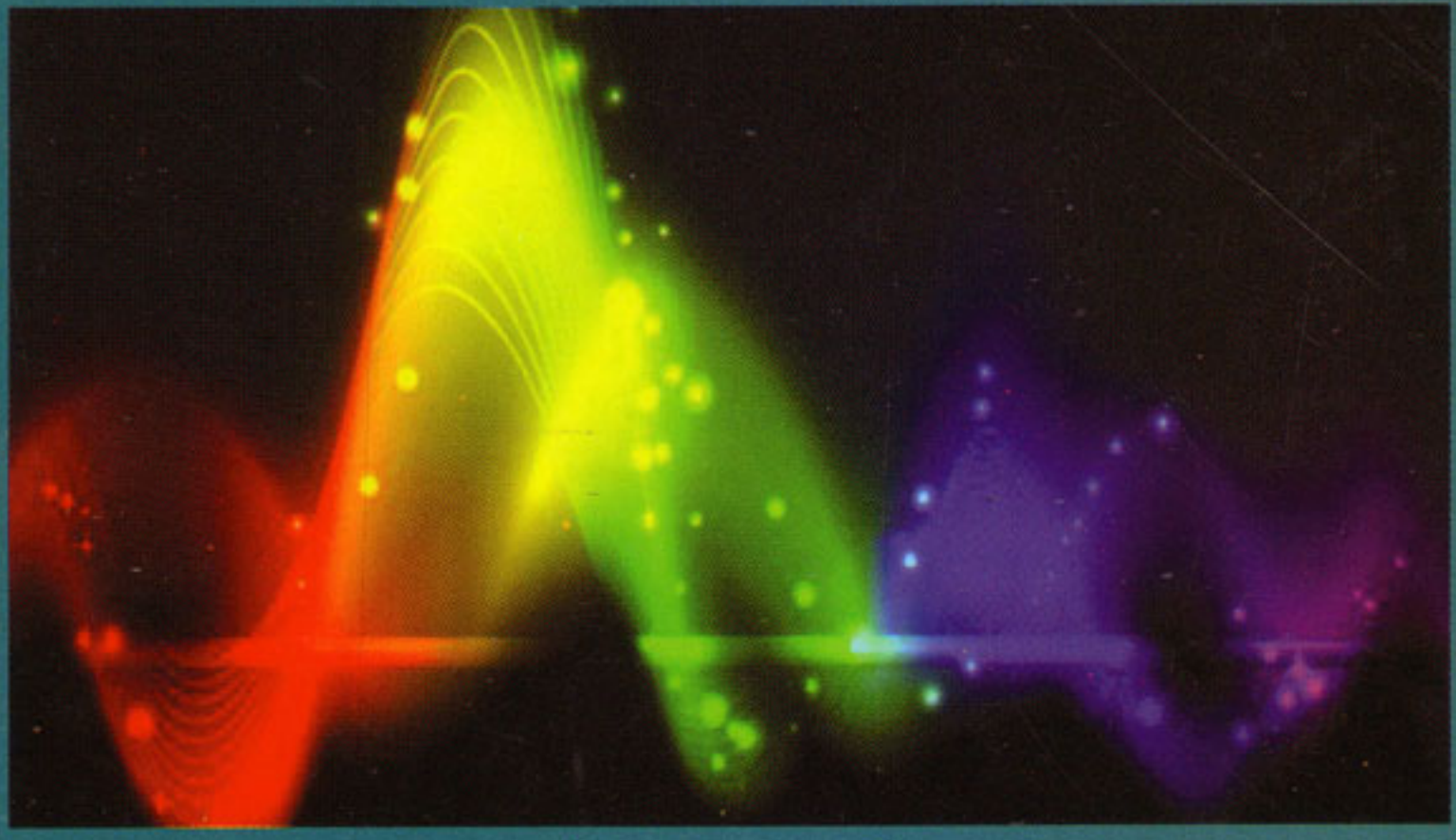
FOURTH REVISED EDITION

An Introduction to STATISTICS AND PROBABILITY

M. Nurul Islam



Mullick & Brothers

ABOUT THE AUTHOR

Dr. M. Nurul Islam is former Selection Grade Professor of Statistics, Faculty of Science at the University of Dhaka, Bangladesh. Currently Dr. Islam is Pro-Vice Chancellor at the World University of Bangladesh. He joined the University of Dhaka as a Lecturer in 1971 and acted as Chairman of the Department of Statistics during 1985-1987. He held the coveted position of the Vice Chancellor of the Mawlana Bhashani Science and Technology University in Tangail during 2009-2013. He is an elected member of famed International Statistical Institute (ISI) of the Netherlands. Dr. Islam is now holding the position of President of the Bangladesh Statistical Association. He also held the same poison of the Association during 2010–2012. Dr. Islam was an elected member of the Dhaka University Senate for the period 2002-2004. He was also a part-time member of University Grants Commission during 2010-2012. During his 45 years' career as a teacher and researcher, he has written nearly one hundred scientific papers and survey reports related to his discipline, which have been published in both local and international journals. He has supervised a number of M.Sc and PhD works. He provided consultancy services to USAID, DANIDA, UNICEF, CIDA, CARE, SIDA, NETZ, Bangladesh Bureau of Statistics (BBS), Ministry of Health and Family Welfare and Ministry of Textile. He independently conducted a number of field surveys as a Principal Investigator. He extensively toured a number of countries including USA, UK, Canada, Egypt, Indonesia, Singapore, Nepal and India and presented papers there in conferences and seminars. He is author of four textbooks entitled (i) **An Introduction to Statistics and Probability, (ii) An Introduction to Research Methods, (iii) An Introduction to Sampling Methods and (IV) An Introduction to Demographic Techniques** published by Mullick & Brothers. The author has also published his entire research works in a single volume entitled **Demographic Research in Bangladesh: An Update,** with the financial assistance from Social Science Research Council, Ministry of Planning. He has been an Executive Editor of several national and international peer reviewed journals. Dr. Islam is also the recipient of the UGC Award and the Dhaka University Faculty Award for 2007 and 2004 respectively for two of his incomparable textbooks.

CONTENTS

Consequently, he required that all persons report to the nearest tax collector. In the ancient Egypt also, such attempt was made to count people and land areas to redistribute the lands among the inhabitants of Egypt. This probably happened at the time of **Ramses** II in about 1400 BC. Further development of statistical records was noted in ancient Chinese, Greek and Roman territories. This feature has tempted many people to label statistics as the **science of king**.

The science of statistics developed gradually and its field of applications widened day by day. Its use and importance continued to be considered indispensable in all spheres of life as time elapsed. At government level, the statistics collected are of the same type as was done at the ancient time and serve almost the same purpose as before viz. the administrative, social, political and military policies of the government. The modern science of statistics really began to develop toward the end of the nineteenth century. It has to day grown into a vast and versatile subject in terms of its uses and applications to all sciences.

In about mid-seventeenth century, the theoretical development in modern statistics came with the introduction of the Theory of Probability. Gambling, in the form of games of chance, led to this theory of probability being originated by Pascal and Fermat because of their interest in the gambling experiences of the Chevalier de Mere.

The role of normal curve of error in the development of statistics can hardly be over emphasized. De Moivre first studied the equation of this curve in 1773. De Moivre's works remained unknown until 1924 when Karl Pearson (1857–1936) found it in a library. However, the same result was later developed independently by two famous mathematical astronomers Laplace (1749–1827) and Gauss (1777–1855).

The works of Charles Darwin (1809–1882), a famous biologist, were largely biomedical or statistical in nature. Mendel, too, extensively used statistical concepts in his study of plant hybrids.

The study of statistics gained impetus when Karl Pearson spent nearly half a century in statistical research. In addition, Karl Pearson founded the **Biometrika**, a famous journal in Statistics. He also established a school of statistics.

W. S. Gosset (1876–1937), a student of Karl Pearson, succeeded in working with samples, when large–sample theory was proving somewhat

inadequate. The works of Gosset appeared in Biometrika in 1908 under the name of 'Student', Gosset's pseudonym. (A famous story has it that Gosset was afraid to publish under his own name for fear that his employers, the Guinness Brewery, would be unhappy to discover that one of its chemists was doing research in statistics.) Today, student's t is a basic tool of statisticians and experimenters.

R. A. Fisher (1890–1962), who is known as the father of statistics, made numerous and significant contributions to statistics. He was influenced by Karl Pearson and Gosset, the two world known statisticians. His pioneering works made valuable contributions to evolve statistical procedures in many fields particularly agriculture, biology and genetics. He is known for his contribution in the field of Analysis of Variance and Experimental Design. Meanwhile, Francis Galton (1822–1921) gave the concept of **regression line** while working on heredity of men and laws governing the transmission of physical and mental characteristics from one generation to another'. His work was confirmed by his friend Karl Pearson.

The development of the theory of statistics was also effected by the pioneering works of a great number of authors, among whom Pareto, Adam, Edgeworth, Bowley, Yule, Stuart and Kendall deserve special mention.

J. Neyman (1894–1981) and E. S. Pearson (1895–1980) will remain ever known for their original contributions to the theory of testing statistical hypothesis in 1936 and 1938. This theory promoted considerable research works of practical use. Indian statistician Mahalanobis (1893–1972) contributed significantly to the field of sample survey. He is well known for his D^2 statistic. He established Indian Statistical Institute (ISI) in 1931 and founded **Sankhya**, an international journal in Statistics. Qazi Motahar Hussain (1897–1981), an authority in statistics, made significant contributions towards the development of statistics in Bangladesh. He is known worldwide for his works **'Hussain Chain'**. He was the founder of the Department of Statistics and the Institute of Statistical Research and Training at the University of Dhaka.

Deming's (1900–1993) philosophical thoughts towards improving the quality of manufactured products through statistical quality control techniques is noteworthy. Genichi Tagguci (1924–) promoted the use of experimental designs for product improvement.

CHAPTER 1

STATISTICS AND ITS ORIGIN

1.1 STATISTICS: ITS ORIGIN AND DEVELOPMENT

It is not precisely known how the word **statistics** was originated. However, most people believe that the term statistics, derived from word **state,** was used to refer to a collection of facts of interest to the state. Some believe that the word **statistic** has been originated from the Italian word **statista**, the French word **statistique** and the German word **statistik**. This background tends to suggest that the term statistics has its origin from the ancient time. At that time, the word statistics had been believed to be in use as an indicator or yardstick of a country's economic, political and social conditions. In this sense, the word statistics serves as an index of a country's overall condition. In modern days too, relevant statistical data are of immense importance to understand the level of development of a country. Presumably, all cultures that recorded history internationally also recorded statistics.

The term **statistics** is an old one. As people abandoned their nomadic way of life, and started to live in groups, they urgently felt the need to know each other's wealth, manpower, strength and position for their survival and safety. The group chiefs or rulers used to collect information on the above aspects in order to impose and levy tax to strengthen their economic condition. Caesar Augustus decreed that the entire world should be taxed.