Some Inference Problems Related to Geometric Distribution

Thesis
Submitted to the
Mahatma Gandhi University, Kottayam
for the Degree of
Doctor of Philosophy in Statistics
Under the Faculty of Science

By

MATHACHAN PATHIYIL

Department of Mathematics
Union Christian College, Aluva
Ernakulam, Kerala, India

July 2007

Dr. E. S. Jeevanand
Reader in Statistics
Union Christian College
Aluva 683 102, India
email: radhajeevanand@yahoo.co.in

Edavazhikkal House
Chiriakkakam
Varapuzha 683 517, India
Phone: 91-0484-2515101

July 10, 2007

Certificate

Certified that the thesis entitled **"Some Inference Problems Related to Geometric Distribution"** is a bonafide record of work done by Mr. Mathachan Pathiyil, under my guidance, in the Department of Mathematics, Union Christian College, Aluva and that no part of it has been included any where previously for the award of any degree or title.

Jeevanand

Dr. E. S. Jeevanand

(Supervising Teacher)

Acknowledgements

I would like to place on record my sincere gratitude to Dr. E. S. Jeevanand, Reader in Statistics, Department of Mathematics, Union Christian College, Aluva for his patience, encouraging comments and valuable guidance throughout my research work.

I am grateful to Dr. N. Unnikrishnan Nair, former Vice-Chancellor and Head of the Department of Statistics, Cochin University of Science and Technology, for his advice and constructive comments.

I thank Rev. Dr. George Thanathuparambil, Principal, Nirmala College, Muvattupuzha and my colleagues in the Department of Statistics for their co-operation and help. I recall with gratitude the help and suggestions given by Dr. Johny Scaria and Dr. T. M. Jacob.

The co-operation rendered by Dr. Varghese John, Principal and Prof. Anu Ommen, Head of the Department of Mathematics, Union Christian College, Aluva is gratefully acknowledged.

I cherish the help rendered by Dr. Vincent Areathu, Newman College, Thodupuzha in the preparation of the thesis and for his linguistic suggestions.

I appreciate the help given by Dr. K. K. Jose, Head of the Department of Statistics, St. Thomas College, Pala and the other faculty members there. A word of thanks to Ms. S. Ajitha, for the lay-out of the thesis in LaTeX.

The blessings and encouragements from my parents, wife and children have been with me throughout the past years. I am indeed appreciative of their patience and whole hearted support.

I thank the University Grants Commission for awarding me leave for two years under the Teacher Fellowship Scheme of the X th plan.

There are many more, too many to mention in this limited space, to be thanked.

Above all I thank God Almighty who has always been an unfailing source of strength and inspiration to me throughout my research work.

Mathachan Pathiyil

Abstract

Reliability analysis is the branch of statistics that deals with collection of data, modeling and analysis of data on lifetimes of units or equipments. The area in which statistics has had its greatest impact in reliability is in the analysis of laboratory and field data on lifetimes or failure times. Statisticians have perhaps concentrated too much in literature on statistical niceties for certain distributions, and too little on innovative methods of life data analysis. Also a large amount of additional research has concerned continuous as opposed to discrete lifetimes. However, discrete lifetimes have important applications. Actuaries and bio-statisticians are interested in the lifetimes of persons or organisms, measured in months, weeks or days. For reliability engineers, 'time' can also be the number of times that a piece of equipment is operated, or the number of miles that a tyre is used. There is a strong case for looking at reliability aspects in the discrete domain.

The geometric distribution, owing to its lack of memory property and constant failure rate, is widely used to model discrete lifetimes. Motivated by the relevance and usefulness of the geometric model, this research aims to obtain some results that have applications in the modeling and analysis of data in the discrete time domain. This thesis is divided into 6 chapters. Chapter 1 serves as an introduction, which proposes a survey of literature relating to the subject matter of our present study, the basic definitions and notations used in the thesis and finally an outline of the work planned. As the geometric model belongs to the class of long tailed distributions, the occurrence of extreme observations is quite common and their identification as outliers or not becomes important. In Chapter 2, procedures for the determination of the number of outliers present in the sample taken from geometric distribution are discussed.

The problem of inferring the parameters of the geometric distribution in the presence of outliers is considered in Chapter 3. Bayes point estimators of the reliability measures of the geometric distribution under the identified outlier model and the ex-

changeable models are derived. Estimators of the reliability measures under the right-censored model are also obtained. Chapter 4 contains the derivation of the Bayes estimators of the reliability measures of geometric distribution using record values. Estimators are proposed under various loss functions. The usual Bayes estimators of the reliability measures using the original data, over which the record values are defined, are also obtained.

The estimation of the residual entropy function of the geometric distribution is considered in Chapter 5. The problem of estimating the function is carried out under three outlier models, namely, the identified model, the exchangeable model and the censored model. This chapter also reports the estimation of the residual entropy function using record values from the geometric model. In Chapter 6, we obtain semiparametric estimators of the survival function, hazard function, mean residual life function and the residual entropy function of the geometric distribution using uncensored and Type I censored samples.

Illustrative examples and simulation study are included in all chapters. The results of the same should give the users some guidance in selecting the estimator or procedure that is more appropriate to their particular application. The thesis concludes with suggestions for future work. Most of the works in the thesis are already communicated and some of them are published after reviewing.

Key words: *Bayesian estimation, Censoring, Credible interval, Geometric distribution, Hazard function, Kaplan-Meier estimator, Least square method of estimation, Loss function, Maximum likelihood estimation, Mean residual life function, Outliers, Prediction, Record values, Residual entropy function, Robustness, Survival function.*

AMS Subject Classification: 62 C20, 62 E25, 62 F, 62 F25, 62 N

Contents