular sense is used to refer to a subject field of study. It is the science that studies methods and procedures for collecting, organising, analysing and interpreting numerical data and for making scientific inferences from such data. It is better described as statistical methods.

The term “Statistics” has again another meaning for those who have been initiated into the mysteries of the subject “Statistics”. In this sense, “statistics” are quantities that have been calculated from sample data. The word is plural when used in this sense. A single quantity that has been so calculated is called a statistic. For example, the average salary of workers is a statistic.

Branches Of Statistics

Statistics is a science because it is a subject field of study and it breaks naturally into two reasonably distinct branches: descriptive statistics and inferential statistics. Descriptive statistics is concerned with the descriptions, presentation and summarisation of a set of data in order to properly describe the various features of that data set. The area of descriptive statistics involves through the use of graphic, tabular, or numerical devices, the abstraction of various properties of sets of observations. Such properties include the frequency with which various values occur, the notion of a typical or usual value, the amount of variability in a set of observations, and the measurement of relationships between two or more variables.

The ultimate objective of the investigator is to draw conclusions, and so the descriptive part of statistics should be looked at as a sort of preliminary to the main bout. Inferential statistics is concerned with drawing conclusions regarding a large data on the basis of the information obtained from a small portion of the larger collection.

It is the process of inferential reasoning that enables the quality control manager to decide whether the quality of a product is and the economist to make predictions of future economic conditions. The terms statistical inference, “inductive statistics” and mathematical statistics are also used for inferential statistics. We may also distinguish between Classical Statistics and Bayesian Statistics (or sometimes referred to as bayesian decision analysis).

Classical Statistics is concerned with the analysis of sampled (objective) data for the purpose of inference, with the exclusion of any personal judgement or opinions. Bayesian statistics incorporate managerial judgements in the statistical analysis and also places emphasis on the possible economic gains or the possible losses associated with alternative decision acts. Just as there are different kinds of statistics so there are different kinds of statisticians. Indeed the subject of statistics is now so wide that it is impossible for one man to become expert in all branches.

Statistics and Mathematics

Statistics differs from subjects like history or chemistry in that it is a methodological field. In this respect it resembles fields like accounting and mathematics and most people equate statistics to mathematics. But this is not correct. Statistics differs from mathematics in several ways.

Origin

Statistics did not originate from within mathematics but in official and private data gathering. The words “Statistics” and “States” are derived from the same root “Statist” which in the sixteenth century was used to describe a politician - a dealer in facts about the state, its government and its people. Rulers of ancient Babylonia, Egypt and Rome and most civilised countries from earliest times gathered detailed information of populations and resources :

- i. to levy taxes to maintain the state and the court;
- ii. to ascertain, for military and fiscal reasons, manpower and material strength of the nation.

Statistics (like physics and economics) certainly makes heavy and essential use of mathematical tools. It was unfortunate that the teaching of statistic at the early stages of the development of the course was captured by academic statisticians who were mathematicians.

Aims and Methods

The aims and methods of statistic are not those of mathematics. Statistics is concerned with the collection, organisation and analysis of data and with inference from data to the underlying reality. The dominant approach in practice is a complex interplay between the data and a mathematical model. The model may be partly validated by carefully designed data collection, or by the data operating through the diagnostic tools. Mathematical derivation of the consequence of a model is an essential prerequisite for this process, but this is an application of mathematics to statistics parallel to the equally essential applications of mathematics in economics and physics. Mathematics (pure) proceeds from a system of axioms, or assumptions, to a collection of theorems (truths) and counterexamples (falsehoods) which can be established within such a system (Spanier). Work in Statistics begins with statement of a problem which has arisen in a fields such as economics, sociology, biology, health etc. It is in this wise that Fisher considered the Science of Statistics as “essentially a branch of Applied Mathematics and may be regarded as mathematics applied to observational data” (Statistical Methods for Research Workers, 1970). (The book was written in 1920.) Thus, an essential distinction

between mathematics and statistic is that mathematical theorems are true, but statistical methods are often effective when used with skill.

Uniqueness of Solutions

There is a part of statistics that is intrinsically non-mathematical. There are not always unique solutions in statistical analysis. Statistical conclusions involve an element of uncertainty which can lead to incorrect conclusions. This element of uncertainty sets statistics apart from other areas of applied mathematics which require conclusion that are deduced with certainty from conclusions and assumptions.

Deduction And The Power Of Abstract Structures Are Central In Mathematics

Whereas In Statistics, It Is The Problems Of Data, Uncertainty And Scientific Inference That Are Central.

a) Probability and Statistics Probability

Theory provides an important tool for the analysis of any situation (in science, in business, or in everyday life) which in some way involves an element of uncertainty or chance. Probability theory came from the investigation of games of chance. Statistical theory and methodology depend heavily of probability theory. In a sense, a problem in probability is the opposite of a problem in statistical inference. In statistical inference, as stated above, statements about a population are made from analysing data obtained about a sample. In probability we start with a population about which we have complete information and then proceed to make statements about a sample. Consider a deck of ten cards marked 1,2,3,4,5,6,7,8,9,10, respectively. The ten cards constitute the population under consideration. If we draw three cards from the deck, the three cards form a sample. Probability theory is concerned with statements which can be made about this sample on the basis of our complete knowledge of the population. Statistical inference is concerned with statements which can be made about the population, on the basis of information on the sample.

It is not exactly proper to describe probability and statistics as being “opposite”, since they cannot really be separated from each other completely. Sometimes it is difficult to tell whether a given problem is strictly a problem in probability or a problem in statistical inference. A problem which is primarily one of statistical inference will use the theory of probability in making the inferences. Although we have oversimplified it, it is convenient for an introduction to elementary probability to give the general description as we have done and which we now repeat: From sample to total population (statistics). From population to sample (probability).

b) Uses Of Statistics

The purpose of statistics is not simply to satisfy curiosity but to enable correct decisions to be taken. Statistics is useful in a wide range of human activities. Now-a-days, statistical devices are used in economics, sociology, medicine and even in corporate law. Solomon Fabricant said some years ago, “The whole world now seems to hold that statistics can be useful in understanding, assessing, and controlling the operations of society”.

The tools of statistics are essential to both the state machinery and the private organisations in areas such as planning, forecasting, and quality control. It is an indispensable tool of the Government in any attempts to regulate the economy. The results of statistical analysis may also be used for description, comparison, projection and prediction.

BEWARE OF STATISTICS

Statistics, when used properly, is a valuable tool to society. Data are used frequently to support a certain position because numerical information is often viewed as being objected and therefore beyond criticism (Tim and Hoog). However, many individuals, business and advertising firms misuse statistics to their own advantage. Some years ago Benjamin Disraeli said that “there are three kinds of lies: lies, damned lies, and statistics”. The first lies are referred to as white lies, which is justifiable; the second are common lies - these have no justification and the third are statistics. What this means is that statistic is the highest level of lies. Others think that “anything can be proved by figures (statistics)”, however, it is important to note that the figures themselves cannot mislead or lie, but those who present the figures certainly can. Thus, “figures don’t lie, liars figure”. Some people use statistics like a drunkard uses a lamppost: “more for support than for illumination”. In fact a great deal of meaningless and unfounded talk is presented to the public in the name of statistics. Many a time we hear people supporting their arguments by the words “statistics show that ...”. Many people seem to believe that a case is proven if statistics can be produced to support it. Thus an endless stream of data is thrown at us in an attempt to impress, persuade or even coerce us into accepting or refusing certain views.

There are so many ways statistics can be misused.

- (i) The data can be used for the wrong purpose, that is, one that is different from the purpose for which they were collected.
- (ii) The data can be collected incorrectly so that they are biased.
- (iii) The data can be analysed carelessly so that the results obtained from them are misleading.

- (iv) The inappropriate statistical technique may be used for correctly collected data.
- (v) The appropriate statistical technique can be used for incorrectly collected data.

Sometimes, the misleading presentation of data is done consciously or unconsciously. The typical abuse has personal objectives and willing to suppress unfavourable data while emphasising supportive data.

The misuse of statistics has prompted people to remark that “statistical analysis has often meant the manipulation of ambiguous data by means of dubious methods to solve a problem that has not been identified”. Other think that statistic analysis are meaningless and should never be accepted to support a particular view.

BE WORRY FOR LACK OF KNOWLEDGE OF STATISTICS

The knowledge of statistics is crucial to every individual of a country. Statistics is an extremely important tool, probably ranking in importance behind only the traditional communication skills of reading, writing, speaking, and listening. Many years ago H.G. wells commented that statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write. Oliver Wendell Holmes Jnr. stated the matter in 1896 as “the man of the future is the man of Statistics and the master of economics.”

Doublespeak exist in statistics as much as in the written or spoken word. The subjective choice of which ‘Statistic’ to report can change the message. In his enjoyable little books “How to lie with Statistics”, Darrel Huff noted that “the crooks already know these tricks; honest men must learn them in self-defence”. When we study Statistic we will become more critical in our analysis of information and will be less susceptible to misleading or deceptive claims. It will help us evaluate numerical facts, when to believe them, when to place tongue in cheek, and when to reject them. It aids us to make correct decisions. There are great quantities of data being collected by a diversity of individuals and agencies, the summaries for which influence major decisions (Health, Education, Energy, Environment, Transportation, etc.) of the government and other institutions that affects us. We also use these data to make decisions, form conclusions, and build our individual ware-house of knowledge. If we want to build a sound knowledge base, make intelligent decisions, assess the reliability of conclusions, and form worthwhile opinions, we must be careful to filter out the incoming information that is erroneous or deceptive. We should also understand the data and the process of collection to be able to appreciate them. When we are data literate we would consume numerical evidence not only with a certain amount of scepticism, but also with a considerable degree of skill; we would also

interpret undistorted data intelligently. This will make us a watchdog against bad statistics.

Statistical analysis sharpens concepts. It is used by scholars in almost all disciplines (economics, sociology, medicine, agriculture, education, etc.) Statistical methodology is a language for scientific social research. According to Fisher “statistical methods are essential to social studies, and it is principally by the aid of such methods that these studies may be raised to the rank of sciences”. It is now common to hear of fields with suffix ‘metrics’ when there is a high degree of “statistisation” of the subject. For example advanced statistical methodology in economics is known as econometrics, in biology as biometrics (biometry), in environment as environmetrics, sociology as sociometrics, in psychology as psychometrics, technology as technometrics. In fact these days scientific papers in social studies are considered serious by its degree of statistics - the extent to which the idea and techniques of statistics are used in the paper. We can not therefore afford not to study statistics because we do not want to be figure illiterate and stay behind time.