Forum for Interdisciplinary Mathematics

Ashis SenGupta
Barry C. Arnold *Editors*

# Directional Statistics for Innovative Applications

## A Bicentennial Tribute to Florence Nightingale

FORUM FOR INTERDISCIPLINARY MATHEMATICS
FIM
शास्त्राणां गणितम् मूर्धनि स्थितम्
MATHEMATICS IS AT THE ZENITH OF ALL DISCIPLINES

Springer

# Forum for Interdisciplinary Mathematics

The **Forum for Interdisciplinary Mathematics** is a Scopus-indexed book series. It publishes high-quality textbooks, monographs, contributed volumes and lecture notes in mathematics and interdisciplinary areas where mathematics plays a fundamental role, such as statistics, operations research, computer science, financial mathematics, industrial mathematics, and bio-mathematics. It reflects the increasing demand of researchers working at the interface between mathematics and other scientific disciplines.

More information about this series at https://link.springer.com/bookseries/13386

Ashis SenGupta · Barry C. Arnold
Editors

# Directional Statistics for Innovative Applications

## A Bicentennial Tribute to Florence Nightingale

Springer

*Editors*
Ashis SenGupta
Department of Population Health Sciences,
MCG
Augusta University
Augusta, GA, USA

Indian Statistical Institute
Kolkata, WB, India

Department of Statistics
Middle East Technical University
Ankara, Turkey

Barry C. Arnold
Department of Statistics
University of California
Riverside, CA, USA

ISSN 2364-6748 ISSN 2364-6756 (electronic)
Forum for Interdisciplinary Mathematics
ISBN 978-981-19-1043-2 ISBN 978-981-19-1044-9 (eBook)
https://doi.org/10.1007/978-981-19-1044-9

Mathematics Subject Classification: 62-XX, 62H11, 62Jxx, 62H05, 62Hxx



This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

*It is by the aid of Statistics that law in the social sphere can be ascertained and codified, and certain aspects of the character of God thereby revealed. The study of Statistics is thus a religious service.*

*Florence Nightingale*
*(12 May 1820–13 August 1910)*

*Source* https://en.wikipedia.org/wiki/Florence_Nightingale

# Preface

The fascinating story of Florence Nightingale, OM, RRC, DStJ, (1820–1910), The Lady with the Lamp, on her contributions to the care of wounded soldiers during the Crimean war and her subsequent efforts to elevate the status of the nursing profession and to modernize and improve institutional care for military participants and retirees, remains vividly implanted in our culture. While recognizing the importance of her work in that direction, we statisticians also celebrate her contributions to the field of Statistics. Her meticulous compilation of data to support her arguments for change in military hospitals was recognized as an important factor in the development of a positive image of Statistics as a science in its own right. The second celebrated statistical contribution of Miss Nightingale was the use of "coxcombs" as per her nomenclature, to illustrate periodic data, e.g., monthly mortality figures over 1854–1856 during the Crimean war. This graphical representation was the genesis of modern-day Rose diagrams representing one of the earliest instances of the analysis of circular data, and a sizeable body of research in that field has subsequently followed her pioneering efforts. Statisticians may thus well acknowledge her as The Lady with the Rose (diagram). The third revered contribution of Miss Nightingale was her indefatigable efforts through non-trivial data collection and statistical presentation toward the improvement of health conditions and mortality rates in rural areas, especially in India, during the British regime. There have been many international and national awards created in her name, including the Florence Nightingale Medal from the International Committee of the Red Cross and the National Florence Nightingale Nurses Award conferred annually by usually the Honorable President of India. As an appropriate statistical celebration of her 200th birthday birth anniversary, we have chosen to publish the present collection of research articles dealing directly or indirectly with circular data. It is hoped that this volume will serve not only to celebrate the work of Florence Nightingale, but will also encourage further theoretical research and applied analyses in circular, and its emerging generalizations to manifold, data settings that we believe would have been recognized by her as being of importance to our society.

This volume consists of both original research and review papers received only through invitations. There are 25 papers authored by 58 renowned researchers spanning 5 continents. Forty-nine eminent scholars kindly served as Referees. These contributors belong to 25 countries—Australia, Belgium, Canada, China, Denmark, France, Germany, Hungary, India, Iran, Italy, Japan, Malaysia, Mexico, Nigeria, New Zealand, Norway, Russia, South Africa, South Korea, Spain, Switzerland, Turkey, UK, and USA.

The 25 papers in this volume have been grouped into eight parts. Part One titled Quality of Life, though it contains but a single paper, seems particularly appropriate to be included first. Mukherjee, in this paper 1, focuses on Florence Nightingale's work related to public health and links it to modern research on quality of life.

There are five papers in Part Two on Innovative Applications. In paper 1, Mardia, Barber, Burdett, Kent and Hamelryck discuss spherical mixture models in the context of protein folding problems. Arnold, Jupp and Schaeben, in paper 2, considers statistics associated with orientation relationships in the study of crystals and pairs of crystals. In paper 3, Hundrieser, Klatt and Munk address issues related to optimal transport in the context of circular distributions and data together with related inference issues. An updated model for forecasting the movement of the magnetic North Pole is developed and investigated in paper 4 by Di Marzo, Panzera, Fensore and Taylor. In paper 5, Ameijeiras-Alonso and Crujeiras discuss the use of flexible parametric families of circular distributions in the analysis of car accident data.

Part Three on Data Visualization, Simulation and Transformations, includes two papers. Jammalamadaka and Terdik, in paper 1, discuss simulation and visualization of spherical distributions in three dimensions. They provide algorithms for simulation in a MATLAB package. In paper 2, Johnson proposes use of a parametric family of transformations that can be used to improve the fit of a von Mises model to particular data sets.

In Part Four, six papers deal with Distribution Theory and Parametric Inference. Bekker, Rad, Arashi and Ley discuss generalized asymmetric distributions on the circle and the torus and compare them with previously available models in paper 1. Gatto, in paper 2, utilizing maximum entropy ideas, highlights the role of the generalized von Mises spectral distribution in the analysis of stationary time series in the frequency domain. Shimizu and Imoto consider distributions constructed via products of cardioid densities. The resulting distributions, described in paper 3, can exhibit asymmetry and multimodality. In paper 4, Abe, Imoto, Shiohama and Miyata discuss new circular, toroidal and cylindrical models. Inference and identifiability issues for the models are discussed. Ong and SenGupta, in paper 5, investigate bivariate cardioid distributions constructed using mixtures. Appropriate inference and tests of simplifying hypotheses are given attention. Kim and Asare-Kumi, in paper 6, discuss inference and outlier detection for a three parameter generalized von Mises distribution that can adapt to model asymmetric data configurations.

The part on Regression Analysis, Part Five, includes four papers. In paper 1, Guttorp and Lockhart investigate how wind speed affects the parameters of a mixture of von Mises distribution. Lagona, in paper 2, introduces autoregressive models for spatial circular data with parameters that are covariate dependent. In paper 3, Zhan, Ma and Liu consider circular data as unit complex numbers and propose a complex multiplication framework for circular regression. Jha and Biswas, in paper 4, survey a wide variety of regression models for directional variables together with discussion of associated inferential and computational issues.

Part Six includes two papers dealing with Non-parametric Inference for directional data. In paper 1, Chaubey reviews available smoothing methods for circular distributions. Use of transformations to adapt linear algorithms for use with circular data is also investigated. Verdebout, in paper 2, considers the use of weighted sign tests for investigating and testing rotational symmetry against skewed alternatives.

In Part Seven, three papers are focused on Time Series and Change-Point Analysis. Beran, Steffens and Ghosh, in paper 1, present results on circular time series exhibiting long range dependence. In paper 2, Ugwuowo discusses a variety of circular time series models and illustrates their use in modeling wind direction data. Potgieter, Lombard and Hawkins, in paper 3, survey statistical process control methodology in the context of circular data for identifying changes in location and/or concentration which might indicate that a process is out of control.

The final Part Eight includes two papers on Statistical Machine Learning. In paper 1, Laha and Majumdar consider neural network models for angular–angular and angular–linear regression. In paper 2, Tugac and Yildirak use deep learning algorithms and circular principle component analysis for forecasting multivariate wind data time series.

We take this opportunity to thank all of the authors who have contributed their work for inclusion in this Florence Nightingale 200 volume. We also acknowledge and profusely thank the scholars who contributed to the volume by serving as referees, including some of them who obliged us with reports on two papers, and graciously devoting their valuable time and sincere efforts for evaluating and farther enriching the volume.

Finally, our thanks go to the editorial staff at Springer Nature for their encouragement of this project and preparation of the final version of the book. As a final word, we hope that this volume with its emphasis on directional statistics will motivate and promote this research area, while serving in some small way to keep evergreen the appreciation for Florence Nightingale's pioneering contributions to it.

Kolkata, India — Ashis SenGupta
Riverside, USA — Barry C. Arnold

# Contents

# Editors and Contributors

## About the Editors

**Ashis SenGupta** is Adjunct Professor at Augusta University, Georgia, USA; Distinguished Professor at Middle East Technical University, Turkey; and Advisor/Consultant at the Indian Statistical Institute, Kolkata, India. He completed his Ph.D. from Ohio State University, USA, and was visiting faculty at renowned universities across the world, including Stanford University, USA; University of California, Riverside, USA; University of California, Santa Barbara, USA; University of Wisconsin, Madison, USA; Michigan State University, East Lansing, USA; Concordia University, Montreal, Canada; Hebrew University of Jerusalem, Israel; Institute of Statistical Mathematics, Tokyo, Japan; Hacettepe University, Ankara, Turkey; and the University of Malaya, Malaysia. He is on the editorial board of the Forum for Interdisciplinary Mathematics (a Springer book series) and has been Editor-in-Chief of two international journals.

Professor SenGupta has supervised the Ph.D. thesis of 14 scholars in India, Turkey and USA, and has authored more than 100 publications, including 12 books and edited volumes. His research interests include big data analytics, directional statistics, distribution theory, financial statistics, multivariate statistical analysis, reliability inference and statistical machine learning. He visited Mainland China as a Citizen Ambassador from the American Statistical Association and is the recipient of several international and national recognitions, including two lifetime achievements and one distinguished statistician awards. He is a member of several Expert, Project Advisory and Monitoring committees of the Department of Science and Technology, Government of India. He served as Vice-President of the Forum for Interdisciplinary Mathematics, USA; President (India chapter, 3 successive terms) of the International Indian Statistical Association, and President of the Mathematical Sciences section of the Indian Science Congress, and is recognized by International Statistical Institute, the Netherlands (Elected Member); Indian Society of Probability and Statistics, India (Fellow); National Academy of Sciences, India (Fellow); and American Statistical Association, USA (Fellow).

**Barry C. Arnold** is Distinguished Professor Emeritus at the Department of Statistics, University of California, Riverside, USA. He received his Ph.D. in Statistics from Stanford University, USA, in 1965. He has authored 14 books and more than 275 research papers in reputed peer-reviewed journals and contributed volumes. Professor Arnold has guided 17 Ph.D. scholars and has been on editorial boards of renowned journals. His research interests include estimation theory, probability, stochastic processes, mathematical learning models, biological models, characterizations, income distributions, order statistics, inequality measurement, record values, conditionally specified distributions, and Bayesian inference. He has been invited for scholarly lectures from across the world and is recognized by American Statistical Association, USA (Fellow); American Association for the Advancement of Science, USA (Fellow); Institute of Mathematical Statistics, USA (Fellow); Royal Statistical Society, UK (Fellow); International Statistical Institute, the Netherlands (Elected Member); and Forum for Interdisciplinary Mathematics, USA (Vice-President).

## Contributors

**Toshihiro Abe** Faculty of Economics, Hosei University, Tokyo, Japan

**Jose Ameijeiras-Alonso** Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, de Compostela Santiago, Spain

**Mohammad Arashi** Department of Statistics, University of Pretoria, Pretoria, South Africa;
Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

**Richard Arnold** School of Mathematics and Statistics, Victoria University of Wellington, Wellington, New Zealand

**Abeku A. Asare-Kumi** University of Louisiana at Lafayette, Lafayette, USA

**Stuart Barber** School of Mathematics, University of Leeds, Leeds, UK

**Andriette Bekker** Department of Statistics, University of Pretoria, Pretoria, South Africa

**Jan Beran** Department of Mathematics and Statistics, University of Konstanz, Konstanz, Germany

**Atanu Biswas** Applied Statistics Unit, Indian Statistical Institute, Kolkata, India

**Philippa M. Burdett** School of Mathematics, University of Leeds, Leeds, UK

**Yogendra P. Chaubey** Department of Mathematics and Statistics, Concordia University, Montréal, Canada

**Rosa M. Crujeiras** Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, de Compostela Santiago, Spain

**M. Di Marzio** DMQTE, University of Chieti-Pescara, Chieti-Pescara, Italy

**S. Fensore** DMQTE, University of Chieti-Pescara, Chieti-Pescara, Italy

**Riccardo Gatto** Department of Mathematics and Statistics, Institute of Mathematical Statistics and Actuarial Science, University of Bern, Bern, Switzerland

**Sucharita Ghosh** Statistics Lab, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

**Peter Guttorp** Norwegian Computing Center, Oslo, Norway

**Thomas Hamelryck** Department of Biology/Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

**Douglas M. Hawkins** Scottsdale Scientific LLC, Scottsdale, AZ, USA

**Shayan Hundrieser** Institute for Mathematical Stochastics, University of Göttingen, Göttingen, Germany

**Tomoaki Imoto** School of Management and Information, University of Shizuoka, Shizuoka, Japan

**S. Rao Jammalamadaka** Department of Statistics and Applied Probability, University of California, Santa Barbara, USA

**Jayant Jha** Institut de Neurosciences des Systèmes, Aix-Marseille University, Marseille, France

**Richard A. Johnson** Department of Statistics, University of Wisconsin, Madison, WI, USA

**Peter Jupp** School of Mathematics and Statistics, University of St Andrews, St Andrews, UK

**John T. Kent** School of Mathematics, University of Leeds, Leeds, UK

**Sungsu Kim** University of Louisiana at Lafayette, Lafayette, USA

**Marcel Klatt** Institute for Mathematical Stochastics, University of Göttingen, Göttingen, Germany

**Francesco Lagona** Department of Political Sciences, University of Roma Tre, Rome, Italy;
Department of Mathematics, University of Bergen, Bergen, Norway

**Arnab Kumar Laha** Indian Institute of Management Ahmedabad, Ahmedabad, India

**Christophe Ley** Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium;
Department of Mathematics, University of Luxembourg, Esch-sur-Alzette, Luxembourg

**Shuangzhe Liu** Faculty of Science and Technology, University of Canberra, Canberra, ACT, Australia

**Richard Lockhart** Simon Fraser University, Burnaby, BC, Canada

**F. Lombard** Department of Statistics, University of Johannesburg, Johannesburg, South Africa

**Tiefeng Ma** School of Statistics, Center of Statistical Research, Southwestern University of Finance and Economics, Chengdu, China

**Sourav Majumdar** Indian Institute of Management Ahmedabad, Ahmedabad, India

**Kanti V. Mardia** School of Mathematics, University of Leeds, Leeds, UK;
Department of Statistics, University of Oxford, Oxford, UK

**Yoichi Miyata** Faculty of Economics, Takasaki City University of Economics, Gunma, Japan

**Shyama Prasad Mukherjee** Department of Statistics, University of Calcutta, Kolkata, India;
Indian Association for Productivity, Quality and Reliability, Kolkata, India

**Axel Munk** Institute for Mathematical Stochastics, University of Göttingen, Göttingen, Germany;
Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

**Najmeh Nakhaei Rad** Department of Statistics, University of Pretoria, Pretoria, South Africa;
Department of Mathematics and Statistics, Mashhad Branch, Islamic Azad University, Mashhad, Iran;
DSI-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), Johannesburg, South Africa

**Seng-Huat Ong** Institute of Actuarial Science and Data Analytics, UCSI University, Kuala Lumpur, Malaysia

**A. Panzera** DISIA, University of Florence, Florence, Italy

**Cornelis J. Potgieter** Department of Mathematics, Texas Christian University, Texas, TX, USA

**Helmut Schaeben** Geophysics and Geoscience Informatics, TU Bergakademie Freiberg, Freiberg, Germany

**Ashis SenGupta** Augusta University, Augusta, GA, USA;
Indian Statistical Institute, Kolkata, India;
Middle East Technical University, Ankara, Turkey

**Kunio Shimizu** School of Statistical Thinking, The Institute of Statistical Mathematics, Tokyo, Japan

**Takayuki Shiohama** Department of Data Science, Nanzan University, Nagoya, Japan

**Britta Steffens** Department of Mathematics and Statistics, University of Konstanz, Konstanz, Germany

**C. C. Taylor** Department of Statistics, University of Leeds, Leeds, UK

**Gyorgy Terdik** Faculty of Informatics, University of Debrecen, Debrecen, Hungary

**Serdar Tugac** Koç Information and Defence Systems, Ankara, Turkey

**Anietie Edem Udokang** Department of Mathematics and Statistics, School of Applied Science and Technology, Federal Polytechnic, Offa, Nigeria

**Fidelis Ifeanyi Ugwuowo** Department of Statistics, University of Nigeria, Nsukka, Nigeria

**Thomas Verdebout** Department of Mathematics and ECARES, Université Libre de Bruxelles (ULB), Brussels, Belgium

**Kasirga Yildirak** Actuarial Sciences, Hacettepe University, Ankara, Turkey

**Xiaoping Zhan** Law School, Sichuan University, Chengdu, China

# Quality of Life

# Florence Nightingale, Quality of Life and Statistics

**Shyama Prasad Mukherjee**

**Abstract** Florence Nightingale, the 'Lady with the Lamp,' was a pioneer reformer who substantially raised the content, impact as also the status of nursing and public healthcare service. Nightingale, the 'Lady with the Data,' was a passionate statistician and a strong advocate of statistics for reducing the burden of disease and improving the quality of life of people at large. She wrote profusely and developed innovative diagrams to convey messages underlying assiduously collected data in a simple, clear and convincing manner to decision-makers and administrators.

## 1 Introduction

Florence Nightingale is respectfully remembered and even venerated for her exemplary, dedicated and ingenious services to improve quality of life of wounded soldiers in British Army hospitals. Avoiding anecdotal details, it can be easily said that it was her singular devotion to nursing of the sick and the wounded beyond these hospitals that could establish nursing as a noble and humanitarian profession to attract generations of young persons committed to take care of the sick, the wounded and the diseased. 'The lady with the lamp' will continue to inspire humanity in providing relief and succor to the ailing and, beyond that, to improve the quality of life of the general population.

Less known, however, are the significant contributions of Nightingale toward improvement in positive aspects of health and welfare including sanitation and proper hygienic conditions in institutions providing medical care. And to muster support for her initiatives in promoting positive aspects of health—both physical and mental health—she developed and deployed some simple but meaningful numerical presentations, aggregation of facts and figures she painstakingly documented during her devoted career. These 'statistical' tools used by her went a long way to convince the

S. P. Mukherjee (✉)
Department of Statistics, University of Calcutta, Kolkata, India
e-mail: prof.mukherjee@gmail.com

Indian Association for Productivity, Quality and Reliability, Kolkata, India

 
A. SenGupta and B. C. Arnold (eds.), *Directional Statistics for Innovative Applications*, Forum for Interdisciplinary Mathematics,
https://doi.org/10.1007/978-981-19-1044-9_1

concerned administration in moving beyond providing treatment to the diseased to promoting positive aspects of health care by paying due attention to sanitation and public hygiene.

Florence Nightingale did not mention ‘Quality of Life’ explicitly in her dossiers. It is clear, however, that what she emphasized upon had a strong bearing on enhancing what has been later on branded as ‘Quality of Life’ in international parlance on human development, not merely for people undergoing treatment in healthcare institutions, but for the entire population. It is worth mentioning that Nightingale took great interest in improving sanitation and related aspects of positive health in India.

Health as an important constituent of ‘quality of life’ has to focus on both positive and negative aspects of health. And, in our bid to enhance this quality, we should focus on identifying major causes of not only mortality but also morbidity among members of the population. Such an orientation in terms of a quantitative analysis of the burden of disease will help a nation or a community to plan for effective prevention and control of major diseases within available resources.

The present article is an attempt to briefly review the ‘statistical’ work done by Florence Nightingale to get an insightful glimpse into possible means and mechanisms for improving institutionalized health care as also her concern for and contribution to enhancement of positive aspects of health. The second part points to the role of health in quality of life, and the third part provides a concise treatment of the concept and use of burden of disease.

## 2 Contributions of Florence Nightingale

Florence Nightingale and her team of 38 nurses came to run hospitals for wounded soldiers evacuated from Crimea. Her pioneering work in the Scutari barracks has been recorded as a legend in the annals of modern medical care. On arrival at Crimea in November 1854, the team found soldiers suffering from frostbite, dysentery, cholera and typhus, living in utterly chaotic, unsanitary and inhumane living conditions. There were no blankets, beds, furniture, food or cooking utensils, but there were rats and fleas everywhere. On top of this dismal situation, the nurses found inadequate medical records. There, as no systematic record, hundreds of soldiers were buried with a record being made of their deaths and a bureaucratic inertia prevented nurses and administrators from spotting obvious flaws in the system.

Nightingale and her dedicated team adopted a two-pronged approach to improve appalling situation. They approached administration to provide some basic amenities for the patients with possibly a limited expectation that their demands would be met forthwith. For the second approach, they took up a strong awareness campaign among the soldiers to maintain personal hygiene, wash their clothing, clean their utensils and keep the hospital environment garbage-free to the extent possible. Her team members also helped the soldiers in this task. They realized that, until the patients are convinced that the efforts on their part would help improve their conditions of living in the hospital, the campaign would not yield the desired results.

While Florence Nightingale will always be remembered as the ‘lady with the lamp’ who appeared as an angel to the dying British soldiers in Army hospitals, she has been also recognized as the ‘Lady with the Data’ who passionately used statistics to comprehend the causes of death in those hospitals and to bring about a change in the outlook of health administration. In the words of [4], ‘no more forceful example of the value of using health statistics to understand and improve health condition exists than (that) displayed by Florence Nightingale.’

Born to progressive parents (her father William Nightingale was known for his strong emphasis on women’s education) Nightingale learnt several European languages at an early age and picked up a lot of interest in Mathematics and in data, collection, representation and interpretation. Some unconfirmed reports indicate that she received her mathematics education from the great mathematician Sylvester. In fact, she had recorded in details her experiences about prevailing social systems, legal administrations and peoples’ behavior in different countries in Europe through which the Nightingale family had moved when Florence was a child. This knack for careful and systematic data collection and display picked up early in life served her well throughout her life.

Nightingale set about collecting statistics in Crimea on the number of soldiers killed, wounded or diseased, just a biological scientist does collect data on specimens of butterflies and fossils to study biodiversity. She led a team of people for gathering such data. She could unravel the truth behind the unusually high mortality among the soldiers by painstakingly recording medical data and streamlining documentation of medical and health data and establishing the fact that the number of deaths among soldiers due to preventable unsanitary conditions and lack of basic amenities in the hospitals was seven times more than the fatalities due to battlefield wounds. All this she could do because of her early training in mathematics and her childhood skills to collect, organize and diagrammatize data.

Improvement in health care and its positive impact on quality of life were the key motivation behind Nightingale’s revolutionary efforts chronicled in history. Keeping proper records and paying due attention to the factors contributing to ill health or morbidity of a patient, she could strike at the root of the malady afflicting the patient and could arrange for the appropriate medical care. She also realized the need to keep up the morale of the patients and to involve them in exercises to improve their personal hygiene. In fact, she appeared to be a Beacon of Light or as an angel to the dying and the diseased because of her personal attention to each one of them and her empathy that would ameliorate the feeling of pain and depression among the patients.

## 3 Florence Nightingale—The ‘Passionate’ Statistician

When she returned from Crimea, she started on her work of publishing her statistical findings as also her proposed reforms in healthcare management. Her mission was to reach out to people who could put her ideas and concrete suggestions for reforms into practice. Her emphasis on data, data visualization and data interpretation was not to get a detailed account of deaths but to save lives through a data segregation approach to find out causes of death which could be prevented through reforms.

One of Nightingale's most significant innovations was the polar-area diagram, also known as the 'rose' diagram or the 'cox-comb', which showed the causes of death for every month over two successive years in Crimea (Fig. 1). The first year, viz. 1854-55, was following her arrival in the region and the second was the next year after she could implement a series of reforms in the hospital and nursing practices. Each wedge in the diagram represented a month, and the area of the wedge stood for the number of soldiers who died during that month. The blue area showed deaths from preventable diseases contracted in the dismal conditions prevailing in Crimea while the red sections showed deaths from battlefield wounds. Black areas represented deaths from other causes. A comparison of the diagrams for the two consecutive years could be easily convince one of the huge positive difference in mortality through the reforms implemented and the campaigned carried out by Nightingale. The second was the stark revelation that preventable diseases led to more deaths than battlefield wounds.

The diagram was drawn on a polar coordinate grid. Each category or interval in the data is divided into equal segments on this radial chart. How far each segment extends from the center of the polar axis depends on the magnitude it represents. So, each ring from the center of the polar grid can be used as a scale to plot the segment size and represent a higher value. Thus, it is important to note that the area, rather than the radius of a segment that represents the magnitude in this diagram invented by Nightingale. A deficiency that can be noted is that the outer segments are given more emphasis because of their larger areas. This represents increases in magnitude rather disproportionately.

Let us remind ourselves of the fact we are speaking about the effectiveness of a statistical diagram—that may appear to-day to be quite simplistic–thought of and constructed nearly 170 years back when there were countably few statisticians to offer some relatively simple statistical tools to draw inferences from data. Even systematic data collection had not been introduced those days. And, most statisticians used tables to represent the data they looked at. Further, the diagram suggested by Nightingale is not to be confused with the Pi-diagram. In fact, each wedge of the diagram she had used was like a Pi-diagram. Placing one diagram drawn by Nightingale over the other, a month-on-month comparison could be made along with the shifts in causes of death. The effectiveness of this diagram could be judged by the fact members of British Parliament and Army officials could understand the scale of the problem, the impact of reforms implemented by Nightingale and the truth about the real factors behind the high mortality among the wounded.

To carry her message for reforms in healthcare administration including standardization in medical records and logical analysis of those records to get an insight into diseases, specially those which are preventable, Nightingale began collaborating with radical sociologist and noted journalist Harriet Martineaux who wrote a popular book 'England and Her Soldiers' published in 1859, revealing Nightingale's findings to the public. No wonder, the book did not get official entry into Army Hos-

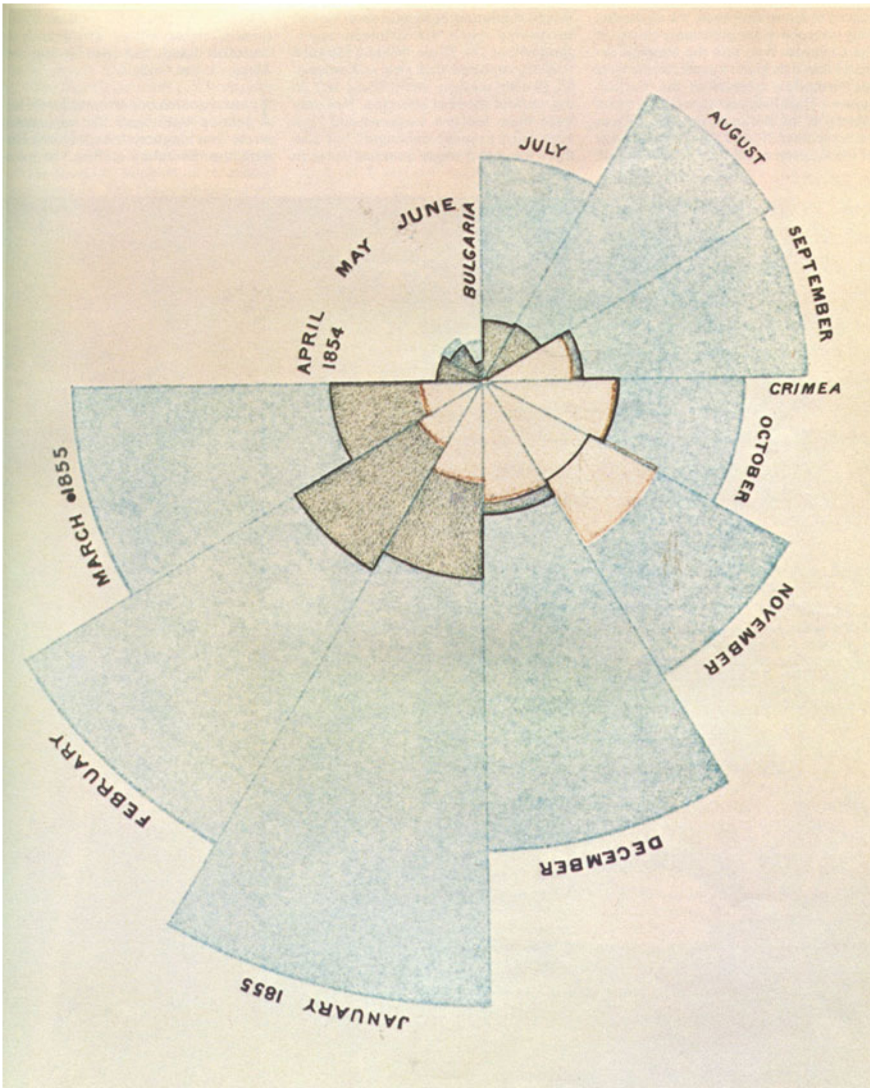


**Fig. 1** Rose diagram of the causes of mortality in the army in the East from April 1854 to March 1855, adopted and modified from the Source: Modern reproduction of the image from Nightingale (1858) http://hugh-small.co.uk/

pital Libraries. Her book on 'Measuring Hospital care Outcomes' bears testimony to her persistent use of data and data analysis for assessing healthcare outcomes and improvements therein.

Contributions of Florence Nightingale did remain confined to British soldiers or even to people in Britain only. Once the British Parliament passed the British Public

Health Act 1874-75 as a follow-up of her strong advocacy of data-based reforms in healthcare administration, Nightingale looked to the healthcare systems in some British colonies, notably in India. She studied the causes behind the famine in India, and she wrote on the principal factors responsible for the famine pointing out the need for drastic improvement in sanitation and other public health improvement measures. She also collected statistics of irrigation in India and the effect of irrigation on the life and health of the people, mainly confirm her concerns for quality of life.

In recognition of her contributions in the field of statistics, Nightingale was elected the first female member of the Statistical Society (now the Royal Statistical Society) in 1858 just two years after her return from the Crimea. She was elected to the Statistical Congress the same year. In 1874, she was made an honorary foreign member of the American Statistical Association. Her pioneering work in Modern Nursing and in Public Health Statistics brought her the first Royal Red Cross Award in 1883 and the Lady of Grace Award in 1904. The Order of Merit was conferred on her in 1907.

## 4 A Call on Quality of Life

Quality of life has been formally defined in several ways: Attempts have even made to indirectly quantify it to the extent required to compare quality of life over time and space. Some exponents emphasize on a holistic view that links quality of life to human development. Different aspects or determinants or quality of life have come to be widely recognized, and quality of health care has been duly accorded an important place. However, indirect reflections of quality of health care—which goes beyond quality of medical care or quality of health provided by care-givers distinct from quality of health care as is enjoyed by the people at large or, in particular, by the sick and the afflicted—have been usually considered in assessing quality of life. Somewhat marginalized or missed out are steps needed to improve this quality. In fact, the human development index computed and published by the United Nations Development Programme has a component index based only on one outcome variable, viz. expectation of life at birth to bear on 'health' with no determining or influencing variables included in the computation. And, this expectation derived from the life table takes account of mortality only, though morbidity has a significant effect on the quality of life lived by an individual and on the collective quality of life for a population.

To celebrate the bi-centenary of the birth of the 'lady with the lamp' who enlightened the contemporary society to recognize nursing as a noble profession that can contribute effectively to improve quality of life of patients in hospitals and similar institutions, we should remember that this Noble lady focused on improving basic amenities enjoyed by patients even making their own efforts through an awareness campaign and an enabling environment coupled with a competent guidance.

Quality of health care to the people at large by the health administrators includes consideration of adequacy and efficacy of public health measures like sanitation as well as quality of environment—quality of air we breathe or the quality of water we drink and use for cleaning. We have to remember that a hospital or a hospitex or a nursing home is not just an attractive and well-maintained building, sometimes with an impressive array of security men or administrative staff members. However, it may even maintain some standard of cleanliness. But it may lack adequate manpower in terms of doctors, nurses, paramedical staff, technicians and other health workers including attendants and cleaning staff. It may not have adequate equipments and accessories which are functional. It may not have an adequate stock of medicines, testing kits, medical appliances, measuring devices, control equipments and commodities to meet dietary requirements of all concerned. Above all, there could be a lack of patient-friendly and humane environment prevailing within the institution.

Nightingale's first paper espousing the cause of hospital reforms in 1858 specified a mer four common defects: agglomeration of a large number of sick under the same roof, deficiency of space, of ventilation and of illumination. Later, she expanded the list to include 16 defects in hospital design and administration. She followed up her work, and two years later, she presented her paper 'Health Statistics' at the International Statistical Congress held in London chaired by von Quetlet where a follow-up on uniform hospital statistics was approved unanimously.

The pavilion method of hospital construction was an important recommendation of Nightingale to prevent cross-infection and thus to bring down fatalities. In her Notes on Hospitals, she asserted that non-infectious patients could be accommodated safely in wards with infectious patients, provided there was adequate ventilation. And, her keenness to learn and to change made her an advocate for isolation rooms in hospitals for infectious patients on the advice on anonymous reviewer in Medical Times and Gazette.

Much of Nightingale's data collection was meant to bring out comparisons between different sizes and types of institutions (hospitals, infirmaries, home deliveries, etc.). She could point out in terms of her assiduously collected data clearly displayed that death rates were influenced by practices for hand-washing, clothing worn by doctors and nurses, furnishings, wall and floor coverings and their upkeep.

## 5 Burden of Disease

Nightingale was convinced that measures to improve public health conditions by creating greater awareness among the people as also among the healthcare providers to address positive aspects of health by mitigating the incidence and impact of preventable diseases is much more important than focusing attention on institutional health care only. Thus, her emphasis on morbidity as a priority over mortality has been a pointer to the development of concepts and measures of 'burden of disease' much later, though this phrase was not coined or articulated distinctly by her. In due remembrance of her contribution to improvement of health of the population, it will not be out of place to enter into a discussion on burden of disease.

As a concept, burden of disease on a population group during a certain time period is a summary measure that takes care of both fatal and non-fatal outcomes of diseases, causing (premature) deaths and disabilities, respectively. The impact of morbidity itself is understood in terms of years of ‘healthy’ life lost by virtue of being in states of poor health or disability (YLD). To estimate YLD for a particular cause in a particular time period, the number of incident cases in that period is multiplied by the average duration of the disease and a weight factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (dead). The basic formula for YLD, without applying social preferences, from a given disease or group of diseases is given by

$$\text{YLD} = I \times \text{DW} \times L$$

where $I$ = number of incident cases, DW = disability weight and $L$ = average duration of the case until remission or death (years). Disability weights are taken as mid-points of seven different ranges for disability scores in seven categories or states used in the GBD 2000 study. The Global Burden of Disease Study used the same disability weights for all regions of the world. It used the same life expectancy ‘ideal’ standard for all population subgroups and it excluded all non-health characteristics (such as race, socio-economic status or occupation) apart from age and sex from consideration in calculating lost years of healthy life. Most importantly, it used the same ‘disability weight’ for everyone living a year in a specified health state. Additionally, 3% time discounting to account for the fact time lived now is preferred to time in future. Age weights giving less weight to years lived at very young and older ages compared to ages of young adults were used in calculating DALYs for the original Global Burden of Disease study.

Important methodological issues in estimating the burden of disease, sometimes attempted for a particular disease, arise in the following contexts: Segregation of deaths which are caused by diseases (as distinct from homicides, suicides, accidental deaths and natural deaths at mature ages) and, similarly, segregation of disabilities associated with diseases (sickness) from those one to accidents. Diagnosis of disabilities caused by diseases and putting them in well-defined categories. Assigning weights to different disability categories to calculate YLD averaging over duration of sickness and over diseases. Choice of time-discounting rates and age weights. Within each of these issues, there are finer points which tend to be glossed over, like the disability suffered by an individual at the end of sickness and the disability during its existence may not be the same, making it not so logical to multiply the duration by the disability scale value, as is currently done.

To estimate the burden of a particular disease on a large population group, we have to take a suitable sample of new cases of incidences. The (sub-) population of new incidences is not easy to identify, and in the absence of a sampling frame, we may be compelled to use some non-random sampling procedure like snowball or chain referral sampling to arrive at a sample of some size, not pre-determined. We may use either linear or exponential referrals and have to work out the estimate of YLD. The estimate of YLL can be conveniently obtained from a household survey using random sampling with appropriate stratification, if needed. Properties of the estimated YLD are yet to be established. The two components of this burden, viz. Years of Life Lost due to deaths (YLL) and Years of Life Lost due to disability (YLD), are combined to yield Disability Adjusted Life Years lost indicating the gap between the ideal state of (public) health and the existing state of health. Data requirements obviously include records of deaths, classified according to (primary) cause of death, age, sex and such other classificatory variables which may be pertinent in the context as well records of incidences of different diseases as also of their health outcomes in terms of disabilities. Of course, we need the life table for the population under consideration.

Being a summary measure, burden of disease, as is or even can be computed ignores a lot of variations among individuals as well as among diseases and disabilities. Several assumptions—not all tenable in all situations— are made about discounting for time and age weighting to arrive at measures which are advocated as being more informative.

## 6 Concluding Remarks

Statistics should have a purpose and preferably a noble one that serves humanity well. Improving healthcare delivery to the sick and the dying is, beyond any shade of doubt, a very noble cause and using statistics to this effect—howsoever simple the methods and tools may appear—and that too during the nascent phase of development of statistics should be hailed as one of the most remarkable achievements in applications of statistics. What counts more in this context is the enrichment provided by statistics to the substantive content of a study meant to benefit the population at large rather than adding to the technical content and its sophistication that appeals to a small group of specialists. Also, to be appreciated is the role of advocacy of statistics for a purpose beyond contributing to the development of the discipline so that benefits from such developments can be simply, clearly and persuasively conveyed to the common people.

For further details, the interested reader is referred to [1–3, 5].

These lessons from the life and works of Florence Nightingale will continue to inspire many knowledge-seekers and activists striving to raise the quality of life for all.

**Acknowledgements** The author gratefully acknowledges the suggestions made by one of the editors for making the content more appealing and the help rendered by Dr. Radhakanta Das of the Presidency University, Kolkata, in preparing the final version of the paper.

## References

1. Jose, K.K.: Biostatistician Florence Nightingale. Jour. Ind. Stat. Assoc. **58**(1), 105–112 (2020)
2. McDonald, L.: Florence Nightingale: The making of a hospital reformer. Health Environ. Res. Des. J. **13**(2), 25–31 (2020)
3. Nightingale, F.: Introductory Notes on Lying-in Institutions. Longmans, Green (1871)

4. Schloman, B.F.: Using health statistics: a Nightingale legacy. Publ. Med. **6**(3), 7 (2001)
5. Science Museum London: Florence Nightingale: The Pioneer Statistician (2018). https://www.sciencemuseum.org.uk/objects-and-stories/florence-nightingale-pioneer-statistician

# Innovative Applications

# Mixture Models for Spherical Data with Applications to Protein Bioinformatics

**Kanti V. Mardia, Stuart Barber, Philippa M. Burdett, John T. Kent, and Thomas Hamelryck**

**Abstract** Finite mixture models are fitted to spherical data. Kent distributions are used for the components of the mixture because they allow considerable flexibility. Previous work on such mixtures has used an approximate maximum likelihood estimator for the parameters of a single component. However, the approximation causes problems when using the EM algorithm to estimate the parameters in a mixture model. Hence, the exact maximum likelihood estimator is used here for the individual components. This paper is motivated by a challenging prize problem in structural bioinformatics of how proteins fold. It is known that hydrogen bonds play a key role in the folding of a protein. We explore this hydrogen bond geometry using a data set describing bonds between two amino acids in proteins. An appropriate coordinate system to represent the hydrogen bond geometry is proposed, with each bond represented as a point on a sphere. We fit mixtures of Kent distributions to different subsets of the hydrogen bond data to gain insight into how the secondary structure elements bond together, since the distribution of hydrogen bonds depends on which secondary structure elements are involved.

K. V. Mardia (✉) · S. Barber · Philippa M. Burdett · John T. Kent
School of Mathematics, University of Leeds, Leeds LS2 9JT, UK
e-mail: mardia@stats.ox.ac.uk; K.V.Mardia@leeds.ac.uk

S. Barber
e-mail: S.Barber@leeds.ac.uk

John T. Kent
e-mail: J.T.Kent@leeds.ac.uk

K. V. Mardia
Department of Statistics, University of Oxford, Oxford OX1 3LB, UK

T. Hamelryck
Department of Biology/Department of Computer Science, University of Copenhagen, Copenhagen, Denmark
e-mail: thamelry@binf.ku.dk

 
A. SenGupta and B. C. Arnold (eds.), *Directional Statistics for Innovative Applications*, Forum for Interdisciplinary Mathematics,
https://doi.org/10.1007/978-981-19-1044-9_2