Encyclopaedia of Application of Statistics

Rakesh Kumar Dwivedi

ALEXIS PRESS
JERSEY CITY, USA

ENCYCLOPAEDIA OF APPLICATION OF STATISTICS

ENCYCLOPAEDIA OF APPLICATION OF STATISTICS

Rakesh Kumar Dwivedi

ALEXIS PRESS

Published by: Alexis Press, LLC, Jersey City, USA
www.alexispress.us

First Published 2022

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication Data

Includes bibliographical references and index.

Encyclopaedia of Application of Statistics by *Rakesh Kumar Dwivedi*

ISBN 978-1-64532-582-6

CONTENTS

CHAPTER 1
ACCOUNTING INFORMATIZATION CLOUD STATISTICS BASED ON STATISTICS MINING

Rakesh Kumar Dwivedi, Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id-r_dwivedi2000@yahoo.com

ABSTRACT:

The number of all types of data information is growing quickly along with the quick growth of information technology. Big data technology offers management decision makers a thorough strategic foundation that the conventional processing mode cannot match. It is a significant tool for gathering, storing, and managing massive amounts of data, as well as for analyzing and predicting the habits and characteristics of specific groups of people and even the development trend of an industry. Modern management accounting produces a wide range of different data since it supports the whole business internal control process. More and more people are starting to study data mining as a result of the rapid expansion of network data and the size of databases, and the classification algorithm the primary technology used in data mining has drawn a lot of attention as well. Many businesses started paying more attention to data mining in order to advance the information technology level of enterprise management accounting and deepen information application, and deep data mining improved the scope and depth of enterprise data analysis. Data mining technology makes up approximately 50% of the informatization in the study on data and accounting informatization, which is the direction for future growth. The dependency of businesses on information technology in the process of accounting management has increased with the advent of the information age. Businesses need to focus more on management accounting information technology and enhance their staff members' information application skills if they wish to prosper in the information era.

KEYWORDS:

Accounting Informatization, Cloud Computing,Cloud Statistics, Technology.

INTRODUCTION

Numerous classify algorithms exist. The decision tree, Bayesian, genetic, artificial neural network, and classification algorithms based on association rules will be the main topics of this research. We must first establish the adaptability of statistics mining to management accounting and the essential relationship between statistics mining techniques and management accounting concepts in order to build a functioning method system for management accounting based on statistics mining. An in-depth examination of the relationships between statistics, the discovery of these relationships, analysis of the business circumstances of enterprises, the identification of operational issues, and promotion of an improvement in the level of financial analysis and decision-making of enterprises are all parts of statistics mining. The statistics classification algorithm has grown in importance in statistical analysis technologies. More and more statistics categorize algorithms have been suggested as statistical analysis study has progressed. Building a statistics classifier, which is used to precisely identify certain unknown kinds of statistics, is an essential stage in the statistics classification process. A considerable number of hardware and software items had to be purchased in the past in order to build an accounting information system, and ongoing

maintenance and updating were also quite expensive. Numerous factors demanded a lot of labor and material resources, which raised operational expenses for businesses. The fundamental component of massive statistics mining is the statistics classification algorithm. Its primary purpose is to extract important information and message from massively disordered statistics using a variety of procedures, examine the characteristics of all types of message, and give statistical support for researchers to further anticipate a certain trend. The quantity of statistics that businesses must account for is growing along with the big data age, which practically increases the workload and difficulties of accountants [1], [2].

Cloud computing has enough network storage space since it is an outcome of the Internet's advanced growth. A new computer paradigm based on shared resources, cloud computing has grown significantly in recent years. Technology for cloud storage is the main component of this. A mass storage device is brought together by app software to operate together through RAC, networking, or multitiered file storage system, offering practical and affordable mass storage services. The amount of statistics used in accounting work is growing, and the difficulty of statistics calculation and analysis is getting higher and higher in the age of big statistics, which inevitably has a significant impact on the accounting work of businesses. The natural result of the growth of statisticsbase technology is statistics mining, which is the process of extracting valuable information and value from a vast amount of statistics. Numerous industries, including retail, banking, insurance, health care, and communication, have made extensive use of statistics mining. One of the most crucial technologies in statistics mining is classify, and several algorithms have so far been suggested. In order to categorize samples of unclassified categories, the technology Classify builds a classifier based on the properties of statistical sets.

Accounting's objective is to offer communication support for an organization's internal management and control. The management accounting team's primary objective in gathering and compiling all messages is to assess the firm's future production and operation condition based on historical production and operation outcomes, in order to support the strategic decision of the enterprise. The management accounting message technology is a requirement for the statistics mining app. Through the use of messaging technology, statistics' breadth and depth, business organizations' capacity for financial analysis, and their degree of financial management may all be enhanced. The early investment in hardware and software has been dramatically reduced since the statistical period and the development of cloud service technology, and just a few computers are often required, which significantly reduces the cost of organizations. Statistics classify algorithms are continuously developing and being gradually optimized as a result of the widespread use of statistics mining technology. Among these are the classical classify algorithms, such as the decision tree classify algorithm, naive Bayes algorithm, support vector machine classify algorithm, artificial neural network classify algorithm, etc[3], [4].

DISCUSSION

People are becoming more concerned about security and privacy problems as popular technologies like cloud computing and storage have significantly fulfilled the rising need for storage space. Due to improper management, insufficient security, bit rot, disk controller error, and tape failure, as well as the fact that cloud storage service providers are not entirely reliable, it is possible for the integrity of user statistics to be lost rapid and efficient scaling. When used in corporate internal accounting tasks, it may increase accounting efficiency. It is necessary to use statistics mining technology to conduct statistics analysis during the enterprise management accounting message processing process in order to provide reliable statistics support for management accounting and provide reliable assurance for the

development of enterprises and the improvement of message processing capability. Accounting is a kind of accounting job that focuses mostly on the internal administration and control of businesses. The message that management accounting pays attention to and collects is frequently not only limited to the single financial message, but also the reflected content is not only the post-reaction and supervision of the enterprise's operating results.

This is because there is a fundamental difference between financial accounting and financial accounting in the service object. A computer technology called statisticsbase-based knowledge discovery has been put out with the quick advancement of artificial intelligence and statisticsbase. Many fields, including machine learning, pattern recognition, statistics, knowledge acquisition, intelligent statisticsbase, expert systems, and high-performance computing are closely related to this technology because it searches the hidden useful message from a large amount of statistics using some algorithm. In order to significantly lower the cost of calculating hashes for large-scale statistical structures, Chen proposed the idea of the Merkle Hash Tree. The BLS signature has less signatures than the RSA and DSA signature schemes under the same security conditions and a modulus of 1024 bits, according to a short message signature system presented by Hu, Chen, and Ling. The Zheng approach fills the gap left by the lack of sample data that the Bayesian classify method requires. Shi et al. proposed a number of enhanced Bayesian classification algorithms, including the semi-naive Bayesian algorithm, candidate compressed Bayesian network building algorithm, TAN algorithm, and other efficient techniques that may decrease autonomy. A CBA classification technique was presented by Yunyang et al. Two processes make up the majority of the CBA algorithm. Yang suggested a sentinel-based statistics recoverability proof technique that, if the statistics are destroyed, may partially restore the original statistics in addition to verifying the integrity of the statistics on the distant node [5]–[7].

Statistics Mining Technology Research

Finding relevant bits of statistical messages and interpreting them in light of their development trend and composition style are the main goals of statistics mining. The knowledge learning stage is often referred to as statistics mining. Statistics mining is an important phase in knowledge discovery based on statisticsbase. Sorting the training set based on the attribute value is time-consuming when dealing with continuous attributes since each internal node must find its own appropriate splitting criteria. An inductive learning technique called decision tree categorize predicts a collection of random and disorganized sample data in the form of a tree structure. The decision tree classification algorithm, which is made up of root nodes, internal nodes, leaf nodes, and directed edge nodes, may naturally depict the issues and major issues faced by decision-making classes in each decision-making stage. Currently, the traditional classifiers used in big data analysis and statistics mining mostly consist of decision trees, naive Bayes, support vector machines, neural network classifiers, and others. The non-C4.5 approach is prevalent among decision tree classifiers, however as computer and messaging technology has advanced, the 4C5 technique is no longer able to keep up with the complexity of statistical classifiers. Various categorize techniques thus have their own features and various challenges that need to be solved in order to fit to the processing of large-scale statistical collections.

Pre-sorting technology is used by the SL IQ algorithm to do away with the requirement to sort statistics sets at each node of the decision tree. One of the inductive learning algorithms is the decision tree classify algorithm, which primarily refers to the classification rules that infer "tree" structure from a collection of irregular and unordered sample statistical messages. Decision tree classify algorithms have numerous clear benefits over more conventional statistics classify algorithms like statistical techniques and neural network methods. For

instance, the decision tree classify algorithm's statistical classification principles are straightforward, simple to comprehend, and tough to erroneously apply in practice. The decision tree may learn several concepts by evaluating and summarizing case sets. It is simple to use and has many different applications. When categorizing large-scale case data represented by unstructured attribute-value pairs, the decision tree technique is a superior option. The most popular decision tree learning algorithms at the moment are ID3, C4.5, SLIQ, and SPRINT. Figures 1 illustrate how the researches associated model diagrams are set up to study and explain them.

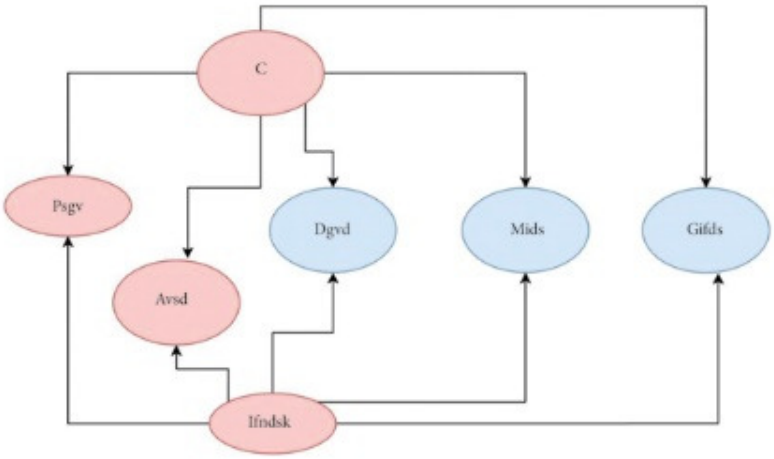

Figure 1: Shows a model representation of the sorting concept in statistics mining

The data mining principle model diagram in Figure 1 further exemplifies the data mining application principle and directs the development of the data mining algorithm. Big statistics mining technology primarily involves the process of gathering and segmenting statistics messages from enormous message statistics in accordance with a certain characteristic, and then acquiring and amassing some useful messages over time. In the age of large statistics, network message technology has evolved into statistics mining technology, which primarily uses artificial intelligence, statistics base, statistics, and other techniques. Depending on whether fault-tolerant preprocessing is used on the statistics, the methods for ensuring statistical integrity in cloud storage may be separated into two categories: statistics holding proof mechanisms and statistics recoverability proof mechanisms. According to specialized study on the classification algorithm, there is a high association between the algorithm's efficacy and the traits of statistics, which include vacancy value, loud noise, and dispersion. Some data are dispersed and jumbled, whereas others have continuous attribute features. The creation of a classifier typically involves two steps: training and testing. An accurate description or model of the related statistics set is created for each category once the properties of the training statistics set are examined [8], [9].

During the testing step, the test is categorized using the category's description or model, and the classification's correctness is evaluated. In general, the testing phase costs substantially less than the training phase does. Big statistics is only the mining and integration of big statistics message as an abstract term. These data kinds are varied, there is a large volume of statistics, a low value density, and a rapid growth rate. The app value hidden beneath them can only be found via suitable statistics mining and statistical analysis. Numerous sectors are developing their output, which means that daily statistics production will increase. This

message has a subtle impact on people's daily lives and even the growth of a certain business thanks to big data technologies.

Statistical Mining Algorithm Research

The decision tree categorize algorithm may visually display the issues and crucial elements of various decision-making classes at various stages of the whole decision-making process. Root nodes, internal nodes, leaf nodes, and directed edges linking nodes make up the decision tree. The root node is distinct; it represents a collection of samples that have been categorized, while the internal group represents an attribute of the item and the terminal node the classification's outcome. The process continues this process until the last node and category of the route are put in the leaf node, starting with the root node and selecting the relevant attribute value from top to bottom. The decision tree classify algorithm may now provide a collection of rules that are simple to comprehend since each branch of the decision tree corresponds to a classify rule. The decision tree statistics classification technique offers a lot of benefits, but it also has a lot of drawbacks.

Because the decision tree determines the statistics analysis process, it is inevitable that the statistics will be scanned and sorted several times in succession during the statistics classification process, particularly during the tree construction process. This will inevitably cause the entire statistics analysis process to become slow. The Bayes classify algorithm, which is based on the Bayes formula, is a classification method that makes use of probability statistical information. The classify method uses the Bayes theorem to compute the likelihood that a given sample belonging to an unknown class belongs to each class when the prior probability and class conditional probability are known. It then chooses the class with the greatest probability as the determined class of the sample. SLIQ algorithm technology has high scalability for expanding the number of records and characteristics to some level because Li algorithm is upgraded in many ways based on 4C5 algorithm technology and also incorporates sorting and breadth-first strategy technologies [10], [11].

Since of the C45 algorithm's features, the decision tree's construction is finished using the depth-first approach, which has an incredibly poor efficiency since each important node must be examined throughout the statistics' classification and analysis. However, once the breadth-first strategy technology is used, it is possible to get the best splitting criterion for each leaf node in the present decision tree by simply scanning each attribute list once for each layer. Because of the many properties of the statistics themselves, the naive Bayesian method is rather stable and won't significantly affect the outcomes of the classification. The accuracy of the classification findings increases with the strength of the independence between the naive Bayesian statistics. But it's important to keep in mind that the conditional independence hypothesis, which is the ideal situation, must serve as the foundation for the classification algorithm. The accuracy of the classification will be decreased by linkages between statistical characteristics in a real program, making it challenging to achieve the theoretical maximum impact of this strategy. It aims to boost the value of statistics and the processing capacity of statistics. From a technological perspective, big statistics and cloud computing are similar to the front and back of a coin in their interdependence. Massive statistics may be mined in a multitiered manner because enormous statistics cannot be handled by a single computer and must be implemented in a multitiered architecture. It must, however, depend on the multitiered processor and cloud storage virtualization technologies of cloud computing. Big statistics has increasingly become more popular as the cloud age has progressed.

The decision tree algorithm's statistics analysis structure has undergone corresponding changes thanks to the SPRIN algorithm. In particular, the list of statistics categories that the

SLI Q algorithm needed to store in memory has been deleted, and its place has been taken by a list of attributes for each statistics number. The benefit of this technique is that it may avoid repeating statistics analysis when it traverses each attribute list to determine the best splitting standard for the current node while studying a large number of statistics. The statistical computation of the category distribution message of the statistics set corresponding to each non-terminal node and the splitting of the statistics set using splitting criteria are the two procedures that take the longest to complete while creating a decision tree. The genetic algorithm is an effective search and random optimization method that developed from the notion of biological evolution. Both operations are accomplished by UDF in M IND. It represents a significant advancement in the use of computer algorithms to natural science. The method converts problem-solving into the process of identifying chromosomes with high fitness in accordance with the genes on chromosomes by using the concept of natural evolution. This algorithm overcomes the drawback that most optimization techniques are simple to fall into local optimum because it combines the benefits of directional search and random search, giving it superior global search ability. The method overcomes the flaw that the conventional optimum solution is difficult to accomplish local optimization by combining random search with directed search, improving its global optimization performance. Genetic algorithms may fix issues without being aware of them, much like nature. Its primary responsibility is to assess each chromosome produced by the algorithm and choose the appropriate chromosomes based on fitness, making the chromosomes simpler to replicate. Although researchers in the subject of statisticsbase have contributed to the growth of statistics mining research, most of the algorithms put out to far have not used the technology associated with statisticsbase, and it is challenging for statistics mining applications to interact with statisticsbase platforms. One of the most important problems in this area is now this one [12], [13].

Research Statistics Mining Accounting Management Message Processing

Financial sharing has favorable technological circumstances thanks to the advancement of communication technology. The growth of the Internet makes it possible to share business financial data, which benefits the timely delivery of financial information in corporate accounting messaging systems. Multitiered computing includes cloud computing as a key component. In particular, it breaks down enormous statistics processing and computing programs into many smaller ones, processes and analyzes them via a system made up of several servers, and then relays the findings to consumers. Early cloud computing likewise used a straightforward multitiered work allocation and integrated computing outcomes. Advanced software and hardware facilities cannot be isolated from message management. To make existing library resources into digital message content that can be easily searched for and downloaded on the platform, we should first strengthen the construction and improvement of basic management facilities.

Next, we should upload the converted library resources to the designated library resources platform. A key factor in enhancing the timeliness of the financial statistics message in the enterprise accounting message system is the message platform's comprehensive management of the financial statistics of each region of the enterprise. As a result, the financial statistics can appear on the financial sharing platform in due time after they occur, allowing the financial staff to handle the accounts in due time after obtaining the pertinent statistics. In order to implement the management accounting message system and realize the overall "messagization" of enterprises, we should further broaden the scope of message processing, assist businesses in doing so, and incorporate the operation statistics of other departments into

the accounting message system. This will enable us to manage and supervise the entire enterprise and will enhance the management impact of the accounting message system.

About 90% of the total accounting informatization is accounted for by the management choice accounting informatization. Budgeting is a crucial component of an organization's prior management, and it helps to increase the efficiency and effectiveness of management accounting message processing. Therefore, in the management accounting message processing based on statistics mining, enterprises should integrate the budget work into the accounting message system of enterprises, use accounting software to analyze the business situation of enterprises, formulate a scientific and reasonable budget in conjunction with the future development strategy of enterprises, improve the budget management level of enterprises, and reduce the gap between budget and actual spending. You should update your ideas on time, be aware of the most recent trends, and keep up with the growth of your own business and adjacent industries as a manager. The greater use of cloud computing technology to internal financial and accounting administration of businesses in the age of big statistics requires the improvement of relevant rules and regulations to guarantee message security. Faced with the overall trend of message management, we should be willing to take a chance, be willing to depart from the conventional model of the past, be aware of people and technological advancements, and recognize the evolution of accounting messages.

CONCLUSION

The development of cloud computing technology in the age of big statistics has increased the opportunities for business accounting informatization, but there are also still a number of concerns. In the age of big numbers, how to take advantage of the chance to deal with risks and problems is still a significant issue that every organization is facing. By methodically examining the classify algorithm used in statistics mining, we can learn about the benefits and drawbacks of the decision tree classify algorithm, naive Bayes classify algorithm, support vector machine classify algorithm, and neural network classify algorithm as well as the scenarios in which they are useful. We can then make targeted improvements and optimizations to address their drawbacks. Big data and cloud computing technologies make it much easier to build accounting information technology, which not only lowers costs and increases efficiency but also achieves resource sharing and has excellent growth potential. The significance of the statistics categorize algorithm will become more clear with the emergence of the age of large statistics, and the statistical analysis trend will shift toward the features of algorithms like execution speed, scalability, and understandability of output findings. Enterprises need to pay more attention to the level of financial personnel, increase the level of financial informatization, and provide a human resource foundation for better development of financial informatization when applying statistics mining to management accounting message processing.

REFERENCES:

[1] A. Zimek and P. Filzmoser, “There and back again: Outlier detection between statistical reasoning and data mining algorithms,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2018. doi: 10.1002/widm.1280.

[2] K. Moriya, S. Matsushima, and K. Yamanishi, “Traffic Risk Mining from Heterogeneous Road Statistics,” *IEEE Trans. Intell. Transp. Syst.*, 2018, doi: 10.1109/TITS.2018.2856533.

[3] E. G. M. Hui, *Learn R for applied statistics: With data visualizations, regressions, and statistics*. 2018. doi: 10.1007/978-1-4842-4200-1.

[4] R. Li *et al.*, "Meningococcal conjugate vaccine safety surveillance in the Vaccine Safety Datalink using a tree-temporal scan data mining method," *Pharmacoepidemiol. Drug Saf.*, 2018, doi: 10.1002/pds.4397.

[5] S. Koseoglu and A. Bozkurt, "An exploratory literature review on open educational practices," *Distance Educ.*, 2018, doi: 10.1080/01587919.2018.1520042.

[6] I. Majeed and S. Naaz, "Current State of Art of Academic Data Mining and Future Vision," *Indian J. Comput. Sci. Eng.*, 2018, doi: 10.21817/indjcse/2018/v9i2/180902026.

[7] D. Sarkar, R. Bali, and T. Sharma, *Practical Machine Learning with Python*. 2018. doi: 10.1007/978-1-4842-3207-1.

[8] Y. Dagasan, P. Renard, J. Straubhaar, O. Erten, and E. Topal, "Automatic parameter tuning of multiple-point statistical simulations for lateritic bauxite deposits," *Minerals*, 2018, doi: 10.3390/min8050220.

[9] N. Koriat and R. Gelbard, "Knowledge sharing motivation among external and internal IT workers," *J. Inf. Knowl. Manag.*, 2018, doi: 10.1142/S0219649218500260.

[10] K. R. Jumiyanti, "Analisis Location Quotient dalam Penentuan Sektor Basis dan Non Basis di Kabupaten Gorontalo," *Gorontalo Dev. Rev.*, 2018, doi: 10.32662/golder.v1i1.112.

[11] D. Sarkar, R. Bali, and T. Sharma, *Brownlee, Jason, Master Machine Learning Algorithms: Discover How They Work and Implement them from Scrach, 2017*. 2018.

[12] A. Ristea, J. Kurland, B. Resch, M. Leitner, and C. Langford, "Estimating the spatial distribution of crime events around a football stadium from georeferenced tweets," in *ISPRS International Journal of Geo-Information*, 2018. doi: 10.3390/ijgi7020043.

[13] M. Marwan, A. Kartit, and H. Ouahmane, "A Secured Data Processing Technique for Effective Utilization of Cloud Computing," *J. Data Min. Digit. Humanit.*, 2018, doi: 10.46298/jdmdh.4085.

CHAPTER 2
USING BAYESIAN STATISTICS FOR CALIBRATION AND MEASUREMENT CONTROL

Ashendra Kumar Saxena, Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id-ashendrasaxena@gmail.com

ABSTRACT:

This study describes a Bayesian methodology that has the intrinsic ability to select candidates from calibration-type curves that are credible in light of observed data, expert knowledge, and theoretical models including the nature of the measurement errors. In this study, the fundamental procedures of Bayesian calibration are reviewed, and potential applications of the results are discussed. The proposed technique is illustrated using a calibration based on head-space gas chromatographic data. A log-normal distributed measurement error has been applied to the linear calibration situation. Such a treatment of noise emphasizes how crucial it is to model the random elements of any issue. Bayesian statistics is a statistical theory based on the Bayesian interpretation of probability, in which probability indicates a degree of belief in an occurrence. The level of belief may depend on past information about the event, such as the findings of earlier trials, or it may depend on the individual's beliefs about the event. This is different from a few other interpretations of probability, like the frequentist interpretation, which sees probability as the upper bound on the relative frequency of an event after numerous trials.

KEYWORDS:

Bayesian, Candidates, Frequency, Statistics.

INTRODUCTION

Bayesian statistical approaches compute and update probabilities after getting new data using Bayes' theorem. According to data, prior knowledge, and assumptions about the event or circumstances surrounding it, Bayes' theorem calculates the conditional probability of an event. The Bayes theorem, for instance, can be employed in Bayesian inference to determine the parameters of a probability distribution or statistical model. The Bayes theorem can immediately ascribe a probability distribution that quantifies the belief to the parameter or collection of parameters since Bayesian statistics treats probability as a degree of belief. Thomas Bayes, who developed a particular application of Bayes' theorem in a paper published in 1763, is the name given to the field of Bayesian statistics. The Bayesian interpretation of probability was established by Pierre-Simon Laplace over the course of numerous publications published in the late 18th and early 19th century. Laplace solved several statistical issues using techniques that are today regarded as Bayesian. Later writers created several Bayesian approaches, but it wasn't until the 1950s that the phrase "Bayesian method" became widely used to refer to these techniques.

Many statisticians have negative opinions about Bayesian approaches for much of the 20th century because of both philosophical and practical reasons. The majority of methods that were widely utilized during the century were based on the frequentist interpretation, and many Bayesian approaches took a lot of computation to complete. However, Bayesian approaches have become more popular in statistics in the twenty-first century as a result of the development of strong computers and fresh algorithms like Markov chain Monte Carlo.

Bayesian inference is a statistical inference method where probability is used to quantify inference uncertainty. Traditional frequentist inference assumes that model parameters and hypotheses are constant. In frequentist inference, probabilities are not ascribed to parameters or hypotheses. For instance, it would not be logical to immediately give a probability to an event that can only occur once, such as the outcome of the following fair coin toss. However, it would be reasonable to claim that as the quantity of coin flips rises, the fraction of heads approaches a half.

A set of statistical assumptions and procedures that characterize the methods used to produce the sample data are specified in statistical models. There are several parameters that can be changed in statistical models. For instance, samples from a Bernoulli distribution, which represents two potential outcomes, can be used to represent a coin. The chance of one result, which is typically the probability of landing on heads, is the only parameter of the Bernoulli distribution. Devising a good model for the data is central in Bayesian inference. Most of the time, models can only approximate the underlying process and may not account for all the variables affecting the data. Model parameters can be given probability in Bayesian inference. Random variables can be used to express parameters. The Bayes theorem is used in Bayesian inference to update probability as new information becomes available or is understood. 'Influence of previous beliefs' is a term used in the Bayesian experimental design. With this method, the results of earlier tests are incorporated into the design of the subsequent experiment using sequential analysis techniques. This is accomplished by revising 'beliefs' using prior and posterior distribution. This makes it possible to efficiently utilize resources of various kinds while designing tests. The issue with multi-armed bandits is one illustration of this. Exploratory analysis of Bayesian models is an extension of exploratory data analysis that takes into account the specific requirements and characteristics of Bayesian modeling.

Exploratory data analysis aims to make simple descriptions or structures in the data visible. We examine data or graphs and look for trends. We follow up on leads generated by historical data, imagination, noticed trends, and previous experience with various data analysis. Together with other distributions like the posterior predictive distribution and the prior predictive distribution, the inference process creates the posterior distribution, which plays a key role in Bayesian statistics. To effectively respond to the issues that drive the inference process, these distributions must be shown, analyzed, and interpreted in the right way. In addition to inference itself, a number of associated activities must be attended to while dealing with Bayesian models: inference quality assessments are necessary when employing numerical methods like Markov chain Monte Carlo approaches Evaluations of model assumptions and predictions are included in model criticism.

DISCUSSION

The accomplishment of each of these activities is essential to the interactive and iterative modeling process. They are all a part of the exploratory analysis of Bayesian model's approach. Both numerical and visual summaries are needed for these activities. There are two major ways that the statistical term calibration is used to describe particular kinds of statistical inference issues. "Calibration" can refer to a variety of things. Regression performed in reverse, where a known observation of the dependent variables is used to predict a corresponding explanatory variable rather than a future dependent variable being predicted from known explanatory factors. Techniques used in statistical classification to estimate the likelihood that a given new observation will belong to each of the pre-existing classes. Additionally, the term "calibration" is used in statistics to refer to calibration in its conventional sense. Model calibration, for instance, can also apply to Bayesian inference on

the parameters of a model given a set of data, or more broadly, to any sort of fitting of a statistical model.

According to Philip Dawid, "a forecaster is well calibrated if, for example, the long-run proportion of those events to which he assigns a probability 30 percent turns out to be 30 percent “Transforming convert classifier scores into class membership probability is known as calibration in classification. Gebel provides a summary of calibrating techniques for two-class and multi-class classification applications. There are many metrics available to gauge how well-calibrated a classifier's probabilities are produced. A Brier score is sometimes used in prediction and forecasting to evaluate the accuracy of a collection of forecasts, specifically the degree to which the assigned probability' magnitudes correspond to the relative frequency of the actual events. In his 2015 book Super forecasting, Philip E. Tet lock uses the term "calibration" in this sense. This is distinct from precision and accuracy. According to Daniel Kahneman, for instance, "your discrimination is perfect but your calibration is miserable if you give all events that happen a probability of.6 and all the events that don't happen a probability.

Definition of BIPM

The International Bureau of Weights and Measures (BIPM) has the following definition of calibration in its official regulations: "Operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties of the calibrated instrument or secondary standard), and in a second step, uses this information to establish According to this definition, the calibration procedure is just a comparison, but it also introduces the idea of measurement uncertainty when comparing the accuracies of the test instrument and the standard [1]–[3].

current calibration techniques

National laboratories were established as a result of the growing demand for standards with uniform international comparability, recognized accuracy, and uncertainty. A National Metrology Institute (NMI) will be present in many nations, and it will be responsible for maintaining the primary standards of measurement (the primary SI units as well as a number of derived units) that will be used to offer traceability to client instruments through calibration. By constructing an uninterrupted chain from the highest level of standards to a measurement equipment, the NMI supports the metrological infrastructure in that nation (and frequently in others). National Metrology Institutes include, among others, NIST in the United States, PTB in Germany, and NPL in the United Kingdom. Since the signing of the Mutual Recognition Agreement, obtaining traceability from any participating NMI has become simple. Companies are no longer required to obtain traceability for measurements from the NMI of the nation in which they are located, such as the National Physical Laboratory in the UK [4]–[6].

Quality

It is preferable for the calibration and future measurements to be "traceable" to the globally specified measurement units in order to enhance the calibration's quality and have the findings acknowledged by other organizations. A formal comparison to a standard that is directly or indirectly tied to national standards (like NIST in the USA), international standards, or certified reference materials is how traceability is established. National standards laboratories run by the government or commercial companies that provide metrology services may carry this out. An efficient metrology system that includes the

formal, regular, and recorded calibration of all measuring instruments is required for use with quality management systems. The ISO and ISO standards specify how to quantify these traceable actions and demand that they be of a high caliber. The calibration value is frequently accompanied by a traceable uncertainty statement to a given confidence level in order to communicate the calibration's quality. This is assessed via a rigorous study of uncertainty. It is occasionally necessary to run machinery that is out of specification (DFS). When this does occur, it must be documented and approved in writing by a manager with the help of a calibration professional [7], [8].

The physical quantities that measuring tools and equipment are intended to measure are classified. Internationally varying examples include NABL-141 in India and NIST 150-2G in the United States Together, these standards cover instruments that measure a variety of physical quantities, including mechanical quantities (limit switch, pressure gauge, pressure switch), thermodynamic or thermal properties thermometer, temperature controller, sound level meter, noise dosimeter, time and frequency (intervalometer), ionizing radiation (Geiger counter), light (light meter), and electromagnetic radiation (RF probes). A dead weight tester for pressure gauge calibration and a dry block temperature tester for temperature gauge calibration are two examples of the standard instruments used for each test device, respectively.

Goal and range

The design of the measuring device that needs to be calibrated is the first step in the calibration procedure. Through its calibration interval, the design must be able to "hold a calibration". The design must, in other words, be able to produce measurements that are "within engineering tolerance" when utilized within the specified environmental conditions for a sufficient amount of time. These design features increase the possibility that the actual measuring devices will function as predicted. In essence, calibration serves to preserve the accuracy of measurements and to guarantee that a certain equipment is functioning correctly [9]–[11].

Frequency

The precise method for determining tolerance values differs depending on the nation and the type of industry. Manufacturers typically provide measurement tolerances, recommend calibration intervals (CI), and specify the usage and storage environments when measuring equipment. The actual calibration period is often determined by the utilizing organization and is based on the anticipated amount of usage for this particular measuring equipment. Based on the outcomes of earlier calibrations, the assignment of calibration intervals may be a formal process. The recommended CI values are not made apparent in the standards themselves. The calibration process definition comes next. The calibration procedure's selection of a standard or standards is its most obvious step. The measurement uncertainty of the standard should ideally be less than 1/4 of that of the instrument being calibrated. When this objective is achieved, and the final measurement is also made using the 4:1 ratio, it is thought that the total measurement error of all the standards involved is negligible. The Handbook 52 that accompanied MIL-STD-45662A, an early US Department of Defense metrology program standard, is likely where this ratio was initially established. Up until the 1970s, it was 10:1, but as technology advanced, 10:1 became impractical for the majority of electronic measurements.

It is challenging to keep modern equipment's accuracy ratio at 4:1. Both the operating standard and the test equipment being calibrated might be equally precise. The calibration tolerance can be adjusted to make up the difference if the accuracy ratio is less than 4:1.

When 1:1 is attained, the only calibration that is 100 percent accurate is one in which the standard and the instrument being calibrated are an identical match. Reducing the accuracy of the calibrating device is another typical approach for handling this capability mismatch. To apply a 1% accuracy standard at 4:1, for instance, a gauge with a manufacturer-stated accuracy of 3% can be modified to 4%. The accuracy of the final readings won't change if the gauge's accuracy is dropped to 4% if the application calls for 16% accuracy. We refer to this as a restricted calibration. But the 3% gauge can never be better than 3.3:1 if the final measurement demands 10% precision. The gauge's calibration tolerance might then be changed, which might be a preferable option. The 1% standard would actually be between 99 and 101 units if the calibration is done at 100 units. 96 to 104 units, inclusive, would be the range of calibration values that are considered acceptable when the test equipment is set to a 4:1 ratio. All of the standards' potential contributions might be eliminated while maintaining a 3.3:1 ratio by changing the allowable range to 97 to 103 units. Further lowering the allowed range to 98 to 102 recovers a final ratio of greater than 4:1.

This is a condensed illustration. It is possible to argue against the example's mathematics. Whatever reasoning underpinned this procedure in a real calibration must be documented and made available. Informality is a factor in tolerance stacking and other challenging post calibration issues. In the aforementioned example, the calibration value of 100 units would ideally be the perfect location within the range of the gauge to carry out a single-point calibration. It can be the manufacturer's suggestion or the standard calibration procedure for similar devices. Also utilized are calibrations with multiple points. A zero unit state, or the lack of the phenomena being measured, may also be a calibration point, depending on the instrument. Or the user may be able to reset zero; there are various alternatives. Once more, the calibration points should be recorded. The standard and the device being calibrated may be connected using particular methods, which could have an impact on the calibration. The impedance of the cable connections, for instance, can have a direct impact on the outcome of electronic calibrations involving analog phenomena.

Automatic and manual calibrations

Modern gadgets can be calibrated manually or automatically. US serviceman calibrating a pressure gauge manually. On his left is the equipment being tested, and on his right is the test standard. An illustration of a manual procedure would be pressure gauge calibration. The process entails a number of procedures, including connecting the gauge being tested to a reference master gauge and an adjustable pressure source, applying fluid pressure to both gauges at predetermined places along the gauge's length, and comparing the readings of the two. To make sure the gauge's zero point and response to pressure adhere as closely as feasible to the intended accuracy, adjustments may be made. Manual record keeping is necessary for each stage of the procedure. An American serviceman using a 3666C auto pressure calibrator for automatic calibration An electronic control unit, a pressure intensifier used to compress a gas like nitrogen, a pressure transducer used to detect desirable levels in a hydraulic accumulator, and accessories like liquid traps and gauge fittings are all included in an automatic pressure calibrator A system that collects data automatically may also provide tools for automating data collecting for record-keeping.

Documentation and process descriptions

A calibration procedure, which is a particular test method, is used to gather all of the data previously mentioned. These instructions include each step required to complete a calibration successfully. One may be offered by the manufacturer, or an organization may create one that includes all of its other criteria. Calibration methods are centralized at clearinghouses like the

Government-Industry Data Exchange Program (GIDEP) in the US. Until transfer standards, certified reference materials, or natural physical constants the measurement standards with the lowest level of laboratory uncertainty are reached, this precise procedure is repeated for each of the standards utilized. This proves the calibration's traceability. Other elements that are taken into account when developing the calibration process are listed in metrology.

Individual instruments of the particular type covered above can then be calibrated after all of this. Usually, the process starts with a simple damage inspection. Before doing any routine maintenance, some businesses, like nuclear power facilities, collect "as-found" calibration data. An "as-left" calibration is carried out after routine maintenance and any flaws found during calibration have been corrected. More frequently, the entire procedure is left in the hands of a calibration professional, who also signs the calibration certificate that proves the calibration was successful. The fundamental procedure described above is a challenging and costly endeavor. As a rough rule of thumb, the annual cost of routine equipment support is typically around 10% of the initial purchase price. Even more expensive to maintain are exotic machines like scanning electron microscopes, gas chromatographs, and laser interferometers.

CONCLUSION

The ‘single measurement' gadget mentioned in the previous explanation of the fundamental calibration method does exist. However, depending on the organization, the majority of calibration-required devices may have many functionalities and ranges in a single instrument. A typical modern oscilloscope is a nice illustration. There are limits on how much of a comprehensive calibration can be automated, and there could potentially be 200,000 different configurations to calibrate. Usually used after calibration, tamper-proof seals prevent unauthorized access to an instrument. The oscilloscope rack picture demonstrates this and demonstrates that the instrument has not been removed since it was last calibrated as doing so could allow unwanted access to the instrument's adjusting components. There are labels that display the date of the most recent calibration as well as the period between calibrations, which indicates when the following one is required. To standardize record keeping and keep track of the accessories that are essential to a certain calibration condition, some organizations also give each instrument a special identification number.

REFERENCES

[1] W. T. Chyla, “On the structure of the New SI definitions of base units,” *Metrologia*, 2012, doi: 10.1088/0026-1394/49/4/L17.

[2] J. W. Lovell-Smith *et al.*, “Metrological challenges for measurements of key climatological observables. Part 4: Atmospheric relative humidity,” *Metrologia*. 2015. doi: 10.1088/0026-1394/53/1/R40.

[3] M. J. T. Milton, R. Davis, and N. Fletcher, “Towards a new SI: A review of progress made since 2011,” *Metrologia*. 2014. doi: 10.1088/0026-1394/51/3/R21.

[4] S.-H. Kim *et al.*, “Bolometer-Type Uncooled Infrared Image Sensor Using Pixel Current Calibration Technique,” *J. Sens. Sci. Technol.*, 2016, doi: 10.5369/jsst.2016.25.5.349.

[5] S. H. Kim *et al.*, “Reduction of fixed pattern noise in bolometer-type uncooled infrared image sensors using pixel current calibration technique,” in *Proceedings of IEEE Sensors*, 2017. doi: 10.1109/ICSENS.2017.8234089.

[6] M. Rezaiee-Pajand and S. Sinaie, "On the calibration of the Chaboche hardening model and a modified hardening rule for uniaxial ratcheting prediction," *Int. J. Solids Struct.*, 2009, doi: 10.1016/j.ijsolstr.2009.04.002.

[7] E. K. Bünemann *et al.*, "Soil quality – A critical review," *Soil Biology and Biochemistry*. 2018. doi: 10.1016/j.soilbio.2018.01.030.

[8] N. Gorla, T. M. Somers, and B. Wong, "Organizational impact of system quality, information quality, and service quality," *J. Strateg. Inf. Syst.*, 2010, doi: 10.1016/j.jsis.2010.05.001.

[9] S. Staats and L. Pierfelice, "Travel: A long-range goal of retired women," *J. Psychol. Interdiscip. Appl.*, 2003, doi: 10.1080/00223980309600630.

[10] A. M. Scherer, H. O. Witteman, J. Solomon, N. L. Exe, A. Fagerlin, and B. J. Zikmund-Fisher, "Improving the understanding of test results by substituting (Not Adding) goal ranges: Web-based between-subjects experiment," in *Journal of Medical Internet Research*, 2018. doi: 10.2196/11027.

[11] J. Tynjälä, P. Kangastupa, T. Laatikainen, M. Aalto, and O. Niemelä, "Effect of age and gender on the relationship between alcohol consumption and serum GGT: Time to recalibrate goals for normal ranges," *Alcohol Alcohol.*, 2012, doi: 10.1093/alcalc/ags072.

CHAPTER 3
HIGH ORDER STATISTICS FOR DOPPLER MICRO EMBOLIC SIGNAL DETECTION

Rajendra P. Pandey, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id- panday_004@yahoo.co.uk

ABSTRACT:

Strokes, which are the second leading cause of death worldwide, can be effectively avoided with the robust identification of the smallest circulating cerebral micro emboli. Most people agree that transcranial Doppler ultrasound is the most practical method for finding micro emboli. The Doppler energy signal is used to detect standards most frequently, and it is dependent on a constant threshold that has been determined empirically. On the other hand, higher order statistics, which are descriptive statistics that can be used to find signal outliers, have attracted a lot of attention in recent years. Based on the windowed calculation of the energy signal's third moment skewness and fourth moment kurtosis, we suggest novel forms of micro embolic detectors in this study. The energy distribution is unaltered and the skewness and kurtosis signals don't show any peak values during energy embolus-free times. The energy distribution is altered in the presence of emboli, and the skewness and kurtosis signals show peaks that correspond to the latter emboli. When used on actual signals, the skewness and kurtosis signals beat conventional approaches for the detection of micro emboli. For the skewness and kurtosis detectors, the sensitivities and specificities were 78%, 91%, 80%, and 90%, respectively.

KEYWORDS:

Doppler, Energy, Signal, Ultrasound.

INTRODUCTION

The Transcranial Doppler (TCD) signal's abrupt intensity increases are primarily interpreted as cerebral emboli's characteristics. Brain blood vessels can become blocked if cerebral emboli flow through them, which would cause a stroke, the second leading cause of death worldwide. Up to 14% of all strokes can be described as embolic. Because embolic strokes pose such a serious threat to life, it is crucial to find effective solutions for the problem of early identification of the smallest micro emboli. This early identification would serve as a foundation for an early diagnosis of stroke, preventing its occurrence. TCD is currently regarded as the most reliable embolic stroke diagnostic method. Even though the physical properties of embolic signals in the TCD signal have been thoroughly established, it is still difficult to detect embolic signals, especially small microembolic signals. The most reliable way for identifying the passage of emboli is the simultaneous visual and aural detection of the time-frequency representation (spectrogram) shown on the TCD screen and the abrupt "chirp" or "moan" caused by emboli. Due to temporal and frequency masking effects in audio files, one major drawback of the gold standard is its inability to acoustically detect microembolic impulses placed at the systolic phase [1], [2].

According to Rayleigh theory emboli backscatter ultrasound energy that is higher than that backscattered by the surrounding blood, so the standard signal processing method for detecting embolic signals is based on calculating the energy from the spectrogram and applying constant thresholds to pick up the emboli. The main drawbacks of conventional

approaches include their inability to identify microscopic microembolic signals that are weaker than the background blood around them, especially at the systolic peak. Numerous research projects have been carried out in an effort to find the smallest microemboli. Here are a few of the most prompt techniques.

The first frequency filtering techniques appeared. High detection sensitivity and specificity rates were reported by the study. A frequency filtering-based online automatic embolic signal detection system was subsequently created. With respect to sensitivity and specificity for a certain case (postcarotid endarterectomy), the latter method shown strong performances. However, the system's sensitivity and specificity were significantly reduced in other circumstances (atrial fibrillation). Additionally, it's arguable that the system performed less effectively and with significantly lower sensitivity and specificity when it came to the identification of low energy microembolic signals. Methods based on sudden change detection were first introduced. The parametric autoregressive methods and nonparametric detection techniques, primarily the Fourier, Wigner-Ville, and wavelet approaches, were contrasted. The identification of tiny microemboli has proven to be extremely effective and performant using the new parametric approaches. The approaches, however, were never evaluated on a set of genuine signals; instead, they were tested on synthetic, generated Doppler signals. Another extremely successful wavelet-based system was created.

The method successfully combined high levels of sensitivity and specificity. However, with low energy micro embolic signals, the system's rates dropped. In an impressive offline detection method was suggested. When it came to emboli with high intensities in comparison to background blood clutter, the system performed excellently. It should be noted, nonetheless, that the study did not explore the possibility of picking up faint embolic signals. Another extremely effective detection method based on the discrete wavelet transform (DWT) was introduced by the authors in. Major gains in specificity and sensitivity were possible with DWT. However, the reduced frequency resolution at low frequency scales, where embolic signals are primarily detected, was a significant flaw in the DWT implementation. The authors of suggested embolic detection using nervously classification and an adaptive wavelet packet basis. We created a sparse representation of the Doppler ultrasonography blood flow signals using the adaptive wavelet packet basis. The technique yielded extremely precise and reliable results. The correlated specificity, however, was not measured when the method was compared to other approaches; only the sensitivity was considered. Instead of using the short time Fourier transform, which is the typical method of detection in TCD systems, the paper filed in suggested the use of the fractional Fourier transform. The findings demonstrated that the Fractional Fourier Transform-based discriminating characteristics facilitate the analysis and detection of embolic signals [3], [4].

DISCUSSION

This method was not shown to be reliably adequate for the detection of the smallest micro emboli, despite its simplicity and acceptable findings. Although the approach suggested in obtained extremely high sensitivity and specificity, significant detection mistakes were made as a result of the modest reflected signals from gaseous emboli. The key drawback of the majority of the previously introduced articles is that the detection is based on time-varying information while the threshold is fixed. Two ways can be suggested to fit the time-varying information with the threshold. The first involves suggesting a time-varying threshold, as in that reflects the decision information's time-varying trend. The second step is to suggest a fixed threshold that corresponds to the decision data and eliminates the time-varying trend. The methods for comparing an energy free of its time-varying trend to a fixed threshold that we suggested in this work are based on the use of high order statistics (HOS) of windowed

Doppler energy signals. When asymptomatic carotid artery patients are tracked with a Holter TCD, we frequently demonstrate the skewness and kurtosis as two reliable ways to detect micro embolic signals. A Doppler radar is a specialized radar that generates velocity information about distant objects using the Doppler Effect. This is accomplished by sending a microwave signal at a target and measuring how the motion of the target affects the frequency of the returned signal. The radial component of a target's velocity in relation to the radar can be directly and extremely accurately measured using this variation.

The phrase refers to radar systems used in a variety of fields, including meteorology, police radar detectors, navigation, and aircraft. The difference between a wave's seen frequency and its emitted frequency for an observer traveling in relation to the wave source is known as the Doppler Effect also known as the Doppler shift and is named after Austrian physicist Christian Doppler who first postulated it in 1842. It frequently occurs as a siren-sounding vehicle approaches, passes, and moves away from a viewer. When approaching, the received frequency is greater than the emitted frequency; when passing by, it is the same; and when reaccessioning, it is lower. This frequency variation is also influenced by the motion of the wave source relative to the observer; it is greatest when the source is moving directly in front of or behind the observer and decreases with increasing angle between motion and wave direction until there is no shift when the source is moving perpendicular to the observer.

Consider a baseball pitcher who pitches a ball to the catcher once per second (1 ball per second). The catcher catches one ball every second if the pitcher is stationary and the balls move at a steady speed. The catcher, however, catches balls more frequently because they are less spaced out (the frequency increases) when the pitcher is jogging towards the direction of the catcher. If the pitcher is moving away from the catcher, the opposite is true. The frequency of catches by the catcher drops as a result of the pitcher's backward motion. The frequency variation at which the receiver catches balls is lower if the pitcher moves at an angle but maintains the same pace because the distance between the two changes less rapidly. When tossing balls or sending out microwaves, the frequency is the same from the pitcher's perspective. The wavelength of the waves is also impacted because frequency and wavelength are inversely correlated in electromagnetic radiation, such as microwaves or sound. Thus, the Doppler Effect results from the relative velocity difference between a source and an observer.

Navigation

Doppler radars were utilized by both aircraft and spacecraft as a navigational aid. For the first time, the wind speed could be precisely calculated by using the radar to measure the movement of the ground and then comparing that measurement to the airspeed returned by the aircraft instruments. Dead reckoning was then performed using this value with extreme accuracy. The Green Satin radar employed in the English Electric Canberra is one of the earliest instances of such a system. In order to employ a single antenna for both transmission and reception, this technology emitted pulsed signals with an extremely low repetition rate. For comparing the received signal to the reference frequency, an oscillator was used. The first "fix" was really obtained via a radio navigation system, usually Gee, and the Green Satin then gave precise long-distance navigation outside of Gee's 350-mile range. Similar systems were employed in a number of aircraft of the time, and by the 1960s, they had been integrated into the primary search radars of fighter designs [5], [6].

Prior to being completely replaced by inertial navigation systems in the 1960s, doppler navigation was widely used in commercial aircraft. A transmitter/receiver unit, a processing unit, and an antenna platform with gyro stabilization made up the apparatus. A servo motor

rotated the antenna, which produced four beams and balanced the Doppler shift from the left and right antennas to line with the aircraft's track. An indication of the "drift angle" was provided by a synchro that relayed the platform angle to the flight deck. The Doppler shift between the forward and aft facing beams was used to calculate the ground speed. These were shown on a single instrument on the flying deck. There was an additional "Doppler Computer" in some planes. This mechanical device contained a steel ball that was revolved by a motor whose speed was managed by the ground speed as estimated by the Doppler Effect.

This motor's 'drift angle' determined its angle. Two fixed wheels, one driving forward and the other driving backward, drove counters to produce track distance and track difference. A desired track could be set between two waypoints on a great circle route over water thanks to the integration of the aircraft's compass into the computer. Readers in the twenty-first century might find it odd, but it actually worked rather well and was a significant improvement over other "dead reckoning" techniques available at the time. Typically, it was supported by position fixes from the Loran system, or as a last option, a sextant and chronometer. When within range of a few VORs or NDBs, it was possible to cross the Atlantic with a handful of mile error. The sea state was its main weakness in practice since a calm sea produced low radar returns and, as a result, inaccurate Doppler readings. But in the North Atlantic, this wasn't common.

Micro embolic Detection Typically Used

A material that has traveled from another part of the body through the bloodstream (embolism) to block an artery in the lungs is called a pulmonary embolism chest pain, especially while inhaling, and bloody coughing are all signs of a] Leg symptoms, such as a red, heated, swollen, and painful leg, may also be present in blood clots in the legs. Low blood oxygen levels, fast breathing, a quick heartbeat, and occasionally a minor temperature are symptoms of severe cases can include obstructive shock, passing out, extremely low blood pressure, and rapid death. PE is typically brought on by a blood clot that starts in the leg and moves to the lung.

Advanced age, cancer, extended bed rest and immobility, smoking, stroke, long-distance travel over four hours, specific hereditary diseases, estrogen-based medication, pregnancy, obesity, trauma or bone fracture, and after various forms of surgery all raise the risk of blood clots. The embolization of air, fat, or amniotic fluid accounts for a very tiny percentage of cases. Test results are combined with symptoms and indicators to make a diagnosis. A blood test called a D-dimer may rule out the disorder if the risk is low. Otherwise, the diagnosis could be confirmed by a CT pulmonary angiography, a lung ventilation/perfusion scan, or a leg ultrasound. Venous thromboembolism (VTE) is the collective name for deep vein thrombosis and PE[7]–[9].

PE can be avoided by moving as soon as feasible following surgery, performing lower leg movements while seated, and using blood thinners after certain procedures.[16] Anticoagulants such heparin, warfarin, or one of the direct-acting oral anticoagulants (DOACs) are used as a form of treatment. These are advised for a minimum of three months. Patients with severe cases may need surgery (a pulmonary thrombectomy) or thrombolysis using drugs such tissue plasminogen activator (tPA) either intravenously or through a catheter. A temporary vena cava filter may be employed in place of blood thinners if necessary. In Europe, pulmonary emboli injure roughly occur annually in the United States, resulting in at least 40,000 fatalities. Male and female rates are comparable. As people age, they appear more frequently.

Symptoms and signs

The signs and symptoms of a pulmonary embolism can include one or more of the following: hemoptysis (coughing up blood), dyspnea (shortness of breath), tachypnea (rapid breathing), and chest discomfort that is "pleuritic" in character (made worse by breathing), cough, and shortness of breath. Cyanosis (blue staining, typically of the lips and fingers), collapse, and circulatory instability due to decreased blood flow via the lungs and into the left side of the heart are symptoms of more severe cases. PE is responsible with 15% of all cases of sudden death. Less than 1% of syncope instances are caused by PE, despite the fact that PE can present with syncope. The lungs typically appear healthy after a physical examination. On rare occasions, the damaged portion of the lung may produce a pleural friction rub (more often in PE with infarct). There may occasionally be an exudative pleural effusion, indicated by a diminished percussion note, audible breath noises, and vocal resonance. A left parasternal heave, a loud pulmonary component of the second heart sound, and/or elevated jugular venous pressure are all signs of right ventricle strain. If pulmonary hemorrhage or infarction is present, there may be a low-grade fever as well. Smaller pulmonary emboli are more likely to result in lung infarction and minor effusions, both of which are uncomfortable, but not hypoxia, dyspnea, or hemodynamic instability like tachycardia since they have a tendency to lodge in more peripheral locations without collateral circulation. Larger PEs, which frequently lodge in the center, frequently result in dyspnea, hypoxia, low blood pressure, a rapid heartbeat, and fainting, although they frequently cause little pain because there is no lung infarction because of collateral circulation [10]–[12].

A huge fragmented embolism that causes both large and tiny PEs is most likely what causes the characteristic PE presentation of pleuritic discomfort, dyspnea, and tachycardia. Because they are painless and frequently mirror other disorders, little PEs are frequently ignored because they only cause pleuritic pain and no other symptoms, and large PEs are frequently overlooked because they are often painless and cause small increases in troponin and brain natriuretic peptide levels. Depending on the clinical signs and symptoms, PEs can be classified as large, submissive, or nonpassive. Although the precise definitions of these are unclear, hemodynamic instability is generally considered as a sign of large PE. Obstructive shock is caused by this, which manifests as persistently low blood pressure, a sluggish heartbeat, or pulselessness.

Diagnosis

An individual with a right lower lobe pulmonary embolism has a Hampton hump. It is advised to study the clinical criteria to determine whether further testing is necessary to diagnose a pulmonary embolism. Further testing is often not required in people who are at low risk, defined as those who are under 50, have a heart rate under 100 beats per minute, have an oxygen saturation of over 94% on room air, and do not exhibit limb edema, bloody coughing, recent surgery or trauma, blood clots, or estrogen use. Additional testing is required in cases when there are more high-risk individuals. Due to its simplicity and accuracy, a CT pulmonary angiography (CTPA) is the procedure of choice for diagnosing a pulmonary embolism although a CTPA is recommended, additional tests can be performed as well. A proximal lower limb compression ultrasound (CUS) is one possible method. This test is typically utilized as a confirmatory test, which means it validates results from an earlier investigation that indicated the presence or a possible presence of a pulmonary embolism. A cross-sectional investigation found that the sensitivity and specificity of CUS tests were both 41% and 96%.

Imaging is then performed if previous tests have indicated that a PE diagnosis is likely if there are any concerns, in order to assess the chances of being able to confirm the diagnosis by imaging Because the typical clinical presentation (chest pain, shortness of breath) cannot be clearly distinguished from other causes of chest pain and shortness of breath, the diagnosis of PE is mostly dependent on validated clinical criteria supplemented with selected testing. The medical history, symptoms, and physical examination findings are used to make the choice to undertake medical imaging, which is then followed by an evaluation of clinical likelihood. The outcomes were evident. The typical approach based on second-order statistics was substantially outperformed by the methods based on HOS. The HOS sensitivity in altering the distribution form is what explains this superiority. We suggest using the skewness and kurtosis as novel techniques for micro embolus detection as we are aware that the presence of a microembolus superimposed on the Doppler energy signal imposes changes in the distribution of this signal. The distribution of the Doppler energy signals remains fixed during embolus-free times, and its skewness and kurtosis are never changed. They don't exhibit any differences. The skewness and kurtosis signals are changed when a microembolus is superimposed on the energy signal, and the embolus is given a peak with a peaked Ness level that is higher than all the other points in the signal. The detection can outperform conventional techniques. The skewness and kurtosis-based detection produced notable gains after being evaluated on a set of genuine signals, including very high specificity reaching up to 91% and 90%, respectively, compared to the 60% attained by the traditional method. Additionally, the sensitivity is raised for skewness and kurtosis-based detectors, respectively, from 65% for traditional approaches to 78% and 80% [13], [14].

In light of this, we may state that skewness and kurtosis can provide a robust and more reliable detection than conventional detection methods, and can thus be thought of as new ways for boosting micro embolic detection systems. Given that we have two detectors, one based on skewness detection and the other on kurtosis detection, it is useful to notice that they both work fairly similarly and produce results that are quite near. The only difference that can be seen is that the skewness signal swings more strongly around the embolic peaks than the kurtosis signal does around the detected embolic peak. As a result, the detection threshold for kurtosis detection may be set more quickly and robustly, giving it a slight advantage.

CONCLUSION

In this paper, we suggest two detectors for improved cerebral microembolus detection based on calculations of the skewness and kurtosis of the Doppler energy signal. The skewness and kurtosis-based detectors allow increase of both the sensitivity and specificity in comparison to the traditional detector where the detection is carried out directly on the energy signal. This work highlights the significant challenges that traditional micro embolic energy detectors with empirical threshold still have for the reliable detection of micro emboli. It also demonstrates how much more effectively micro emboli, which are precursors of larger upcoming emboli with high stroke risks, can be detected using detectors that incorporate detection based on skewness and kurtosis calculation from the energy. Therefore, adopting these easy-to-use detectors would be an additional tool enhancing attempts to lower the frequency of strokes.

The next stage would involve making an effort to improve the approaches' overall performance, particularly in terms of sensitivity, and validating the created algorithms on a larger database. Additionally, we plan to use automatic artifact rejection approaches rather than manual ones in the entire detection system.

REFERENCES

[1] C. A. Vollgraff Heidweiller-Schreurs *et al.*, "Prognostic accuracy of cerebroplacental ratio and middle cerebral artery Doppler for adverse perinatal outcome: systematic review and meta-analysis," *Ultrasound in Obstetrics and Gynecology*. 2018. doi: 10.1002/uog.18809.

[2] A. A. Oglat, M. Z. Matjafri, N. Suardi, M. A. Oqlat, M. A. Abdelrahman, and A. A. Oqlat, "A review of medical doppler ultrasonography of blood flow in general and especially in common carotid artery," *J. Med. Ultrasound*, 2018, doi: 10.4103/JMU.JMU_11_17.

[3] A. K. P. Lim, K. Satchithananda, E. A. Dick, S. Abraham, and D. O. Cosgrove, "Microflow imaging: New Doppler technology to detect low-grade inflammation in patients with arthritis," *Eur. Radiol.*, 2018, doi: 10.1007/s00330-017-5016-4.

[4] A. Gaspar, P. Azevedo, and R. Roncon-Albuquerque, "Non-invasive hemodynamic evaluation by Doppler echocardiography," *Revista Brasileira de Terapia Intensiva*. 2018. doi: 10.5935/0103-507X.20180055.

[5] H. Eichenbaum, "The role of the hippocampus in navigation is memory," *Journal of Neurophysiology*. 2017. doi: 10.1152/jn.00005.2017.

[6] C. Seguin, M. P. Van Den Heuvel, and A. Zalesky, "Navigation of brain networks," *Proc. Natl. Acad. Sci. U. S. A.*, 2018, doi: 10.1073/pnas.1801351115.

[7] C. Boulon *et al.*, "[Detection of micro-embolic signals: a review of the literature].," *J. Mal. Vasc.*, 2012.

[8] O. Seguchi *et al.*, "The Clinical Relevance of Transcranial Doppler Detection of Micro-Embolic Signals in Patients with Heartmate II," *J. Hear. Lung Transplant.*, 2016, doi: 10.1016/j.healun.2016.01.687.

[9] M. A. Topcuoglu, "Transcranial Doppler ultrasound in neurovascular diseases: Diagnostic and therapeutic aspects," in *Journal of Neurochemistry*, 2012. doi: 10.1111/j.1471-4159.2012.07942.x.

[10] O. A. Gbinigie, J. M. Ordóñez-Mena, T. R. Fanshawe, A. Plüddemann, and C. Heneghan, "Diagnostic value of symptoms and signs for identifying urinary tract infection in older adult outpatients: Systematic review and meta-analysis," *J. Infect.*, 2018, doi: 10.1016/j.jinf.2018.06.012.

[11] D. V. Skjelstad, U. F. Malt, and A. Holte, "Symptoms and signs of the initial prodrome of bipolar disorder: A systematic review," *Journal of Affective Disorders*. 2010. doi: 10.1016/j.jad.2009.10.003.

[12] C. Fonseca *et al.*, "The diagnosis of heart failure in primary care: Value of symptoms and signs," *Eur. J. Heart Fail.*, 2004, doi: 10.1016/j.ejheart.2004.08.002.

[13] C. Z. Zhang *et al.*, "Saliva in the diagnosis of diseases," *International Journal of Oral Science*. 2016. doi: 10.1038/ijos.2016.38.

[14] I. K. Sand, "Classification, diagnosis, and differential diagnosis of multiple sclerosis," *Current Opinion in Neurology*. 2015. doi: 10.1097/WCO.0000000000000206.

CHAPTER 4
COMPARING ORDER STATISTICS AND SPACINGS STOCHASTICALLY

Rupal Gupta, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id-r4rupal@yahoo.com

ABSTRACT:

Reviewing some recent advancements in stochastic comparisons of order statistics and sample spacings, we look at some of the most recent changes. We take into account both the same and nonidentically dispersed parent observations. However, we will frequently assume that the observations are unrelated. The proportional hazard rate model and the case of independent exponentials with different scale parameters are both thoroughly examined. The Oxford English Dictionary defines the word stochastic as initially being used as an adjective with the connotation "pertaining to conjecturing" and coming from a Greek word that meant "to aim at a mark, guess." Regarding Bernoulli, this adage was employed. Physics researchers Stanislaw Ulam, Enrico Fermi, John von Neumann, and Nicholas Metropolis made the stochastic Monte Carlo method mainstream. The repeated nature of the operation and the use of randomization are similar to the actions carried out in a casino. Methods of simulation and statistical sampling typically do the opposite, utilizing simulation to assess a deterministic problem that was already known to exist. Although there are historical instances of an "inverted" strategy, it wasn't until the Monte Carlo method gained prominence that it became widely accepted.

KEYWORDS:

Conjecturing, Metropolis, Medians, Statistics.

INTRODUCTION

The means and medians of two random variables are the easiest and most widely used way to compare the magnitudes of two random variables. The median of occasionally may be higher than that of while the mean occasionally may be lower than that of. However, if the random variables are stochastically ordered, this mistake won't occur. The same may also occur if one were to compare the variability of with that of just using numerical measurements like standard deviation, etc. Furthermore, in other circumstances, these distributional properties may not be present. Most often, knowledge about the underlying distributions can be expressed in terms of their survival functions, quantile functions, hazard rate functions, mean residual functions, and other appropriate functions of probability distributions. These techniques offer a lot more information than those that rely just on a few number of numerical distribution features. Comparing random variables using these functions typically results in the establishment of partial ordering. They are referred to as stochastic orders [1], [2].

In many areas of statistics, stochastic models are usually sufficiently complex. It is practical significance to obtain boundaries and approximations for their properties. In other words, a stochastic model's approximation by a simpler model or a model with simple constituent components may result in useful bounds and approximations for some specific and desirable

model properties. It is also quite interesting to investigate how a model's attributes change when its component parts change. Because random variables are used in the stochastic parts of models, the subject of stochastic ordering among random variables is crucial in these fields. This subject is well-treated in books by Muller and Stoyan and Shaked and Shanthi Kumar. Many fields of statistics are quite interested in order statistics, and scholars have given them a lot of attention. Let the variables be random. The smallest of all's, the order statistic, is represented by. In reliability engineering, a component system is referred to as an "out of" system if and only if at least one of the components functions.

A -out-of- system's lifetime can be expressed as. Series systems are out-of-systems, whereas parallel systems are out-of-systems. Therefore, researching lifetimes of systems that are not in use is analogous to researching the stochastic characteristics of order statistics. In statistics, generally speaking, and specifically in the context of life testing and reliability models, spacings, the disparities between consecutive order statistics, and their roles are also significant. Excellent sources of information on this subject include the books by David and Nagaraj, and two volumes of papers by Balakrishnan and Rao However, the majority of this research has been limited to the scenario in which the observations have a same distribution and are independent. The observations are not always i.i.d. in many real-world circumstances, such as reliability theory. Researchers have only started to pay attention to this subject in the last 20 years or so. Pledger and Proustian have discovered some intriguing partial ordering results on order statistics and spacings from independent but nonidentically random variables. We will examine some recent advancements in the field of stochastic comparisons of order statistics and spacings in this review paper [3]–[5].

We shall assume that all distributions under investigation are perfectly continuous and that rising means nondecreasing and decreasing means nonincreasing throughout this chapter. Enrico Fermi's 1930 use of a random approach to determine the properties of the newly found neutron may be its most well-known early application. Although they were severely constrained by the available computer resources at the time, Monte Carlo methods were essential to the simulations needed for the Manhattan Project. Therefore, extensive research into Monte Carlo methods didn't start until after the first electronic computers were constructed they were first applied at Los Alamos for early work on the hydrogen bomb's development in the 1950s, and they later gained popularity in the disciplines of physics, physical chemistry, and operations research. During this period, Monte Carlo methods were funded and disseminated by a number of significant institutions, including the RAND Corporation and the U.S. Air Force, and they started to find extensive use in a variety of sectors.

DISCUSSION

The use of Monte Carlo methods necessitated the creation of pseudorandom number generators, which were far quicker to employ than the random number tables that had previously been used for statistical sampling since they required enormous quantities of random values. It has been discovered that adding stochastic "noise" to biological systems might enhance the signal intensity of the internal feedback loops that regulate balance and other vestibular functions. It has been discovered to aid in balance control in diabetic and stroke patients. Numerous biological processes are amenable to stochastic analysis. For instance, the molecular collisions that occur during the binding and unbinding of RNA polymerase to a gene promoter, as well as other stochastic events, contribute to the expression of genes. Stochastic models are used in the financial markets to depict the behavior of various financial assets that appears to be random, including the behavior of interest rates and the price of one currency relative to another. Quantitative analysts use these

models to value options on stock prices, bond prices, and interest rates; for more information, see Markov models. Additionally, it is the foundation of the insurance sector.

Concepts and Orderings of Dependence

According to dependency theory, resources move from a "periphery" of underdeveloped and impoverished nations to a "core" of developed states, enriching the latter at the expense of the former. According to dependence theory, the way in which poor states are included into the "world system" causes rich states to become richer while poor states become poorer. Following World War II, this idea was formally formulated in the late 1960s as researchers looked for the underlying cause of Latin America's lack of progress. The theory emerged in response to modernization theory, an earlier development theory that claimed all societies go through similar stages of development, that today's underdeveloped areas are therefore in a similar situation to that of today's developed areas at some point in the past, and that, therefore, the task of helping the underdeveloped areas out of poverty is to accelerate them along this supposed common path of development, through various means such as investments. Dependency theory disagreed, contending that poor nations have distinct characteristics and systems that set them apart from affluent nations and, more crucially, that they are the weaker participants in a global market economy [6]–[8].

Some authors have defended its continued applicability as a conceptual framework for the distribution of wealth globally. Liberal reformists and neo-Marxists are the two main subgroups of dependency theorists. While neo-Marxists support a command-centered economy, liberal reformists frequently support targeted policy interventions. In his analysis, Baran put capital accumulation and surplus extraction at the forefront. A population must produce more than it needs for bare subsistence in order for development to occur (a surplus). Furthermore, if development is to take place, some of that excess must be used for capital accumulation, or the purchase of new means of production; using the surplus for things like luxury consumption does not result in progress. Baran identified two major categories of economic activity in underdeveloped nations. In the more traditional of the two, plantation agriculture, which dates back to colonial times, the majority of the surplus goes to the landowners, who use it to mimic the consumption patterns of wealthy people in developed countries. As a result, much of it is used to buy luxury goods made abroad, such as cars and clothing, and little is saved for investing in development. Industry, although of a specific kind, is the more modern kind of economic activity in the periphery. Foreigners typically carry it out, though frequently in coordination with regional interests. It frequently falls under special tariff protection or other forms of government assistance. The surplus from this production primarily goes to two places: a portion is returned to the foreign stockholders as profit, and a second portion is used for ostentatious expenditure in a manner akin to the plantation aristocracy. Once more, not much is used for development. Baran believed that in order to end this cycle, political upheaval was required [9]–[11].

The Latin American Structuralist school asserted that the system is more flexible than the Marxists thought in the 1960s. As support for this theory, they pointed to the period attempts at industrialization in Latin America (including those in Argentina, Brazil, and Mexico), which were only partially successful. They were persuaded to believe that dependence exists between nations with varying levels of industrialization, not between commodity exporters and industrialized nations. Their strategy makes a contrast between the political and economic sectors. The former may be developed or underdeveloped, while the latter even if (slightly) developed economicallymay be independent or dependent on the former Page not found in more recent times, Guillermo O'Donnell has claimed that the military coups in Latin

America that came to support development under authoritarian disguise eliminated the restrictions on growth imposed by neoliberalism.

The Latin American Structuralists emphasized the significance of multinational enterprises and government-sponsored technological advancement. Tanzler has distinguished between spurious competitiveness, which is based on low wages, and systemic or true competitiveness, which is the capacity to compete on the basis of superior productivity. The viability or desirableness of "dependent development" was questioned in part by the third-world debt crises of the 1980s and the 1990s' persistence of stagnation in Africa and Latin America. Contrary to what traditional dependency theorists assert, the difference in financial standing between the core and peripheral countriesparticularly the latter's inability to borrow in its own currency is the sine qua non of the dependency relationship. Because of the significance of its financial markets and the fact that it controls the international reserve currency the US dollar hegemony of the United States is quite powerful, in his opinion. He thinks that the United States' position was significantly boosted by the termination of the Bretton Woods international financial agreements in the early 1970s since it freed them from some financial restraints.

In contrast to Marxism, "standard" dependence theory argues against internationalism and any optimism that less developed countries will make progress toward industrialization and a freeing revolution. Theotokion dos Santos defined a "new dependency" based on a Marxian analysis that concentrated on the internal and external relationships of less developed peripheral nations. While in political exile in the 1960s, former Brazilian President Fernando Henrique Cardoso (in power from 1995 to 2002) published a great deal about dependence theory and argued that it was a method for examining the economic inequalities between the center and the periphery. According to Cardoso, his interpretation of dependency theory is as follows [12]–[14].

Majoritarian Thoughts and Related Orderings

The MM algorithm is an iterative optimization technique that uses a function's convexity to discover its maximum and minimum values. Depending on whether the desired optimization is a minimization or a maximization, the MM stands for "Majorize-Minimization" or "Minorize-Maximization". MM is not an algorithm despite its name; rather, it is a guide on how to build an optimization method. One can think of the expectation-maximization algorithm as a particular instance of the MM algorithm. Convexity and inequality are the main focus of the MM method, which is generally easier to learn and use. Conditional expectations are typically involved in the EM algorithm. The origins of the MM algorithm can be traced at least as far back as 1970, when Ortega and Rheingold were researching line search techniques. Hunter and Lange proposed "MM" as a generic framework in 2000. The approach has been used in recent studies in a variety of academic areas, including mathematics, statistics, machine learning, and engineering.

When the observations are not evenly distributed, dependence orderings among order statistics might occur. The Kolmogorov-Smirnov test (K-S test or KS test), which can be used to compare one sample with a reference probability distribution or two samples with each other (two-sample K-S test), is a nonparametric test of the equality of continuous one-dimensional probability distributions. In essence, the test responds to the questions, "How likely is it that we would see two sets of samples like this if they were drawn from the same (but unknown) probability distribution?" and "How likely is it that we would see a collection of samples like this if they were drawn from that probability distribution?" It bears the names of Nikolai Smirnov and Andrey Kolmogorov.

The Kolmogorov-Smirnov statistic measures the separation between the empirical distribution functions of two samples, or between the empirical distribution functions of the sample and the cumulative distribution function of the reference distribution. When using one sample, or two samples when using two samples, the null distribution of this statistic is derived under the null hypothesis that the sample was taken from the reference distribution. The distribution taken into account in the one-sample scenario can be continuous completely discrete. The distribution taken into account for the null hypothesis in the two-sample situation (see Section 3) is continuous but otherwise unrestricted. The two-sample test can also be conducted under more open-ended circumstances that take discontinuity, heterogeneity, and sample dependence into account [15].

As it is sensitive to variations in both the location and form of the empirical cumulative distribution functions of the two samples, the two-sample K-S test is one of the most helpful and broad nonparametric approaches for comparing two samples. A goodness of fit test can be added to the Kolmogorov-Smirnov test. Samples are normalized and compared to a reference normal distribution in the special situation of testing the distribution's normality. It is well known that utilizing these to establish a particular reference distribution affects the null distribution of the test statistic (see Test with estimated parameters), and doing so is equivalent to setting the mean and variance of the reference distribution equal to the sample estimates. Studies have shown that the test is less effective than the Shapiro-Wilk test or the Anderson-Darling test for determining normalcy, even in its corrected form. These additional tests do, however, have their own drawbacks. For instance, it is commonly known that the Shapiro-Wilk test performs poorly in samples containing lots of identical values.

Test using approximation parameters

The critical values obtained in this manner are invalid if either the form or the parameters of F(x) are determined from the data Xi. Although tables have been developed for some situations, Monte Carlo or other techniques may be needed in others. The critical values for the normal distribution, the exponential distribution, and the necessary adjustments to the test statistic have all been reported in detail, and later publications have also included the Gumbel distribution. An example of this for the normal distribution is the Lilliefors test. When the Kolmogorov test data does not appear to suit the expectation that it came from a normal distribution, the logarithm transformation may be able to help.

Which estimating technique should be utilized is a question that arises when using estimated parameters. Normally, the maximum likelihood approach would be used, however for the normal distribution, for instance, MLE has a significant bias error on sigma. Instead, employing a moment fit or KS minimization has a significant effect on the critical values and a limited effect on test power. A ML estimate based on H0 data is normal, thus use the standard deviation for scale) would produce a substantially bigger KS distance than a fit with minimum KS if the student-T data with df = 2 was to be determined whether the data could be normal or not. In this situation, we should reject H0, as is frequently the case with MLE, because the sample standard deviation for T-2 data may be quite big, but KS minimization may still result in a KS that is too low to reject H0. In the Student-T scenario, the KS test is in fact somewhat poorer when the KS estimate is used in place of the MLE. In some circumstances, nevertheless, a modified KS test achieves marginally higher test power.

CONCLUSION

According to report a multivariate Kolmogorov-Smirnov goodness of fit test without distribution is available. In order to compute the statistic in the bivariate case, an algorithm was created. The statistic is constructed using Rosenblatt's transformation. It also includes a

rough test that may be calculated with ease in any dimension. If a comparable test is to be used on multivariate data, the Kolmogorov-Smirnov test statistic needs to be changed. This is complicated by the fact that, usually speaking, the maximum difference of any complementary distribution function is not the same as the greatest difference between two joint cumulative distribution functions. So, depending on which of them is chosen, the maximum difference will change. The arrangement is P or either of the other two alternatives. The outcome of the test being utilized might not need to depend on the decision made, if one so desire. Comparing the CDFs of the two samples with all feasible orderings and choosing the largest of the set of resulting KS statistics is one method for generalizing the Kolmogorov-Smirnov statistic to higher dimensions that addresses the aforementioned concern. The number of such orderings in d dimensions is 2d 1. Two of these variations are credited to Peacock and Fasano and Franceschini see Lopes et al. for a comparison and computational details Simulations can produce critical values for the test statistic, but they are subject to the joint distribution's dependence structure. Kolmogorov-Smirnov statistics and the so-called star discrepancy equivalent in one dimension, therefore using D would be a straightforward way to extend KS to higher dimensions in a native manner. Sadly, high-dimensional calculations of the star discrepancy are challenging.

REFERENCES:

[1] B. Kim, J. Kim, and S. Lee, "Strong unimodality of discrete order statistics," *Stat. Probab. Lett.*, 2018, doi: 10.1016/j.spl.2018.04.018.

[2] K. Davies and A. Dembińska, "Computing moments of discrete order statistics from non-identical distributions," *J. Comput. Appl. Math.*, 2018, doi: 10.1016/j.cam.2017.07.017.

[3] S. Calik and A. T. Bugatekin, "Moments of sample extremes of order statistics from discrete uniform distribution and numerical results," *Therm. Sci.*, 2018, doi: 10.2298/TSCI170612291C.

[4] S. Li, R. He, B. Lin, and F. Sun, "DOA estimation based on sparse representation of the fractional lower order statistics in impulsive noise," *IEEE/CAA J. Autom. Sin.*, 2018, doi: 10.1109/JAS.2016.7510187.

[5] N. Ben Rached, Z. Botev, A. Kammoun, M. S. Alouini, and R. Tempone, "On the Sum of Order Statistics and Applications to Wireless Communication Systems Performances," *IEEE Trans. Wirel. Commun.*, 2018, doi: 10.1109/TWC.2018.2871201.

[6] G. Kimeldorf and A. R. Sampson, "A framework for positive dependence," *Ann. Inst. Stat. Math.*, 1989, doi: 10.1007/BF00049108.

[7] J. Rodríguez-Lallena and M. Úbeda-Flores, "A new class of bivariate copulas," *Stat. Probab. Lett.*, 2004, doi: 10.1016/j.spl.2003.09.010.

[8] J. Zhang, W. Klingsch, A. Schadschneider, and A. Seyfried, "Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram," *J. Stat. Mech. Theory Exp.*, 2012, doi: 10.1088/1742-5468/2012/02/P02002.

[9] K. D. Priddy, "Is there logic behind fetal monitoring?," *JOGNN - J. Obstet. Gynecol. Neonatal Nurs.*, 2004, doi: 10.1177/0884217504268942.

[10] A. Lou and C. Grosvenor, "Selective Laser Sintering , Birth of an Industry," *Univ. Texas Austin, Dep. Mech. Eng. 100 Years*, 2012.

[11] J. Southworth *et al.*, "Parks, people and pixels: Evaluating landscape effects of an East African national park on its surroundings," *Trop. Conserv. Sci.*, 2010, doi: 10.1177/194008291000300202.

[12] M. Boianovsky and R. Solís, "The Origins and Development of the Latin American Structuralist Approach to the Balance of Payments, 1944-1964," *Rev. Polit. Econ.*, 2014, doi: 10.1080/09538259.2013.874191.

[13] L. C. Bresser-Pereira and P. Gala, "Macroeconomia estruturalista do desenvolvimento," *Rev. Econ. Política*, 2010, doi: 10.1590/s0101-31572010000400007.

[14] C. Kay, "Latin American Structuralist School," in *International Encyclopedia of Human Geography*, 2009. doi: 10.1016/B978-008044910-4.00105-X.

[15] H. Nurmi, *Comparing Voting Systems*. 1987. doi: 10.1007/978-94-009-3985-1.

CHAPTER 5
NEUROMORPHIC STATISTICS FOR PRODUCT ACCEPTANCE DETERMINATION WITH MEASUREMENT ERROR

Vineet Saxena, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id-tmmit_cool@yahoo.co.in

ABSTRACT:

The measurement method that produced the variable data was truncated or lacking in clarity due to measurement error. When data or parameters are ambiguous, imprecise, or unclear, the industry might use neuromorphic statistics, which is an extension of classical statistics. In this publication, a brand-new sampling strategy for measurement error is developed utilizing neuromorphic statistics. The sample size and acceptance number are two neuromorphic criteria in the suggested sampling scheme. Additionally included is the neuromorphic operational function. The neuromorphic optimization problem will be used to calculate the neuromorphic plan parameters. For a few supplied parameters, some tables are provided. According to the comparative analysis, the proposed sample plan is superior to the current sampling plan under classical statistics in terms of flexibility, adequacy, and effectiveness in an uncertain environment. To serve as an example, a real-world situation is used. The discrepancy between a quantity's measured value and its true value is known as observational error also known as measurement error an error is not always a "mistake" in statistics. Both the outcomes of measurements and the measurement procedure are inherently subject to variation.

KEYWORDS:

Error, Measurement, Sampling, Statistics.

INTRODUCTION

During the production process, it is hoped to provide a product of good quality at a reasonable price. Plans for sampling have been used frequently to inspect finished or undetermined goods. Every inspection method is built on a carefully thought-out sample strategy, which has been used extensively in the sector for lot punishment. Any inspection system has the potential to result in the acceptance of a nonconforming unit or the rejection of a confirming unit. According to the authors of this inspection mistake should be calculated and, if it is significant, an industrialist should take corrective action. An inspection error sampling plan was created by the authors. The association between inspection mistake and lot sentencing was the focus of the authors in states that "the requirement that the measurement of a single item not exceed some specified limit is sometimes more important than the requirement that the mean and variability for the items be at or near some predetermined value." In order to determine if the measurement of quality interest exceeds the specified specification limits or not, it is crucial to use the acceptance sampling strategy. As the cost of inspection is strongly correlated with sample size, the sample selection process is crucial during the production process. Based on sample data, a lot's fate is determined; an accepted lot will be delivered to the marker, whilst failed lots are recertified or have inferior products substituted. To explore inspection error, several writers developed sampling designs for various features, for instance.

In actual practice, the experimenters frequently lack certainty regarding the percentage of defective product. A strategy known as fuzzy sampling plans can be used in this situation for product inspection and lot sentencing. The industry has employed fuzzy sample plans extensively for a variety of situations. This topic was covered by a number of authors, such the current measurement-error prone sample designs are created using traditional statistics. The measurement method is predicated on precision, or clear, precise, and sharp observations. stated that "all measurements and observations of continuous variables are more or less nonpractice numbers rather than precise numbers. Unlike variability and mistakes, this imprecision is different. As a result, lifetime data are also less precise and more ambiguous than precise figures. 'Nonpractice numbers' are the best modern mathematical model for this error. When the observations or plan parameters are vague or uncertain, one can use eutrophic statistics. The eutrophic statistics were first introduced by authors, and they were then used. Aslam recently added neuromorphic statistics to the acceptance sampling plan field. Reference suggested a sample plan that took into account neuromorphic process loss and had ambiguous plan parameters. You may get more information regarding the neuromorphic sampling plan.

According to our research of the literature, no work has been done on developing a sampling plan with measurement error using neuromorphic statistics. In this publication, a brand-new sampling strategy for measurement error is developed utilizing neuromorphic statistics. The sample size and acceptance number are two neuromorphic criteria in the suggested sampling scheme. Additionally included is the neutrosophic operational function. The neuromorphic optimization problem will be used to calculate the neutrosophic plan parameters. For a few supplied parameters, some tables are provided. To serve as an example, a real-world situation is used.There are two types of measurement errors: random and systematic. Random mistakes are measurement errors that cause measurable values to differ when a constant attribute or quantity is measured repeatedly. Errors that are not caused by chance but rather by repetitive, system-inherent processes are known as systematic errors. Systematic error can also refer to an error that has a non-zero mean and whose impact is maintained after averaging the data.

Accuracy and precision are two metrics that can be used to characterize measurement mistakes. Measurement uncertainty should not be mixed up with measurement error.Statistical bias is another name for systematic error. Using established methods can frequently help to eliminate it. Learning how to use standardized tools and procedures to reduce systematic mistake is a part of the learning process in the various disciplines. Factors that cannot or will not be controlled are what cause random error or random variation. It might be prohibitively expensive to regulate them each time the experiment is run or the measurements are taken, which is one possible justification for not doing so. Other explanations include the fact that the object we are attempting to measure is dynamic see dynamic models or probabilistic in nature. When instruments are stretched to the limit, random mistake frequently happens. For instance, random mistake in the least significant digit of digital balances is typical. A single object's three measurements might look like this: The two types of measurement errors are systematic error and random error.

DISCUSSION

A measurement will always contain random error. It results from naturally unpredictable oscillations in measurement device readings or in the experimenter's interpretation of the instrument reading. Different findings for what appears to be the same repeated measurement are the consequence of random mistakes. They can be lowered by averaging several measurements and calculated by comparing several measurements. Systematic mistake is predictable, usually proportional to the real value, and constant. The systematic error can

typically be eradicated if its root cause is found. Systematic mistakes always have a predictable impact on experiment outcomes and are brought about by inaccurate measuring device calibration, inaccurate observational techniques, or intervention from the environment. An example of a systematic instrumentation error is incorrect instrument zeroing. Uneven measuring instrument calibration (zero error), environmental changes that interfere with the measurement process, and occasionally unreliable observational techniques can all be sources of systematic error. If an experimenter were to use a stopwatch or timer to measure the length of time it took for a pendulum to swing past a fiducial marker, all of their results would be inaccurate by one second.

The calculated average of the data will contain a % error if the researcher runs this experiment 20 times, starting at 1 second each time; the end result will be a little longer than the actual timeframe. If the minor slowing of the waves in the air is not taken into account, distances reported by radar will be consistently overstated. An example of a systematic instrumentation error is incorrect instrument zeroing. The outcome of an estimate based on a mathematical model or physical law may also contain systematic mistakes. For instance, if minute movement of the support is not taken into account, the estimated oscillation frequency of a pendulum will be consistently.

The proposed plan's design

A design is a precise and, most often, comprehensive notion for an item, a procedure, or a system. Design, though occasionally used to refer to something's nature, describes something that was or was purposely made by a thinking agent. The action of creating a design is expressed by the verb design. In some circumstances, the direct building of an objectas in some artwork and craftworkwithout an explicit prior plan, may also be seen as a design. The design typically has to adhere to a set of objectives and restrictions, as well as possible aesthetic, functional, economic, and socio-political factors, and is anticipated to interact with a specific context. Circuit diagrams, sewing patterns, architectural and technical drawings, and less tangible artifacts like business process models are typical examples of designs. Designers are those who create designs. A professional in one of the many design fields is often referred to as a "designer." The term "designer" within a profession is typically defined by the field of expertise (for example, one may be a fashion designer, product designer, web designer, or interior designer), but it can also denote others such as architects and engineers (see below: Types of designing). The series of steps taken by a designer to create a design, sometimes including design thinking and design techniques, is referred to as the design process [1]–[3].

A design can be created in a short amount of time (a rapid sketch), or it might take a long time and be quite complex, requiring extensive study, negotiation, contemplation, modeling, interactive adjustment, and re-design. Aside from those who are formally recognized as designers, many other people engage in the practice of designing. Interdisciplinary scientist Herbert A. Simon stated that "Everyone designs who devises courses of action aimed at changing existing situations into preferred ones" in his landmark book The Sciences of the Artificial. In addition, everyone has some degree of design skill because it is ingrained in our brains as a natural cognitive function, says design researcher Nigel Cross.

Comparison Research

Cross-cultural or comparative studies, which try to compare across various nations or cultures, are good examples of comparative research in the social sciences. The fact that data sets from other nations may not utilize the same categories or may define categories differently (for instance, by using different definitions of poverty is a significant issue in

comparative research. The act of comparing two or more items with the goal of learning more about one or all of the objects being compared is known as comparative research. This method frequently combines several fields in a single study. The general consensus about technique is that there is no methodology unique to comparative research. The flexibility provided by the multidisciplinary approach is beneficial, but comparative programs do have a defense to make when it is claimed that their research lacks a "seamless whole [4]–[6].

However, there are some techniques that are used in comparative research much more frequently than others. The bulk of comparison studies that use quantitative data demonstrate that quantitative analysis is far more frequently conducted than qualitative analysis. Comparative study follows the same general comparison process as our everyday comparison activities. Similar cases are handled similarly, and distinct cases are handled differently depending on how dissimilar they are from one another. If two carries can be sufficiently distinguished from one another, the research's findings won't be particularly useful. In comparative research, secondary analysis of quantitative data is relatively common, undoubtedly in part due to the expense of gathering primary data for such broad topics as a country's policy environment. In general, this study is an analysis of aggregate data. Large amounts of data are frequently compared, especially data from the government. Comparing welfare states often involves balancing the amount of money spent on social welfare.

Comparative study does not frequently look into "grand theories," such as Marxism, in keeping with how much theorizing has developed over the past century. Instead, it focuses on moderate hypotheses that only attempt to describe a portion of our social system as a whole. A notable illustration of this is the widespread study strategy that seeks to identify differences between two or more social systems before examining if those differences are related to any other characteristic that coexists in those cultures. The study on social welfare systems by Esping-Andersen is a prime example of this. He observed that several social welfare system types existed and contrasted them according to the extent to which social welfare products were recommodified. Based on the degree of decommodification, he discovered that he could divide welfare states into three categories. From this, he further hypothesized that decommodification was driven by a mix of class coalitions, mobilization, and regime legacies. Esping-Andersen is employing comparative research in this case; he evaluates the degree of decommodification in a number of western nations before formulating a theory of the divergence based on his conclusions.

A comparative study may take many different shapes. Space and time are two important aspects. Cross-national comparisons are by far the most common in terms of geography, but within-country comparisons that contrast various regions, cultures, or governments also exist and are very beneficial, particularly in a country like New Zealand where policies frequently change depending on which race they apply to. Comparing comparable or dissimilar countries or groups of countries, as well as one's own country to others or the entire world, are common interregional studies.

The historical comparative research compares several time periods. The two basic options in this model are simply comparing the same thing over time to evaluate if a policy's impacts change over time, or comparing two points in time either snapshots or time-series Many claim that there is no particular topic area for comparative inquiries. This may be the case, however a quick look at comparative studies shows that some topics are discussed more frequently than others. It's a common question to ask if socioeconomic or political variables influence government behavior more. However, the existence of differences to be analyzed is generally the only aspect of comparative research concerns that can be relied upon.

Development

Cross-cultural and comparative study, according to Mousie's, ought to be viewed as a component of the scientific mindset that emerged in Greece in the sixth century and the general appreciation of learning and knowledge that characterized the fifth century. In other words, it is a component in the development of episteme and philosophic, which is a love of learning that is unrelated to monetary gain. As a form and activity in the area of logos, episteme signaled the end of cognitive closure and advanced rational debate, empirical inquiry, and the pursuit of truth. The high regard for intellectual pursuits also sparked a sincere interest in other civilizations, which has since been at the core of comparative study.

Additionally, the philosophical and political questioning that characterized the life of the democratic polis lay behind the Greek comparative gaze. From the Milesians to the Sophists, philosophical inquiry challenged the perceptions and cognitive traditions of their own people. Herodotus' Histories show that philosophical inquiry into other people's traditions was a practice connected to the ethos of philosophical critique that characterized democratic life in Greece. Similar to this, the first historians' attempt to consider their own institutions while studying those of others is connected to the questioning of Greek laws, institutions, and the values and practices that go along with them (such as signoria and parrhesia), as part of Greek politics. Also according to Karl Deutsch, we have been conducting investigations in this manner for more than 2,000 years. Basic scientific and philosophical investigation, which has been done for a very long period, depends on comparisons of things. In estimating how long comparative research has been around, most authors are more conservative. With many arguing that comparing objects does not qualify as comparative study, the discussion over the definition of the tradition is essentially academic in nature [7]–[9][10].

By the 1880s, textbooks on this field of study were starting to appear, but it wasn't until World War II that they really started to gain ground. There are many reasons why social scientists now consider comparative research to be one of their most important tools. The desire and possibility for educational exchanges as well as intellectual curiosity about different cultures have increased significantly as a result of globalization. Greater quantitative data creation for comparison has been made possible by information technology, and global communications technology has made it easier to disseminate this information.

CONCLUSION

The goal of comparative cultural studies is to study culture and culture products, including but not limited to literature, communication, media, art, and so on. Selected tenets of comparative literature are combined with selected tenets of the field of cultural studies, including culture theories, (radical) constructivism, communication theories, and systems theories. This is done within a contextual and relational framework using a variety of interdisciplinary methods and approaches, as well as teamwork where necessary. The main goals of research and study in comparative cultural studies are the processes of communicative action(s) in culture and the how of these processes. However, textual analysis appropriate to other well-established disciplines of study is not excluded from comparative cultural studies scholarship. The framework and tools available in the systematic and empirical study of culture are ideally preferred in comparative cultural studies. A fundamental and defining component of the framework and its application is the attention to other cultures in opposition to essentialist notions and practices and beyond the paradigm of the nation-state. This postulate is included in theoretical, methodological, and applied research in comparative cultural studies.

REFERENCES

[1] G. Gavriilidis, N. N. ataraja. Gavriilidou, E. Pettersson, E. Renhammar, A. Balkfors, and P. O. Östergren, "Empowerment evaluation of a Swedish gender equity plan," *Glob. Health Action*, 2014, doi: 10.3402/gha.v7.23710.

[2] M. D. M. García-Calvente, E. Castaño-López, I. Mateo-Rodríguez, G. Maroto-Navarro, and M. T. Ruiz-Cantero, "A tool to analyse gender mainstreaming and care-giving models in support plans for informal care: Case studies in Andalusia and the United Kingdom," *J. Epidemiol. Community Health*, 2007, doi: 10.1136/jech.2007.060665.

[3] J. Zhang, H. Y. Li, Y. Z. Luo, and G. J. Tang, "The orbital reconfiguring design and analysis after maneuver failure for rendezvous phasing," *Guofang Keji Daxue Xuebao/Journal Natl. Univ. Def. Technol.*, 2009.

[4] B. P. Buunk and T. Mussweiler, "New directions in social comparison research," *Eur. J. Soc. Psychol.*, 2001, doi: 10.1002/ejsp.77.

[5] C. Feng, M. Ding, and B. Sun, "A comparison research on industry-university-research strategic alliances in countries," *Asian Soc. Sci.*, 2011, doi: 10.5539/ass.v7n1p102.

[6] K. Bruchmann, "Compared to What? The Importance of Control Groups in Social Comparison Research," *Basic Appl. Soc. Psych.*, 2017, doi: 10.1080/01973533.2017.1281808.

[7] J. Gupta and C. Vegelin, "Sustainable development goals and inclusive development," *Int. Environ. Agreements Polit. Law Econ.*, 2016, doi: 10.1007/s10784-016-9323-z.

[8] Y. Zheng, M. Hatakka, S. Sahay, and A. Andersson, "Conceptualizing development in information and communication technology for development (ICT4D)," *Information Technology for Development*. 2018. doi: 10.1080/02681102.2017.1396020.

[9] J. O'Flaherty and M. Liddy, "The impact of development education and education for sustainable development interventions: a synthesis of the research," *Environ. Educ. Res.*, 2018, doi: 10.1080/13504622.2017.1392484.

[10] V. Gaete, "Adolescent psychosocial development," *Revista Chilena de Pediatria*. 2015. doi: 10.1016/j.rchipe.2015.07.005.

CHAPTER 6
STATISTICS AND DOMINATING DAVID DERIVED NETWORKS ENTROPY CALCULATION

Shambhu Bharadwaj, Associate Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id- shambhu.bharadwaj@gmail.com

ABSTRACT:

Information entropy, often known as Shannon entropy, was first established by Claude Shannon in his 1948 paper "A Mathematical Theory of Communication. A data source, a communication channel, and a receiver make up the three components of a data communication system according to Shannon's theory. According to Shannon, the "fundamental problem of communication" is the receiver's ability to determine which data was produced by the source based on the signal it receives across the channel. In his renowned source coding theory, Shannon demonstrated that the entropy reflects an absolute mathematical limit on how well data from the source can be listlessly compressed onto a perfectly noiseless channel. Shannon studied several techniques to encode, compress, and transmit messages from a data source. In his noisy-channel coding theorem, Shannon greatly improved this conclusion for noisy channels.

KEYWORDS:

Analogy, Communication, Machine Learning, Mathematics.

INTRODUCTION

Entropy in statistical thermodynamics and information theory are directly comparable. The analogy occurs when the values of the random variable identify the energies of microstates, and as a result, the Gibbs formula for entropy and Shannon's formula are formal equivalents. Other branches of mathematics, such combinatorics and machine learning, are relevant to entropy. A set of axioms that state that entropy should be a measurement of how informative the average result of a variable is can be used to construct the definition. Differential entropy is equivalent to entropy for a continuous random variable[1], [2]. Statistics that descriptive statistics are most frequently concerned with are central tendency and dispersion. Central tendency seeks to characterize the distribution's central or typical value, while dispersion characterizes the degree to which members of the distribution deviate from it and from one another. Under the framework of probability theory, which deals with the analysis of random processes, inferences on mathematical statistics are produced [3]–[5].

Data collection leading to a test of the link between two statistical data sets, or a data set and synthetic data derived from an idealized model, is a common statistical method. An alternative to the idealized null hypothesis, which states that there is no association between the two data sets, is a hypothesis that is put forth regarding the statistical relationship between the two data sets. Statistical tests that quantify the extent to which the null can be shown wrong, given the data provided in the test, are used to reject or disprove the null hypothesis. Working from a null hypothesis, two main types of mistakes are recognized: Type I errors, which result in a "false positive" when the null hypothesis is incorrectly rejected, and Type II errors, which result in a "false negative" when the null hypothesis is correctly rejected but an actual relationship between the populations is missed. This paradigm has been linked to a

number of issues, including difficulty in acquiring a big enough sample size and difficulty in defining a good enough null hypothesis.

DISCUSSION

The data that are produced by statistical measuring procedures are likewise subject to error. Many of these errors are categorized as random or systematic, but there are other types of errors that might happen, such gaffe, like when an analyst reports the wrong units. Biased estimates may be caused by missing data or censoring, and to deal with these issues, particular methodologies have been created. Statistics is a field of mathematics or a body of mathematics that deals with the gathering, analysis, interpretation, and presentation of data Instead of being a subfield of mathematics, some people view statistics as a separate mathematical discipline. While data are used in many scientific projects, statistics is concerned with how data are used when there is ambiguity and how to make decisions when there is doubt. It is customary to begin with the population or process to be examined when applying statistics to a problem. Populations can refer to a variety of things, such as "every person living in a nation" or "each atom making up a crystal." Ideally, statisticians conduct a census to gather data on the entire population. Governmental statistical institutes may organize this. The demographic data can be summarized using descriptive statistics.

For continuous data, numerical descriptors like mean and standard deviation are useful, whereas frequency and percentage are better at defining categorical data. When a census is not possible, a sample, or selected subset of the population, is examined. Data on the sample participants are gathered in an observational or experimental environment once a sample that is representative of the population has been identified. The sample data can once more be summarized using descriptive statistics. However, there is a random element in the sample selection process, thus the sample's numerical descriptors are also subject to uncertainty. Inferential statistics are required to make inferences about the entire population. While controlling for randomness, it makes assumptions about the population represented by the sample by using patterns in the data. These conclusions can be drawn by answering affirmative or negative questions about the data hypothesis testing, estimating their numerical properties, identifying relationships within the data, and modeling those relationships. Inference can also be used to anticipate, predict, and estimate unobserved variables that are part of or connected to the population under study. Data mining, as well as extrapolation and interpolation of time series or spatial data, might be included.

Topological Indices Based on Degree

A topological index, also called a connectedness index, is a form of molecular descriptor that is derived based on the molecular graph of a chemical molecule in the fields of chemical graph theory, molecular topology, and mathematical chemistry. Topological indices are numerical variables that describe the topology of a graph and are typically graph invariant. The discovery of quantitative structure-activity relationships, in which the biological activity or other features of molecules are connected with their chemical structure, is one application of topological indices. Hydrogen-suppressed molecular networks, in which the atoms are represented by vertices and the bonds by edges, are the source of topological descriptors. Different types of topological matrices, which may be mathematically manipulated to obtain a single integer, typically known as graph invariant, graph-theoretical index or topological index, can be used to characterize the connections between the atoms. As a result, the topological index can be thought of as two-dimensional descriptors that are simple to calculate from molecular graphs, do not depend on how the graph is represented or labeled, and do not require the chemical structure to be minimized in energy. The simplest topological

indices are defined for linked undirected molecular graphs only and disregard hydrogen atoms, double bonds, and various atom kinds. The hybridization state of each atom in the molecule is taken into account by more complex topological indices. The Hosoya index, often known as "the" topological index, was the first topological index to be recognized in chemical graph theory. The Wiener index, Randi's molecular connection index, Balaban's J index, and the TAU descriptors are more examples. Based on the improvement of TAU descriptors, the extended topochemical atom indices have been created [6]–[8].

Degree-Based Graph Entropy

Entropy is a measurable physical characteristic and a scientific notion that is frequently connected to a condition of disorder, unpredictability, or uncertainty. From classical thermodynamics, where it was originally recognized, through the microscopic description of nature in statistical physics, to the fundamentals of information theory, the phrase and concept are utilized in a variety of disciplines. It has numerous applications in physics and chemistry, biological systems and how they relate to life, cosmology, economics, sociology, weather science, and information systems, especially the exchange of information. William Rankine, a Scottish scientist and engineer, first used the terms' thermodynamic function and heat potential to describe the thermodynamic notion in 1850. The definition of thermodynamics given in 1865 by German physicist Rudolf Clausius, one of the principal pioneers of the discipline, is the ratio of an infinitesimal amount of heat to the immediate temperature. He first referred to it as transformation-content, or Verwandlungsinhalt, in German, and then adopted the Greek word for transformation to create the term entropy. Clausius defined the term as "desegregation" in 1862 by referring to microscopic constitution and structure. The second law of thermodynamics, which asserts that an isolated system's entropy cannot decrease over time if left to spontaneous evolution, is fundamentally based on entropy. So, isolated systems go toward thermodynamic equilibrium, which has the largest entropy. Because of the second rule of thermodynamics, some processes cannot be reversed [9]–[11].

Ludwig Boltzmann, an Austrian physicist, defined entropy as the count of microscopic configurations or states of individual atoms and molecules of a system that satisfy the macroscopic condition of the system. By doing so, he created a new branch of thermodynamics known as statistical mechanics and introduced the ideas of statistical disorder and probability distributions. He also discovered the connection between microscopic interactions, which fluctuate around an average configuration, and macroscopically observable behavior in the form of a straightforward logarithmic law with a proportionality constant, the Boltzmann constant, which has since become one of the defining universal constants for the modern Interna Similar statistical techniques for quantifying microscopic uncertainty and multiplicity were created by Bell Labs scientist Claude Shannon in 1948 to address the issue of random information losses in telecommunication communications. Shannon created the area of information theory by naming this missing information object entropy, which is similar to how it is used in statistical mechanics, on the advice of John von Neumann. The definition of entropy that is accepted worldwide fits this description.

Entropy of a Graph Based on Edge Weight

A graph is a structure that amounts to a set of items where some pairs of the objects are in some manner "related" in discrete mathematics, more specifically in graph theory. The items are represented by mathematical abstractions known as vertices, and each pair of connected vertices is known as an edge also known as a link or a line. A graph is typically shown

diagrammatically as a collection of dots or circles representing the vertices and lines or curves representing the edges. One of the topics studied in discrete mathematics is graphs. Both directed and undirected edges are possible. For instance, if the edges between two individuals are handshakes, then the graph is undirected since any individual A can only shake hands with an individual B if B also shakes hands with A. The graph is directed, however, if an edge from person A to person B indicates that A owes money to B because borrowing money is not always returned. The fundamental topic of graph theory is graphs. Due to a direct connection between mathematics and chemical structure, J. J. Sylvester first used the word "graph" in this sense in 1878. Edges are a type of edge in a graph and are used to distinguish it from a directed graph. The endpoints of an edge are the vertices x and y of an edge x, y. The edge is described as incident on x and y and as joining x and y. A vertex may not be linked to any other vertex if it belongs to no edge. A generalization that enables several edges to share the same pair of endpoints is a multigraph. Multigraphs are sometimes simply referred to as graphs in writings [12]–[14].

The edges that connect a vertex to itself are known as loops, and they are occasionally permitted in graphs. The pairs of vertices in E must be permitted to share the same node more than once in order to support loops. When it is obvious from the context that loops are permitted, such generalized graphs are referred to as graphs with loops or just graphs. The set of edges must also be finite because the set of vertices V is typically assumed to be finite. Although infinite graphs are occasionally taken into consideration, they are typically seen as a particular form of binary relation since most conclusions about finite graphs do not apply to infinite graphs or require a very different kind of justification.

Any graph with an empty set of vertices and, consequently, an empty set of edges is said to be empty. A graph's order is determined by its vertex count, or |V|. A graph's |E| number of edges measures its size. A non-empty graph could have a size of 0, however in some situations, like when defining the computational complexity of algorithms, the size is |V| + |E|. The number of edges that are incident to a vertex determines its degree or valency; for graphs with loops, an edge is counted twice for each loop[15]–[17].

CONCLUSION

When data compression is used, information theory can be used to determine how little information is needed to transmit a message. Take the transmission of sequences made up of the four letters over a binary channel, for instance. If the likelihood of each of the four letters is equal one cannot do better than to encode each letter with two bits. A might be coded. However, one might use variable length codes if the probabilities of each letter are different, for example, if "A" occurs with a 70% likelihood, "B" with a 26% probability, and "C" and "D" each with a 2% probability.

In this instance, the letters "A" would be classified According to this illustration, only one bit is required to be communicated 70% of the time, two bits only 26% of the time, and three bits only 4% of the time. Due to the low entropy, less than 2 bits are typically needed. This effect is measured and captured by computing the sum of probability-weighted log probabilities. When seen as a string of characters, English text is relatively predictable because of its low entropy. For instance, we may reasonably predict that the combination 'e' will be significantly more common than 'z', the combination 'Qu' will be significantly more common than any other combination including a 'q', and the combination 'the' will be significantly more common than 'z', 'q', or 'Qu'. One can frequently guess the rest of the term after the first few letters. Between 0.6 and 1.3 bits of entropy are included in each character of English text.

REFERENCES

[1] G. Menichetti, G. Bianconi, G. Castellani, E. Giampieri, and D. Remondini, "Multiscale characterization of ageing and cancer progression by a novel network entropy measure," *Mol. Biosyst.*, 2015, doi: 10.1039/c5mb00143a.

[2] P. Bereziński, B. Jasiul, and M. Szpyrka, "An entropy-based network anomaly detection method," *Entropy*, 2015, doi: 10.3390/e17042367.

[3] Z. Chen, M. Dehmer, and Y. Shi, "Bounds for degree-based network entropies," *Appl. Math. Comput.*, 2015, doi: 10.1016/j.amc.2015.06.003.

[4] J. Sun, D. Taylor, and E. M. Bollt, "Causal network inference by optimal causation entropy," *SIAM J. Appl. Dyn. Syst.*, 2015, doi: 10.1137/140956166.

[5] B. Sen Chen, S. W. Wong, and C. W. Li, "On the calculation of system entropy in nonlinear stochastic biological networks," *Entropy*, 2015, doi: 10.3390/e17106801.

[6] C. Betancur, R. Cruz, and J. Rada, "Vertex-degree-based topological indices over starlike trees," *Discret. Appl. Math.*, 2015, doi: 10.1016/j.dam.2014.12.021.

[7] I. Gutman, "Degree-based topological indices," *Croat. Chem. Acta*, 2013, doi: 10.5562/cca2294.

[8] R. Khalid and N. Idrees, "Degree based topological indices of Hanoi graph," *arXiv Prepr. arXiv1804.09431. 2018 Apr 25*, 2018.

[9] G. Lu, "Generalized degree-based graph entropies," *Symmetry* ., 2017, doi: 10.3390/sym9030029.

[10] G. Lu, B. Li, and L. Wang, "Some new properties for degree-based graph entropies," *Entropy*, 2015, doi: 10.3390/e17127871.

[11] S. Cao, M. Dehmer, and Y. Shi, "Extremality of degree-based graph entropies," *Inf. Sci.* ., 2014, doi: 10.1016/j.ins.2014.03.133.

[12] J. Gomez-Pilar *et al.*, "Quantification of Graph Complexity Based on the Edge Weight Distribution Balance: Application to Brain Networks," *Int. J. Neural Syst.*, 2018, doi: 10.1142/S0129065717500320.

[13] S. T. Jawaid and S. L. Smith, "Informative path planning as a maximum traveling salesman problem with submodular rewards," *Discret. Appl. Math.*, 2015, doi: 10.1016/j.dam.2015.01.004.

[14] L. Zhang, L. Qiao, and S. Chen, "Graph-optimized locality preserving projections," *Pattern Recognit.*, 2010, doi: 10.1016/j.patcog.2009.12.022.

[15] V. Adiyasuren, T. Batbold, and M. A. Khan, "Refined arithmetic-geometric mean inequality and new entropy upper bound," *Commun. Korean Math. Soc.*, 2016, doi: 10.4134/CKMS.2016.31.1.095.

[16] A. K. Singh, H. P. Singh, and Karmeshu, "Analysis of finite buffer queue: Maximum entropy probability distribution with shifted fractional geometric and arithmetic means," *IEEE Commun. Lett.*, 2015, doi: 10.1109/LCOMM.2014.2377236.

[17] K. P. Nelson, "Assessing probabilistic inference by Comparing the generalized mean of the model and source probabilities," *Entropy*, 2017, doi: 10.3390/e19060286.

CHAPTER 7
EMPIRICAL TEST BASED ON EPS GLOBAL STATISTICS: INNOVATION OR IMITATION

Ajay Rastogi, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id-ajayrahi@gmail.com

ABSTRACT:

Technology advancement is a key example of innovation-driven development and the primary factor in long-term economic growth. The introduction of technology, learning, imitation, and independent research and development are all examples of technological advancements. Technology transfer to underdeveloped nations has the potential to accelerate technological advancement, close the technology gap, and foster quick and sustainable economic growth. If the technological divide between developed and developing nations is indeed narrowing over time, then the ongoing adoption of new technologies in wealthy nations may easily result in "path dependence" and technical bottlenecks. Additionally, this trend could not be good for the preservation of intellectual property rights. There is an internal Nash equilibrium point between the R & D strategy and the imitation strategy because of the externality of knowledge and the cost of imitation. In this article, we optimize the enterprise's investment path using population dynamics, the replicon dynamics of R & D and imitation, as well as equilibrium circumstances. The ratio of the output elasticity of the R & D firm determines the balance between the number of innovators and copycats. The pace of technological advancement and the rate of capital growth were found to increase proportionally as a result of the presence of this balance. This conclusion was supported by our empirical investigation, which also revealed that stability is impacted by the factor's international flows. Therefore, it was discovered that innovation-driven technical advancement is a key factor in determining long-term economic growth.

KEYWORDS:

Advancement, Development, Investigating, Stability.

INTRODUCTION

Prior to 2008, China's rapid economic expansion was mostly driven by its low cost of production and export. China's marginal labor and capital output rates kept dropping after the 2008 global financial crisis, which slowed down economic development. In this case, realizing the shift from traditional factors and investment-driven to innovation-driven growth is important to ensure sustainable economic growth. Technology advancement is a key example of innovation-driven development and a key factor in the stability and speed of economic growth. Technology introduction, education, and independent research and development all contribute to technological advancement. Technology introduction can speed up technological advancement and encourage sustained and rapid economic development if there is a significant technological gap between developing and industrialized nations. The steady closing of the technological gap between the two will make it easy for the so-called "path dependence" and technology bottleneck to arise, which is detrimental to the preservation of intellectual property rights.

The only source of long-term economic growth, according to empirical study and practical observation analysis, is internal technological advancement. This has been confirmed and examined using various models that economists have built. For instance, using particular empirical data, Tang Yong and Fan Xin conducted a quantitative examination of the mechanism and impact of technology advancement on economic growth. The findings indicated that the economic growth rate would improve by 0.347% for every 1% increase in technological advancement. Technological advancement has a substantially bigger impact on economic growth than constant capital, variable capital, and surplus value ratio. By building a structural equation model, Ma Weiwei remeasured the extent to which China's technology advancement contributed to economic growth. The analysis's findings indicate that during the past 20 years, China's economic growth has largely followed a trend of "fluctuating upward" and that supply capacity for technology has been steadily increased. It serves as a key gene for China's long-term and steady economic growth in the new historical era.

Li used the geographical weighted regression model (GWR) to examine how technological innovation affects regional economic development. The findings indicate that the growth of technical innovation has a sizable beneficial effect on the local economy. Local governments must build local innovation development strategies and regulations, enhance the quality of patents and the pathways for transformation, boost investments in innovation research and development, and promote the level of economic openness. According to Liu et al. independent innovation is the primary means of advancing the performance of green growth China among the many means of technical innovation. It is challenging for technology import from technology trade channels to support green growth performance, while imports of technology from foreign capital and import trade channels can both contribute to green growth performance. In order for businesses to fully benefit from the favorable impact of technological innovation on the performance of green growth, they must be encouraged to improve independent innovation and the external introduction of green technology.

According to Zhang et al.'s theory technological innovation may be broadly categorized into independent innovation and introduction of technology. The author develops a local equilibrium model based on the ABBGH (2005) model in order to demonstrate that foreign technology import has a lagging promoting effect on the independent innovation activities of local enterprises, and that the promoting effect only happens in the enterprises with relatively lagging innovation ability and some LL industries. Meng and Zhang built an econometric model based on the traits of sample data and theoretical logic and examined the effects of natural resource endowment on innovation, which is the mode of technological advancement, as well as on the introduction of new technologies. While the adoption of technology does not help to increase regional green growth efficiency, innovation does. According to Yang the introduction of technology and independent research and development are the two primary means of technical innovation to support economic progress. The technological gap cannot be closed if just replication of existing technologies can be accomplished without independent research and development innovation. Based on threshold model analysis and comparison, China's economic growth from 2001 to 2015 was primarily driven by the import of technology. When there is inadequate intellectual property protection, technology innovation in the field of market failure is serious. Even though the short-term promoting effect on economic growth is relatively obvious, it will be beneficial due to the lack of independent innovation for a long time.

Using factory-level data from Mexico, Tracy Collins examined the dynamic link between innovation and imitation at the microlevel. The empirical findings demonstrate that, in the context of developing nations, creativity and imitation complement one another. In the UK,

Xia TJ and Liu XH examined longitudinal firm-level data and discovered an inverted U-shaped association between imitation and local firm innovation performance as well as a mediating role for imitation in the relationship between international competition and local firm innovation performance. At the moderate imitation level, the absorptive capacity diminishes the return on invention and moderates the mediating effect of imitation. Lina and Yu experimentally evaluate the effects of several small technical innovations on regional economic development using a spatial econometric model. The empirical findings reveal a strong spatial relationship between technology advancement and local economic growth. Additionally, the introduction of technology has a good impact on regional economic development, as do imitation innovation, independent innovation, and independent innovation.

The impact of new technology and imitation innovation on the expansion of economic development gradually slows down as the industrial structure is upgraded. Lin et al. used intermediary effect and dynamic threshold effect models to examine the point of change of the ICT capital structure and its influence mechanism to support high-quality economic development. Optimizing the course of technological advancement, particularly through the transformation of copycat innovation into independent innovation, is the key to striking a balance between high-quality economic development and consistent growth. From the viewpoint of an original innovation value network, Liang and Yu explain how a country or region chooses the path of high-quality industrial development. According to the analysis, the input-output ratios of the various industries driven by different innovation modes differ significantly from one another. The input-output efficiency of industrial innovation driven by original innovation is also clearly superior to that of other industries based on imitation-based innovation development.

In actuality, imitation is expensive. According to Acemoglu and Gigliotti's theory of technology direction, emerging countries' comparative advantage will be harmed by their technological mimicry of developed nations. The overall profit area will become more smaller as a result of a large number of imitation activities. According to this theory, the cost of imitation rises as the number of imitations increases. Therefore, it is evident from both the analysis of the technological direction model and the intellectual property protection model that neither comprehensive R&D nor complete imitation would occur in any given nation or region. Even industrialized nations with robust intellectual property rights protection tend to prioritize new products, quality, and testing above soft technologies like innovative production methods and cutting-edge business strategies. According to the global economic model, developing nations can adopt the appropriate technological techniques based on the current scenario.

We are aware that free-riding imitation results from knowledge externalities. Companies will cut internal R & D spending if they use the output-input ratio as their optimization goal function. Additionally, if every business takes an imitation strategy, market rivalry will force them to raise the caliber of their products. As a result, the R&D strategy and the imitation strategy are in a stable inner point Nash equilibrium. It's critical to understand the existence of this balance to understand why many developing nations have seen stable growth. In this article, research is approached from a broad technical standpoint, where creation and experimentation with soft technologies like new procedures and business models are just as important as improving new goods and their quality. As a result, the model does not take into account the impact of intellectual property protection. Instead, it optimizes the investment path of businesses in a closed economic system to obtain the overall dynamics and replicator dynamics of R & D and imitation, and it studies the equilibrium path through the change in

the number of R & D and imitators. In light of the analysis above, we chose data from 54 nations from 1976 to 2013 as our sample. We draw the conclusion that there is an evolutionary balance between R & D and imitation only when the returns to scale of the R & D group that determines the technological frontier are increasing, and on this basis, the solution and explanation to practical problems are provided.

DISCUSSION

The "Babel Experiment," created by Vladimir L. Pavlov in 2001 as a training tool for software engineering students, is where the P-Modeling Framework got its start. Its goal was to give students experience using UML to solve "condensed" versions of communication issues that are common in software development. The following method was used to conduct this experiment. The job of creating a software system with the following restriction factor was given to a group of students: The sole language that could be used to communicate while working on the project had to be UML. The idea was to expose students to a "condensed" version of the communication issues that arise frequently in software development and give them practice using UML to solve these issues. Following this exercise, students created models that were incredibly clear and succinct [1]–[3].

Later, during a design session, the identical assignment was being worked on by two separate teams. The first team's means of communication were limited to UML as previously indicated, whereas the other team was free to speak in a language other than UML. It ended up being the case that the first, more constrained crew completed the work more effectively than the second. The first team's UML diagrams were better constructed, more thorough, understandable, and elaborated. The purpose of these subsequent studies by Vladimir L. Pavlov was to determine whether the "silent" modeling sessions are more effective than the conventional ones. Silent teams did at least as well as the others in these experiments, and in some situations, they even exceeded the conventional ones.

Semantic Reverse Traceability

A quality assurance technique called reverse semantic traceability enables testing of each translation step's output. The present artifacts are "reverse engineered" and the restored text is compared to the original before moving on to the next stage. In the event that there is a discrepancy between these two texts, the tested artifacts are fixed to fix the issue (or the original text is fixed). As a result, each step is validated by taking a step back and ensuring that progress stays on course. Issues can be found and resolved right away in this fashion, preventing them from building up and cascading into later stages of the development cycle. The word "Semantic" in the method's name is crucial. Its foundation is the idea that original and restored copies of a text should be compared semantically, with an emphasis on the "meaning" of the text rather than on the "words" used in it.

Early users of the Reverse Semantic Traceability approach cited the following usage situations as the most common the original and restored textual descriptions of a domain are compared as part of the UML model validation process by quality engineers. Validating model changes for a new need entail restoring the textual description of the requirement given the original and modified versions of the model, then comparing the original and restored descriptions. When validating a bug fix, quality engineers restore a textual description of the bug that was fixed and compare the original and restored descriptions. Reverse Semantic Traceability for the important artifacts from the present projects is the assignment given to the new team member who is being integrated as a new software engineer [4]–[7].

Evolutionary stability conditions and technological advance rates

An evolutionarily stable strategy (ESS) is a tactic (or group of tactics) that, once accepted by a population in response to a particular environment, is impermeable and cannot be replaced by another tactic (or set of tactics), even if it is innovative or originally uncommon. A key idea in behavioral ecology, evolutionary psychology, mathematical game theory, and economics, with applications in anthropology, philosophy, and political science, it was first introduced by John Maynard Smith and George R. Price in 1972/3.

An ESS, or "evolutionarily stable system," is a game-theoretical term for a Nash equilibrium that has been refined and is also "evolutionarily stable." Therefore, once a strategy has been fixed in a population, natural selection alone can keep it there without the need for substitute (mutant) strategies (although this does not rule out the possibility that a better strategy, or set of strategies, will emerge in response to selective pressures brought on by environmental change. The classic concept for a solution in game theory is the Nash equilibrium. It depends on the players' mental faculties. Players are believed to be mindful of the game's structure and actively attempt to anticipate their opponents' moves in order to increase their own rewards. Additionally, it is assumed that everyone involved is aware of this (see common knowledge. These presumptions are then applied to explain why Nash equilibrium strategies are chosen by players.

Strategies that are evolutionary stable are driven very differently. It is assumed that the players' tactics are physiologically encoded and inherited in this situation. A person need not be aware of the game and has no influence over their tactics. The game's rewards indicate reproductive success biological fitness, and they reproduce while being susceptible to natural selection's forces. It is hypothesized that different game strategies occasionally appear through a process akin to mutation. A strategy must be resistant to these alternatives in order to qualify as an ESS. It may come as a surprise that Esses and Nash equilibria frequently coincide given the drastically different motivating assumptions. In actuality, each ESS is a Nash equilibrium, however not all Nash equilibria are Esses.

Nash balance

A Nash equilibrium is refined or altered to become an ESS. Examples contrasting the two are provided in the following section.) If everyone adopts their respective components, there is a Nash equilibrium where no player can gain from adopting a different tactic. It is a strategy pair in a two-player game. be the result of using strategy S versus strategy T. In a two-player game, the strategy pair (S, S) is a Nash equilibrium if and only if for both players, for any strategy T: The terms evolutionarily stable state (ESS) and evolutionarily stable strategy (ESS) have a tight relationship in population biology but refer to separate phenomena. If every member of a population adopts an evolutionary stable strategy, no mutant strategy can invade. There is no 'logical' alternative once almost the entire population has adopted this tactic. A component of traditional game theory is ESS.

If the disruption is not too severe, selection can restore the genetic makeup of a population in an evolutionarily stable condition. When a population is disrupted from its starting condition and returns to utilizing a strategyor a combination of strategiesit is said to be in an evolutionarily stable state. It is a component of evolutionary game theory, dynamical systems, or population genetics. Convergent stability is the current term for this. According to B. Thomas (1984), an individual strategy that is mixed may be referred to as an ESS, and an evolutionarily stable population state may refer to a population mixture of pure strategies that is formally comparable to the mixed ESS. Genetic diversity does not necessarily indicate how evolutionary stable a population is; it might also be genetically monomorphic [8]–[10].

Dilemma facing prisoners

The Prisoner's dilemma is a popular example of social cooperation and compassion. Here, playing Cooperate would benefit everyone more as a group, but since Defect performs better, each player has an incentive to play Defect. Making it possible for players to retaliate by playing the game against the same opponent repeatedly is one way to address this issue. The same two people repeatedly act out the prisoner's dilemma in the so-called iterated Prisoner's Dilemma. The iterated Prisoner's Dilemma has an enormous number of potential strategies, whereas the original Prisoner's Dilemma only has two (Cooperate and Defect). There may actually be an endless number of such contingency plans because each history can have a unique contingency plan and the game can be played an infinite number of times.

Always Defect, Always Cooperate, and Tit for Tat are three straightforward contingency strategies that have drawn a lot of attention. The latter responds on the following round by performing what was done to it on the previous roundit responds to Cooperate with Cooperate and Defect with Defect. The first two strategies carry out the identical activities regardless of the opposing player's actions. Tit-for-Tat will perform better than Always Defect if the whole population plays Tit-for-Tat and a mutation develops that plays Always Defect. The percentage of mutants will be maintained low if the mutant population grows too great. Tit for Tat is an ESS since it solely considers only two tactics. On the other hand, an island of Always Defect players will withstand an invasion by a small number of Tit-for-Tat players but not by a sizable one. A population of Tit-for-Tat is no longer an ESS if we include Always Cooperate. Because Tit-for-Tat players always cooperate, the Always Cooperate strategy performs the same in this population. A mutant that plays Always Cooperate won't be eliminated as a result. Even though a population of Always Cooperate and Tit-for-Tat individuals can coexist, the selective pressure is against Always Cooperate and in favor of Tit-for-Tat if a small portion of the population is Always Defect. This is because, in the event that the opponent defects, cooperation has smaller payoffs than defecting. This highlights the challenges in translating the official definition of an ESS to games with broad strategy spaces, which has led some to think about working around the problem [11], [12].

CONCLUSION

This study concludes that there is a stable internal point Nash equilibrium between the R&D strategy and the model strategy because imitation cost exists. The scale of compensation for research and development companies determines the ratio between the two. The proportionate link between the output elasticity of input and development and the output elasticity of production is what drives social technological advancement. The magnitude of returns for R & D-type enterprises will become increasingly obvious as society continues to improve technologically since the rate of technological development is directly proportional to the rate of capital growth. As a result, increasing amounts of investment are required to develop new technologies. The cost of imitation rises as a result of intellectual property rights protection, which also ensures the size of R & D-type businesses, allowing for a progressive development in the number of these businesses. Due to the balance between R&D and imitation, for developing nations with highly competitive labor and product markets, capital inflows will inevitably be the driving force behind domestic R&D operations. The article's primary results can be used to explain economic issues caused by endogenous growth. We respond to the query of steady growth first. Why, for instance, does a nation's GDP growth rate persist throughout time without changing? The foundation of a country's overall economic volume determines this, according to the conventional theory. The article's model demonstrates that technical advancement through consistent R&D and imitation efforts

determines this. The rate of technological advancement and capital growth increase proportionately to one another because of the balance between research and development and imitation. Second, we address a query regarding the middle-income trap: Why may certain nations pass through the middle-income stage but not others? This phenomenon can be explained by the equilibrium path and saddle point path of this article. The elastic coefficient, which is determined and impacted by the human capital variable, serves as the primary indicator to calculate the equilibrium path and the saddle point path even though it is not a factor in this study. Additionally, the index developed in this document is useful for assessing a country's potential for growth.

REFERENCES

[1] J. Chen, L. Sun, C. Guo, W. Wei, and Y. Xie, "A data-driven framework of typical treatment process extraction and evaluation," *J. Biomed. Inform.*, 2018, doi: 10.1016/j.jbi.2018.06.004.

[2] R. L. Shafer, K. M. Newell, M. H. Lewis, and J. W. Bodfish, "A cohesive framework for motor stereotypy in typical and atypical development: The role of sensorimotor integration," *Front. Integr. Neurosci.*, 2017, doi: 10.3389/fnint.2017.00019.

[3] N. Zhang, Y. Wang, C. Kang, J. Cheng, and D. He, "Blockchain technique in the Energy Internet: preliminary research framework and typical applications," *Zhongguo Dianji Gongcheng Xuebao/Proceedings Chinese Soc. Electr. Eng.*, 2016, doi: 10.13334/j.0258-8013.pcsee.161311.

[4] V. L. Pavlov, N. Boyko, A. Babich, O. Kuchaiev, and S. Busygin, "Applying pantomime and reverse engineering techniques in software engineering education," in *Proceedings - Frontiers in Education Conference, FIE*, 2007. doi: 10.1109/FIE.2007.4418054.

[5] V. V. Rogalchuk, A. V. Tyrva, and A. D. Khomonenko, "Estimation of program reverse semantic traceability influence at program reliability with assistance of object-oriented metrics," in *2009 5th Central and Eastern European Software Engineering Conference in Russia, CEE-SECR 2009*, 2009. doi: 10.1109/CEE-SECR.2009.5501175.

[6] J. I. Maletic, E. V Munson, A. Marcus, and T. N. Nguyen, "Using a hypertext model for traceability link conformance analysis," *Proc. 2nd Int. Work. Traceability Emerg. Forms Softw. Eng. TEFSE03*, 2003.

[7] Y. Zhang, R. Witte, J. Rilling, and V. Haarslev, "An ontology-based approach for traceability recovery," in *3rd International Workshop on Metamodels, Schemas, Grammars, and Ontologies for Reverse Engineering (ATEM 2006)*, 2006.

[8] J. Hu and M. P. Wellman, "Nash Q-learning for general-sum stochastic games," *J. Mach. Learn. Res.*, 2004, doi: 10.1162/1532443041827880.

[9] C. García-Ruiz and J. C. Fernández-Checa, "Mitochondrial Oxidative Stress and Antioxidants Balance in Fatty Liver Disease," *Hepatology Communications*. 2018. doi: 10.1002/hep4.1271.

[10] M. Idrissou, D. F. Ngom, R. Malou, B. Vayssade, and A.-L. Courbis, "Water balance simulation for resource evaluation at the catchment scale: application to the Nema (Sudano-Sahelian zone, Senegal)," *Hydrol. Sci. J.*, 2015, doi: 10.1080/02626667.2014.932055.

[11] M. Hearson, "The Challenges for Developing Countries in International Tax Justice," *Journal of Development Studies*. 2018. doi: 10.1080/00220388.2017.1309040.

[12] Y. Yan, R. Zhao, and H. Chen, "Prisoner's dilemma on competing retailers' investment in green supply chain management," *J. Clean. Prod.*, 2018, doi: 10.1016/j.jclepro.2018.02.229.

CHAPTER 8
GRAPHENE NANO SCROLL CARRIER STATISTICS AND QUANTUM CAPACITANCE MODELS

Manish Joshi, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id- gothroughmanish@gmail.com

ABSTRACT:

The extraordinary electrical characteristics of graphene nano scroll (GNS), a new class of quasi-one-dimensional materials, have recently captured the attention of researchers. Additionally, compared to their corresponding planar configurations like nanoribbon, nanotube, and bilayer graphene, it is acknowledged that the scrolling shapes of graphene suggest greater energy stability. The current study introduces modeling of the density of state (DOS), carrier concentration, and quantum capacitance for graphene nano scroll (illustrated schematic ideal scroll-like Archimedes spiral). This is done by employing a novel analytical approach. The DOS model was initially developed, and the carrier concentration and quantum capacitance models were later computed using it. In addition, the effect of structural factors and chirality number on the density of state and carrier concentration were studied, as well as the carrier concentration and quantum capacitance for both degenerate and nondegenerate regimes. Later, researchers looked into how temperature affected quantum capacitance. The carbon allotrope known as graphene is a single-layer of atoms organized in a hexagonal lattice nanostructure. Since the carbon allotrope in graphite contains a lot of double bonds, the name is derived from "graphite" and the suffix -end.

KEYWORDS

Density, Researches, Temperature, Quantum.

INTRODUCTION

It is generally acknowledged that materials based on graphene exhibit superior electrical transmission. Due to recent developments in the creation of atomic-sized conductors, it is now possible to fabricate electronic devices with extremely small dimensions (less than 100 nanometers). The majority of previous surveys have, up until recently, mostly concentrated on cnt and gnr. The existence of microscopic graphene nanoribbons that have been coiled up into spirals like the graphene nano scrolls is well known. The principles behind the development of gns share striking similarities with those underlying the recognizable big graphene and boron nitride nano scrolls. The rolled layers' overlapping surfaces in gnss have the power to increase structural stability. According to reports, graphene nano scrolls are promising materials for the next generation of nanoelectronics devices, such as the channel and connectivity in fets and mosfets. The electron microscope and diffraction can be regarded as highly effective methods for delineating the structure of the nano scrolls. Gnss have the potential to be employed as carriers for electrons. Currently, the hydrothermal technique or plasma accelerated chemical vapor deposition (cvd) are being used to create the quasi-one-dimensional nanocarbons, specifically the nano wall, nanowire, nanobelt, and nano scroll.

The creation of high-quality carbon Nano scrolls (CNSs) has been thought of as being made possible by using the isotropy alcohol solution to wrap up monolayer graphene that has been predefined on substrates. Furthermore, it is established that the GNS obtained using this technique would be able to support a high current density up to. This demonstrates that it is a

suitable option for microcircuit connectivity, in turn. Boron nitride nanotubes (BNNTs) and boron nitride Nano scrolls (BNSs) are a significant class of nanostructures. BNSs have not yet been investigated, despite the fact that many theoretical and experimental works have been reported to BNNTs. In theory, cubic boron nitride crystals can be utilized as a starting material to create BNSs utilizing the previously described experimental techniques for creating CNSs. Similar to CNSs, BNS formation is dominated by two major energetic contributions: the free energy decrease brought on by the van der Waals interaction energy of overlapping regions of the layer increasing stability and the elastic energy increase brought on by the bending of the BN layer decreasing stability. According to this, the van der Waals interactions apply more to the BNS instance than the CNS case.

A typical graphene nano scroll is its length is given by where stands for the right angle for the spiral from to for as the starting value from,, (is the number of turns), and (is the distance from the origin). The energy dispersion relation and conductance of graphene nano scroll can be obtained by using a third nearest-neighbor tight binding analysis. It has been confirmed that the energy band gap depends on the geometrical structure and chirality number. Therefore, in future research for using GNS as electronic devices, similar to the FET and MOSFET, the controlled band gap in GNS can stand as the key factor. The parabolic energy dispersion approximation for the ZGNS is used in the current research to describe the state density, carrier concentration, and quantum capacitance. The quantum capacitance is appropriately approximated by applying the first derivative of the Taylors expansion approximation on the energy dispersion equation. At low concentrations, the Maxwell-Boltzmann distribution law was also applied. The quantum capacitance was only thought to exist within parabolic bands at the time. Additionally, while analyzing the impacts of temperature fluctuations on the quantum capacitance, the carrier statistics dependence on the spiral and chiral vector lengths was demonstrated.

The impact of the electro actuation phenomena on the quantum capacitance model has not been considered in this article. A valence band covers the whole surface of a graphene sheet because each atom is joined to its three closest neighbors by -bonds and a delocalized -bond. The same kind of bonding is present in glassy carbon, fullerenes, polycyclic aromatic hydrocarbons, carbon nanotubes, and (partially) polycyclic aromatic hydrocarbons. Graphene is a semimetal with peculiar electrical characteristics that are best explained by theories for massless relativistic particles because the valence band is touched by a conduction band. The energy-to-momentum relationship of charge carriers in graphene is linear rather than quadratic, and the material can be used to create field-effect transistors with bipolar conduction. Long-range charge transport is ballistic, and the material displays substantial quantum oscillations as well as substantial and nonlinear diamagnetism Although graphite is black because of the material's significant absorption of light at all visible wavelengths, a single graphene sheet is virtually transparent due to its exceptional thinness. The material has a strength that is approximately 100 times greater than the strongest steel of the same thickness.

Transmitted light image of a floating graphene membrane. Because it absorbs around 2.3% of light, this one atom thick substance can be viewed with the naked eye. For many years, researchers conjectured about graphene's probable existence and manufacture. It has probably been made inadvertently for ages in modest amounts through the usage of pencils and other graphite-based products. In 1962, it may have been seen using electron microscopes, but it was only examined when supported on metal surfaces. Andre Gemi and Konstantin NovoLog found, isolated, and studied the substance in 2004 at the University of Manchester Gleim and Novello received the 2010 Nobel Prize in Physics for their "groundbreaking experiments

regarding the two-dimensional material graphene It turned out to be surprisingly simple to extract high-quality graphene.

Due to its unusually high tensile strength, electrical conductivity, transparency, and status as the world's thinnest two-dimensional material, graphene has emerged as a valued and practical nanomaterial. In 2012, there was a $9 million global market for graphene with the majority of the demand coming from research and development in semiconductor, electronics, electric batteries, and composites. The term "graphite" should be used to describe the three-dimensional substance, according to the IUPAC (International Union for Pure and Applied Chemistry, while "graphene" should only be used when discussing the reactions, structural relationships, or other aspects of individual layers. According to a more specific definition, "isolated or free-standing graphene" must be adequately isolated from its surroundings however this definition would also include layers suspended in or transferred to silicon dioxide or silicon carbide..

DISCUSSION

A Report released transmission electron microscopy (TEM) photographs of thin graphite samples made up of a few graphene layers in 1948. Single layers were eventually also directly detected. By using transmission electron microscopy on bulk materials, particularly within of soot produced by chemical exfoliation, single layers of graphite were also discovered. Hanns-Peter Boehm's study of incredibly thin graphite flakes was published in 1961–1962, and he came up with the name "graphene" for the fictitious single-layer structure. This study presents graphitic flakes that provide additional contrast equivalent to 3 atomic layers of amorphous carbon, or 0.4 nm, in thickness. For 1960 TEMs, this was the highest resolution achievable. However, it is impossible to debate how many layers were included in those flakes, either back then or now. We now understand that the focusing settings have the greatest influence on the TEM contrast of graphene. The only method now known to discriminate between suspended monolayer and multilayer graphene, for instance, is to compare the relative brightness of different diffraction sites. Gleim and Novello’s 2007 review's references 24 and 26 most likely contain the first trustworthy TEM observations of monolayers [1]–[3].

Beginning in the 1970s, C. Oshima and others described the epitaxial growth of single layers of carbon atoms on top of other materials. Similar to free-standing graphene, this "epitaxial graphene" is made of a hexagonal lattice of sp2-bonded carbon atoms that is one atom thick. The electrical structure of the two materials is notably different from that of graphene that is free-standing because to the large charge transfer between the two materials and, in some situations, hybridization between the d-orbitals of the substrate atoms and the orbitals of graphene. As a component of graphite intercalation compounds, which can be thought of as crystalline salts of the intercalant and graphene, single sheets of graphite were once again referred to as "graphene" in 1987. Additionally, it was utilized in the descriptions of polycyclic aromatic hydrocarbons in 2000 by S. Wang and others as well as carbon nanotubes in 1992 by R. Saito, Mildred, and Gene Dresselhaus.

In 1990, work to produce thin graphite sheets using mechanical exfoliation began. Early attempts used exfoliation methods akin to the sketching process. It was possible to obtain multilayer samples with a thickness of 10 nm. A method to create graphene by periodically peeling off layers from a graphite flake glued to a substrate, producing a graphite thickness of 0.00001 inches was described by Robert B. Rutherford and Richard L. Dudman in a patent application in the US in 2002. High-throughput visual recognition of graphene on a carefully selected substrate, which offers a slight but perceptible optical difference, was the key to

success. The same year, Boor Z. Jang and Wen C. Huang submitted a second U.S. patent application for a process based on exfoliation and attrition to manufacture graphene. the patenting of a method for making single-layer graphene sheets by inventor Larry Fullerton.

complete characterization and isolation

At the Nobel Laureate press presentation, Royal Swedish Academy of Sciences, 2010, Andre grim and Konstantin Novello were present. At the University of Manchester, Andre grim and Konstantin Novello successfully isolated and described graphene in 2004. They employed the Scotch tape technique or micromechanical cleavage to separate graphene layers from graphite using a standard adhesive tape. Then, a thin coating of silicon dioxide (silica) on a silicon plate (a "wafer") was covered with the graphene flakes. In order to create nearly charge-neutral graphene layers, the silica electrically separated the graphene and very slightly interacted with it. beneath the Sio is the silicon

To change the charge density in the graphene over an extensive range, electrode 2 might be employed as a "back gate" electrode. These two were awarded the 2010 Nobel Prize in Physics "for pioneering experiments regarding the two-dimensional material graphene" as a result of their work.

A "graphene gold rush" was prompted by their publication and the unexpectedly simple preparation procedure they revealed. Quantum mechanical, electrical, chemical, mechanical, optical, magnetic, and other unusual features of the material were explored as the research grew and branched off into several subfields [4]–[6].

Bonding

The hybrid orbital sp2 with three major lobes at 120° is formed from the carbon orbitals 2s, 2px, and 2py. The last orbital, pz, is protruding from the plane of the graphene. graphene's sigma and pi bonds. Sigma bonds are created when sp2 hybrid orbitals overlap, while pi bonds are created when projecting pz orbitals tunnel through one another. In order to establish -bonds, three of each atom's four outer-shell electrons in a graphene sheet occupy three sp2 hybrid orbitals, which are a combination of orbitals s, px, and py. These bonds are 0.142 nanometers long on average. The final outer-shell electron is located in an orbital called a pz that is parallel to the plane. These orbitals combine to create the two half-filled bands of free-moving electrons, and, which are in charge of the majority of the significant electronic features of graphene. Recent quantitative assessments of aromatic stabilization and limiting size based on hydrogenation enthalpies (Hydron) are in good agreement with published data [7]–[9].

A stack of graphene sheets with an interplanar spacing of 0.335 nm (3.35) results in graphite. The (002) layering of graphite is typically visible in the diffraction patterns of solid graphene sheets. Some single-walled nanostructures exhibit this property. However, the center of premolar graphite onions has been shown to include unlayered graphene with only (hk0) rings. Flat graphene sheets exhibit faceting at faults, according to TEM research, and this suggests that two-dimensional crystallization from a melt may be involved. Transmission electron microscopy (TEM) of sheets of single-layer graphene hung between bars of a metallic grid allows for the direct observation of the hexagonal lattice structure. A flat sheet appeared to be "rippling" in several of these photos, with an amplitude of roughly one nanometer. The instability of two-dimensional crystals may cause these waves to be inherent to the material, or could be a result of the constant dirt that can be observed in TEM photos of graphene. The "adsorbates" shown in TEM images could be photoresist residue, which needs to be removed to achieve atomic-resolution images. This would explain the rippling that

Images of graphene supported on silicon dioxide substrates obtained using a scanning tunneling microscope (STM) similarly exhibit the hexagonal structure. The rippling visible in these photos is not intrinsic; rather, it results from graphene's conformity to the substrate's lattice.

Electronic

Graphene's electronic characteristics Graphene's electronic band structure. At the six points where the valence and conduction bands intersect, linearly spreading Dirac cones are created. zero-gap semiconductor. Six sites in momentum space, near the Brillouin zone, that are separated into two distinct groups of three points each are known as the Dirac points. The labels for the two sets are K and K'. The sets result in a valley degeneracy of gv = 2 for graphene. In contrast, the main area of interest for conventional semiconductors is typically, where momentum is zero. It differs from other condensed matter systems in four key ways. However, its electrical structure would alter if the in-plane direction was constrained as opposed to unlimited. Graphene nanoribbons are the name given to them. The bandgap would remain 0 even if it were to "zig-zag". The bandgap would be non-zero if it is "armchair".

The hexagonal lattice of graphene can be compared to two interlocking triangle lattices. With the use of a tight-binding approximation, this viewpoint was successfully employed to determine the band structure for a single layer of graphite. Epitaxial graphene nanoribbons with a width of 40 nanometers exhibit distinct changes in electrical resistance. The conductivity of the ribbons outperforms expectations by a factor of 10. Electrons can flow easily along the ribbon edges by acting more like optical waveguides or quantum dots in the case of the ribbons. Resistance in copper rises according to length as electrons come into contact with contaminants. Two kinds of transportation are most common. One is thermally activated, whereas the other is ballistic and temperature-independent. Ballistic electrons are similar to carbon nanotubes, which are spherical. At a specific lengththe ballistic mode at 16 micrometers and the other at 160 nanometers (1% of the former length)resistance rises sharply at ambient temperature [10], [11].

Even at ambient temperature, graphene electrons can travel micrometer distances without scattering. Despite having no carriers close to the Dirac points, graphene has a minimum conductivity of It is still unknown where this low conductivity came from. However, ionized impurities in the Sio or rippling of the graphene sheet substrate could result in small pockets of carriers that permit conduction. According to some ideas, the minimum conductivity should be minimal carrier density At high carrier densities, graphene displays both positive and negative photoconductivity. The interaction between photoinduced variations in the carrier scattering rate and the Drude weight controls this. By gently heating in vacuum, graphene that has been doped with different gaseous species (both acceptors and donors) can be restored to its undoped condition. Carrier mobility shows no discernible change even at dopant concentrations greater than 1012 cm2. In a low-temperature, ultra-high vacuum environment, graphene doped with potassium can limit mobility by a factor of 20. When the potassium is removed from the graphene by heating it, the mobility loss can be reversed. It is believed that charge fractionalization, which occurs when individual pseudo particles in low-dimensional systems have apparent charges that are less than one quantum happens as a result of graphene's two dimensions. Therefore, it might be a good material for building quantum computers that use anionic circuits.

CONCLUSION

The Hall effect is the development of transverse (perpendicular to the main current) conductivity in the presence of a magnetic field. The quantum Hall effect is a quantum

mechanical variation of the Hall effect. The Hall effect's quantization where e is the fundamental electric charge and h is the Planck constant, is a fundamental quantity. Typically, it can only be seen in extremely pure gallium arsenide or silicon crystals at 3 K temperatures and extremely strong magnetic fields. With regard to conductivity quantization, graphene exhibits the quantum Hall effect. The effect is unusual in that the steps are displaced by half relative to the conventional steps and by an additional factor of four. The Hall conductivity of graphene the unit cell of graphene contains two identical carbon atoms and two zero-energy states, one of which has the electron on atom A and the other of which has the electron on atom B. The scenario is different, though, if the unit cell's two atoms aren't exactly the same. According to research by Hunt et al., hexagonal boron nitride (h-BN) in contact with graphene can change the potential between atoms A and B to the point where the electrons generate a mass and band gap of around 30 me and other properties.

REFERENCES

[1] R. Tenne, L. Margulis, M. Genut, and G. Hodes, "Polyhedral and cylindrical structures of tungsten disulphide," *Nature*, 1992, doi: 10.1038/360444a0.

[2] D. B. Snyder, J. A. Craven, M. Pilkington, and M. J. Hillier, "The 3-dimensional construction of the Rae craton, central Canada," *Geochemistry, Geophys. Geosystems*, 2015, doi: 10.1002/2015GC005957.

[3] S. Wu *et al.*, "Observation of the quantum spin Hall effect up to 100 kelvin in a monolayer crystal Downloaded from," *Science (80-.).*, 2018.

[4] J. Schageman *et al.*, "The complete exosome workflow solution: From isolation to characterization of RNA cargo," *Biomed Res. Int.*, 2013, doi: 10.1155/2013/253957.

[5] N. C. Kim, T. N. Graf, C. M. Sparacino, M. C. Wani, and M. E. Wall, "Complete isolation and characterization of silybins and isosilybins from milk thistle (Silybum marianum)," *Org. Biomol. Chem.*, 2003, doi: 10.1039/b300099k.

[6] V. E. Chizhikov, C. F. Spiropoulou, S. P. Morzunov, M. C. Monroe, C. J. Peters, and S. T. Nichol, "Complete genetic characterization and analysis of isolation of Sin Nombre virus," *J. Virol.*, 1995, doi: 10.1128/jvi.69.12.8132-8136.1995.

[7] M. Wuttig, V. L. Deringer, X. Gonze, C. Bichara, and J. Y. Raty, "Incipient metals: Functional materials with a unique bonding mechanism," *Adv. Mater.*, 2018, doi: 10.1002/adma.201803777.

[8] E. S. A. de Cock *et al.*, "Longitudinal Associations Between Parental Bonding, Parenting Stress, and Executive Functioning in Toddlerhood," *J. Child Fam. Stud.*, 2017, doi: 10.1007/s10826-017-0679-7.

[9] C. W. Tsao and D. L. DeVoe, "Bonding of thermoplastic polymer microfluidics," *Microfluidics and Nanofluidics*. 2009. doi: 10.1007/s10404-008-0361-x.

[10] D. Slavko, "Electronic Commerce," *ECONOMICS*, 2016, doi: 10.1515/eoik-2017-0003.

[11] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene," *Rev. Mod. Phys.*, 2009, doi: 10.1103/RevModPhys.81.109.

CHAPTER 9
USING MOVING STATISTICS THEORY, THE SIGNAL NON-STATIONARY DEGREE EVALUATION METHOD

Namit Gupta, Associate Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id- namit.k.gupta@gmail.com

ABSTRACT:

Nonstationary signals are those whose statistics change over time, and determining their degree of nonstationary can help determine the signal source's operational status. This study analyzed the applicability of various evaluation methods and offered a variety of typical signal global and local nonstationary degree evaluation techniques. Based on the theory of moving statistics, this paper proposed the concepts and calculation methods of the moving mean, moving standard deviation, moving variation coefficient, and moving Hurst exponent in light of the limitations of the existing evaluation methods in the range of application and taking the influence of adjacent signal points into consideration. Three distinct fields of signals sinusoidal signal, mechanical failure signal, and ECG signal are examined using various nonstationary degree evaluation techniques. The findings demonstrate that the signal nonstationary degree evaluation method put forth in this paper has distinct advantages over other nonstationary signal processing techniques by being able to reveal the time-varying details of nonstationary signals with high precision and strong stability.

KEYWORDS:

Determining, Evaluation, Nonstationary, Signals, Techniques.

INTRODUCTION

Numerous measurements and associated studies demonstrate that the signals derived from natural phenomena and technical applications are typically stochastic or random because of their unique properties and noise effects. The statistical method must be used to study the random signal because it cannot be correctly recreated nor can it be characterized by a deterministic time function. Researchers in various branches of science and engineering must acquire physical parameters or statistics that, depending on their specific requirements, can represent time-varying random signals in order to perform signal identification or diagnosis. Therefore, it is important to identify statistical features that can accurately capture the random properties of the signal [1], [2].

The typical statistics for random signals include the mean value first-order statistics, variance and correlation function second-order statistics, and higher-order statistics like third-order and fourth-order or higher-order moments, higher-order cumulants, and higher-order spectrum. The signal is referred to as a stationary signal if the statistic's change is time independent. The signal is referred to as a nonstationary signal if the statistic changes over time in engineering, random signals are frequently nonstationary signals. When a stationary signal and nonstationary noise are combined, the nonstationary component can be eliminated using filtering or smoothing technologies, leaving only the signal's true information. It is necessary to conduct extensive research on the time-frequency variation characteristics and extract significant information to provide technical support for feature recognition and spectrum analysis when the nonstationary degree of the signal can represent the physical or essential characteristics of the signal source. The time-varying properties of the signal, a local

distortion degree, trend term, and envelope characteristics must all be taken into consideration when the nonstationary of the signal can indicate the short- or long-term variation of the signal. These characteristics can serve as a foundation for defect identification and signal classification. Therefore, it is crucial to research the level of signal nonstationary.

Many assumptions made about stationary signals in engineering applications are no longer true as nonstationary signals are more understood. Modern signal processing has been increasingly concerned with the analysis and handling of nonstationary signals. Nonparametric method and parametric method are the two broad categories that nonstationary signal analysis falls under, depending on the many technological approaches used. The aforementioned two categories are further classified into time-domain methods, frequency domain methods, and time-frequency-domain methods, respectively, based on the various analytic domains of each method The nonparametric approach immediately transforms the signal before processing it. Total energy, cumulative energy, Teager energy operator (TEO), temporal moment of energy, power, signal envelope, and average time are the primary time-domain nonparametric approaches. The Fourier spectrum, the traditional spectral estimate method of the power spectrum, high-order spectral analysis, average frequency, spectral moment, and other spectral parameters are the major frequency domain methods in the nonparametric method. The wavelet decomposition, Wigner-Ville distribution (WVD) decomposition, and Hilbert-Huang transform (HHT) approach are the basic time-frequency-domain nonparametric methods. The time-frequency analysis approach can be used to get domain decomposition signals, but its rigor, accuracy, and adaptability need to be further enhanced due to frequency band aliasing or cross term interference [3]–[5].

Model-based method is another name for the parametric approach. The signal is first used to construct a parametric model, which is then used to analyze and process the signal in accordance with the model's parameters. The time-domain method is primarily based on the time-varying ARMA model and the time-varying basis function ARMA model of instantaneous spectrum estimation method, while the frequency domain method is primarily based on the signal power spectrum parametric modeling method and the time-frequency-domain method is primarily based on the ARMA model and the AR model of modern spectrum estimation method.It should be noted that the approaches above typically employ statistics, transformation, decomposition, and other signal processing. To extract the signal's time-frequency-domain features and information is the goal. However, the quantification of nonstationary signals, or nonstationary degree evaluation, can extract the characteristics of the signal from its core and pinpoint its intrinsic property, which is also required for many technical applications.

Although the definitions of signal stationarity and nonstationarity are rather straightforward, the actual procedure of determining the degree of signal nonstationarity is somewhat difficult, and there isn't a universally useful quantitative index. Numerous nonstationary signal processing techniques are very simple or have other drawbacks as a result of theoretical development limits; both qualitative and quantitative analysis are not thorough and systematic, necessitating extensive research. In light of this, this research offered various approaches for evaluating the degree of signal nonstationarity and contrasted their applicability and limits. Numerous signal nonstationary degree evaluation approaches, such as moving variation coefficient, are proposed based on moving statistical features and take into account the time-varying properties of nonstationary signals. The examination of various signal types demonstrates the effectiveness of the approaches suggested in this paper and demonstrates their capacity to precisely identify the nonstationary degree of nonstationary

signals. The study block When shifted in time, the unconditional joint probability distribution of a stochastic process known as a stationary process (also known as a strict/strictly stationary process, strong/strongly stationary process, or strongly stationary process) does not change. As a result, variables like mean and variance also remain constant over time. A stationary process, if you could draw a line across the middle of it, would be flat; it might have'seasonal' cycles around the trend line, but it does not generally trend upward or downward.

Non-stationary data are frequently modified to become stationary since stationarity is an assumption underlying many statistical processes used in time series analysis. A trend in the mean, which can be caused by either the existence of a unit root or by a deterministic trend, is the most frequent reason for stationarity to be violated. Stochastic shocks have long-lasting consequences in the former unit root example, and the process is not mean-reverting. Stochastic shocks only temporarily affect the variable in the latter case of a deterministic trend, which is known as a trend-stationary process, after which the variable tends toward a deterministically evolving (non-constant) mean. By eliminating the underlying trend, which is only a function of time, a trend stationary processwhich is not precisely stationarycan be converted into a stationary process. Similar to this, differencing can make processes stationary that have one or more unit roots. Cyclisation processes, which are stochastic processes that vary cyclically across time, are a significant subset of non-stationary processes that do not exhibit trend-like behavior. Strict-sense stationarity is too limiting for many applications. The use of wide-sense stationarity or N-th-order stationarity or other types of stationarity is then made. Different writers' definitions of various types of stationarity differ from one another [6], [7].

DISCUSSION

There is an urgent need for discussion in several academic domains about the evaluation and analysis of nonstationary signals. To determine if a particular signal exhibits nonstationary or not and to assess the degree of no stationarity, or nonstationary degree, there are numerous approaches in diverse domains. For instance, average entropy theory and the unit root test method are frequently employed in econometrics. Nonlinear system analysis typically uses recurrence plots. Stock analysis employs the Hurst exponent. Seismic signal analysis employs the crossing-zero rate and nonlinear cross prediction method. The aforementioned approaches each have their own benefits and drawbacks, thus the best approach should be chosen based on the engineering specifications and characteristics for actual use. The global assessment method and the local evaluation technique can be used to categorize the aforementioned methodologies. The former employs only one value for evaluation and concentrates on the signal's entire nonstationary degree. The latter, whose value is a collection of vectors, can effectively indicate the local or signal's degree of no stationarity. This paper introduces, examines, and talks about the application scope and features of typical algorithms in order to choose an appropriate nonstationary degree evaluation method according to the signal characteristic requirements.

Signal Global Nonstationary Degree Evaluation Method

Using relevant indicators, the nonstationary degree of a signal is assessed using the global nonstationary degree evaluation method. The approach for determining how nonstationary a signal is globally is straightforward to use and may qualitatively and immediately describe how nonstationary a signal is at the global level. The autocorrelation function technique, unit root test, information entropy method, and recurrence plot method are the basic approaches for evaluating signal global nonstationary degree. The autocorrelation function can indicate the degree of correlation at various points throughout the same process and can be used to

characterize the correlation between neighboring variables in time series. A random time series signal can be regarded as stationary if the autocorrelation function rapidly approaches zero as the delay increases. A sequence is deemed nonstationary if the autocorrelation function does not rapidly fall to zero as the delay increases. Although the autocorrelation function approach is straightforward and easy to understand, the amount of data that can be used in the calculation will decrease as delay increases since the actual signal has a finite duration and is bounded. At this point, it has been challenging for the autocorrelation function to completely and precisely assess the signal's degree of stability across time.

The unit root test refers to establishing the sequence's nonstationary degree by determining whether the sequence contains a unit root. The signal is unstable if the sequence has a unit root. Unit root testing can be done using a variety of techniques, such as the ADF, PP, and NP tests. The unit root test can only judge whether the signal is stable or nonstationary according to weather is equal to 1, so it cannot be used to quantitatively evaluate the nonstationary degree of various signals. This is because the unit root test is a type of qualitative judgment for the evaluation of signal no stationarity, but not a quantitative analysis. By removing the upper bound of the information entropy from the original signal's stable information structure, the nonstationary degree based on information entropy measures the signal's no stationarity. Although the distinction between stationary and nonstationary is not sufficiently obvious, the signal is thought to be nonstationary when the value approaches 0, and stationary when the value approaches 1. Additionally, the upper bound of information entropy varies slightly for different signal types, making it hard to assess the nonstationary degree of various signals using information entropy. This results in the application and expansion of information entropy being constrained [8]–[10].

Recurrence plot analysis is a signal processing technique that illustrates the system's recursive properties using the idea of phase space reconstruction. Through the periodic trajectory of phase space representation, it primarily assesses the nonstationary level of time series. The change in system trajectory over time in the state space is particularly sensitive to the recurrence plot. With this technique, the trajectory in a high-dimensional space can be immediately analyzed. The image representation is not entirely suitable in some domains, and the method's capacity to evaluate some long-term signals with strong nonstationary is still constrained.

Local Nonstationary Degree of Signals Evaluation Method

The global nonstationary degree evaluation approach may qualitatively assess the signal's nonstationary, but it is unable to expose the signal's time-varying properties and cannot satisfy the precision criteria. Some random signals exhibit nonstationary and mutation in the local segment but are stationary overall. It is required to simultaneously monitor the entire signal fluctuation and the local signal mutation in order to judge the stationary change of the signal comprehensively and thoroughly. The statistic change can be quantitatively calculated using the local nonstationary degree evaluation method. The nonlinear cross prediction method, the stationary degree method based on the Hilbert-Huang Transform (HHT), the crossing-zero rate method, and the Hurst graph approach are the most used techniques for evaluating local nonstationary degree of signals.

The signal is separated into a number of adjacent segments based on the nonlinear cross prediction method, and the cross-prediction errors between segments are utilized to describe the degree of variation and stability of the time series of each segment. The benefit of this method is that it is simple to produce more reliable calculation results since the nonstationary degree calculation takes into consideration both periodic fluctuations and the time series trend

changes over time. However, due to its amplitude sensitivity and unclear judgment threshold, this method can only be utilized as an auxiliary method. the number of times a signal crosses across zero points in a unit of time is known as the crossing-zero rate. It can be used to show how a signal's characteristics, such as its frequency or period, change with time. As a result, it can quickly display a frequency that is nonstationary to some extent. However, the crossing-zero rate, which is susceptible to high-frequency omission, makes it difficult to distinguish whether the high-frequency portion of the signal is prominent and does not entirely intersect with the zero line. As a result, the zero-crossing rate cannot accurately capture the degree of nonstationary in the signal. Typically, it needs to be used in conjunction with other techniques to produce a thorough evaluation.

The Hurst exponent, which is determined using the rescaled range analysis approach is a type of indicator used to determine whether time series data follow a biased random walk or random walk process. Additionally, the Hurst exponent helps gauge how noise levels and signal intensity are changing over time. As a result, it can serve as the foundation for evaluating the degree of signal nonstationary. The cumulative deviation of the signal should be calculated using the following formula to determine the Hurst exponent:3. Moving Statistics-Based Nonstationary Degree Evaluation Method Each point in the signal is related to the neighboring points even if the actual signals in many engineering domains exhibit stochastic features due to the influence of the generating process and noise characteristics. The signal close by typically has a bigger effect on the signal point itself, whereas the signal far away typically has a less effect on the center point. By taking into account the relationship between various signal points and their capacity for interaction, the nonstationary degree of random signals may be evaluated more thoroughly.

The weighted moving statistics method can typically be used to calculate each point in the signal and highlight the aforementioned qualities. In order to acquire more effective signal features, data smoothing can be accomplished by giving weights to the signal points in the given range and replacing the points with statistics in the appropriate range. When signals are processed using weighted moving statistics, it is the same as low-pass filtering because the corresponding filters have a smooth transition from passband to stopband. As a result, it can reduce the impact of random fluctuations on signal trends, reflect local variation characteristics, and emphasize the general change rule. In this paper, a technique for determining the degree of signal nonstationary is proposed. It is based on the weighted moving statistics method. The method appropriately chooses and adjusts the weight value and weight range in accordance with the various precision requirements of signal analysis, and the moving statistics are parameters like mean value, variance and standard deviation, coefficient of variation, crossing-zero rate, and Hurst exponent. The fluctuation and mutation of signals may be determined by the examination and comparison of the aforementioned statistics, and the original signals can be quantitatively and clearly described. To assess the local and global nonstationary degree of signals, it is possible to define in fully the occurrence time, sequence, severity, and development process of signal feature changes. The approach has good application, sensitivity, and antinomies and is easy to use.

CONCLUSION

Mechanical fault signals, heart disease signals, and various groups of sinusoidal signals mixed with white noise are chosen, and the aforementioned methods are each applied in turn to determine the characteristics and applicability of various nonstationary degree evaluation methods. This is done in order to compare the effectiveness of traditional methods and the moving statistical calculation method. The simplest of the three signals is the sinusoidal signal. Any complex signal can be thought of as the superposition of several sinusoidal

signals of various amplitudes and frequency. The other two signals are typical complex signals seen in the mechanical and biological fields, respectively. The applicability of the suggested methodologies for simple signals and complicated signals, as well as their broad applicability in various domains, can be checked to some extent by using these three signals as examples for verification analysis. Additionally, the nonstationary degree analysis can be performed after removing the mean value for signals whose mean value is not zero and the nonstationarity is not strong. The accuracy of the calculating results is not significantly impacted by this procedure. The difficulty of nonstationary degree analysis is lower than that of the signal whose mean value is zero for the signal whose overall mean value is not zero and the nonstationarity is very strong because the signal itself has a very obvious trend term. Therefore, in this paper, the signal whose nonstationary evaluation is more difficult to analyze is not chosen. It is important to note that is taken as 0.1 in this study, and the number m of points on the two sides of the data point participating in the weighting computation is equal to 5. According to professional requirements and signal sampling rate, the values of and m must be speculatively established for other various signals. In general, the value ranges of and m should be between 0 and 0.2 and 1/100 and 1/10 of the signal sampling rate, respectively.

REFERENCES:

[1] J. Fawcett, "Thoughts About Theories and Statistics," *Nurs. Sci. Q.*, 2015, doi: 10.1177/0894318415585635.

[2] G. B. Weller and W. F. Eddy, "Multivariate Order Statistics: Theory and Application," *Annu. Rev. Stat. Its Appl.*, 2015, doi: 10.1146/annurev-statistics-062614-042835.

[3] C. Zhang, J. Yu, H. Zheng, G. Zhang, T. Gong, and W. Zeng, "System framework of land consolidation based on systematic engineering and statistics theory in China," *Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric. Eng.*, 2015, doi: 10.11975/j.issn.1002-6819.2015.15.035.

[4] T. Watari, "Statistics of F-theory flux vacua for particle physics," *J. High Energy Phys.*, 2015, doi: 10.1007/JHEP11(2015)065.

[5] D. Danner, D. Hagemann, and K. Fiedler, "Mediation analysis with structural equation models: Combining theory, design, and statistics," *Eur. J. Soc. Psychol.*, 2015, doi: 10.1002/ejsp.2106.

[6] "Probability Theory and Statistics," in *Machine Learning*, 2015. doi: 10.1016/b978-0-12-801522-3.09983-3.

[7] F. B. Christiansen, "A. Probability Theory and Statistics," in *Theories of Population Variation in Genes and Genomes*, 2015. doi: 10.1515/9781400866656-013.

[8] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*. 2015. doi: 10.1126/science.aaa8415.

[9] Q. Zhou and J. Luo, "The risk management using limit theory of statistics on extremes on the big data era," *J. Comput. Theor. Nanosci.*, 2015, doi: 10.1166/jctn.2015.4661.

[10] R. Angulo, M. Fasiello, L. Senatore, and Z. Vlah, "On the statistics of biased tracers in the Effective Field Theory of Large Scale Structures," *J. Cosmol. Astropart. Phys.*, 2015, doi: 10.1088/1475-7516/2015/09/029.

CHAPTER 10
BLACK HISPANIC BREAST CANCER SURVIVAL DATA INFERENTIAL STATISTICS

Pradeep Kumar Shah, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id- pradeep.rdndj@gmail.com

ABSTRACT:

In this study, we evaluate the statistical probability models for racial and ethnic differences in breast cancer survival rates. Information was gathered from women who were diagnosed with breast cancer in the US between 1973 and 2009. To create the statistical probability models, we randomly selected a stratified sample of Black and Hispanic female patients from the Surveillance Epidemiology and End Results (SEER) database. When measuring the goodness of fit tests, we used three commonly used model-building criteria: the Akaike Information Criteria (AIC), the Bayesian Information Criteria (BIC), and the Deviance Information Criteria (DIC). We discovered that the survival data for Black and Hispanic female patients fit the exponentiated exponential probability model better. The posterior density function for the model parameters and the predictive inference for future response were derived using a novel Bayesian technique. We paid particular attention to the Black Hispanic race. The summary results of posterior parameters were obtained using the Markov Chain Monte Carlo (MCMC) method. We also provided intervals that predicted future survival times. The allocation of healthcare resources and the design of treatments would both greatly benefit from these findings. The most prevalent invasive malignancy in women worldwide is breast cancer. Breast cancer accounts for 16% of all female cancers and 22.9% of all invasive malignancies in women Breast cancer claimed 458,503 lives worldwide in 2008, accounting for 6.0% of all cancer deaths in both men and women and 13.7% of all cancer deaths in women. Lung cancer, the second most frequent cause of cancer-related mortality in women, was responsible for 12.8% of those fatalities, or 18.2% of all cancer-related deaths in both men and women.

KEYWORDS:

Breast cancer, Bayesian, epidemiology, Information.

INTRODUCTION

Inducing permanent changes in cellular genetic pathways that lead to unchecked development and proliferations is what is known as cancer. Any aberrant growth of cancer cells that manifests as a lump or mass is referred to as a tumor. Fat, connective tissues, lymphatic veins, and arranged lobules of milk-secreting glands make up the majority of the human breast. These lobules have secretory ducts that connect them to the nipple on the outside. Because they are restricted by epithelial borders to the duct (ductal carcinoma in situ, or DCIS), or to the lobule (lobular carcinoma in situ, or LCIS), the majority of breast tumors are carcinoma in situ (CIS). One of the most prevalent tumors that can be fatal in women of all ages is breast cancer. In accordance with a World Health Organization (WHO) data from 2004, breast cancer accounts for roughly 16% of all cancer types and kills 519,000 people worldwide every year. Unexpectedly, 69% of these fatalities occurred in developing nations, dispelling the myth that breast cancer only affects wealthy nations. Over 226,000 new cases

of breast cancer are identified in the USA each year, and 40,000 American women pass away from the disease, according to the American Institute [1], [2].

Among all racial and ethnic groups in the US, breast cancer is one of the most frequently diagnosed malignancies. For all racial and ethnic groupings between 2004 and 2008, the incidence rates that are particular to race and ethnicity remained largely stable. Formerly, it was thought that some of the key factors influencing the prognosis of the disease included family history, socioeconomic status, levels of education, how frequently mammograms were performed, and access to health care resources. Recent research has however demonstrated that racial and ethnic characteristics can significantly increase the risk for the disease's prognosis. According to data gathered by the American Cancer Society, there are noticeable disparities in breast cancer death rates among states, socioeconomic divisions, and racial/ethnic groupings. Race and ethnicity may also be a significant risk factor for breast cancer, even if age is the best predictor. All ethnic groups except American Indians and Alaska Natives have seen a decline in breast cancer mortality rates since the early 1990s, highlighting yet another racial gap related to the illness. White women in the US have a higher risk of breast cancer than Black, Hispanic, Asian, or American Indian/Alaska Native women

Black Hispanic Women

Despite being the minority group in America with the fastest rate of growth, there are little particular breast cancer statistics for Black Hispanics. Hispanics' breast cancer statistics are typically reported under one ethnic category (Hispanics), hence it is difficult to get Hispanic breast cancer statistics by race. Overall, Hispanic (Black and White) women have lower rates of breast cancer incidence and mortality than non-Hispanic White women. When compared to women of any other ethnicity, Hispanic women (Black and White) exhibit poorer levels of awareness about the risk factors linked with the disease and have fewer access to healthcare facilities. Sadly, there aren't many research that explain differences in breast cancer prevalence across different racial groups within the Hispanic ethnic group. Typical research findings do not mention racial disparities within the ethnicity when describing breast cancer incidence, mortality, and death rates among Hispanic women, among other critical statistics linked with the disease. According to Banegas and Li, additional research on specific breast cancer outcomes among Hispanic women of varied racial backgrounds could significantly advance understanding of the prevalence and risk factors for the disease in this high-risk ethnic group. Their research revealed that, as compared to non-Hispanic Whites, non-Hispanic Blacks had a 1.5–2.5-fold higher risk of developing stage IV breast cancer types and a 10–50% higher risk of breast cancer-specific mortality The need for a study that attempts to comprehend the existing situation regarding breast cancer survival in this subgroup within the United States is once again demonstrated by this conclusion [3], [4].

DISCUSSION

The enormous volumes of phenomics and genetic data that healthcare professionals have gathered should be fully utilized from a research standpoint. It is important to test these enormous databases using more recent statistical techniques and statistical probability models. To better understand the severity and effects of the disease, it would be highly helpful to be able to forecast future patterns of morbidity and mortality. Exponentiated exponential (EE), exponentiated Weibull (EW), beta generalized exponential (BGE), beta inverse Weibull (BIW), and other statistical probability models may be used to describe data, in addition to exponential, gamma, Weibull, normal, half-normal, log-normal, Rayleigh, and

inverse Gaussian models. To describe the data and provide trustworthy scientific findings, statistical models are incredibly helpful.

Exponentiated exponential models (EEM) are frequently used in the engineering and biomedical disciplines for data modeling. Gupta and Kundu presented the EEM, a generalization of the exponential distribution, and it was quickly and broadly embraced. Two factors are taken into account by the EEM: "shape" and "scale." Furthermore, Gupta and Kundu pointed out that the EEM is comparable to the Weibull family and proposed the idea of substituting the EE distribution for the Weibull model. If a random variable's probability density function (pdf) is given by, then it is said to follow the exponentiated exponential distribution. where the scale parameter and the shape parameter are both. The aforementioned density function for exponentiated exponential distribution was first presented by Gupta and Kundu. You can write the random variable as. Designing rainfall estimation in the Coast of Chiapas, analyzing rainfall data in Los Angeles, and using software reliability growth models for essential quality indicators are a few intriguing examples of EEM. Kannan et al. established a cure rate model based on the generalized exponential distribution that takes into account the influence of risk variables or covariates on the likelihood of an individual being a long-term survivor. Additionally, the Gompertz form of the exponentiated exponential model was utilized to forecast the conductance of the squid axon voltage clamp.

The probability density function for a beta generalized exponential model is defined

where the scale parameter and the shape parameter are both. Additionally, and are two more parameters. These metrics serve as a description of skewness and tail weight. Beta exponential and generalized exponential models are two well-known models that the BGE model generalizes as specific examples. The Weibull distribution was developed by Swedish physicist Weibull particularly for testing a material's breaking strength. Mud Holkar and Srivastava proposed the first EW model with bathtub-shaped distribution and unimodal failure rates. Since then, Nassar and Eissa and Choudhury have both advocated using the EW model to examine lifetime data.

The probability density function (pdf) for the EW

Numerous issues in the disciplines of engineering, health, and medicine can be solved using the BIW paradigm. It displays the best fit for a variety of data sets, such as the length of time it takes for insulating fluids to break down after being subjected to tension. The probability density function (pdf) for the BIW model is given by where denotes the shape parameter and the two parameters, skewness and tail weight, are represented by and. The posterior probability for the parameters needed to calculate posterior inference can be derived using a novel Bayesian technique. A Bayesian estimate method treats model parameters and data as random variables. A probabilistic model describes their combined probability distribution. In a Bayesian approach, data are referred to as "observed variables" and parameters as "unobserved variables." In order to obtain the joint distribution for the parameters, multiply likelihood and prior. The parameter's information is contained in the "prior". The likelihood is expressed as a conditional distribution that specifies the probability of the observed data and is dependent on the model of the underlying process. Prior and likelihood incorporate all the data that is known about the parameters [5]–[7].

It is possible to deduce information about the probability model's parameters from the given data by modifying the joint distribution of prior and likelihood. For specific sets of observable data, the Bayesian inference aims to create the posterior distribution of the parameters. This study's goals are to examine a few socioeconomic and demographic variables; review right-skewed models EE, EW, BGE, and BIW demonstrate through

goodness-of-fit tests that the sample data fits a particular model conduct a Bayesian analysis of the posterior distribution of the parameters; and derive a Bayesian predictive model for future behavior.

Real-World Data Illustration

an actual example of data, breast cancer data from the Surveillance, Epidemiology and End Results website has been used. The SEER database contains information on breast cancer patients that has been gathered from twelve different states. Nine states were randomly chosen from among these twelve states using a stratified random selection approach. Race and ethnicity distribution categories for these nine states were included in the data. There were 4 269 men and 653 443 women in these statistics. Males rarely develop breast cancer, so we exclusively used data from females in our analysis. The Black Hispanic data were used to choose 298 female respondents using the simple random sampling (SRS) method. The pedigree chart for the grouping of Black Hispanic breast cancer patients among all female breast cancer patients. There were 300 Black Hispanic female patients overall, however information was lacking for 2 individuals. The nine randomly chosen states (dark blue zones), which were then randomly chosen Black Hispanic breast cancer patients In the United States, the lifetime risk for breast cancer is typically stated as 1 in 8 (12%) of women by age 95, with a 1 in 35 (3%) probability of dying from the disease. Except for those who pass away from breast cancer prior to becoming 95, it is assumed that all women will live to at least that age in this computation. Recent studies based on actual data suggest that the actual risk is likely less than half of the theoretical risk.

The highest yearly incidence rates of breast cancer are seen in the United States, where white people have a rate and African Americans have a rate of 112.6 per It is the second most prevalent cancer in women, just behind skin cancer, and the second leading cause of cancer death. In the US, 40,910 deaths from breast cancer were anticipated in 2007 (7% of cancer fatalities; nearly 2% of all deaths). Out of 2000 cancer cases, 450–500 annual fatalities among men are included in this number. Incidence and mortality rates for breast cancer have been falling in the US during the past few years. The age-adjusted incidence of breast cancer per 100,000 women in the US increased from roughly 102 cases annually in the 1970s to approximately 141 cases annually in the late 1990s. Since then, the incidence has decreased, remaining constant at around 125 since 2003. The number of age-adjusted breast cancer deaths per 100,000 women increased slightly before slowly declining Even though heart disease is a significantly more frequent cause of mortality for women, a US survey from 2005 found that breast cancer is still the most feared illness According to studies, women overestimate their chance of developing breast cancer [8]–[10].

Techniques for Determining

A statistical model's goodness of fit measures how well it matches a set of observations. The difference between observed values and the values anticipated by the in-question model is often summarized by measures of goodness of fit. Such measurements can be used to test for residual normality, determine whether two samples were taken from the same distributions (see Kolmogorov-Smirnov test), or determine whether outcome frequencies follow a particular distribution (see Pearson's chi-square test), among other statistical hypothesis testing techniques. A lack-of-fit sum of squares could be one of the components into which the variance is divided in the analysis of variance.The most popular models for evaluating the goodness of fit are the Akaike Information Criterion (AIC), Deviance Information Criterion (DIC), and Bayesian Information Criterion (BIC). A Bayesian measure of fit called DIC is used to compare various models, such as Congdon's usage of open data. Both positive and

negative numbers are possible with DIC. Models with lower values are deemed superior than those with higher values. When models with just fixed effects are fitted, DIC is comparable to AIC and yields findings that are identical to those of AIC. BIC is an asymptotic finding that relies on the assumption that the data distribution belongs to an exponential family. It can only be used to compare estimated models when the dependent variable's numerical values are the same throughout all comparisons. More than AIC, the BIC penalizes free parameters. Given any two estimated models, the model with a lower BIC value is favored, much like with AIC. The Black Hispanic Survival Data Results of Goodness of Fit Tests and Posterior Inference for the Parameters

The outcomes of the posterior distribution of the parameters from the EE for the survival data for Black Hispanic breast cancer patients the posterior distribution of the parameters is what is produced in the Bayesian technique when the knowledge of the parameter distribution is updated using observed data. Assuming that the observed random sample forms an adequate statistical probability distribution, we are interested in estimating the posterior distribution of the parameters for the breast cancer data. The posterior distribution parameters' findings and the data are used to construct the values of the land, which are then estimated using the MCMC method. A class of techniques called Markov Chain Monte Carlo is used in statistics to create samples from probability distributions. The log-likelihood function is derived from the EE model, and the parameter values are allotted to the proper theoretical probability distributions. Using the Windbags program, the parameters' summary results mean, SD, MC error, median, and confidence intervals are obtained. To eliminate any biases in the predicted values of the parameters due to the value used to begin the chain, the early iterations up to 1,000 are discarded in this process. This procedure is called burn-in. After the burn-in samples have been removed, the remaining samples are handled as if they came from the original distribution. The inference for the posterior parameters, was created using 50,000 Monte Carlo trials. The kernel densities for both the shape and size parameters exhibit roughly symmetric distributions after 50,000 Monte Carlo iterations.

The summary findings of the posterior distribution of the exponentiated Weibull parameters are presented. These findings were obtained using survival data from Black Hispanic female breast cancer patients. The data was used to generate the values, and. The outcomes of the posterior distribution parameters, and are estimated by setting the generated values using the MCMC method. The EW model serves as the foundation for the log-likelihood function. The parameter values that are subsequently allocated to the relevant probability distributions are then derived. Using the Windbags program, the parameters' summary results (mean, SD, MC Error, median, and confidence intervals) are calculated. summarizes the graphical representation of the parameter behavior distributions. While other model parameters display skewed distributions, the shape parameters and show a normal distribution. The results of the posterior distribution of the parameters from the BGE model are summarized. These findings were based on data collected from Black Hispanic female breast cancer patients. The summary findings (mean, SD, MC error, median, and confidence intervals) of the parameters were obtained using the Windbags software. The parameters and demonstrate a symmetrical pattern of distribution, while other parameters demonstrate a nonsymmetrical pattern. The results of the posterior distribution of the BIW model's parameter values. The data on breast cancer patient survival among Black and Hispanic women has been used. Using the Windbags program, the summary results of the parameters, including mean, SD, MC error, median, and confidence intervals, have been determined. It should be observed that although other parameters display roughly uniform distributions, parameters like the BIW display a skewed distribution pattern.

Health care expenditures are soaring as a result of the current economic crisis. It's crucial for medical researchers and service providers to quickly pinpoint the factors that increase a population's risk of contracting various diseases. The objective is to locate and offer preventive treatments without considerably raising management costs. Predictive modeling is currently a well-liked method for evaluating high-risk situations at extremely low prices. Predictive modeling will be extremely helpful to researchers and healthcare professionals in order to enhance current healthcare services and lower future healthcare expenses. Predictive modeling is a procedure that can be used using healthcare data that is already available, for example, to identify persons who have a high medical demand and who are "at risk" for utilizing medical services above average in the future. On the basis of historical patient data acquired in the past, we are developing a unique Bayesian algorithm that can forecast the breast cancer survival days. The Bayesian prediction technique is becoming increasingly well-liked and is finding new uses in a variety of disciplines, including business and economics, social sciences, engineering, environmental science, and health sciences. In the health sciences, survival research studies are designed and analyzed using the Bayesian predictive technique. It is frequently employed to lower healthcare expenses and distribute resources wisely.

CONCLUSION

Four different types of statistical probability models were applied to the data on cancer survival among Black and Hispanic women. In comparison to other commonly used models, the exponentiated exponential model was determined to be the best fitted model to the Black Hispanic female cancer survival data. It was discovered that the shape of the future survival model for Black Hispanics is positively skewed after the results of the predictive inference under the fitted model were obtained. These models aid academics and healthcare professionals in forecasting a patient's future survival outcomes based on the patient's reported present and prior circumstances. So, it is possible to strengthen and improve the justifications for better usage of current facilities and planned allocation of future resources by combining current knowledge and future predictions. Using the SPSS program 19.0, descriptive statistics were obtained. The "Google fusion table" was used to create the geographic maps of the nine randomly chosen states out of a total of twelve states. Basic summary data for the Black Hispanic subset's breast cancer survival times were obtained using SPSS version 19.0 software [39]. We utilized cutting-edge computational software named "Mathematica version 8.0" [40] to provide the graphical representations of the predictive density for a single future answer for Black Hispanic women. Using the same method, we were able to forecast extra survival durations for the ethnicity. We performed all necessary computations, checked the goodness of fit tests, summarized the posterior parameter values, calculated the parameter kernel densities, and more using the Windbags software.

REFERENCES:

[1] A. Kühberger, A. Fritz, E. Lermer, and T. Scherndl, “The significance fallacy in inferential statistics Psychology,” *BMC Res. Notes*, 2015, doi: 10.1186/s13104-015-1020-4.

[2] B. G. Gibbs, K. Shafer, and M. J. Dufur, “Why infer? The use and misuse of population data in sport research,” *Int. Rev. Sociol. Sport*, 2015, doi: 10.1177/1012690212469019.

[3] T. Ledermann and D. A. Kenny, “A toolbox with programs to restructure and describe dyadic data,” *J. Soc. Pers. Relat.*, 2015, doi: 10.1177/0265407514555273.

[4] U. Simonsohn, J. P. Simmons, and L. D. Nelson, "Specification Curve: Descriptive and Inferential Statistics on All Reasonable Specifications," *SSRN Electron. J.*, 2015, doi: 10.2139/ssrn.2694998.

[5] Z. J. Zimmermann, E. E. Watkins, and A. Poling, "JEAB Research over Time: Species Used, Experimental Designs, Statistical Analyses, and Sex of Subjects," *Behav. Anal.*, 2015, doi: 10.1007/s40614-015-0034-5.

[6] S. C. Mat, S. M. Maat, and N. Mohd, "Identifying Factors that Affecting the Entrepreneurial Intention among Engineering Technology Students," *Procedia - Soc. Behav. Sci.*, 2015, doi: 10.1016/j.sbspro.2015.11.135.

[7] H. Yang *et al.*, "Appraisal of within- and between-laboratory reproducibility of non-radioisotopic local lymph node assay using flow cytometry, LLNA: BrdU-FCM: Comparison of OECD TG429 performance standard and statistical evaluation," *Toxicol. Lett.*, 2015, doi: 10.1016/j.toxlet.2015.02.014.

[8] J. Wolstenholme, "A Statistical Primer: Understanding Descriptive and Inferential Statistics," *Evid. Based Libr. Inf. Pract.*, 2015.

[9] S. Ghavifekr and W. A. W. Rosdy, "Teaching and learning with technology: Effectiveness of ICT integration in schools," *Int. J. Res. Educ. Sci.*, 2015, doi: 10.21890/ijres.23596.

[10] S. P. De La Cruz and M. V. Abellan, "Professional burnout, stress and job satisfaction of nursing staff at a university hospital," *Rev. Lat. Am. Enfermagem*, 2015, doi: 10.1590/0104-1169.0284.2586.

CHAPTER 11
QUANTUM STATISTICS AND THERMODYNAMIC ENTROPY FOR STOCK MARKET NETWORKS

Rohaila Naaz, Assistant Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id-rohailanaaz2@gmail.com

ABSTRACT:

The stock market is a dynamic system made up of complex connections between many types of financial organizations, including banks, firms, and institutions. The network structure can be used to describe such an intricate interactive system. The stock exchange's fundamental mechanism creates a network of businesses and persons that changes over time, which is what drives stock prices to correlate in time-sequential exchanges. Here, we create a brand-new quantum statistical method to examine the development of the financial markets. Starting with a heat bath analogy, the Hamiltonian operator of the network is represented by the normalized Laplacian matrix. The Hamiltonian's eigenvalues define the network's energy states. These states are either filled by fermions with matching Bose-Einstein and Fermi-Dirac statistics or by indistinguishable bosons. We create the thermodynamic entropy using the pertinent partition functions to investigate dynamic network characterizations. We run tests to put this cutting-edge technique to use in locating the substantial variations in network topology that occurred during the financial crisis. The thermodynamic entropy offers a great foundation for describing the changes in the stock market.

KEYWORDS:

Hamiltonian, Fundamental, Organizations, Network.

INTRODUCTION

One of the primary indicators of economic activity on the financial market is typically the stock price. It depicts how each person and business connect with one another. The relationship between various financial institutions is an intricate system that changes throughout time. Researchers from several disciplines are drawn to study the dynamic growth of such a complex system because it shows the financial market's internal workings Tools from complex networks have been used to analyze the time-sequential stock market values in order to characterize such a dynamic system. The majority of existing network techniques typically map time series into the network domain so that it displays the topological and structural characteristics of the system. For instance, the topological space in the correlation coefficients of the economic taxonomy is provided by the hierarchical structure of a minimum spanning tree. The stock market networks' community structure serves as a representation of the structural changes that occurred during the financial crisis. However, the topological structure of the financial networks is the main focus of the majority of the existing research. They just present the general data for a given time. Since there is a considerable correlation between the stock market's time evolution and the statistical features of dynamic networks, this research is important, particularly at this time of financial crisis. Entropic measurement has recently been introduced by a reliable method to quantify network characterization [1], [2].

By way of illustration, the von Neumann entropy provides a qualitative expression for the entropy connected to the degree combinations of nodes forming edges. The goal of this

research is to provide accurate and efficient methods for measuring thermodynamic entropy in time-evolving networks, which serves as the impetus for this kind of analysis. We analyze the New York Stock Exchange stock market networks in particular. We demonstrate that the presence of well-defined increases in thermodynamic entropy characterizes financial disasters, although outside of these critical periods, this characterization is stable for extended periods of time. To do this, we utilize a novel paradigm in quantum statistics pertaining to the generation of partition functions in Bose-Einstein and Fermi-Dirac statistics using the normalized Laplacian matrix. The study of the interactions between thermodynamics and quantum mechanics is known as quantum thermodynamics two distinct theories discuss the fundamental properties of matter and light. In 1905, Albert Einstein stated that the need for consistency between electromagnetism and thermodynamics led to the conclusion that light is quantized, yielding the connection.

The beginning of quantum theory is this publication. In a few decades, quantum theory developed its own set of recognized laws. At the moment, quantum thermodynamics deals with how thermodynamic laws come from quantum mechanics. In that it places more focus on dynamical processes out of equilibrium, it varies from quantum statistical mechanics. In addition, it is hoped that the theory will hold true for a single quantum system. Quantum thermodynamics and the theory of open quantum systems are closely related. Finite-time thermodynamics is supported by quantum mechanics, which adds dynamics to thermodynamics. The primary premise is that because the entire planet is a vast closed system, a global Hamiltonian-generated unitary transformation governs temporal evolution. The global Hamiltonian can be broken down into the following parts for the combined system bath scenario: Adiabatic thermodynamic processes don't change the entropy. Normally, the status is changed by an external control. An externally controlled time dependent Hamiltonian can simulate the quantum version of an adiabatic process [3]–[5].

being constant Therefore, there is no net change in the population of the instantaneous energy levels, which is the quantum adiabatic condition. The Hamiltonian should commute with itself at various moments, according to this. To obtain the final control value, more effort is needed if the adiabatic conditions are not met. This work can be recovered for an isolated system since the dynamics is unitary and reversible. In this situation, as shown in the lab using a unitary Fermi gas in a time-dependent trap, quantum friction can be controlled utilizing shortcuts to adiabaticity. The necessary information to recoup the additional energy cost and reverse the dynamics is carried by the coherence stored in the density operator's off-diagonal elements. Most of the time, this energy cannot be recovered because of interaction with a bath, which results in energy dephasing. In this instance, the bath serves as an instrument for measuring energy. This energy is wasted due to friction at the quantum level.

DISCUSSION

The fundamental tenet of quantum typicality is that a large proportion of all pure states with a common expectation value of a certain generic observable at a particular time would produce relatively comparable expectation values of the same observable at any subsequent time. This is intended to apply to high-dimensional Hilbert spaces with Schrödinger type dynamics. As a result, the ensemble average usually does a good job of describing the individual dynamics of expectation values. John von Neumann's quantum ergodic theorem is a powerful conclusion drawn from the fundamental mathematical framework of quantum mechanics. The QET is a precise description of what is known as normal typicality, which is the idea that every initial wave function exists for typical large systems. The micro-canonical density matrix is macroscopically identical to statistically implausible state changes can be quantified by the second law of thermodynamics to the point where they are essentially excluded. Quantum

thermodynamics resource theory is a description of thermodynamics in the regime where it may be applied to a small number of particles interacting with a heat bath. The second law often applies to systems made of many particles interacting. The second law for microscopic systems takes on a fundamentally different shape than it does at the macroscopic scale for processes that are cyclic or very near to cyclic, putting not just one limitation on what state changes are feasible but an entire family of constraints. These second laws apply to individual macroscopic systems engaging via long-range interactions as well as tiny systems, which only generally meet the usual second law. The laws of thermodynamics take on a form by being more precise in their definition of thermal operations, with the first law defining the category of thermal operations, the zeroth law emerging as a singular condition guaranteeing the theory is nontrivial, and the remaining laws being a monotonicity property of generalized free energies.

Set of stock market data

In the 12th century in France, and in the 13th century in Bruges and Italy, stock exchanges first appeared. In ancient days, information on deals was probably recorded by scribes and sent by messenger. Reuters used carrier pigeons to send data between Germany and Belgium in the early 19th century. Early exchanges in London were situated close to coffee shops, which might have influenced trading. Young men known as "runners" in New York posted prices between broker's offices and the exchange in the late 1860s, and frequently, these prices were written by hand on enormous chalkboards inside the buildings. Many traders began their careers in the financial markets by updating chalkboards, and as described in the book Reminiscences of a Stock Operator, those updating the boards would wear fur sleeves to prevent mistakenly erasing values [6], [7].

The "Big Board" moniker for the New York Stock Exchange may have originated from these substantial chalk boards. Some nations continued to utilize these chalkboards up until recently. Up until 1967, dealers in Chicago sent information to clerks known as "board markers" using the Morse code. A stock telegraph printing device was created in 1863 by Edward A. Calahan of the American Telegraph Company. It allowed information on stocks, bonds, and commodities to be delivered directly from exchanges to broker offices across the nation. On 0.75-inch (1.9-cm) wide paper tape wrapped on substantial reels, the data was produced. It was given the term "stock ticker" because of the noise it made when printing. Later, Thomas Edison patented a "universal stock ticker" and sold over 5,000 of them in the late 19th century after other inventors made improvements to it.

Early in the 20th century, Western Union gained the rights to a better ticker that could handle the daily increase in the number of equities being sold. Trading volumes were so high at the time of the stock market crash in October 1929 that the tickers lagged, adding to the fear. The New York Quotation Stock Ticker gained popularity in the 1930s. In 1960, a new enhancement was implemented. Trans-Lux Corporation invented a back projection device in 1923 that allowed everyone in a brokerage office to view the ticker's movement on a screen. As a result of its enormous success, there were more than 1400 stock-ticker projectors in the United States and 200 more in Canada by 1949. A Trans-Video system dubbed CCTV, which gave a customer a small video desk display where he could watch the tickers, began selling in 1959.

Electrician, the first wall-mounted all-electric ticker display system, was introduced by Optronics in August 1963. Over 1100 units were running in stock broker offices in the United States and Canada by 1964. Trans-Lux launched the Trans-Lux Jet in response to competition, including Optronics' Electrician electron wall technology. Lighted disks that

moved on a belt on the broker's wall were controlled by air jets. By the middle of 1969, more than 3000 units had been ordered by brokers and were in use throughout the United States and Canada.

Quotron vs. Aptronics

Robert S. Sinn was drawn in by Quotron's success and noted one of its shortcomings: it could only provide a last price. There was no information available regarding the opening price, daily high and low, or share volume. The ticker transmissions from the various stock and commodity exchanges were received by his system. A hardwired digital computer then automatically translated them, computing and updating each stock's highs, lows, and total volume while simultaneously updating a drum memory with the last selling prices. A data packet would be generated as these elements were updated and sent over AT&T Dataphone at 1000 bits per second to identical magnetic drum storage devices in each major American city. Desk units in nearby stock brokerage offices may then access these slave drum memory units that were situated in the US's major metropolitan areas utilizing Dataphone communication.

The desk units would mechanically activate micro switches to set the ticker symbol code for the desired stock. Each desk unit would be continuously queried in order by the local control box, which would then send a request data packet via Dataphone to the nearby drum memory, which would then transmit reply data packets back to the neighborhood brokerage office. The request stock alphabetic symbols and a desk unit identifier would be included in each packet for both the request and the response. The control unit could complete the interrogation of all desk units in the office in about every one or two seconds because the desk units set up the requested stock code statically, so the desk unit would automatically update the stock price, volume, and highs and lows without any operator intervention. This is the first use of data packet transmission with the sender's identification imbedded in the data packet in order to avoid switching- a forerunner of the interoffice protocol. The entire system operated only on data packets including sender identification; there was no switching at all [8]–[10].

Sinn established Ultronic Systems Corp. in December 1960, serving as its president and CEO until his resignation from the organization in 1970. Ultronics had installed its first desk units (Stockmaster) by the fall of 1961 in New York and Philadelphia, with San Francisco and Los Angeles following. The Stockmaster desk devices gave the user instant access to and constant monitoring of information on each stock's last sale, bid, ask, high, low, total volume, open, close, earnings, and dividends on the NYSE and AMEX, as well as commodities from the different U.S. commodity exchanges. Eventually, 10,000 devices were installed globally by Ultronics and General Telephone (who acquired Ultronics in 1967). A joint venture agreement was signed in June 1964 by Ultronics and the British news organization Reuters to promote Stockmaster outside of North America. This endeavor, which lasted for ten years and was tremendously successful, dominated the global market for data on U.S. stock and commodity prices. Time division multiplex technology was developed by Ultronics to send U.S. stock and commodity information along with Reuters' teletype news channels across Reuters' voice grade lines to Europe and the Far East. The last sale and the bid and ask ticker lines were the only real-time information available from the various stock and commodity exchanges at the time these first desk top quotation systems were established in the early 1960s. Each trade was listed in the last sale ticker, along with its price and volume. There were only two prices and no size in the bid-ask ticker. Therefore, the last sale ticker had far more data than the bid-ask ticker did. As a result, on days with strong trading activity, the final sale ticker would lag the bid-ask ticker by up to fifteen minutes. Even though there was an additional fee to the exchange for this information, the stockbroker found the bid-ask on their desk top unit to be of utmost importance due to the time difference.

Desk unit for a Quotron

The Optronic danger prompted an urgent response from Scantlin Electronics. They started developing their own computer-based system in early 1962 and launched it in December of that year. In New York, it made use of four Control Data CDC 160A computers, which stored trading information in magnetic core memory. A broker could request the price and net change from the opening for any stock, as well as a summary that included highs, lows, and volumes, on newly designed Quotron II desk units in brokerage offices in major cities (later SEI added other features like dividends and earnings). The queries were sent to a central office, which consolidated and sent them on to the computer in New York. Reversed order of responses came next. The high-speed telephone service Dataphone from AT&T was used to send the data. Numerous brokers adopted the new system in 1963, and it was placed in hundreds of their locations.

The same technology delivered stock market pages to United Press International at the end of each day, and it then sold those pages to its newspaper clients all over the world. In 1961, Ultronics released their Stockmaster desk units, pricing the service at a level that was comparable to that of the Quotron desk units. Only a performance-based competition was what they desired. All of these stock quote gadgets were offered for sale as rentals with monthly fees. Therefore, the seller alone bore the expense of the system, desk units, and installation, not the client (broker). The units were quite profitable at the time they were priced, allowing the businesses to finance the cost and take advantage of quick accounting depreciation on the machinery. In comparison to Stockmaster or Quotron, Teleregister'sTelequote desk units were debuted in 1964 at much lower prices. Ultronics and Scantlin were compelled to lower the cost of their Stockmaster and Quotron systems as a result. Although the Telequote desk units were never successful in gaining a sizable portion of the desk top quotation market, their price reductions significantly lowered the industry's overall profitability in the United States. The joint venture Ultronics formed with Reuters for the stock quotation industry outside of North America, where this price reducing was not an issue, proved fortunate.

Ultronics was denied a historic first due to the Cuban Missile Crisis in 1962. In order to transmit television and telephone voice channels between the United States and England and France, AT&T launched the first commercial satellite (Telstar) in July 1962. This non-synchronous satellite had an elliptical orbit around the globe that took it roughly 2.5 hours to complete, giving the United States and Europe only about 20 minutes of contact on each pass. In October 1962, Ultronics made a deal with AT&T to send American stock prices to Paris over one of its voice channels. All of the American stockbrokers, the press, and television were prepared for this historic moment. A Stockmaster unit was put in the Bache brokerage office on Rue Royale. The stock transmission was canceled about two hours prior to the pass because U.S. President John F. Kennedy intended to use the pass to send his speech to France over the Cuban missile crisis [11], [12].

CONCLUSION

We are currently doing studies on the evolutions of the stock market network using thermodynamic entropy. This offers an insightful characterization for examining stock market volatility. We first look into how well this sort of entropy can identify the structural network variance in time series. depicts the New York Stock Exchange in the Bose-Einstein and Fermi-Dirac statistics thermodynamic entropy. Sharp peaks in the time sequential data show where important financial events, including Black Monday, the mini-crash on Friday the 13th, the Early 1990s Recession, the 1997 Asian Crisis, the 9.11 Attacks, the Downturn of

2002–2003, the 2007 Financial Crisis, Lehman Brothers' bankruptcy, and the European Debt Crisis, occurred. Each financial crisis highlights the large entropy variation linked to abrupt network structure changes. We use the recession of 2002–2003 as an illustration. Investors lost faith in the American economy following the September 11 attacks because of terrorism. In the aftermath, numerous Internet businesses failed. This led to the restatement of results by a significant number of firms and restored investor confidence This significantly changed how stocks interacted with one another and significantly changed the market's overall structure..

REFERENCES:

[1] Y. Zheng and D. Poletti, "Quantum statistics and the performance of engine cycles," *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 2015, doi: 10.1103/PhysRevE.92.012110.

[2] A. Marchewka and E. Granot, "Role of quantum statistics in multi-particle decay dynamics," *Ann. Phys. (N. Y).*, 2015, doi: 10.1016/j.aop.2015.02.022.

[3] D. Zhao, Y. Gu, H. Chen, J. Ren, T. Zhang, and Q. Gong, "Quantum statistics control with a plasmonic nanocavity: Multimode-enhanced interferences," *Phys. Rev. A - At. Mol. Opt. Phys.*, 2015, doi: 10.1103/PhysRevA.92.033836.

[4] S. V. Mousavi, A. S. Majumdar, and D. Home, "Effect of quantum statistics on the gravitational weak equivalence principle," *Class. Quantum Gravity*, 2015, doi: 10.1088/0264-9381/32/21/215014.

[5] S. Hartmann, F. Friedrich, A. Molitor, M. Reichert, W. Elsäßer, and R. Walser, "Tailored quantum statistics from broadband states of light," *New J. Phys.*, 2015, doi: 10.1088/1367-2630/17/4/043039.

[6] A. Javadi *et al.*, "Single-photon non-linear optics with a quantum dot in a waveguide," *Nat. Commun.*, 2015, doi: 10.1038/ncomms9655.

[7] A. Crespi *et al.*, "Particle statistics affects quantum decay and Fano interference," *Phys. Rev. Lett.*, 2015, doi: 10.1103/PhysRevLett.114.090201.

[8] O. P. D. S. Neto, M. C. Deoliveira, and G. J. Milburn, "Temperature measurement and phonon number statistics of a nanoelectromechanical resonator," *New J. Phys.*, 2015, doi: 10.1088/1367-2630/17/9/093010.

[9] A. G. A. Pithis and H. C. Ruiz Euler, "Anyonic statistics and large horizon diffeomorphisms for loop quantum gravity black holes," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.*, 2015, doi: 10.1103/PhysRevD.91.064053.

[10] C. H. Woo, H. Wen, A. A. Semenov, S. L. Dudarev, and P. W. Ma, "Quantum heat bath for spin-lattice dynamics," *Phys. Rev. B - Condens. Matter Mater. Phys.*, 2015, doi: 10.1103/PhysRevB.91.104306.

[11] T. Kazimierczuk *et al.*, "Photon-Statistics Excitation Spectroscopy of a Quantum-Dot Micropillar Laser," *Phys. Rev. Lett.*, 2015, doi: 10.1103/PhysRevLett.115.027401.

[12] D. Fehrenbacher *et al.*, "Free-running performance and full control of a passively phase-stable Er:fiber frequency comb," *Optica*, 2015, doi: 10.1364/optica.2.000917.

CHAPTER 12
MODELING ERROR STATISTICS FOR COLD-FORMED STEEL COLUMNS

Ramesh Chandra Tripathi, Professor
College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India
Email Id- rctripathig@gmail.com

ABSTRACT:

In this article, an effort has been made to estimate the Modelling Error (ME) linked to compression capacity models for various failure modes of compression members constructed from Cold-Formed Steel lipped channel sections that are accessible in international standards. For the first time, a database for the compression capacities of CFS lipped-channel sections has been constructed utilizing test findings that have been published in the literature. The database provides information on 273 compression member tests that failed in a variety of ways, including yielding and flexural, torsional, flexural-torsional, local, and distortion buckling. The database contains just those sources that provide all the information necessary to calculate capacities using various standards. This test database also contains the findings of experimental studies conducted at the CSIR-Structural Engineering Research Centre, Chennai. international codes of practice were utilized to calculate the compression capabilities of the database columns taken into consideration in this research. Different design recommendations for failure modes are provided by design standards. For instance, the) standard sets strict restrictions for the maximum width to thickness ratio for stiffened and unstiffened parts. standard examines combined local and distortional buckling mode. As a result, instructions for the distortional buckling mode are not supplied. Additionally, depending on the design criterion and failure mechanism, the sample size for each design standard varies. Studies on the statistical analysis of ME point to the substantial difference in compression capacity forecasting models for the flexural-torsional buckling mode, regardless of the design standard. Similar findings are reported for the standards' distortional buckling models.

KEYWORDS:

Compression, Database, Engineering, Information.

INTRODUCTION

The partial safety factors available in design standards are ignored when evaluating the compression capacities for test database parts. The significance of taking ME into account as a random variable is highlighted by the probabilistic analysis used to estimate the statistical properties of compression capacity. The ME results will therefore be helpful in studies of code calibration and may have potential application to design practice.The demand for power supply has multiplied due to the exponential growth experienced across all sectors, including industrial, housing, transportation, communication, and services. According to Juette and Raffaella the transmission line towers must hold heavier and larger conductors in bundle arrangement to meet this increased demand]. As a result, the tower configurations get taller and wider, which makes the tower system heavier. Traditionally, Hot-Rolled Steel (HRS) angle sections have been used to construct the transmission line towers. It is required to either develop lighter steel sections with higher strengths or take into account eco-friendly alternate sustainable materials (such as GFRP sections) in order to lower the tower weights. CFS

sections are typically employed to increase structural member and structure strength without sacrificing overall drift requirements [1], [2].

Engineers can create affordable transmission line tower designs because to the benefits of CFS membership that have been outlined above. The CFS members can be manufactured to perfectly match the design specifications. Gaylord and Wilhoite made recommendations for the design of plain and lipped angles created by the cold-forming technique for transmission towers. are a few more international standards that offer design guidance for the usage of CFS sections. According to geometrical, material, and boundary condition details, these codes offer equations or models for the prediction of the compression and tension capabilities of CFS members failing in various modes. The estimation of ME related with equations defined in standards for compression capacity assessment of CFS members is the primary subject of this research. The following are some of the highlights of the international design standards taken into consideration.Design guidelines for CFS sections for transmission line towers are provided under the standard. The CRC curve is used in this standard to estimate compression capacity without incorporating any partial safety factors.

The models utilized in the calculation of structural reactions are typically not full and perfect, hence the real results cannot be anticipated without mistake, according to the JCSS probabilistic model code. As a result of the mathematical relations' simplification, the variables employed in model functions are linked to potential uncertainties, which take into account random influences that are ignored in the models. This work focuses on establishing the ME related with the compression capacity equations found in regulations, as has already been mentioned. A database of experimental column (pinned terminated) capacities made of CFS lipped channel sections breaking in various modes is first established in order to estimate the same. In addition to the test results of three ostensibly comparable columns evaluated at CSIR-SERC, this database also contains experimental findings for CFS lipped channel columns that were taken from the literature. Here we provide a brief review of the literature on column testing, probabilistic analysis, and reliability index determination for CFS lipped channel sections [3], [4].

In order to investigate the effects of residual stress, cross-sectional dimensions, yield type (gradual or sharp) of stress strain curve, forming methods (press brake or cold rolling), local buckling, load application (concentric or eccentric), etc. on CFS lipped channel columns, a number of researchers have conducted experiments: Young and Rasmussen are just a few of the authors who have written about Weng and Pekoz. In order to create the database, the compression capacity results for CFS lipped channel portions from the aforementioned experimental experiments were employed. According to the literature analysis, recent research has focused on experimental and analytical studies of built-up CFS channel sections under axial compression. To investigate the buckling behavior of back-to-back built-up CFS channel sections covering stub columns to thin columns, Ting et al. conducted numerical and experimental research. LVDT equipment with a 0.01 mm accuracy was used to measure the geometric defects, which were then taken into account in the finite element modeling. It has been noted that for short, intermediate, and slender columns that failed as a result of a combination of local and global buckling and/or global buckling, the FEA and test findings were in good agreement and conservative with the estimated strengths as per AISI and AS/NZS standards. Experimental compression tests on built-up, closed-section members made of intermediately welded CFS channel sections were carried out by Whittle and Ramseyer, and the test capacities were compared to theoretical buckling capacities based on the AISI standards adjusted slenderness ratio. The modified slenderness ratio was applied with extreme caution. The slenderness ratio provisions' unmodified capacities were less

conservative. In order to better understand how local and global flexural buckling interact, Ye et al. resented the findings of experimental and numerical studies performed on CFS plain and lipped channels under axial compression. The creation of finite element models uses the measured initial geometric flaws and material attributes. The compression capacity determined using the direct strength method and Eurocode was compared to the experimental and finite element results. The compressive capacity of plain and lipped channel sections is conservatively predicted by the Eurocode, however lipped channel forecasts by the direct strength approach are more accurate. These studies, according to some reports, are the most recent ones.

Studies on the probabilistic analysis of the strength of steel defective columns were conducted by Balaji Rao and Appa Rao. This research has led to the proposal of a characteristic strength equation that can be used to rationally construct defective columns. The researchers used formulae/model functions from in several research reports from the University of Missouri- and Supomsilaphachai et al, the determination of reliability index for CFS elements or members is presented. Both the fundamental research data and the reliability index inherent in the AISI Specification are presented in great detail. Ellingwood et al. Galambos et al., and Ellingwood et al. reanalyzed the entire set of data for HR steel and CFS using (a) updated load statistics and a more sophisticated level of probability analysis that was able to incorporate probability distributions and more accurately describe the true distributions [5], [6].

The above-mentioned assessment of the literature reveals that there is few research that deal with ME estimate for CFS compression individuals. Care must be taken, though, to avoid using any partial factors that may be embedded in the coal equation when estimating the ME. The main objectives of this paper are creation of database for compression capacities of CFS lipped-channel sections from the test results available in literature estimation of ME for database results for the various failure modes as per standards studies on statistical characteristics of ME estimated at fitting a statistical distribution for the ME. The following part provides information on the experimental studies conducted at the CSIR-Structural Engineering Research Centre, Chennai, that were used to create the test database.

DISCUSSION

On lipped channel sections measuring LC 90 50 20 3.15 mm (Lipped Channel Web depth Flange width Lip depth Thickness, element-level compression tests were performed. The experiments fell under the category of concentrically loaded members, which correspond to the latticed tower panel's leg members. The 250 Ton displacement control UTM with spherical joints and a ball bearing head at the top end and a fixed base at the bottom was used for the tests. The manufactured CFS lipped channel specimens were 2.225 m in length and joined with custom fixtures at the ends utilizing three different 16 mm bolts. In order to accurately replicate the member end condition for the leg member of an X-braced panel, this is done. Using displacement dial gauges (0.01 mm least count), the lateral displacements of the flanges and the web were measured during the testing. To evaluate the longitudinal strain fluctuations throughout the load application, the specimens' cross sections were instrumented with linear foil strain gauges. Using an HBM data logger, the test results for dial and strain gauges were recorded. The specimens were instrumented at the maximum lateral displacement portions, or the middle of the specimens, in order to assess the stresses and displacement responses at important locations [7]–[9].

Given that the specimen's thickness is 3.15 mm and the end fittings are formed of 16 mm thick HRS plates, the end fixtures are more rigid than the specimen. Therefore, it is believed

that the test specimen's end condition will function as pin ends. As a result, the end connection's center to center is taken into account while determining the effective length for analysis. The observed test capacities are with an average compression capacity of the test specimens failed in flexural-torsional buckling mode. The capacities computed based on the assumed effective length and in accordance with are well in line with the capacity tested. This finding demonstrates the suitability of the effective length calculated under pinned end conditions. It should be emphasized that it can be challenging to precisely determine effective length. The capacity estimates are derived using the design standards under consideration for an effective length equal to the bolted connection's center-to-center distance without taking into account the safety or partial safety factors that are included in the model equations.

Methodology for Modeling Error Estimation

For some of the factors included in the prediction equation, the mathematical relations or model functions utilized to forecast the capacity of the compression component are based on simplifying assumptions and/or disregarded random influences. As a result, the model function might not accurately capture the relevant physical phenomenon. This, together with the modeler’s incomplete knowledge, is what accounts for the discrepancy between the expected and real compression member capacities. Using ME analysis, the capacity differential can be put into numbers. In light of this, the ME should be described as the proportion of calculated to real capacity. Even so, it still provides a close estimate to the genuine capacity. It's likely that the test capacity is not the true capacity because there may be some inevitable flaws involved in the test and measurement procedure. In this study, the ME is defined as the ratio of the compression member's calculated capacity to its test capacity. Additionally, model functions are taken into account in ME analysis for a variety of modes of failure, including flexure, flexure-torsion, local and distortion buckling, and yielding of sections based on ASCE standards as described in the preceding section.

Accordingly, the member compression capacity has been calculated for the balance test database members for governing failure modes given in each design standard after removing the safety or partial safety factors present in the model function, after passing through the filtering criteria as discussed in "Database of Compression Capacity for CFS Lipped Channel Members". The ME then calculated an estimate using the ratio of the database's test capacity to the calculated capacity. The estimated ME findings and its statistical features for the test database compression participants in various failure scenarios in accordance with the aforementioned design requirements.Flexure, flexure-torsion, local and distortion buckling, and yielding are the failure modes taken into account for column members. Local and distortion buckling and yielding are the failure mechanisms for stub columns [10], [11]

Compression Capacity Analysis Using Probability

The purpose of this section is to emphasize how crucial it is to take ME into account when assuming the distribution of compression capacity. Simulating compression capacity with and without ME as the random variable allows researchers to examine how ME affects that capacity. The material parameters of CFS are the yield strength, Fu, ultimate strength, and elastic modulus, E, which are employed as random variables in the compression capacity calculation of CFS columns. The lognormal distribution for Fee and E and the normal distribution for Fu for high-tensile steel plates were proposed by Ravindra and Galambos [59] and Hess et al. [60]. According to the aforementioned studies, Fy's coefficient of variation standards, simulation of 106 cycles is used in the present studies to estimate the statistical properties of compression capacity of a typical CFS lipped channel column (database column no. 18), with and without taking ME into account as a random variable along with the other

random variables. The column is chosen so that its mean ME and COV are both greater than 1.15 and 0.25, respectively. In the flexure torsion mode of buckling, the column is crucial according to the aforementioned requirements. show the histograms and statistical characteristics of the simulated compression capacity for the aforementioned standards evaluated with and without ME as the random variable. Additionally, a Chi-square goodness-of-fit test on a simulated compression capacity was run to predict statistical distributions.

Skewness

All of the buckling modes' nonzero values of skewness suggest a non-Gaussian distribution for the ME data. Additionally, the majority of skewness values are positive, indicating a positively skewed or right-skewed distribution, in which the majority of values are concentrated near the left tail of the distribution while the right tail of the distribution is longer. Negative skewness values are seen for a select few failure modes, including local buckling of stub columns for the ASCE 10-15 standard, yielding of stub columns for the AISI S100-16 & AS/NZS 4600:2018 standard, and local and distortional buckling of columns for the EN 1993-1-3:2006 standard. The detailed examination of test data for these failure modes reveals that there are fewer data points in the sample and that more experimental research is needed to obtain the accurate statistics.

Kurtosis

Kurtosis values are found to be greater than three for the failure modes of flexure and flexure-torsion buckling for columns regardless of the design standard, as well as local buckling mode for column and stub columns as per standard and combined local and distortion buckling mode as per EN 1993-1-3:2006 standard. The positive excess kurtosis in these circumstances suggests that the ME data have a non-Gaussian distribution and are leptokurtic with heavy tails and extreme values. metheg data is platykurtic, has flat tails, and has a low likelihood of extreme values when the kurtosis values for other types of failure are close to three with negative excess kurtosis. The statistical distribution for ME data is no Gaussian in this instance as well.

Distribution

To ascertain the makeup of the probability distribution of ME, chi-square goodness-of-fit tests are carried out. For the test, three alternative fictitious distributionsNormal, Lognormal, and Uniform distributions are taken into consideration. displays the outcomes of these examinations. According to the Chi-square goodness-of-fit test results, the hypothesis of the assumed distributions taken into consideration cannot be ruled out at the 1% and 5% level of significance. For the specific mode of failure, a distribution with the lowest Chi-square value is taken into account. In some circumstances, the normal distribution is controlling, but when nonzero skewness values and kurtosis values less than 3.0 are taken into account, the lognormal distribution is assumed. The assumption of a lognormal distribution for ME is also supported by the fact that negative values for ME, no matter how small, have no engineering significance. Since the sample size is limited, it may be anticipated that the ME for the failure modes will follow a uniform distribution.

CONCLUSION

The compression capacity of a set of lipped channel compression members is predicted using the model functions provided for various modes of failure in standards. The literature is where the test results for compression members are found. A database is established for the first time and reported in this study using the test findings from this experiment as well as

those from three nominally identical compression members that were also evaluated at CSIR-SERC. Each test compression member has enough data to calculate its compression capacity in accordance with the aforementioned requirements. ME analysis was conducted using the ratio of the test to forecast compression capacity to evaluate the model functions' accuracy in calculating compression capacity for the various failure modes present in these standards and to propose probability distributions for the ME. The following is a succinct summary of the statistical study' findings It can be concluded from the results shown in Tables 6 and 7 that the ME for different failure modes, with the exception of the cases when the number of data points is less, follows a lognormal distribution at a 5% significance level, with means about equal to 1.10 and. To counteract the effect of the standard's overestimation of the shield's capacity, a greater value of standard appears to function adequately in the assessment of compression member capacity generally. The results of this investigation's measurements of the coefficient of skewness and kurtosis further support the usage of the lognormal distribution, an asymmetrical distribution about the mean.

REFERENCES:

[1] R. Stanev, "Early stopping of RCTs: two potential issues for error statistics," *Synthese*, 2015, doi: 10.1007/s11229-014-0602-3.

[2] Y. He, O. S. Salih, C. X. Wang, and D. Yuan, "Deterministic process-based generative models for characterizing packet-level bursty error sequences," in *Wireless Communications and Mobile Computing*, 2015. doi: 10.1002/wcm.2356.

[3] G. Descombes, T. Auligné, F. Vandenberghe, D. M. Barker, and J. Barré, "Generalized background error covariance matrix model (GEN-BE v2.0)," *Geosci. Model Dev.*, 2015, doi: 10.5194/gmd-8-669-2015.

[4] J. Burger, A. van Delden, and S. Scholtus, "Sensitivity of mixed-source statistics to classification errors," *J. Off. Stat.*, 2015, doi: 10.1515/JOS-2015-0029.

[5] T. E. Nightingale, J. P. Walhin, D. Thompson, and J. L. J. Bilzon, "Influence of accelerometer type and placement on physical activity energy expenditure prediction in manual wheelchair users," *PLoS One*, 2015, doi: 10.1371/journal.pone.0126086.

[6] S. T. Schultz, C. Kruschel, T. Bakran-Petricioli, and D. Petricioli, "Error, power, and blind sentinels: The statistics of seagrass monitoring," *PLoS One*, 2015, doi: 10.1371/journal.pone.0138378.

[7] C. Providakis *et al.*, "A new damage identification approach based on impedance-type measurements and 2D error statistics," *Struct. Monit. Maint.*, 2015, doi: 10.12989/smm.2015.2.4.319.

[8] L. Mitchell and A. Carrassi, "Accounting for model error due to unresolved scales within ensemble Kalman filtering," *Q. J. R. Meteorol. Soc.*, 2015, doi: 10.1002/qj.2451.

[9] A. Berchet, I. Pison, F. Chevallier, P. Bousquet, J. L. Bonne, and J. D. Paris, "Objectified quantification of uncertainties in Bayesian atmospheric inversions," *Geosci. Model Dev.*, 2015, doi: 10.5194/gmd-8-1525-2015.

[10] Z. Zang *et al.*, "Background error statistics for aerosol variables from WRF/Chem predictions in Southern California," *Asia-Pacific J. Atmos. Sci.*, 2015, doi: 10.1007/s13143-015-0063-8.

[11] A. Gagin and I. Levin, "Accounting for unknown systematic errors in Rietveld refinements: A Bayesian statistics approach," *J. Appl. Crystallogr.*, 2015, doi: 10.1107/S1600576715011322.