SOME GENERALIZATIONS OF CAUCHY DISTRIBUTION

Thesis submitted to the

UNIVERSITY OF CALICUT

for the award of the degree of

DOCTOR OF PHILOSOPHY

in

STATISTICS

under the Faculty of Science

by

Fasna, K.

Under the Guidance of

Prof. (Dr.) K. Jayakumar

Department of Statistics

University of Calicut

Kerala-673 635

India

April 2024

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT

Prof. (Dr.) K. Jayakumar
Senior Professor

Calicut University (P.O.)
Kerala, India 673 635.
Phone : 0494 2407340
Email : jkumar19@rediffmail.com

Date : 09.01.2025

CERTIFICATE

I hereby certify that the work presented in this thesis entitled 'SOME GENERALIZATIONS OF CAUCHY DISTRIBUTION' is a bonafide work done by Mrs. Fasna K under my guidance for the award of the degree of Doctor of Philosophy in the Department of Statistics, University of Calicut and that this work has not been included in any other thesis submitted previously for the award of any degree. Also certify that the contents of the thesis have been checked using anti-plagiarism data base and no unacceptable similarity was found through the software check.

09.01.2025

DEPARTMENT OF STATISTICS UNIVERSITY OF CALICUT KERALA - 673635

Prof.(Dr.) K. Jayakumar

(Supervising Teacher)

DR. K. JAYAKUMAR
Senior Professor
Department of Statistics
University of Calicut
Kerala - 673635

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT

Prof. (Dr.) K. Jayakumar
Senior Professor

Calicut University (P.O.)
Kerala, India 673 635.
Phone : 0494 2407340
Email : jkumar19@rediffmail.com

Date : 09.01.2025

CERTIFICATE

Certified that the corrections/suggestions from the adjudicators of the thesis have been incorporated in this copy of the thesis.

09.01.2025

DR. K. JAYAKUMAR
Senior Professor
Department of Statistics
University of Calicut
Kerala - 673635

Prof. (Dr.) K. Jayakumar
(Supervising Teacher)

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT
KERALA - 673635

DECLARATION

I hereby declare that the work presented in this thesis entitled 'SOME GENERALIZATIONS OF CAUCHY DISTRIBUTION' is an original work done by me under the guidance of Dr. K. Jayakumar, Department of Statistics, University of Calicut, and has not been included in any other thesis submitted previously for the award of any degree.

University of Calicut

Date: 09.01.2025

Fasna K

Acknowledgements

Throughout the work of the thesis, I have received a great deal of support and assistance.

First, I would like to thank my supervisor, Dr. K. Jayakumar, Senior Professor, Department of Statistics, University of Calicut for his guidance and the freedom to choose and pursue the problem. He has been accommodating, patient, and encouraging.

I am extremely thankful to my teachers Dr. Krishnarani S. D., Head of the Department, Dr. M. Manoharan, Dr. C. Chandran, and Dr. Dileep Kumar, Department of Statistics, University of Calicut for the constant encouragement and valuable support.

I am very much thankful to Dr. Z.A. Ahamad Asharaf., Associate Professor, Head, Department of Statistics, Govt. College, Mokeri., Dr. Sankaran K.K., Head, Department of Statistics, Sree Narayana College, Nattika., Dr. Girish Babu M., Head, Department of Statistics, Govt. Arts and Science College, Kozhikode for their moral support and encouragement during the period of this research work.

My sincere thanks are also due to Librarian and the non-teaching staff of the Department of Statistics, University of Calicut for their help and co-operation. Also, I express my sincere thanks to all the research scholars, M.Phil. and M.Sc. students of the Department of Statistics, University of Calicut for their help and support.

I am grateful to my parents Asharaf and Sahira. They have supported me throughout my academic life. A huge thanks extended towards my brothers. I am thankful to my in-laws and brother-in-law, they made my research life smooth. I am very much grateful to my sweet little children for their understanding.

Finally my love and gratitude to my husband Junaid Muringattu, without his motivation, encouragement, and continous support, I definitely not able to complete this research work.

Lastly, but not least, I remember with sincere thanks to my extended family members and friends for bearing with me throughout this research work with their cooperation. This thesis would not have been possible without the support and the help of several individuals in the preparation and completion of the study.

Abstract

Some Generalizations of Cauchy Distribution

This thesis is mainly concerned with study of some generalizations of Cauchy distribution. The distributions commonly used for modelling of insurance losses, financial returns, file sizes on the network servers, etc. are subject to some sort of deficiencies. Also, there are only few probability distributions capable of modelling heavy tailed data sets and none of them are flexible enough to provide greater accuracy in fitting complex forms of data. Furthermore, in financial and actuarial risk management problems, the data sets are usually unimodal, skewed to the right, and possess thick right tail. The distributions that exhibit such characteristics can be used quite effectively to model insurance loss data to estimate the business risk level.

To address the problems stated above, we have an interest in defining new families of distributions through different approaches such as introducing additional, location, scale, shape, and transmuted parameters, to generalize the existing distributions.

We adopted six estimation approaches for estimating parameters of our models, and assess the performance of these estimators. Therefore, this study is addressed scientific computation challenge by performing numerical comparison between several estimators for the model parameters and identified which of them perform better in terms of estimation efficiency. The comparison is based on Monte Carlo simulations and the outcomes of a real data analysis. The simplicity of the proposed distributions and the great flexibility in modelling real life data will attract researchers to use these distributions as an alternative of the Cauchy distribution in modelling different scenarios. Our proposed families are useful for modelling insurance claim data sets, better models for financial returns because the normal model does not capture the large fluctuations seen in real assets. Also, our families of distributions have received considerable attention due to the heavy tail property.

സംഗ്രഹം

കോഷി വിതരണത്തിന്റെ ചില പൊതുവൽക്കരണങ്ങൾ

ഈ പ്രബന്ധം പ്രധാനമായും കോഷി വിതരണത്തിന്റെ ചില പൊതുവൽക്കരണങ്ങളെക്കുറിച്ചുള്ള പഠനമാണ്. ഇൻഷുറൻസ് നഷ്ടം, സാമ്പത്തിക വരുമാനം, നെറ്റ്‌വർക്ക് സെർവറുകളിലെ രേഖാസമാഹാരങ്ങളുടെ വലുപ്പം മുതലായവയുടെ മാതൃകയാക്കാൻ സാധാരണയായി ഉപയോഗിക്കുന്ന വിതരണങ്ങൾ ഒരുതരം പോരായ്മകൾക്ക് വിധേയമാണ്. കൂടാതെ, ദീർഘവാലുള്ള ദത്ത ഗണങ്ങളെ മാതൃകയാക്കാൻ കഴിവുള്ള ചുരുക്കം ചില സാധ്യത വിതരണങ്ങൾ മാത്രമേ ഉള്ളൂ, അവയൊന്നും ദത്തങ്ങളുടെ സങ്കീർണ്ണമായ രൂപങ്ങൾ ഘടിപ്പിക്കുന്നതിൽ കൂടുതൽ കൃത്യത നൽകാൻ പര്യാപ്തമല്ല. കൂടാതെ, സാമ്പത്തിക മാതൃക ഉണ്ടാക്കൽ പ്രശ്നങ്ങളിൽ ദത്ത ഗണങ്ങൾ സാധാരണയായി ഏകീകൃതവും, വലതുനിന്നും വലിച്ചുള്ളതും, കട്ടിയുള്ള വലത് വാലുള്ളതുമാണ്. അത്തരം സ്വഭാവസവിശേഷതകൾ പ്രകടിപ്പിക്കുന്ന വിതരണങ്ങൾ ഇൻഷുറൻസ് നഷ്ട ദത്തങ്ങൾ മാതൃകയാക്കാൻ വളരെ ഫലപ്രദമായി ഉപയോഗിക്കാം.

മുകളിൽ പറഞ്ഞിരിക്കുന്ന പ്രശ്നങ്ങൾ പരിഹരിക്കുന്നതിന്, നിലവിലുള്ള വിതരണങ്ങളെ പൊതുവൽക്കരിക്കാൻ വ്യത്യസ്ത സമീപനങ്ങളിലൂടെ വിതരണത്തിന്റെ പുതിയ കുടുംബങ്ങളെ നിർവചിക്കുന്നതിൽ ഞങ്ങൾക്ക് താൽപ്പര്യമുണ്ട്. അതിനാൽ, കോഷി വിതരണങ്ങളുടെ ചില പൊതുവൽക്കരണങ്ങൾ നിർദ്ദേശിക്കുകയും അതിന്റെ ഗുണങ്ങൾ ചർച്ച ചെയ്യുകയും ചെയ്യുന്നു.

കൂടാതെ, ഈ വിതരണങ്ങളുടെ സ്ഥിരരാശികൾ കണക്കാക്കുന്നതിനും ഈ മതിപ്പ്കളുടെ പ്രകടനം വിലയിരുത്തുന്നതിനും ഞങ്ങൾ ആറ് എസ്റ്റിമേഷൻ സമീപനങ്ങൾ സ്വീകരിച്ചു. അതിനാൽ, മാതൃക സ്ഥിരരാശികൾക്കായി നിരവധി മതിപ്പുകൾ തമ്മിലുള്ള സംഖ്യാപരമായ താരതമ്യം നടത്തി, അവയിൽ ഏതാണ് എസ്റ്റിമേഷൻ കാര്യക്ഷമതയുടെ കാര്യത്തിൽ മികച്ച പ്രകടനം കാഴ്ചവയ്ക്കുന്നതെന്ന് കണ്ടെത്തുന്നതിലൂടെ ശാസ്ത്രീയമായ കണക്കുകൂട്ടലെന്ന വെല്ലുവിളിയെ അഭിമുഖീകരിക്കുന്നു. മോണ്ടെ കാർലോ സിമുലേഷനുകളും, യഥാർത്ഥ ദത്ത വിശകലനത്തിന്റെ ഫലങ്ങളും അടിസ്ഥാനമാക്കിയുള്ളതാണ് താരതമ്യം. നിർദിഷ്ട വിതരണങ്ങളുടെ ലാളിത്യവും യഥാർത്ഥ ജീവിത ദത്ത മാതൃക ചെയ്യുന്നതിലെ മികച്ച വഴക്കവും വ്യത്യസ്ത സാഹചര്യങ്ങളെ മാതൃകയാക്കുന്നതിൽ കോഷി വിതരണത്തിന് പകരമായി ഈ വിതരണങ്ങൾ ഉപയോഗിക്കാൻ ഗവേഷകരെ ആകർഷിക്കും.

അതിനാൽ ഞങ്ങളുടെ നിർദ്ദിഷ്ട കുടുംബങ്ങൾ ഇൻഷുറൻസ് നഷ്ട ദത്ത ഗണങ്ങളുടെ മാതൃക ചെയ്യുന്നതിനും, യഥാർത്ഥ ആസ്തികളിൽ കാണുന്ന വലിയ ഏറ്റക്കുറച്ചിലുകൾ സാധാരണ മാതൃക പിടിച്ചെടുക്കുന്നില്ലാത്ത കാരണം സാമ്പത്തിക വരുമാനത്തിനായുള്ള മികച്ച മാതൃകകളായും ഉപയോഗിക്കുന്നു. ദീർഘവാലുള്ള സവിശേഷത കാരണം ഞങ്ങളുടെ വിതരണ കുടുംബങ്ങൾക്ക് ഗണ്യമായ ശ്രദ്ധ ലഭിച്ചു.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike Information Criterion
AD	Anderson Darling
AR	Autoregressive
BIC	Bayesian Information Criterion
BTC	Beta Transformed Cauchy
CVM	Cramer-Von Mises
AICC	Corrected Akaike Information Criterion
cdf	cumulative distribution function
DIC	Deviance Information Criterion
DML	Discrete Mittag-Leffler
DMLC	Discrete Mittag-Leffler Cauchy
GCD	Generalized Cauchy Distribution
hrf	hazard rate function
HQIC	Hannan-Quinn Information Criterion
i.i.d.	independent and identically distributed

KS	Kolmogorov Smirnov
LS	Least Square
LCL	Lower Credible Limit
MCMC	Markov Chain Monte Carlo
ML	Maximum Likelihood
MPS	Maximum Product Spacings
MSE	Mean Square Error
NGP	New Generalized Pareto
pdf	probability density function
pgf	probability generating function
pmf	probability mass function
r.v.	random variable
RHS	Right Hand Side
RTAD	Right Tailed Anderson Darling
SC	Skew Cauchy
SE	Standard Error
sf	survival function
TC	Transmuted Cauchy
TNB	Truncated Negative Binomial
QRTM	Quadratic Rank Transmutation Mapping

UCL	Upper Credible Limit

Chapter 1

INTRODUCTION

1.1 Introduction

A statistical distribution is a parameterized mathematical formula that provides the probabilities of various outcomes for a random variable (r.v.). There are discrete and continuous distributions based on the random value it models. Sampling distributions are crucial in statistics as they enable us to gather samples and estimate the parameters of the population distribution. Therefore, distribution is essential for drawing inferences about the entire population.

In statistical theory, data modeling is a captivating research subject in engineering, medical, and financial sciences. Thus, probability distributions are essential for modeling such datasets. The most commonly used probability distributions include exponential, Weibull, Rayleigh, gamma, beta, Pareto, log-normal, Lomax, and Burr, among others. However, these distributions may not adequate for complex datasets. For instance, in reliability engineering and biomedical sciences, datasets often exhibit a unimodal distribution skewed to the right; see Demicheli et al. (2004), Lai and Xie (2006) and Almalki and Yuan (2013). Therefore, in such scenarios, commonly used statistical models may not be an appropriate option. Conversely, log-normal, gamma and beta distributions do not have closed forms for the cumulative distribution function (cdf), which posses challenges in estimation of parameters. Furthermore, the data sets are usually uni-modal, skewed to the right and possess thick right tail in financial and actuarial risk management problems; for details see, Cooray and Ananda (2005) and Eling (2012), among others. Distributions displaying such characteristics can be effectively utilized to model insurance loss data for estimating the level of business risk. The distributions commonly used in the literature include Pareto by Cooray and Ananda (2005), Lomax by Scollnik (2007), Burr by Nadarajah and Bakar (2014) and Weibull by Bakar et al. (2015), which are particularly suitable for modeling insurance losses, financial returns, file sizes on network servers, and similar phenomena. Regrettably, these distributions are susceptible to certain deficiencies. For instance, the Pareto distribution, owing to its monotonically decreasing density shape, may not offer the optimal fit in numerous applications. Conversely, the Weibull distribution can effectively capture the behavior of small losses but may fall short in representing the behavior of large losses. As a result, only a limited number of probability distributions can effectively model heavy-tailed datasets, and none provide enough flexibility to achieve high accuracy when fitting complex forms of data.

To address the problems identified above, researchers are keen on defining new families of distributions by augmenting one or more parameters to the well known distributions. The new families have been formulated using various methods including the incorporation of additional parameters such as location, scale, shape, and transmuted parameters, to generalize the existing distributions. These generalizations are mainly based on the transformation of the variable and compounding of two or more models.

Distributions obtained by compounding a parent distribution with a discrete distribution is a common technique in statistics. The fundamental concept behind introducing compound families is that in systems with discrete r.v. represented by N components and a positive continuous r.v., denoted as X_i, the lifetime of the i^{th} component can be expressed as the non-negative r.v. $Y = \min(X_1, X_2, \ldots, X_N)$ or $Z = \max(X_1, X_2, \ldots, X_N)$ depending on whether the components are structured in series or parallel, respectively.

Many compound classes can be constructed by choosing discrete models namely the geometric, Poisson, logarithmic, binomial, negative-binomial and power series with probability mass function (pmf) $P(N = n)$. The r.v.s $Y = \min(X_1, X_2, \ldots, X_N)$ and $Z = \max(X_1, X_2, \ldots, X_N)$ give rise to numerous models applicable to series and parallel systems with identical components, and they have wide applications in both industrial and biological contexts.

Heavy-tailed phenomena represent a distinct qualitative behavior of the underlying models, characterized by deviations from normal behavior, often influenced by extreme values of the sample. In Probability Theory, heavy-tailed distributions are those whose tails decay to zero at a slower rate than exponential decay. The exponential distribution is usually considered as the borderline between heavy and light tails. Among heavy-tailed distributions, two classes have been particularly successful: distributions with regularly varying tails and subexponential distributions.

All commonly used heavy-tailed distributions are belong to the subexponential class. One-tailed distributions include the Weibull distribution with a shape parameter greater than 0 but less than 1, Pareto distribution, Burr distribution, Log-normal distribution, Levy distribution, among others. Additionally, two-tailed distributions include Cauchy distribution (itself a special case of both the stable distribution and t-distribution), the family of stable distributions, financial models with long-tailed distributions and volatility clustering, the t-distribution, the skew log-normal cascade distribution, and more.

Statistical modeling plays an important role in engineering problem solving, providing tools to characterize stochastic events and derive theories and methods for sample processing. Cauchy distribution is particularly advantageous due to its closed form pdf expression across the whole family, which makes it suitable for modeling various real-life processes in engineering applications.

Cauchy distribution's history was first observed and studied in works of Pierre de Fermat and then studied by prominent researchers such as Isaac Newton, Gottfried Leibniz and others (see, Johnson et al. (1994)). It was formally introduced and published in 1824 by the French mathematician Poisson (see, Poisson (1827)). According to Johnson et al. (1994), this distribution is attributed to Cauchy (1853) when Cauchy responded to an article by Bienayme (1853) criticizing a method of interpolation proposed by Cauchy. Physicists often refer to Cauchy distribution as Lorentz distribution due to Hendrink Lorentz. One notable characteristic of Cauchy distribution is its lack of a defined mean, yet it possesses well defined mode and median values.

The probability density function (pdf) of a r.v. X following a Cauchy distribution, denoted by $C(\mu, \theta)$ with parameters $\mu \in \mathbb{R}$ and $\theta > 0$, is given by

$$g(x) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}; \quad -\infty \leq x \leq \infty, -\infty \leq \mu \leq \infty, \theta > 0 \tag{1.1}$$

and the cdf of X is

$$G(x) = \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right) + 0.5, \tag{1.2}$$

where μ and θ are the location and scale parameters respectively.

Early interest in the distribution focused on its value as a counterexample highlighting the necessity of regularity conditions for proving important limit theorems (see Stigler (1974)). While sometimes considered a pathological case, Cauchy distribution finds application in describing a wide range of phenomena. It represents a continuous probability distribution with heavier tails compared to normal distributions, making it suitable for modeling scenarios such as the ratio of two normal r.v.s. One of the notable characteristics of Cauchy distribution is its long-tailed nature, which makes it relevant in various statistical contexts. Despite its usefulness, Cauchy distribution poses challenges for estimation methods such as method of moments estimation and minimum variance unbiased estimation due to its long tails and lack of finite moments.

Although a graph of Cauchy density looks like the normal density, the tails of Cauchy distribution approaches to the x-axis much more slowly than the normal indicating higher probabilities for extremely large or small values and hence the mean does not exist.

This distribution is a special case of the Pearson type VIII distribution (Johnson et al. (1994)). Cauchy distribution is a limiting case of a Pearson distribution of type IV. It is also a special case of a Pearson distribution of type VII. If X and Y are independent and identically distributed (i.i.d.) standard Cauchy r.v.s $C(0,1)$, then $X+Y$ is a $C(0,2)$ r.v.. Cauchy r.v.s are such that the characteristic function of their sum is the product of their characteristic functions and yet they are not independent. In the case of a sequence of i.i.d. Cauchy r.v.s, the central limit theorem does not hold. Cauchy distribution belongs to the class of stable distribution and hence is infinitely divisible. The standard Cauchy distribution is connected with Student's t-distribution with one degree of freedom. Several results in Probability Theory concerning expected values such as the strong law of large numbers do not apply to Cauchy distribution.

In particle and nuclear Physics, the energy profile of a resonance is explained by Breit-Wigner distribution, while Cauchy distribution is the BreitWigner distribution. Cauchy distribution is a singular limit of a hyperbolic distribution. Wrapped Cauchy distribution taking values on a circle is derived from Cauchy distribution by wrapping it around the circle. In spectroscopy, Cauchy distribution describes the shape of spectral lines which are subject to homogeneous broadening in which all atoms interact in the same way with the frequency range contained in the line shape. Many mechanisms cause homogeneous broadening, most notably collision broadening, see Hecht (2002). Applications of Cauchy distribution can be found in exponential growth fields, see John (1958). Cauchy distribution is often the distribution of observations for objects that are spinning. In hydrology, Cauchy distribution is applied to extreme events such as annual maximum one-day rainfalls and river discharges. Cauchy distributions can be used to model fat tails in computational finance. It has been applied in various fields such as electrical and mechanical theory, measurements and calibration problems, physical anthropology, modeling depth map data and prices of speculative assets such as stock returns and the phase derivative of air components in an urban environment.

This distribution describes the probability of finding a particle in a particular energy state in quantum mechanics applicable to both excited states of atoms and resonant states of elementary particles. It can be shown that whenever one has a state which decays exponentially with time, the energy width of the state is described as Cauchy distribution, see Roe (1992). Winterton et al. (1992) showed that the source of fluctuations in contact window dimensions is variation in contact resistivity and the contact resistivity is distributed as Cauchy r.v.. Kagan (1992) found that Cauchy distribution is the distribution of hypocenters on focal spheres of earthquakes. Stapf et al. (1996) described an application of this distribution to study the polar and non-polar liquids in porous glasses. Min et al. (1996) described Cauchy distribution is the distribution of velocity differences induced by different vortex elements.

One of the non-normal distributions used for studying the behavior of price data is Cauchy distribution. A major difference between normal and Cauchy distributions is that the latter has a longer tail than the former. This makes a better modeling of the price data using Cauchy distribution. General Insurance companies aim to utilize past or present claim amounts to predict the seriousness of future claims. They must identify an appropriate statistical distribution for their extensive datasets of claim amounts and evaluate the goodness of fit of this distribution to their actual claim data. Most data in general insurance problems exhibit right skewness, distributions with similar characteristics are typically employed to model claim severity. Moreover, insurance data often consists of large claim amounts and large claims hold significant financial importance. However, assessing the distribution of these large claims presents challenges. Therefore, there is a clear necessity to employ statistical distributions that are highly skewed and exhibit relatively heavy tails.

For these probabilistic models, finite moments of the first order and above do not exist. But, they still have applications in other areas of science including analysis of data with extremes and Statistical Inference. Estimating distributional properties such as reliability, hazard rate and others involves determining suitable parameter estimates for the distribution model using appropriate estimation methods. Parameter estimation has recieved significant attention in statistical literature with many researchers conducting comparative studies to assess the performance of various estimation methods numerically. See, for example, Gupta and Debasis (2001); Alkasasbeh and Raqab (2009); Mazucheli et al. (2013); Do Espirito Santo and Mazucheli (2015); Qoshja and Hoxha (2016); Dey et al. (2017); Balakrishnan and Alam (2019) and Alam and Nassar (2021).

Cauchy probability distribution has special features due to its very thick tails. This lead to the problem of estimating the center of the distribution from a sample. It's known that the sample mean is an inconsistent estimator because its sampling variability doesn't decrease as the sample size increases. The maximum likelihood (ML) estimator is consistent and efficient asymptotically. The sample median is the simplest consistent estimator and is inefficient to use in practice. Rothenberg et al. (1964) showed that the average of the middle one-quarter of the ordered observations is an unbiased estimator of θ with a smaller variance than the median itself. Blom (1958), Barnett (1966, 1968) and Bloch (1966) considered the linear estimation of θ based on the sample order statistics, while Sarhan and Greenberg (1962), Chan (1970) and Balmer et al. (1974) discussed the estimation of θ and σ based on a small number of the sample quantiles taken from a large sample.

The one parameter ML estimation is complicated in this case. The joint ML estimation of both location and scale is often simpler than the estimation of location alone. Ferguson (1978) discussed closed form solutions when $n < 5$ for the ML estimators of both θ and σ. Such closed form solutions are not known to exist for $n > 5$. Other charecteristics of ML estimation are discussed in Haas et al. (1970). Koutrouvelis (1982) discussed the estimation of location and scale using the empirical characteristic function.

Ghosh et al. (1982) gave conditions under which the integrated risk of an approximate Bayes estimate (calculated as in Lindley (1980)) approximates the Bayes risk. Franck (1981) considered the problem of testing of normal versus Cauchy and Spiegelhalter (1985) used the methods of Franck to obtain exact expressions for the Bayes estimators of θ and σ under a non informative prior density. These exact expressions are fairly complex and difficult to compute especially for large n. Howlader and Weiss (1985) discussed the exact expressions due to Spiegelhalter (1985) and the imprecise expressions derived by using a method due to Lindley (1980) for Bayes estimation of the location and scale parameters of Cauchy distribution.

Pekasiewicz (2014) introduced the quantile methods to estimate Cauchy distribution. Mahdizadeh and Zamanzade (2017, 2019) discussed some goodness of fit tests for Cauchy distribution and revealed that their tests have a better performance than the existing tests. Also, Ebner et al. (2022) discussed a new characterization of Cauchy distribution and proposed a class of goodness of fit tests for Cauchy family. Alam (2022) discussed another pathological model known as Cauchy Birnbaum-Saunders distribution and some of its goodness of fit tests.

The absence of moments in Cauchy distribution which leads to the breakdown of the law of large numbers has motivated researchers to explore and develop generalizations of this distribution. Indeed, some generalizations of Cauchy distribution have been proposed and studied in the literature; Rider (1957) proposed a generalization of Cauchy distribution, Batschelet (1981) proposed Wrapped-up Cauchy distribution, the skew-Cauchy distribution was introduced by Arnold and Beaver (2000), another class of skew Cauchy distribution was discussed in Behboodian et al. (2006), Huang and Chen (2007) discussed a generalization of the skew Cauchy distribution and Alshawarbeh et al. (2013) used the beta family introduced by Eugene et al. (2002) to generate beta Cauchy distribution. Dahiya et al. (2001) and Nadarajah and Kotz (2006) proposed a truncated Cauchy density to overcome the problem of the non-existence of the moments and the ML estimates. Tahir et al. (2017) studied Weibull power-Cauchy distribution.

Indeed, there is a clear need for more extended distributions in many applied areas like lifetime analysis, insurance and finance where real world data often exhibit high degrees of skewness and kurtosis. Real-life data analysis often involves highly positively or negatively skewed data. These datasets frequently contain extreme observations. Many common distributions lack the flexibility to model these characteristics accurately. As a result, there is a growing demand for more flexible distributions that can effectively handle such data. While there are some distributions available that can accommodate either positive or negative skewness to some extent, there remains a scarcity of distributions that can adequately capture both types of skewness simultaneously. This limitation underscores the necessity for the development of extended forms of distributions that offer greater flexibility and applicability in diverse fields of study and practice.

In this thesis, we obtain some generalization of Cauchy distribution and investigate the performance of some estimation methods for the model parameters. The obtained estimators for the model parameters in this study include the maximum likelihood estimators (MLEs), the least-squares estimators (LSEs), the maximum product spacing estimators (MPSEs), the Cramer-von Mises estimators (CVMEs), the Anderson-Darling estimators (ADEs) and Right-tail Anderson-Darling estimators (RTADEs). Thus, our aim of this study is to tackle this scientific computational challenge by conducting numerical comparison among these estimators for the model parameters to identify which ones perform better in terms of estimation efficiency. This comparison is facilitated through Monte Carlo simulations and the analysis of real data outcomes.

The present research work is mainly concerned with the study of some generalizations of Cauchy distribution.

1.2 Review of Literature

Statistical distributions often exhibit heavy tails that are commonly approximated using either log-normal or power-law functions (Newman (2005)). Since the entries in a heavy tail have disproportionate significance, understanding the precise shape of the tail is very important.

Definition 1.2.1. *A function $f \geq 0$ is said to be heavy-tailed if and only if*

$$\limsup_{x \to \infty} f(x) e^{mx} = \infty \quad for\ all\ m > 0 \tag{1.3}$$

In recent years heavy tail distribution have played an important role in fields such as insurance, finance and risk management. Financial returns are commonly observed to exhibit heavy tails, leading to a significant impact on risk management strategies, which often depend on heavy-tail analysis. For an in depth study on heavy-tailed distributions and their practical applications, we refer to Embrechts et al. (1997), Rachev (2003), Newman (2005), Bingham et al. (1987), Resnick (2007) and Foss et al. (1970). One of the most significant classes of heavy tail distributions is the regular variation class encompassing distributions with tails that exhibit regular variation. This class has been extensivly used in modeling various heavy-tailed phenomena.

1.2.1 Transmuted family of distributions

Recently there has been interest in adding parameters to a distribution for obtaining more flexible new families of distributions. Shaw and Buckley (2009) introduced a method for enhancing the flexibility of existing distributions by adding a new parameter. They employed the quadratic rank transmutation map (QRTM) to create a flexible family of distributions. According to this approach, a r.v. X is said to have transmuted probability distribution if it's cdf is

$$F(x) = (1+\lambda)G(x) - \lambda G(x)^2; \quad |\lambda| \leq 1 \tag{1.4}$$

and the corresponding pdf is

$$f(x) = g(x)[(1+\lambda) - 2\lambda G(x)], \tag{1.5}$$

where $G(x)$ is the cdf of the base distribution, $g(x)$ and $f(x)$ are the corresponding pdf of $G(x)$ and $F(x)$ respectively.

It is note that when $\lambda = 0$, (1.4) reduces to $G(x)$.

A transmuted model provide flexibility in modeling different types of real life data particularly when the parent model fails to provide a better fit. Aryal and Tsokos (2009) introduced transmuted generalized extreme value distribution. Aryal and Tsokos (2011) proposed a generalized Weibull distribution called transmuted Weibull distribution. Khan and King (2012) developed the transmuted generalized inverse Weibull distribution. Aryal (2013) proposed and studied various structural properties of the transmuted log-logistic distribution. Khan and King (2013) introduced the transmuted modified Weibull distribution which extends the transmuted Weibull distribution and studied its mathematical properties. Elbatal (2013) proposed the transmuted modified inverse Weibull distribution. Elbatal and Aryal (2013) proposed and studied transmuted additive Weibull model.

Many transmuted distributions have been proposed in literature (see, Ashour and Eltehiwy (2013), Merovci (2013a,b, 2014),Afify et al. (2015), Ahmad et al. (2015), Khan et al. (2016) and Adeyinka (2019)).

1.2.2 Marshall-Olkin family of distributions

Marshall and Olkin (1997) introduced a family of distributions by adding a new parameter to an existing one is known as Marshall-Olkin family of distributions.
Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.s with survival function (sf) $\bar{F}(x)$. Let N be a geometric r.v. with probability mass function (pmf) $P(N = n) = p(1-p)^{n-1}$, for $n = 1, 2, \ldots$ and $0 < p < 1$ and independent of $X_i's$.
Consider

$$U_N = \min(X_1,\ X_2,\ ..., X_N). \tag{1.6}$$

Let $\bar{G}(.)$ be the sf of U_N. Then,

$$\begin{aligned}
\bar{G}(x) &= P(U_N > x) \\
&= P(\min(X_1,\ X_2,\ ...,\ X_N) > x) \\
&= \sum_{n=1}^{\infty} P(X > x|N = n)P(N = n) \\
&= \sum_{n=1}^{\infty} P^n(X > x)p(1-p)^{n-1} \\
&= p\bar{F}(x)\sum_{n=1}^{\infty}\left[\bar{F}(x)(1-p)\right]^{n-1} \\
&= \frac{p\bar{F}(x)}{1-(1-p)\bar{F}(x)}. \qquad (1.7)
\end{aligned}$$

If $p > 1$ and N is a geometric r.v. with pmf

$$P(N = n) = \frac{1}{p}\left(1 - \frac{1}{p}\right)^{n-1}, \quad n = 1, 2, 3, \dots\ ,$$

then the r.v.

$$V_N = \max(X_1, X_2, \dots, X_N)$$

also having the sf (1.7).

If $X_1, X_2, \dots$ is a sequence of i.i.d. r.v.s with distributions from Marshall-Olkin family (1.7), and if N follows a geometric distribution on $\{1, 2, \dots\}$, then the minimum $\min(X_1, \dots, X_N)$ and the maximum $\max(X_1, \dots, X_N)$ also have distributions within the same family. It is evident that when $p = 1$, we get $\bar{G} = \bar{F}$. Whenever F has a density, the sf $\bar{G}$ given by (1.7) has easily computable densities. Specifically, if F has a density f

and hazard rate r_F, then G has the density g given by

$$g(x) = \frac{pf(x)}{(1-(1-p)\bar{F}(x))^2}, \quad x \in \mathbb{R},\ p > 0,$$

and hazard rate

$$r_G(x) = \frac{r_F(x)}{1-(1-p)\bar{F}(x)}, \quad x \in \mathbb{R},\ p > 0.$$

Sankaran and Jayakumar (2008) gave a physical explanation of Marshall-Olkin family of distributions using proportional odds model. In the analysis of lung cancer data, Bennett (1983) used a proportional odds model where the odds ratio served as the covariates.

Let X be a random sample with cdf $F(x)$ and pdf $f(x)$. Then the proportional odds model with covariates can be expressed as

$$\frac{\bar{G}(x, p(x))}{1-\bar{G}(x, p(x))} = p(x)\frac{\bar{F}(x)}{1-\bar{F}(x)}.$$

Then

$$\bar{G}(x; p(x)) = \frac{p(x)\bar{F}(x)}{1-(1-p(x))\bar{F}(x)},$$

where $p(x)$ is a non negative function of the covariates and $\bar{G}(x; p(x))$ is the sf incorporating these covariates. Treating $p(x)$ as a constant p, yields,

$$\bar{G}(x; p) = \frac{p\bar{F}(x)}{1-(1-p)\bar{F}(x)}.$$

This representation of the sf of Marshall-Olkin family of distributions demonstrates a close relationship between Marshall-Olkin family and the proportional odds model in survival analysis. Some special cases of Marshall-Olkin distribution in the literature include Marshall-Olkin extensions of the Weibull by Cordeiro and Lemonte (2013), Pareto by Alice and Jose (2003) and Ghitany (2005), Logistic by Alice and Jose (2005), Lomax by Ghitany et al. (2007), gamma by Ristic et al. (2007), Burr by Jayakumar and Mathew (2008), uniform by Jose and Krishna (2011), Frechet distribution by Krishna et al. (2013), generalized exponential by Ristic and Kundu (2015), additive Weibull by Afify et al. (2018), inverse power Lindley by Hibatullah et al. (2018), inverse Weibull by Pakungwati et al. (2018) and half logistic distribution by Yegen and Gamze (2018).

1.2.3 Discrete Mittag-Leffler distribution

The function $E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1+\alpha k)}, \; z \in (0,\infty)$ was first proposed by Mittag-Leffler in 1903 (see Erdelyi et al. (1955)). In Feller (1971), the Laplace transforms of $E_\alpha(-x^\alpha)$ with $0 < \alpha \leq 1$, is shown to be $\frac{\lambda^{\alpha-1}}{1+\lambda^\alpha}, \lambda > 0$. But $E_\alpha(-x^\alpha)$ is not a probability distribution. Pillai (1990) showed that $F_\alpha(x) = 1 - E_\alpha(-x^\alpha)$ is a cdf. Hence he named $F_\alpha(x)$ as Mittag-Leffler distribution. We have

$$F_\alpha(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^{\alpha k}}{\Gamma(1+\alpha k)}, \; 0 < \alpha \leq 1, \; x \geq 0 \tag{1.8}$$

and the corresponding pdf is

$$f_\alpha(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^{\alpha k - 1}}{\Gamma(\alpha k)}.$$

The Laplace of transform of (1.8) is

$$\phi(\lambda) = \frac{1}{1+\lambda^\alpha}, \; \lambda \geq 0. \tag{1.9}$$

Mittag-Leffler distribution has received much attention of many researchers recently (see, Jaykumar and Suresh (2003), Lin (1998) and the references there in). Note that when $\alpha = 1$, equation (1.9) reduces to the Laplace transform of exponential distribution.

A discrete version of Mittag-Leffler distribution was proposed by Pillai and Jayakumar (1995).

A r.v. X on $\{0, 1, 2, \ldots\}$ is said to follow discrete Mittag-Leffler (DML) distribution if its pgf is

$$\Phi(s) = \frac{1}{1 + c(1-s)^{\alpha}}, \quad 0 < \alpha \leq 1, \; c > 0, \; |s| \leq 1. \tag{1.10}$$

The DML distribution can be viewed as the distribution of geometric sum of i.i.d. Sibuya r.v.s. In a sequence of independent trails of Bernoulli, let $\frac{\alpha}{k}$ be the success probability in k^{th} trail. Then the number of trails required to obtain the first success has Sibuya distribution (see, Devroye (1993)).

DML is a generalization of geometric distribution, since in equation (1.10), when $\alpha = 1$, we get geometric. Pillai and Jayakumar (1995) obtained some distributional properties of the DML. It is geometrically infinitely divisible, falls within the discrete class L and is normally attracted to stable law. Pillai and Jayakumar (1995) developed autoregressive models with marginals as DML distribution. Jayakumar and Sreenivas (2003) pointed out that if X has Mittag-Leffler distribution and $N_c(.)$ is a unit Poisson process with parameter c, independent of X, then $Y = N_c(X)$ has DML distribution.

1.3 Some basic statistical concepts

1.3.1 Compounding

Let us consider a discrete r.v. N with the pmf given by $P(N = n),\ n \in \mathbb{N} \cup \{0\}$. Denoted by $\Phi(s) = P(N = 0) + sP(N = 1) + s^2P(N = 2) + \dots,\ |s| < 1$, the pgf of the r.v. N. Then the zero truncated distribution of the r.v. Z has the pmf given by

$$P(Z = z) = \frac{P(N = z)}{1 - \Phi(0)},\ z \in \mathbb{N}.$$

If, N is a positive r.v., then $\Phi(0) = 0$.

Let F be the parent distribution. Then the sf of the r.v. $Y = \min(X_1,\ X_2, \dots,\ X_Z)$ is

$$\begin{aligned}
\bar{G}(x) &= P(\min(X_1,\ X_2,\ \dots,\ X_Z) > x) \\
&= \sum_{n=1}^{\infty} P(\min(X_1,\ X_2,\ \dots,\ X_Z) > x) P(Z = n) \\
&= \sum_{n=1}^{\infty} \frac{(\bar{F}(x))^n P(N = n)}{1 - \Phi(0)} \\
&= \frac{\Phi(\bar{F}(x)) - \Phi(0)}{1 - \Phi(0)}.
\end{aligned}$$

The corresponding cdf and pdf are given by

$$G(x) = \frac{1 - \Phi(\bar{F}(x))}{1 - \Phi(0)} \tag{1.11}$$

$$g(x) = \frac{f(x)\Phi'(\bar{F}(x))}{1 - \Phi(0)} \tag{1.12}$$

By inverting the equation (1.11), we reach at the quantile function

$$G^{-1}(u) = F^{-1}\left(1 - \Phi^{-1}(1 - u(1 - \Phi(0)))\right), \quad u \in (0, 1),$$

where F^{-1} represents quantile function of the parent distribution F. The function $G^{-1}(u)$ can be utilized for generating random samples from G. (see, Nadarajah et al. (2013)). By using (1.12), with substitution $v = F(x)$ and properties of cdf, we get an expression for the moment of order $k \in \mathbb{N}$,

$$\begin{aligned} E(X^k) &= \int_a^b x^k g(x)dx = \frac{1}{1-\Phi(0)} \int_a^b x^k \Phi'(\bar{F}(x)) f(x)dx \\ &= \frac{1}{1-\Phi(0)} \int_0^1 \left(F^{-1}(v)\right)^k \Phi'(1-v)dv. \end{aligned}$$

Geometric compounding

If $X \sim$ exponential(λ) with cdf $F(x;\lambda) = 1 - e^{-\lambda x}$, then the cdf of the exponential-geometric model introduced by Adamidis and Loukas (1998) is

$$G(x;p,\lambda) = \frac{1-e^{-\lambda x}}{1-pe^{-\lambda x}}; \quad x, \lambda > 0,\ p \in (0,1).$$

Extended exponential-geometric by Adamidis et al. (2005), Weibull-geometric by Souza et al. (2011), beta-exponential geometric by Bidram (2012), generalized linear failure rate-geometric by Nadarajah et al. (2014), modified Weibull-geometric by Wang and Elbatal (2014), generalized exponential geometric extreme by Ristic and Kundu (2016), are some works related to geometric compounding.

Truncated negative binomial compounding

Nadarajah et al. (2013) proposed a family of life time models based on the truncated negative binomial (TNB) distribution having pmf

$$P(N=n) = \frac{p^\nu}{1-p^\nu}\binom{\nu+n-1}{\nu-1}(1-p)^n; \quad p \in (0,1),\ \nu > 0,\ n = 1, 2, \ldots$$

The authors demonstrated that the random minimum, $U_N = \min(X_1,\ X_2,\ \ldots,\ X_N)$ has the sf of the form

$$\bar{G}(x;p,\nu) = \frac{p^{\nu}}{1-p^{\nu}}[(F(x)+p\bar{F}(x))^{-\nu}-1], \tag{1.13}$$

when X_i's are i.i.d. r.v.s having cdf $F(x)$ and N is TNB with parameter p and ν. Note that if $p \to 1$ then $\bar{G}(x;p,\nu) \to \bar{F}(x)$. When $\nu = 1$, equation (1.13) reduces to equation (1.7). Thatmeans the family of distributions provided in equation (1.13) constitutes a generalization of Marshall-Olkin family of distribution.

Using TNB model, Nadarajah et al. (2013) introduced and studied exponential-TNB, Jose and Remya (2015) studied Rayleigh-TNB, Jayakumar and Sankaran (2016) studied uniform-TNB, Babu (2016) studied Weibull-TNB and Jayakumar and Sankaran (2017) studied generalized exponential-TNB distribution.

1.3.2 Regular variation

A positive measurable function f is called regularly varying [at infinity] with index $\alpha \in \mathbb{R}$ if

- It is defined on some neighbourhood $[x_0, \infty]$ of infinity
- $\lim_{x\to\infty} \frac{f(tx)}{f(x)} = t^{\alpha}$ for all $t > 0$.

If $\alpha = 0$, f is said to be slowly varying (at infinity).

Regularly varying random variables

The regular variation appears in different fields of applied probability including queuing theory, renewal theory, extreme value theory, the theory of summation of r.v.s and point process theory.

A non negative r.v. X and its distribution are considered regularly varying with index $\alpha \geq 0$ if the tail distribution $\bar{F}_X$ is regularly varying with index $-\alpha$, where $\bar{F}_X = 1 - F_X(x), x \in \mathbb{R}$.

Now, we consider the following definitions discussed in Foss et al. (1970) and Omey et al. (2018).

Definition 1.3.1. *An ultimately positive function f is called regularly varying at infinity with index $\alpha \in \mathbb{R}$ if for any fixed $c > 0$,*

$$\lim_{x \to \infty} \frac{f(cx)}{f(x)} = c^{\alpha}. \tag{1.14}$$

Definition 1.3.2. *An ultimately positive function f is said to belongs to the class of long tailed distribution $\mathbb{L}$ if*

$$\lim_{x \to \infty} \frac{f(x+y)}{f(x)} = 1, \quad for\ all\ y > 0. \tag{1.15}$$

if f exhibits a long-tailed behavior, then we can replace y by $-y$ in (1.15).

Remark 1.3.1. *A distribution $F \in \mathbb{L}$ if and only if $\lim_{x \to \infty} h(x) = 0$, where h(x) is the hazard rate function (hrf).*

For details, see Kluppelberg (1988).

Definition 1.3.3. *An ultimately positive function f belong to the class $\mathbb{D}$ of dominated variation distributions if there exists $c > 0$*

$$\lim_{x \to \infty} \frac{f(x)}{f(2x)} = c, \quad for\ all\ \ x > 0. \tag{1.16}$$

Two distributions F and G are said to be tail-equivalent if

$$\lim_{x\to\infty} \frac{\bar{F}(x)}{\bar{G}(x)} = c \ \in \ (0, \infty).$$

An encyclopaedic treatment of regular variation can be found in Bingham et al. (1987). A survey of regular variation is discussed in Seneta (1976). Additionally, several books provide surveys on regularly varying functions and their properties; see for example Feller (1971), Ibragimov (1975), Resnick (1987) and Embrechts et al. (1997).

1.3.3 Subexponential distribution

Let $X_1, X_2, \ldots, X_n$ be i.i.d. non negative regularly varying r.v.s,

Then

$$P(S_n > x) \sim nP(X > x) \sim P(M_n > x), \quad as \quad x \to \infty, n = 2, 3, \ldots \tag{1.17}$$

where

$S_n = X_1 + X_2 + \cdots + X_n$ and $M_n = \max_{1\le i\le n} X_i$.

A non negative r.v. X and its distribution is said to be subexponential if i.i.d. copies X_i of X satisfy relation (1.17).

That is for large x the event $\{S_n > x\}$ is fundamentally due to the event $\{M_n > x\}$. The relation $P(M_n > x) \sim nP(X > x)$ as $x \to \infty$ also implies that M_n is regularly varying with the index same as X.

The subexponential class of distributions was independently introduced by Chistyakov (1964) and Chover et al. (1973) in the context of branching processes. A textbook treatment is given in Athreya et al. (2004). An independent introduction of subexponential distributions through questions in queuing theory is to be found in Borovkov (2012). See also Pakes (1975).

1.3.4 Entropy

Entropy is a measure of variation or uncertainty within a system. In information theory and statistics, entropy quantifies the uncertainty associated with a r.v. or a probability distribution. Higher entropy indicates higher uncertainty or unpredictability whereas lower entropy implies greater predictability or certainty.

The Renyi entropy of a r.v. with pdf $f(.)$ is defined as

$$I_R(\gamma) = \frac{1}{\gamma} \log \int_0^\infty f^\gamma(x) dx, \ \gamma > 0, \ \gamma \neq 1.$$

Various entropy measures are developed by Mathematicians, Physicists and Engineers to discuss several phenomena in the context of communication theory.
The Shannon entropy of a r.v. X is defined by $E[-\log \ f(X)]$. It is a particular case of the Renyi entropy for $\gamma = 1$. The concept of entropy applied in different fields including thermodynamics, statistics, queuing theory, image analysis, stock market analysis and reliability estimation (see, Kapur (1989)). Ebrahimi (2000) described the maximum entropy method for life time distributions.

1.3.5 Stochastic ordering

Stochastic orders have been used over the past forty years at an accelerated rate across numerous areas of probability and statistics. These include reliability theory, survival analysis, biology, economics, queueing theory, insurance and actuarial science (see, Shaked and Shanthikumar (2007)).

Let X and Y be two r.v.s with distribution functions F and G respectively. Denote by $\bar{F}$ $= 1 - F$ and $\bar{G} = 1 - G$ their survival functions with corresponding pdf's f and g. The r.v. X is said to be smaller than Y in the the following stochastic orderings:

(i) Stochastic order (denoted as $X \leq_{st} Y$) if $\bar{F}(x) \leq \bar{G}(x)$ for all x;

(ii)Likelihood ratio order (denoted as $X \leq_{lr} Y$) if $f(x)/g(x)$ is decreasing in $x \geq 0$;

(iii) Hazard rate order (denoted as $X \leq_{hr} Y$)if $\bar{F}(x)/\bar{G}(x)$ is decreasing in $x \geq 0$;

(iv) Reversed hazard rate order (denoted as $X \leq_{rhr} Y$) if $F(x)/G(x)$ is decreasing in $x \geq 0$.

The stochastic orders defined above are interconnected, and they have certain implications among them (see,Shaked and Shanthikumar (2007)):

$$X \leq_{rhr} Y \Leftarrow X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y. \tag{1.18}$$

1.3.6 Information criterion

The fitting of the data set will be usually improved when increasing the number of parameters and likelihood value will also increase. When comparing statistical models that may have different numbers of parameters, using ML estimation alone might favor models with more parameters leading to overfitting. To address this issue, information criteria such as the Akaike Information Criterion (AIC), Corrected Akaike information criterion (AICC), Bayesian Information Criterion (BIC) and Hannan-Quinn information criterion (HQIC) are commonly used. Akaike information criterion due to Akaike (1974) is defined as

$$AIC = -2 \log \ L + 2k,$$

where L is the likelihood function evaluated at the ML estimates and k is the number of unknown parameters. The model which has smallest AIC be the most suitable model to fit the given data set. When the number of parameters k is large or when the sample size n is not large, AIC may tend to favor models with more parameters. To address this issue, Hurvich and Tsai (1989) introduced a corrected version of AIC denoted as AICC which is defined as follows:

$$AICC = AIC + \frac{2k(k+1)}{n-k-1},$$

where n is the sample size.

The Bayesian information criteria due to Schwarz (1978) is defined by

$$BIC = -2\log\ L + k\log(n).$$

The Hanna-Quinn information criteria due to Hannan and Quinn (1979) is defined by

$$HQIC = -2\log\ L + 2k\log(\log(n)).$$

The Deviance Information Criterion (DIC) serves as a hierarchical modeling generalization of the AIC. It is employed in Bayesian model selection problems where the posterior distributions of the models are obtained via Markov chain Monte Carlo (MCMC) simulation. DIC provides an asymptotic approximation as the sample size becomes large, similar to AIC. See details in Spiegelhalter et al. (2002) and Spiegelhalter et al. (2014).

1.3.7 Goodness of fit

There are various methods used for testing whether a given random sample is coming from a specified distribution or for comparing the underlying model for fitting a given data set.

Cramer-von Mises and Anderson-Darling test

Cramer-von Mises (W^*) and Anderson-Darling (A^*) statistic are commonly used to assess the goodness of fit of a specified distribution to a dataset. In general, the smaller the values of W^* and A^*, indicate better fit to the data. Let $G(x;\theta)$ be the cdf, where the form of G is known but the parameter θ is unknown, we compute the statistics W^* and A^* as follows:

(i) Calculate $\psi_i = G(x_i;\hat{\theta})$, where the $x_i's$ are in ascending order;

(ii) Calculate $x_i = \Phi^{-1}(\psi_i)$, where $\Phi(.)$ is the normal cdf and $\Phi^{-1}(.)$ its inverse;

(iii) Calculate $u_i = \Phi\{(x_i - \bar{x})/s_x\}$, where $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$ and $s_x^2 = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$;

(iv) Compute

$$W^2 = \sum_{i=1}^{n}\left\{u_i - \frac{(2i-1)}{2n}\right\}^2 + \frac{1}{12n}$$

and

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}\{(2i-1)\log(u_i) + (2n+1-2i)\,\log(1-u_i)\};$$

(v) Modify W^2 into $W^* = W^2(1+0.5/n)$ and A^2 into $A^* = A^2(1+0.75/n+2.25/n^2)$. For further details, see Chen and Balakrishnan (1995).

Kolmogrov-Smirnov test

Kolmogorov (1933) introduced Kolmogrov-Smirnov (K-S) test for assessing whether a given random sample $x_1,\ x_2,\ \ldots,\ x_n$ belongs to a population with a specific distribution. K-S test compute the distance between the empirical distribution function of the given sample and the estimated cdf of the candidate distribution. The null and alternative hypothesis are stated as follows:

H_0: sample follow specific distribution versus H_1 : H_0 is false.

Let $G(x_i)$ denotes the value of the cdf of the candidate distribution at x_i and $\hat{G}(x_i)$ denotes the value of the empirical distribution function at x_i. The value of K-S test statistic is defined by

$$D = \sup_{x_i}(|G(x_i) - \hat{G}(x_i)|); \quad i = 1, 2, \ldots, n,$$

where $\hat{G}(x_i) = \frac{\#\{x_j : x_j \leq x_i\}}{n}$.

After computing K-S statistic, it is compared with the tabulated critical value of K-S statistic at a significance level α to determine whether to reject the null hypothesis. Additionally, when comparing multiple distributions, the distribution with the smaller K-S statistic value is considered more appropriate.

1.4 Objectives of the present work

The present study has been undertaken with the following specific objectives:

1. To develop new families of Cauchy distributions.

2. To study heavy tail properties and different methods of estimation of these proposed generalized distributions.

3. To introduce new transformation method for generating distributions, study the estimation and application of this new method of transformation by considering Cauchy model as a parent distribution.

4. To study the feasibility of these models in modeling real data.

1.5 Organization of the present work

The present research work is mainly concerned with study of some generalizations of Cauchy distribution and their applications. Chapter 1 introduces the area of research and contains review of literature, basic statistical concepts and a brief summary of work done.

In Chapter 2, we study a three parameter Cauchy distribution called transmuted Cauchy (TC) distribution. We study its tail behavior, discuss truncated TC distribution, derive the expression for characteristic function, quantiles, mode and mean deviation. We discuss the estimation methods such as maximum likelihood (ML) estimation, method of maximum product spacings (MPS), least square (LS) estimation, Cramer-von Mises (CVM) method, Anderson-Darling (AD) method, right-tail Anderson-Darling estimation (RTAD) and Bayesian framework for the estimation of parameters of TC distribution. The mathematical findings are then verified through simulation studies. A first order autoregressive (AR(1)) minification model with TC distribution as marginal is developed. Two real data applications of the TC distribution are presented.

In Chapter 3, we introduce a new generalization of Cauchy distribution as a competitor for several generalizations of Cauchy distribution. The shape properties, quantiles, mode, Mean deviation and pdf of order statistics are derived. Study the tail properties of the model and it is shown that the distribution has regularly varying tail. It belongs to the class of long-tailed distributions and is a member of the dominated variation distribution. Hence it belongs to the class of subexponential distributions. ML, MPS, LS, CVM, AD and RTAD estimation methods are used to estimate the parameters of the new model. Monte Carlo simulation is conducted to investigate the performance of the estimates. The existence and uniqueness of the ML estimates are proved. A first order autoregressive (AR(1)) minification model with GCD as marginal is developed. Two sets of real data are analyzed to illustrate the use of the proposed distribution.

We introduce another generalization of Cauchy distribution using truncated discrete Mittag-Leffler distribution called Discrete Mittag-Leffler Cauchy (DMLC) distribution in Chapter 4. Expressions for the quantiles, mode, mean deviation and distribution of order statistics are derived. The tail behaviour of DMLC model is studied. Parameters of the distribution are estimated by the method of ML, MPS, LS, CVM, AD and RTAD. Monte Carlo simulation is carried out to investigate the performance of the estimates. Applications of two sets of real data are presented to exhibit the performance of the new model over other generalizations of Cauchy distribution.

In Chapter 5, we introduce a transformation that yields new models by using a given baseline distribution known as beta transformation. It contains only one new parameter other than the parameters involved in the baseline distribution. To demonstrate the use of this new transformation, we choose Cauchy as the baseline distribution namely beta transformed Cauchy (BTC). We derive the quantiles, mode, mean deviation and pdf of order statistics. Certain characterizations of the proposed distribution are obtained. We demonstrate the flexibility of the model using a real data set. The method of ML, LS, CVM and AD estimation are applied to estimate the model parameters.

The recommendations and future works of the thesis are presented in Chapter 6.

1.6 Papers published

1. Jayakumar, K. and Fasna, K. (2023). Half Cauchy - exponential distribution: Estimation and Applications. *Reliability Theory and Applications*, 18, 387-401.

2. Fasna, K. (2023). Discrete Mittag- Leffler Cauchy distribution: Estimation and Applications. *Journal of Scientific Research*, 67, 110-118.

3. Jayakumar, K. and Fasna, K. (2022). On a new generalization of Cauchy distribu-

tion. *Asian Journal of Statistical Sciences*, 2, 61-81.

4. Fasna, K. (2022). A method for generating lifetime models and its application to real data. *Reliability Theory and Applications*, 17, 344-357.

1.7 Papers presented

1. Presented the paper entitled “A new generalization of Cauchy distribution and its applications” in the International Conference on Statistics for Twenty-first Century-2020 (ICSTC-2020), organized by the Department of Statistics, University of Kerala, Trivandrum during December 16-19, 2020.

2. Presented the paper entitled “Transmuted Cauchy Distribution and Its Applications” in the International Conference in Statistics in the Era of Pandemic (STEP), organized by the Research Department of Statistics, Nehru Arts ad Science College, Kanhankad, Kannur University during March 04-06, 2021.

3. Presented the paper entitled “Discrete Mittag- Leffler Cauchy distribution: Estimation and Applications” in the International Conference on Emerging Trends in Statistics and Data Science, organized by the Departments of Statistics of Cochin University of Science and Technology, Cochin, M.D. University, Rohtak, University of Kerala, Trivandrum, Bharathiar University, Coimbatore, and Madura College (Autonomous), Madurai during September 07-10, 2021.

Chapter 2

ESTIMATION AND APPLICATIONS OF TRANSMUTED CAUCHY DISTRIBUTION

2.1 Introduction

An interesting idea of generalization where the distribution is derived using the QRTM was introduced by Shaw and Buckley (2009). A theoretical interpretation of the construction through transmuted mapping was given by Kozubowski and Podgorski (2016). They showed that the transmuted distributions can be viewed as the distribution of maxima (or minima) of a random number N of i.i.d. r.v.s with the base distribution G(x), where N has Bernoulli distribution shifted up by one. So that the transmuted models are the special case of extremal distributions defined through a more general N.

More specifically, for $\lambda \in [-1, 0]$,

$$Y = \begin{cases} max(X_1, X_2) & \text{with probability } p = \lambda \\ X_1 & \text{with probability } 1 - p \end{cases} \tag{2.1}$$

and for $\lambda \in [0, 1]$

$$Y = \begin{cases} min(X_1, X_2) & \text{with probability } p = \lambda \\ X_1 & \text{with probability } 1 - p \end{cases} \tag{2.2}$$

where F(x) is the cdf of Y and X_1, X_2 are i.i.d. random variables having cdf G(x).
In this chapter, we study a three parameter Cauchy distribution called Transmuted Cauchy (TC) distribution. The TC distribution was introduced by Ball et al. (2021). They only studied some mathematical properties and the ML estimation of the unknown parameters. The aim of this chapter is to study its tail behaviour, discuss truncated TC distribution, derive the expression for characteristic function, quantiles and mean deviation. Also we focus on different estimation methods of the TC model. The parameters of the probability distributions are estimated using ML, MPS, LS, CVM, AD, RTAD and also Bayesian framework is explored using different priors upon parameters. An application in autoregressive time series modeling is presented. We present two real data applications which exhibits the performance of TC model compared to other models. Finally, the summary is presented.

2.2 Transmuted Cauchy distribution

2.2.1 Probability density function

A r.v X on $(-\infty, \infty)$ is said to have transmuted Cauchy probability distribution with parameters $-1 \leq \lambda \leq 1$, $-\infty \leq \mu \leq \infty$ and $\theta > 0$ denoted by $TC(\lambda, \mu, \theta)$ if its pdf is given by

$$f(x) = \frac{1}{\pi\theta} \frac{1}{(1 + (\frac{x-\mu}{\theta})^2)} \left[(1+\lambda) - 2\lambda \left(\frac{1}{\pi} \arctan \left(\frac{x-\mu}{\theta} \right) + 0.5 \right) \right], \tag{2.3}$$

where λ is the transmutation parameter.

The pdf plots of $TC(\lambda, \mu, \theta)$ for various values of the parameters are presented in Figure 2.1.

2.2.2 Distribution function

The cdf of the $TC(\lambda, \mu, \theta)$ distribution is given by

$$F(x) = (1+\lambda) \left[\frac{1}{\pi} \arctan \left(\frac{x-\mu}{\theta} \right) + 0.5 \right] - \lambda \left[\frac{1}{\pi} \arctan \left(\frac{x-\mu}{\theta} \right) + 0.5 \right]^2. \tag{2.4}$$

The cdf plots of $TC(\lambda, \mu, \theta)$ for various values of the parameters are given in Figure 2.2.

Remark 2.2.1. *When $\lambda = 0$, TC reduces to Cauchy distribution with parameters μ and θ.*

It can be seen that no moments of TC(λ, μ, θ) distribution exist. However the median and mode do exist.

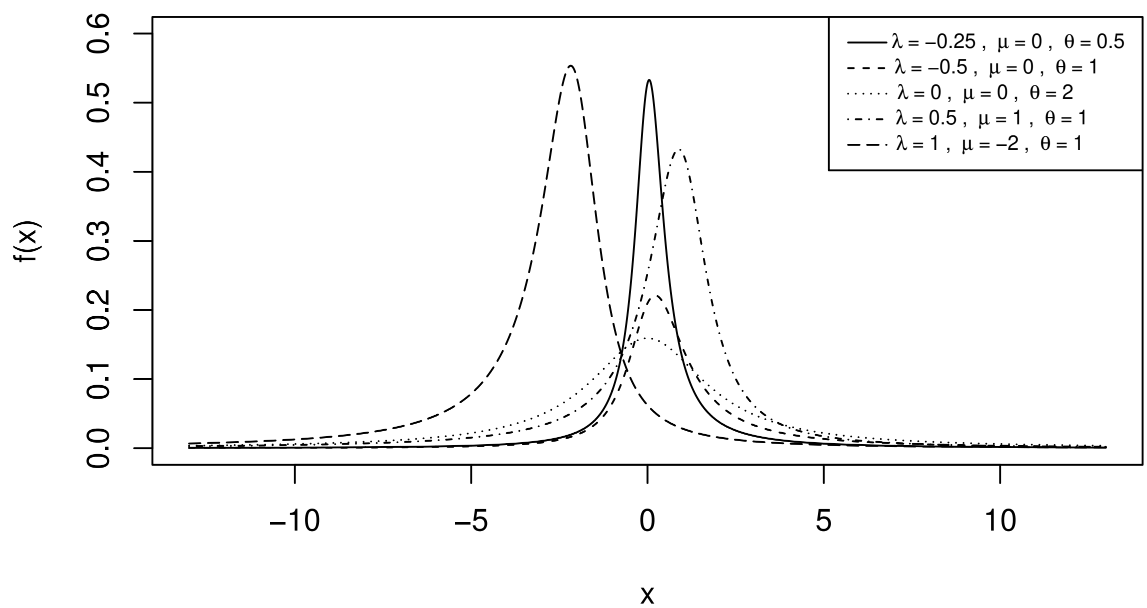


Figure 2.1: Plots of the pdf of $TC(\lambda, \mu, \theta)$ distribution

2.2.3 Tail behaviour

The TC distribution has heavy tail, that is it takes extreme values with high probability. This feature empirically distinguishes TC from the normal and many other distributions. Figure 2.3 plots the right tails of density of TC and compare it with Cauchy and normal densities. TC distribution has tails thicker than both Cauchy and normal.

We can easily shows that $\limsup_{x\to\infty} f(x)e^{mx} = \infty$ for any $m > 0$, Hence the TC density f is heavy tailed.

Theorem 2.2.1. *The TC distribution belongs to the class of long tailed distribution* $\mathbb{L}$.

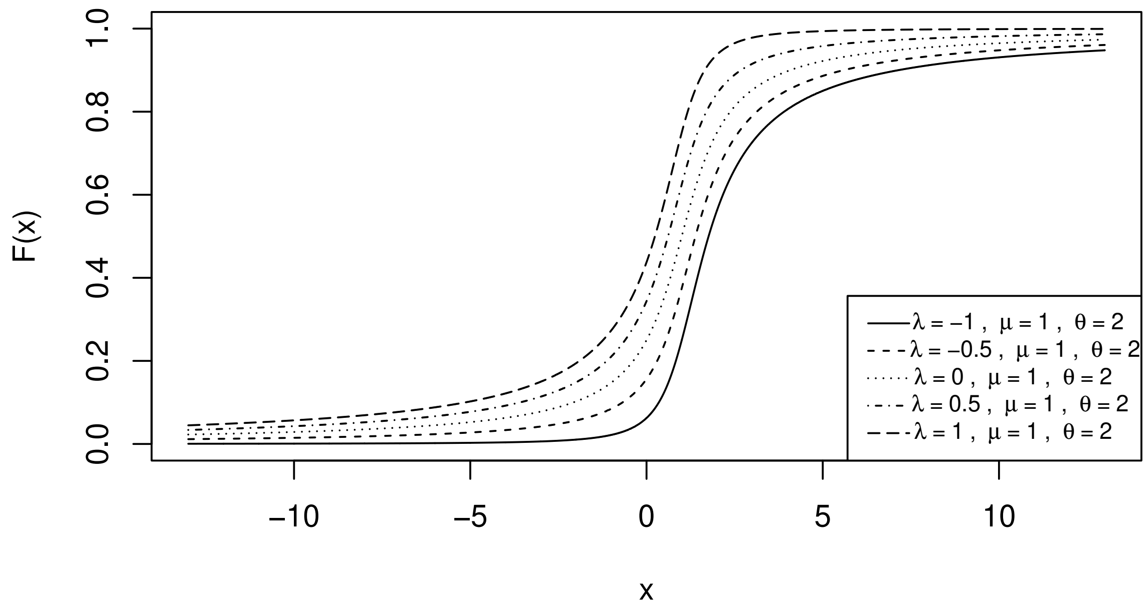


Figure 2.2: Plots of the cdf of $TC(\lambda, \mu, \theta)$ distribution

Proof.

$$\begin{aligned}\lim_{x\to\infty} \frac{f(x+y)}{f(x)} &= \lim_{x\to\infty} \frac{\frac{1}{\pi\theta}\frac{1}{(1+(\frac{(x+y)-\mu}{\theta})^2)}[(1+\lambda) - 2\lambda(\frac{1}{\pi}\arctan(\frac{(x+y)-\mu}{\theta}) + 0.5)]}{\frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}[(1+\lambda) - 2\lambda(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5)]} \\ &= 1.\end{aligned}$$

Therefore f belongs to the class $\mathbb{L}$, the class of long tailed distributions. □

Two distributions F and G are said to be tail-equivalent if

$$\lim_{x\to\infty} \frac{\bar{F}(x)}{\bar{G}(x)} = c \in (0, \infty).$$

Hence it can be shown that TC and Cauchy distribution are tail-equivalent.

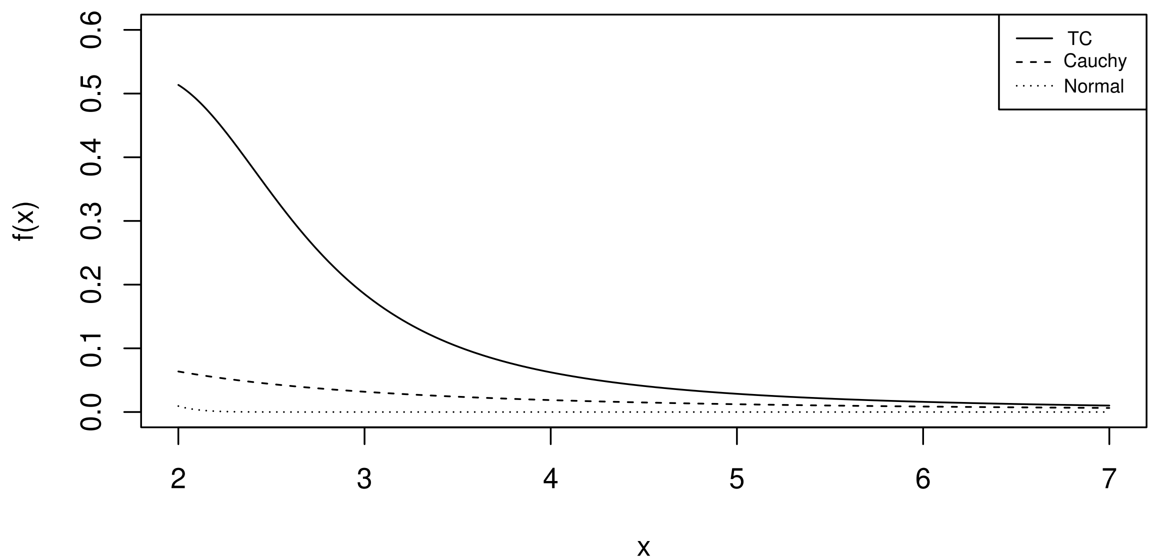


Figure 2.3: Comparison of right tails of Cauchy, Normal and TC densities.

2.2.4 Characteristic function

Theorem 2.2.2. *If X be the TC(λ, μ, θ), then the characteristic function of X is*

$$\phi_X(t) = e^{it\mu - \theta|t|} - \frac{2\lambda}{\pi^2\theta}\int_{x=-\infty}^{\infty} \frac{e^{itx}\arctan(\frac{x-\mu}{\theta})}{(1+(\frac{x-\mu}{\theta})^2)}dx.$$

Proof. Let X has TC(λ, μ, θ) distribution, Then the characteristic function of X is $\phi_X(t) = E(e^{itX})$

$$= \int_{-\infty}^{\infty} e^{itx}\frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}\left[(1+\lambda) - 2\lambda\left(\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right) + 0.5\right)\right]dx$$

$$= \frac{(1+\lambda)}{\pi\theta}\int_{-\infty}^{\infty}\frac{e^{itx}}{(1+(\frac{x-\mu}{\theta})^2)}dx - \frac{2\lambda}{\pi^2\theta}\int_{-\infty}^{\infty}\frac{e^{itx}\arctan(\frac{x-\mu}{\theta})}{(1+(\frac{x-\mu}{\theta})^2)}dx$$

$$-\frac{\lambda}{\pi\theta}\int_{-\infty}^{\infty}\frac{e^{itx}}{(1+(\frac{x-\mu}{\theta})^2)}dx$$

$$=(1+\lambda)[e^{it\mu-\theta|t|}]-\frac{2\lambda}{\pi^2\theta}\int_{-\infty}^{\infty}\frac{e^{itx}\arctan(\frac{x-\mu}{\theta})}{(1+(\frac{x-\mu}{\theta})^2)}dx-\lambda[e^{it\mu-\theta|t|}]$$

On simplification, we get

$$\phi_X(t)=e^{it\mu-\theta|t|}-\frac{2\lambda}{\pi^2\theta}\int_{-\infty}^{\infty}\frac{e^{itx}\arctan(\frac{x-\mu}{\theta})}{(1+(\frac{x-\mu}{\theta})^2)}dx, \tag{2.5}$$

which completes the proof. □

2.2.5 Quantile function

Theorem 2.2.3. *The q^{th} quantile x_q of the $TC(\lambda,\mu,\theta)$ distribution is given by*

$$x_q=\mu+\theta\tan\left[\pi\left[\frac{(1+\lambda)\pm\sqrt{(1+\lambda)^2-4q\lambda}}{2\lambda}-0.5\right]\right] \tag{2.6}$$

Proof. Let the q^{th} quantile x_q of the TC distribution is defined as

$$q=P(X\leq x_q)=F(x_q), \quad x_q\in\mathbb{R},$$

which implies

$$(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]-\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2-q=0.$$

That is,

$$\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+q=0 \tag{2.7}$$

Considering this as quadratic in $[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5]$, we get

$$x_q = \mu + \theta \tan\left[\pi\left[\frac{(1+\lambda) \pm \sqrt{(1+\lambda)^2 - 4q\lambda}}{2\lambda} - 0.5\right]\right].$$

This completes the proof. □

Using the inversion method, we can generate r.v.s from the TC distribution by using equation (2.6) when the parameters λ, μ and θ are known.
The median of TC is given by,

$$x_{0.5} = \mu + \theta \tan\left[\pi\left[\frac{(1+\lambda) \pm \sqrt{(1+\lambda)^2 - 2\lambda}}{2\lambda} - 0.5\right]\right]. \tag{2.8}$$

We consider measures based on quantiles to show the effect of the transmuted parameter λ on skewness and kurtosis. There are many heavy tailed distributions for which the measures are infinite and uniformative.
Bowley's measure of skewness is given by

$$S = \frac{Q_{3/4} + Q_{1/4} - 2Q_{1/2}}{Q_{3/4} - Q_{1/4}},$$

and
Moor's measure of kurtosis is given by

$$K = \frac{Q_{3/8} - Q_{1/8} + Q_{7/8} - Q_{5/8}}{Q_{6/8} - Q_{2/8}},$$

where $Q(.)$ denotes the quantile function of X. These measures are exist even for distributions without moments. Skewness measures the degree of asymmetry or long tail in the distribution, while kurtosis quantifies the degree of peakedness or fatness in the tails.

When the distribution is symmetric, $S = 0$ and when the distribution is right (or left) skewed, $S > 0$ (or $S < 0$). As K increases, the tail of the distribution becomes heavier. Table 2.1, below presents Bowley's skewness and Moor's kurtosis for selected values of μ, θ and λ. Also, the results from Table 2.1 indicate that the TC distribution can be left-skewed or right-skewed. Kurtosis of the model showing that the proposed distribution is leptokurtic.

Table 2.1: Bowley's skewness and Moor's kurtosis for some values of μ, θ and λ

μ	θ	λ	*Skewness*	*Kurtosis*
5	0.5	-0.2	-0.1595	5.5886
		-0.4	0.0148	4.9669
		-0.6	0.2596	3.2892
		-0.8	0.0840	3.1927
4.5	0.2	0.2	-0.0110	4.7965
		0.4	0.0609	4.0412
		0.6	-0.1424	4.9449
		0.8	-0.0768	4.1966

2.2.6 Mode

Theorem 2.2.4. *The mode of the* $TC(\lambda, \mu, \theta)$ *is the solution of the equation* $k(x) = 0$, *where*

$$k(x) = \pi(x-\mu)\left(1 - \frac{2\lambda}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right) + \lambda.$$

Proof. The critical point of TC density function are the roots of the equation

$$\frac{\partial \log(f(x))}{\partial x} = 0$$

That is

$$\frac{\partial \log(f(x))}{\partial x} = \frac{2(x-\mu)}{\theta} + \frac{2\lambda}{\pi(1 - \frac{2\lambda}{\pi \arctan(\frac{x-\mu}{\theta})})}. \tag{2.9}$$

The critical values of equation (2.9) are the solution of $k(x) = 0$. Hence the proof. $\square$

2.2.7 Mean deviation

The mean deviation about the median can be used as measures of the degree of scatter in a population. Let M be the median of TC distribution given by equation (2.8).
The mean deviation about the median can be calculated as

$$\delta(X) = E|X - M| = \int_{-\infty}^{\infty} |x - M| f(x) dx.$$

Hence we obtain the mean deviation about median, $\delta = \mu - 2J(M)$
where, J(q) is

$$J(q) = \frac{1}{\pi\theta} \int_{-\infty}^{q} \frac{x}{(1 + (\frac{x-\mu}{\theta})^2)} \left[(1 + \lambda) - 2\lambda \left(\frac{1}{\pi} \arctan\left(\frac{x - \mu}{\theta}\right) + 0.5\right)\right] dx. \quad (2.10)$$

We can compute this integral numerically using softwares such as R, MATLAB or Mathcad and hence obtain the mean deviation about the median.

2.2.8 Truncated transmuted Cauchy distribution

A truncation version of Transmuted Cauchy distribution denoted by $TTC(\lambda, \mu, \theta)$ has the pdf given by

$$f(x) = \frac{1}{\theta} \frac{1}{(1 + (\frac{x-\mu}{\theta})^2)} \frac{[(1 + \lambda) - 2\lambda(\frac{1}{\pi} \arctan(\frac{x-\mu}{\theta}) + 0.5)]}{\arctan(\frac{B-\mu}{\theta}) - \arctan(\frac{A-\mu}{\theta})}, \quad (2.11)$$

for $-\infty < A \leq x \leq B < \infty$,$-\infty < \mu < \infty, \theta > 0$ and $-1 \leq \lambda \leq 1$.
The pdf plots of $TTC(\lambda, \mu, \theta)$ for various values of the parameters are presented in Figure 2.4.

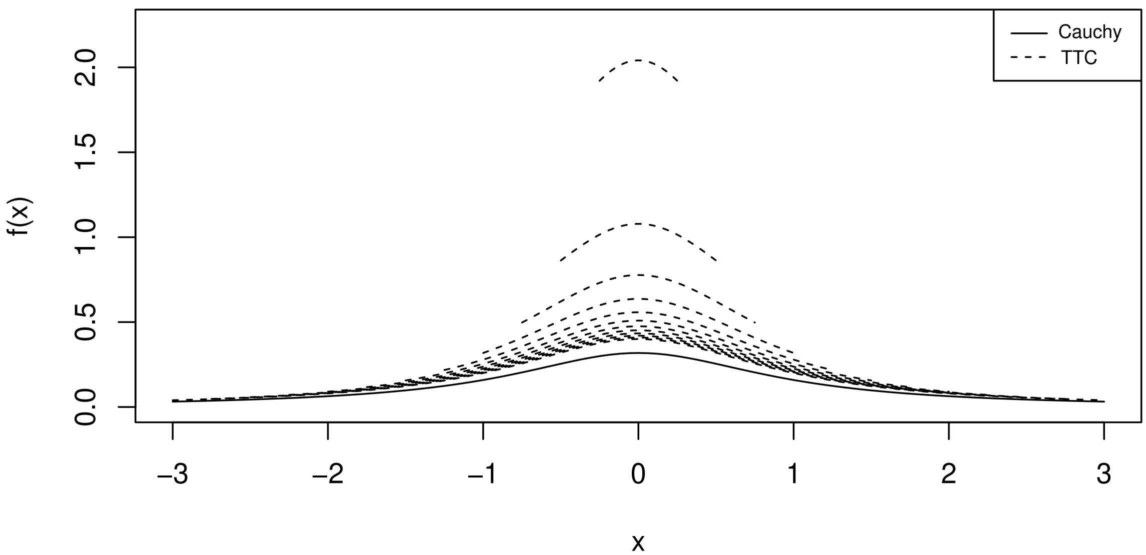


Figure 2.4: Plots of the pdf of $TTC(\lambda, \mu, \theta)$ distribution.

2.3 Estimation of parameters

In this section, we use ML, MPS, LS, CVM, AD, RTAD and Bayesian approach for estimation.

2.3.1 Maximum likelihood estimation

If the parameters of the TC distribution are unknown, then the ML estimates of the parameters are given as follows.
Consider a random sample $x_1, x_2, \ldots, x_n$ of size n from the $TC(\lambda, \mu, \theta)$ distribution. Then, the log likelihood function is given by,

$$\log(L) = -n\log(\pi\theta) - \sum_{i=1}^{n} \log\left(1 + \left(\frac{x_i - \mu}{\theta}\right)^2\right)$$

$$+\sum_{i=1}^{n}\log\left((1+\lambda)-\frac{2\lambda\arctan(\frac{x_i-\mu}{\theta})}{\pi}-\lambda\right). \tag{2.12}$$

Therefore, the ML estimates of λ,μ and θ are derived from the derivatives of $\log L$. They should satisfy the following equations:

$$\frac{\partial\log(L)}{\partial\lambda}=-\frac{2}{\pi}\sum_{i=1}^{n}\frac{\arctan(\frac{x_i-\mu}{\theta})}{\left[1-\frac{2\lambda\arctan(\frac{x_i-\mu}{\theta})}{\pi}\right]}=0, \tag{2.13}$$

$$\begin{aligned}\frac{\partial\log(L)}{\partial\mu}&=\frac{1}{\theta^2}\sum_{i=1}^{n}\frac{2(x_i-\mu)}{1+(\frac{x_i-\mu}{\theta})^2}+\frac{1}{\pi\theta}\sum_{i=1}^{n}\frac{2\lambda}{(1+(\frac{x_i-\mu}{\theta})^2)\left[1-\frac{2\lambda\arctan(\frac{x_i-\mu}{\theta})}{\pi}\right]}\\&=0,\end{aligned} \tag{2.14}$$

and

$$\begin{aligned}\frac{\partial\log(L)}{\partial\theta}&=\frac{-n}{\theta}+\frac{2}{\theta^3}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{1+(\frac{x_i-\mu}{\theta})^2}+\frac{2\lambda}{\pi\theta^2}\sum_{i=1}^{n}\frac{(x_i-\mu)}{[1+(\frac{x_i-\mu}{\theta})^2]\left[1-\frac{2\lambda\arctan(\frac{x_i-\mu}{\theta})}{\pi}\right]}\\&=0.\end{aligned} \tag{2.15}$$

These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

2.3.2 Method of maximum product spacings

This method was introduced by Cheng and Amin (1983) as an alternative to method of ML estimation. The cdf of TC distribution is given by equation (2.4), and the uniform spacing are defined as follows:

$$D_1 = F(x_1) = (1+\lambda)\left[\tfrac{1}{\pi}\arctan\left(\tfrac{x_1-\mu}{\theta}\right)+0.5\right] - \lambda[\tfrac{1}{\pi}\arctan(\tfrac{x_1-\mu}{\theta})+0.5\big]^2,$$

$$D_{n+1} = 1 - F(x_n) = 1 - \left[(1+\lambda)\left[\tfrac{1}{\pi}\arctan\left(\tfrac{x_n-\mu}{\theta}\right)+0.5\right] - \lambda\left[\tfrac{1}{\pi}\arctan\left(\tfrac{x_n-\mu}{\theta}\right)+0.5\right]^2\right],$$

and the general term of spacing is given by

$$D_i = F(x_i) - F(x_{i-1})$$

such that $\sum D_i = 1$.

MPS choose the estimates which maximizes the product of spacings or which maximizes the geometric mean of the spacing. That is, we find estimates suchthat

$$G = \left[\prod_{i=1}^{n+1} D_i\right]^{\frac{1}{n+1}},$$

is maximized. By taking the logarithm of G, we get

$$H = \frac{1}{n+1}\sum_{i=1}^{n+1}\log(D_i).$$

Partially differentiating the above equations with respect to the parameters λ, μ and θ respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

2.3.3 Method of least square estimation

The LS estimators were proposed by Swain et al. (1988) to estimate the parameters of beta distributions. Here, we use the same technique for the TC distribution. The LS estimators of the unknown parameters λ, μ and θ of TC distribution can be given by minimizing

$$\sum_{i=1}^{n}\left[F(x_i \mid \lambda,\mu,\theta)-\tfrac{i}{n+1}\right]^2$$

$$=\sum_{i=1}^{n}\left[\left((1+\lambda)\left[\tfrac{1}{\pi}\arctan\left(\tfrac{x_i-\mu}{\theta}\right)+0.5\right]-\lambda\left[\tfrac{1}{\pi}\arctan\left(\tfrac{x_i-\mu}{\theta}\right)+0.5\right]^2\right)-\tfrac{i}{n+1}\right]^2$$

Partially differentiating the above equations with respect to the parameters λ, μ and θ respectively and equating them to zero, we get the normal equations. They do not have exact solutions and have to be obtained numerically by using statistical softwares like *optim* package in R programming.

2.3.4 Method of Cramer-von Mises

CVM type minimum distance estimators are based on minimizing the distance between the theoretical and empirical cumulative distribution functions. Macdonald (1971) provided empirical evidence that the bias of these estimators is smaller than the bias of other minimum distance estimators. The CVM estimators $\hat{\lambda}_{CME}$, $\hat{\mu}_{CME}$ and $\hat{\theta}_{CME}$ are the values of λ, μ and θ minimizing

$$\begin{aligned}C(\lambda,\mu,\theta)&=\frac{1}{12n}+\sum_{i=1}^{n}\left[F(x_i \mid \lambda,\mu,\theta)-\frac{2i-1}{2n}\right]^2\\&=\frac{1}{12n}+\sum_{i=1}^{n}\left[\left((1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)+0.5\right]-\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)+0.5\right]^2\right)-\frac{2i-1}{2n}\right]^2.\end{aligned}$$

Partially differentiating the above equations, with respect to the parameters λ, μ and θ respectively and equating them to zero, we get the normal equations. They do not have exact solutions and have to be obtained numerically by using statistical softwares like *optim* package in R programming.

2.3.5 Methods of Anderson-Darling and right-tail Anderson-Darling

The method of AD test was developed by Anderson and Darling (1952) as an alternative to statistical tests for detecting sample distributions departure from normality.
The AD estimators $\hat{\lambda}_{ADE}$, $\hat{\mu}_{ADE}$ and $\hat{\theta}_{ADE}$ are the values of λ, μ and θ minimizes

$$\begin{aligned}
A(\lambda,\mu,\theta) &= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log F(x_i \mid \lambda,\mu,\theta) + \log \bar{F}(x_{n+1-i} \mid \lambda,\mu,\theta)\} \\
&= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\Bigg\{\log\Bigg[(1+\lambda)\left(\frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)+0.5\right) \\
&- \lambda\left(\frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)+0.5\right)^2\Bigg] \\
&+ \log\Bigg[1-\Bigg[(1+\lambda)\left(\frac{1}{\pi}\arctan\left(\frac{x_{n+1-i}-\mu}{\theta}\right)+0.5\right) \\
&- \lambda\left(\frac{1}{\pi}\arctan\left(\frac{x_{n+1-i}-\mu}{\theta}\right)+0.5\right)^2\Bigg]\Bigg]\Bigg\}.
\end{aligned}$$

The RTAD estimators $\hat{\lambda}_{RTADE}$, $\hat{\mu}_{RTADE}$ and $\hat{\theta}_{RTADE}$ are the values of λ, μ, and θ minimizes

$$
\begin{aligned}
R(\lambda, \mu, \theta) &= \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i \mid \lambda, \mu, \theta) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log \bar{F}(x_{n+1-i} \mid \lambda, \mu, \theta) \\
&= \frac{n}{2} - 2\sum_{i=1}^{n}\left[(1+\lambda)\left(\frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)+0.5\right)\right. \\
&\left. -\lambda\left(\frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)+0.5\right)^2\right] \\
&-\frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\left[1-\left[(1+\lambda)\left(\frac{1}{\pi}\arctan\left(\frac{x_{n+1-i}-\mu}{\theta}\right)+0.5\right)\right.\right. \\
&\left.\left. -\lambda\left(\frac{1}{\pi}\arctan\left(\frac{x_{n+1-i}-\mu}{\theta}\right)+0.5\right)^2\right]\right].
\end{aligned}
$$

Partially differentiating the above equations with respect to the parameters λ, μ and θ respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

2.3.6 Simulation study

By conducting Monte Carlo simulation to compare the performance of the estimators discussed in the previous sections and the process is repeated 1000 times. We analyze the performance of the estimators based on bias and mean squared error (MSE). Methods are compared for sample sizes $n = 100$, 300 and $n = 500$.

We calculate the bias, MSE for each estimate. The statistics are obtained by the following formulae.

$Bias(\hat{\lambda}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\lambda} - \lambda)$ $\quad$ $Bias(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu} - \mu)$

$Bias(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta} - \theta)$

$MSE(\hat{\lambda}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\lambda} - \lambda)^2$ $\quad$ $MSE(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu} - \mu)^2$

$MSE(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta} - \theta)^2$

The bias (estimate-actual) and the MSEs of the parameter estimates for the ML, MPS, LS, CVM, AD and RTAD estimation are presented in Tables 2.2 and 2.3.

From Table 2.2 and 2.3, we note that the ML method works well for estimating the model parameters. Also, as the sample size increases the biases and the MSEs of the average estimates of ML estimates decreases.

The following observations can be drawn from Tables 2.2 and 2.3.

1. All the estimators exhibit the property of consistency, i.e., the MSE decreases as the sample size increases.

2. The bias of all parameters decreases with an increasing n for all the estimation methods.

3. The bias of $\hat{\lambda}$, $\hat{\mu}$ generally increases with an increasing lamda, mu for any given lamda, mu and n and for all estimation methods.

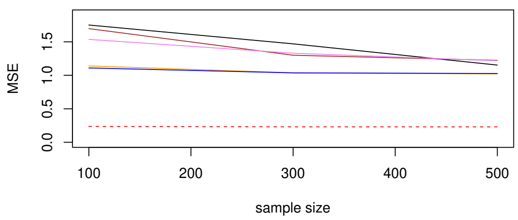
4. In all the methods of estimation produce smaller MSE for $\hat{\theta}$ compared to that of other parameters.

Furthermore, the results in these tables are depicted graphically in Figures 2.5 and 2.6

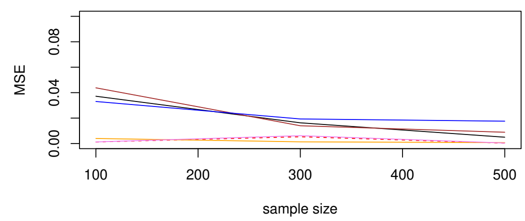
Table 2.2: Simulation results for $\lambda = 0.5$, $\mu = 5$, and $\theta = 0.3$.

n	Parameters	ML		MPS		LS		CVM		AD		RTAD	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
100	λ	-0.4867	0.2371	1.0365	1.1405	1.1074	1.7506	1.1127	1.6970	0.9364	1.1097	1.1488	1.5362
	μ	0.0094	0.0026	0.00679	0.0040	0.0290	0.0372	0.0338	0.0438	0.0457	0.03307	0.0103	0.0013
	θ	-0.0057	0.0010	-0.0014	0.0023	-0.0115	0.0027	-0.0056	0.0033	-0.1054	0.0023	0.0007	0.0020
300	λ	-0.4905	0.2308	1.0080	1.0367	1.0692	1.4696	1.0578	1.2996	0.8697	1.0362	1.1012	1.3297
	μ	0.0013	0.0053	0.0008	0.0013	0.0151	0.0163	0.0158	0.0140	0.0272	0.01934	0.0068	0.0062
	θ	-0.0039	0.0004	-0.0008.	0.0006	-0.0030	0.0007	-0.0011	0.0008	-0.3029	0.0007	0.0006	0.00063
500	λ	-0.4920	0.2302	1.0047	1.0208	1.0294	1.1528	1.0476	1.2234	0.7130	1.0267	1.0657	1.2173
	μ	0.0002	0.0003	0.00041	0.0007	0.0071	0.0050	0.0126	0.0089	0.01728	0.0176	0.0043	0.0004
	θ	-0.0009	0.0002	0.0011	0.0003	-0.0028	0.0005	-0.0020	0.0004	-0.4702	0.0006	0.0010	0.0006

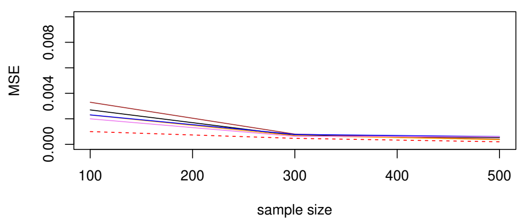
show respectively the MSE of the simulated estimates of λ, μ and θ.



(a) MSEs of $\hat{\lambda}$

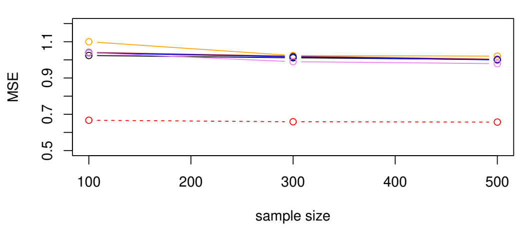


(b) MSEs of $\hat{\mu}$

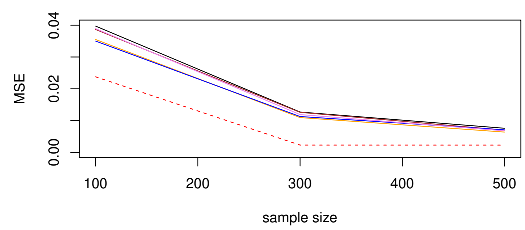


(c) MSEs of $\hat{\theta}$

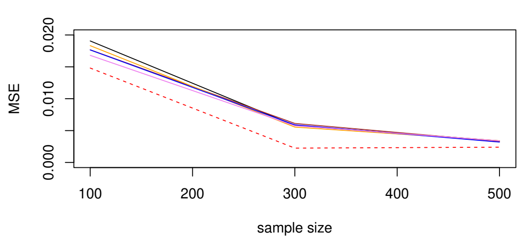
Figure 2.5: MSEs of the estimates of $\lambda = 0.5$, $\mu = 5$ and $\theta = 0.3$.



(a) MSEs of $\hat{\lambda}$



(b) MSEs of $\hat{\mu}$



(c) MSEs of $\hat{\theta}$

Figure 2.6: MSEs of the estimates of $\lambda = -0.5$, $\mu = 1$ and $\theta = 0.9$.

Table 2.3: Simulation results for $\lambda = -0.5$, $\mu = 1$, and $\theta = 0.9$.

n	Parameters	ML		MPS		LS		CVM		AD		RTAD	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
100	λ	-0.8036	0.6671	-1.0181	1.0999	-0.9787	1.0244	-0.9866	1.0414	-0.9906	1.0383	-0.9906	1.0393
	μ	0.0048	0.0237	-0.0092	0.0355	0.0146	0.0397	0.0038	0.0386	0.0094	0.0350	0.0079	0.0389
	θ	-0.0039	0.0148	0.0132	0.0183	-0.0256	0.0190	-0.0071	0.0176	-0.0103	0.0176	-0.0044	0.0168
300	λ	-0.8116	0.6587	-1.0019	1.0237	-0.9650	1.0113	-0.9995	1.0204	-0.9989	1.0165	-0.9845	0.9899
	μ	0.0049	0.0023	0.0026	0.0109	0.0034	0.0127	-0.0032	0.0126	0.0025	0.0112	0.0119	0.0119
	θ	-0.0020	0.0022	0.0047	0.0055	-0.0067	0.0058	0.0012	0.0061	-0.0007	0.0059	0.0029	0.0057
500	λ	-0.8118	0.6570	-1.0043	1.0203	-0.9941	1.0016	-1.005	1.0034	-0.9940	0.9996	-0.9739	0.9790
	μ	0.0491	0.00231	-0.0002	0.0064	0.0028	0.0076	-0.0029	0.0070	0.0051	0.0071	0.0052	0.0068
	θ	-0.0011	0.0024	0.0009	0.0034	-0.0031	0.0033	0.0035	0.0033	-0.0029	0.0032	0.0001	0.0033

2.3.7 Bayesian analysis

In this section, we discuss a parameter estimation of the proposed model through Bayesian techniques. The utility of the Bayesian approach in handling complexities is well established, see, Zellner (1971) and Khan et al. (2016).

The likelihood function of the TC distribution is given as

$$L = \prod_{i=1}^{n} \frac{1}{\pi\theta} \frac{1}{\left(1 + \left(\frac{x-\mu}{\theta}\right)^2\right)} \left[(1+\lambda) - 2\lambda \left(\frac{1}{\pi} \arctan\left(\frac{x-\mu}{\theta}\right) + 0.5\right)\right] \tag{2.16}$$

The following independent prior densities are assumed: $\lambda \to Uniform(l_1, l_2)$, $\mu \to Normal(m_1, m_2)$, $\theta \to Lognormal(n_1, n_2)$ where $l_1, l_2 \epsilon [-1, 1]$, $m_1 \epsilon R, m_2 > 0$ and n_1, n_2 are positive.

The joint posterior density of λ, μ, θ has the form

$$f(\lambda, \mu, \theta \mid x) = \frac{L(x \mid \lambda, \mu, \theta) p(\lambda) p(\mu) p(\theta)}{\int_\theta \int_\mu \int_\lambda L(x \mid \lambda, \mu, \theta) p(\lambda) p(\mu) p(\theta) d\lambda d\mu d\theta}. \tag{2.17}$$

Also the Bayes estimator of transmuting parameter λ under squared error loss function is

$$\hat{\lambda}_{Bayes} = E(\lambda \mid x) = \frac{\int_\theta \int_\mu \int_\lambda \lambda L(x \mid \lambda, \mu, \theta) p(\lambda) p(\mu) p(\theta) d\lambda d\mu d\theta}{\int_\theta \int_\mu \int_\lambda L(x \mid \lambda, \mu, \theta) p(\lambda) p(\mu) p(\theta) d\lambda d\mu d\theta}. \tag{2.18}$$

Moreover the Bayes risk λ is estimable through the relationship $Var(\lambda \mid x) = E(\lambda^2 \mid x) - (E(\lambda \mid x))^2$, where

$$E(\lambda^2 \mid x) = \frac{\int_\theta \int_\mu \int_\lambda \lambda^2 L(x \mid \lambda, \mu, \theta) p(\lambda) p(\mu) p(\theta) d\lambda d\mu d\theta}{\int_\theta \int_\mu \int_\lambda L(x \mid \lambda, \mu, \theta) p(\lambda) p(\mu) p(\theta) d\lambda d\mu d\theta}. \tag{2.19}$$

The Bayes estimates for other parameters can also be written in a similar way. We notice that the estimation is not analytically tractable. Therefore we use adaptive Metropolis Hasting Algorithm of MCMC to obtain the estimates. The estimation is demonstrated in the coming section through the use of two data sets.

2.4 Autoregressive TC minification process

In the literature, models with minification structures have been introduced as an alternative to the non-Gaussian time series models. The study on minification process began with the work of Tavares (1980). Sim (1986) developed a first order autoregressive Weibull process. Jose et al. (2010)developed different types of autoregressive processes with minification structure and max-min structure. Jose (2011) considered various Marshall-Olkin distributions and developed autoregressive minification processes with stationary marginals as exponential, Weibull, uniform, Pareto, Gumbel etc. We develop an AR(1) minification process with TC distribution as marginal distribution.

Consider an AR(1) minification process with structure

$$X_n = \begin{cases} \epsilon_n & \text{w.p} \quad \rho \\ \min(X_{n-1}, \epsilon_n) & \text{w.p} \quad 1-\rho \end{cases} \tag{2.20}$$

where $0 < \rho < 1,\ n \geq 1$ and $\{\epsilon_n\}$ is a sequence of i.i.d. r.v.s

To develop time series models with TC marginals, we need the following definition.

Definition 2.4.1. *A r.v. X on $(-\infty, \infty)$ is said to have Marshall-Olkin transmuted Cauchy (MOTC) distribution and write as $X \stackrel{d}{=}$ MOTC $(p, \lambda, \mu, \theta)$ if it has the sf*

$$\bar{F}(x) = \frac{p\left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right]}{1-(1-p)\left(1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right)}.$$

Theorem 2.4.1. *The AR(1) process given by* (2.20) *defines a stationary AR(1) minification process with TC(λ, μ, θ) as marginal distribution if and only if ϵ_n's are i.i.d. MOTC$(\rho^{-1}, \lambda, \mu, \theta)$ with $X_0 \stackrel{d}{=}$ TC(λ, μ, θ).*

Proof. We have, sf for TC(λ, μ, θ),

$$\bar{F}_X(x) = \left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right]$$

The model (2.20) can be rewritten in terms of sf as

$$P(X_n > x) = P(\epsilon_n > x)[\rho + (1-\rho)P(X_{n-1} > x)].$$

That is,

$$\bar{G}_{X_n}(x) = \bar{G}_{\epsilon_n}(x)[\rho + (1-\rho)\bar{G}_{X_{n-1}}(x)]. \tag{2.21}$$

If $\{X_n\}$ is stationary with TC(λ, μ, θ) marginals, then

$$\begin{aligned}\bar{F}_{\epsilon_n}(x) &= \frac{\bar{F}_X(x)}{\rho + (1-\rho)\bar{F}_X(x)} \\ &= \frac{\frac{1}{\rho}\left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right]}{1-(1-\frac{1}{\rho})\left(1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right)}.\end{aligned}$$

That is, ϵ_n's are i.i.d. MOTC$(\rho^{-1}, \lambda, \mu, \theta)$.

Conversely, if ϵ_n's are i.i.d. MOTC$(\rho^{-1}, \lambda, \mu, \theta)$ with $X_0 \stackrel{d}{=} TC(\lambda, \mu, \theta)$, then from (2.21), we have

$$\begin{aligned}\bar{F}_{X_1}(x) &= \rho\bar{F}_{\epsilon_1}(x) + (1-\rho)\bar{F}_{\epsilon_1}(x)\bar{F}_{X_0}(x) \\ &= \rho\frac{\frac{1}{\rho}\left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right]}{1-(1-\frac{1}{\rho})\left(1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right)}+ \\ &\quad (1-\rho)\frac{\frac{1}{\rho}\left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right]}{1-(1-\frac{1}{\rho})\left(1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right)} \\ &\quad \left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right] \\ &= \left[1-(1+\lambda)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]+\lambda\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right]^2\right].\end{aligned}$$

That is, $X_1 \stackrel{d}{=} TC(\lambda, \mu, \theta)$.

If we assume that $X_{n-1} \stackrel{d}{=} TC(\lambda, \mu, \theta)$, then by induction, we can establish that $X_n \stackrel{d}{=} TC(\lambda, \mu, \theta)$.

Hence the process $\{X_n\}$ is stationary with TC marginals.

This completes the proof. □

2.5 Applications

In this section, we show how the TC distribution works in practice. For this we consider two real data sets and compare the fit of the TC distribution with the following distributions:
(a) Two parameter Cauchy distribution having pdf

$$f(x;\mu,\theta)=\frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)};$$

where $-\infty<x<\infty, -\infty<\mu<\infty, \theta>0$
(b)Three parameter Skew Cauchy (SC) distribution introduced by Behboodian et al. (2006) with pdf

$$f(x;\mu,\theta,\lambda)=\frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}\left[1+\frac{\lambda(x-\mu)}{\sqrt{\theta^2+(1+\lambda^2)(x-\mu)^2}}\right];$$

where $-\infty<x<\infty, -\infty<\mu,\lambda<\infty, \theta>0$.
The values of the log-likelihood functions ($-\ln(L)$), AIC, AICC, and BIC are calculated for the three distributions in order to verify which distribution fits better to two sets of data. The better distribution corresponds to smaller $-\ln(L)$, AIC, AICC and BIC values. Under different parameter settings, we quantify Bayes estimates, lower credible limit (LCL), upper credible limit (UCL) of the credible intervals, whereas DIC is used to compare the performance of the different competing distributions. The better distribution corresponds to smaller DIC.

2.5.1 First data set

The first data set (`http://www.ibge.gov.br/seriesestatisticas/exibedados.php?idnivel=-BR&idserie=PRECO101`), is the INPC data which represents the national index of consumer prices of Brazil since 1979. The INPC index measures the cost of living of households with heads employees.

The data set is given in Table 2.4. The data is skewed to the right with skewness= 1.800

Table 2.4: INPC data set.

0.69	0.44	0.13	0.03	0.17	0.37	2.47	0.62	0.57	1.39
0.39	0.97	0.42	0.12	-0.11	0.50	0.39	2.70	0.31	0.84
0.30	0.55	0.43	0.49	0.27	0.70	0.73	0.82	3.39	1.07
0.48	-0.05	0.74	0.30	0.62	0.23	0.91	0.50	0.18	1.57
0.74	0.49	0.09	0.07	0.25	0.42	0.38	0.73	0.40	0.04
0.83	1.29	0.77	0.13	0.05	0.59	0.43	0.40	0.44	0.41
-0.06	0.86	0.94	0.55	0.05	0.47	0.32	0.16	0.54	0.57
0.57	0.99	1.15	0.44	0.29	0.61	1.28	0.31	-0.02	0.58
0.86	0.39	1.38	0.61	0.79	0.16	0.74	1.29	0.26	0.11
0.15	0.44	0.83	1.37	0.09	1.11	0.43	0.94	0.65	0.26
-0.07	0.00	0.17	0.54	1.46	0.68	0.60	1.21	0.96	0.42
-0.18	-0.28	0.49	0.15	0.18	0.68	0.34	1.20	0.29	1.51
2.46	0.11	0.15	0.54	0.29	0.35	0.45	0.38	1.33	0.71
1.40	2.18	-0.31	0.72	0.85	0.10	0.11	0.81	0.02	1.28
1.46	1.17	2.10	-0.49	0.45	0.57	-0.03	0.60	0.33	0.50
0.93	1.65	1.02	2.49	1.62	1.01	1.44			

and kurtosis= 4.183.

The descriptive statistics of the above data set are given in Table 2.5. Figure 2.7, shows

Table 2.5: Descriptive statistics of first data set.

Min	1st Q	Median	Mean	3rd Q	Max
0.00000	0.290	0.500	0.6646	0.8600	3.3900

the fitted density curves for the data set.

From Tables 2.6 and 2.7, we can see that TC distribution gives a better fit to the data set.

2.5.2 Second data set

The second real data set corresponds to data set from Nichols and Padgett (2006) on breaking stress of carbon fibres (in Gba): The second data set is given in Table 2.8. The

Table 2.6: Parameter estimates for various models fitted for the first data set.

Model	parameter estimates	log L	AIC	AICC	BIC
Cauchy	$\hat{\mu} = 0.4792$ $\hat{\theta} = 0.2656$	-139.3542	284.7083	284.8653	293.8771
SC	$\hat{\lambda} = 1.1888$ $\hat{\mu} = 0.2424$ $\hat{\theta} = 0.3275$	-132.7465	271.4929	271.6499	280.6617
TC	$\hat{\lambda} = 0.2527$ $\hat{\mu} = 0.4995$ $\hat{\theta} = 0.1400$	-132.0012	270.0025	270.1593	276.1711

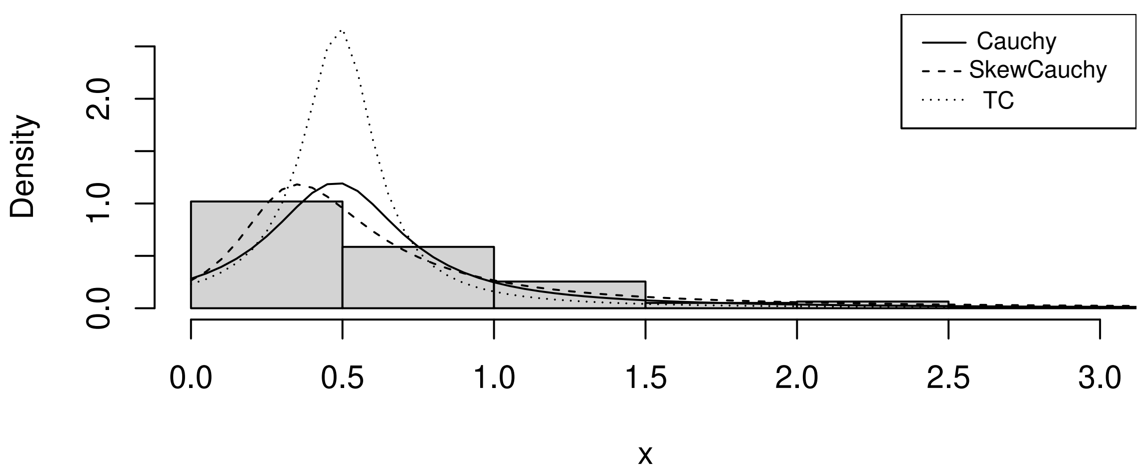


Figure 2.7: Fitted pdf plots of first data set

data is approximately symmetric with skewness=0.541 and kurtosis=0.141.

The descriptive statistics of the above data set are given in Table 2.9. From Table 2.10, it can be seen that the TC distribution provides best fit to the data set.

Figure 2.8, shows the fitted density curves for the second data set.

In Table 2.7 and 2.11, we present the results of all considered distributions under Bayesian inference. Here we use adaptive Metropolis Hasting Algorithm of MCMC to obtain the

Table 2.7: Posterior results of the TC and other models using first data set.

Model	parameter estimates	LCL	UCL	DIC
Cauchy	$\hat{\mu} = 0.4741$	0.4155	0.5360	658.939
	$\hat{\theta} = 0.2546$	0.2052	0.3127	
SC	$\hat{\mu} = 0.3781$	0.3132	0.4491	625.076
	$\hat{\theta} = 0.2590$	0.2099	0.3152	
	$\hat{\lambda} = 0.4154$	0.2175	0.4979	
TC	$\hat{\mu} = 2.8004$	2.7369	2.8669	623.258
	$\hat{\theta} = 0.0662$	0.0366	0.1053	
	$\hat{\lambda} = 0.1783$	0.0128	0.3996	

Table 2.8: Carbon fibres data set.

3.70	2.74	2.73	2.50	3.60	3.11	3.27	2.87	1.47	3.11
4.42	3.22	1.69	3.28	3.09	1.87	3.15	4.90	3.75	2.43
2.95	2.97	3.39	2.67	2.93	3.22	3.39	2.81	4.20	3.33
2.55	3.31	3.31	2.85	2.56	2.35	2.55	2.59	2.38	2.81
2.77	2.17	2.83	1.92	1.41	3.68	2.97	2.76	4.91	3.68
1.84	1.59	3.19	1.57	0.81	5.56	1.73	1.59	2.00	1.71
2.17	1.17	5.08	2.48	1.18	3.51	2.17	1.69	1.25	4.38
1.84	2.48	0.85	1.61	2.79	4.70	2.03	1.80	1.57	1.08
2.03	1.61	2.12	1.89	2.05	3.65				

Table 2.9: Descriptive statistics of second data set.

Min	1st Q	Median	Mean	3rd Q	Max
0.810	1.875	2.700	2.673	3.257	5.560

Table 2.10: Parameter estimates for various models fitted for the second data set.

Model	parameter estimates	log L	AIC	AICC	BIC
Cauchy	$\hat{\mu} = 2.3966$ $\hat{\theta} = 0.1000$	-239.3104	484.6208	484.9134	491.9838
SC	$\hat{\lambda} = 0.6136$ $\hat{\mu} = 2.5586$ $\hat{\theta} = 0.1100$	-136.5362	279.0724	279.365	286.4354
TC	$\hat{\lambda} = 0.6023$ $\hat{\mu} = 2.6911$ $\hat{\theta} = 0.1100$	-136.478	278.956	279.2486	286.319

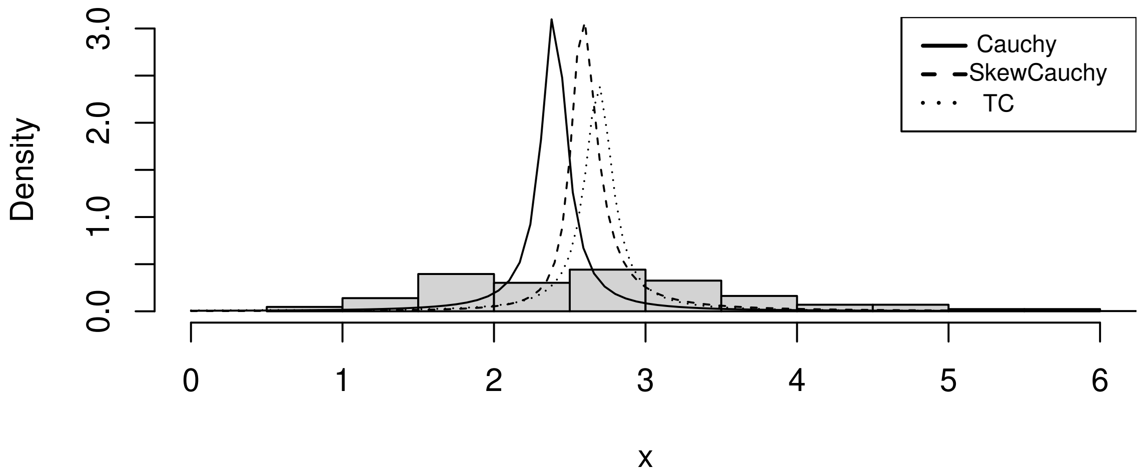


Figure 2.8: Fitted pdf plots of second data set

Table 2.11: Posterior results of the TC and other models using second data set.

Model	parameter estimates	LCL	UCL	DIC
Cauchy	$\hat{\mu} = 2.5967$	2.3698	2.8111	711.015
	$\hat{\theta} = 0.6209$	0.4703	0.7985	
SC	$\hat{\mu} = 2.2175$	1.8452	2.6398	709.135
	$\hat{\theta} = 0.6965$	0.5165	0.9133	
	$\hat{\lambda} = 0.5569$	0.0393	0.9819	
TC	$\hat{\mu} = 2.1981$	1.8443	2.8354	706.8782
	$\hat{\theta} = 0.6884$	0.5141	0.8889	
	$\hat{\lambda} = -0.6485$	−0.9951	0.3832	

estimates.

The following independent prior densities are considered: $\lambda \rightarrow Uniform(-1, 1)$, $\mu \rightarrow Normal(0, 1)$, $\theta \rightarrow Lognormal(0, 1)$.

Under different parameter settings, we quantify Bayes estimates, LCL, UCL of the credible intervals. The value of the DIC statistics for our proposed model is lower than that of the other models. This finding again emphasizes on the efficient nature of the suggested TC distribution.

2.6 Summary

Transmuted family of distributions have been of great attention among researchers. In the present work, we studied some basic statistical and mathematical properties of the model. We considered estimation of parameters of TC distribution using ML, MPS, LS, CVM, AD, RTAD estimation and Bayesian inference. We have performed a simulation study to compare these methods. The applicability of the model is established by using two real data sets and this indicates the flexibility and capacity of the distribution in data modelling.

Chapter 3

A GENERALIZATION OF CAUCHY DISTRIBUTION: ESTIMATION AND APPLICATIONS

3.1 Introduction

Cauchy distribution is used as the canonical example of pathological distribution in statistics. In mathematics, it is related to Poisson kernel which is the fundamental solution of Laplace equation. It is one of the distributions that are stable.

Cauchy distribution resembles the normal distribution family of curves, it has a taller peak than normal. That means it is a heavy tail probability distribution and unlike the normal distribution its fat tails decay more slowly.

This chapter introduces a new four parameter Cauchy distribution called the new

[1]This Chapter is based on Jayakumar and Fasna (2022)

Generalized Cauchy distribution (GCD). The GCD distribution has regularly varying tails, it belongs to the class of long-tailed distributions and is a member of the dominated variation distribution. Hence it belongs to the class of subexponential distributions. Also, it can be seen that $\lim_{x\to\infty} h(x) = 0$, where $h(x)$ is the hazard rate of GCD.

As a consequence of these, we have the following:
If $X_1, X_2, \ldots, X_n$ are i.i.d. r.v.s and if $S_n = X_1 + X_2 + \ldots + X_n$,
then

$$P(S_n > x) \sim nP(X > x), \quad as \quad x \to \infty.$$

That is, if

$$M_n = \max_{1\le i\le n} X_i,$$

then

$$P(S_n > x) \sim nP(X_i > x) \sim P(M_n > x).$$

Hence, for large x, the event $\{S_n > x\}$ is due to the event $\{M_n > x\}$. That is, exceedances of the high threshold by the sum are due to the exceedances of this threshold by the largest value in the sample.

One of the applications for subexponential distributions are insurance mathematics. Such distributions are used as a realistic models for expressing the sizes of real life insurance claims which can have distributions with heavy tails. This interpretation suggests one way of defining heavy-tailed distribution.

The importance of subexponential distributions as a useful class of heavy tailed distribution functions in the context of applied probability and insurance mathematics was discussed by Teugels (1975). A survey paper is by Goldie and Kluppelberg (1998). A textbook treatment of subexponential distributions is given in Embrechts et al. (1997).

The aim of this chapter is to introduce a new generalized Cauchy distribution, discuss the shape of the density function and distribution function of the model. We derive the quantiles, mode, Mean deviation and pdf of order statistics. Study the tail properties of distributions and it is shown that the distribution has a regularly varying tail, and belongs to the class of subexponential distributions. ML, MPS, LS, CVM, AD and RTAD estimation methods are explored and we evaluate the performance of these estimates using simulation. An autoregressive minification process with GCD marginals are developed. We analyze two real data sets to discuss the use of proposed distribution. Summary of the work done is presented.

3.2 New generalized Cauchy distribution

We introduce a new generalization of Cauchy distribution as a competitor for several generalizations of Cauchy distribution. Marshall and Olkin (1997) introduced a method of adding parameters to distributions leading to the development of several extensions of existing distributions in the literature. Building upon this method, Nadarajah et al. (2013) generalized the Marshall-Olkin scheme and introduced a family of distributions generated through the TNB distribution. It is note that both Marshall-Olkin scheme and its generalization via TNB originated from modeling scenarios involving random minimum or random maximum. For the applications of random minimum or random maximum in various fields, see Marshall and Olkin (1997).

Some generalizations of Uniform distributions have appeared in the literature. Jose and Krishna (2011) introduced Marshall-Olkin extended Uniform distribution with pdf

$$g(y;\alpha,\beta) = \frac{\alpha\beta}{(\alpha\beta+(1-\alpha)y)^2}; \quad 0 \le y \le 1, \alpha > 0, \beta > 0, \tag{3.1}$$

and expressed it as a mixture distribution with exponential distribution as mixing density.

Nadarajah et al. (2013) introduced a new family of life time models as follows:
Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.s with sf $\bar{F}(x)$. Let N be a TNB r.v. with parameters $\alpha \in (0,1)$ and $\theta > 0$. That is,
$Pr(N = n) = \frac{\alpha^\theta}{1-\alpha^\theta}\binom{\theta+n-1}{\theta-1}(1-\alpha)^n, n = 1, 2, \ldots$
Consider $U_N = min\{X_1, X_2, \ldots X_N\}$. Then,

$$Pr(U_N > x) = \bar{G}_U(x)$$

$$= \frac{\alpha^\theta}{1-\alpha^\theta}\sum_{n=1}^{\infty}\binom{\theta+n-1}{\theta-1}((1-\alpha)\bar{F}(x))^n.$$

That is,

$$\bar{G}_U(x) = \frac{\alpha^\theta}{1-\alpha^\theta}[(F(x)+\alpha\bar{F}(x))^{-\theta}-1]. \tag{3.2}$$

Similarly, if $\alpha > 1$ and N is a TNB r.v. with parameters $\frac{1}{\alpha}$ and $\theta > 0$, then $V_N = max\{X_1, X_2, \ldots, X_N\}$ also has the sf (3.2). Which gives a new family of distributions given by the sf

$$\bar{G}_U(x;\alpha,\theta) = \frac{\alpha^\theta}{1-\alpha^\theta}[(F(x)+\alpha\bar{F}(x))^{-\theta}-1]; \quad \alpha > 0, \quad \theta > 0 \quad and \quad x \in \mathbb{R}.$$

Note that

$$\bar{G}_U(x;\alpha,\theta) \longrightarrow \bar{F}(x) \text{ as } \alpha \longrightarrow 1.$$

This family is reduced to the Marshall-Olkin family of distributions when $\theta = 1$. so this is a generalization of the Marshall-Olkin family. Using the approach of Nadarajah et al. (2013), Jayakumar and Sankaran (2016) defined a generalized Uniform distribution with pdf

$$g(y;\alpha,\beta)=\frac{(1-\alpha)\beta\alpha^{\beta}}{(1-\alpha^{\beta})(y(1-\alpha)+\alpha)^{\beta+1}};\quad 0\le y\le 1,\alpha>0,\beta>0, \tag{3.3}$$

and studied its properties.

Note that, a uniform r.v. Y can be transformed to a Cauchy r.v X along the transformation $X=\mu+\theta\tan[(Y-0.5)\pi]$.

By using this transformation, where Y has the pdf (3.3), we obtain a new distribution which we call Generalized Cauchy distribution denoted by $GCD(\alpha,\beta,\mu,\theta)$ with parameters $-\infty\le\mu\le\infty,\alpha,\beta,\theta>0$.

3.2.1 Probability density function

The pdf of $GCD(\alpha,\beta,\mu,\theta)$ distribution thus obtained is

$$f(x)=\frac{\beta\alpha^{\beta}(1-\alpha)}{\pi(1-\alpha^{\beta})}\frac{[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-(\beta+1)}}{(1+(\frac{x-\mu}{\theta})^{2})}, \tag{3.4}$$

where $\quad -\infty\le x\le\infty,-\infty\le\mu\le\infty,\alpha,\beta,\theta>0$.

The graph of $f(x)$ for different values of parameters are given in Figure 3.1.

3.2.2 Distribution function

The cdf of X is given by

$$F(x)=\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}. \tag{3.5}$$

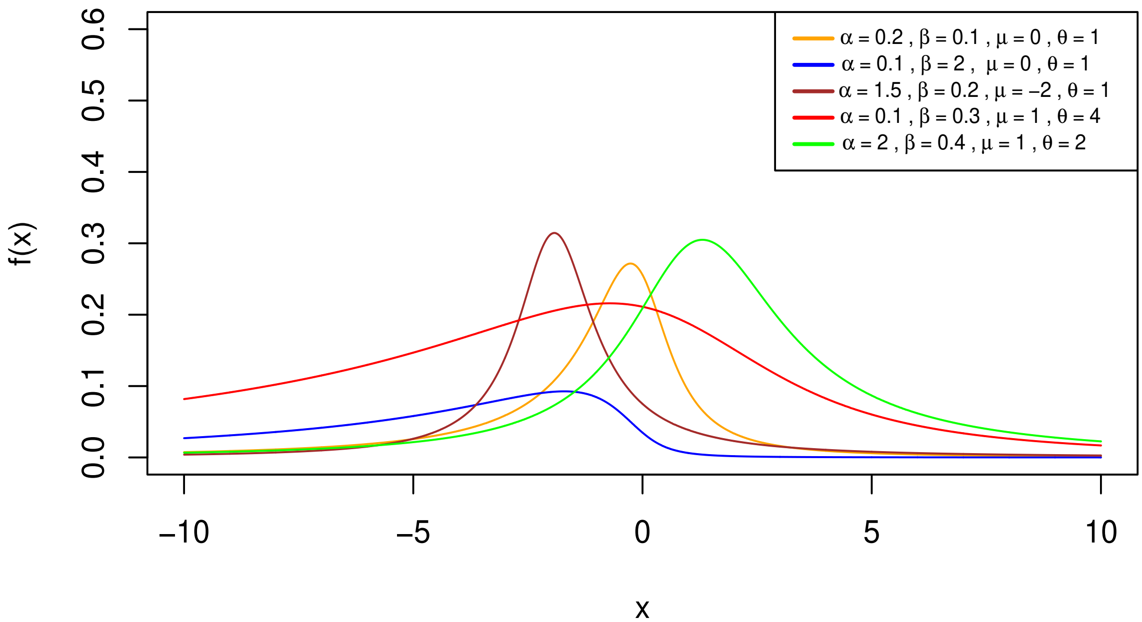


Figure 3.1: Plots of the pdf of $GCD(\alpha, \beta, \mu, \theta)$ distribution

For convenience, let $\theta = 1$ and $\mu = 0$.

Then

$$f(x) = \frac{\beta\alpha^{\beta}(1-\alpha)}{\pi(1-\alpha^{\beta})} \frac{[(0.5 + \frac{1}{\pi}\arctan(x))(1-\alpha) + \alpha]^{-(\beta+1)}}{(1+x^2)}, \tag{3.6}$$

where $-\infty \leq x \leq \infty, \alpha, \beta > 0$.

The graph of $F(x)$ for different values of parameters are given in Figure 3.2.

Remark 3.2.1. *When $\beta = 1$ and $\alpha \to 1$, GCD reduces to Cauchy distribution with parameters μ and θ.*

Theorem 3.2.1. *The limit of the GCD density function as $x \to \pm\infty$ is zero.*

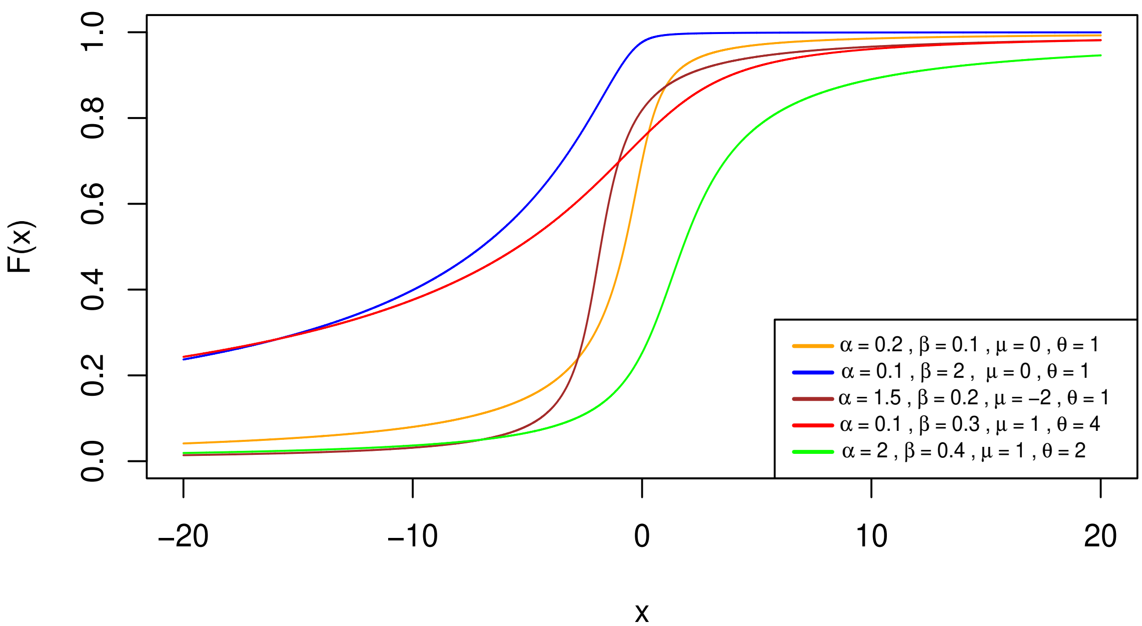


Figure 3.2: Plots of the cdf of $GCD(\alpha, \beta, \mu, \theta)$ distribution

Proof. Trivial and hence omitted. □

3.2.3 Quantile function

Theorem 3.2.2. *The q^{th} quantile x_q of the GCD r.v. is given by*

$$x_q = \mu + \theta \tan\left[\pi\left[\frac{\alpha}{1-\alpha}\left[(1 - q(1-\alpha^{\beta}))^{-\frac{1}{\beta}} - 1\right] - 0.5\right]\right]. \tag{3.7}$$

Proof. The q^{th} quantile x_q of the GCD r.v is defined as

$$q = P(X \leq x_q) = F(x_q), \quad x_q \in \mathbb{R}$$

Using the cdf of the GCD distribution, we have

$$q = F(x_q) = \frac{1 - \alpha^{\beta}[(0.5 + \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha) + \alpha]^{-\beta}}{1 - \alpha^{\beta}}$$

That is,

$$1 - \alpha^{\beta}\left[\left(0.5 + \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right)(1-\alpha) + \alpha\right]^{-\beta} = q(1-\alpha^{\beta}) \tag{3.8}$$

which implies

$$\left(0.5 + \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right) = \frac{\alpha}{1-\alpha}[(1 - q(1-\alpha^{\beta}))^{-\frac{1}{\beta}} - 1] \tag{3.9}$$

Hence

$$x_q = \mu + \theta\tan\left[\pi\left[\frac{\alpha}{1-\alpha}\left[(1 - q(1-\alpha^{\beta}))^{-\frac{1}{\beta}} - 1\right] - 0.5\right]\right]$$

This completes the proof. □

Using the inversion method, we can generate r.v.s from the GCD distribution. We can use equation (3.7) to generate random numbers when the parameters α, β, μ and θ are known.

Hence, the median of GCD is given by,

$$x_{0.5} = \mu + \theta\tan\left[\pi\left[\frac{\alpha}{1-\alpha}\left[(1 - 0.5(1-\alpha^{\beta}))^{-\frac{1}{\beta}} - 1\right] - 0.5\right]\right]. \tag{3.10}$$

Theorem 3.2.3. *The mode of the* $GCD(\alpha, \beta, \mu, \theta)$ *is the solution of the equation* $k(x) = 0$, *where*

$$k(x) = 2(\mu - x)\pi\left[\left(\left(0.5 + \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right)(1-\alpha)\right) + \alpha\right] - \theta(1-\alpha)(1+\beta).$$

Proof. The critical point of GCD density function are the roots of the equation:

$$\frac{\partial \log(f(x))}{\partial x} = 0$$

But

$$\frac{\partial \log(f(x))}{\partial x} = \frac{(1-\alpha)(1+\beta)}{\pi\theta[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha](1+(\frac{x-\mu}{\theta})^2)} + \frac{2(x-\mu)}{\theta^2(1+(\frac{x-\mu}{\theta})^2)} \quad (3.11)$$

The critical values of (3.11) are the solution of $k(x) = 0$. Hence the proof. □

3.2.4 Mean deviation

Let M be the median of the *GCD* distribution.

The mean deviation about the median can be calculated as

$$\delta(X) = E|X - M| = \int_{-\infty}^{\infty} |x - M| f(x) dx.$$

Hence, we obtain the following equation $\delta = \mu - 2J(M)$ where J(q) is

$$J(q) = \frac{(1-\alpha)\beta\alpha^{\beta}}{\pi(1-\alpha^{\beta})} \int_{-\infty}^{q} \frac{x[(0.5 + \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-(\beta+1)}}{(1+(\frac{x-\mu}{\theta})^2)} dx. \quad (3.12)$$

One can easily compute this integral numerically using softwares such as R,MATLAB, Mathcad and others and hence obtain the mean deviation about the median as desired.

3.2.5 Reliability function

The Reliability function of $GCD(\alpha, \beta, \mu, \theta)$ is given by,

$$R(t) = 1 - \left[\frac{1 - \alpha^{\beta} \left[(0.5 + \frac{1}{\pi} \arctan(\frac{t-\mu}{\theta}))(1-\alpha) + \alpha \right]^{-\beta}}{1 - \alpha^{\beta}} \right]. \tag{3.13}$$

The plot of reliability function of $GCD(\alpha, \beta, \mu, \theta)$ for various choices of the values of the parameters are presented in Figure 3.3.

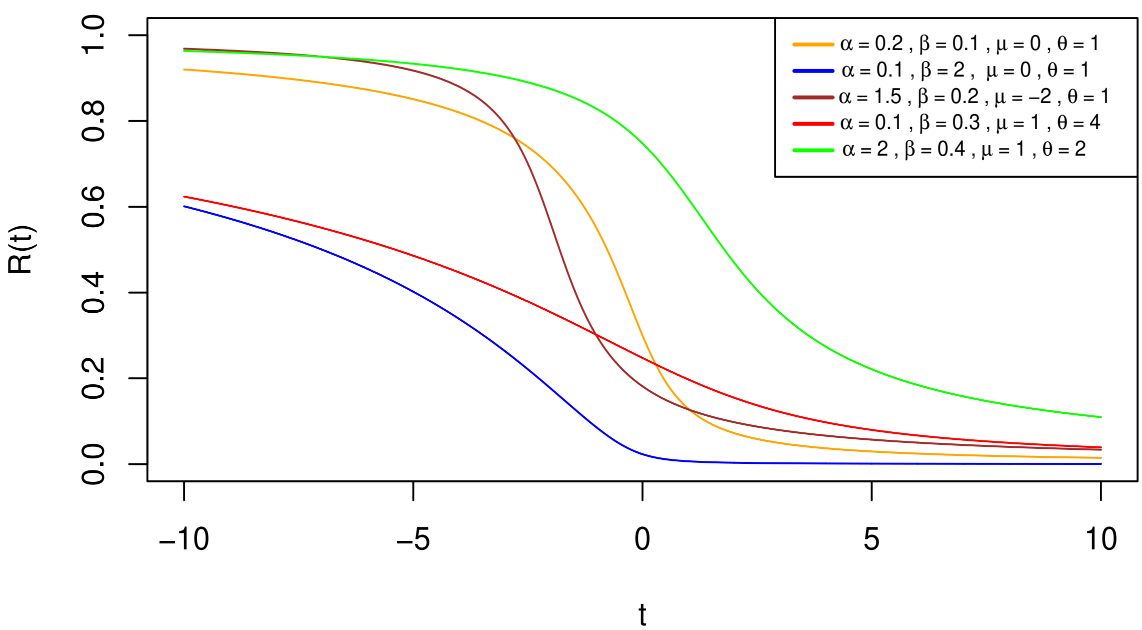


Figure 3.3: Plots of the reliability function of $GCD(\alpha, \beta, \mu, \theta)$ distribution

3.2.6 Hazard rate function

The hrf of $GCD(\alpha, \beta, \mu, \theta)$ is

$$h(t) = \frac{\frac{\beta\alpha^{\beta}(1-\alpha)}{\pi(1-\alpha^{\beta})} \frac{[(0.5+\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta}))(1-\alpha)+\alpha]^{-(\beta+1)}}{(1+(\frac{t-\mu}{\theta})^2)}}{1-\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}. \quad (3.14)$$

The plot of hrf of $GCD(\alpha, \beta, \mu, \theta)$ for various choices of the values of the parameters is given in Figure 3.4.

The cumulative hrf of GCD distribution $H(t)$ is given by,

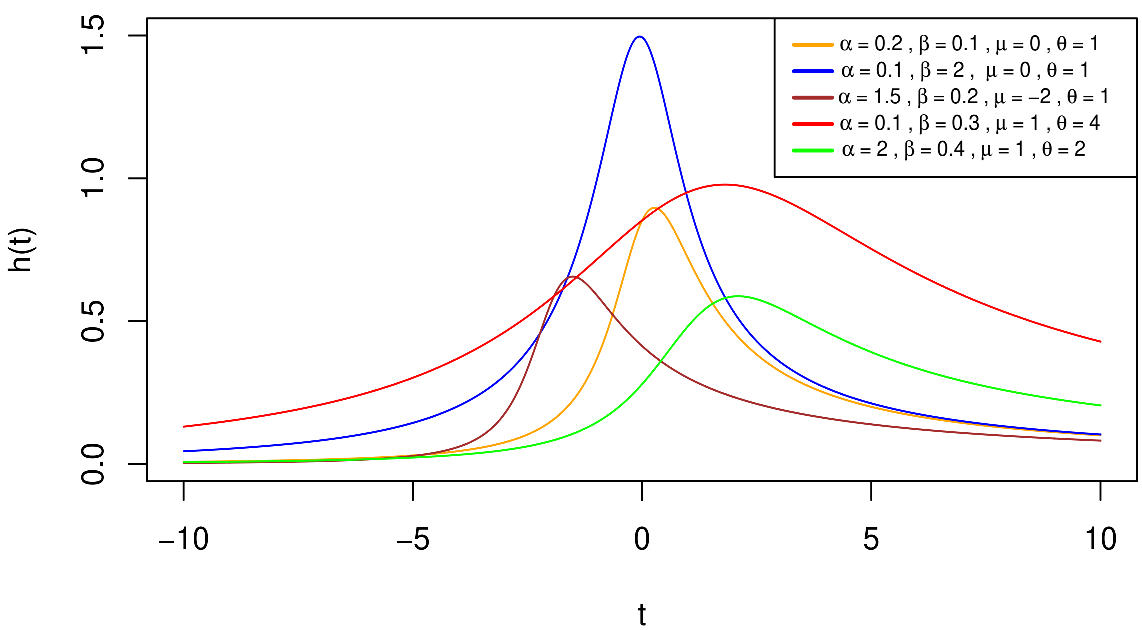


Figure 3.4: Plots of the hazard rate function of $GCD(\alpha, \beta, \mu, \theta)$ distribution

$$
\begin{aligned}
H(t) &= -\ln R(t) \\
&= -\ln\left[1-\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]\right].
\end{aligned} \tag{3.15}
$$

Theorem 3.2.4. *The limit of the hrf of GCD distribution as $t \to \pm\infty$ is zero.*

Proof. Trivial and hence omitted. □

3.2.7 Order statistics

Let $X_1, X_2, \ldots, X_n$ be a random sample from $GCD(\alpha, \beta, \mu, \theta)$ and let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the corresponding order statistics. Then the pdf and cdf of k^{th} order statistics are given by

$$
\begin{aligned}
f_X(x) &= \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}\,[1-F(x)]^{n-k}\,f(x) \\
&= \frac{n!}{(k-1)!(n-k)!}\frac{1}{1-\alpha^{\beta}}\frac{\beta\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta}))(1-\alpha)+\alpha]^{-(\beta+1)}(1-\alpha)}{\pi(1+(\frac{t-\mu}{\theta})^2)} \\
&\left[\frac{1-\alpha^{\beta}[(0.5+1/\pi\arctan(x))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]^{k-1} \\
&\left[1-\frac{1-\alpha^{\beta}[(0.5+1/\pi\arctan(\frac{t-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]^{n-k}
\end{aligned} \tag{3.16}
$$

and

$$
\begin{aligned}
F_X(x) &= \sum_{j=k}^{n}\binom{n}{j}[F(x)]^{j}[1-F(x)]^{n-j} \\
&= \sum_{j=k}^{n}\binom{n}{j}\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]^{j} \\
&\left[1-\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]^{n-j}
\end{aligned} \tag{3.17}
$$

respectively

The pdf of the minimum is,

$$f_{X_{(1)}}(x) = \frac{n\beta\alpha^\beta[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-(\beta+1)}(1-\alpha)}{\pi(1-\alpha^\beta)(1+(\frac{x-\mu}{\theta})^2)} \left[1-\frac{1-\alpha^\beta[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^\beta}\right]^{n-1} \tag{3.18}$$

and the pdf of the maximum is,

$$f_{X_{(n)}}(x) = \frac{n\beta\alpha^\beta[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-(\beta+1)}(1-\alpha)}{\pi(1-\alpha^\beta)(1+(\frac{x-\mu}{\theta})^2)} \left[\frac{1-\alpha^\beta[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^\beta}\right]^{n-1}. \tag{3.19}$$

3.2.8 Stochastic ordering

The GCD is ordered with respect to the strongest likelihood ratio ordering as shown in the following theorem.

Theorem 3.2.5. *Let $X \sim GCD(\alpha_1, \beta_1, 0, 1)$ and $Y \sim GCD(\alpha_2, \beta_2, 0, 1)$. If $\beta_1 = \beta_2 = \beta$ and $\alpha_1 < \alpha_2$; then $X \leq_{lr} Y$ hence $X \leq_{rhr} Y$, $X \leq_{hr} Y$ and $X \leq_{st} Y$.*

Proof. The likelihood ratio is

$$\frac{g_X(y)}{g_Y(y)} = \frac{{\alpha_1}^\beta(1-\alpha_1)(1-\alpha_2^\beta)[(0.5+\frac{1}{\pi}\arctan(x))(1-\alpha_1)+\alpha_1]^{-(\beta+1)}}{{\alpha_2}^\beta(1-\alpha_2)(1-\alpha_1^\beta)[(0.5+\frac{1}{\pi}\arctan(x))(1-\alpha_2)+\alpha_2]^{-(\beta+1)}}$$

Thus,

$$\frac{d\log}{dx}\left[\frac{g_X(y)}{g_Y(y)}\right]$$

$$= \frac{(1+\beta)}{\pi(1+x^2)} \left[\frac{(1-\alpha_2)}{[(0.5+\frac{1}{\pi}\arctan(x))(1-\alpha_2)+\alpha_2]} - \frac{(1-\alpha_1)}{[(0.5+\frac{1}{\pi}\arctan(x))(1-\alpha_1)+\alpha_1]} \right]$$

$$< 0.$$

Now, if $\beta_1 = \beta_2 = \beta$ and $\alpha_1 < \alpha_2$, then $\frac{d\log}{dx}\left[\frac{g_X(y)}{g_Y(y)}\right] < 0$, which implies that $X \leq_{lr} Y$ hence $X \leq_{rhr} Y$, $X \leq_{hr} Y$ and $X \leq_{st} Y$. □

3.2.9 Tail behaviour

The GCD has heavy tail, that is it takes extreme value with high probability. This feature empirically distinguishes GCD from the normal and many other distributions.

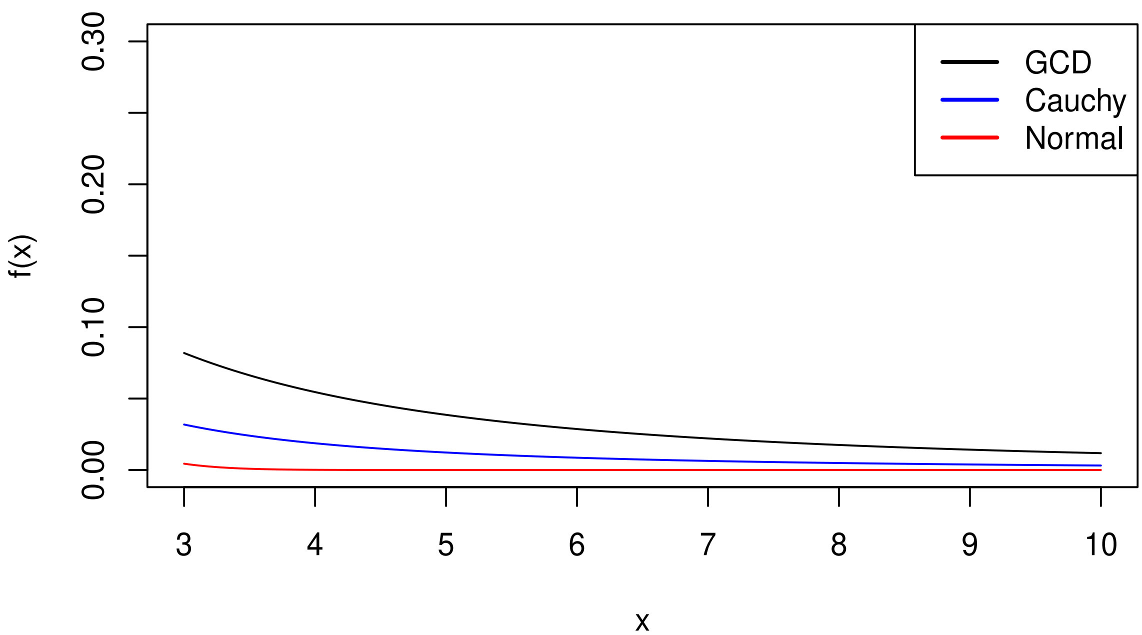


Figure 3.5: Comparison of right tails of Cauchy, Normal and GCD densities.

Figure 3.5 plots the right tails of the density of GCD and compare them with Cauchy and normal densities. It can be seen that GCD distribution has tails thicker than both Cauchy and normal.

We can show that $\limsup_{x\to\infty} f(x)e^{mx} = \infty$ for any $m > 0$ and hence the density f is heavy tailed.

The following theorem establishes that the GCD density function given in (3.6) is a function with regularly varying tails.

Theorem 3.2.6. *The density function of GCD distribution is a function with regularly varying tails.*

Proof. Using the density function (3.6), we have

$$\lim_{x\to\infty} \frac{f(cx)}{f(x)} = \lim_{x\to\infty} \frac{[(0.5 + \frac{1}{\pi}\arctan(cx))(1-\alpha)+\alpha]^{-(\beta+1)}(1+x^2)}{[(0.5 + \frac{1}{\pi}\arctan(x))(1-\alpha)+\alpha]^{-(\beta+1)}(1+(cx)^2)},$$

Applying limits, the above simplifies to

$$\lim_{x\to\infty} \frac{f(cx)}{f(x)} = \frac{1}{c^2},$$

and hence we have the desired result. □

Theorem 3.2.7. *The GCD distribution belongs to the class of long tailed distribution $\mathbb{L}$.*

Proof.

$$\lim_{x\to\infty} \frac{f(x+y)}{f(x)} = \frac{[(0.5 + \frac{1}{\pi}\arctan(x+y))(1-\alpha)+\alpha]^{-(\beta+1)}(1+x^2)}{[(0.5 + \frac{1}{\pi}\arctan(x))(1-\alpha)+\alpha]^{-(\beta+1)}(1+(x+y)^2)} = 1,$$

and hence f belongs to the class $\mathbb{L}$. □

Theorem 3.2.8. *The GCD distribution belongs to the class $\mathbb{D}$ dominated variation distributions .*

Proof.

$$\lim_{x\to\infty}\frac{f(x)}{f(2x)}=\lim_{x\to\infty}\frac{[(0.5+\frac{1}{\pi}\arctan(x))(1-\alpha)+\alpha]^{-(\beta+1)}(1+(2x)^2)}{[(0.5+\frac{1}{\pi}\arctan(2x))(1-\alpha)+\alpha]^{-(\beta+1)}(1+x^2)},$$

Applying limits, the above simplifies to

$$\lim_{x\to\infty}\frac{f(x)}{f(2x)}=4,$$

and hence f belongs to the class of dominated variation distributions. □

Theorem 3.2.9. *The GCD distribution belongs to the class of subexponential distributions.*

Proof. A distribution F is subexponential if

$$\lim_{x\to\infty}\frac{1-F*F(x)}{1-F(x)}=2;$$

where $*$ denotes the convolution operation.

From theorem 3.2.7 and theorem 3.2.8, the GCD distribution belongs to $\mathbb{D}\cap\mathbb{L}$. By Kluppelberg (1988), $\mathbb{D}\cap\mathbb{L}\subset S$, where S is a class of subexponential distributions. Hence the theorem. □

3.3 Estimation of parameters

In this section, we use the ML, MPS, LS, CVM, AD and RTAD estimation to estimate parameters of GCD distributions.

3.3.1 Method of maximum likelihood

If the parameters of the GCD distribution are unknown, then the ML estimates of the parameters can be obtained as follows:
For analytical simplicity, let assume that $\mu = 0$ and $\theta = 1$.

Consider a random sample $x_1, x_2, \ldots, x_n$ of size n from the $GCD(\alpha, \beta, \mu, \theta)$ distribution where $\mu = 0$ and $\theta = 1$. Then, the log likelihood function is given by,

$$\begin{aligned} \log L &= n \log \beta - n \log(1-\alpha^{\beta}) + n\beta \log \alpha + n \log(1-\alpha) - n \log \pi \\ &-(\beta+1) \sum_{i=1}^{n} \log[(0.5 + \frac{1}{\pi} \arctan(x_i))(1-\alpha) + \alpha] + \sum_{i=1}^{n} \log(1 + {x_i}^2). \end{aligned} \tag{3.20}$$

The likelihood equations are,

$$\begin{aligned} \frac{\partial \log L}{\partial \alpha} &= \frac{n\beta\alpha^{\beta-1}}{1-\alpha^{\beta}} + \frac{n\beta}{\alpha} - \frac{n}{1-\alpha} - (\beta+1) \sum_{i=1}^{n} \frac{(0.5 - \frac{1}{\pi} \arctan(x_i))}{[(0.5 + \frac{1}{\pi} \arctan(x_i))(1-\alpha) + \alpha]} \\ &= 0, \end{aligned} \tag{3.21}$$

and

$$\begin{aligned} \frac{\partial \log L}{\partial \beta} &= \frac{n\alpha^{\beta} \log \alpha}{1-\alpha^{\beta}} + \frac{n}{\beta} + n \log \alpha - \sum_{i=1}^{n} \log[(0.5 + \frac{1}{\pi} \arctan(x_i))(1-\alpha) + \alpha] \\ &= 0. \end{aligned} \tag{3.22}$$

These equations do not have explicit solutions and they have to be obtained numerically using statistical software like *optim* package in R
Now as in Popovic et al. (2015), we study the existence and uniqueness of the ML estimates when the other parameters are known (or given).

Theorem 3.3.1. *Let* $g_1(\alpha; \beta, x)$ *denote the function on the right-hand side (RHS) of equation* (3.21) *, where* β *is the true value of the parameter. Then there exists a unique* *solution for $g_1(\alpha;\beta,x)=0$, for $\hat{\alpha}\in(0,\infty)$.*

Proof. We have

$$g_1(\alpha;\beta,x)=\tfrac{n\beta\alpha^{\beta-1}}{1-\alpha^\beta}+\tfrac{n\beta}{\alpha}-\tfrac{n}{1-\alpha}-(\beta+1)\textstyle\sum_{i=1}^{n}\tfrac{(0.5-\frac{1}{\pi}\arctan(x_i))}{[(0.5+\frac{1}{\pi}\arctan(x_i))(1-\alpha)+\alpha]}.$$

Now

$$\lim_{\alpha\to 0} g_1(\alpha;\beta,x)=0+\infty-n-(\beta+1)\sum_{i=1}^{n}\frac{(0.5-\frac{1}{\pi}\arctan(x_i))}{(0.5+\frac{1}{\pi}\arctan(x_i))}=\infty,$$

On the other hand

$$\lim_{\alpha\to\infty} g_1(\alpha;\beta,x)=-\infty.$$

Therefore there exists atleast one root, say $\hat{\alpha}\in(0,\infty)$ such that $g_1(\alpha;\beta,x)=0$

To show uniqueness, the first derivative of $g_1(\alpha;\beta,x)$ is

$$\frac{\partial g_1(\alpha;\beta,x)}{\partial\alpha}<0,$$

Hence there exist a solution for $g_1(\alpha;\beta,x)=0$, and root $\hat{\alpha}$ is unique. □

Theorem 3.3.2. *Let $g_2(\beta;\alpha,x)=0$ denote the function on the RHS of equation* (3.22) *, where α is the true value of the parameter. Then there exists a unique solution for $g_2(\beta;\alpha,x)=0$, for $\hat{\beta}\in(0,\infty)$.*

Proof. We have

$$g_2(\beta;\alpha,x)=\frac{n\alpha^\beta\log\alpha}{1-\alpha^\beta}+\frac{n}{\beta}+n\log\alpha-\sum_{i=1}^{n}\log[(0.5+\frac{1}{\pi}\arctan(x_i))(1-\alpha)+\alpha].$$

Now

$$\lim_{\beta\to 0} g_2(\beta;\alpha,x) = \infty + n\log(\alpha) - \sum_{i=1}^{n}\log[(0.5+\frac{1}{\pi}\arctan(x_i))(1-\alpha)+\alpha] + 0 = \infty,$$

On the other hand

$$\lim_{\beta\to \infty} g_2(\beta;\alpha,x) = 0 + n\log(\alpha) - \sum_{i=1}^{n}\log[(0.5+\frac{1}{\pi}\arctan(x_i))(1-\alpha)+\alpha] + 0 < 0.$$

Therefore there exists at least one root, say $\hat{\beta} \in (0,\infty)$ such that $g_2(\beta;\alpha,x) = 0$

To show uniqueness, the first derivative of $g_2(\beta;\alpha,x)0$ is

$$\frac{\partial g_2(\beta;\alpha,x)}{\partial \beta} < 0,$$

Hence there exists a solution for $g_2(\beta;\alpha,x) = 0$, and root $\hat{\beta}$ is unique. □

3.3.2 Method of maximum product spacings

The cdf of GCD distribution is given by equation (3.5), and the uniform spacing are defined as follows:

$$D_1 = F(x_1) = \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_1-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}},$$

$$D_{n+1} = 1 - F(x_n) = 1 - \left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_n-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right],$$

and the general term of spacing is given by

$$D_i = F(x_i) - F(x_{i-1})$$

such that $\sum D_i = 1$.

MPS choose the estimates which maximizes the product of spacings or which maximizes the geometric mean of the spacing. That is, we find estimates such that

$$G = \left[\prod_{i=1}^{n+1} D_i\right]^{\frac{1}{n+1}},$$

is maximized. Taking the logarithm of G, we get

$$H = \frac{1}{n+1}\sum_{i=1}^{n+1}\log(D_i).$$

Differentiating the above equation partially, with respect to the parameters α, β, μ and θ respectively and them equating to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

3.3.3 Method of least square estimation

The LS estimators of the unknown parameters α, β, μ and θ of GCD distribution can be obtained by minimizing

$$\sum_{i=1}^{n}\left[F(x_i \mid \alpha,\beta,\mu,\theta) - \frac{i}{n+1}\right]^2$$

$$= \sum_{i=1}^{n}\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}} - \frac{i}{n+1}\right]^2.$$

Differentiating the above equation partially, with respect to unknown parameters α, β, μ and θ respectively and them equating to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

3.3.4 Method of Cramer-von Mises

The CVM estimators $\hat{\alpha}_{CME}$, $\hat{\beta}_{CME}$, $\hat{\mu}_{CME}$ and $\hat{\theta}_{CME}$ are the values of α,β,μ and θ minimizing

$$C(\alpha,\beta,\mu,\theta) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_i \mid \alpha,\beta,\mu,\theta,) - \frac{2i-1}{2n}\right]^2$$

$$= \frac{1}{12n} + \sum_{i=1}^{n}\left[\left(\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right) - \frac{2i-1}{2n}\right]^2.$$

Partially differentiating the above equations, with respect to the parameters α, β, μ and θ respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

3.3.5 Methods of Anderson-Darling and right-tail Anderson-Darling

The AD estimators $\hat{\alpha}_{ADE}$, $\hat{\beta}_{ADE}$, $\hat{\mu}_{ADE}$ and $\hat{\theta}_{ADE}$ are the values of α, β, μ and θ minimizes

$$\begin{aligned} A(\alpha,\beta,\mu,\theta) &= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log F(x_i \mid \alpha,\beta,\mu,\theta) + \log \bar{F}(x_{n+1-i} \mid \alpha,\beta,\mu,\theta)\} \\ &= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\Bigg\{\log\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right] \\ &+ \log\left[1-\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_{n+1-i}-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]\right]\Bigg\}. \end{aligned}$$

The RTAD estimators $\hat{\alpha}_{RTADE}$, $\hat{\beta}_{RTADE}$, $\hat{\mu}_{RTADE}$ and $\hat{\theta}_{RTADE}$ are the values of α, β, μ and θ minimizes

$$\begin{aligned} R(\alpha,\beta,\mu,\theta) &= \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i \mid \alpha,\beta,\mu,\theta) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log \bar{F}(x_{n+1-i} \mid \alpha,\beta,\mu,\theta) \\ &= \frac{n}{2} - 2\sum_{i=1}^{n}\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right] \\ &- \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\left[1-\left[\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x_{n+1-i}-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]\right]. \end{aligned}$$

Differentiating the above equation partially, with respect to the parameters α, β, μ, and θ respectively and them equating to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

3.3.6 Simulation study

We conduct Monte Carlo simulation to compare the performance of the estimators discussed in the previous sections and the process is repeated 1000 times. We evaluate the performance of the estimators based on bias and MSE. Methods are compared for sample sizes $n = 150$, $n = 250$ and $n = 500$.

We calculate the bias, MSE for each estimate. The statistics are obtained using the following formulae.

$Bias(\hat{\alpha}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha}-\alpha)$ $\quad$ $Bias(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\beta}-\beta)$

$Bias(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}-\mu)$ $\quad$ $Bias(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}-\theta)$

$MSE(\hat{\alpha}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha}-\alpha)^2$ $\quad$ $MSE(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\beta}-\beta)^2$

$MSE(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}-\mu)^2$ $\quad$ $MSE(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}-\theta)^2$

The bias (estimate-actual), and the MSE of the parameter estimates for the ML, MPS, LS, CVM, AD and RTAD methods are presented in Tables 3.1 and 3.2.

From Tables 3.1 and 3.2 , we note that the ML method performs well for estimating the model parameters. Also, as the sample size increases, the biases and the MSEs of the average estimates of ML estimates decrease as expected.

The following observations can be drawn from Tables 3.1 and 3.2.

1. All the estimators show the property of consistency, i.e., the MSE decreases as the sample size increases.

2. The bias of all parameters decreases with an increasing n for all the methods of estimations.

3. The bias of $\hat{\mu}$, $\hat{\theta}$ generally increases with an increasing mu, theta for any given mu,theta and n and for all methods of estimation. Figures 3.6 and 3.7 show respectively the MSE

Table 3.1: Simulation results for $\alpha = 0.1$,$\beta = 0.5$,$\mu = 2$ and $\theta = 0.3$.

n	Parameters	ML		MPS		LS		CVM		AD		RTAD	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
150	α	-0.0399	0.0049	-0.1424	0.0544	-0.6971	0.0151	0.0111	0.0060	-0.0298	0.0146	-0.0259	0.0164
	β	0.1287	0.0078	-1.0147	3.001	0.1501	0.0091	0.1798	0.03715	-0.0823	0.6673	-0.1372	0.5934
	μ	0.0063	0.0023	0.0345	0.0245	0.6526	0.4355	0.0689	0.02012	0.0383	0.0161	0.0038	0.0048
	θ	-0.0104	0.0037	-0.0452	0.0054	0.1474	0.0499	-0.0272	0.0410	-0.0154	0.0042	-0.0400	0.0042
250	α	-0.0175	0.0030	0.02606	0.0025	-0.6782	0.0147	0.0368	0.0029	-0.0743	0.0084	-0.0189	0.0095
	β	-0.0380	0.0014	0.1546	0.1837	0.0111	0.0073	-0.0152	0.0022	-0.6317	0.2129	-0.1429	0.5922
	μ	0.0061	0.0020	0.0088	0.0047	0.6294	0.4030	0.0264	0.0023	0.0085	0.0122	0.0092	0.0046
	θ	-0.0019	0.0017	0.0055	0.0039	0.1768	0.0449	0.1869	0.0398	-0.0182	0.0021	-0.0039	0.0023
500	α	-0.0136	0.0028	-0.0435	0.0023	-0.6518	0.0142	-0.0797	0.0079	0.0213	0.0042	-0.0201	0.0074
	β	-0.03781	0.0009	-0.3627	0.1607	-0.0018	0.0004	-0.0222	0.0007	-0.0177	0.0020	-0.1477	0.4296
	μ	0.0060	0.0011	-0.0247	0.0012	0.6193	0.3868	-0.1276	0.0176	-0.0478	0.0227	0.0034	0.0025
	θ	-0.0003	0.0007	0.0063	0.0002	0.1992	0.0443	-1.1553	0.0304	-0.4273	0.0015	-0.0046	0.0012

of the simulated estimates of α, β, μ and θ.

3.4 Autoregressive GCD minification process

We develop an AR(1) minification process with GCD distribution as marginal distribution. Consider an AR(1) minification process with structure

$$X_n = \begin{cases} \epsilon_n & \text{w.p} \quad \rho \\ \min(X_{n-1}, \epsilon_n) & \text{w.p} \quad 1-\rho \end{cases} \tag{3.23}$$

Table 3.2: Simulation results for $\alpha = 0.2$,$\beta = 0.1$,$\mu = 1$ and $\theta = 0.1$.

n	Parameters	ML		MPS		LS		CVM		AD		RTAD	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
150	α	0.1954	0.0382	-0.1660	-0.0089	-0.3112	0.0969	-0.0117	0.0908	-0.0919	0.0948	-0.1756	0.0438
	β	-0.3385	0.1236	0.0395	0.0003	-0.4978	0.2480	0.0291	0.2869	0.0461	0.2904	-0.7341	0.6505
	μ	-0.0068	0.0003	-0.3982	-0.05115	-0.0847	0.0004	0.0023	0.0003	0.0069	0.0003	-0.0055	0.0004
	θ	0.0088	0.0001	0.1585	0.0026	-0.3967	0.0015	0.0010	0.0003	0.0011	0.0001	-0.0029	0.0002
250	α	0.1945	0.0378	-0.1057	-0.0024	-0.3080	0.0951	-0.1915	0.0742	0.0202	0.0421	-0.1789	0.0407
	β	-0.3451	0.1204	0.0384	0.0002	-0.5071	0.2077	0.0716	0.2821	0.1156	0.2587	-0.7143	0.6096
	μ	0.0061	0.0002	-0.5161	-0.0018	-0.0859	0.0007	0.0023	0.0002	0.0063	0.0002	-0.0027	0.0002
	θ	0.0080	0.0001	0.1314	0.0001	-0.3897	0.0015	-0.0016	0.0001	0.0010	0.0001	-0.0055	0.0001
500	α	0.1933	0.0374	-0.0894	-0.0016	-0.3114	0.0900	0.0242	0.0074	0.0363	0.0384	-0.1814	0.0372
	c	-0.3605	0.1054	0.0381	0.0001	-0.4981	0.2004	0.0955	0.0924	0.1612	0.1709	-0.7223	0.5772
	μ	-0.0061	0.0002	-0.4279	-0.0018	-0.0844	0.0004	0.0014	0.0001	0.0059	0.0002	-0.0026	0.0001
	θ	0.0079	0.0001	0.0729	0.0001	-0.3970	0.0015	0.0012	0.0007	0.0013	0.0002	-0.0047	0.0008

where $0 < \rho < 1,\ n \geq 1$ and $\{\epsilon_n\}$ is a sequence of i.i.d. r.v.s.

In order to develop time series models with GCD marginals, we need the following definition.

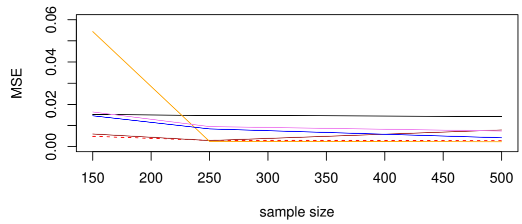
Definition 3.4.1. *A r.v. X on $(-\infty, \infty)$ is said to have Marshall-Olkin generalized Cauchy (MOGCD) distribution and write as $X \stackrel{d}{=}$ MOGCD $(p, \alpha, \beta, \mu, \theta)$ if it has the survival function*

$$\bar{F}(x) = \frac{p\left[1-\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}{1-(1-p)\left[1-\frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}.$$

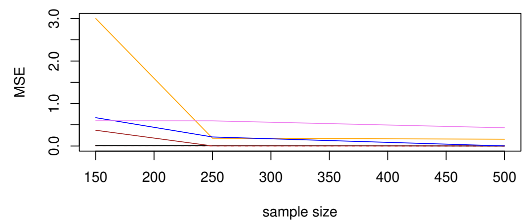
Theorem 3.4.1. *The AR(1) process given by* (3.23) *defines a stationary AR(1) minification process with GCD$(\alpha, \beta, \mu, \theta)$ as marginal distribution if and only if ϵ_n's are i.i.d. MOGCD $(\rho^{-1}, \alpha, \beta, \mu, \theta)$ with $X_0 \stackrel{d}{=}$ GCD$(\alpha, \beta, \mu, \theta)$.*

Proof. We have, sf for GCD$(\alpha, \beta, \mu, \theta)$,

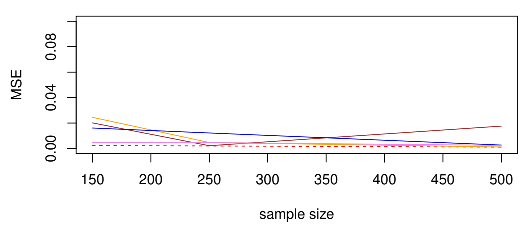
$$\bar{F}_X(x) = \left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]$$



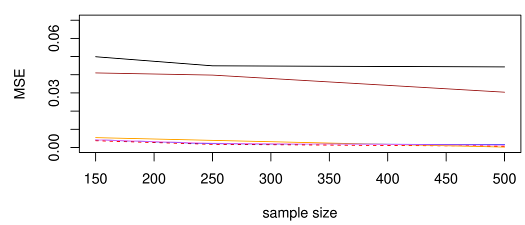
(a) MSEs of $\hat{\alpha}$



(b) MSEs of $\hat{\beta}$



(c) MSEs of $\hat{\mu}$



(d) MSEs of $\hat{\theta}$

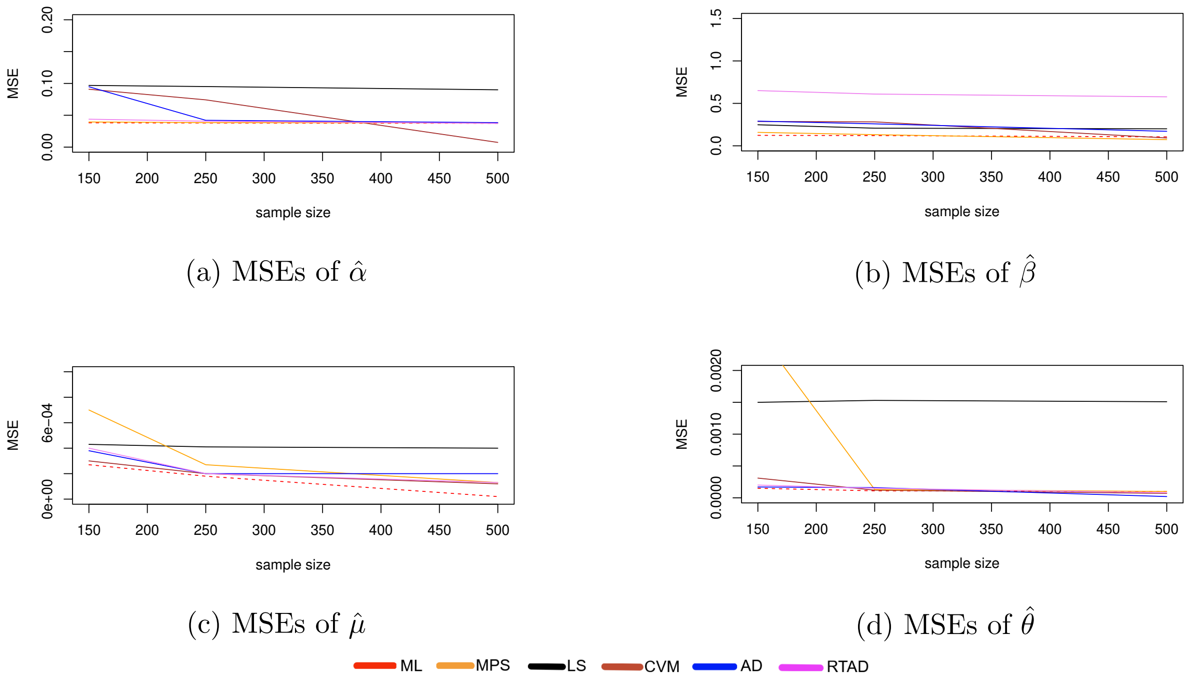
Figure 3.6: MSEs of the estimates of $\alpha = 0.1$, $\beta = 0.5$, $\mu = 2$ and $\theta = 0.3$.

The model (3.23) can be rewritten in terms of sf as

$$P(X_n > x) = P(\epsilon_n > x)[\rho + (1 - \rho)P(X_{n-1} > x)].$$

That is,

$$\bar{G}_{X_n}(x) = \bar{G}_{\epsilon_n}(x)[\rho + (1 - \rho)\bar{G}_{X_{n-1}}(x)]. \tag{3.24}$$



(a) MSEs of $\hat{\alpha}$

(b) MSEs of $\hat{\beta}$

(c) MSEs of $\hat{\mu}$

(d) MSEs of $\hat{\theta}$

Figure 3.7: MSEs of the estimates of $\alpha = 0.2$, $\beta = 0.1$, $\mu = 1$ and $\theta = 0.1$.

If $\{X_n\}$ is stationary with GCD$(\alpha, \beta, \mu, \theta)$ marginals, then

$$\bar{F}_{\epsilon_n}(x) = \frac{\bar{F}_X(x)}{\rho + (1-\rho)\bar{F}_X(x)}$$
$$= \frac{\frac{1}{\rho}\left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}{1-(1-\frac{1}{\rho})\left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}.$$

That is, ϵ_n's are i.i.d. MOGCD $(\rho^{-1}, \alpha, \beta, \mu, \theta)$.

Conversely, if ϵ_n's are i.i.d. MOGCD $(\rho^{-1}, \alpha, \beta, \mu, \theta)$ with $X_0 \stackrel{d}{=} GCD(\alpha, \beta, \mu, \theta)$ then from

(3.24), we have

$$
\begin{aligned}
\bar{F}_{X_1}(x) &= \rho \bar{F}_{\epsilon_1}(x) + (1-\rho)\bar{F}_{\epsilon_1}(x)\bar{F}_{X_0}(x) \\
&= \rho \frac{\frac{1}{\rho}\left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}{1-(1-\frac{1}{\rho})\left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]} + \\
&\quad (1-\rho)\frac{\frac{1}{\rho}\left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]}{1-(1-\frac{1}{\rho})\left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right]} \\
&\quad \left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right] \\
&= \left[1 - \frac{1-\alpha^{\beta}[(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))(1-\alpha)+\alpha]^{-\beta}}{1-\alpha^{\beta}}\right].
\end{aligned}
$$

That is, $X_1 \stackrel{d}{=} GCD(\alpha, \beta, \mu, \theta)$.

If we assume that $X_{n-1} \stackrel{d}{=} GCD(\alpha, \beta, \mu, \theta)$ then by induction, we can establish that $X_n \stackrel{d}{=} GCD(\alpha, \beta, \mu, \theta)$.

Hence the process $\{X_n\}$ is stationary with GCD marginals.

This completes the proof. □

3.5 Applications

In this section, we consider two real data sets to compare fit of the GCD distribution with the following distributions:

(a) Two parameter Cauchy distribution having pdf

$$
f(x; \mu, \theta) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)},
$$

where $-\infty < x < \infty, -\infty < \mu < \infty, \theta > 0$.

(b) Three parameter SC distribution introduced by Behboodian et al. (2006) with pdf

$$f(x;\mu,\theta,\lambda) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}\left[1+\frac{\lambda(x-\mu)}{\sqrt{\theta^2+(1+\lambda^2)(x-\mu)^2}}\right],$$

where $-\infty < x < \infty, -\infty < \mu, \lambda < \infty, \theta > 0$.

(c) TC distribution introduced by Ball et al. (2021) with pdf

$$f(x;\lambda,\mu,\theta) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}\left[(1+\lambda)-2\lambda\left(\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right)\right],$$

where $-\infty < x < \infty, -1 \leq \lambda \leq 1, -\infty < \mu < \infty, \theta > 0$.

(d) New generalized Pareto (NGP) distribution introduced by Jayakumar et al. (2020) with pdf

$$f(x;\alpha,\beta,\gamma,\theta) = \frac{\alpha\beta^\alpha\theta(1-\gamma)\gamma^\theta}{1-\gamma^\theta}\frac{x^{\alpha\theta-1}}{(\gamma x^\alpha+(1-\gamma)\beta^\alpha)^{\theta+1}},$$

where $x > \beta, \alpha, \beta, \gamma, \theta > 0$

The values of $-\ln(L)$, AIC, AICC and BIC are calculated for the five distributions in order to verify which distribution fits better to data. The better distribution corresponds to smaller $-\ln(L)$, AIC, AICC and BIC values.

3.5.1 First data set

The real data set corresponds to the data set from Nichols and Padgett (2006) on breaking stress of carbon fibers (in Gba): The data set are given in Table 3.3.

The data is approximately symmetric with skewness= 0.541 and kurtosis= 0.141

The descriptive statistics of the above data set are given in Table 3.4. The values in Table

Table 3.3: Carbon fibres data set.

3.70	2.74	2.73	2.50	3.60	3.11	3.27	2.87	1.47	3.11
4.42	3.22	1.69	3.28	3.09	1.87	3.15	4.90	3.75	2.43
2.95	2.97	3.39	2.67	2.93	3.22	3.39	2.81	4.20	3.33
2.55	3.31	3.31	2.85	2.56	2.35	2.55	2.59	2.38	2.81
2.77	2.17	2.83	1.92	1.41	3.68	2.97	2.76	4.91	3.68
1.84	1.59	3.19	1.57	0.81	5.56	1.73	1.59	2.00	1.71
2.17	1.17	5.08	2.48	1.18	3.51	2.17	1.69	1.25	4.38
1.84	2.48	0.85	1.61	2.79	4.70	2.03	1.80	1.57	1.08
2.03	1.61	2.12	1.89	2.05	3.65				

3.5 indicate that the GCD distribution provide better fit than other four models.

Figure 3.8, shows the fitted density curves for the first data set.

Table 3.4: Descriptive statistics of first data set.

Min	1st Q	Median	Mean	3rd Q	Max
0.810	1.875	2.700	2.673	3.257	5.560

Table 3.5: Parameter estimates and goodness of fit for various models fitted for the first data set.

Model	parameter estimates	log L	AIC	AICC	BIC	K-S	p-value
Cauchy	$\hat{\mu} = 2.3966$ $\hat{\theta} = 0.1000$	-239.3104	484.6208	484.9134	491.9838	0.5806	0.002
SC	$\hat{\mu} = 2.5586$ $\hat{\theta} = 0.1100$ $\hat{\lambda} = 0.6136$	-136.5362	279.0724	279.365	286.4354	0.2203	0.004
TC	$\hat{\lambda} = 0.6023$ $\hat{\mu} = 2.6911$ $\hat{\theta} = 0.1100$	-136.478	278.956	279.2486	286.319	0.3177	0.005
NGP	$\hat{\alpha} = 0.7017$ $\hat{\beta} = 1.4363$ $\hat{\gamma} = 0.9996$ $\hat{\theta} = 1.1291$	-141.719	291.42	291.913	301.237	0.5060	0.002
GCD	$\hat{\mu} = 2.6474$ $\hat{\theta} = 0.6047$ $\hat{\alpha} = 0.9954$ $\hat{\beta} = 10.2729$	-134.1803	276.3606	276.8544	286.1776	0.1035	0.3147

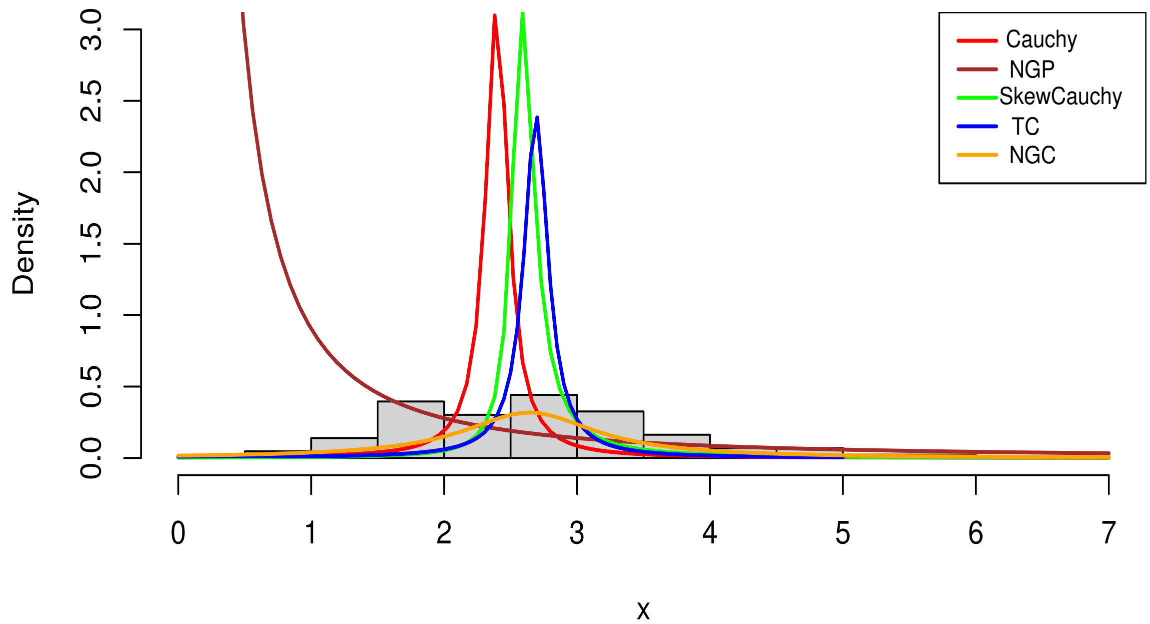


Figure 3.8: Fitted pdf plots of first data set

3.5.2 Second data set

The second data set (`http://www.ibge.gov.br/seriesestatisticas/exibedados.php?idnivel=-BR&idserie=PRECO101`), is the INPC data which represents the national index of consumer prices in Brazil since 1979. The INPC index measures the cost of living of households with head employees. The data set are given in Table 3.6.

The data is skewed-to-the right with skewness= 1.800 and kurtosis= 4.183.

The descriptive statistics of the above data set are given in Table 3.7. The values in Table 3.8 shows that the GCD distribution provide better fit than other four models.

Figure 3.9, shows the fitted density curves for the second data set.

Table 3.6: INPC data

0.69	0.44	0.13	0.03	0.17	0.37	2.47	0.62	0.57	1.39	0.39
0.97	0.42	0.12	-0.11	0.50	0.39	2.70	0.31	0.84	0.30	0.55
0.43	0.49	0.27	0.70	0.73	0.82	3.39	1.07	0.48	-0.05	0.74
0.30	0.62	0.23	0.91	0.50	0.18	1.57	0.74	0.49	0.09	0.07
0.25	0.42	0.38	0.73	0.40	0.04	0.83	1.29	0.77	0.13	0.05
0.59	0.43	0.40	0.44	0.41	-0.06	0.86	0.94	0.55	0.05	0.47
0.32	0.16	0.54	0.57	0.57	0.99	1.15	0.44	0.29	0.61	1.28
0.31	-0.02	0.58	0.86	0.39	1.38	0.61	0.79	0.16	0.74	1.29
0.26	0.11	0.15	0.44	0.83	1.37	0.09	1.11	0.43	0.94	0.65
0.26	-0.07	0.00	0.17	0.54	1.46	0.68	0.60	1.21	0.96	0.42
-0.18	-0.28	0.49	0.15	0.18	0.68	0.34	1.20	0.29	1.51	2.46
0.11	0.15	0.54	0.29	0.35	0.45	0.38	1.33	0.71	1.40	2.18
-0.31	0.72	0.85	0.10	0.11	0.81	0.02	1.28	1.46	1.17	2.10
-0.49	0.45	0.57	-0.03	0.60	0.33	0.50	0.93	1.65	1.02	2.49
1.62	1.01	1.44								

Table 3.7: Descriptive statistics of second data set.

Min	1st Q	Median	Mean	3rd Q	Max
0.00000	0.290	0.500	0.6646	0.8600	3.3900

Table 3.8: Parameter estimates and goodness of fit for various models fitted for the second data set.

Model	parameter estimates	log L	AIC	AICC	BIC	K-S	p-value
Cauchy	$\hat{\mu} = 0.4792$ $\hat{\theta} = 0.2656$	-139.3542	284.7083	284.8653	293.8771	0.5372	0.0303
SC	$\hat{\lambda} = 1.1888$ $\hat{\mu} = 0.2424$ $\hat{\theta} = 0.3275$	-132.7465	271.4929	271.6499	280.6617	0.4158	0.2013
TC	$\hat{\lambda} = 0.2527$ $\hat{\mu} = 0.4995$ $\hat{\theta} = 0.1400$	-132.0012	270.0025	270.1593	276.1711	0.2076	0.2628
NGP	$\hat{\alpha} = 0.0490$ $\hat{\beta} = 0.0020$ $\hat{\gamma} = 0.0867$ $\hat{\theta} = 0.0001$	-138.308	284.616	284.879	296.840	0.8579	0.3260
GCD	$\hat{\mu} = 0.4858$ $\hat{\theta} = 0.2193$ $\hat{\alpha} = 0.9993$ $\hat{\beta} = 3.7722$	-125.3175	258.635	258.8982	270.86	0.1249	0.6541

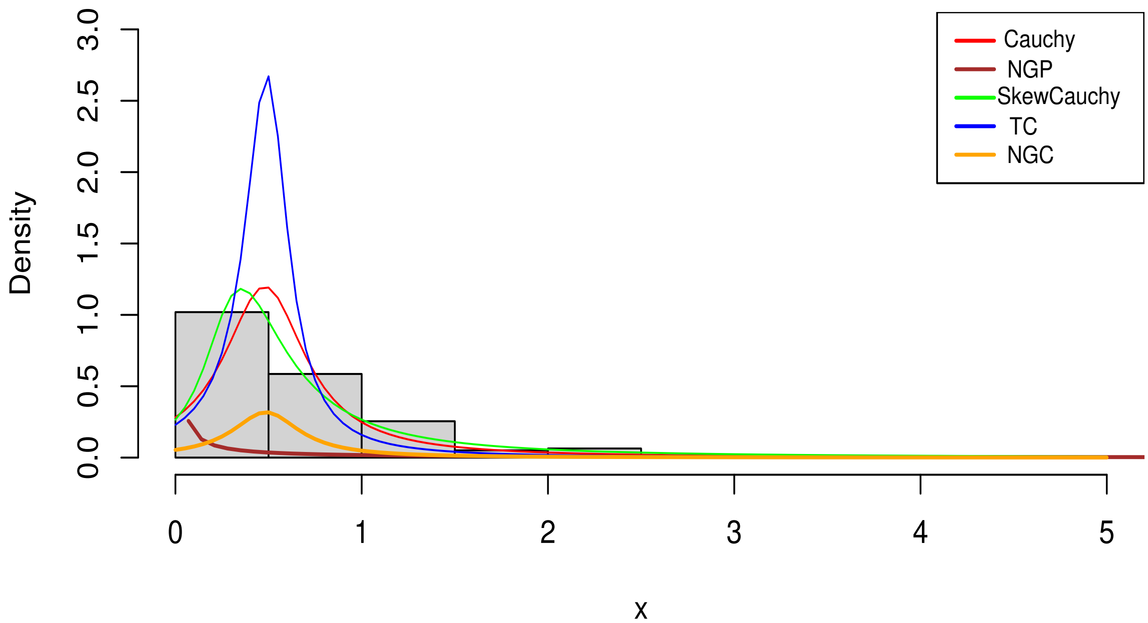


Figure 3.9: Fitted pdf plots of second data set.

3.6 Summary

In this chapter, we proposed and studied a new family of distribution called the GCD which extends Cauchy distribution. We studied various distributional characteristics of the model. The GCD distribution is heavy-tailed and belongs to the class of subexponential distributions. It has regularly varying tails and is a competitor of a number of existing generalizations of Cauchy. We considered estimation of parameters of GCD distribution using the method of ML, MPS, LS, CVM, AD and RTAD estimation. The heavy-tailed properties of the model make it appropriate for modeling a number of real-life situations. The applicability of the model is established by using two real data sets. From Tables 3.5 and 3.8, we observed a better performance of our model than the existing distributions.

Chapter 4

A NEW FAMILY OF CAUCHY DISTRIBUTION GENERATED THROUGH DISCRETE MITTAG LEFFLER DISTRIBUTION

4.1 Introduction

Cauchy distribution is one of the distributions that is stable. Stable distributions are a particular family of probability distributions suitable for modeling data that are heavy tailed and skewed. A detailed study of the generalized Cauchy family of distributions with applications has been studied by Alzaatreh et al. (2016).

Marshall and Olkin (1997) introduced a new family of distributions by adding a parameter to a given family of distributions. They started with a parent sf $\bar{F}(x) = 1 - F(x)$

[1]This Chapter is based on Fasna (2023)

and considered a family of sf given by

$$\bar{G}(x) = \frac{\alpha \bar{F}(x)}{F(x) + \alpha \bar{F}(x)}, \quad \alpha > 0. \tag{4.1}$$

The construction of their family of distributions are in the following way:

Let $X_1, X_2, \ldots.$ be a sequence of i.i.d. r.v.s with sf $\bar{F}(x)$. Let N be a geometric r.v. with p.m.f $Pr(N = n) = \alpha(1-\alpha)^{n-1}$, for $n = 1, 2, \ldots$ and $0 < \alpha < 1$. Then the random variable $U_N = min\{X_1, X_2, \ldots, X_N\}$ has the sf given by (4.1). If $\alpha > 1$ and N is a geometric r.v. with p.m.f $Pr(N = n) = \frac{1}{\alpha}(1 - \frac{1}{\alpha})^{n-1}$, $n = 1, 2, \ldots$ then the r.v. $V_N = max\{X_1, X_2, \ldots, X_N\}$ also has the sf as in (4.1).

Many authors have studied various univariate distributions belonging to the Marshall-Olkin family of distributions; see Ristic et al. (2007), Jose et al. (2010) and Cordeiro and Lemonte (2013). Jayakumar and Mathew (2008) proposed a generalization of the family of Marshall-Olkin distribution as

$$\bar{G}(x; \alpha, \gamma) = \left[\frac{\alpha \bar{F}(x)}{1 - (1-\alpha)\bar{F}(x)}\right]^{\gamma}, \ for\ \alpha > 0,\ \gamma > 0\ and\ x \in \mathbb{R}. \tag{4.2}$$

Nadarajah et al. (2013) introduced a new family of life time models as follows:

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.s with sf $\bar{F}(x)$ and N be a TNB r.v. with parameters $\alpha \in (0, 1)$ and $\theta > 0$. That is,

$Pr(N = n) = \frac{\alpha^{\theta}}{1-\alpha^{\theta}} \binom{\theta+n-1}{\theta-1} (1-\alpha)^n,\ n = 1, 2, \ldots$

Consider $U_N = min\{X_1, X_2, \ldots, X_N\}$. Then,

$$\begin{aligned} Pr(U_N > x) &= \bar{G}_U(x) \\ &= \frac{\alpha^{\theta}}{1-\alpha^{\theta}} \sum_{n=1}^{\infty} \binom{\theta+n-1}{\theta-1} ((1-\alpha)\bar{F}(x))^n. \end{aligned}$$

That is,

$$\bar{G}_U(x) = \frac{\alpha^\theta}{1-\alpha^\theta}[(F(x) + \alpha\bar{F}(x))^{-\theta} - 1]. \tag{4.3}$$

Similarly, if $\alpha > 1$ and N is a TNB r.v. with parameters $\frac{1}{\alpha}$ and $\theta > 0$, then $V_N = max\{X_1, X_2, \ldots, X_N\}$ also has the sf (4.3). This implies that we can consider a new family of distributions having the sf

$$\bar{G}_U(x; \alpha, \theta) = \frac{\alpha^\theta}{1-\alpha^\theta}[(F(x) + \alpha\bar{F}(x))^{-\theta} - 1],\ \alpha > 0,\ \theta > 0\ and\ x \in \mathbb{R}.$$

Note that $\bar{G}_U(x; \alpha, \theta) \longrightarrow \bar{F}(x)$ as $\alpha \longrightarrow 1$. This family is reduced to the Marshall-Olkin family of distributions, when $\theta = 1$. That means it is a generalization of the Marshall-Olkin family.

Pillai and Jayakumar (1995) introduced the DML distribution and studied its properties. The mathematical origin of the DML distribution can be described as follows: Consider a sequence of independent Bernoulli trails in which the k^{th} trail has probability of success $\frac{\alpha}{k}$ with $0 < \alpha < 1$ and $k = 1, 2, 3, \ldots$. Let N be the trail number in which the first success occurs. Then the probability that $\{N = r\}$ is given by

$$\begin{aligned} P(N = r) &= (1-\alpha)\left(1 - \frac{\alpha}{2}\right)\left(1 - \frac{\alpha}{3}\right)\ldots\left(1 - \frac{\alpha}{r-1}\right)\frac{\alpha}{r} \\ &= \frac{(-1)^r\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-r+1)}{r!}. \end{aligned} \tag{4.4}$$

Probability generating function (pgf) of N is given by $G(z) = 1-(1-z)^\alpha$. Let $X_1, X_2, \ldots, X_n$ be i.i.d. r.v.s as N and let $X_0 = 0$. Let M be geometric distributed r.v. with parameter p, ie. $Pr(M = k) = q^k p,\ k = 0, 1, 2, \ldots;\ 0 < p < 1,\ q = 1 - p$.

Then $X_1 + X_2 + \ldots + X_M$ has generating function

$$P(z) = \frac{p}{1 - q(1 - (1 - z)^{\alpha})} = \frac{1}{1 + c(1 - z)^{\alpha}} \tag{4.5}$$

with $p = \frac{1}{(1+c)}$. The distribution with pgf (4.5) is known as DML distribution with parameters α and c.

Sankaran and Jayakumar (2016) introduced the truncated DML distribution as follows:
Let a new r.v. Y such that

$$P(Y = x) = \frac{P(X = x)}{1 - p_0}; \quad x = 1, 2, ...$$

Then

$$\begin{aligned} H(s) = E(s^Y) &= \sum_{y=1}^{\infty} \frac{s^y p(X = y)}{1 - p_0} \\ &= \frac{1 + c}{c} \left[\frac{1}{1 + c(1 - s)^{\alpha}} \right] - \frac{1}{c} \end{aligned}$$

Therefore

$$H(\bar{F}(x)) = \frac{1 + c}{c} \left[\frac{1}{1 + c(1 - \bar{F}(x))^{\alpha}} \right] - \frac{1}{c}.$$

Hence the sf of this family of distributions with parameters α and c is given by

$$\bar{G}(x) = \frac{1 - F^{\alpha}(x)}{1 + cF^{\alpha}(x)}, \tag{4.6}$$

The corresponding cdf is given by

$$G(x) = \frac{(1 + c)F^{\alpha}(x)}{1 + cF^{\alpha}(x)}. \tag{4.7}$$

This truncated DML distribution reduces to Marshall-Olkin family when $\alpha = 1$. So it can be considered as a generalization of Marshall-Olkin family of distributions. In equation (4.7), when F(x) is exponential, G(x) becomes the Marshall-Olkin generalized exponential distribution studied in Ristic and Kundu (2015). When F(x) in equation (4.7) is Weibull, G(x) reduces to Marshall-Olkin exponentiated Weibull distribution discussed in Bidram et al. (2015). Hence equation (4.7) is a rich class which leads to various generalizations of existing distributions that have the capability of modeling real data sets.

In this chapter we introduce a new family of univariate distribution which is a generalization of Marshall-Olkin family of distribution. We discuss the shape of the density function and distribution function of the model under study. Also derive its quantiles, mode, mean deviation and distribution of order statistics. We discuss the tail behaviour of this new distribution and study the estimation of parameters. The parameters of the probability distributions are estimated by the method of ML, MPS, LS, CVM, AD and RTAD estimation. We analyze two real data sets to explore the usefullness of our proposed distribution and finally conclusions are presented.

4.2 Discrete Mittag- Leffler Cauchy distribution

In this section, we introduce a new four parameter Cauchy distribution called truncated Discrete Mittag- Leffler Cauchy distribution.

A r.v. X is said to have Cauchy distribution with parameters μ and θ if its pdf is given by

$$f(x) = \frac{1}{\pi\theta}\frac{1}{(1 + (\frac{x-\mu}{\theta})^2)}; \quad -\infty \leq x \leq \infty, -\infty \leq \mu \leq \infty, \theta > 0 \tag{4.8}$$

and the cdf of X is given by

$$F(x) = \frac{1}{\pi} \arctan\left(\frac{x-\mu}{\theta}\right) + 0.5. \tag{4.9}$$

4.2.1 Distribution function

Using (4.7), we get the distribution function G(x), for F(x) in (4.9) as

$$G(x) = \frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}, \quad x \in \mathbb{R}, -\infty \leq \mu \leq \infty, \alpha, c, \theta > 0. \tag{4.10}$$

We refer to this distribution as truncated Discrete Mittag- Leffler Cauchy (DMLC) distribution with parameters α, c, μ and θ; and it is denoted by $DMLC(\alpha, c, \mu, \theta)$.
The cdf plots of $DMLC(\alpha, c, \mu, \theta)$ for various choices of the values of the parameters are given in Figure 4.1.

4.2.2 Probability density function

The pdf of $DMLC(\alpha, c, \mu, \theta)$ is obtained is

$$g(x) = \frac{\alpha(1+c)[0.5 + \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1 + (\frac{x-\mu}{\theta})^2)[1 + c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}]^2}. \tag{4.11}$$

The pdf plots of $DMLC(\alpha, c, \mu, \theta)$ for various values of the parameters are given in Figure 4.2.

Remark 4.2.1. *When $\alpha = 1$ and $c \longrightarrow 0$, DMLC reduces to Cauchy distribution with parameters μ and θ.*

$DMLC(\alpha, c, \mu, \theta)$ distribution for the fact that it's expected value and other moments

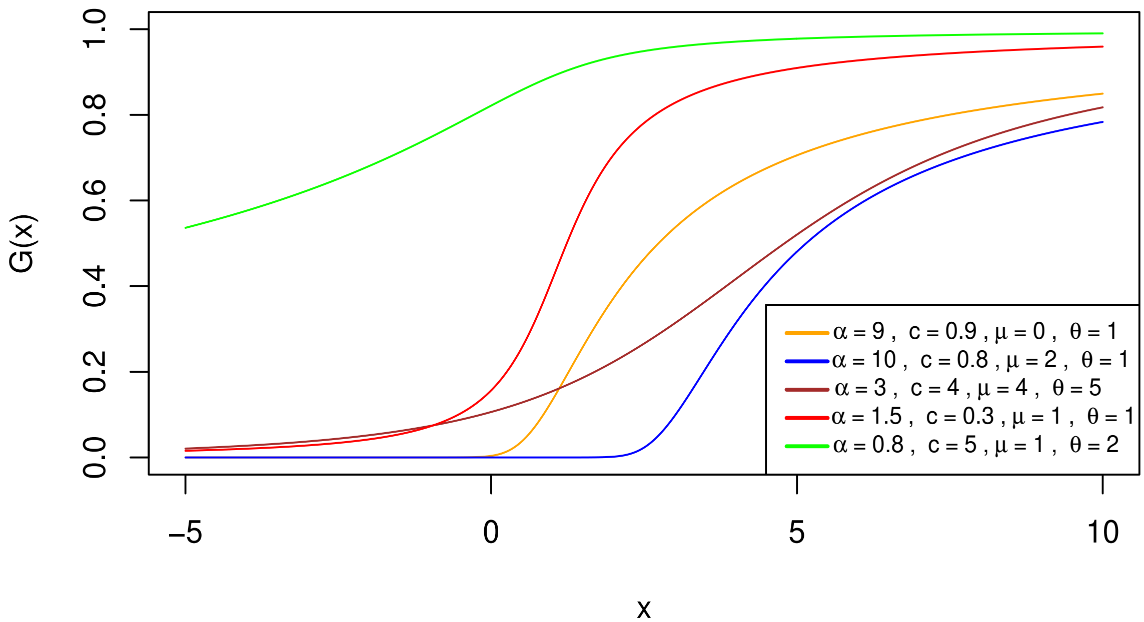


Figure 4.1: Plots of the cdf of $DMLC(\alpha, c, \mu, \theta)$ distribution

do not exist. The median, mode do exist.

Theorem 4.2.1. *The limit of the DMLC density function as* $x \to \pm\infty$ *is zero.*

Proof. Trivial and hence omitted. □

4.2.3 Quantile function

Theorem 4.2.2. *The* q^{th} *quantile* x_q *of the DMLC r.v. is given by*

$$x_q = \mu + \theta \tan\left[\pi\left[\left(\frac{q}{1+c-qc}\right)^{\frac{1}{\alpha}} - 0.5\right]\right]. \tag{4.12}$$

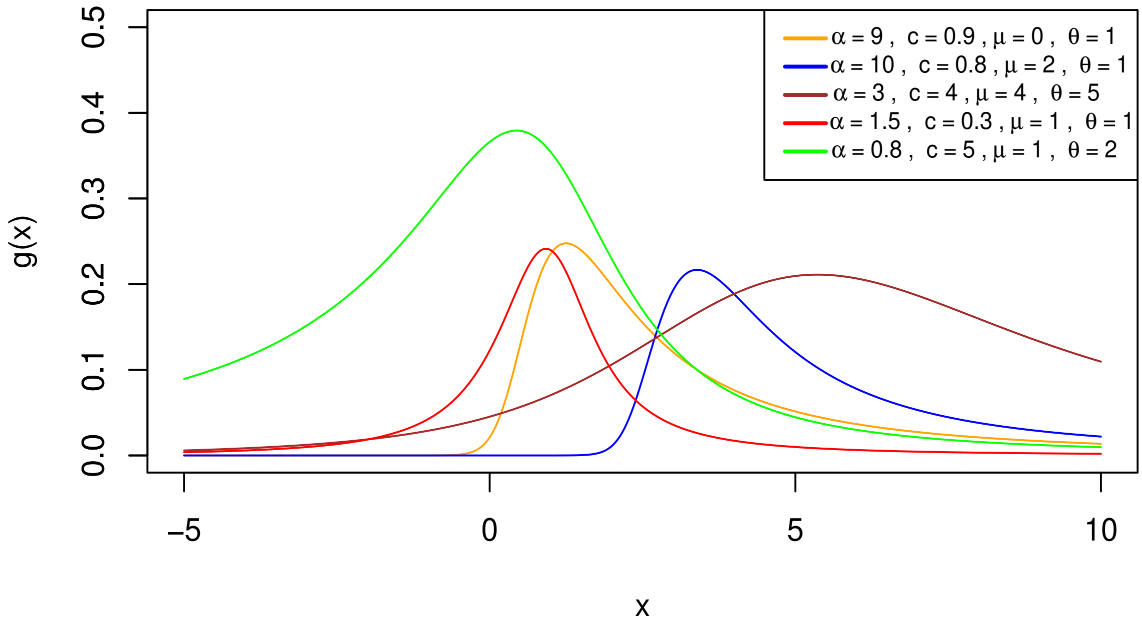


Figure 4.2: Plots of the pdf of $DMLC(\alpha, c, \mu, \theta)$ distribution

Proof. The q^{th} quantile x_q of the DMLC r.v. is defined as

$$q = P(X \leq x_q) = G(x_q), \quad x_q \in \mathbb{R}$$

Using the cdf of the DMLC distribution, we have

$$q = G(x_q) = \frac{(1+c)\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right) + 0.5\right]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right) + 0.5]^{\alpha}}$$

That is,

$$(1+c)\left[0.5 + \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right]^{\alpha} = q\left[1 + c\left[\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right) + 0.5\right]^{\alpha}\right] \quad (4.13)$$

Which implies

$$\left[0.5 + \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right] = \left[\frac{q}{1+c-qc}\right]^{\frac{1}{\alpha}} \tag{4.14}$$

We get

$$x_q = \mu + \theta \tan\left[\pi\left[\left(\frac{q}{1+c-qc}\right)^{\frac{1}{\alpha}} - 0.5\right]\right].$$

This completes the proof. □

Using the inversion method, we can generate r.v.s from the DMLC distribution by using the equation (4.12) when the parameters α, c, μ and θ are known.

The median of DMLC is given by,

$$x_{0.5} = \mu + \theta \tan\left[\pi\left[\left(\frac{0.5}{1+0.5c}\right)^{\frac{1}{\alpha}} - 0.5\right]\right]. \tag{4.15}$$

4.2.4 Mode

Theorem 4.2.3. *The mode of the* $DMLC(\alpha, c, \mu, \theta)$ *is the solution of the equation* $k(x) = 0$*, where*

$$\begin{aligned} k(x) = (\alpha-1)\pi\left[0.5+\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right]\left[1+c\left(0.5+\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)^{\alpha}\right)\right] \\ -2c\alpha\left[0.5+\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right]^{\alpha-1}. \end{aligned}$$

Proof. The critical point of DMLC density function are the roots of the equation

$$\frac{\partial \log(f(x))}{\partial x} = 0$$

That is

$$\frac{\partial \log(g(x))}{\partial x} = \frac{(\alpha-1)}{\pi\theta^2[0.5+\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)](1+(\frac{x-\mu}{\theta})^2)} - \frac{1}{\theta^2(1+(\frac{x-\mu}{\theta})^2)} - \frac{2c\alpha[0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{\pi\theta^2(1+(\frac{x-\mu}{\theta})^2)[1+c[0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha}]} \tag{4.16}$$

The critical values of equation (4.16) are the solution of $k(x) = 0$.

Hence the proof. □

4.2.5 Mean deviation

Let M be the median of DMLC distribution given by equation (4.15).

The mean deviation about the median can be calculated as

$$\delta(X) = E|X - M| = \int_{-\infty}^{\infty} |x - M| g(x)dx,$$

Hence we obtain the following equation $\delta = \mu - 2J(M)$ where J(q) is

$$J(q) = \frac{\alpha(1+c)}{\pi\theta}\int_{-\infty}^{q} \frac{x[0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{(1+(\frac{x-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5]^{\alpha}]^2}dx. \tag{4.17}$$

One can easily compute this integral numerically in software such as R, MATLAB, Mathcad, and others and hence obtain the mean deviation about the median.

4.2.6 Reliability function

The reliability function of $DMLC(\alpha, c, \mu, \theta)$ is given by,

$$R(t) = 1 - \left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta})+0.5]^{\alpha}}{1+c[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta})+0.5]^{\alpha}}\right]. \tag{4.18}$$

The reliability plot of $DMLC(\alpha, c, \mu, \theta)$ for various choices of the values of the parameters are presented in Figure 4.3.

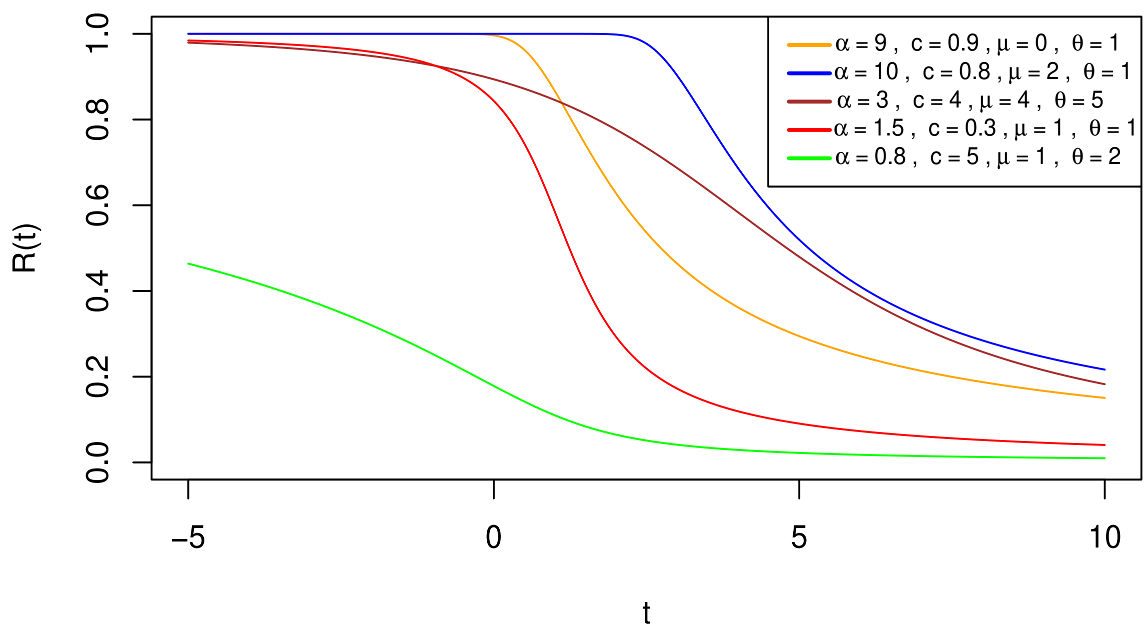


Figure 4.3: Plots of the reliability function of $DMLC(\alpha, c, \mu, \theta)$ distribution

4.2.7 Hazard rate function

The hrf of $DMLC(\alpha, c, \mu, \theta)$ is given by,

$$h(t) = \frac{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{t-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta})+0.5]^{\alpha}]^2}}{1-\left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta})+0.5]^{\alpha}}{1+c[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta})+0.5]^{\alpha}}\right]}. \tag{4.19}$$

The plot of hrf of $DMLC(\alpha, c, \mu, \theta)$ for various choices of the values of the parameters are presented in Figure 4.4. The cumulative hrf of a DMLC distribution, $H(t)$ is given by,

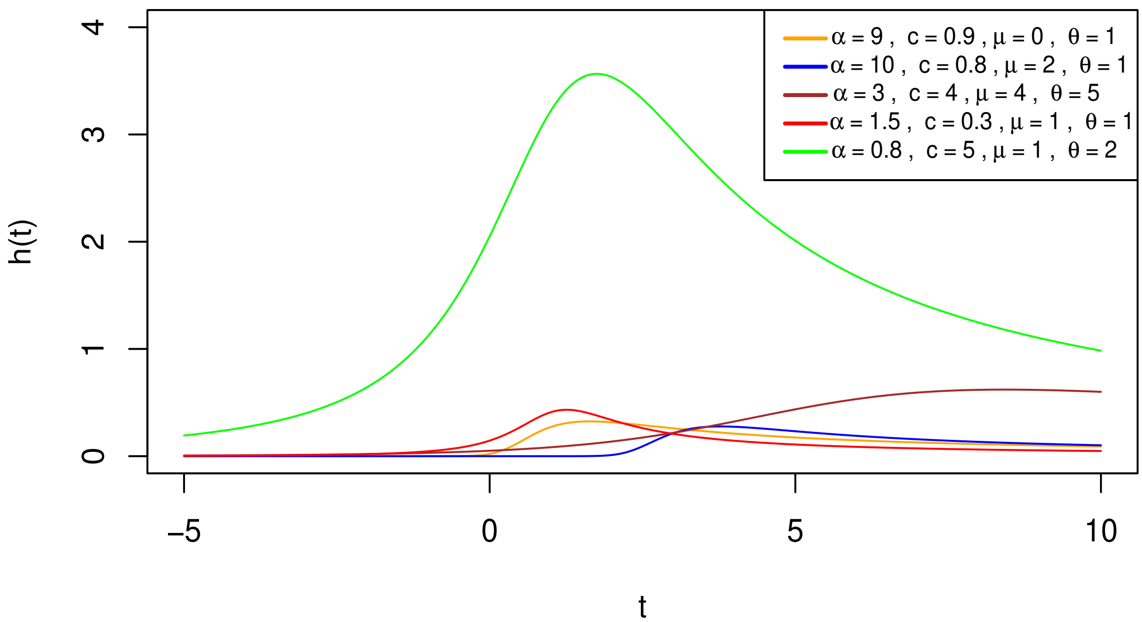


Figure 4.4: Plots of the hazard rate function of $DMLC(\alpha, c, \mu, \theta)$ distribution

$$H(t) = -\ln R(t) = -\ln\left[1 - \left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{t-\mu}{\theta}) + 0.5]^{\alpha}}\right]\right]. \tag{4.20}$$

Theorem 4.2.4. *The limit of the DMLC hrf as $t \to \pm\infty$ is zero.*

Proof. Trivial and hence omitted. □

4.2.8 Order statistics

Let $X_1, X_2, \ldots, X_n$ be a random sample from $DMLC(\alpha, c, \mu, \theta)$ and let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the corresponding order statistics. Then the pdf of k^{th} order statistic is given by

$$\begin{aligned} g_X(x) &= \frac{n!}{(k-1)!(n-k)!} g(x)[G(x)]^{k-1}\left[1 - G(x)\right]^{n-k} \\ &= \frac{n!}{(k-1)!(n-k)!} \frac{\alpha(1+c)[0.5 + \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1 + (\frac{x-\mu}{\theta})^2)[1 + c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}]^2} \\ &\left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}\right]^{k-1} \left[1 - \frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}\right]^{n-k}. \end{aligned} \tag{4.21}$$

4.2.9 Model identifiability

To prove model identifiability only with respect to the parameters α and c, since the other two parameters (μ and θ) are from the parent distribution. Suppose that $G(x; \alpha_1, c_1) = G(x; \alpha_2, c_2)$ for all $x \in \mathbb{R}$. We have to show this condition implies that $\alpha_1 = \alpha_2$ and $c_1 = c_2$.

We know, F is the parent distribution and the mixing distribution is truncated discrete Mittag-Leffler distribution.

Sankaran and Jayakumar (2016) introduced truncated discrete Mittag-Leffler distributions with cdf given by

$$G(x) = \frac{(1+c)F^{\alpha}(x)}{1 + cF^{\alpha}(x)}. \tag{4.22}$$

Now, if the parent distribution F transformed to Cauchy distribution by adding parameters μ and θ, then

$$G(x) = \frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5]^{\alpha}}, \quad x \in \mathbb{R}, -\infty \leq \mu \leq \infty, \alpha, c, \theta > 0, \tag{4.23}$$

This is the cdf of our proposed distribution.

For proving model identifiability, we use Theorem 2.4 of Chandra (1977)

PROPOSITION 4.2.1. *The class of all mixing distributions relative to the DMLC distribution is identifiable.*

Proof. If N_i is truncated discrete Mittag-Leffler r.v., then the pgf is

$$\phi_i(s) = \frac{1}{1 + c(1-s)^{\alpha}}, \; i = 1, 2$$

From the cdf of N_i, we have

$G_1 < G_2$ when $c_1 = c_2$ and $\alpha_1 < \alpha_2$

and

$G_1 < G_2$ when $\alpha_1 = \alpha_2$ and $c_1 < c_2$.

Let the domain of definition $D_{\phi_1(s)} = (-\infty, c_1)$, $D_{\phi_2(s)} = (-\infty, c_2)$ and $s = 1 + (\frac{1}{c_1})^{\frac{1}{\alpha_1}}$. Hence,

$$\lim_{s \to 1 + (\frac{1}{c_1})^{\frac{1}{\alpha_1}}} \phi_1(s) = \frac{1}{1 + c_1(1 - (1 + (\frac{1}{c_1})^{\frac{1}{\alpha_1}}))^{\alpha_1}} = \infty$$

When $\alpha_1 = \alpha_2$, and $c_2 < c_1$, we obtain

$$\lim_{s \to 1 + (\frac{1}{c_1})^{\frac{1}{\alpha_1}}} \phi_2(s) = \frac{1}{1 + c_2(1 - (1 + (\frac{1}{c_1})^{\frac{1}{\alpha_1}}))^{\alpha_1}} > 0$$

So

$$\lim_{s \to 1 + (\frac{1}{c_1})^{\frac{1}{\alpha_1}}} \frac{\phi_2(s)}{\phi_1(s)} = 0,$$

and thus the model identifiability is proved. Hence the cdf G is identifiable with respect to α and c.

This completes the proof. □

4.2.10 Tail behaviour

In this section, we study the tail behaviour of DMLC distribution. Figure 4.5 plots the

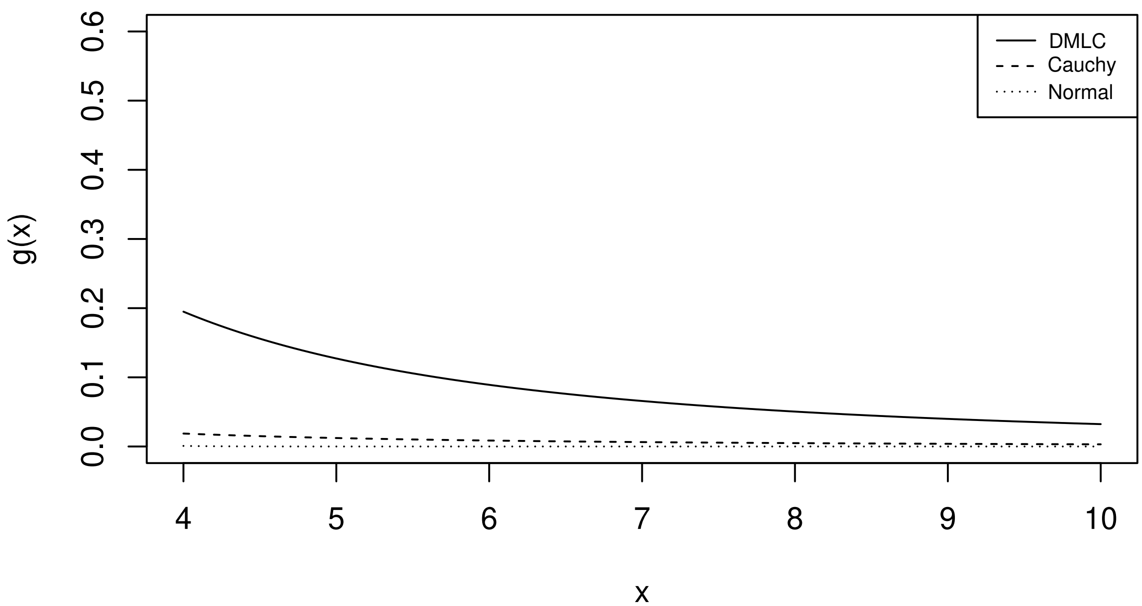


Figure 4.5: Comparison of right tails of Cauchy, Normal and DMLC densities.

right tails of density of DMLC and compare them with Cauchy and normal densities. DMLC distribution has tails thicker than both Cauchy and normal.

We can easily shows that $\limsup_{x\to\infty} g(x)e^{mx} = \infty$ for any $m > 0$, Hence the density f is heavy tailed.

The following theorem establishes that the density function given in equation (4.11) is a function with regularly varying tails.

Theorem 4.2.5. *The density function of DMLC distribution is a function with regularly varying tails.*

Proof. Using the density function (4.11), we have

$$\lim_{x\to\infty}\frac{g(ax)}{g(x)}=\lim_{x\to\infty}\frac{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{ax-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{ax-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{ax-\mu}{\theta})+0.5]^{\alpha}]^2}}{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{x-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5]^{\alpha}]^2}}.$$

Applying limits, the above simplifies to

$$\lim_{x\to\infty}\frac{g(ax)}{g(x)}=\frac{1}{a^2},$$

Hence we arrive at the desired result. □

Theorem 4.2.6. *The DMLC distribution belongs to the class long tailed distribution* $\mathbb{L}$.

Proof.

$$\lim_{x\to\infty}\frac{g(x+y)}{g(x)}=\frac{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{(x+y)-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{(x+y)-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{(x+y)-\mu}{\theta})+0.5]^{\alpha}]^2}}{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{x-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5]^{\alpha}]^2}}=1,$$

then f belongs to the class $\mathbb{L}$, the class of long tailed distributions. □

Theorem 4.2.7. *The DMLC distribution belongs to the class* $\mathbb{D}$ *dominated variation distributions.*

Proof.

$$\lim_{x\to\infty}\frac{g(x)}{g(2x)}=\lim_{x\to\infty}\frac{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{x-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5]^{\alpha}]^2}}{\frac{\alpha(1+c)[0.5+\frac{1}{\pi}\arctan(\frac{2x-\mu}{\theta})]^{\alpha-1}}{\pi\theta(1+(\frac{2x-\mu}{\theta})^2)[1+c[\frac{1}{\pi}\arctan(\frac{2x-\mu}{\theta})+0.5]^{\alpha}]^2}},$$

Applying limits, the above simplifies to

$$\lim_{x\to\infty}\frac{g(x)}{g(2x)}=2^2,$$

then f belongs to the class of dominated variation distributions. □

Two distributions G and F are said to be tail-equivalent if

$$\lim_{x\to\infty} \frac{\bar{G}(x)}{\bar{F}(x)} = a \in (0, \infty).$$

Hence it can be shown that DMLC and Cauchy distribution are tail-equivalent.

4.3 Estimation of parameters

In this section, we use ML, MPS, LS, CVM, AD and RTAD procedure for estimation.

4.3.1 Method of maximum likelihood

If the parameters of the DMLC distribution are unknown, then the ML estimates of the parameters are given as follows. For analytical simpilicity, let assume that $\mu = 0$ and $\theta = 1$.

Consider a random sample $x_1, x_2, \ldots, x_n$ of size n from the $DMLC(\alpha, c, \mu, \theta)$ distribution where $\mu = 0$ and $\theta = 1$. Then, the log likelihood function is given by

$$\begin{aligned} \log L = {} & n \log \alpha + n \log(1+c) - n \log(\pi\theta) + (\alpha - 1) \sum_{i=1}^{n} \log[0.5 + \frac{1}{\pi} \arctan(x_i)] \\ & - \sum_{i=1}^{n} \log(1 + x_i^2) - 2 \sum_{i=1}^{n} \log[1 + c[0.5 + \frac{1}{\pi} \arctan(x_i)]^{\alpha}]. \end{aligned} \tag{4.24}$$

The likelihood equations are,

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log[0.5 + \frac{1}{\pi} \arctan(x_i)] \left[1 - \frac{2c(0.5 + \frac{1}{\pi} \arctan(x_i))^{\alpha}}{1 + c(0.5 + \frac{1}{\pi} \arctan(x_i))^{\alpha}}\right]$$

$$= 0, \tag{4.25}$$

and

$$\frac{\partial \log L}{\partial c} = \frac{n}{1+c} - \sum_{i=1}^{n} \frac{2\left[0.5 + \frac{1}{\pi}\arctan(x_i)\right]^{\alpha}}{1 + c\left[0.5 + \frac{1}{\pi}\arctan(x_i)\right]^{\alpha}} = 0. \tag{4.26}$$

These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

4.3.2 Method of maximum product spacings

The cdf of DMLC distribution is given by equation (4.10), and the uniform spacing are defined as follows:

$$D_1 = F(x_1) = \frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_1-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{x_1-\mu}{\theta}) + 0.5]^{\alpha}},$$

$$D_{n+1} = 1 - F(x_n) = 1 - \left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_n-\mu}{\theta}) + 0.5]^{\alpha}}{1 + c[\frac{1}{\pi}\arctan(\frac{x_n-\mu}{\theta}) + 0.5]^{\alpha}}\right],$$

and the general term of spacing is given by

$$D_i = F(x_i) - F(x_{i-1})$$

such that $\sum D_i = 1$.

Method of product spacing choose the estimates which maximizes the product of spacings or which maximizes the geometric mean of the spacing. That is, we find estimates such that

$$G = \left[\prod_{i=1}^{n+1} D_i\right]^{\frac{1}{n+1}},$$

is maximized. Taking the logarithm of G, we get

$$H = \frac{1}{n+1} \sum_{i=1}^{n+1} \log(D_i).$$

Differentiating the above equation partially, with respect to the parameters α, c, μ and θ respectively and them equating to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

4.3.3 Method of least square estimation

The LS estimators of the unknown parameters α, c, μ and θ of DMLC distribution can be obtained by minimizing

$$\sum_{i=1}^{n} \left[F(x_i \mid \alpha, c, \mu, \theta) - \frac{i}{n+1} \right]^2$$

$$= \sum_{i=1}^{n} \left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^{\alpha}}{1+c[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^{\alpha}} - \frac{i}{n+1} \right]^2 .$$

Partially differentiating the above equation with respect to unknown parameters α, c, μ and θ respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

4.3.4 Method of Cramer-von Mises

The CVM estimators $\hat{\alpha}_{CME}$, $\hat{c}_{CME}$, $\hat{\mu}_{CME}$ and $\hat{\theta}_{CME}$ are the values of α, c, μ and θ minimizing

$$C(\alpha, c, \mu, \theta) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_i \mid \alpha, c, \mu, \theta,) - \frac{2i-1}{2n}\right]^2$$

$$= \frac{1}{12n} + \sum_{i=1}^{n}\left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^\alpha}{1+c[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^\alpha} - \frac{2i-1}{2n}\right]^2.$$

Differentiating the above equation partially, with respect to the parameters α, c, μ and θ respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

4.3.5 Methods of Anderson-Darling and right-tail Anderson-Darling

The AD estimators $\hat{\alpha}_{ADE}$,$\hat{c}_{ADE}$,$\hat{\mu}_{ADE}$ and $\hat{\theta}_{ADE}$ are the values of α, c, μ and θ minimizes

$$\begin{aligned} A(\alpha, c, \mu, \theta) &= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log F(x_i \mid \alpha, c, \mu, \theta) + \log \bar{F}(x_{n+1-i} \mid \alpha, c, \mu, \theta)\} \\ &= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log\left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^\alpha}{1+c[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^\alpha}\right]\right. \\ &\left. + \log\left[1 - \left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^\alpha}{1+c[\frac{1}{\pi}\arctan(\frac{x_{n+1-i}-\mu}{\theta})+0.5]^\alpha}\right]\right]\right\}. \end{aligned}$$

The RTAD estimators $\hat{\alpha}_{RTADE}$,$\hat{c}_{RTADE}$,$\hat{\mu}_{RTADE}$ and $\hat{\theta}_{RTADE}$ are the values of α, c, μ and θ minimizes

$$
\begin{aligned}
R(\alpha,\beta,\mu,\theta) &= \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i \mid \alpha,c,\mu,\theta) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log \bar{F}(x_{n+1-i} \mid \alpha,c,\mu,\theta) \\
&= \frac{n}{2} - 2\sum_{i=1}^{n}\left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^{\alpha}}{1+c[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^{\alpha}}\right] \\
&- \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log\left[1-\left[\frac{(1+c)[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^{\alpha}}{1+c[\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5]^{\alpha}}\right]\right].
\end{aligned}
$$

Partially differentiating the above equation with respect to the parameters α, c, μ and θ respectively and them equating to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

4.3.6 Simulation study

By conducting Monte Carlo simulation to compare the performance of the estimators discussed in the previous sections and the process is repeated 1000 times. We evaluate the performance of the estimators based on bias and MSE. Methods are compared for sample sizes $n = 100$, $n = 250$ and $n = 500$.

For each estimate we calculate the bias, MSE. The statistics are obtained using the following formulae.

$Bias(\hat{\alpha}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha}-\alpha)$ $\quad$ $Bias(\hat{c}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{c}-c)$

$Bias(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}-\mu)$ $\quad$ $Bias(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}-\theta)$

$MSE(\hat{\alpha}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha}-\alpha)^2$ $\quad$ $MSE(\hat{c}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{c}-c)^2$

$MSE(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mu}-\mu)^2$ $\quad$ $MSE(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}-\theta)^2$

The biases and the MSEs of the parameter estimates for the ML, MPS, LS, CVM, AD and RTAD estimation are presented in Table 4.1 and 4.2.

From Table 4.1 and 4.2, we can see that the ML method performs well for estimating the model parameters. Also, as the sample size increases the biases and the MSEs of the average estimates of ML estimates decrease as expected.

The following observations can be drawn from Tables 4.1 and 4.2.

1. All the estimators exhibit the property of consistency, i.e., the MSE decreases as the sample size increases.

2. The bias of all parameters decreases with an increasing n for all the methods of estimations.

3. The bias of $\hat{\mu}$,$\hat{c}$ increases with an increasing mu, c for any given mu, c and n and for all the methods of estimation.

4. All the methods of estimation produce smaller MSE for $\hat{\theta}$ than that of other parameters.

Table 4.1: Simulation results for $\alpha = 0.8$, $c = 1.2$, $\mu = 0.6$, and $\theta = 0.2$.

n	Parameters	ML		MPS		LS		CVM		AD		RTAD	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
100	α	-0.0039	0.03201	-0.0353	0.0152	0.0659	0.0456	0.0571	0.0399	-0.0044	0.0332	0.0258	0.0510
	c	1.0581	1.1196	-0.5985	4.5855	0.1349	1.7005	0.0469	1.8266	-0.3503	2.8755	-0.2966	3.2983
	μ	0.0128	0.0026	0.0010	0.0053	0.0195	0.0043	0.0199	0.0044	0.0030	0.0034	0.0046	0.0040
	θ	0.0026	0.0012	-0.0022	0.0018	-0.0033	0.0017	0.00217	0.0016	0.0009	0.0015	0.0037	0.0017
250	α	-0.0014	0.0103	0.0028	0.0245	0.0364	0.02149	0.0271	0.0200	-0.0009	0.0127	0.0203	0.0275
	c	-0.1832	0.8025	-0.0979	1.1546	0.0881	0.9199	0.0338	0.8894	-0.1514	0.8455	-0.0847	1.2252
	μ	0.0052	0.0012	0.0066	0.0024	0.0124	0.0021	0.0093	0.0019	0.0008	0.0014	0.0043	0.0018
	θ	0.0015	0.0003	-0.0015	0.0007	-0.0018	0.00068	-0.0008	0.0007	0.0014	0.0006	0.0019	0.0006
500	α	0.0026	0.0058	0.0185	0.0181	0.0278	0.0113	0.0101	0.0108	0.0044	0.0071	0.0158	0.0147
	c	-0.0005	0.3019	0.0588	0.5977	0.0842	0.4717	-0.0183	0.4675	-0.0382	0.3599	-0.0006	0.5492
	μ	0.0007	0.0009	0.0096	0.0017	0.0082	0.0011	0.0040	0.0010	0.0008	0.0008	0.0014	0.0043
	θ	0.0005	0.0001	-0.0017	0.0003	0.0009	0.0003	-0.0002	0.0003	0.0001	0.0003	0.0014	0.0003

Figures 4.6 and 4.7 show respectively the MSE of the simulated estimates of α, c, μ and θ for different values of n. Based on these figures, the MSEs of all estimates tends to zero for large n.

Table 4.2: Simulation results for $\alpha = 0.5$, $c = 2$, $\mu = 1$, and $\theta = 1.5$.

n	Parameters	ML		MPS		LS		CVM		AD		RTAD	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
100	α	-0.0017	0.0059	0.0044	0.0117	0.0348	0.0097	0.0128	0.0100	-0.0013	0.0081	0.0060	0.011
	c	-0.1709	2.3902	-0.2943	5.5597	0.1846	2.6119	-0.1844	3.6851	-0.2986	3.1976	-0.3446	4.2437
	μ	0.0408	0.1670	0.3427	1.9234	0.2493	0.6126	0.1917	0.6366	0.0678	0.3273	0.0935	0.22719
	θ	-0.0128	0.1039	-0.0829	0.3078	-0.2101	0.4661	-0.1354	0.4405	-0.0702	0.2092	-0.0194	0.1864
250	α	0.0019	0.0021	0.03124	0.0110	0.0206	0.0054	0.0108	0.0053	0.0022	0.0028	0.0035	0.0042
	c	-0.0596	0.7301	0.2396	2.9268	0.1581	1.2398	0.0150	1.4775	-0.0600	0.8655	-0.1153	1.1827
	μ	0.0315	0.1000	0.4622	1.0162	0.1369	0.3061	0.1248	0.2705	0.0446	0.1269	0.0324	0.1008
	θ	-0.0039	0.0581	-0.1012	0.1951	-0.1147	0.1980	-0.0873	0.1816	-0.0314	0.0871	-0.0139	0.0731
500	α	0.0011	0.0019	0.0429	0.0109	0.0200	0.0045	0.0134	0.0040	0.0003	0.0015	0.0003	0.00236
	c	-0.0561	0.5992	0.4201	2.3069	0.2139	0.8504	0.1369	0.8475	-0.0490	0.4275	-0.0243	0.5673
	μ	0.0214	0.0466	0.5353	0.0795	0.1387	0.2330	0.1269	0.2348	0.0099	0.0594	0.0236	0.0549
	θ	-0.0045	0.0290	-0.1252	0.1594	-0.1234	0.1481	-0.1250	0.1523	-0.0138	0.0398	-0.0165	0.03858

4.4 Applications

In this section we considered two sets of real data and compare the fit of the DMLC distribution with the following distributions:

(a) Two parameter Cauchy distribution having pdf

$$f(x;\mu,\theta) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)};$$

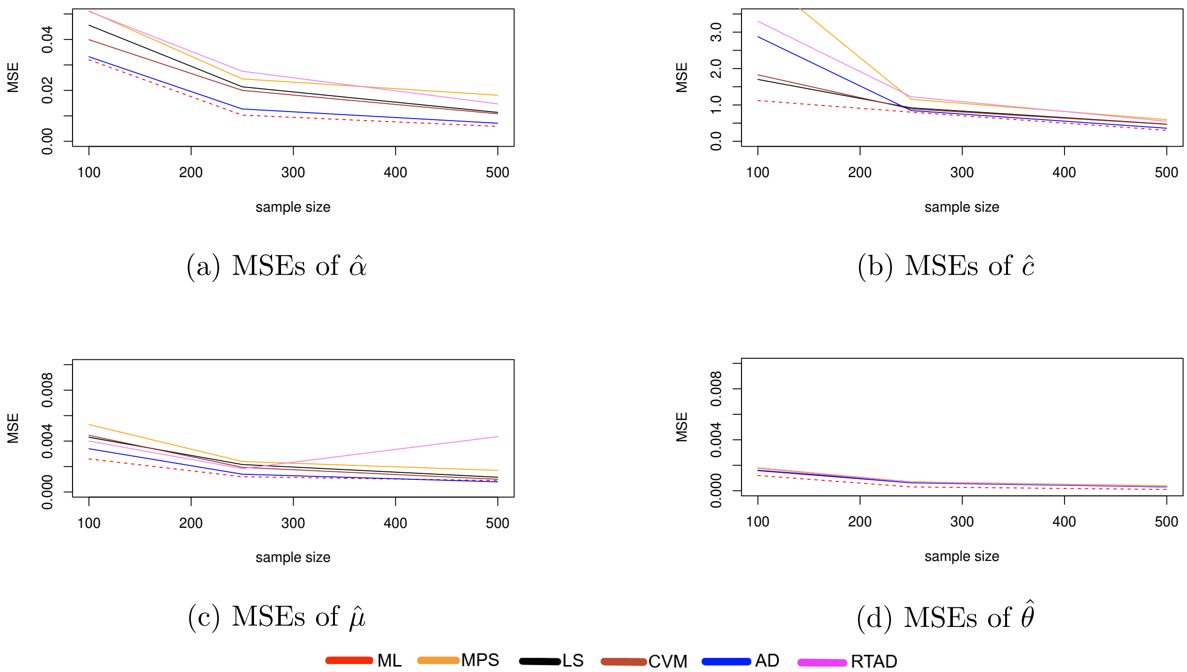
where $x \in \mathbb{R}, -\infty < \mu < \infty, \theta > 0$,

(b) Three parameter SC distribution introduced by Behboodian et al. (2006) with pdf

$$f(x;\mu,\theta,\lambda) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}\left[1+\frac{\lambda(x-\mu)}{\sqrt{\theta^2+(1+\lambda^2)(x-\mu)^2}}\right];$$

where $x \in \mathbb{R}, -\infty < \mu, \lambda < \infty, \theta > 0$.

The values of $-\ln(L)$, AIC, AICC and BIC are calculated for the three distributions to verify which distribution fits better to real data set. The better distribution corresponds to smaller $-\ln(L)$, AIC, AICC and BIC values.



(a) MSEs of $\hat{\alpha}$

(b) MSEs of $\hat{c}$

(c) MSEs of $\hat{\mu}$

(d) MSEs of $\hat{\theta}$

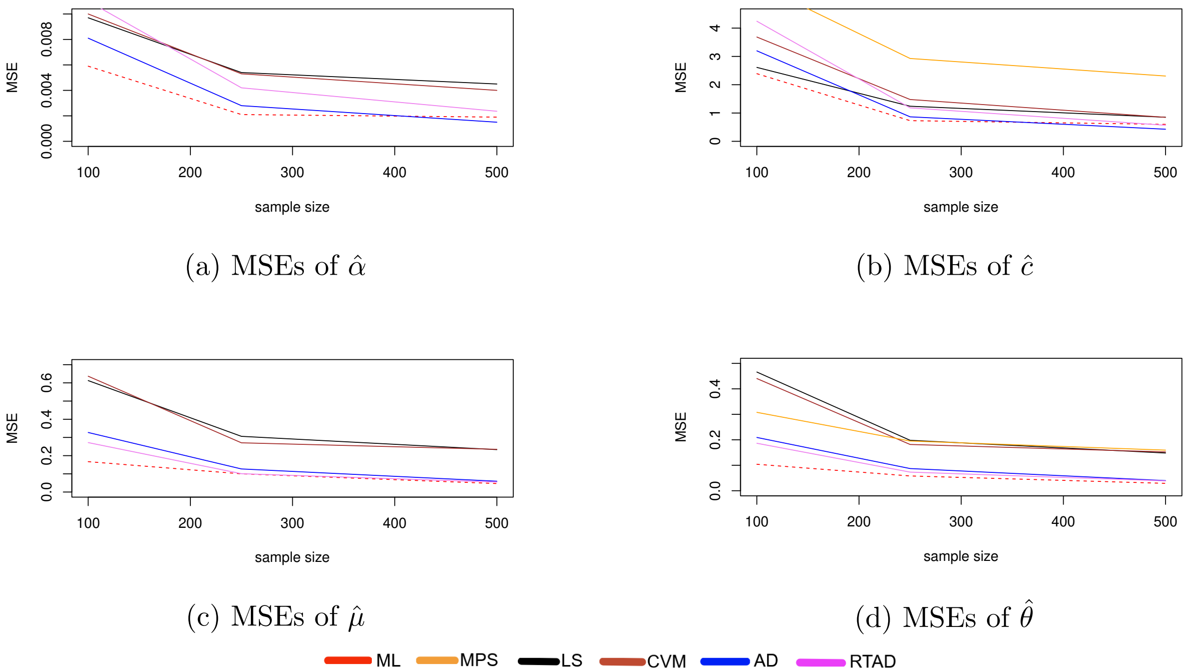
Figure 4.6: MSEs of the estimates of $\alpha = 0.8$,$c = 1.2$,$\mu = 0.6$, and $\theta = 0.2$.

4.4.1 First data set

The real data set corresponds to data set from Weisberg (2005) represents the sum of skin folds in 102 male and 100 female athletes collected at the Australian Institute of Sports. The data set is given in Table 4.3. The data is skewed to the right with skewness=1.175 and kurtosis=1.365.

The descriptive statistics of the first data set are given in Table 4.4. The values in Table 4.5 shows that the DMLC distribution gives a better fit than other two models.

Figure 4.8 shows the fitted density curves for the first data set. Figure 4.9 shows the empirical and the fitted cdfs for the first data set.



(a) MSEs of $\hat{\alpha}$

(b) MSEs of $\hat{c}$

(c) MSEs of $\hat{\mu}$

(d) MSEs of $\hat{\theta}$

Figure 4.7: MSEs of the estimates of $\alpha = 0.5$,$c = 2$,$\mu = 1$, and $\theta = 1.5$.

Table 4.3: Skin folds data.

28	98	89	68.9	69.9	109	52.3	52.8	46.7	82.7	42.3	109.1	96.8	98.3	103.6
110.2	98.1	57	43.1	71.1	29.7	96.3	102.8	80.3	122.1	71.3	200.8	80.6	65.3	78
65.9	38.9	56.5	104.6	74.9	90.4	54.6	131.9	68.3	52	40.8	34.3	44.8	105.7	126.4
83	106.9	88.2	33.8	47.6	42.7	41.5	34.6	30.9	100.7	80.3	91	156.6	95.4	43.5
61.9	35.2	50.9	31.8	44	56.8	75.2	76.2	101.1	47.5	46.2	38.2	49.2	49.6	34.5
37.5	75.9	87.2	52.6	126.4	55.6	73.9	43.5	61.8	88.9	31	37.6	52.8	97.9	111.1
114	62.9	36.8	56.8	46.5	48.3	32.6	31.7	47.8	75.1	110.7	70	52.5	67	41.6
34.8	61.8	31.5	36.6	76	65.1	74.7	77	62.6	41.1	58.9	60.2	43.0	32.6	48
61.2	171.1	113.5	148.9	49.9	59.4	44.5	48.1	61.1	31.0	41.9	75.6	76.8	99.8	80.1
57.9	48.4	41.8	44.5	43.8	33.7	30.9	43.3	117.8	80.3	156.6	109.6	50.0	33.7	54.0
54.2	30.3	52.8	49.5	90.2	109.5	115.9	98.5	54.6	50.9	44.7	41.8	38.0	43.2	70.0
97.21	23.6	181.7	136.3	42.3	40.5	64.9	34.1	55.7	113.5	75.7	99.9	91.2	71.6	103.6
46.1	51.2	43.8	30.5	37.5	96.9	57.7	125.9	49.0	143.5	102.8	46.3	54.4	58.3	34.0
112.5	49.3	67.2	56.5	47.6	60.4	34.9								

Table 4.4: Descriptive statistics of first data set.

Min	1st Q	Median	Mean	3rd Q	Max
28.00	43.85	58.60	69.02	90.35	200.80

Table 4.5: Parameter estimates for various models fitted for the first data set.

Model	parameter estimates	$-\log L$	AIC	AICC	BIC
Cauchy	$\hat{\mu} = 55.5789$ $\hat{\theta} = 16.9283$	1011.7310	2027.4630	2027.5223	2034.0785
SC	$\hat{\mu} = 30.1404$ $\hat{\theta} = 27.9345$ $\hat{\lambda} = 29.6768$	972.6959	1951.3920	1957.8225	1977.2414
DMLC	$\hat{\mu} = -10.2494$ $\hat{\theta} = 11.6394$ $\hat{\alpha} = 82.2378$ $\hat{c} = 85.5034$	964.0236	1936.047	1936.2088	1949.2802

4.4.2 Second data set

Here, we consider a heavy-tailed real data set from the insurance field to explore the usefulness of the DMLC distribution. This data set represents monthly metrics on unemployment insurance from July 2008 to April 2013 including 58 observations, and it is reported by the Department of Labor, Licensing and Regulation, State of Maryland, USA. The data consist of 21 variables and we particularly analyze the variable number 13(% of claims filed via the internet). The data are available at: `https://catalog.data.gov/dataset/unemployment-insurance-data-july-2008-to-april-2013`.

The data is skewed to the left with skewness=-1.3764 and kurtosis=2.8259.

The descriptive statistics of the second data set are given in Table 4.6. The values in Table 4.7 shows that the DMLC distribution gives a better fit than other two models.

Figure 4.10 shows the fitted density curves for the second data set.

Figure 4.11 shows the empirical and the fitted cdfs for the second data set.

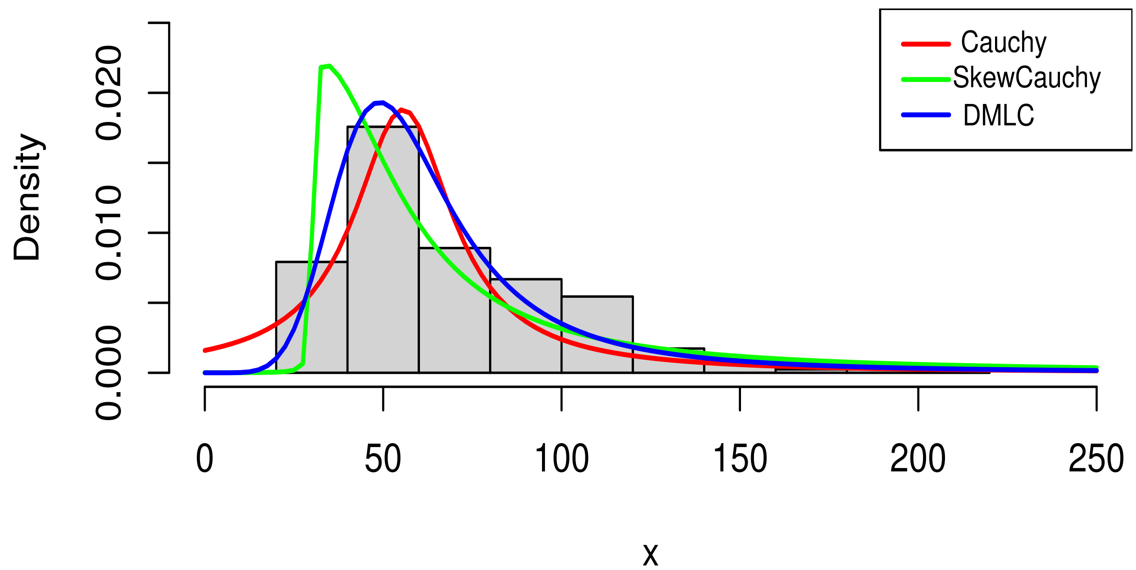


Figure 4.8: Fitted pdf plots of first data set

Table 4.6: Descriptive statistics of second data set.

Min	1st Q	Median	Mean	3rd Q	Max
18.80	40.00	44.00	43.22	48.98	54.50

4.5 Summary

In this chapter, we introduced and studied a new family of distribution called the DMLC distribution which extends Cauchy distribution. In the present work, we studied some basic statistical and mathematical properties of this new model. The DMLC distribution is heavy-tailed and it has regularly varying tails. The model parameters are estimated by the methods of estimation namely ML, MPS, LS, CVM, AD and RTAD. We performed simulation study to compare these methods. The simulation results showed that ML

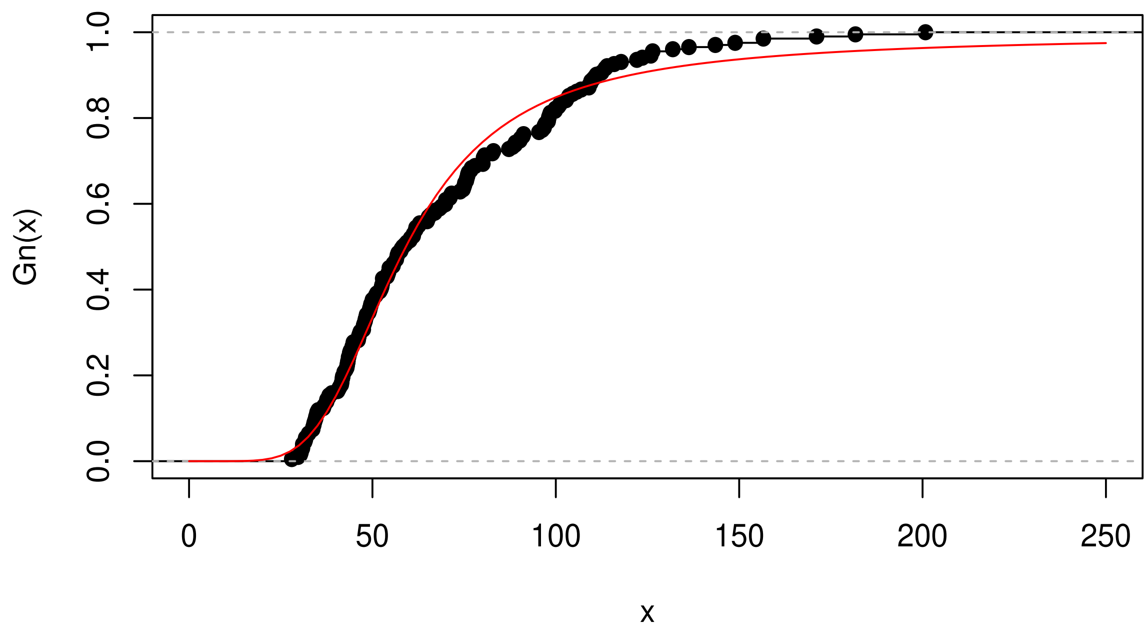


Figure 4.9: Empirical and the fitted cdfs for the first data set

Table 4.7: Parameter estimates for various models fitted for the second data set.

Model	parameter estimates	$-\log L$	AIC	AICC	BIC
Cauchy	$\hat{\mu} = 3.9147$ $\hat{\theta} = 43.6130$	201.1985	406.3969	406.6151	410.5178
SC	$\hat{\mu} = 41.4275$ $\hat{\theta} = 0.5385$ $\hat{\lambda} = 3.8896$	200.9152	407.8305	408.2748	414.0117
DMLC	$\hat{\mu} = 52.7924$ $\hat{\theta} = 7.5182$ $\hat{\alpha} = 3.3043$ $\hat{c} = 120.6444$	193.6542	395.3085	396.0631	403.550

estimators are the better performing estimator in terms of biases and MSEs. Fitting the DMLC model with real data sets indicates the flexibility and capacity of this new distribution in data modeling.

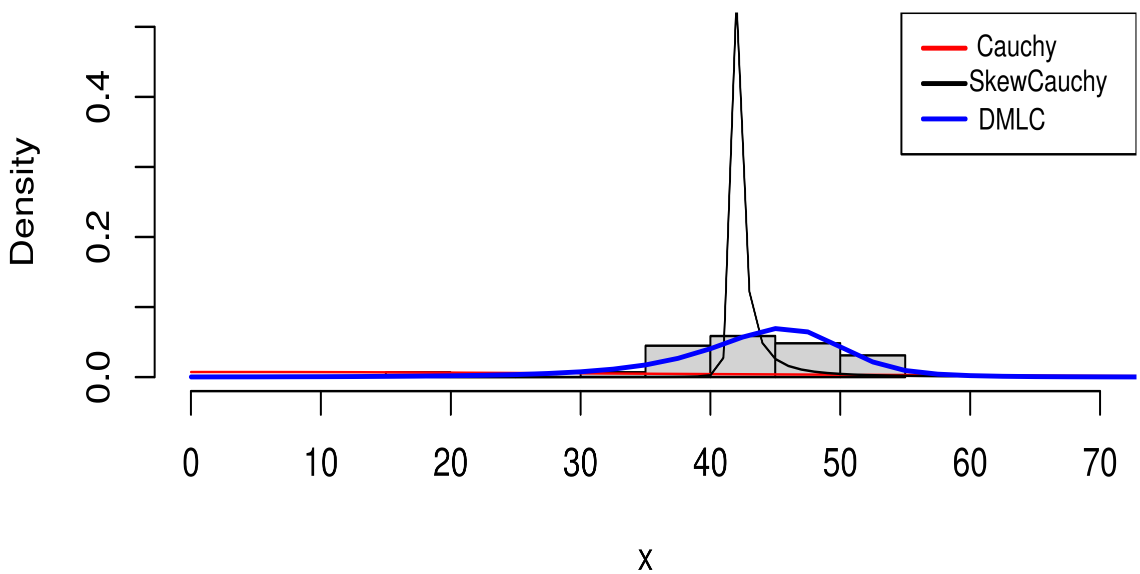


Figure 4.10: Fitted pdf plots of second data set

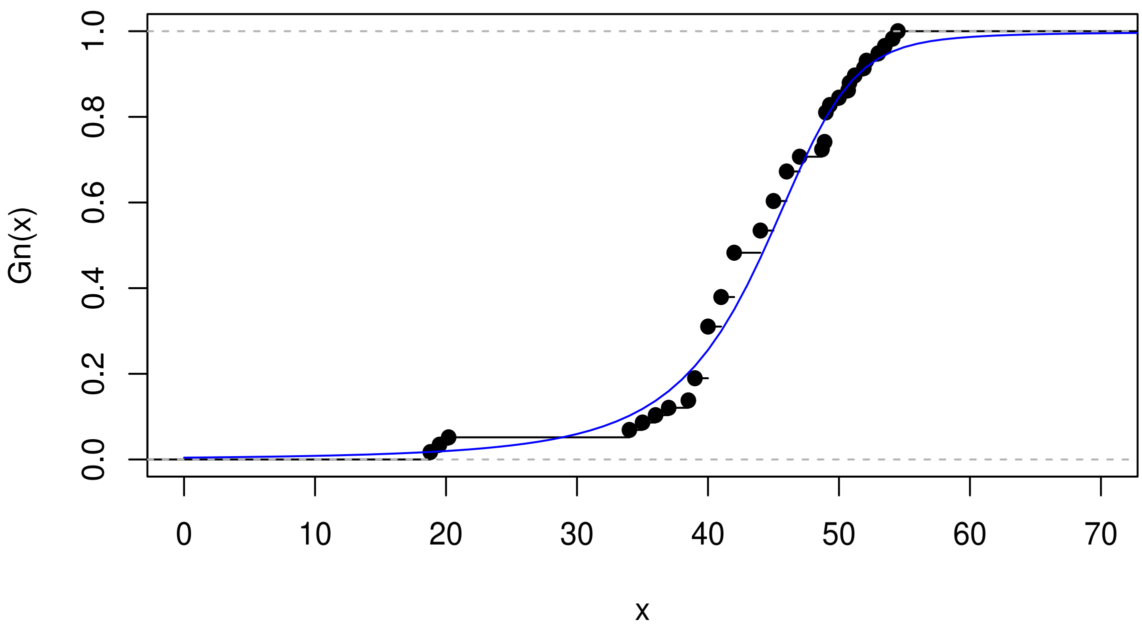


Figure 4.11: Empirical and the fitted cdfs for the second data set

Chapter 5

A NEW METHOD FOR GENERATING DISTRIBUTIONS WITH AN APPLICATION TO CAUCHY DISTRIBUTION

5.1 Introduction

The development of new methods of elaborating the existing distributions is a wide area in distribution theory. Various approaches have been proposed in statistical literature to generate new distributions based on a baseline distribution. This has been accomplished through various approaches.

In Statistical literature, number of transformations are accessible to produce new cdf corresponding to a given cdf. Suppose, we have a cdf F(x), then the related proposed cdf will be $G_i(x)$.

[1] This Chapter is based on Fasna (2022)

- The power transformation initiated by Gupta et al. (1998) having cdf of the form

$$G_1(x) = [F(x)]^{\alpha}; \alpha > 0$$

- QRTM introduced by Shaw and Buckley (2009) having cdf of the form

$$G_2(x) = (1 + \lambda)F(x) - \lambda F^2(x); \mid \lambda \mid \leq 1$$

- DUS transformation proposed by Kumar et al. (2015a) having cdf of the form

$$G_3(x) = \frac{e^{F(x)} - 1}{e - 1}; e = exp(1)$$

- SS-transformation proposed by Kumar et al. (2015b) having cdf of the form

$$G_4(x) = sin\left(\frac{\pi}{2}F(x)\right)$$

- Minimum Guarantee (MG) distribution introduced by Kumar et al. (2017) having the form

$$G_5(x) = e^{1-\frac{1}{F(x)}}$$

- Log-transformation proposed by Maurya et al. (2016) and having cdf of the form

$$G_6(x) = 1 - \frac{\ln(2 - F(x))}{\ln 2}$$

- Transformation based on the generalization of Kumar et al. (2015a) called GDUS transformation proposed by Maurya et al. (2017) having cdf of the form

$$G_7(x) = \frac{e^{F^\alpha(x)} - 1}{e - 1}; \alpha > 0$$

- New transformation proposed by Kyurkchiev (2017) to develop a sigmoid family of functions for Verhulst Logistic function is

$$G_9(x) = \frac{2F(x)}{1 + F(x)}$$

- New trigonometry transformation called PCM proposed by Kumar et al. (2021) and having cdf of the form

$$G_{10}(x) = tan\left(\frac{\pi}{4}F(x)\right)$$

In this continuation after taking motivations and ideas from the above discussed transformations, we introduce a new transformation called the beta transformation for $x \in \mathbb{R}$ is given below

$$G(x) = \begin{cases} \frac{\beta}{\beta-1}[1 - \beta^{-F(x)}] & \text{if } \beta > 0, \beta \neq 1 \\ F(x) & \text{if } \beta = 1 \end{cases} \tag{5.1}$$

Where, $G(x)$ and $F(x)$ are the cdfs of this new transformation and baseline distribution. On differentiating (5.1) with respect to x, we get the pdf $g(x)$ and is given by

$$g(x) = \begin{cases} \frac{\beta \log \beta}{\beta-1} f(x)\beta^{-F(x)} & \text{if } \beta > 0, \beta \neq 1 \\ f(x) & \text{if } \beta = 1 \end{cases} \tag{5.2}$$

For $\beta \neq 1$, $g(x)$ is a weighted version of $f(x)$, where the weight function

$$w(x) = \beta^{-F(x)},$$

and $g(x)$ can be written as

$$g(x) = \frac{f(x)w(x)}{c}.$$

where constant $c = E(w(X))$,

Here $c = \frac{\beta-1}{\beta \log \beta}$.

The sf $s(x)$ and the hrf $h(x)$ are obtained as

$$s(x) = \begin{cases} \frac{\beta^{1-F(x)}-1}{\beta-1} & \text{if } \beta \neq 1 \\ 1 - F(x) & \text{if } \beta = 1 \end{cases} \tag{5.3}$$

and

$$h(x) = \begin{cases} f(x)\frac{\log \beta \beta^{1-F(x)}}{\beta^{1-F(x)}-1} & \text{if } \beta \neq 1 \\ \frac{f(x)}{S(x)} & \text{if } \beta = 1 \end{cases} \tag{5.4}$$

Lifetime models are widely used in various fields to analyze and explain the lifespan or failure behavior of systems or devices. These models are apt in biological field, reliability, insurance, engineering, etc. The motivations for introducing our beta transformed model is that it is efficient to analyze lifetime data and very easy method of inducting an additional parameter to a family of distributions functions. It improve the characteristics, bring more flexibility to the given family and provide better fits than the other models having the same or higher number of parameters. The proposed method is very interesting with a closed form for the cdf and capable of modeling heavy tailed data sets.

The aim of this chapter is to introduce a transformation that yields new distributions by using a given baseline distribution. It contains only one new parameter other than the parameters involved in the baseline distribution. To illustrate the usefulness of this new transformation, we choose Cauchy as the baseline distributions in the present work.

This chapter is organized as follows. We introduce a special sub-case of (5.1) called a beta transformed Cauchy (BTC) distribution by considering Cauchy model as a parent distribution and discuss the shape of the density function and distribution function of the model. We derive the quantiles, mode and pdf of order statistics. Certain characterizations of the proposed distribution are to be addressed. The method of ML, CVM, AD and LS estimations are discussed. We examine the usefulness of the proposed distribution by means of real data set which exhibits the performance of BTC model compared with other Cauchy generated models. Estimation techniques are applied to calculate the model parameters. Finally, the summary is presented.

5.2 Beta transformed Cauchy distribution

In this section, a sub model of the beta transformed family called the beta transformed Cauchy (BTC) distribution is proposed.

5.2.1 Distribution function

Let $F(x;\mu,\theta)$ be cdf of Cauchy r.v. given by $F(x) = \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5$. Using this in equation (5.1), then the cdf of the BTC for $x \in \mathbb{R}$ with the parameters $\mu \in \mathbb{R}$, $\theta > 0$ and

$\beta > 0$ denoted by $BTC(\mu, \theta, \beta)$ has the following form

$$G(x) = \begin{cases} \frac{\beta}{\beta-1}[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)}] & \text{if } \beta \neq 1 \\ \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) + 0.5 & \text{if } \beta = 1 \end{cases} \tag{5.5}$$

Figure 5.1 provides the plots of the cdf of the model for various choices of the values of the parameters.

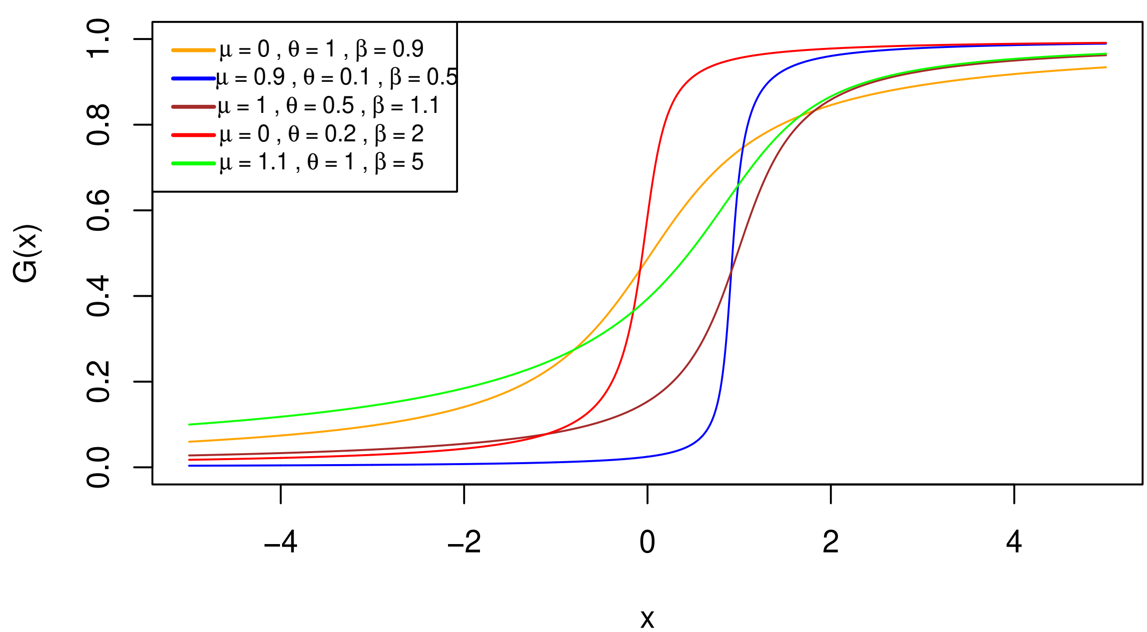


Figure 5.1: Plots of the cdf of $BTC(\mu, \theta, \beta)$ distribution

5.2.2 Probability density function

The pdf g(x) is given by

$$g(x) = \begin{cases} \frac{\log\beta}{\pi\theta(\beta-1)} \frac{1}{1+(\frac{x-\mu}{\theta})^2} \beta^{0.5-\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})} & \text{if } \beta \neq 1 \\ \frac{1}{\pi\theta} \frac{1}{(1+\frac{x-\mu}{\theta})^2} & \text{if } \beta = 1 \end{cases} \tag{5.6}$$

Figure 5.2 provides the plots of the pdf of the model for various choices of the values of the parameters.

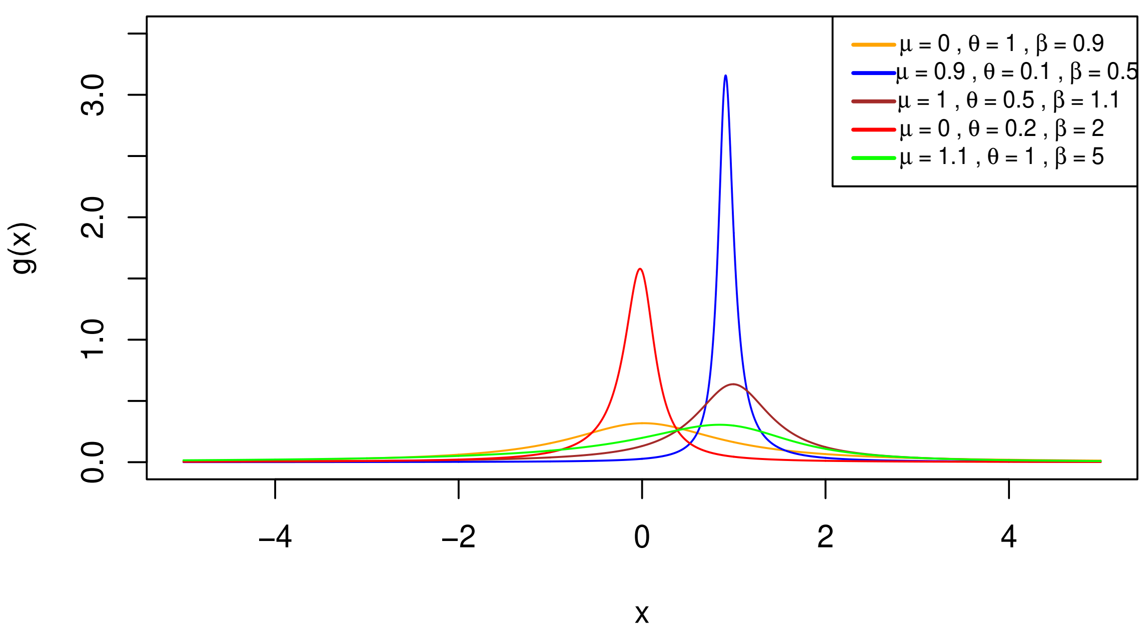


Figure 5.2: Plots of the pdf of $BTC(\mu, \theta, \beta)$ distribution

5.2.3 Survival function

The sf $s(x)$ is obtained as

$$s(x) = \begin{cases} \frac{\beta^{0.5-\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})}-1}{\beta-1} & \text{if } \beta \neq 1 \\ 0.5 - \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}) & \text{if } \beta = 1 \end{cases} \tag{5.7}$$

Figure 5.3 provides the plots of the sf of the model for various choices of the values of the parameters.

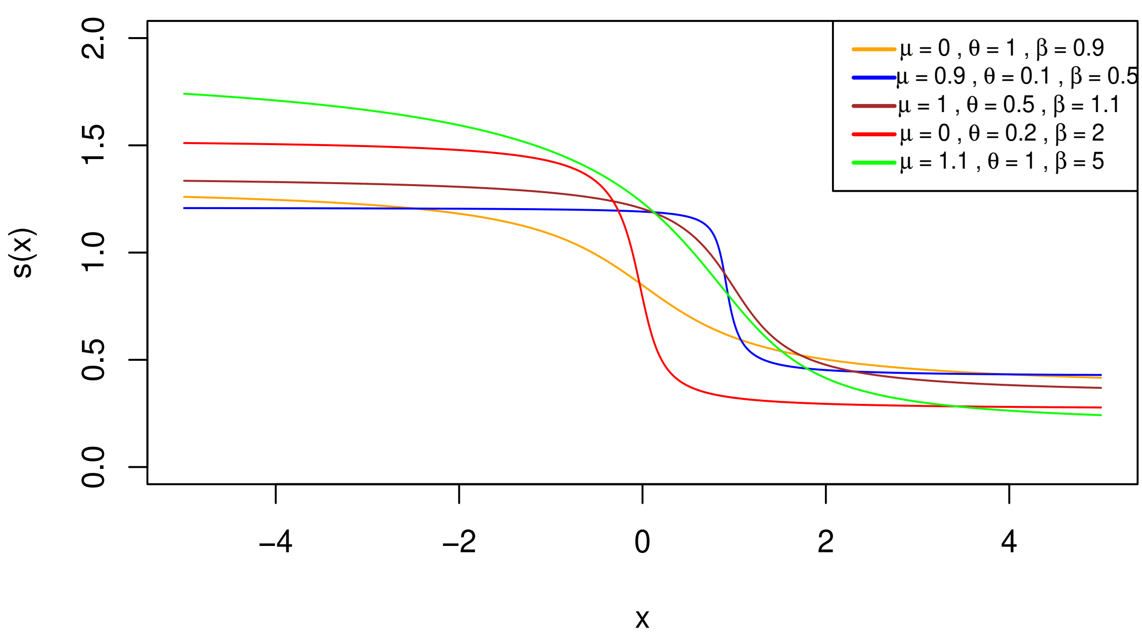


Figure 5.3: Plots of the survival function of $BTC(\mu, \theta, \beta)$ distribution

5.2.4 Hazard rate function

The hrf $h(x)$ is

$$h(x) = \begin{cases} \frac{\log\beta}{\pi\theta} \frac{\beta^{0.5-\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})}}{(1+(\frac{x-\mu}{\theta})^2)(\beta^{0.5-\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})}-1)} & \text{if } \beta \neq 1 \\ \frac{1}{\pi\theta} \frac{1}{(1+(\frac{x-\mu}{\theta})^2)(0.5-\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))} & \text{if } \beta = 1 \end{cases} \tag{5.8}$$

Figure 5.4 provides the plots of the hrf of the model for various choices of the values of the parameters.

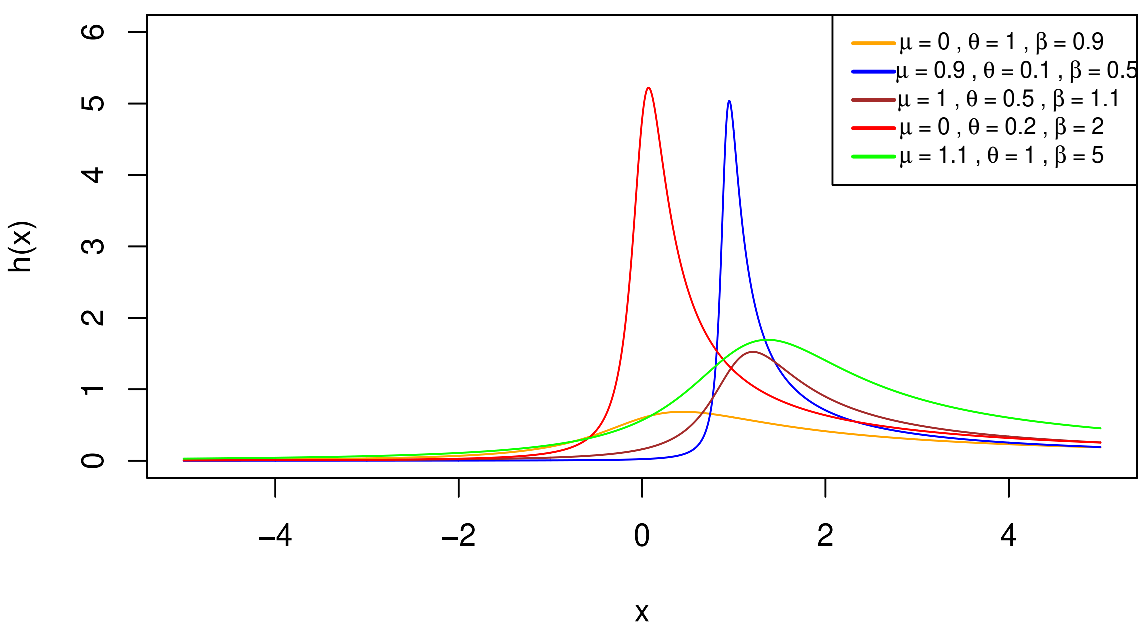


Figure 5.4: Plots of the hazard rate function of $BTC(\mu, \theta, \beta)$ distribution

5.2.5 Quantile function

Theorem 5.2.1. *The q^{th} quantile x_q of the BTC r.v. is given by*

$$x_q = \mu - \theta \tan\left(\pi \frac{\log\left(\frac{\beta - q(1-\beta)}{\sqrt{\beta}}\right)}{\log \beta}\right). \tag{5.9}$$

Proof. The q^{th} quantile x_q of the BTC r.v is defined as

$$q = P(X \leq x_q) = G(x_q), \quad x_q \in \mathbb{R}$$

Using the cdf of the BTC distribution, we have

$$q = G(x_q) = \frac{\beta}{\beta - 1}[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)}]$$

That is,

$$\beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)} = 1 - \frac{q(\beta-1)}{\beta}$$

Hence

$$x_q = \mu - \theta \tan\left(\pi \frac{\log\left(\frac{\beta - q(1-\beta)}{\sqrt{\beta}}\right)}{\log \beta}\right).$$

This completes the proof. □

Using the inversion method, we can generate r.v.s from the BTC distribution. We can use equation (5.9) to generate random numbers when the parameters μ, θ and β are known.

Hence, the median of BTC is given by,

$$x_{0.5} = \mu - \theta \tan\left(\pi \frac{\log\left(\frac{\beta - 0.5(1-\beta)}{\sqrt{\beta}}\right)}{\log \beta}\right). \tag{5.10}$$

It can be seen that $BTC(\mu, \theta, \beta)$ is a uni-modal function with mode at $\mu - \frac{\theta \log \beta}{2\pi}$.

5.2.6 Order statistics

Let $X_1, X_2, \ldots, X_n$ be a random sample from $BTC(\mu, \theta, \beta)$. Also, let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denote the corresponding order statistics. Then the pdf and cdf of k^{th} order statistics are given by

$$\begin{aligned} g_X(x) &= \frac{n!}{(k-1)!(n-k)!}[G(x)]^{k-1}\left[1 - G(x)\right]^{n-k} g(x) \\ &= \frac{n!}{(k-1)!(n-k)!}\frac{\log\beta}{\pi\theta(\beta-1)}\frac{1}{1+(\frac{x-\mu}{\theta})^2}\beta^{\left[0.5 - \frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})\right]} \\ &\left[\frac{\beta}{\beta-1}\left[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)}\right]\right]^{k-1}\left[1 - \left[\frac{\beta}{\beta-1}\left[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)}\right]\right]\right]^{n-k} \end{aligned} \tag{5.11}$$

and

$$\begin{aligned} G_X(x) &= \sum_{j=k}^{n}\binom{n}{j}[G(x)]^j[1-G(x)]^{n-j} \\ &= \sum_{j=k}^{n}\binom{n}{j}\left[\frac{\beta}{\beta-1}\left[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)}\right]\right]^j \\ &\left[1 - \left[\frac{\beta}{\beta-1}\left[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})+0.5)}\right]\right]\right]^{n-j} \end{aligned} \tag{5.12}$$

respectively.

The pdf of the minimum and maximum of order statistics are obtained by putting $X = X_1$

and $X = X_n$ respectively in equation (5.11).

5.3 Characterizations of beta transformed Cauchy distribution

In this section, we discuss certain characterizations of the BTC distribution based on a relationship between two truncated moments. This characterization result employs a theorem due to Glanzel (1987) which stated as follows:

Theorem 5.3.1. *Let $(\Omega, \mathcal{F}, \boldsymbol{P})$ be a given probability space and let $H = [a, b]$ be an interval for some $a < b$ ($a = -\infty$, $b = \infty$ might as well be allowed). Let $X : \Omega \to \mathbf{H}$ be a continuous r.v. with the distribution function F and let q_1 and q_2 be two real functions defined on $\mathbf{H}$ such that*

$$E[q_2(X)|X \geq x] = E[q_1(X)|X \geq x]\eta(x), x \in \mathbf{H},$$

is defined with some real function η. Assume that q_1, q_2 are continuous functions, η has continuous derivative and F is twice continuously differentiable and strictly monotone function on the set $\mathbf{H}$. Finally, assume that the equation $\eta q_1 = q_2$ has no real solution in the interior of $\mathbf{H}$. Then F is uniquely determined by the functions q_1, q_2 and η particularly

$$F(x) = \int_a^x \mathcal{C} \mid \frac{\eta^{'}(\mu)}{\eta(\mu)q_1(u) - q_2(u)} \mid \exp(-s(u))du,$$

where the function s is a solution of the differential equation $s^{'} = \frac{\eta^{'} q_1}{\eta q_1 - q_2}$ and $\mathcal{C}$ is a constant, chosen to make $\int_{\mathbf{H}} dF = 1$.

Proposition 5.3.1. *Let $X : \Omega \to (0, \infty)$ be a continuous r.v. and let $q_1(x) = e^{-x}(1 + (\frac{x-\mu}{\theta})^2)$*

$\beta^{-(0.5-\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))}$ *and* $q_2(x) = 2q_1(x)e^{-x}$ *for* $x \in \mathbb{R}$. *The r.v X has pdf* (5.6) *if and only if the function* η *defined in Theorem 5.3.1 has the form*

$$\eta(x) = 4e^{-x},\ x \in \mathbb{R}$$

Proof. Let X be a r.v with pdf (5.6), then

$(1 - G(x))E[q_1(X) \mid X \geq x] = \frac{\log\beta}{\pi\theta(\beta-1)}e^{-x},\ x \in \mathbb{R}$

and

$(1 - G(x))E[q_2(X) \mid X \geq x] = 4\frac{\log\beta}{\pi\theta(\beta-1)}e^{-2x},\ x \in \mathbb{R}$

and finally

$\eta(x)q_1(x) - q_2(x) = 2q_1e^{-x},\ for\ x \in \mathbb{R}.$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta^{'}(x)q_1(x)}{\eta q_1(x) - q_2(x)} = -2,\ x \in \mathbb{R}$$

and hence

$s(x) = -2x, x \in \mathbb{R}$, or $e^{-s(x)} = e^{2x},\ x \in \mathbb{R}$. Now, in view of Theorem 5.3.1, X has density (5.6). □

Corollary 5.3.1. *Let* $X : \Omega \to (0, \infty)$ *be a continuous r.v. and let* $q_1(x)$ *be as in Proposition 5.3.1. The pdf of X is* (5.6) *if and only if there exist functions* q_2 *and* η *defined in Theorem 5.3.1 satisfying the differential equation*

$$\frac{\eta^{'}(x)q_1(x)}{\eta q_1(x) - q_2(x)} = -2, x \in \mathbb{R}$$

Remark 5.3.1. *The general solution of the differential equation in Corollary 5.3.1 is*

$$\eta(x) = 2e^{-2x}\left[\int (e^{2x})[q_1(x)]^{-1}q_2(x)dx + D\right]$$

where D is a constant. Note that a set of functions satisfying the above differential equation is given in Proposition 5.3.1 with $D = 0$. However, it should be also noted that there are other triplets (q_1, q_2, η) satisfying the conditions of Theorem 5.3.1.

5.4 Estimation of parameters

In this section, we describe the methods of ML, CVM, AD and LS estimation to estimate the parameters μ, θ and β of BTC distribution.

5.4.1 Method of maximum likelihood

If $x_1, x_2, \dots, x_n$ is a random sample of size n from the BTC distribution with unknown parameters μ, θ and β. Then its log likelihood function will be as follows

$$\begin{aligned}\log L = n\log\left(\frac{\log\beta}{\beta-1}\right) - n\log(\pi\theta) + \sum_{i=1}^{n}\log\frac{1}{1+(\frac{x_i-\mu}{\theta})^2} \\ + \log\beta\sum_{i=1}^{n}\left(0.5 - \frac{1}{\pi}\arctan\left(\frac{x_i-\mu}{\theta}\right)\right).\end{aligned} \tag{5.13}$$

The likelihood equations are,

$$\frac{\partial\log L}{\partial\mu} = \frac{2}{\theta^2}\sum_{i=1}^{n}\frac{x_i-\mu}{1+(\frac{x_i-\mu}{\theta})^2} - \frac{\log\beta}{\pi\theta}\sum_{i=1}^{n}\frac{1}{1+(\frac{x_i-\mu}{\theta})^2} = 0, \tag{5.14}$$

$$\frac{\partial\log L}{\partial\theta} = -\frac{n}{\theta} + \frac{2}{\theta^3}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{1+(\frac{x_i-\mu}{\theta})^2} + \frac{\log\beta}{\pi\theta^2}\sum_{i=1}^{n}\frac{x_i-\mu}{1+(\frac{x_i-\mu}{\theta})^2} = 0, \tag{5.15}$$

and

$$\frac{\partial \log L}{\partial \beta} = \frac{n}{\beta \log \beta} - \frac{n}{\beta - 1} + \frac{1}{\beta}\sum_{i=1}^{n}\left(0.5 - \frac{1}{\pi}\arctan\left(\frac{x_i - \mu}{\theta}\right)\right)$$

$$= 0. \tag{5.16}$$

These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

5.4.2 Method of Cramer-von Mises

The CVM estimators $\hat{\mu}_{CME}$, $\hat{\theta}_{CME}$ and $\hat{\beta}_{CME}$ are the values of μ, θ and β minimizing

$$C(\mu, \theta, \beta) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_i \mid \mu, \theta, \beta) - \frac{2i-1}{2n}\right]^2$$

$$= \frac{1}{12n} + \sum_{i=1}^{n}\left[\left(\frac{\beta}{\beta - 1}[1 - \beta^{-(\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5)}]\right) - \frac{2i-1}{2n}\right]^2.$$

Partially differentiating the above equations, with respect to the parameters μ, θ and β respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

5.4.3 Method of Anderson-Darling

The AD estimators $\hat{\mu}_{ADE}$, $\hat{\theta}_{ADE}$ and $\hat{\beta}_{ADE}$ are the values of μ, θ and β minimizes

$$A(\mu,\theta,\beta) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\{\log G(x_i \mid \mu,\theta,\beta) + \log \bar{G}(x_{n+1-i} \mid \mu,\theta,\beta)\}$$

$$= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\Bigg\{\log\left(\frac{\beta}{\beta-1}[1-\beta^{-(\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5)}]\right)$$
$$+\log\left(1-\frac{\beta}{\beta-1}[1-\beta^{-(\frac{1}{\pi}\arctan(\frac{x_{n+1-i}-\mu}{\theta})+0.5)}]\right)\Bigg\}.$$

Partially differentiating the above equations, with respect to the parameters μ, θ and β respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

5.4.4 Method of least square estimation

The LS estimators of the unknown parameters μ, θ and β of BTC distribution can be obtained by minimizing

$$\sum_{i=1}^{n}\left[F(x_i \mid \mu,\theta,\beta) - \frac{i}{n+1}\right]^2$$

$$= \sum_{i=1}^{n}\left[\left(\frac{\beta}{\beta-1}[1-\beta^{-(\frac{1}{\pi}\arctan(\frac{x_i-\mu}{\theta})+0.5)}]\right) - \frac{i}{n+1}\right]^2$$

with respect to unknown parameters μ, θ and β.

Differentiating the above equation partially with respect to the parameters μ, θ and β respectively and equating them to zero, we get the normal equations. These equations do not have exact solutions and they have to be obtained numerically by using statistical softwares like *optim* package in R programming.

5.5 Applications

In this section, we consider a real life data set to explore the importance of our proposed model. We compare the fit of the BTC distribution with different transformation of distribution where Cauchy distribution as the baseline distribution.

1. Two parameter Cauchy distribution having pdf

$$f(x;\mu,\theta) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)},$$

where $-\infty < x < \infty, -\infty < \mu < \infty, \theta > 0$.

2. The pdf of the new distribution obtained by DUS transformation by Kumar et al. (2015a), is the DUS Cauchy (DUSC) distribution having pdf

$$f(x;\mu,\theta) = \frac{1}{\pi\theta(e-1)}\frac{1}{1+(\frac{x-\mu}{\theta})^2}e^{(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))},$$

where $-\infty < x < \infty, -\infty < \mu < \infty, \theta > 0$.

3. The pdf of the distribution obtained by minimum guarantee (MG) distribution by Kumar et al. (2017), is the MG Cauchy (MGC) distribution with pdf

$$f(x;\mu,\theta) = \frac{1}{\pi\theta}\frac{e^{\left(1-\frac{1}{0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})}\right)}}{(1+(\frac{x-\mu}{\theta})^2)(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))^2},$$

where $-\infty < x < \infty, -\infty < \mu < \infty, \theta > 0$.

4. The pdf of the distribution obtained by Pi-Exponentiated Transformation (PET) by Lone and Jan (2023), is the PET Cauchy (PETC) distribution with pdf

$$f(x;\alpha,\mu,\theta) = \frac{\alpha\log\pi}{\pi\theta(\pi-1)}\pi^{(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta}))^{\alpha}}\left(0.5+\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)\right)^{\alpha-1},$$

where $-\infty < x < \infty, \alpha > 0, -\infty < \mu < \infty, \theta > 0.$

5. The pdf of the distribution obtained by QRTM suggested by Shaw and Buckley (2009), is the TC distribution with pdf

$$f(x;\lambda,\mu,\theta) = \frac{1}{\pi\theta}\frac{1}{(1+(\frac{x-\mu}{\theta})^2)}\left[(1+\lambda)-2\lambda\left(\frac{1}{\pi}\arctan\left(\frac{x-\mu}{\theta}\right)+0.5\right)\right],$$

where $-\infty < x < \infty, -1 < \lambda < 1, -\infty < \mu < \infty, \theta > 0.$

6. The pdf of the distribution obtained by Alpha power transformation (APT) by Mahdavi and Kundu (2017), is the Alpha power transformed Cauchy distribution with pdf

$$f(x;\alpha,\mu,\theta) = \begin{cases} \frac{\log\alpha}{\pi\theta(\alpha-1)}\frac{1}{1+(\frac{x-\mu}{\theta})^2}\alpha^{\left(0.5+\frac{1}{\pi}\arctan(\frac{x-\mu}{\theta})\right)} & \text{if } \alpha > 0,\ \alpha \neq 1 \\ \frac{1}{\pi\theta}\frac{1}{(1+\frac{x-\mu}{\theta})^2} & \text{if } \alpha = 1 \end{cases}$$

where $-\infty < x < \infty, -\infty < \mu < \infty, \theta > 0$

For the data set, we estimate the unknown parameters of each distribution by the method of CVM, ML, AD and LS. With these obtained estimates, we obtain the values of $-\ln(L)$, AIC, AICC and BIC as well as Kolmogorov-Smirnov statistic (K-S), Cramer-von Mises statistic (W^*), Anderson Darling statistic (A^*) and their corresponding p-values. The required computations are carried out in the R-language introduced by Team (2009).

5.5.1 Data set

The real data set corresponds to data set from Nichols and Padgett (2006) on breaking stress of carbon fibres (in Gba): The data set is given in the Table 5.1.

The data is approximately symmetric with skewness=0.541 and kurtosis=0.141.

Table 5.1: Carbon fibres data set

3.70	2.74	2.73	2.50	3.60	3.11	3.27	2.87	1.47	3.11
4.42	3.22	1.69	3.28	3.09	1.87	3.15	4.90	3.75	2.43
2.95	2.97	3.39	2.67	2.93	3.22	3.39	2.81	4.20	3.33
2.55	3.31	3.31	2.85	2.56	2.35	2.55	2.59	2.38	2.81
2.77	2.17	2.83	1.92	1.41	3.68	2.97	2.76	4.91	3.68
1.84	1.59	3.19	1.57	0.81	5.56	1.73	1.59	2.00	1.71
2.17	1.17	5.08	2.48	1.18	3.51	2.17	1.69	1.25	4.38
1.84	2.48	0.85	1.61	2.79	4.70	2.03	1.80	1.57	1.08
2.03	1.61	2.12	1.89	2.05	3.65				

The descriptive statistics of this data set are given in Table 5.2. The ML estimate for μ, θ and β are given in Table 5.3 along with their standard errors (S.E.). The values in Table 5.4 expose that the BTC model leads to a better fit than the other models. Based on the values of the AIC, AICC and BIC criteria, we observe that the BTC model provides the best fit for these data among all the models considered.

Table 5.2: Descriptive statistics of second data set.

Min	1st Q	Median	Mean	3rd Q	Max
0.810	1.875	2.700	2.673	3.257	5.560

Table 5.3: ML Estimates and S.E for μ, θ and β

Parameters	ML Estimate	S.E.
μ	2.6479	0.2685
θ	0.6047	0.0853
β	1.0543	0.9856

Figure 5.5 shows the fitted density curves for the data set. Figure 5.6, shows the

Table 5.4: Parameter estimates for various models fitted for the data set.

Model	parameter estimates	log L	AIC	AICC	BIC
Cauchy	$\hat{\mu} = 2.3966$ $\hat{\theta} = 0.1000$	-239.3104	484.6208	484.9134	491.9838
DUSC	$\hat{\mu} = 2.900$ $\hat{\theta} = 0.05$	-187.471	378.942	381.235	383.850
MGC	$\hat{\mu} = 1.749$ $\hat{\theta} = 0.362$	-175.036	354.071	354.215	358.98
PETC	$\hat{\alpha} = 5.371$ $\hat{\mu} = 3.649$ $\hat{\theta} = 2.079$	-155.458	316.916	317.21	324.277
TC	$\hat{\lambda} = 0.689$ $\hat{\mu} = 3.100$ $\hat{\theta} = 0.95$	-138.752	281.504	281.648	286.412
APTC	$\hat{\alpha} = 0.9484$ $\hat{\mu} = 2.647$ $\hat{\theta} = 0.6047$	-137.62	279.16	281.532	288.603
BTC	$\hat{\beta} = 1.0543$ $\hat{\mu} = 2.6479$ $\hat{\theta} = 0.6047$	-136.48	278.96	279.252	286.323

empirical and the fitted cdfs for the data set.

The goodness of fit of CVM, ML, AD and LS methods are observed by the test statistics values and their p-values for K-S, W^* and A^* for the dataset which are displayed in Table 5.5. Figure 5.7 shows fitted distribution's histogram and the density function using CVM,

Table 5.5: Statistics values and their associated p-values for the dataset

Estimation method	Estimates	K-S(p-value)	W^*(p-value)	A^* (p-value)
CVM	2.7058 0.5889 1.2645	0.1063(0.285)	0.1397(0.423)	1.3776(0.2086)
ML	2.6479 0.6047 1.0543	0.1035(0.3146)	0.1417(0.4165)	1.416(0.1979)
AD	2.1711 0.5480 1.2961	0.1002(0.3535)	0.1471(0.3989)	1.333(0.2219)
LS	2.7025 0.6020 1.2483	0.1081(0.2671)	0.1404(0.4027)	1.4088(0.1998)

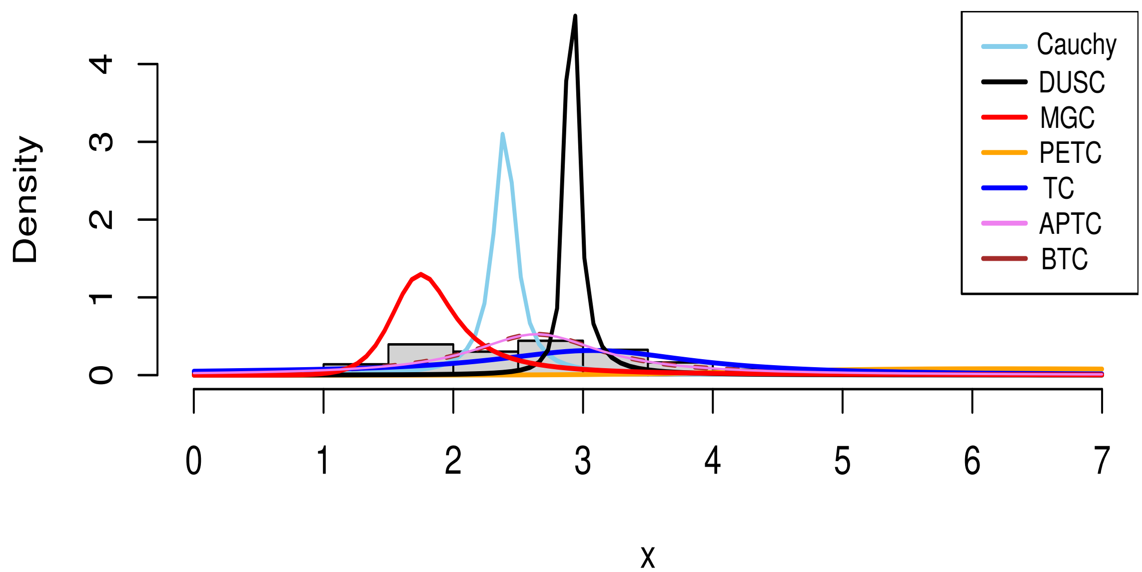


Figure 5.5: Fitted pdf plots of data set

ML, AD and LS estimation methods for the data set of BTC distribution.

5.6 Summary

In this chapter, we discussed beta transformation to get a transformed distribution of some available baseline distribution. Beta transformation of Cauchy distribution considered to check its application to the real problem called the BTC distribution. We provided expressions for quantiles, hazard rates and order statistics. A useful characterizations based on truncated moments are stated for the BTC distribution. The advantage of these characterizations is that the cdf is not required to have a closed form and are given in terms of an integral whose integrand depends on the solution of a first order differential

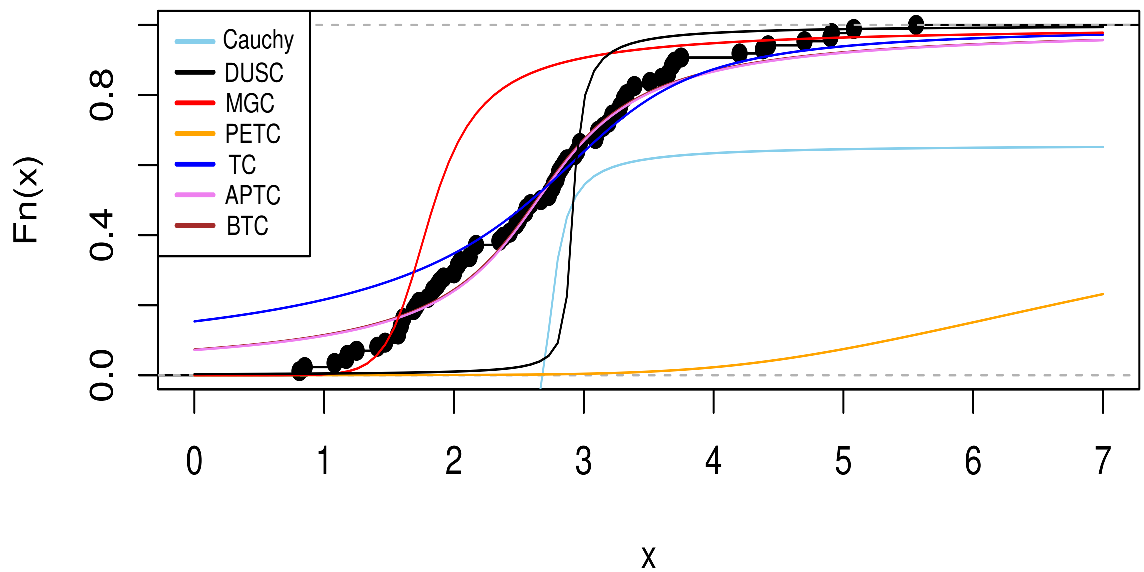


Figure 5.6: Empirical and the fitted cdfs for the data set

equation.

The applicability of the model is validated by using a real data set. From Table 5.4, we observed a better performance of our distribution than the other competing models. Based on these, the new suggested model can be considered as a more efficient, flexible and therefore may be an alternative to other distributions for modeling real data sets. The model parameters are estimated by CVM, ML, AD and LS method of estimation. Estimation of the model parameters under the Bayesian Inference is currently underway. We hope that our new model will attract extensive application in various fields.

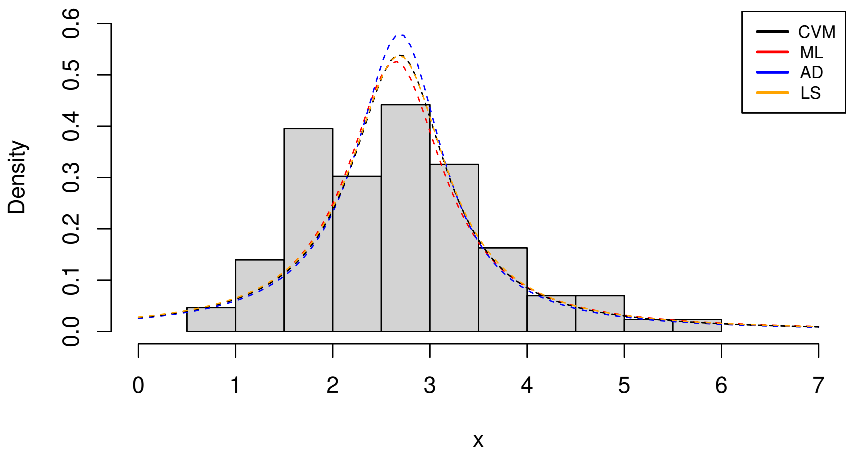


Figure 5.7: Fitted distribution's histogram and the density function using CVM, ML, AD and LS estimation methods for the data set.

Chapter 6

RECOMMENDATIONS

Large number of probability distributions are developed in the literature. However, there still remains various practical problems that could not be modelled using any of the existing probability distributions. This necessitates the introduction of new probability distributions.

In this thesis, we have studied some generalizations of Cauchy distribution namely Transmuted Cauchy distribution, Generalized Cauchy distribution, Discrete Mittag-Leffler Cauchy distribution and Beta transformed Cauchy distribution. These distributions are recommended for data modelling when the data is heavy tailed and could not be modelled using the usual heavytailed distributions such as Laplace distribution, t distribution, logistic distribution, extreme value distribution, etc.

We hope that our proposed models will attract extensive applications in the insurance and financial sciences and would be more helpful for the new comers in the field of distribution theory.

Some future works to be carried out are given below:

- Characterizations of the proposed models are to be addressed.

- Extend the models to multivariate case.
- We have mainly concentrated on non-censored data alone. Future studies should consider the use of censored data in establishing the applications of the developed models.
- Bayesian estimation of the parameters under different types of loss functions based on complete and censored samples will be addressed.

REFERENCES

Adamidis, K., Dimitrakopoulou, T., and Loukas, S. (2005). On an extension of the exponential-geometric distribution. *Statistics and Probability Letters*, 73(3):259–269. 10.1016/j.spl.2005.03.013.

Adamidis, K. and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics and Probability Letters*, 39(1):35–42. 10.1016/S0167-7152(98)00012-1.

Adeyinka, F. S. (2019). On the tractability of transmuted type I generalized logistic distribution with application. *International Journal of Theoretical and Applied Mathematics*, 5(2):31–36. 10.11648/j.ijtam.20190502.12.

Afify, A. Z., Cordeiro, G. M., Yousof, H. M., Saboor, A., and Ortega, E. M. M. (2018). The Marshall-Olkin additive Weibull distribution with variable shapes for the hazard rate. *Hacettepe Journal of Mathematics and Statistics*, 47(2):365–381. 10.15672/hjms.201612618532.

Afify, A. Z., Nofal, Z. M., Yousof, H. M., El Gebaly, Y. M., and Butt, N. S. (2015). The transmuted Weibull Lomax distribution: properties and application. *Pakistan Journal of Statistics and Operation Research*, 11(1):135–152. 10.18187/pjsor.v11i1.956.

Ahmad, A., Ahmad, S. P., and Ahmed, A. (2015). Characterization and estimation of transmuted Kumaraswamy distribution. *Mathematical Theory and Modeling*, 5(9):168–174. api.semanticscholar.org/CorpusID:212586124.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723. 10.1109/TAC.1974.1100705.

Alam, F. M. A. (2022). On comparing different methods of estimation for the parameters of a pathological distribution with application to climate data. *Austrian Journal of Statistics*, 51(4):120–147. 10.17713/ajs.v51i4.1331.

Alam, F. M. A. and Nassar, M. (2021). On modeling concrete compressive strength data using Laplace Birnbaum-Saunders distribution assuming contaminated information. *Crystals*, 11(7):830. 10.3390/cryst11070830.

Alice, T. and Jose, K. K. (2003). Marshall-Olkin Pareto processes. *Far East Journal of Theoretical Statistics*, 9(2):117–132.

Alice, T. and Jose, K. K. (2005). Marshall-Olkin logistic processes. *STARS International Journal*, 6(1):1–11.

Alkasasbeh, M. R. and Raqab, M. Z. (2009). Estimation of the generalized logistic distribution parameters: Comparative study. *Statistical Methodology*, 6(3):262–279. 10.1016/j.stamet.2008.10.001.

Almalki, S. J. and Yuan, J. (2013). A new modified Weibull distribution. *Reliability Engineering and System Safety*, 111:164–170. 10.1016/j.ress.2012.10.018.

Alshawarbeh, E., Famoye, F., and Lee, C. (2013). Beta-Cauchy distribution: Some properties and applications. *Journal of Statistical Theory and Applications*, 12(4):378–391. 10.2991/jsta.2013.12.4.5.

Alzaatreh, A., Lee, C., Famoye, F., and Ghosh, I. (2016). The generalized Cauchy family of distributions with applications. *Journal of Statistical Distributions and Applications*, 3(1):1–16. 10.1186/s40488-016-0050-3.

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2):193–212. jstor.org/stable/2236446.

Arnold, B. C. and Beaver, R. J. (2000). The skew-Cauchy distribution. *Statistics and Probability Letters*, 49(3):285–290. 10.1016/S0167-7152(00)00059-6.

Aryal, G. R. (2013). Transmuted log-logistic distribution. *Journal of Statistics Applications and Probability*, 2(1):11–20. 10.12785/jsap/020102.

Aryal, G. R. and Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods and Applications*, 71(12):e1401–e1407. 10.1016/j.na.2009.01.168.

Aryal, G. R. and Tsokos, C. P. (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 4(2):89–102. ejpam.com/index.php/ejpam/article/view/1170.

Ashour, S. K. and Eltehiwy, M. A. (2013). Transmuted Lomax distribution. *American Journal of Applied Mathematics and Statistics*, 1(6):121–127. 10.12691/ajams-1-6-3.

Athreya, K. B., Ney, P. E., and Ney, P. (2004). *Branching Processes*. Courier Corporation.

Babu, M. G. (2016). On a generalization of Weibull distribution and its applications. *International Journal of Statistics and Applications*, 6(3):168–176. 10.5923/j.statistics.20160603.10.

Bakar, S. A. A., Hamzah, N. A., Maghsoudi, M., and Nadarajah, S. (2015). Modeling loss data using composite models. *Insurance: Mathematics and Economics*, 61:146–154. 10.1016/j.insmatheco.2014.08.008.

Balakrishnan, N. and Alam, F. M. A. (2019). Maximum likelihood estimation of the parameters of students t Birnbaum-Saunders distribution: a comparative study. *Communications in Statistics-Simulation and Computation*, 51(3):793–822. 10.1080/03610918.2019.1659359.

Ball, C., Rimal, B., and Chhetri, S. (2021). A new generalized Cauchy distribution with an application to annual one day maximum rainfall data. *Statistics, Optimization and Information Computing*, 9(1):123–136. 10.19139/soic-2310-5070-1000.

Balmer, D. W., Boulton, M., and Sack, R. A. (1974). Optimal solutions in parameter estimation problems for the Cauchy distribution. *Journal of the American Statistical Association*, 69(345):238–242. 10.2307/2285534.

Barnett, V. D. (1966). Order statistics estimators of the location of the Cauchy distribution. *Journal of the American Statistical Association*, 61(316):1205–1218. 10.2307/2283210.

Barnett, V. D. (1968). Correction: Order statistics estimators of the location of the Cauchy distribution. *Journal of the American Statistical Association*, 63(321):383–385. 10.2307/2283869.

Batschelet, E. (1981). Circular Statistics in Biology. *Academic Press, 111 Fifth Ave., New York*.

Behboodian, J., Jamalizadeh, A., and Balakrishnan, N. (2006). A new class of skew-Cauchy distributions. *Statistics and Probability Letters*, 76(14):1488–1493. 10.1016/j.spl.2006.03.008.

Bennett, S. (1983). Analysis of survival data by the proportional odds model. *Statistics in Medicine*, 2(2):273–277. 10.1002/sim.4780020223.

Bidram, H. (2012). The beta exponential-geometric distribution. *Communications in Statistics - Simulation and Computation*, 41(9):1606–1622. 10.1080/03610918.2011.611309.

Bidram, H., Alamatsaz, M. H., and Nekoukhou, V. (2015). On an extension of the exponentiated Weibull distribution. *Communications in Statistics-Simulation and Computation*, 44(6):1389–1404. 10.1080/03610918.2013.819918.

Bienayme, J. (1853). Remarques sur les differences qui distinguent l'interpolation de M. Cauchy de la methode des moindres carres, et qui assurent la superiorite de cette methode. *Journal de Mathematiques Pures et Appliquees*, 18:299–308.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987). *Regular variation.* Cambridge university press, New York.

Bloch, D. (1966). A note on the estimation of the location parameter of the Cauchy distribution. *Journal of the American Statistical Association*, 61(315):852–855. 10.2307/2282794.

Blom, G. (1958). *Statistical Estimates and Transformed Beta-Variables.* John Wiley and Sons, New York.

Borovkov, A. (2012). *Stochastic Processes in Queueing Theory*, volume 4. Springer Science and Business Media.

Cauchy, A. L. (1853). Sur les resultats moyens dobservations de meme nature, et sur les resultats les plus probables. *Comptes Rendus de l'Academie des sciences dey Paris*, 37:198–206.

Chan, L. K. (1970). Linear estimation of the location and scale parameters of the Cauchy distribution based on sample quantiles. *Journal of the American Statistical Association*, 65(330):851–859. 10.2307/2284592.

Chandra, S. (1977). On the mixtures of probability distributions. *Scandinavian Journal of Statistics*, 4(3):105–112. jstor.org/stable/4615660.

Chen, G. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, 27(2):154–161. 10.1080/00224065.1995.11979578.

Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(3):394–403. 10.1111/j.2517-6161.1983.tb01268.x.

Chistyakov, V. P. (1964). A theorem on sums of independent positive random variables and its applications to branching random processes. *Theory of Probability and Its Applications*, 9(4):640–648. 10.1137/1109088.

Chover, J., Ney, P., and Wainger, S. (1973). Functions of probability measures. *Journal danalyse mathematique*, 26(1):255–302. 10.1007/BF02790433.

Cooray, K. and Ananda, M. M. A. (2005). Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal*, 2005(5):321–334. 10.1080/03461230510009763.

Cordeiro, G. M. and Lemonte, A. J. (2013). On the Marshall-Olkin extended Weibull distribution. *Statistical Papers*, 54(2):333–353. 10.1007/s00362-012-0431-8.

Dahiya, R. C., Staneski, P. G., and Chaganty, N. R. (2001). Maximum likelihood estimation of parameters of the truncated Cauchy distribution. *Communications in Statistics - Theory and Methods*, 30(8-9):1737–1750. 10.1081/STA-100105695.

Demicheli, R., Bonadonna, G., Hrushesky, W. J. M., Retsky, M. W., and Valagussa, P. (2004). Menopausal status dependence of the timing of breast cancer recurrence after surgical removal of the primary tumour. *Breast Cancer Research*, 6(6):R689–R696. 10.1186

Devroye, L. (1993). A triptych of discrete distributions related to the stable law. *Statistics and Probability Letters*, 18(5):349–351. 10.1016/0167-7152(93)90027-G.

Dey, S., Raheem, E., and Mukherjee, S. (2017). Statistical properties and different methods of estimation of transmuted Rayleigh distribution. *Revista Colombiana de Estadistica*, 40(1):165–203. 10.15446/rce.v40n1.56153.

Do Espirito Santo, A. P. J. and Mazucheli, J. (2015). Comparison of estimation methods for the Marshall- Olkin extended Lindley distribution. *Journal of Statistical Computation and Simulation*, 85(17):3437–3450. 10.1080/00949655.2014.977904.

Ebner, B., Eid, L., and Klar, B. (2022). Cauchy or not Cauchy? new goodness-of-fit tests for the Cauchy distribution. *Statistical Papers*, 65:45–78. 10.1007/s00362-022-01382-0.

Ebrahimi, N. (2000). The maximum entropy method for lifetime distributions. *Sankhya: The Indian Journal of Statistics, Series A*, 62(2):236–243. jstor.org/stable/25051309.

Elbatal, I. (2013). Transmuted modified inverse Weibull distribution: A generalization of the modified inverse Weibull probability distribution. *International Journal of Mathematical Archive*, 4(8):117–129.

Elbatal, I. and Aryal, G. (2013). On the transmuted additive Weibull distribution. *Austrian Journal of Statistics*, 42(2):117–132. 10.17713/ajs.v42i2.160.

Eling, M. (2012). Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models? *Insurance: Mathematics and Economics*, 51(2):239–248. 10.1016/j.insmatheco.2012.04.001.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events.* Springer-Verlag, Berlin. 10.1007/978-3-642-33483-2.

Erdelyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F. G. (1955). *Higher Transcendental Functions, volume 3.* McGraw-Hill, New York.

Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, 31(4):497–512. 10.1081/STA-120003130.

Fasna, K. (2022). A method for generating lifetime models and its application to real data. *Reliability: Theory and Applications*, 17(4 (71)):344–357. 10.24412/1932-2321-2022-471-344-357.

Fasna, K. (2023). Discrete Mittag- Leffler Cauchy distribution: Estimation and applications. *Journal of Scientific Research*, 67:110–118. 10.37398/JSR.2023.670213.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, Volume II.* Wiley, Newyork.

Ferguson, T. S. (1978). Maximum likelihood estimates of the parameters of the Cauchy distribution for samples of size 3 and 4. *Journal of the American Statistical Association*, 73(361):211–213. 10.2307/2286549.

Foss, S., Korshunov, D., and Zachary, S. (1970). *An Introduction to Heavy-Tailed and Subexponential Distributions*. Springer. 10.1007/978-1-4614-7101-1.

Franck, W. E. (1981). The most powerful invariant test of normal versus Cauchy with applications to stable alternatives. *Journal of the American Statistical Association*, 76(376):1002–1005. 10.1080/01621459.1981.10477755.

Ghitany, M. E. (2005). Marshall-Olkin extended Pareto distribution and its application. *International Journal of Applied Mathematics*, 18(1):17–32.

Ghitany, M. E., Al-Awadhi, F. A., and Alkhalfan, L. A. (2007). Marshall-Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics - Theory and Methods*, 36(10):1855–1866. 10.1080/03610920601126571.

Ghosh, J. K., Sinha, B. K., and Joshi, S. N. (1982). Expansions for posterior probability and integrated Bayes risk. *Statistical Decision Theory and Related Topics III,Academic Press*, 1:403–456.

Glanzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families. In *Mathematical Statistics and Probability Theory*, volume B, pages 75–84. Springer.

Goldie, C. M. and Kluppelberg, C. (1998). *Subexponential Distributions. A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhauser Boston.

Gupta, R. D. and Debasis, K. (2001). Generalized exponential distribution: Different method of estimations. *Journal of Statistical Computation and Simulation*, 69(4):315–337. 10.1080/00949650108812098.

Gupta, R. G., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by

Lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4):887–904. 10.1080/03610929808832134.

Haas, G., Bain, L., and Antle, C. (1970). Inferences for the Cauchy distribution based on maximum likelihood estimators. *Biometrika*, 57(2):403–408. jstor.org/stable/2334849.

Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2):190–195. jstor.org/stable/2985032.

Hecht, E. (2002). *Optics*. Addison-Wesley. books.google.co.in/books?id=T3ofAQAAMAAJ.

Hibatullah, R., Widyaningsih, Y., and Abdullah, S. (2018). Marshall-Olkin extended inverse power Lindley distribution with applications. *AIP Conference Proceedings*, 2021(1). 10.1063/1.5062789.

Howlader, H. A. and Weiss, G. M. T. (1985). *A Survey of Small-Sample Estimation of Cauchy Location and Scale*. Technical Report 1, Department of Statistics, University of Winnipeg, Winnipeg, Manitoba.

Huang, W. J. and Chen, Y. H. (2007). Generalized skew-Cauchy distribution. *Statistics and Probability Letters*, 77(11):1137–1147. 10.1016/j.spl.2007.02.006.

Hurvich, C. M. and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307. 10.1093/biomet/76.2.297.

Ibragimov, I. (1975). *Independent and Stationary Sequences of Random Variables*. Groningen.

Jayakumar, K. and Fasna, K. (2022). On a new generalization of Cauchy distribution. *Asian Journal of Statistical Sciences*, 2:61–81.

Jayakumar, K., Krishnan, B., and Hamedani, G. G. (2020). On a new generalization of Pareto distribution and its applications. *Communications in Statistics-Simulation and Computation*, 49(5):1264–1284. 10.1080/03610918.2018.1494281.

Jayakumar, K. and Mathew, T. (2008). On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution. *Statistical Papers*, 49(3):421–439. 10.1007/s00362-006-0024-5.

Jayakumar, K. and Sankaran, K. K. (2016). On a generalisation of uniform distribution and its properties. *Statistica*, 76(1):83–91. 10.6092/issn.1973-2201/6090.

Jayakumar, K. and Sankaran, K. K. (2017). Generalized exponential truncated negative binomial distribution. *American Journal of Mathematical and Management Sciences*, 36(2):98–111. 10.1080/01966324.2017.1295892.

Jayakumar, K. and Sreenivas, P. C. (2003). On discrete stable laws. *Advances and Applications in Statistics*, 3:255–266.

Jaykumar, K. and Suresh, R. P. (2003). Mittag-Leffler distribution. *Journal of the Indian Society for Probability and Statistics*, 7:51–71. hdl.handle.net/2259/457.

John, S. W. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *The Annals of Mathematical Statistics*, 29(4):1188 – 1197. 10.1214/aoms/1177706450.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions, Volume 2*. John wiley and sons, Inc., New York.

Jose, K. K. (2011). Marshall-Olkin family of distributions and their applications in reliability theory, time series modeling and stress-strength analysis. *Proc. ISI 58th World Statist. Congr Int Stat Inst, 21st-26th August*, 201:3918–3923.

Jose, K. K. and Krishna, E. (2011). Marshall-Olkin extended uniform distribution. *ProbStat Forum*, 4:78–88.

Jose, K. K., Naik, S. R., and Ristic, M. M. (2010). Marshall-Olkin q-Weibull distribution and maxmin processes. *Statistical Papers*, 51(1):837–851. 10.1007/s00362-008-0173-9.

Jose, K. K. and Remya, S. (2015). Negative binomial Marshall-Olkin Rayleigh distribution and its applications. *Economic Quality Control*, 30(2):89–98. 10.1515/eqc-2015-0009.

Kagan, Y. Y. (1992). Correlations of earthquake focal mechanisms. *Geophysical Journal International*, 110(2):305–320. 10.111/j.1365-246X.1992.tb00876.x.

Kapur, J. N. (1989). *Maximum-Entropy Models in Science and Engineering*. John Wiley and Sons.

Khan, A., Hussain, Z., Khalil, A., and Tahir, M. (2016). On the posterior distribution of location parameter of log-normal distribution. *Pakistan Journal of Statistics*, 32(6):497–515.

Khan, M. S. and King, R. (2012). Transmuted generalized inverse Weibull distribution. *Journal of Applied Statistical Science*, 20(3):213–230.

Khan, M. S. and King, R. (2013). Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. *European Journal of Pure and Applied Mathematics*, 6(1):66–88. ejpam.com/index.php/ejpam/article/view/1606.

Khan, M. S., King, R., and Hudson, I. L. (2016). Transmuted Kumaraswamy distribution. *Statistics in Transition. New Series*, 17(2):183–210.

Kluppelberg, C. (1988). Subexponential distributions and integrated tails. *Journal of Applied Probability*, 25(1):132–141. 10.2307/3214240.

Kolmogorov, A. (1933). Sulla determinazione empirica di una lgge di distribuzione. *Giornale dellIstituto Italiano degli Attuari*, 4:83–91.

Koutrouvelis, I. A. (1982). Estimation of location and scale in Cauchy distributions using the empirical characteristic function. *Biometrika*, 69(1):205–213. 10.1093/biomet/69.1.205.

Kozubowski, T. J. and Podgorski, K. (2016). Transmuted distributions and random extrema. *Statistics and Probability Letters*, 116:6–8. 10.1016/j.spl.2016.04.001.

Krishna, E., Jose, K. K., and Ristic, M. M. (2013). Applications of Marshall-Olkin Frechet distribution. *Communications in Statistics-Simulation and Computation*, 42(1):76–89. 10.1080/03610918.2011.633196.

Kumar, D., Kumar, P., Kumar, P., Singh, S. K., and Singh, U. (2021). Pcm transformation: properties and their estimation. *Journal of Reliability and Statistical Studies*, 14(2):373–392. 10.13052/jrss0974-8024.1421.

Kumar, D., Singh, S. K., and Singh, U. (2017). Life time distributions: Derived from some minimum guarantee distribution. *Sohag Journal of Mathematics*, 4(1):7–11. 10.18576/sjm/040102.

Kumar, D., Singh, U., and Singh, S. K. (2015a). A method of proposing new distribution and its application to bladder cancer patients data. *Journal of Statistics Applications and Probability Letters*, 2(3):235–245. 10.12785/jsapl/020306.

Kumar, D., Singh, U., and Singh, S. K. (2015b). A new distribution using sine function-its application to bladder cancer patients data. *Journal of Statistics Applications and Probability*, 4(3):417–427. digitalcommons.aaru.edu.jo/jsap/vol4/iss3/9.

Kyurkchiev, N. (2017). A new transmuted cumulative distribution function based on the Verhulst logistic function with application in population dynamics. *Biomath Communications*, 4(1). 10.11145/bmc.2017.05.132.

Lai, C. D. and Xie, M. (2006). *Stochastic Ageing and Dependence for Reliability*. Springer NewYork. 10.1007/0-387-34232-X.

Lin, G. D. (1998). On the Mittag-Leffler distributions. *Journal of Statistical Planning and Inference*, 74(1):1–9. 10.1016/S0378-3758(98)00096-2.

Lindley, D. V. (1980). Approximate Bayesian methods. *Trabajos de estadistica y de investigacion operativa*, 31(1):223–245. 10.1007/BF02888353.

Lone, M. A. and Jan, T. R. (2023). A new pi-exponentiated method for constructing distributions with an application to Weibull distribution. *Reliability: Theory and Applications*, 18(1 (72)):94–109. 10.24412/1932-2321-2023-172-94-109.

Macdonald, P. D. M. (1971). Comments and queries comment on an estimation procedure for mixtures of distributions by choi and bulgren. *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(2):326–329. jstor.org/stable/2985013.

Mahdavi, A. and Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics - Theory and Methods*, 46(13):6543–6557. 10.1080/03610926.2015.1130839.

Mahdizadeh, M. and Zamanzade, E. (2017). New goodness of fit tests for the Cauchy distribution. *Journal of Applied Statistics*, 44(6):1106–1121. 10.1080/02664763.2016.1193726.

Mahdizadeh, M. and Zamanzade, E. (2019). Goodness-of-fit testing for the Cauchy distribution with application to financial modeling. *Journal of King Saud University - Science*, 31(4):1167–1174. 10.1016/j.jksus.2019.01.015.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3):641–652. jstor.org/stable/2337585.

Maurya, S. K., Kaushik, A., Singh, R. K., Singh, S. K., and Singh, U. (2016). A new method of proposing distribution and its application to real data. *Imperial Journal of Interdisciplinary Research*, 2(6):1331–1338.

Maurya, S. K., Kaushik, A., Singh, S. K., and Singh, U. (2017). A new class of distribution having decreasing, increasing, and bathtub-shaped failure rate. *Communications in Statistics-Theory and Methods*, 46(20):10359–10372. 10.1080/03610926.2016.1235196.

Mazucheli, J., Louzada, F., and Ghitany, M. E. (2013). Comparison of estimation methods for the parameters of the weighted Lindley distribution. *Applied Mathematics and Computation*, 220:463–471. 10.1016/j.amc.2013.05.082.

Merovci, F. (2013a). Transmuted Lindley distribution. *International Journal of Open Problems in Computer Science and Mathematics*, 6(2):63–72. 10.12816/0006170.

Merovci, F. (2013b). Transmuted Rayleigh distribution. *Austrian Journal of Statistics*, 42(1):21–31. 10.17713/ajs.v42i1.163.

Merovci, F. (2014). Transmuted generalized Rayleigh distribution. *Journal of Statistics Applications and Probability*, 3(1):9–20. 10.12785/jsap/030102.

Min, I. A., Mezic, I., and Leonard, A. (1996). Levy stable distributions for velocity and velocity difference in systems of vortex elements. *Physics of Fluids*, 8(5):1169–1180. 10.1063/1.868908.

Nadarajah, S. and Bakar, S. A. A. (2014). New composite models for the

Danish fire insurance data. *Scandinavian Actuarial Journal*, 2014(2):180–187. 10.1080/03461238.2012.695748.

Nadarajah, S., Jayakumar, K., and Ristic, M. M. (2013). A new family of lifetime models. *Journal of Statistical Computation and Simulation*, 83(8):1389–1404. 10.1080/00949655.2012.660488.

Nadarajah, S. and Kotz, S. (2006). A truncated Cauchy distribution. *International Journal of Mathematical Education in Science and Technology*, 37(5):605–608. 10.1080/00207390600595223.

Nadarajah, S., Popovic, B. V., and Ristic, M. M. (2013). Compounding: an R package for computing continuous distributions obtained by compounding a continuous and a discrete distribution. *Computational Statistics*, 28:977–992. 10.1007/s00180-012-0336-y.

Nadarajah, S., Shahsanaei, F., and Rezaei, S. (2014). A new four-parameter lifetime distribution. *Journal of Statistical Computation and Simulation*, 84(2):248–263. 10.1080/00949655.2012.705284.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351. 10.1080/00107510500052444.

Nichols, M. D. and Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, 22(2):141–151. 10.1002/qre.691.

Omey, E., Van Gulck, S., and Vesilo, R. (2018). Semi-heavy tails. *Lithuanian Mathematical Journal*, 58(1):480–499. 10.1007/s10986-018-9417-0.

Pakes, A. G. (1975). On the tails of waiting-time distributions. *Journal of Applied Probability*, 12(3):555–564. 10.2307/3212870.

Pakungwati, R. M., Widyaningsih, Y., and Lestari, D. (2018). Marshall-Olkin extended inverse Weibull distribution and its application. In *Journal of Physics: Conference Series*, volume 1108. IOP Publishing. 10.1088/1742-6596/1108/1/012114.

Pekasiewicz, D. (2014). Application of quantile methods to estimation of Cauchy distribution parameters. *Statistics in Transition. New Series*, 15(1):133–144.

Pillai, R. N. (1990). On Mittag-Leffler functions and related distributions. *Annals of the Institute of statistical Mathematics*, 42(1):157–161. 10.1007/BF00050786.

Pillai, R. N. and Jayakumar, K. (1995). Discrete Mittag-Leffler distributions. *Statistics and Probability Letters*, 23(3):271–274.

Poisson, S. D. (1827). Sur la probabilite des resultats moyens des observations. *Connaissance des tems*, 1824:273–302.

Popovic, B. V., Ristic, M. M., and Cordeiro, G. M. (2015). A two-parameter distribution obtained by compounding the generalized exponential and exponential distributions. *Mediterranean Journal of Mathematics*, 13(5):2935–2949. 10.1007/s00009-015-0665-5.

Qoshja, A. and Hoxha, F. (2016). Different methods of estimation for generalized inverse Lindley distribution. *Mathematical Theory and Modeling*, 6(5):132–141.

Rachev, S. T. (2003). *Handbook of Heavy Tailed Distributions in Finance*. Elsevier, Amsterdam.

Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag.

Resnick, S. I. (2007). *Heavy-tail Phenomena: Probabilistic and Statistical Modeling*. Springer Science and Business Media, New York.

Rider, P. R. (1957). Generalized Cauchy distributions. *Annals of the Institute of Statistical mathematics*, 9(1):215–223. 10.1007/BF02892507.

Ristic, M. M., Jose, K. K., and Ancy, J. (2007). A Marshall-Olkin gamma distribution and minification process. *Stars: Stress Anxiety Research Society*, 11:107–117.

Ristic, M. M. and Kundu, D. (2015). Marshall-Olkin generalized exponential distribution. *Metron*, 73(3):317–333. 10.1007/s40300-014-0056-x.

Ristic, M. M. and Kundu, D. (2016). Generalized exponential geometric extreme distribution. *Journal of Statistical Theory and Practice*, 10(1):179–201. 10.1080/15598608.2015.1101628.

Roe, B. P. (1992). *Basic Probability Concepts. Probability and Statistics in Experimental Physics*. Springer, New York. 10.1007/978-1-4757-2186-7_1.

Rothenberg, T. J., Fisher, F. M., and Tilanus, C. B. (1964). A note on estimation from a Cauchy sample. *Journal of the American Statistical Association*, 59(306):460–463. 10.2307/2282999.

Sankaran, K. K. and Jayakumar, K. (2016). A new extended uniform distribution. *International Journal of Statistical Distributions and Applications*, 2(3):35–41. 10.11648/j.ijsd.20160203.12.

Sankaran, P. G. and Jayakumar, K. (2008). On proportional odds models. *Statistical Papers*, 49:779–789. 10.1007/s00362-006-0042-3.

Sarhan, A. E. and Greenberg, B. G. (1962). *Contributions to Order Statistics*. John Wiley and Sons, New York.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464. jstor.org/stable/2958889.

Scollnik, D. P. M. (2007). On composite lognormal-Pareto models. *Scandinavian Actuarial Journal*, 2007(1):20–33. 10.1080/03461230601110447.

Seneta, E. (1976). *Functions of Regular Variation. In: Regularly Varying Functions*. Springer. 10.1007/BFb0079659.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer.

Shaw, W. T. and Buckley, I. R. C. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*.

Sim, C. H. (1986). Simulation of weibull and gamma autoregressive stationary process. *Communications in Statistics - Simulation and Computation*, 15(4):1141–1146. 10.1080/03610918608812565.

Souza, W. B., de Morais, A. L., and Cordeiro, G. M. (2011). The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, 81(5):645–657. 10.1080/00949650903436554.

Spiegelhalter, D. J. (1985). Exact Bayesian inference on the parameter of a Cauchy distribution with vague prior information. *Bayesian Statist.*, 2:743–749.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(4):583–639. 10.1111/1467-9868.00353.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 76(3):485–493. jstor.org/stable/24774528.

Stapf, S., Kimmich, R., Seitter, R. O., Maklakov, A. I., and Skirda, V. D. (1996). Proton and deuteron field-cycling NMR relaxometry of liquids confined in porous glasses. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 115:107–114. 10.1016/0927-7757(96)03610-2.

Stigler, S. M. (1974). Studies in the history of Probability and Statistics. XXXIII: Cauchy and the witch of Agnesi: An historical note on the Cauchy distribution. *Biometrika*, 61(2):375–380. 10.1093/biomet/61.2.375.

Swain, J. J., Venkatraman, S., and Wilson, J. R. (1988). Least-squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation*, 29(4):271–297. 10.1080/00949658808811068.

Tahir, M. H., Zubair, M., Cordeiro, G. M., Alzaatreh, A., and Mansoor, M. (2017). The Weibull-power Cauchy distribution: model, properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 46(4):767–789.

Tavares, L. V. (1980). An exponential Markovian stationary process. *Journal of Applied Probability*, 17(4):1117–1120. 10.2307/3213224.

Team, R. D. C. (2009). A language and environment for statistical computing. *R-project. org.*

Teugels, J. L. (1975). The class of subexponential distributions. *The Annals of Probability*, 3(6):1000–1011. jstor.org/stable/2959204.

Wang, M. and Elbatal, I. (2014). The modified Weibull geometric distribution. *METRON*, 73(3):303–315. 10.1007/s40300-014-0052-1.

Weisberg, S. (2005). *Applied Linear Regression*. John Wiley and Sons, New York. 10.1002/0471704091.

Winterton, S. S., Smy, T. J., and Tarr, N. G. (1992). On the source of scatter in contact resistance data. *Journal of Electronic Materials*, 21(9):917–921. 10.1007/BF02665549.

Yegen, D. and Gamze, O. (2018). Marshall-Olkin half logistic distribution with theory and applications. *Alphanumeric Journal*, 6(2):407–416. 10.17093/alphanumeric.409992.

Zellner, A. (1971). Bayesian and Non-Bayesian analysis of the log-normal distribution and log-normal regression. *Journal of the American Statistical Association*, 66(334):327–330. 10.1080/01621459.1971.10482263.