NON-LINEAR TIME SERIES MODELS AND THEIR APPLICATIONS

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

VIDYA V. P.

under the guidance of

Dr. KRISHNARANI S. D.

PG & RESEARCH DEPARTMENT OF STATISTICS
FAROOK COLLEGE (AUTONOMOUS)
KOZHIKODE, KERALA- 673 632

PG & RESEARCH DEPARTMENT OF STATISTICS
FAROOK COLLEGE (AUTONOMOUS)
KOZHIKODE, KERALA-673 632

CERTIFICATE

I hereby certify that the work reported in this thesis entitled **'NON-LINEAR TIME SERIES MODELS AND THEIR APPLICATIONS'** that is being submitted by **Smt. Vidya V. P** for the award of **Doctor of Philosophy,** to the University of Calicut, is based on the bonafide research work carried out by her under my supervision and guidance in the PG & Research Department of Statistics, Farook College (Autonomous), Kozhikode, Kerala. The results embodied in this thesis have not been included in any other thesis submitted previously for the award of any degree or diploma of any other university or institution. 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.

Dr. KRISHNARANI S. D.
(Research Guide)

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT

Dr. Krishnarani S.D.
Associate Professor & Head

CALICUT UNIVERSITY (P. O.)
KERALA, INDIA 673 635.
Phone : 0494 – 2407341
Mob : 9496127836
Email : stathod@uoc.ac.in

……

Date: 03.05.2024

CERTIFICATE

Certified that the corrections/suggestions, recommended by the adjudicators, of the Ph. D. Thesis entitled - **'NON-LINEAR TIME SERIES MODELS AND THEIR APPLICATIONS'** submitted by **Smt. Vidya V. P.,** Research scholar, Department of Statistics, Farook College (Autonomous) under my supervision and guidance, have been incorporated in this copy of the thesis and the contents in the thesis and the soft copy are one and the same.

Dr. Krishnarani S.D.

Research Guide

DECLARATION

I hereby declare that this thesis entitled **'NON-LINEAR TIME SERIES MODELS AND THEIR APPLICATIONS'** submitted to the University of Calicut for the award of the degree of **Doctor of Philosophy** in Statistics under the Faculty of Science is an independent work done by me under the guidance and supervision of **Dr. Krishnarani S. D.,** Associate Professor & Head, Department of Statistics, University of Calicut, Kozhikode, Kerala.

I also declare that this thesis contains no material which has been accepted for the award of any other degree or diploma of any university or institution and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due reference made in the text of thesis.

Place: Farook College VIDYA V. P.

Date: 03.05.2024

ACKNOWLEDGEMENTS

For the fulfillment of the thesis, many people have put in their effort and help. I would like to express my heartfelt gratitude to each and every one of them for their relentless help.

First of all, I would like to express my sincere gratitude to my supervising guide, Dr. Krishnarani S. D., Associate Professor & Head, Department of Statistics, University of Calicut, Kozhikode for her invaluable guidance, support, and encouragement throughout my research. Her expertise and insights have been instrumental in shaping my work.

I wish to express my deep sense of gratitude to the teaching and non-teaching staffs of PG & Research Department of Statistics, Farook College (Autonomous), Kozhikode, Kerala for their helpful feedback and discussions. I express my indebtedness to all my teachers, who provided whole hearted support to me.

Healthy and fruitful interactions with my friends and other research scholars of the department often helped me during the work. I express my sincere thanks to all of them for their valuable suggestions and help.

I express my thanks to the Principal, Teaching and Non-teaching staff of Govt. Rajah's Higher Secondary School, Kottakkal, Malappuram for their whole hearted support and co-operation in completing my research work.

I also indebted to my family and relatives for their support, prayers and blessings.

My sincerest gratitude to the Almighty for granting me the health, strength, wisdom, and determination to bring this work to fruition. Above all, I bow before the grace of the Almighty.

Vidya V. P.

Abstract of the thesis

The thesis is primarily concerned with the construction of non-linear time series models and their applications in real-world data. Non-linear models excel at accommodating non-Gaussian and heavy tailed distributions enabling more precise modelling of extreme events and outliers. So the analysis of financial time series requires non-linear modelling using non-Gaussian distributions. Within the range [0,1], we propose a novel distribution termed the uniform truncated Poisson distribution (UTPD) and investigate its features, parameter estimates, and applicability in real-world scenarios. There is also a comparison with the power function distribution and generalization to this distribution. The non-linear applicability of this distribution is investigated by introducing processes with the UTPD under a variety of assumptions. We build a uniform truncated Poisson autoregressive process of order 1 (UTPAR(1)) with UTPD as the marginal function. Investigates the new process's attributes, estimating methodologies, and real-world application. Another process is the uniform truncated autoregressive conditional duration process (UTPACD(1,1)). We talk about analytical characteristics and traditional techniques. Estimation and application are also looked upon. We address the analytical characteristics, traditional estimating methodologies, and real-world applications of the process. This thesis also includes spatial analysis of child mortality data using spatial lag models, spatial Durbin models, and spatial error models that incorporate non-linearity. Minification procedures with two distinct structures are presented, with UTPD acting as a marginal. The processes are known as Type I uniform truncated minification process (Type I UTPM) and Type II uniform truncated minification process (Type II UTPM). The key attributes, estimation methods, and application are also investigated. The relevance of non-linear non-Gaussian time series model is emphasized at the end of this thesis. This thesis concludes by underlining the significance of non-linear non-Gaussian time series models in time series analysis and suggesting future directions.

Key Words: Truncated uniform distribution, Non-linear time series, Spatial auto-correlation, ACD, Minification process.

സംഗ്രഹം

ഈ പ്രബന്ധം പ്രാഥമികമായി അരേഖീയ സമയ ശ്രേണി മോഡലുകളുടെ നിർമ്മാണവുമായും യഥാർത്ഥ ലോക ഡാറ്റയിലെ അവയുടെ പ്രായോഗികതമായും ബന്ധപ്പെട്ടിരിക്കുന്നു. അരേഖീയ സമയശ്രേണി മോഡലുകൾ നോൺ ഗൗസിയൻ, ഹെവിടെയിൽഡ് വിതരണങ്ങൾ എന്നിവ ഉൾക്കൊള്ളുന്നതിൽ മികവ് പുലർത്തുന്നു. അത് അങ്ങേയറ്റത്തെ സംഭവങ്ങളുടെയും ഔട്ട് ലെയറുകളുടെയും കൂടുതൽ കൃത്യമായ മോഡലിംഗ് സാധ്യമാക്കുന്നു. അതിനാൽ സാമ്പത്തിക സമയ ശ്രേണിയുടെ വിശകലനത്തിന് നോൺ ഗൗസിയൻ വിതരണങ്ങൾ ഉപയോഗിച്ച് അരേഖീയ മോഡലിംഗ് ആവശ്യമാണ്. യൂണിഫോം ട്രങ്കേറ്റഡ് പോയിസൺ വിതരണം (യു ടി പി ഡി) എന്ന് വിളിക്കുന്ന ഒരു പുതിയ വിതരണം [0,1] എന്ന പരിധിക്കുള്ളിൽ ഞങ്ങൾ മുന്നോട്ട് വെക്കുകയും അതിൻറെ സവിശേഷതകൾ, പരാമീറ്ററുകളുടെ മതിപ്പ്, യഥാർത്ഥ ലോക സാഹചര്യങ്ങളിലെ പ്രായോഗികക്ഷമത എന്നിവ അന്വേഷിക്കുകയും ചെയ്യുന്നു. ഈ വിതരണത്തിന് പവർ ഫംഗ്ഷൻ വിതരണവുമായുള്ള താരതമ്യവും ഇതിന്റെ സാമാന്യവത്കരണവും നടത്തുന്നുണ്ട്. ഈ വിതരണത്തിന്റെ അരേഖീയ പ്രായോഗികക്ഷമത വിവിധ അനുമാനങ്ങളിൽ യു ടി പി ഡി യുമായി പ്രക്രിയകൾ അവതരിപ്പിച്ചുകൊണ്ട് അന്വേഷിക്കുന്നു. യു ടി പി ഡി നാമമാത്ര വിതരണമായി ഓർഡർ 1 ആയ യൂണിഫോം ട്രങ്കേറ്റഡ് പോയിസൺ ഓട്ടോറിഗ്രസ്സീവ് പ്രക്രിയ (യു ടി പി എ ആർ (1)) നിർമ്മിക്കുന്നു. പുതിയ പ്രക്രിയയുടെ സവിശേഷതകൾ, എസ്റ്റിമേറ്റിങ്ങ് രീതിശാസ്ത്രങ്ങൾ, യഥാർത്ഥ ലോകത്തിലുള്ള പ്രായോഗികത എന്നിവയെ കുറിച്ച് അന്വേഷിക്കുന്നു. യു ടി പി ഡിയുമായി നാമമാത്ര വിതരണമായി നിർമിക്കുന്ന മറ്റൊരു പ്രക്രിയ ഓർഡർ (1,1) ആയ യൂണിഫോം ട്രങ്കേറ്റഡ് ഓട്ടോറിഗ്രസ്സീവ് കണ്ടീഷണൽ ഡ്യൂറേഷൻ (യു ടി പി എ സി ഡി (1,1)) പ്രക്രിയയാണ്. പുതിയ പ്രക്രിയയുടെ വിശകലന സവിശേഷതകൾ, പരമ്പരാഗത എസ്റ്റിമേറ്റിംഗ് രീതികൾ, യഥാർത്ഥ ലോകത്തിലെ പ്രായോഗികതകൾ, എന്നിവയിലും ശ്രദ്ധ ചെലുത്തിയിട്ടുണ്ട്. സ്പേഷ്യൽ ലാഗ് മോഡലുകൾ, സ്പേഷ്യൽ ഡർബിൻ മോഡലുകൾ, സ്പേഷ്യൽ എറർ മോഡലുകൾ എന്നിവയിൽ അരേഖീയത ഉൾപ്പെടുത്തി ശിശു മരണനിരക്ക് ഡാറ്റയുടെ സ്ഥാനിക വിശകലനവും ഈ പ്രബന്ധത്തിൽ ഉൾപ്പെടുന്നു. യു ടി പി ഡി ഒരു നാമമാത്ര വിതരണമായി രണ്ടു വ്യത്യസ്ത ഘടനകൾ ഉള്ള മിനിഫിക്കേഷൻ പ്രക്രിയകൾ അവതരിപ്പിച്ചിരിക്കുന്നു. ടൈപ്പ് I യൂണിഫോം ട്രങ്കേറ്റഡ് പോയിസൺ മിനിഫിക്കേഷൻ (ടൈപ്പ് I യു ടി പി എം) പ്രക്രിയ എന്നും ടൈപ് II യൂണിഫോം ട്രങ്കേറ്റഡ് മിനിഫിക്കേഷൻ (ടൈപ്പ് II യു ടി പി എം) പ്രക്രിയ എന്നും ഈ പ്രക്രിയകൾ അറിയപ്പെടുന്നു. പ്രധാന സ്വഭാവ സവിശേഷതകൾ, എസ്റ്റിമേറ്റ് രീതികൾ, പ്രായോഗികതകൾ എന്നിവയും അന്വേഷിക്കുന്നു. സമയ ശ്രേണി വിശകലനത്തിൽ അരേഖീയ നോൺ ഗൗസിയൻ മോഡലുകളുടെ പ്രാധാന്യം അടിവരയിടുകയും ഭാവി ഗവേഷണ ദിശകൾ നിർദ്ദേശിക്കുകയും ചെയ്തുകൊണ്ട് പ്രബന്ധം ഉപസംഹരിച്ചിരിക്കുന്നു.

പ്രധാന പദങ്ങൾ : ട്രങ്കേറ്റഡ് യൂണിഫോം വിതരണം, അരേഖീയ സമയ ശ്രേണി, സ്പേഷ്യൽ ഓട്ടോ-കോറിലേഷൻ, എ സി ഡി, മിനിഫിക്കേഷൻ പ്രക്രിയ.

NON-LINEAR TIME SERIES MODELS AND THEIR APPLICATIONS

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

The data obtained from successive observations over a long span of time are very common. Monthly interest rates, annual sales data, hourly wind speeds, crop production etc. are some examples of time series data. Main area of focus in recent research on time series analysis has been on exploring and studying the assumption that the structure of the series can be effectively described by linear time series models. However, there are instances in which the subject, theory, or facts indicate that the linear models are not always reliable. It is preferable to consider non-linear alternatives in those circumstances. Most often, linear time series models use Gaussian distributions for errors or residuals, which may not accurately reflect many sets of

data from the real world. In contrast, non-linear models excel at accommodating non-Gaussian and heavy tailed distributions enabling more precise modelling of extreme events and outliers. So the analysis of financial time series requires non-linear modelling using non-Gaussian distributions. Non-linear time series models exhibit great significance across multiple fields since they capture intricate patterns and dynamics that linear models fail to capture.

Non-linear features such as cycles, asymmetries, bursts, jumps, chaos, thresholds, heteroscedasticity etc. were exhibited by many of the time series encountered in practice. Simple time series models typically do not provide comprehensive explanations for the different aspects of economic and financial data. The economic and financial systems experience both structural and behavioral changes over time, so it is logical to consider the necessity of employing distinct time series models to elucidate the empirical data at various points in time. Non-linear time series models are essential for analyzing and capturing the dynamics of these types of changes in economic and financial data.

One of the objectives of the study is to construct non-linear autoregressive models with different marginal distributions. This prompted us to build several non-linear time series models connected to the newly created distribution and to investigate classical estimation techniques for real-world data applications. In the following sections, we present a compilation of fundamental concepts that aid in the systematic progression of the thesis.

1.2 Stochastic process

A collection of random variables $\{X_t, t \geq 0\}$ defined over the sample space is called a stochastic process. In this sense, a stochastic process is a collection $\{X_t, t \in T\}$ of random variables X_t, where T is some indexed set, usually an interval of real numbers. For the stochastic process $\{X_t, t \geq 0\}$, $\omega : X_t(\omega)$ is a function of T, namely sample function or realisation of the process. It is usually represented by $X(t)$ or X_t. The mean function of the process is defined as $\mu_t = E(X_t)$. The variance function of the process is, $\sigma_t^2 = Var(X_t) = E(X_t - \mu_t)^2$.

To draw inferences about the structure of a stochastic process, on the basis of an observed record of that process, some assumptions have to be made regarding the structure. The most important assumption discussed below is stationarity.

1.2.1 Strong stationarity

The fundamental concept of stationarity is that the probability laws that govern the behavior of the process do not change over time. In other words the process is in statistical equilibrium. A process $\{X_t\}$ is said to be strictly stationary if the joint distribution of $X_{t_1}, X_{t_2}, ..., X_{t_n}$ is identical with the joint distribution of $X_{t_{1-k}}, X_{t_{2-k}}, ..., X_{t_{n-k}}$ that of for all choices of time points $t_1, t_2, ..., t_n$ and for all choices of time lag k. If a process with strict stationarity conditions has a finite variance, then the covariance function must depend only on time lag.

1.2.2 Weak stationarity

A definition that bears resemblance to the definition of strict stationarity, but is mathematically weaker is weak stationarity. A stochastic process $\{X_t\}$ is said to be weakly (or second-order) stationary if

1. The mean function is constant over time.

2. The covariance function $\gamma_{t,t-k} = \gamma_{0,k}$ for all time t and lag k

1.3 Time series

A time series is a series of data points that are typically measured over time. It can be expressed mathematically as a set of vectors $X(t), t = 0, 1, 2, ...$, where t denotes the time elapsed. The variable $X(t)$ is deemed to be random. In a time series, measurements gathered during an event are grouped in chronological order. A univariate time series is one that contains just records for a single variable. However, if records from more than one variable are included, it is referred to as multivariate. A time series might be discrete or continuous. Observations in a continuous time series are measured at every instance of time, whereas observations in a discrete time series are measured at discrete time points. Hourly temperature readings, monthly rainfall data etc. can be recorded as a continuous time series. On the other hand, discrete time series can represent the population of a certain region, a firm's output, exchange rates between two distinct currencies, and so on. Typically, successive observations

in a discrete time series are recorded at equally spaced time intervals such as hourly, daily, weekly, monthly, or yearly time separations. In a discrete time series, the variable being observed is considered to be measured as a continuous variable on a real number scale. Furthermore, by combining data across a predetermined time interval, a continuous time series may be readily changed to a discrete one. Time series analysis entails evaluating and modelling observations in order to obtain the important information contained within the data.

Mathematical models are now employed to describe the behavior of physical phenomena. In particular, we may determine the value of a time-dependent quantity almost or exactly at a given instant in time. A deterministic model is one in which accurate computations are achievable. In many cases, a time-dependent phenomena must be addressed. As a result, we can build a model that may be used to determine the likelihood of a future value falling within two given boundaries. These models are known as stochastic models or probability models. It is critical to distinguish between the probability model, commonly referred to as a stochastic process, and the observed time series. An observed time series $(x_1, x_2, ..., x_n)$ can be interpreted as a specific realization function from a given stochastic process. We may examine and simulate the data using various statistical and probabilistic approaches by seeing the time series as a representation of a stochastic process. We may make conclusions, discern patterns, identify trends, and estimate future values based on the stochastic process's underlying probabilistic behavior. As a result, the series $(X_t, t = 1, 2, ..., n)$ may be seen as a realization of a particular stochastic process. Time series analysis

has several objectives, subject to its application background. Statisticians consider a time series as an outcome derived from a stochastic process as mentioned earlier. The crucial aim is to reveal the probability distribution that governs the observed time series. By discerning this probability law, we gain insights into the underlying dynamics, enabling us to forecast future events and exercise control over them through interventions. These constitute the three primary goals of time series analysis. When dealing with finite observations, multiple stochastic processes can produce the same observed data, leading to an infinite number of possibilities. However, certain processes are more reasonable and offer better interpretations compared to others. Without additional constraints, identification of the underlying process from a finite set of observations becomes unattainable. To handle this, a common strategy is to restrict the probability law to a defined family of processes and then choose the most plausible member within that family. The former is called modelling and the latter is called estimation, or generally statistical inference.

Modelling of a time series indicates the process of creating a mathematical or statistical representation that depicts the underlying patterns, relationships, and behaviors present in a sequence of data points ordered over time. The main objective of time series modelling is to understand, describe, and forecast the future values of the series based on its historical behavior. The observations in a time series are always dependent and often exhibit patterns or relationships with their past and future values. This inherent structure of dependence is a fundamental feature of time series data and has remarkable implications for the analysis and modelling of

such data. Better understanding of the nature of dependence among observations in a time series is crucial for making accurate predictions and informed decisions in various fields such as finance, economics, environmental science, and more. This requires the development of stochastic and dynamic models for time series data and the use of such models in important area of application. The models for time series that are needed are really stochastic models. When the probability laws within a family are specified, a model is labeled as a parametric model except for certain finite-dimensional defining parameters. A non-parametric model is distinguished by either defining parameters that belong to a subset of an infinite-dimensional space or when the form of the probability laws is not completely specified. The underpinning of time series analysis lies in the utilization of appropriate statistical modelling. When selecting a probability model for time series analysis, the first step involves identifying essential characteristics from the observed data. Subsequently, a suitable model is chosen, considering these identified features. Once parameters or functions in the model are estimated, the next step is to assess whether the model adequately captures the data and to seek potential enhancements if feasible. Different objectives in the analysis may call for the adoption of distinct models. For instance, a model that fits the data well and allows for a clear interpretation may not necessarily be the best choice for forecasting purposes.

1.3.1 White noise

The white noise process is a very important example of a stationary process. It is defined as a sequence of independent and identically distributed (iid) random variables say $\{\varepsilon_t, t \geq 1\}$ with mean zero and variance σ_ε^2. It is also referred to as innovation or shock at time t. Its covariance function is given by

$$\gamma(k) = \begin{cases} \sigma_\varepsilon^2 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases} \tag{1.3.1}$$

1.3.2 Autocovariance function

If $\{X_t, t\epsilon T\}$ is a process such that $Var(X_t) < \infty$ for each $t\epsilon T$, then the mean function, $\mu_t = E(X_t)$ for $t\epsilon T$. In general, μ_t can be different at each time point t. The autocovariance function (ACVF),

$$\gamma_{t,s} = Cov(X_t, X_s) = E[(X_t - \mu_t)(X_s - \mu_s)] = E(X_t X_s) - \mu_t \mu_s \text{ for } t, s \in T.$$

1.3.3 Autocorrelation function

The autocorrelation function (ACF), $\rho_{t,s} = Corr(X_t, X_s)$ for $t, s\epsilon T$

$$\rho_{t,s} = \frac{Cov(X_t, X_s)}{\sqrt{Var(X_t)Var(X_s)}} \tag{1.3.2}$$

1.3.4 Partial autocorrelation function

The partial autocorrelation function (PACF) of a staionary process $\{X_t\}$ denoted by $\phi_{k,k}$, $k = 1, 2, ...$ is defined by

$$\phi_{1,1} = Corr\,(X_1, X_0) = \rho_1$$

and

$$\phi_{k,k} = Corr(X_k - \hat{X}_k, X_0 - \hat{X}_0),\ k \geq 2,$$

where $\hat{X}_k = l_1 X_{k-1} + l_2 X_{k-2} + ... + l_{k-1} X_{k-1}$ is known as the linear predictor. Here both $\left(X_k, \hat{X}_k\right)$ and $\left(X_0, \hat{X}_0\right)$ are correlated with $\{X_1, X_2, ..., X_{k-1}\}$. Under stationarity, the PACF is defined as the correlation between X_t and X_{t-k} obtained by fixing the effect of $X_{t-1}, ..., X_{t-(k-1)}$.

1.4 Box-Jenkins modelling techniques

The Box-Jenkins methodology is a procedure for time series analysis and forecasting. It was developed by Box and Jenkins (1970). It is mainly based on the use of ARIMA models. The four stages in Box-Jenkins approach are listed below.

1. Identification

2. Estimation

3. Diagnostic Checking

4. Forecasting

1.4.1 Identification of the model

The plots of original time series data, autocorrelation, and partial autocorrelation are considered in the primary stage. Model identification can be performed by observing the behavior of these plots.

1.4.2 Estimation of the parameters

The estimation of the parameters involved in the model is considered in the second step. There are several methods of estimation available in the literature. One may refer to Box et al. (1994). The main approaches related to Box and Jenkins models are non-linear least squares and maximum likelihood estimation.

The least squares estimator (LSE) of the parameters is obtained by minimizing the sum of the squared residuals. For autoregressive models, the LSE leads to the linear ordinary least Squares (OLS) estimator. The maximum likelihood estimator (MLE) maximizes the log likelihood functions corresponding to the model specified. Here, explicit distributional assumptions for the innovations are necessary. There are a few additional methods for estimation, like the method of moments (MM), the generalized method of moments (GMM), the Gaussian estimation method, and the Yule-Walker estimation method.

1.4.3 Diagnostic checking

Diagnostic checking refers to the testing of the model's adequacy by checking the validity of the assumptions imposed on the errors. This can be done by using techniques like overfitting and residual plots and checking that the residuals are approximately uncorrelated. A good time series model should be able to produce residuals that are approximately white noise. There are statistical tests like the Box-Pierce test that can be used to check the above. ACF and PACF plots are also used for checking the significance of autocorrelation and partial auto correlation.

Akaike information criteria (AIC) and Bayesian information criteria (BIC) are two model selection criteria commonly used.

$$AIC = -2logL + 2k \tag{1.4.1}$$

$$BIC = -2L + kln\,(n) \tag{1.4.2}$$

where L, k, and n respectively are the likelihood function, the number of parameters, and the number of samples used for fitting respectively. A model with the lowest AIC or BIC values is considered the model with the best fit.

1.4.4 Forecasting

Forecasting refers to the prediction of future values. One of the primary objectives of time series analysis is the forecasting of the future based on the observed values in

the time series data. The minimum mean square error (MMSE) is a useful method for forecasting time series data. In this method, an optimum value of prediction is obtained by minimizing the mean squared error.

1.5 Models of time series

1.5.1 Linear time series models

A time series $\{X_t\}$ is said to be linear, if it can be represented by

$$X_t = \mu + \sum_{i=0}^{\infty} \varphi_i \varepsilon_{t-i} \tag{1.5.1}$$

where μ is the mean of X_t, $\psi_0 = 1$ and $\{\varepsilon_t\}$ is a white noise sequence. The dynamic structure of X_t is dominated by the coefficients ψ_i, named as ψ_i weights of X_t in the literature. The most widely used category of linear time series models includes autoregressive moving average (ARMA) models, encompassing purely autoregressive (AR) and purely moving average (MA) models as special instances. ARMA models are commonly used to represent linear dynamic patterns, explain relationships among lagged variables, and facilitate linear forecasting. Within this group, autoregressive integrated moving average (ARIMA) models are particularly valuable since they incorporate stationary ARMA processes as a subclass.

1.5.1.1 Autoregressive Processes

Autoregressive processes are regressions on themselves. The p^{th}-order autoregressive process $\{X_t\}$ denoted as AR(p), satisfies the equation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t. \tag{1.5.2}$$

Here, the current value of the time series X_t is represented as the linear combination of the most recent past values of itself. The term ε_t incorporates everything new in the series at time t that is not explained by the past values. Also, ε_t is independent of $X_{t-1}, X_{t-2}, X_{t-3}, ...$ The term autoregressive was popularized by the British statistician George Udny Yule. Yule's work focused on the theoretical development of autoregressive models and their application to time series data.

The AR characteristic polynomial of the process in equation (1.5.2) is given as

$$\phi(x) = 1 - \phi_1(x) - \phi_2\left(x^2\right) - ... - \phi_p\left(x^p\right)$$

and the corresponding AR characteristic equation is

$$1 - \phi_1(x) - \phi_2\left(x^2\right) - ... - \phi_p\left(x^p\right) = 0$$

The resulting AR(p) process is weakly stationary if and only if the p roots of the AR characteristic equation each exceed 1 in absolute value.

For a stationary AR(p) process, the autocorrelation function $\rho_x(k)$ can be found by solving a set of difference equations called Yuler-Walker equations given by

$\rho(k) = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2} + ... + \varphi_p\rho_{k-p}$, for $k \geq 1$.

The ACF of an AR(p) process generally exhibits a decreasing pattern, with correlations becoming statistically insignificant (falling within the confidence bounds) after lag p. This decline in autocorrelation assists in determining the appropriate value of p for the model selection process.

The first-order autoregressive process

Due to its simplicity and capability to capture specific patterns in data, the AR(1) model finds various practical applications and uses. An AR(1) model, assuming the stationarity condition is given by

$$X_t = \phi X_{t-1} + \varepsilon_t. \tag{1.5.3}$$

The stationary condition (weakly) for $\{X_t\}$ is $|\phi| \leq 1$. For a stationary AR(1) process, $E(X_t) = 0$, $Var(X_t) = \frac{\sigma_\varepsilon^2}{1-\phi^2}$ and the autocorrelation function is given by

$$\rho_X(k) = \phi^k, \ k = 0, 1, 2, ...$$

Since $\rho_X(k) = \phi^k$, $k = 0, 1, 2, ...$, the magnitude of the autocorrelation function decreases exponentially as the number of lags, k increases. It can follows that the ACF plot of a weakly stationary AR(1) series decays exponentially in k. If we assume the innovation sequence $\{\varepsilon_t\}$ is iid, then the AR(1) sequence is Markovian.

1.5.1.2 Moving average processes

Moving average models were first considered by Slutsky (1927) and World (1938). A moving average of order q, represented by MA(q) is given by

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - ... - \theta_p\varepsilon_{t-q} \tag{1.5.4}$$

where θ_i's are constant and $\{\varepsilon_t\}$ is a white noise sequence. Here X_t is obtained by applying the weights 1, $-\theta_1$, $-\theta_2$, ..., $-\theta_q$ to the variables ε_t, ε_{t-1}, ε_{t-2}, ..., ε_{t-q}, and $\{X_{t+1}\}$ are obtaining by moving the weights and applying them to ε_{t+1}, ε_{t+2}, ..., ε_{t-q+1}.

From (1.5.4),

$$E\left(X_t\right) = 0; \quad Var(X_t) = \sigma_\varepsilon^2 \sum_{i=1}^{q} \theta_i^2.$$

The ACF is

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + ... + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + ... + \theta_q^2} & \text{for } k = 1, 2, ...q, \\ 0 & \text{for } k > q. \end{cases} \tag{1.5.5}$$

It can be seen that ACF cuts off after lag q.

The first-order moving average process

The moving average process of order 1, represented by MA(1) is given by

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}. \tag{1.5.6}$$

Then, $E(X_t) = 0$ and $Var(X_t) = \frac{\sigma_\varepsilon^2}{(1+\theta^2)}$. The ACF of MA(1) process is

$$\rho(k) = \begin{cases} -\frac{\theta}{1+\theta^2} & \text{if } k = 1, \\ 0 & \text{if } k = 2, 3, ... \end{cases} \tag{1.5.7}$$

1.5.1.3 The mixed autoregressive moving average model

If we assume that the time series is a combination of both autoregressive and moving average components, we can represent the model as an autoregressive moving average (ARMA) model. The ARMA model combines the autoregressive terms, which depend on past values of the series, and the moving average terms, which depend on past error terms. The general form of an ARMA model of order (p,q) (ARMA(p,q)) is as follows:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_p \varepsilon_{t-q} \tag{1.5.8}$$

In the context of the ARMA model, the stationarity of the entire model depends on the autoregressive component being stationary, and the invertibility of the model relies on the moving average component being invertible.

ARMA(1,1) Model

Let p=1, q=1 in (1.5.8), The ARMA(1,1) process is given by

$$X_t = \phi X_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \tag{1.5.9}$$

The ACF is given by,

$$\rho_k = \frac{(1-\theta\phi)(\phi-\theta)}{1-2\theta\phi+\theta^2}. \tag{1.5.10}$$

As the lag k increases, the ACF exhibits exponential decay with a damping factor ϕ. However, the decay begins from an initial value ρ_1, which is also influenced by the parameter θ. The specific shape of the ACF, denoted by ρ_k, can vary depending on the signs of ρ_1 and ϕ.

1.5.1.4 Autoregressive integrated moving average

The autoregressive integrated moving average (ARIMA) model, in the context of time series analysis, extends the ARMA model to incorporate differencing for achieving stationarity in the data. In time series analysis, the AR component of ARIMA implies that the variable of interest is regressed upon its own previous values. The MA component indicates that the regression error is a linear combination of error terms occurring simultaneously and at different time points in the past. The integrated (I) part signifies that the original data values have undergone differencing, which involves computing the differences between consecutive values. This process may have been applied more than once to achieve stationarity, making the data more amenable to analysis.

An ARIMA(p,q) model is written as,

$$\begin{aligned} Y_t &= (1+\phi_1)\,Y_{t-1} + (\phi_2-\phi_1)\,Y_{t-2} + (\phi_3-\phi_2)\,Y_{t-3} + ... + (\phi_p-\phi_{p-1})\,Y_{t-p} - \phi_p Y_{t-p-1} \\ &\quad +\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - ... - \theta_q\varepsilon_{t-q}. \end{aligned} \tag{1.5.11}$$

Now the equation (1.5.11) is called the difference equation of the model. It is the form of ARMA(p+1,q) form.

1.6 Non-linear time series models

Linear Gaussian time series models have undergone significant advancements over the course of several decades, which has led to their widespread use in both theoretical research and practical applications. It can be noticed from the work of Yule (1927), where he introduced AR modelling to understand the patterns in sunspot numbers as time progressed, to the work of Box and Jenkins (1970), which exhibits the level of maturity attained by ARMA modelling in terms of theory and methodology, linear Gaussian time series models prospered and ruled both theoretical investigations and empirical applications. In applied research, it has frequently been figured out that basic linear time series models normally leave specific parts of financial data unexplained. Moran (1953) pointed out some limitations of the linear models. He revealed the fact that the residuals for the sample points greater than the mean were significantly smaller than those for the sample points smaller than the mean, which can

now be well explained by the term regime by Tong (1990). The early 1980s denoted a critical turning point in the literature as non-linear time series analysis (NTSA) began acquiring significant consideration. This expanded interest originated from the acknowledgment that linear time series models proved inadequate in capturing numerous crucial real world phenomena. Some of these peculiarities involve assymmetric business cycles, sustained animal population cycles, stock market volatility, regime switching, and various other intricate patterns that demanded more sophisticated analytical approaches. We cannot mathematically define a non-linear process in the same way we have defined a linear one, as mentioned in equation (1.5.1). Since there can be departures from linearity in different directions; we can only define a non-linear phenomenon through those features that cannot be exhibited by linear processes. and that has been observed in various disciplines like modelling of the regime effect or other non-standard features, which include non-normality, assymmetric cycles, bimodality, non-linear relationships between lagged values, variation of prediction performance over the state space, time reversibility, sensitivity to initial conditions, and others. Tong (2022) has given an extensive record of significant improvements to the new subject in the 1980s. It can be said that this period was overwhelmed by the advancement of parametric models. In specific, two classes of models, the threshold autoregressive (TAR) models and the generalized autoregressive restrictive heteroscedastic (GARCH) models, introduced during these early days appear to have endured consideration among professionals in the more extensive scientific and financial networks, which incorporate science, dynamical systems,

nature, financial matters, econometrics, financial engineering, and many others. The huge development of non-linear time series analysis occurred in numerous significant fields in the 1990's. According to Tong (2022), the following are the five most promising directions: the interface between NTSA and chaos, the non-parametric or semi-parametric approach, non-linear state space modelling, financial time series (in both discrete and continuous time), and non-linear modelling of panels (spatially distributed) of time series.

There are two categories of modelling non-linear features: implicit and explicit. In the implicit case, retaining the general ARMA framework, choose the distribution of the white noise appropriately so that the resulting process exhibits a specified non-linear feature. Here, the conditional expectation of the random variables given their lagged values may well be non-linear. Though the modelling capacity of this approach is highly large, identification of the distribution of white noise is a tedious task. So explicit models, which typically express a random variable as a non-linear function of its lagged values, are widely used. According to Tong (1990) and Tjøstheim (1994), the early development of time series analysis has been concentrated on various non-linear parametric forms. Autoregressive conditional hetroscedastic (ARCH) model by Engle (1982), GARCH models by Bollerslev (1986), and threshold modelling of biological and economic data by Tong (1990), Tiao, and Tsay (1994) are examples.

Volatility models like ARCH, GARCH, etc. are used to model data sets showing high volatility over time, such as the financial time series. An ARCH model was introduced by Engel (1982) is given below.

$$X_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = a_0 + b_1 X_{t-1}^2 + ... + b_q X_{t-i}^2 a_0 \geq 0 \ \ , b_j \geq 0 \ \ and \, \{\varepsilon_t\} \sim iid \, (0,1) \tag{1.6.1}$$

A class of non-linear ARCH model were introduced by Higgins and Bera (1992) is given by .

$$h_t = \left[\phi_0 \left(\sigma^2\right)^\delta + \phi_1 \left(\varepsilon_{t-1}^2\right)^\delta + ... + \phi_p \left(\varepsilon_{t-p}\right)^\delta \right]^{\frac{1}{\delta}} \tag{1.6.2}$$

where, $\sigma^2 > 0$, $\phi_i \geq 0$, $\delta > 0$, for i=0, 1, 2, ...,p; $\sum_{i=0}^{p} \phi_i = 1$

Bollerslev (1986) introduced GARCH model by replacing the term σ_t^2 in (1.6.1) with

$$\sigma_t^2 = a_0 + a_1 \sigma_{t-1}^2 + ... + a_p \sigma_{t-p}^2 + b_1 X_{t-1}^2 + ... + b_q X_{t-q}^2 \tag{1.6.3}$$

where $a_j \geq 0$ and $b_j \geq 0$.

To overcome some weaknesses of the GARCH model in handling financial time series, Nelson (1991) proposed the exponential GARCH (EGARCH) model.

Tong (1983) discussed the idea of using probability switching in non-linear time series analysis. Hamilton (1989), emphasizing aperiodic transition between various states of an economy, introduced Markov switching autoregressive(MSA) model given by

$$X_t = \begin{cases} C_1 + \sum_{i=1}^{p} \phi_{1.i x_{i-1}} + a_{1t} & \text{if } S_t = 1 \\ C_2 + \sum_{i=1}^{p} \phi_{2.i x_{i-1}} + a_{2t} & \text{if } S_t = 2 \end{cases} \tag{1.6.4}$$

S_t assumes values in $\{1, 2\}$, and is a first order markov chain with transition probabilities ,

$P\left(S_t = 2/S_{t-1} = 1\right) = \omega_1$

$P\left(S_t = 1/S_{t-1} = 2\right) = \omega_2$

S_t assumes values in $\{1, 2\}$ and $\{a_{1t}\}$ are sequences of random variables with mean zero and finite variance and are independent of each other.

Bilinear model was introduced by Granger and Anderson (1978)

$$X_t = C + \sum_{i=1}^{p} X_{i-1} - \sum_{j=1}^{q} \theta_j a_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{s} B_{ij} X_{t-j} a_{t-j} + a_t \tag{1.6.5}$$

where p,q,m and s are non negative integers.

TAR model, contained within the state-dependent (regime-switching) models family, introduced by Tong (1983). In this model the different linear forms were assumed in different regions of the state space. They are piece wise linear models in which the linear relationship varies with the values of the process. Here the division of the state space is usually based on the threshold variable, denoted by X_{t-d}.

$$X_t = b_0^{(i)} + b_1^{(i)} X_{t-1} + ... + b_p^i X_{t-p} + \varepsilon_t^i, if X_{t-d} \in \Omega_i \tag{1.6.6}$$

where $\{\Omega_i\}$ forms a non overlapping partition of the real line and $\{\varepsilon_t^i\} \sim iid\left(0, \sigma^2\right)$. The simplest class of TAR models, Self Exciting Threshold Autoregressive (SETAR)

models of order p introduced by Tong (1983) is given by,

$$Y_t = \begin{cases} a_0 + \sum_{j=1}^{p} a_j Y_{t-j} + \varepsilon_t \text{ if } Y_{t-d} \leq r \\ (a_0 + b_0) + \sum_{j=1}^{p} (a_j + b_j) Y_{t-j} + \varepsilon_t \text{ if } Y_{t-d} \geq r \end{cases} \tag{1.6.7}$$

Properties of general SETAR models are found very difficult to obtain. Further more some of them can be found in Tong (1990), Chan (1993), Chan and Tsay (1998), and the references therein. The fundamental idea behind these non-linear models involves allowing the time-evolution of the conditional mean, denoted as μ_t, to follow a relatively simple parametric non-linear function. Recent advancements in computational techniques has led to the development of various non-linear models. Notable advancements encompass the non-linear state-space model proposed by Carlin, Polson and Stoffer (1992), the functional-coefficient autoregressive model introduced by Chen and Tsay (1993a), the non-linear additive autoregressive model presented by Chen and Tsay (1993b), and the multivariate adaptive regression spline method developed by Lewis and Stevens (1991). The basic idea of these extensions is either using simulation techniques to describe the conditional distribution of x_t or using data-driven methods to find the non-linear characteristics of a series. Moreover, we can investigate the presence of non-linearity within time series data by employing non-parametric and semi-parametric strategies, including kernel regression and artificial neural networks etc.

In the upcoming sections, we will deal with the non-linear time series models,

such as non-linear AR models, autoregressive conditional duration (ACD) models and minification models. These models will be discussed in detail since our thesis primarily focuses on constructing them, utilizing the newly constructed distribution as marginal distribution.

1.6.1 Non-linear autoregressive models

Time series analysis employs a technique called non-linear autoregressive (NAR) modelling. By taking into account non-linear relationships between the prior observation and the present value, the linear AR models are thus expanded. As a result, they are able to capture the intricate and non-linear patterns found in real-world data. Numerous areas in corporate finance, economics, weather forecasting, and other disciplines where the study of time dependency data is prevalent are covered by the applications of NAR models. Given is a non-linear AR(1) model,

$$X_{n+1} = \lambda(X_n) + Z_{n+1}, \quad n = ..., -1, 0, 1, ... \tag{1.6.8}$$

where $\lambda(.)$ is a fixed real function of real argument, called autoregression function of the process, and the series $\{Z_n\}$ is a sequence of iid random variables.
A generalisation of (1.6.8) is defined as,

$$X_t = C + f_1(X_{t-i}) + f_2(X_2 - i_2) + ... + f_p(X_t - i_p) + \varepsilon_n. \tag{1.6.9}$$

Benrhmach et al. (2020) described the applications of non-linear autoregressive models in neural networks. They examined the use of the non-linear autoregressive neural network method as a prediction technique for financial time series and the intended Kalman filter algorithm for improving the accuracy of the model. Blasques et al. (2020) introduced a non-linear autoregressive model with time-varying coefficients.

Hunt et al. (1995) proved that every non-linear system with a Volterra series expansion can be represented as a non-linear AR model of infinite order.

1.6.2 Autoregressive Conditional Duration models

The majority of empirical studies in finance are carried out using daily data obtained by retaining either the first or the last observation for the variable under study. Here, all the intraday events were neglected. However, due to the advanced growth of automation in financial markets and the rapid developments in computer power, intraday databases that record every single transaction and their characteristics are easily made available. Now the analysis of high-frequency data (HFD) provides a deeper understanding of market activity. The identifying feature of this data is that the observations are irregularly time-spaced. This aspect presents a challenge for researchers, as the conventional econometric methods developed over time are no longer directly suitable. Now, recent models from the market microstructure literature argue that time carries important information and should be taken into account when constructing the models.

Engel and Russel (1998) developed the autoregressive conditional duration (ACD) model to analyze the dynamic behavior of financial duration data. This can be conceived as a marginal model of durations x_i. Consider the time of the occurrence of an event or transaction t_i.

Let $t_0 = 0$, $X_i = t_i - t_{i-1}$, $\quad i = 1, 2, 3, ...$ denotes the i^{th} duration between two events that occur at times t_{i-1} and t_i.

Let

$$\psi_i = E\left(X_i / F_{i-1}\right), \tag{1.6.10}$$

ψ_i represents the conditional expectation of the adjusted duration. It is the information known at time $(i-1)$. Let F_{i-1} be the information set available at the $(i-1)^{th}$ trade, and ψ_i be the expected adjusted duration given F_{i-1}. The main assumption of the model is that the standardized durations

$$\varepsilon_i = \frac{x_i}{\psi_i} \tag{1.6.11}$$

are iid with $E\left(\varepsilon_i\right) = 1$.

For the ACD(1,1) model, Engle and Russell (1998) derived its first two moments, while Bauwens and Giot (2000) computed its autocorrelation function.

The conditional mean of x_i is ψ_i.The unconditional mean,

$$E\left(X_i\right) = \frac{\omega}{1 - \sum_{i=1}^{m} \alpha_i - \sum_{j=1}^{q} \beta_j} \ , \tag{1.6.12}$$

The conditional variance of x_i based on (1.6.10) is that

$$Var(X_i/F_i) = \psi_i^2 V\left(\varepsilon_i\right).$$

The model allows both conditional overdispersion and underdispersion. The basic ACD model, as proposed by Engle and Russell (1998), relies on a linear parameterization of (1.6.10) in which ψ_i depends on m past durations and q past expected durations:

$$\psi_i = \omega + \sum_{j-1}^{m} \alpha_j X_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}. \tag{1.6.13}$$

The equation (1.6.13) is referred to as the ACD (m,q) model. Here $\omega \geq 0, \alpha_j \geq 0$, $\beta_j \geq 0$. Carrasco and Chen (2002) established sufficient conditions to ensure β-mixing and finite higher-order moments for the ACD(m, q) model. Fernandes (2004) derived lower and upper bounds for the density of stationary ACD (m, q) models. The equations (1.6.10) and (1.6.11) are general models that can be constructed by choosing different specifications for the expected duration, ψ and distributions for ε. Engle and Russell (1998) used the standard exponential distribution (the shape parameter is equal to one), which results in the exponential autoregressive conditional duration (EACD) models. It provides quasi-maximum likelihood (QML) estima-

tors for the ACD parameters Engle and Russell(1998), and Engle (2002). Drost and Werker (2004) showed that consistent estimates are obtained when the QML estimation is based on the standard gamma family (hence including the exponential). The exponential specification results in a flat conditional hazard function, which is a highly constrained assumption often invalid in practical financial applications. One may refer to Engle and Russell (1998), Dufour and Engle (2000a,2000b), Feng et al. (2004), and Lin and Tamvakis (2004) for details. To achieve more flexibility in the model, Engle and Russell (1998) use the standardized Weibull distribution with a shape parameter equal to γ and scale parameter equal to one, the resulting model being called Weibull autoregressive conditional duration (WACD). Grammig and Maurer (2000) proposed the use of a Burr distribution that contains the exponential, Weibull, and log-logistic as special cases. The model is then called the Burr-ACD model. However, not all the moments necessarily exist for the Burr distribution without imposing restrictions on the parameters. Lunde (1999) introduced generalized gamma ACD (GACD) and derived their related hazard functions and conditional log-likelihood functions. Bauwens and Giot (2000) introduced the logarithmic version, the log-ACD model.

The ACD model and the GARCH model introduced in Bollerslev (1986) share common features. The autoregressive nature of the ACD model, as given in (1.6.11), enables it to capture the clustering of durations observed in high-frequency data effectively. It implies that, in a manner similar to how the GARCH model explains the clustering of volatility, shorter (longer) durations are often succeeded by consec-

utive shorter or longer durations. In the realm of removing dependencies in squared returns, a GARCH(1,1) model is often found to be adequate. When dealing with temporal dependencies in durations, a relatively simple ACD model of low order frequently proves successful.

ACD model can be formulated as an ARMA(max(m, q), q) model for durations xi, taking $\eta_i \equiv x_i - \psi_i$ which is a martingale difference by construction and rearranging the terms, equation (1.6.13) becomes

$$x_i = \omega + \sum_{j=1}^{max(m,q)} (\alpha_j + \beta_j)\, x_{i-j} + \sum_{j=1}^{q} \beta_j \eta_{i-j} + \eta_i. \tag{1.6.14}$$

Hautsch (2002) proposed different ACD models based on the generalized F distribution that includes, as special cases, the generalized gamma, Weibull and log-logistic distributions.

1.6.3 Minification process

In the literature, minification structures have been proposed as a feasible and effective alternative to non-Gaussian time series models. This model deviates from traditional approaches that depend on generating functions and often lack closed-form expression. We assume a stochastic model and the presence of a stationary sequence of random variables with a specified marginal distribution, subject to certain conditions. Numerous properties of additive autoregressive models are shared

by minimization models. Using the survival function of the underlying random variables, the existence of these models and characteristics can be facilitated.

Tavares (1980) introduced an autoregressive process of the form

$$X_n = \begin{cases} X_0 & \text{if } n = 0 \\ kmin\left(X_{n-1}, \varepsilon_n\right) & \text{if } n \geq 1 \end{cases} \tag{1.6.15}$$

where $k > 1$, is a constant and $\{\varepsilon_n\}$ is a sequence of iid random variables such that $\{X_n\}$ is a stationary markov process with a specified marginal distribution function $F_{X_0}(x)$. Due to its structure (1.6.15) is called the minification process. Hydrological issues served as the primary source of inspiration for Tavare's work, which included modelling runoff data. These data have long tails, making it impossible to model them using an exponential process like the linear autoregressive process of Graver and Lewis (1980). Weibull or extreme value distributions are frequently employed to represent the marginal distribution functions of run-off series, although processes involving these marginal distributions cannot be extended using linear random coefficient models. Therefore, minification procedures play a crucial role as a source of time series for such a procedure.

Sim (1986) developed a first-order autoregressive Weibull process and described its properties. Here $\{X_n\}$ are stationary Weibull random variables with a survival function $e^{\frac{-\theta x^c}{k^c-1}}$ if and only if $\{\varepsilon_n; n = 1, 2, ...\}$ is a sequence of iid Weibull random

variables with survival function $e^{-\theta x^c}$.

Lewis and Mckenzie (1991) gave a detailed description of minification processes and their transformations. If the survival function of X_n in (1.6.15) is $\bar{F}_{X_n}(x)$ for $n \geq 1$, then the survival function of $\{\varepsilon_n\}$ is obtained such that

$$\bar{F}_{\varepsilon_n}(x) = \frac{\bar{F_X}(kx)}{\bar{F}_X(x)}. \tag{1.6.16}$$

If $\bar{F_{\varepsilon_n}}(x)$ is not strictly a proper survivor function, having an atom of probability p located at infinity, then (1.6.15) can be written as

$$X_n = \begin{cases} kX_{n-1} & \text{with probability p ,} \\ kmin\left(X_{n-1}, \varepsilon_n^*\right) & \text{with probability 1-p} \end{cases} \tag{1.6.17}$$

where $0 < p < 1$.

Here,

$$\bar{F}_{\epsilon_n}^*(x) = \frac{\bar{F}_{\varepsilon_n}(x) - p}{1-p}.$$

Now all the properties can be derived similarly using

$$\bar{F}_{\varepsilon_n}(x) = p + (1-p)\bar{F}_{\epsilon_n}^*(x)$$

Arnold and Robertson (1989) constructed a minification process with a logistic

marginal distribution. Another form is

$$X_n = \begin{cases} \varepsilon_n & \text{with probability p} \\ min\left(X_{n-1}, \varepsilon_n\right) & \text{with probability 1-p} \end{cases} \tag{1.6.18}$$

where $0 < p < 1$, and $\{\varepsilon_n\}$ is the innovation series such that $\{X_n\}$ is a stationary Markov process with a given marginal distribution.

Autoregressive semi-logistic process presented in Jayakumar and Thomas (2004). Adke and Balakrishna (1992) have estimated the parameters of the exponential minification model. Balakrishna (1998) discussed the estimation problems in the semi-Pareto and Pareto processes. The minification process with discrete marginals was discussed in Kalamkar (1995). Ristic (2008) provided a three-parameter version of the two-parameter semi-Pareto minification process. The process is,

$$X_n = \begin{cases} \varepsilon_n & \text{with probability q} \\ p^{\frac{-1}{\alpha}} X_{n-1} & \text{with probability p(1-q) } , n \geq 1 \\ min(p^{\frac{-1}{\alpha}} X_{n-1}, \varepsilon_n) & \text{with probability (1-p)(1-q)} \end{cases}$$

where $\{\varepsilon_n, n \geq 1\}$ is a sequence of iid random variables, X_{n-1} and ε_n are independent random variables and $0 < p < 1$, $0 \leq q \leq 1$, $\alpha > 0$.

In recent years, several bivariate minification processes have been defined. Alice and Jose (2004) introduced bivariate minification processes with Marshall-Olkin

bivariate semi-Pareto and Pareto distributions and discussed their properties. Balakrishna and Jayakumar (1997) introduced a bivariate minification process of first order. Ristic (2006) considered a stationary bivariate minification process. The process is

$$X_n = K_1 min\left(X_{n-1}, Y_{n-1}, \varepsilon_n\right)$$

and

$$Y_n = K_2 min\left(X_{n-1}, Y_{n-1}, \eta_n\right)$$

where(ε_n, η_n) is a sequence of iid non- negative non-degenerate random vectors with common survival function $\bar{G}\left(x, y\right)$, random vectors and$(X_0, Y_0)$and (ε_1, η_1) are independent and $K_1 > 1$, $K_2 > 1$.

Krishna et al. (2011) discussed several applications of Marshall Olkin Frechet distribution and process.

1.7 Spatial Analysis

Spatial econometrics encompasses a specialized set of analytical techniques designed to accommodate interdependencies among observations located in close geographical areas, be they spatial points or regions. These techniques build upon the conventional linear regression model by introducing the concept of identifying cohorts of nearest

neighbors. Variables related to location, distance and arrangement are treated explicitly in model specification, estimation, and diagnostic checking and prediction in Anselin (2006). This implementation of spatial methods allows for acknowledging the interdependence that exists among these neighboring regions or observations, as in Anselin (1988) and LeSage and Pace (2009). Thus, spatial econometrics is a subfield of econometrics that portrays spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) in regression models for cross-sectional and panel data, as mentioned in Paelinck and Klaassen (1979). The spatial aspects are considered as spatial effects consisting of spatial dependence and spatial heterogeneity, as defined in Anselin (1988). To address the unique characteristics of spatial dependence, particularly the inherent feedback effects, a specific set of techniques is needed, as highlighted in Anselin (1988). Spatial heterogeneity represents a specific instance of structural instability, a common concern in traditional econometrics. This heterogeneity is spatially organized, meaning it pertains to different spatial locations or regional subsets of observations. Consequently, this spatial structure guides the formulation of heterogeneity, which encompasses spatially varying coefficients, random coefficients, and spatial regimes, akin to spatial fixed effects, as elucidated in Anselin (1988). Whereas early on, apart from Anselin (1980, 1988), Cliff and Ord (1981), and later LeSage and Pace (2009), there was a relative dearth of treatments that provided a comprehensive review of methods and models, this is no longer the case. In recent years, several new texts were published, providing ample access to the breadth of the field, in terms of theoretical results, new methods and a range of

applications. Recent examples include Anselin and Rey (2014), Arbia (2014), Dub´e and Legros (2014), Elhorst (2014a), Kelejian and Piras (2017), and Chi and Zhu (2020). In addition, there are several extensive reviews of the state of the art, both for a cross-sectional setting as well as for spatial panels. Examples include Anselin and Bera (1998), Anselin (2001, 2006, 2021), Anselin et al. (2008), Lee and Yu (2010, 2011, 2015), Elhorst (2012, 2014b), and Bai et al.(2016).

In parallel to time series analysis, spatial stochastic processes are categorized as spatial autoregressive (SAR) and spatial moving average (SMA) processes, although there are several important differences between the crosssectional and time series contexts. A spatial lag model, or a mixed regressive, spatial autoregressive model is expressed as

$$y = \rho W y + X\beta + \varepsilon \tag{1.7.1}$$

$$y = (I_n - \rho W)^{-1} X\beta + I_n (\rho W)^{-1} \varepsilon \tag{1.7.2}$$

where $\varepsilon \sim N(0, \sigma^2 I_n)$. These patterns were often expressed as spatial autocorrelation, which is the tendency for sites that are close together to have more similar values than sites that are farther from each other, as in Sokal and Oden (1978). When spatial autocorrelation exists in the data, ecologists often use spatial statistical models because the assumption of independent errors is violated, making many conventional statistical methods inappropriate (Cliff and Ord (1981), Legendre (1993)). Different spatial autoregressive models are discussed in detail in chapter 5.

In the next section a brief explanation of truncated distribution is given, as we have used a truncated distribution for the construction of the new disribution in the thesis.

1.8 Truncated Distributions

Truncated distributions are a class of probability distributions in which the feasible values of a random variable are confined to a specific interval, and any values falling outside this range are replaced with predefined boundary values.

Let X be a random variable with probability density function (pdf) or probability mass function (pmf) $f(x)$. The distribution of X is said to be truncated at the point $X = a$ if all the values of $X < a$ are discarded. The pdf (or pmf), g(.) of the distribution truncated at $X = a$ are given by:

$$g(x) = \frac{f(x)}{P(X > a)} \tag{1.8.1}$$

$$g(x) = \begin{cases} \frac{f(x)}{\sum_{x>a} f(x)} & \text{if } x > a \quad (for\ discrete\ random\ variable.)) \\ \frac{f(x)}{\int_a^\infty f(x)dx} & \text{if } x > a \quad (for\ continuous\ random\ variable.) \end{cases}$$

The r^{th} moment about origin for the truncated distribution is given by,

$$\mu_r{}' = E(X^r) = \int_a^\infty x^r g(x)\, dx.$$

Numerous scientific disciplines, including communication networks, hydrology, material science, physics, and hydrology, are heavily reliant on long-tailed distributions. For instance, modern communication networks like the internet often exhibit long-tailed distributions, as evidenced by numerous traffic measurement studies. This implies that the way of behaving of this information essentially leaves conventional phone traffic and its connected Markov models with short-range reliance. Specifically, the Poisson arrival process and related analysis in light of the Erlang equation are presently not substantial. The main shortcoming of long-tailed distributions is that they don't have finite moments of all orders. This has restricted their application. Saralees Nadarajah (2008), in his work, introduced truncated versions of five heavy-tailed distributions: Student's t distribution, F distribution, inverted beta distribution, Frechet distribution, and Levy distribution. These truncated versions, which possess finite moments of all orders and could therefore be better models.

1.9 Outline of the thesis

The thesis is organized in to six chapters as follows.

The first chapter is devoted to the introduction of stochastic process and time series. A short account of linear time series models, non-linear time series models and their properties are discussed. A brief description of spatial autoregressive models and truncated distributions is also given.

In the second chapter, we derive a new distribution in the range $[0, 1]$ called the

uniform truncated Poisson distribution using three different methods. The properties of this distribution are studied. We fit the new distribution to the real data set. For estimating the parameter of the distribution, the methods of maximum likelihood and moments are used.

In chapter 3, we construct the time series models in mixture form with the newly derived distribution as the marginal. Important properties of the new process are derived. The estimation of parameter is done using the conditional least squares estimation and maximum likelihood estimation methods. Simulation studies and real-world data applications are also discussed.

In the fourth chapter, we introduce uniform truncated Poisson autoregressive conditional duration process. Important properties of this process are studied. The estimation of parameter is done by using the method of maximum likelihood. Along with simulation studies, real-data applications are done.

Chapter five is devoted for spatial non-linear models. We analyse the spatial effect on mortality through spatial autocorrelations and unifies the regression and spatial effect of GDP. The analysis summarises the results of the recent years to compare the temporal effect. The ordinary least squares, spatial lag model, spatial error model and spatial Durbin model are considered at different time points and comparisons are made.

In the sixth chapter, we construct two types of autoregressive uniform truncated Poisson minification processes. Their properties are discussed. The estimation of parameters is done. The application is illustrated with real data.

Finally recommendations for future work with conclusion are given in chapter 7. Following this, we present a comprehensive bibliography containing a list of references.

CHAPTER 2

UNIFORM TRUNCATED POISSON DISTRIBUTION

2.1 Introduction

[1] A theoretical probability distribution provides a mathematical representation of how different values of random variables are distributed, each with specified probabilities. Recent research on probability distributions has mainly concentrated on distributions with support either on the real line or the positive real line. Researchers have traditionally focused more on probability distributions with support on the real line or positive real line, leaving distributions on finite intervals relatively understudied. However, it can be seen that numerous real-world data sets naturally fall within

[1]This chapter is based on Krishnarani and Vidya (2022)

finite intervals. Research on distributions in the interval $[0, 1]$ is an important area of study in statistics. As the interval $[0, 1]$ represents the set of all possible probabilities or proportions, it is a common range for probability distribution. While it is true that distributions on finite intervals may receive less attention compared to distributions defined on the entire real line or positive real line. For instance, in biology and experimental results in physics, chemistry, and other fields, a uniform pattern is often observed within the interval [0, 1]. Distributions defined on the interval $[0, 1]$ are not as common as some other widely used distributions, and the most widely used distributions belonging to this category are the power function distribution and the beta distribution. While these distributions are commonly used for modelling data on the $[0, 1]$ interval, it is worth noting that there are other distributions defined on this interval as well, including the logit-normal distribution, the arcsine distribution, and the Kumaraswamy distribution, among others. These distributions offer flexibility and are useful for modelling data that is naturally constrained within the $[0, 1]$ interval. Some recent distributions defined on [0,1] are available in the research papers Altawil (2019) and Hassan et al. (2020).

By suitable transformation, any random variable taking values in the real line can be transformed into $[0, 1]$, and we can proceed with further studies. Rescaling data into the range $[0, 1]$ is a valuable preprocessing step in machine learning and image processing. Among the various techniques available, the most elegant and commonly used method is the min-max scaling procedure. This is usually used as an alternative method to z-score normalization. Also, in neural networks, $[0, 1]$ data is

essential for normalizing pixel intensities. As discussed in Weigend and Gershenfield (1993) and Yu et al. (2006) normalization plays an inevitable role in the data management. By this transformation, all the features are kept the same, but it results in smaller standard deviations of the observations, which minimizes the outlier effect. The subsequent rescaled data saves the original data's relationships of the original data; however, all values are currently bound to the minimal and interpretable range of $[0, 1]$. This transformation is especially beneficial while working with algorithms that are sensitive to the scale of input features, as it keeps specific elements from ruling others because of their magnitude.

So in this chapter, we made an attempt to study a distribution with support on $[0, 1]$ which was mentioned in Hao and Godbole (2014). More recently, Quijano Xacur (2019) has introduced a new distribution with support on $[0, 1]$, called the unifed distribution, which can be used as the response distribution for a generalized linear model. Taking the index parameter as unity, this family gives the distribution we study in this chapter. We further investigate this distribution by presenting it in another manner and uniting the pertinent properties and results concerning it. For deriving this, we used the compounding method. Derivations of new discrete and continuous distributions, compounding two distributions, have been described by several authors, see for instance, the uniform-geometric distribution in Akdogan et al. (2016), binomial-Poisson distribution in Hu et al. (2007), and a Weibull-power series distribution in Morais and Barreto-Souza (2011). More distributions can also be seen in Adamidis and Loukas (1998), Kus (2007), Thahmasbi and Rezaei

(2008), and Chahkandi and Ganjali (2009). We have a few notable disseminations, like beta distribution and power function distributions, with support on $[0, 1]$. These distributions are found to have a wide range of applications in several real-world circumstances, including reliability, time series, etc. So we have made a comparison of the distribution studied in this chapter with these well-known distributions. Indeed, the application of the uniform truncated Poisson distribution in modelling time series data opens up new possibilities for advanced-level model diagnosis, specifically in the context of non-linear and volatile time series.

This chapter is organized as follows: In Section 2.2, a uniform truncated Poisson distribution is introduced. Its properties are studied in Section 2.3. Transformations are considered and corresponding distributions are derived in Section 2.4. The estimation of the parameter is discussed in Section 2.5, and numerical illustrations are given therein. The asymptotic properties of the estimators are also outlined in the same section. A generalization of this new distribution with support on a finite interval is done in Section 2.6. Application to real data sets is given in Section 2.7 followed by a concluding section.

2.2 Uniform truncated Poisson distribution

Probability distributions defined on the interval $[0, 1]$ are relatively less common in the literature, and among them, the most widely used distributions are the power

function distribution and the beta distribution. Applications of the distributions in the interval $[0, 1]$ are mentioned in the introduction part. The wide applications of such distributions in various fields like neural networks, pixel intensities, artificial intelligence, physics, engineering, time series, etc. are the motivation behind this study. There have been limited studies conducted on regression and time series models specifically concentrated on variables within the $[0, 1]$ range, such as percentages or fractions. Some of the notable works in this context in the literature are Kieschnick and McCullough (2003), Jara et al. (2013), Ristic and Popovic (2000a), Rocha and Cribari-Neto (2009), and Bayer et al. (2018). So the distribution studied in this paper may be applied in the advanced fields of these areas. We have illustrated some of the applications in Section 2.5. We consider the distribution on $[0, 1]$ mentioned in Hao and Godbole (2014) and construct this distribution in three different ways. These methods are described below.

Method:1

Let U be a random variable following truncated uniform distribution with pdf,

$$g(u) = \frac{1}{e^{\theta} - 1}, \quad 0 \leq u \leq e^{\theta} - 1.$$

Consider the transformation,

$$X = \frac{\log(1 + U)}{\theta}. \tag{2.2.1}$$

Then the pdf of X is

$$f(x) = \frac{1}{e^{\theta} - 1}\theta e^{\theta x}, \quad 0 \leq x \leq 1, \theta \neq 0. \tag{2.2.2}$$

We can see that when $\theta = 0$ the distribution is uniform in [0,1].

Method: 2

The distribution mentioned in (2.2.2) can be written as a solution of the first-order differential equation as given below. Radioactive decay is associated with a differential equation, and an exponential random variable is an example of it. So we are trying to construct a distribution with an initial value at time zero as a function of θ but the limit of the initial value function at time point zero is 1. This initial value function $\frac{\theta}{e^{\theta}-1}$ is monotone decreasing in θ.

Let

$$\frac{\mathrm{d}y}{\mathrm{d}x} - \theta y = \frac{\theta}{e^{\theta} - 1} \tag{2.2.3}$$

be the first order differential equation and choose $y = F(x)$.

That is

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} - \theta F(x) = \frac{\theta}{e^{\theta} - 1}.$$

Solving we get

$$F(x) = \frac{e^{\theta x} - 1}{e^{\theta} - 1}$$

$$\bar{F}(x) = 1 - F(x) = \frac{e^{\theta} - e^{\theta x}}{e^{\theta} - 1}$$

and hence $f(x)$ is of the form (2.2.2). Here we can see that distribution is the solution of the first order differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} - \theta y = \frac{\theta}{e^{\theta} - 1}.$$

Method :3

It can be seen that the random minimum or maximum of N iid random variables is studied in Louzada et al. (2011), Kus (2007), Cancho et al. (2011), and several other papers. It may be noted that Hao and Godbole (2014) have introduced the uniform poisson model; its derivation is given below. They have applied the method mentioned above and explored a few properties in that paper. So using the procedure used there, under the assumption that N is a truncated Poisson with a probability mass function

$$P(N = n) = \frac{e^{-\theta}\theta^{N}}{N!(1 - e^{-\theta})}, \quad N = 1, 2, ...$$

and $X_1, X_2, ..., X_N$ to be of $U[0, 1]$ with distribution function F(.), the distribution of X=min(X_i) is,

$$g(X = x) = f(x) \sum_{n=1}^{\infty} n \left(F(x)\right)^{N-1} P(N = n)$$

which is exactly the same as (2.2.2). Hence we call this random variable X with pdf (2.2.2) as Uniform Truncated Poisson Distribution (θ) denoted as UTPD(θ). It is interesting to note that UTPD(θ) is derived in three distinct ways. In the next

section, we discuss the properties of UTPD.

2.3 Properties

1. The survival function is given by,

$$\bar{F}(x) = \frac{e^{\theta} - e^{\theta x}}{e^{\theta} - 1}. \tag{2.3.1}$$

2. The hazard function,

$$\begin{aligned} h(x) &= \frac{f(x)}{\overline{F}(x)} \\ &= \frac{\theta e^{\theta x}}{e^{\theta} - e^{\theta x}} = \frac{\theta}{e^{\theta(1-x)} - 1} \end{aligned}$$

We have,

$$\frac{\mathrm{d}}{\mathrm{d}x} h(x) = \frac{\theta^2 e^{\theta(x+1)}}{\left(e^{\theta} - e^{\theta x}\right)^2}.$$

It can be seen that for all θ values, the distribution has an increasing failure rate (IFR).

3. The characteristic function is

$$\phi_X(t) = \frac{\theta}{(e^{\theta} - 1)} \frac{1}{(\theta + it)} \left(e^{\theta + it} - 1\right). \tag{2.3.2}$$

4. The r^{th} moment of UTPD is given by

$$\begin{aligned} E(X^r) &= \frac{e^\theta}{e^\theta - 1}\left[1 - \frac{r}{\theta} + \frac{r(r-1)}{\theta^2} - \frac{r(r-1)(r-2)}{\theta^3}\right. \\ &\quad \left. + \frac{r(r-1)(r-2)(r-3)}{\theta^4} - ... + (-1^r)\frac{r(r-1)(r-2)(r-3)...1}{\theta^r}\right] \\ &\quad + \frac{1}{e^\theta - 1}(-1)^{r+1}\frac{r(r-1)(r-2)(r-3)...1}{\theta^r}, \quad \text{for } r\text{=1, 2, ...} \end{aligned}$$

5. Mean= $\frac{e^\theta(\theta-1)+1}{\theta\left(e^\theta-1\right)}$.

6. Variance= $\frac{e^\theta}{e^\theta-1}\left(1 - \frac{2}{\theta} + \frac{2}{\theta^2}\right) - \frac{2}{\theta^3\left(e^\theta-1\right)} - \left(\frac{e^\theta(\theta-1)+1}{\theta\left(e^\theta-1\right)}\right)^2$

   For $\theta = 1$, Mean=$\frac{1}{e-1}$, $Variance = \frac{e^2-3e+1}{(e-1)^2}$.

7. The p^{th} quantile is given by $x_p = \frac{1}{\theta}\log\left\{1 + p\left(e^\theta - 1\right)\right\}$, $0 < p < 1$.

8. Entropy, a measure of the uncertainty associated with the random variable is given by

$$H(X) = \frac{-\theta}{\theta - 1}\left\{\frac{\ln \theta e^\theta}{\theta} - \ln\theta\left(\frac{1}{\theta}\right) + \frac{e^\theta}{\theta} - \frac{1}{\theta^2}\left(e^\theta - 1\right) - \ln\left(e^\theta - 1\right)\right\}.$$

9. Odds ratio : Odds ratios are often used in the medical literature.

   (a) The odds ratio of surviving beyond time, $\phi^+ = \frac{\bar{F}(x)}{F(x)} = \frac{e^\theta - e^{\theta x}}{e^{\theta x}-1}$

   (b) The odds ratio of failure by time, $\phi^- = \frac{F(X)}{\bar{F}(X)} = \frac{e^{\theta x}-1}{e^\theta - e^{\theta x}}$.

The density function, distribution function, and the hazard function for different values of θ are plotted in Figures 2.1 to 2.3 respectively. From the density plots, it is clear that the positive value of the parameter θ confirms the left-skewed behavior and a negative value indicates the right skewed behavior. So it is a distribution on $[0, 1]$, which can be used for modelling left or right skewed data sets. When the value of θ is positive and increases, the density function becomes more peaked but is less left-skewed. The behavior is completely contrary when θ is negative. Even though from Figure 2.3 it is clear that the distribution has IFR for different values of θ, the behavior of the hazard function doesn't vary much. The nature of this distribution is actually very similar to power function distribution. It means a comparison with power function distribution will be quite interesting. For illustrating this, the density plots of UTPD and power function distribution are drawn together in Figure 2.4.

As $\theta > 0$ and increases UTPD, coincides with power function distribution. A comparison with beta distribution is also interesting since beta distribution is a flexible distribution with wide applications. But we know that the failure rate function of the beta (p,q) distribution is increasing only if $p \geq 1$, and the comparison will be meaningful only in this particular case. So we have not considered this part in this study.

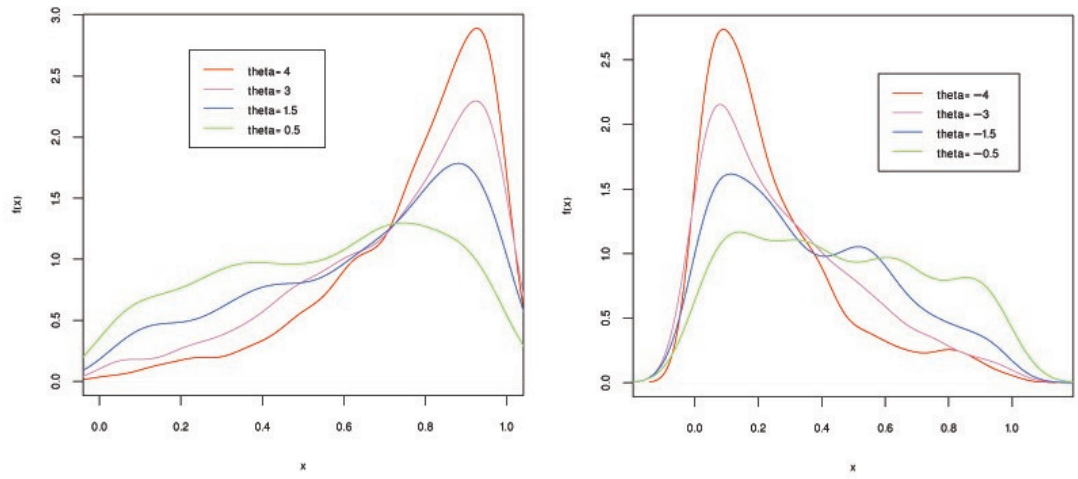

Figure 2.1: Density plots of UTPD for various values of θ

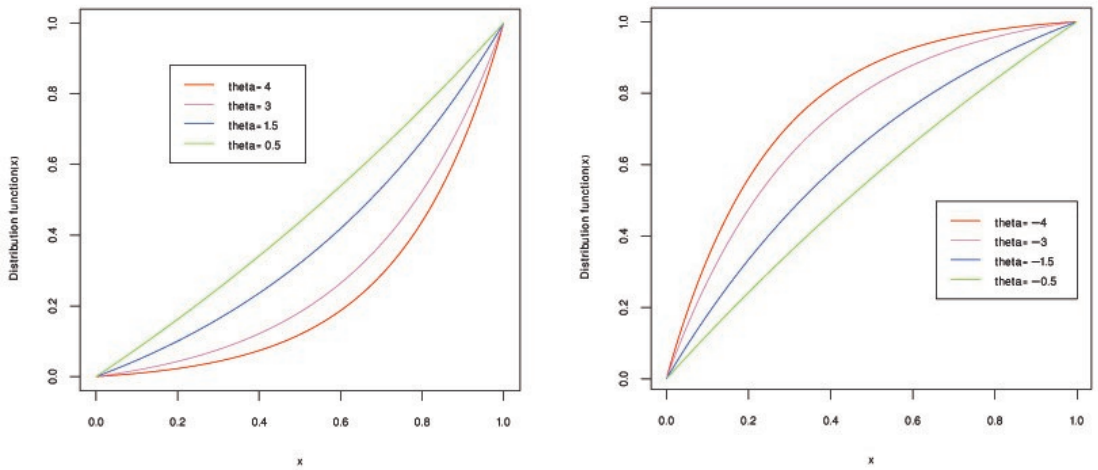

Figure 2.2: Distribution function of UTPD for various values of θ

10. Skewness and Kurtosis

Using the quantile function given in property 7, the first, second and third

quantiles are $x_{0.25}$, $x_{0.50}$, $x_{0.75}$ respectively. Bowley's measure of skewness,

$$\begin{aligned} S &= \frac{x_{0.75} + x_{0.25} - 2x_{0.50}}{x_{0.75} - x_{0.25}} \\ &= \frac{log[\frac{\frac{1}{16}((3e^{\theta}+1)(e^{\theta}+3)}{(\frac{1}{2}(e^{\theta}+1))^2}]}{log[\frac{3e^{\theta}+1}{e^{\theta}+3}]}. \end{aligned}$$

The kurtosis is measured by the method proposed by Moors (1988). He derived this measure using octiles. The octiles E_i are defined as,

$$P(X < E_i) \leq \frac{i}{8}$$

and

$$P(X > E_i) \leq 1 - \frac{i}{8}.$$

Using octiles, the measure of kurtosis ,

$$K = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2}.$$

The skewness and kurtosis for different parameter values are given in Table 2.1 and the observations we made from the density plots regarding skewness and kurtosis are very well established numerically in this table. It can be seen that the distribution is symmetric with respect to θ, but the value of kurtosis is the same for both negative and positive values of the parameter.

Table 2.1: Skewness and Kurtosis

Parameter:θ	Skewness	Kurtosis
0.5	-0.0613	1.0114
2	-0.1953	1.1385
5	-0.2579	1.2925
8	-0.2616	1.3055
-0.5	0.0613	1.0114
-2	0.1953	1.1385
-5	0.2579	1.2925
-8	0.2616	1.3055

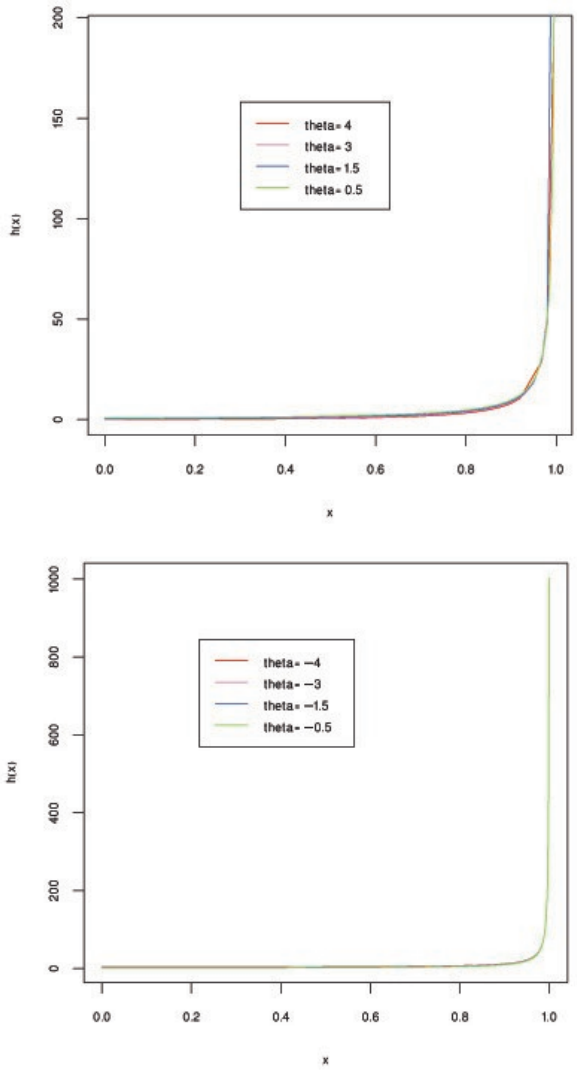

Figure 2.3: Hazard function of UTPD for various values of θ

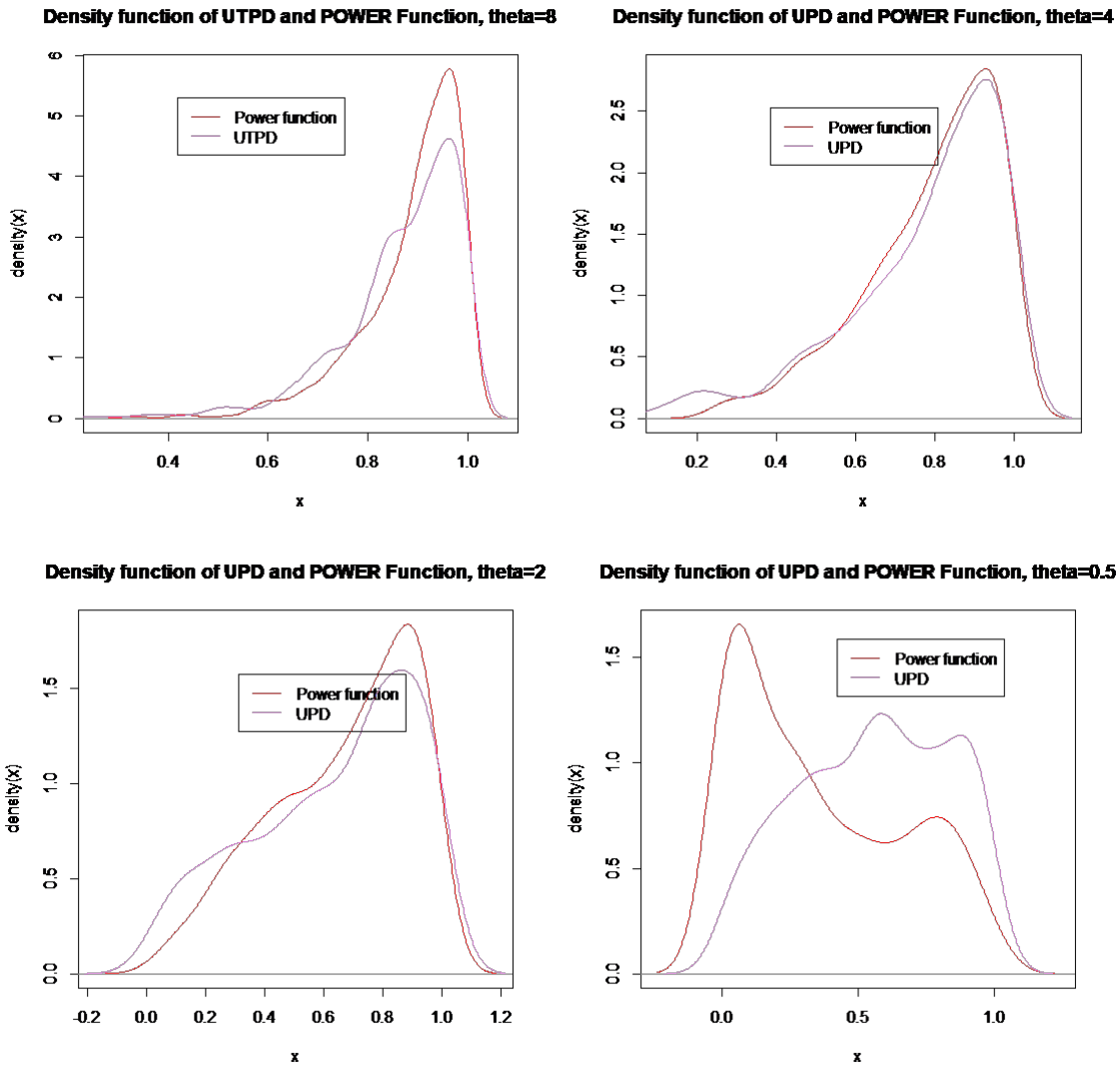

Figure 2.4: Comparison of UTPD and power function distribution

Remark 2.3.1. This distribution finds extensive utility in machine learning, particularly in the context of data normalization for representation, subsequent processing,

and enhancing accuracy. The usual transformation used for this purpose is given by

$$\frac{x_i - min(x_i)}{max(x_i) - min(x_i)}. \tag{2.3.3}$$

Later in our real data analysis part, described in the last section of this chapter, we explain the use of this distribution in such transformations.

2.3.1 Distribution of Order Statistics

Let $X_1, X_2, X_3, ..., X_n$ are independent random variables following UTPD with parameter θ. The pdf of $\min(X_1, X_2, X_3, ..., X_n)$ is given by

$$f_{X_{(1)}}(x) = \frac{n\theta e^{\theta x}}{(e^{\theta} - 1)^n} \left(e^{\theta} - e^{\theta x}\right)^{n-1}$$

and the pdf of $\max(X_1, X_2, X_3..., X_n)$ is given by

$$f_{X_{(n)}}(x) = \frac{n\theta e^{\theta x}}{(e^{\theta} - 1)^n} \left(e^{\theta x} - 1\right)^{n-1}.$$

We describe some transformed distributions in the next section, which seem very similar to some well-known distributions but with different domains.

2.4 Transformed distributions

We consider some random variables generated through transformations of (2.2.2) and derive their distributions.

Result 1:

Considering the transformation

$$U = -\log X,$$

where X follows UTPD with density function given in (2.2.2), the pdf of U is

$$g(u) = \frac{\theta}{e^{\theta} - 1} e^{-u} e^{\theta e^{-u}}, \quad 0 \leq u < \infty, \tag{2.4.1}$$

which is the Weibull-Poisson distribution by Morais and Barreto-Souza (2011).

Result 2:

When we take a power transformation

$$V = X^{\frac{1}{\beta}} \tag{2.4.2}$$

the density function of V becomes

$$g(v) = \frac{\theta\beta}{e^{\theta} - 1} v^{\beta-1} e^{\theta v^{\beta}}, \quad 0 \leq v \leq 1, \tag{2.4.3}$$

which is similar to the Weibull distribution, but the domain is quite different.

Result 3:

The pdf of $W = \frac{1}{X}$, where X follows UTPD is

$$h(w) = \frac{\theta}{e^{\theta} - 1} e^{\theta/w} \frac{1}{w^2}, 1 \leq w < \infty. \tag{2.4.4}$$

The estimation of the parameter of the UTPD is discussed in the next section.

2.5 Estimation of the parameter

Methods of maximum likelihood (ML) and method of moments (MM) are used for the estimation, and comparisons are made with numerical illustrations.

2.5.1 Maximum likelihood estimation

Suppose a sample of size n is taken from UTPD with pdf mentioned in (2.2.2). By taking logarithm of the likelihood function and finding the derivative with respect to θ, we have a non-linear equation

$$\frac{\partial \log L}{\partial \theta} = 0 \Rightarrow \frac{n}{\theta} - \frac{ne^{\theta}}{e^{\theta} - 1} + \sum_{i=1}^{n} x_i = 0,$$

which can be solved numerically to estimate the parameter.

2.5.2 Method of moments

Another method commonly used for the estimation of parameters is the method of moment estimation (MM). Equating the first raw moment to the corresponding central moment, we get the equation given below. Solving the same for θ we will get the estimate.

$$\frac{\sum_{i=1}^{n} x_i}{n} = \frac{e^{\theta}(\theta - 1) + 1}{\theta(e^{\theta} - 1)}.$$

2.5.3 Large sample properties

Large sample properties give significant bits of knowledge into the way estimators behave as the sample size increases, helping researchers understand the quality and reliability of their statistical analysis. The asymptotic properties of the ML estimators under the usual regularity conditions are provided in this section.

Property 1: The ML estimator $\hat{\theta}$ is asymptotically normally distributed with mean θ and variance $\frac{1}{nI(\theta)}$ where $I(\theta)$ is the well known information matrix.

Proof. : The log likelihood function is given by

$$\log L = n \log \theta - n \log\left(e^{\theta} - 1\right) + \theta \sum_{i=1}^{n} x_i.$$

Then

$$\frac{\partial log L}{\partial \theta} = \frac{n}{\theta} - \frac{n}{e^{\theta} - 1} e^{\theta} + \sum_{i=1}^{n} x_i$$

$$\frac{\partial^2 logL}{\partial\theta^2} = \frac{-n}{\theta^2} + \frac{ne^\theta}{(e^\theta - 1)^2}.$$

If we denote the gradient of log L, the score statistic as $S(\theta)$, and $-\frac{\partial^2 logL}{\partial\theta^2}$ as $K(\theta)$, then the above equation can be written as,

$$K(\theta) = -S'(\theta) = \frac{n}{\theta^2} - \frac{ne^\theta}{(e^\theta - 1)^2}.$$

Also we know that,

$$S(\theta) = \frac{\partial logL}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial logf(X_i,\theta)}{\partial\theta} \ \ and$$

$$K(\theta) = \sum_{i=1}^{n} K(X_i, \theta).$$

Then, $E(K(\theta)) = nI(\theta)$ where $I(\theta) = E[\frac{\partial logf(X_i,\theta)}{\partial\theta}]^2$ the information matrix.

Using Taylor's formula,

$0 = S(\hat{\theta}) = S(\theta) - K(\theta)(\hat{\theta} - \theta) + R$, where R tends to zero.

And finally after adjusting the terms, $\sqrt{n}(\hat{\theta} - \theta) = \frac{S(\theta)/\sqrt{n}}{K(\theta)/n}$.

By Slutsky's theorem $\hat{\theta}$ converges in distribution to $N(\theta, \frac{1}{nI(\theta)})$. □

Now the consistency property of $\hat{\theta}$ is stated below, the proof of which readily follows as in Kale (2007).

Property 2: The likelihood equation admits a consistent solution and the consistent estimator is essentially unique.

2.5.4 Numerical Examples

Simulated samples of sizes 20, 60, and 100 from the population following UTPD for selecting the better method of estimation. For the comparison purpose of the two methods discussed above, each sample is generated 1000 times. The estimates of θ, standard error (SE), mean square error (MSE), 95% confidence intervals (CI) for the parameters, and coverage probabilities (CP) are shown in Table 2.2. All the simulation works and other computations are done using R programming. From the table, we can see that SE and MSE are decreasing with an increase in sample size. The coverage probabilities are increasing with increase in the sample size. But for smaller sample sizes, the coverage probabilities of the parameters estimated using the ML method is lesser than that generated by the method of MM. It is clear from the table that both the ML method and MM are equally good for estimation purposes based on the MSE and both the methods give us approximately equal values as parameter estimates.

In the next section, we have made an attempt to generalize the UTPD into a general finite interval (a,b).

Table 2.2: Parameter Estimates

Sample size(n)	Parameter(θ)	Method	Estimate ($\hat{\theta}$)	SE	MSE	CI	CP
20	4	MLE	4.12	1.110	1.307	(3.15,5.09)	0.627
		MM	4.10	0.252	1.288	(1.88,6.37)	0.957
60		MLE	4.07	0.625	0.385	(3.12,5.02)	0.880
		MM	4.03	0.080	0.394	(2.80,5.26)	0.948
100		MLE	4.01	0.482	0.234	(3.07,4.96)	0.955
		MM	4.02	0.048	0.246	(3.06,4.98)	0.954
20	3	MLE	3.06	0.968	1.020	(2.21,3.91)	0.620
		MM	3.09	0.233	1.098	(1.05,5.14)	0.938
60		MLE	3.06	0.553	0.303	(2.22,3.90)	0.876
		MM	3.04	0.072	0.312	(1.95,4.13)	0.950
100		MLE	3.01	0.424	0.183	(2.18,3.84)	0.950
		MM	3.01	0.043	0.190	(2.15,3.86)	0.947
20	1.5	MLE	1.56	0.835	0.740	(0.83,2.29)	0.622
		MM	1.54	0.196	0.770	(-0.17,3.26)	0.950
60		MLE	1.50	0.475	0.231	(0.78,2.22)	0.861
		MM	1.55	0.063	0.147	(0.79,2.23)	0.942
100		MLE	1.50	0.367	0.147	(0.78,2.22)	0.942
		MM	1.49	0.036	0.130	(0.78,2.20)	0.947
20	0.5	MLE	0.52	0.792	0.653	(-0.16,1.22)	0.610
		MM	0.51	0.180	0.653	(-1.07,2.09)	0.950
60		MLE	0.50	0.452	0.209	(-0.18,1.19)	0.868
		MM	0.50	0.058	0.208	(-0.38,1.40)	0.957
100		MLE	0.49	0.349	0.125	(-0.20,1.16)	0.949
		MM	0.48	0.034	0.119	(-0.18,1.16)	0.950
20	-2	MLE	-2.06	0.870	0.780	(-2.83,-1.31)	0.631
		MM	-2.10	0.203	0.840	(-3.89,-0.31)	0.954
60		MLE	-2.04	0.495	0.263	(-2.79,-1.28)	0.865
		MM	-2.04	0.062	0.234	(-2.98,-1.09)	0.948
100		MLE	-2.01	0.382	0.148	(-2.78,-1.26)	0.941
		MM	-2.03	0.036	0.132	(-2.74,-1.33)	0.952

2.6 Generalized UTPD

We have seen in the definition of UTPD that the domain is $[0, 1]$. Now we construct a generalization of UTPD into a distribution defined on a finite interval (a, b). Let X be a continuous random variable defined on (a, b). The pdf of X is given by,

$$f(x) = \frac{\theta}{e^{\theta b} - e^{\theta a}} e^{\theta x}, \quad a < x < b, \theta \neq 0.$$

When $\theta = 0$, it becomes the uniform distribution defined on (a,b).

Properties

1. The r^{th} raw moment is given by,

$$\begin{aligned} E(X^r) \quad &= C\{\frac{b^k e^{\theta b} - a^k e^{\theta a}}{\theta} - \frac{k}{\theta^2}(b^{k-1}e^{\theta b} - a^{k-1}e^{\theta a}) + \frac{(k-1)k}{\theta^3}(b^{k-2}e^{\theta b} - a^{k-2}e^{\theta a}) \\ &\quad +... + \frac{(-1)^k(1.2.3...k)}{\theta^{k+1}}(e^{\theta b} - e^{\theta a})\}, \end{aligned}$$

   where $C = \frac{\theta}{(e^{\theta b} - e^{\theta a})}$.

2. Mean= $\frac{be^{\theta b} - ae^{\theta a}}{\left(e^{\theta b} - e^{\theta a}\right)} - \frac{1}{\theta}$.

3. Variance= $\frac{\theta}{\left(e^{\theta b} - e^{\theta a}\right)} \left\{\frac{b^2 e^{\theta b} - a^2 e^{\theta a}}{\theta} - \frac{2}{\theta^2}\left(be^{\theta b} - ae^{\theta a}\right) + \frac{2}{\theta^3}\left(e^{\theta b} - e^{\theta a}\right)\right\} - \left(\frac{be^{\theta b} - ae^{\theta a}}{\left(e^{\theta b} - e^{\theta a}\right)} - \frac{1}{\theta}\right)^2$.

4. The hazard rate function $h(x) = \frac{\theta e^{\theta x}}{e^{\theta b} - e^{\theta x}}$.

5. The Mean residual life function, (MRL), $\mu(t) = \frac{1}{e^{\theta b} - e^{\theta a} - e^{\theta t} - e^{\theta a}} \left\{e^{\theta b}(b - t - 1) + e^{tb}\right\}$.

6. The quantile function is $x = \frac{1}{\theta} \ln \left[e^{\theta a} (1-u) + u e^{\theta b} \right]$.

More interesting features are considered for further studies.

In the following section, the application of the distribution is demonstrated by fitting UTPD to four distinct datasets. A comparison between the UTPD and power function distributions is carried out for each of these datasets, as mentioned in Section 2.2.

2.7 Real data analysis

Dataset 1: The data set originates from the solar incentive program of the Los Angeles Department of Water and Power (LADWP), which provides incentives to help mitigate the expenses associated with the installation of solar rooftop systems for residential and commercial properties in Los Angeles. This metric measures the Net Energy Metering (NEM) installed capacity (kilowatts), which is available at https://catalog.data.gov/dataset. The data consists of observations from the years 2016 to 2018, which depicts the application of UTPD in time series as well as physics. The data can be transformed using (2.3.3) to transform into the range $[0, 1]$. Now we have made an attempt to fit the power function and UTPD to this transformed data. According to Table 2.3, the Kolmogorov-Smirnov (K-S) distance measure and p-value indicate that both the UTPD and power function distributions are good fit for this data set. The p-value being greater than 0.05 confirms that both distributions are good approximations. However, upon comparison, the UTPD is seemed to

be a better fit as it exhibits a smaller K-S distance and a higher p-value compared to the power function distribution.

Dataset 2: Data set 2 is the total tax and non-tax revenue of Egypt from 2002 to 2018, which is available at https://stats.oecd.org, and these are time series observations from the financial sector. Transform the data using (2.3.3) and Based on the K-S distance and p-value provided in Table 2.3, we find that both the power function and UTPD are appropriate fits for this data. However, considering these two measures, it becomes evident that the UTPD is a better fit compared to the power function distribution for this dataset.

Dataset 3: Now we consider another time series to illustrate the applications of UTPD. This is a set of observations of the Japan consumer confidence index from January 2014 to March 2021. (Ref: https://stats.oecd.org). Upon applying suitable transformation, both the power function and UTPD are found to be suitable fits for the transformed data, as indicated by the values in Table 2.3. However, due to the higher p-value and lower K-S distance associated with the UTPD compared to the power function, it becomes evident that the UTPD provides a good fit for the transformed data.

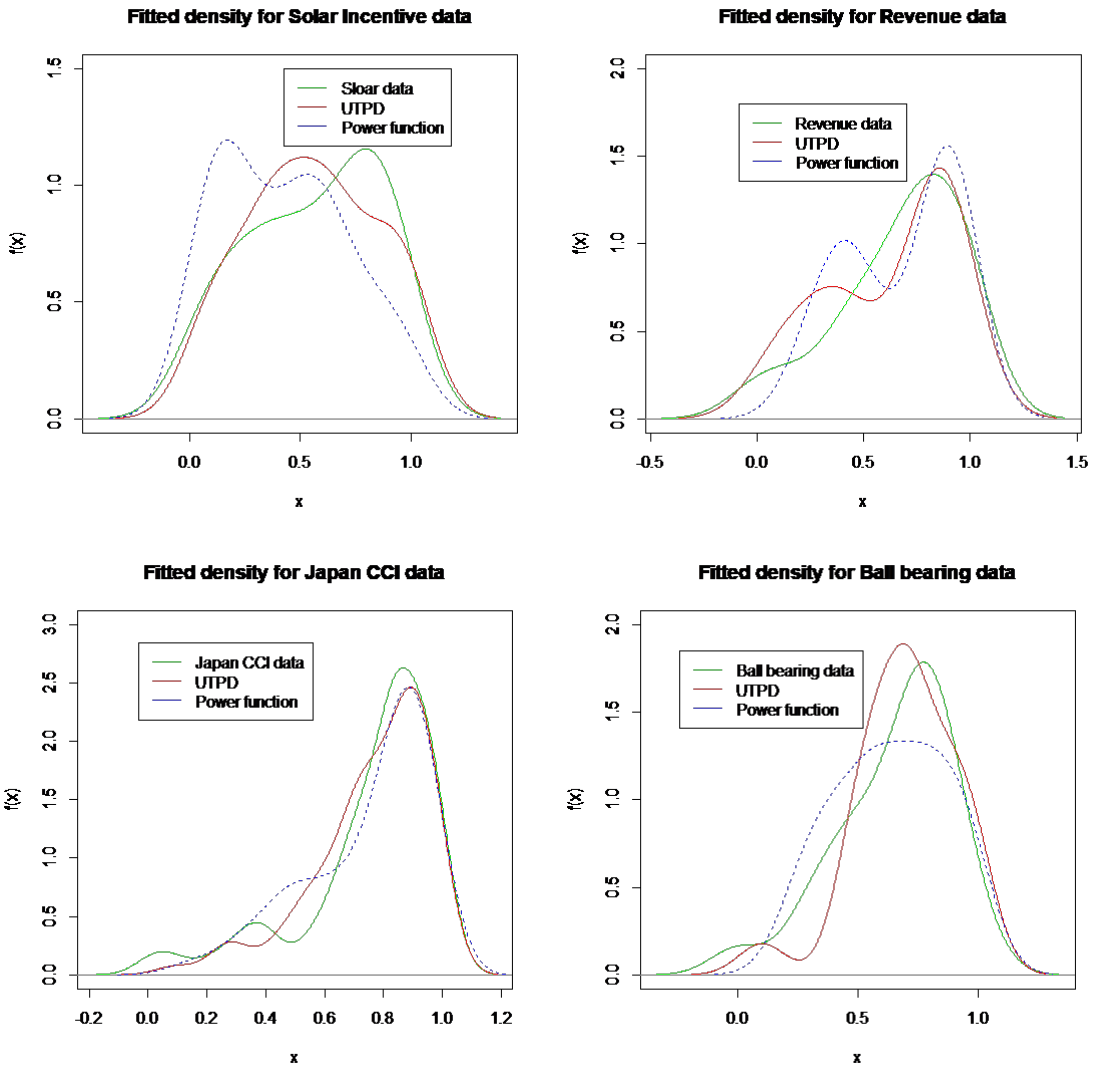

Figure 2.5: Density plots of fitted data sets

Dataset 4: The fourth data set we consider is the ball bearing data taken from Lawless (2003) to employ it in the engineering field. The data are the number of million revolutions before failure for each of the 23 ball bearings in the life test, and they are 17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12,

55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, and 173.40. The data can be transformed to the interval $[0, 1]$ using (2.3.3). By examining the K-S distance and p-value provided in Table 2.3, it becomes evident that the UTPD is a more suitable fit for this transformed data compared to the power function distribution. The densities of the original data sets together with the fitted densities plotted in Figure 2.5 reveal that UTPD is a good fit for all the data sets considered.

Table 2.3: Fitting of real data sets

	Distribution	Parameter	K-S distance	p-value
Dataset 1	UTPD	0.75	0.0967	0.9991
	Power function	1.28	0.1290	0.9634
Dataset 2	UTPD	2.29	0.1176	0.9999
	Power function	2.08	0.1764	0.7631
Dataset 3	UTPD	3.64	0.0919	0.8585
	Power function	3.04	0.1149	0.6164
Dataset 4	UTPD	1.91	0.1421	0.7657
	Power function	1.86	0.2173	0.6487

2.8 Conclusion

In this chapter, we have studied in detail the UTPD as the solution of a first order differential equation and derived the same from the truncated uniform distribution.

Comparisons with some well known distributions are done. The expressions for moments, distributions of the order statistics etc. are further derived. Some transformed distributions are studied. Some of the estimation procedures for the parameters are discussed. The newly constructed distribution is applied on real data.

CHAPTER 3

AUTOREGRESSIVE UNIFORM TRUNCATED POISSON PROCESS

3.1 Introduction

The time series data observed over time exhibit deviations from linear patterns, showing various distinctive features such as cycles, asymmetries, bursts, sudden changes, chaotic behavior, critical points, and varying levels of volatility. Conventional time series models often fail to capture the intricate characteristics present in economic and financial data, disregarding essential elements. In order to facilitate a more comprehensive understanding of the changes that have taken place in the financial and economic systems, including notable structural and behavioral changes, it is appropriate to investigate the necessity of using multiple time series models to explain

empirical data at different time intervals. Therefore, non-linear time series models are employed to analyze and interpret these variations within the data. Different types of non-linear time series models are discussed in Chapter 1.

Probability distributions defined on the interval [0, 1] are essential for many applications in machine learning, risk management, neural networks, etc. In some situations, certain aspects of organisms may exhibit relatively uniform patterns, such as genetic uniformity in clonal organisms, uniformity in certain traits within species, social insects, and ecosocial social behavior. Also, for the variables in $[0, 1]$, like percentages or fractions, only a few studies exist on regression or time series models. This motivates us to examine the non-linear time series applications of UTPD that are discussed in Chapter 2. So this chapter is intended to construct a non-linear AR model with UTPD mentioned in (2.2.2) as the marginal distribution.

This chapter is organized as follows: In Section 3.2, the uniform truncated Poisson autoregressive process is introduced. It's properties are studied in Section 3.3. The parameters involved in the proposed model is estimated using different methods, and simulation studies are also done in Section 3.4. Section 3.5 is devoted to the real data analysis, followed by a concluding section.

3.2 Uniform truncated Poisson autoregressive process of order 1

Autoregressive modelling is one of the techniques used for time-series analysis. An autoregressive model is a time-series model that describes how a particular variable's past values influence its current value. Autoregressive processes are, as their name suggests, regressions on themselves. Autoregressive models are based on the idea that past events can help us predict future events. The study of the AR model has a long history, and several models are available in the literature. We can see some non-linear AR models, like the Beta-Gamma AR(1) process in Lewis et al. (1989). As it's generalization, a stationary Beta-Gamma AR(2) process with Gamma(k, β) as marginal is described in Ristic (2005). A new stationary AR(1) (NUAR(1)) model with marginally continuous uniform (0,1) distribution is presented in Ristic and Popvic (2000a). An estimation of the unknown parameters of the NUAR(1) process was done by Ristic and Popvic (2000b). The uniform autoregressive process of second order (UAR(2)) can be seen in Ristic and Popvic (2002). In this section, we develop an AR(1) model with the UTPD mentioned in equation (2.2.2) as the marginal distribution.

The model is introduced as follows: Let $\{X_n\}$ be a sequence of stationary random variables defined as

$$X_n = \begin{cases} \varepsilon_n & \text{with probability } e^{-|\theta|} \\ e^{-|\theta|}X_{n-1} + \varepsilon_n & \text{with probability } 1 - e^{-|\theta|} \end{cases} \tag{3.2.1}$$

where $\{\varepsilon_n\}$ is a sequence of iid random variables such that X_{n-1} is independent of ε_n, for $n \geq 1$. We can derive the characteristic function of ε_n using (3.2.1). Taking the characteristic function of X_n in (3.2.1) and solving we get,

$$\phi_{\varepsilon_n}(t) = \frac{\phi_{X_n}(t)}{\left(e^{-|\theta|} + (1 - e^{-|\theta|})\,\phi_{X_{n-1}}(e^{-|\theta|}t)\right)}$$

In the model (3.2.1), we assume that $\{X_n\}$ is a stationary sequence following the distribution UTPD. Hence, we call this process as uniform truncated Poisson autoregressive process of order 1 (UTPAR(1)). Now, substituting the characteristic function of $\phi_{X_n}(t)$ given in (2.3.1),

$$\phi_{\varepsilon_n}(t) = \frac{\theta\left(\theta + ie^{-|\theta|}t\right)\left(e^{\theta+it} - 1\right)}{(\theta + it)\left[e^{-|\theta|}\left(e^{\theta} - 1\right)\left(\theta + ie^{-|\theta|}t\right) + \theta\left(1 - e^{-|\theta|}\right)\left(e^{\theta+ie^{-|\theta|}t} - 1\right)\right]}. \tag{3.2.2}$$

Note that the complex structure of the characteristic function makes it difficult to find the pdf of the random variable ε_n using the inversion formula.

We will discuss the properties of the UTPAR(1) process in the next section.

3.3 Properties

Then using the well known result of finding the mean from the characteristic function, the mean of the innovation random variable is $E(\varepsilon_n) = \frac{1}{i}\phi'_{\varepsilon_n}(0)$. After simplification, it is

$$E(\varepsilon_n) = \left[\frac{e^{\theta}(\theta-1)+1}{\theta(e^{\theta}-1)}\right]\left(e^{-2|\theta|} - e^{-|\theta|} + 1\right). \tag{3.3.1}$$

Similarly, from the characteristic function, the variance of the innovation random variable is derived as

$$\begin{aligned} Var(\varepsilon_n) \quad &= \left[\frac{e^{\theta}}{e^{\theta}-1}\left(1 - \frac{2}{\theta} + \frac{2}{\theta^2}\right) - \frac{2}{\theta^3(e^{\theta}-1)}\right]\left[1 - e^{-2|\theta|}\left(1 - e^{-|\theta|}\right)\right] \\ &- \left(\frac{e^{\theta}(\theta-1)+1}{\theta(e^{\theta}-1)}\right)^2 \left(1 - e^{-|\theta|} + e^{-2|\theta|}\right)\left(e^{-|\theta|} - e^{-2|\theta|} + 1\right) \end{aligned}$$

The kth order auto correlation,

$$\rho(k) = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)Var(X_{t-k})}}$$

is obtained as

$$\begin{aligned} \rho(k) \quad &= \frac{\left(1 - e^{-|\theta|} + e^{-2|\theta|}\right)\mu\left(1 + e^{-|\theta|}\left(1 - e^{-|\theta|}\right) + ... + e^{-(k-1)|\theta|}\left(1 - e^{-|\theta|}\right)^{k-1}\right)}{\delta - \mu} \\ &+ \frac{e^{-k|\theta|}\left(1 - e^{-|\theta|}\right)^k \delta - \mu}{\delta - \mu}, \end{aligned}$$

where

$$\mu = \left[\frac{e^{\theta}(\theta-1)}{\theta\left(e^{\theta}-1\right)}\right]^2,$$
$$\delta = \left[\frac{e^{\theta}}{e^{\theta}-1}\left(1-\frac{2}{\theta}+\frac{2}{\theta^2}\right)-\frac{2}{\theta^3\left(e^{\theta}-1\right)}\right].$$

Now we look at the conditional properties. The conditional expectation,

$$\begin{aligned} E\left(X_n/X_{n-1}\right) &= e^{-|\theta|}E\left(\varepsilon_n\right)+\left(1-e^{-|\theta|}\right)\left(e^{-|\theta|}x+E\left(\varepsilon_n\right)\right) \\ &= E\left(\varepsilon_n\right)+e^{-|\theta|}\left(1-e^{-|\theta|}\right)x \\ &= \left[\frac{e^{\theta}\left(\theta-1\right)+1}{\theta\left(e^{\theta}-1\right)}\right]\left(e^{-2|\theta|}-e^{-|\theta|}+1\right)+e^{-|\theta|}\left(1-e^{-|\theta|}\right)x. \end{aligned} \quad (3.3.2)$$

which indicates a consistent pattern or behavior between X_n and X_{n-1}.

The conditional variance can be obtained by using

$$Var\left(X_n/X_{n-1}=x\right) = E\left(X_n^2/X_{n-1}=x\right)-\left(E\left(X_n/X_{n-1}=x\right)\right)^2$$

From (3.2.1),

$$\begin{aligned} E\left(X_n^2/X_{n-1}=x\right) \quad &= \delta\left(1-e^{-2|\theta|}+e^{-3|\theta|}\right)-\mu 2e^{-|\theta|}\left(1-e^{-|\theta|}\right)\left(1-e^{-|\theta|}+e^{-2|\theta|}\right) \\ &\quad +2e^{-|\theta|}\left(1-e^{-|\theta|}\right)\left(e^{-2|\theta|}-e^{-|\theta|}+1\right)\left(\frac{e^{\theta}\left(\theta-1\right)+1}{\theta\left(e^{\theta}-1\right)}\right)x \\ &\quad -\left(1-e^{-|\theta|}\right)e^{-2|\theta|}x^2 \end{aligned}$$

Then,

$$Var(X_n/X_{n-1}=x) = \mu\left(e^{-2|\theta|} - 2e^{-3|\theta|} + e^{-4|\theta|}\right) + \delta\left(1 - e^{-2|\theta|} + e^{-3|\theta|}\right) + \left(e^{-3|\theta|} - e^{-4|\theta|}\right)x^2. \tag{3.3.3}$$

In the next section, we estimate the parameter of the proposed model. A major problem in the analysis of non-Gaussian time series is to find the distribution of the innovation random variable ε_n for a specified marginal distribution of $\{X_n\}$. Due to the structure of (3.2.2), the pdf of the $\{\varepsilon_n\}$ is difficult to identify. It is obvious that the likelihood function for the sequence $\{X_n\}$ do not have a closed-form expression, and hence the ML estimation method fails. We propose the conditional least squares estimation (CLSE) method introduced by Klimko and Nelson (1978) and the Gaussian estimation method for estimating the parameters.

3.4 Estimation and Simulation

3.4.1 Conditional least square estimation

One of the key issues in modelling non-Gaussian time series is parameter estimation in the model. If the innovation random variable has a closed-form density, the maximum likelihood technique of estimation can be utilized. The innovation random variables typically do not have closed-form densities in non-Gaussian AR models. For the proposed AR(1) model, we give a thorough examination of CLSE. A dependent

observation estimation technique called CLSE is based on the idea of minimizing a sum of squared deviations from conditional expectations. Here, we provide a quick explanation of the conditional least squares approach created by Klimko and Nelson (1978).

Let $\{X_t, t = 1, 2, ...\}$ be a stochastic process defined on a probability space (ω, F, P_θ), whose distribution depends on an unknown parameter vector $\theta = (\theta_1, \theta_2, ...\theta_p)'$. Let $\{F_t\}_{t=1}^{\infty}$ denote a sequence of sub-sigma fields with F_{t-1} generated by an arbitrary subset of $\{X_1, X_2, ..., X_{t-1}\}, t > 1$. Then the CLS estimator of the parameters is obtained by minimizing the conditional sum of squares.

$$Q_t(\theta) = \sum_{t=1}^{n} [x_t - g(\theta; F_{t-1})]^2 .$$

with respect to the parameter vector $\theta = (\theta_1, \theta_2, ...\theta_p)'$; $g(\theta; F_{t-1}) = E(X_t/F_{t-1})$. Then the estimates of the parameters are obtained by solving the least square equations given below.

$$\frac{\partial Q_t(\theta)}{\partial \theta_i} = 0, \qquad i = 1, 2, ...p$$

Under a set of regularity conditions, CLS estimators are strongly consistent and asymptotically jointly normally distributed. The assumptions made about the application of strong laws, central limit theorems, and iterated logarithms to sums of dependent variables. Stout (1974), McLeish (1974), and Heyde and Scott (1973) referred to a wide variety of conditions under which these assumptions hold. We

employ this method to estimate the θ of our UTPD AR(1) model. Let $(x_1, x_2, ..., x_T)$ be a realization from the stationary UTPAR(1) sequence. The CLS estimate of the parameter θ is obtained by minimizing the conditional error sum of squares. $Q_t(\theta) = \sum_{t=1}^{T} (x_t - g(\theta, x_{t-1}))^2$, where $g(.)$ is the conditional expectation. Then using (3.3.2),

$$Q_t(\theta) = \sum_{t=2}^{T} \left[x_t - \left[\frac{e^{\theta}(\theta - 1)}{\theta(e^{\theta} - 1)} \right] \left(e^{-2|\theta|} - e^{-|\theta|} + 1 \right) - e^{-|\theta|} \left(1 - e^{-|\theta|} \right) x \right]^2 . \quad (3.4.1)$$

By finding the derivative of $Q_t(\theta)$ with respect to θ and equating to zero, we have a non-linear equation, that can be solved numerically to estimate the parameters.

3.4.2 Gaussian estimation method

Whittle (1962) introduced a method that employed the Gaussian likelihood function as the baseline distribution for estimation. Subsequently, Crowder (1985) applied this method to analyze correlated binomial data. Al-Nachawati et al (1997) and Alwasel et al. (1998) extended this estimation procedure to the context of a first order autoregressive process. Despite its approximate nature, this method yields good estimations for our model. The conditional maximum likelihood function is given by

$$L = f(x_1) \prod_{t=2}^{n} f(x_t / x_{t-1}) . \quad (3.4.2)$$

Here $f(x_t/x_{t-1})$ and $f(x_1)$ are the conditional and marginal probability functions of X_t/X_{t-1} and X_t respectively. We assume Gaussian pdf for $f(x_1)$ and $f(x_t/x_{t-1})$ with conditional mean and conditional variance as the parameters. Then the log-likelihood function can be written as

$$log(L) = nlog\frac{1}{\sqrt{2\pi}} - \frac{1}{2}\sum_{t=2}^{n}\left(log\left(\sigma_{x_{t-1}}\right) + \frac{\left(x_t - m_{x_{t-1}}\right)^2}{\sigma^2_{x_{t-1}}}\right), \tag{3.4.3}$$

$m_{x_{t-1}}$ is the conditional mean and $\sigma^2_{x_{t-1}}$ is the conditional variance. The Gaussian estimators are thus obtained by maximizing the above non-linear equation. But the estimator of the parameter cannot be written in explicit form. These equations can be solved numerically. Crowder (1985) pointed out that under the Gaussian method of estimation of the parameter θ, $\sqrt{n}\left(\hat{\theta} - \theta\right)$ is asymptotically normally distributed with mean 0 and asymptotic variance $[J(\theta)]^{-1}$, where $J(\theta)$ is the conditional expectation information matrix. An approximation of the same using the observed conditional information matrix is described in Bakouch and Popvic (2016). To check the validity of the estimates, we conducted a simulation study, and the MSE is used for comparison purposes.

3.4.3 Simulation

For checking the validity of the model, we simulated 1000 samples of sizes 20, 50, and 100 for different values of the parameter θ for selecting the best method of estimation. The estimates of θ and MSE values are shown in Table 3.1. R-programming is used

for the simulation work and related computations. It can be seen that the MSE is decreasing when the sample size is increasing. From the table, it is clear that both the CLS method and the Gaussian method are equally good for estimation based on the MSE.

3.5 Real Data Analysis

In this section, we illustrate the application of the model using a real-time data set. An important point to be noted is that our marginal distribution is UTPD, and its range is in $[0, 1]$. So we convert the data sets to $[0, 1]$ using the transformation mentioned in (2.3.3). The model is fitted for these transformed data set.

Data set: The data set is taken from the website https://data.worldbank.org. We considered the amount of carbon dioxide emissions in Japan per capita (metric tons) from 1990 to 2020. The data is transformed into $[0, 1]$, using min-max transformation. The ACF plot in Figure 3.1 and PACF plot in Figure 3.2 reveal that the AR(1) model is a good fit, and UTPD also fits well as a distribution with $\hat{\theta}$=0.8940. It is confirmed using the K-S test with a p value of 0.2558 and a test statistic value of 0.4135. The actual and predicted values are plotted in Figure 3.3. The p-value of the Ljung-Box test is 0.1041, implying that the errors are independent. The histogram of errors is plotted in Figure 3.4.

Table 3.1: Estimates of parameter

Sample size(n)	Parameterθ	Method	Estimate ($\hat{\theta}$)	MSE
20	3	CLS	3.0883	0.7008
		Gaussian Estimation	3.5036	0.2535
50		CLS	3.4675	0.4807
		Gaussian Estimation	3.2303	0.0530
100		CLS	2.8395	0.2185
		Gaussian Estimation	2.8921	0.0116
20	2	CLS	2.4689	0.3999
		Gaussian Estimation	2.3553	0.4156
50		CLS	2.3859	0.1489
		Gaussian Estimation	2.8259	0.0303
100		CLS	1.8695	0.0170
		Gaussian Estimation	1.9468	0.0109
20	0.5	CLS	1.9730	1.2060
		Gaussian Estimation	1.5295	1.0598
50		CLS	1.4392	0.8819
		Gaussian Estimation	0.2505	0.0439
100		CLS	0.4431	0.2496
		Gaussian Estimation	0.243	0.0321
20	-1.5	CLS	-3.15	2.7500
		Gaussian Estimation	-2.7378	0.8600
50		CLS	-2.2397	2.5059
		Gaussian Estimation	-2.3608	0.3890
100		CLS	-1.8070	0.0492
		Gaussian Estimation	-1.786	0.2860

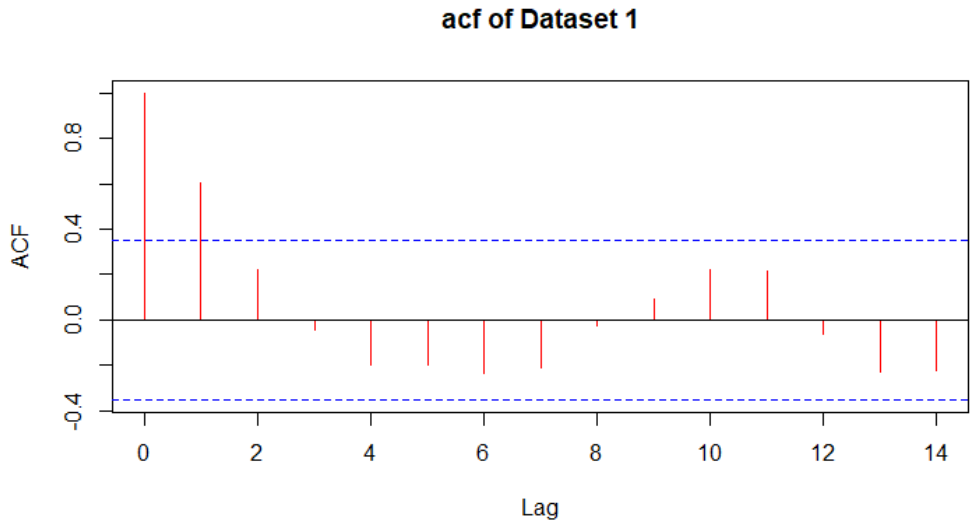

Figure 3.1: ACF of Japan CO2 data

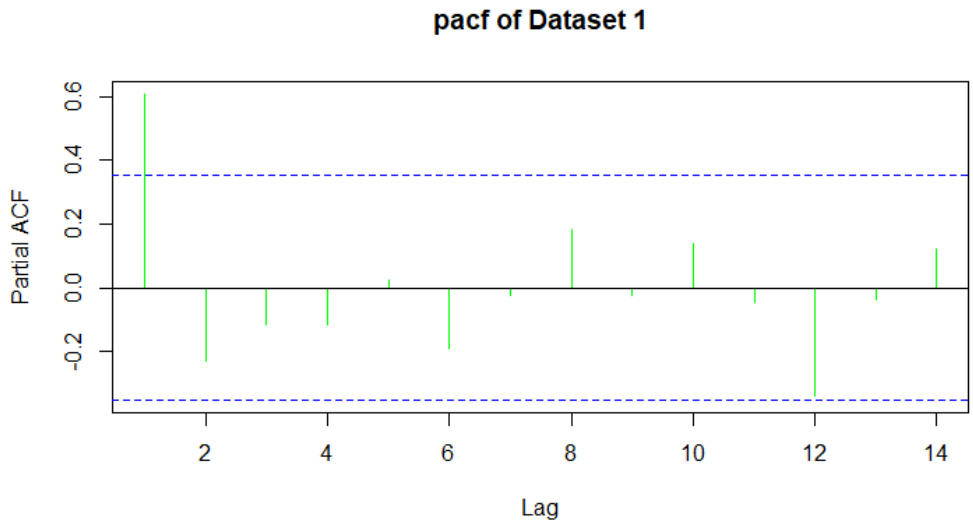

Figure 3.2: PACF of Japan CO2 data

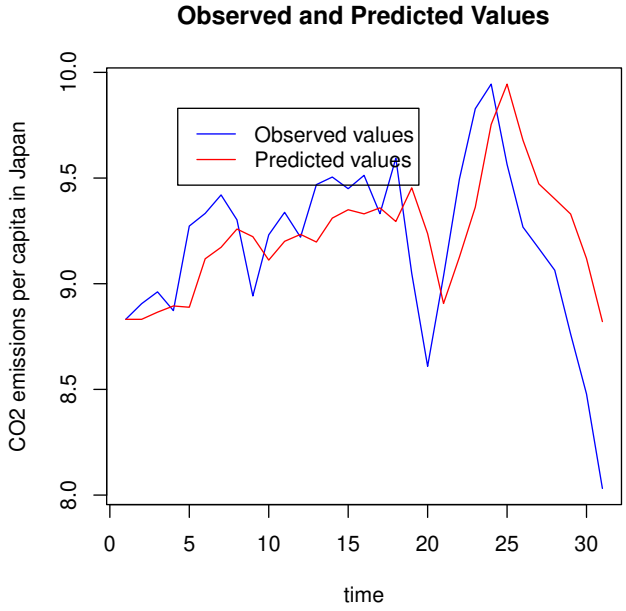

Figure 3.3: Actual and predicted values of Japan Co2 data

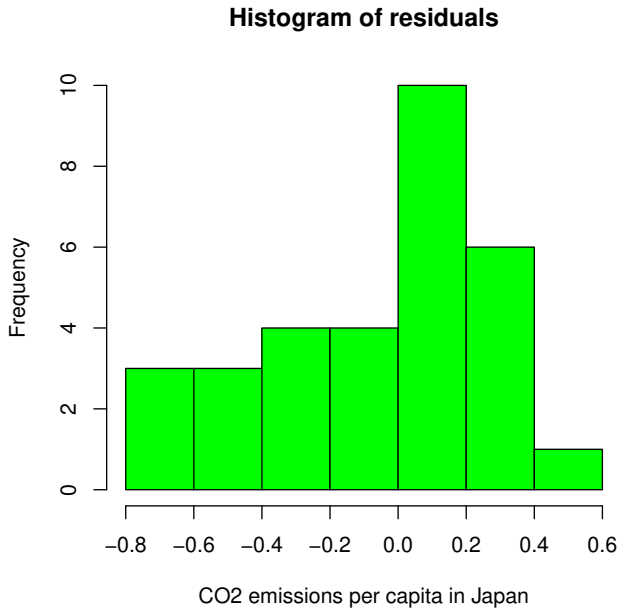

Figure 3.4: Plot of errors of Co2 Japan data

3.6 Conclusion

In this chapter, we constructed the time series models in mixture form, with UTPD as the marginal. Distributional properties and classical procedures for the estimation of the associated parameters of the proposed process are discussed. The simulation studies are done. Applications of the UTPAR(1) process are illustrated with a real data set, and forecasting is done accordingly.

CHAPTER 4

UNIFORM TRUNCATED POISSON AUTOREGRESSIVE CONDITIONAL DURATION PROCESS

4.1 Introduction

Time series data collected at an extremely fine scale is referred to as high-frequency data. High-frequency data is now widely used in financial analysis and high-frequency trading. As a result of advances in computational power in recent decades, high-frequency data can be collected accurately and efficiently for analysis. These data sets were created by accumulating tick by tick market data, where every single event is represented by a tick or one logical unit of information. High-frequency data collec-

tions typically contain a large amount of data and allow for high statistical precision. As a result, the observations are irregularly spaced in time, which is a common feature of this data. Recent models from the market microstructure literature claim that time should be modelled because it can convey a lot of information. To analyse the dynamic behaviour of financial duration data, Engle and Russel (1998) proposed the ACD model as mentioned in (1.6.13). They use the exponential and Weibull distributions to model the conditional distribution of an interval given past data. Since its inception, the ACD model and its various extensions have established themselves as a leading tool for modelling the behaviour of irregularly spaced financial data.

Similar ACD models can be seen in Bauwens and Giot (2000), Grammig and Maurer (2000), Bhatti and Chad (2010), Pacurar (2008), Zhang et al. (2001) etc. In this chapter, we give the ACD model corresponding to the newly introduced distribution where the duration is lying in $[0, 1]$ or by suitably transforming the long duration to (0,1).

The Box and Jenkins method emphasises stationary ARMA models with Gaussian innovations in time series analysis. Non-Gaussian distributions better explain the majority of real-life situations. When dealing with this type of data, some transformations are performed to ensure that the changes result in a normal distribution. However, in several cases, the transformation method produces poor results. As a result, many non-Gaussian models have been developed over the last four decades. This is evident from the research of Nelson and Granger (1979), Weiss (1977), and Yakowitz (1973). Then again, the majority of the data sets we examine fall within

finite ranges. In the physical sciences, there are several data sets in $(0, 1)$ that are employed in experimental study. The literary analysis helped us identify the works of Altawil (2019) and Hassan et al. (2020). In these sectors, the min-max scaling approach indicated in (2.3.3) is the most often used form of rescaling.

In this chapter, we made an attempt to introduce an autoregressive conditional duration model (ACD) with the UTPD mentioned in (2.2.2) as the marginal, where the duration is in $[0, 1]$. In Section 4.2, uniform truncated autoregressive conditional process is introduced and its properties are studied. Estimation and simulation studies are discussed in Section 4.3, and real data applications are illustrated in Section 4.4. Concluding remarks are given in Section 4.5.

4.2 Uniform truncated Poisson autoregressive conditional duration process

Let X follow UTPD with pdf given in (2.2.2) with mean,

$$E(X) = \frac{e^{\theta}(\theta - 1) + 1}{\theta(e^{\theta} - 1)}.$$

Consider the transformation,

$$\epsilon = \frac{X\theta\left(e^{\theta} - 1\right)}{e^{\theta}(\theta - 1) + 1}. \tag{4.2.1}$$

The first and second-order moments of the transformed random variable respectively are,

$$E(\epsilon) = 1 \tag{4.2.2}$$

$$E(\epsilon^2) = \frac{(e^\theta - 1)[e^\theta(\theta^2 - 2\theta + 2) - 2)]}{(e^\theta(\theta - 1) + 1)^2}, \tag{4.2.3}$$

and hence it has a unit mean with pdf

$$f_{\varepsilon_i}(\epsilon) = \frac{e^\theta(\theta - 1) + 1}{(e^\theta - 1)^2} e^{\varepsilon\left[\frac{e^\theta(\theta-1)+1}{e^\theta-1}\right]} \quad ; \qquad 0 < \epsilon < \frac{\theta\left(e^\theta - 1\right)}{e^\theta(\theta - 1) + 1}. \tag{4.2.4}$$

Similar to the basic ACD model in (1.6.11) and (1.6.13), we define a conditional autoregressive duration model as given below.

Let

$$X_i = \psi_i \epsilon_i, \tag{4.2.5}$$

where

$$\psi_i = \omega + \alpha X_{i-1} + \beta\psi_{i-1}, \qquad i = 1, 2, 3, ...n.$$

Here we assume that $\{\epsilon_i\}$ follows (4.2.4). Then the conditional pdf of X_i given ψ_i is

$$f_{\left(\frac{X_i}{\psi_i}\right)}(x_i) = \frac{1}{\psi_i} f_{\varepsilon_i}\left(\frac{x_i}{\psi_i}\right) = \frac{e^\theta(\theta - 1) + 1}{(e^\theta - 1)^2} e^{\left[\frac{e^\theta(\theta-1)+1}{e^\theta-1}\right]\frac{x_i}{\psi_i}} \frac{1}{\psi_i} \quad ; \quad 0 < x_i < \frac{\psi_i\theta\left(e^\theta - 1\right)}{e^\theta(\theta - 1) + 1}. \tag{4.2.6}$$

We call the process defined in (4.2.5) as the uniform truncated Poisson autoregressive conditional duration process of order (1,1) (UTPACD(1,1)). Analytical properties

such as moments, autocorrelations will be investigated next, allowing us to gain a thorough understanding of the model.

4.2.1 Properties

Here we discuss the properties of the UTPACD(1,1) process.

1. **Conditional Mean:** Assuming stationarity, the conditional mean of the UTPACD process is given by

$$\mu_x = E\left(\frac{X_i}{\psi_i}\right) = \frac{\omega}{1-\alpha-\beta}. \tag{4.2.7}$$

2. **Second order moments and variance:**

   For weak stationarity of $\{X_i\}$, the condition $0 \leq \alpha + \beta < 1$ must be satisfied. Second order moment,

$$E\left(X_i^2\right) = E\left[E\left(\psi_i^2 \varepsilon_i^2 / F_{i-1}\right)\right]. \tag{4.2.8}$$

   Now using weak stationarity of ψ_i and X_i, we have

$$\begin{aligned} E\left(\psi_i^2\right) &= E\left(\omega + \alpha X_{i-1} + \beta\psi_{i-1}\right)^2 \\ &= E\left(\omega^2 + \alpha^2 X_{i-1}^2 + \beta^2\psi_{i-1}^2 + 2\alpha\omega X_{i-1} + 2\beta\omega\psi_{i-1} + 2\alpha\beta X_{i-1}\psi_{i-1}\right) \\ &= \omega^2 + \alpha^2 E\left(X_{i-1}^2\right) + \beta^2 E\left(\psi_{i-1}^2\right) + 2\alpha\omega\mu_x + 2\omega\beta\mu_x + 2\alpha\beta\mu_x^2 \\ &= \frac{\mu_x^2\left[1-(\alpha^2+\beta^2)\right]}{1-\alpha^2 E[\epsilon_i^2]-\beta^2}. \end{aligned}$$

Therefore

$$E\left(X_i^2\right) = \frac{\mu_x^2\left[1-(\alpha^2+\beta^2)\right]}{1-\alpha^2 E[\epsilon_i^2]-\beta^2} E[\epsilon_i^2], \tag{4.2.9}$$

3. **Variance:**

Using $Var\left(X_i\right) = E\left(X_i^2\right) - \left(E\left(X_i\right)\right)^2$ and (4.2.9), we have the unconditional variance,

$$\begin{aligned} Var\left(X_i\right) &= \frac{\mu_x^2\left[1-(\alpha^2+\beta^2)\right]}{1-\alpha^2 E[\epsilon_i^2]-\beta^2} E[\epsilon_i^2] - \mu_x^2 \\ &= \frac{\mu_x^2(1-\beta)^2[E(\epsilon_i^2)-1]}{1-\alpha^2 E(\epsilon_i^2)-\beta^2} \end{aligned} \tag{4.2.10}$$

4. **Autocorrelation function:**

The recurrence relation for k^{th} order ACF is $(k>1)$

$$\rho_k = (\alpha+\beta)\,\rho_{k-1}. \tag{4.2.11}$$

Proof:

The k^{th} order autocovariance function of $\{X_i\}$ is,

$$\begin{aligned} \gamma_k &= Cov\left(X_i, X_{i-k}\right) \\ &= Cov\left(\psi_i, X_{i-k}\right) \\ &= Cov\left(\omega + \alpha X_{i-1} + \beta\psi_{i-1}, X_{i-k}\right) \\ &= \alpha Cov\left(X_{i-1}, X_{i-k}\right) + \beta Cov\left(\psi_{i-1}, X_{i-k}\right) \\ &= (\alpha+\beta)\,\gamma_{k-1} \end{aligned}$$

Now the ACF of lag k, when $k > 1$ is

$$\rho_k = (\alpha + \beta)\, \rho_{k-1}.$$

The first-order autocovariance function of X_i is

$$\begin{aligned} \gamma_1 \quad &= Cov\,(X_i, X_{i-1}) \\ &= Cov\,(\psi_i, X_{i-1}) \\ &= Cov\,(\omega + \alpha X_{i-1} + \beta\psi_{i-1}, X_{i-1}) \\ &= \alpha Var\,(X_i) + \beta Var\,(\psi_{i-1}) \end{aligned}$$

where

$$\begin{aligned} Var\,(\psi_i) \quad &= E\left(\psi_i^2\right) - \left(E\,(\psi_i)\right)^2 \\ &= \frac{\mu_x^2\,[1 - (\alpha^2 + \beta^2)]}{1 - \alpha^2 E[\epsilon_i^2] - \beta^2} - \mu_x^2 \\ &= \frac{\alpha^2 \mu_x^2\,(E\,(\varepsilon_i^2 - 1))}{1 - \alpha^2 E\,(\varepsilon_i^2) - \beta^2} \end{aligned}$$

Therefore,

$$\rho_1 = \frac{\alpha(1 - \beta^2 + \alpha\beta)}{[1 - \beta^2]}.$$

The parameter estimation of the UTPACD(1,1) process is discussed in the following section. The simulations are also used to evaluate the precision of the estimation methods.

4.3 Estimation of parameters

The Gaussian method of estimation is proposed here for the estimation of the parameters. But, due to the complicated structure of the likelihood function, the estimation yields some problems in the estimates of one of the parameters α and β. So we use a heuristic procedure for evaluation of the initial values of these parameters.

Gaussian estimation

Let $X_1, X_2, ..., X_n$ be a sample. The likelihood function is defined as

$$L\left(X \mid \Theta\right) = f\left(x_1 \mid \Theta\right) \prod_{i=2}^{n} f_{x_i|\psi_i}\left(X_i \mid F_{i-1}; \Theta\right), \tag{4.3.1}$$

where $f\left(x_1 \mid \Theta\right)$ is the density function of the initial random variable. The conditional log-likelihood function after eliminating the density of X_1 is

$$logL = \sum_{i=2}^{n} log\left\{\frac{e^{\theta}(\theta-1)+1}{(e^{\theta}-1)^2} e^{[\frac{e^{\theta}(\theta-1)+1}{e^{\theta}-1}]\frac{x_i}{\psi_i}} \frac{1}{\psi_i}\right\}$$

$$= nlog[e^{\theta}(\theta-1)+1] - 2nlog(e^{\theta}-1) + [\frac{e^{\theta}(\theta-1)+1}{e^{\theta}-1}]$$

$$\sum_{i=2}^{n} \frac{x_i}{(\omega + \alpha X_{i-1} + \beta\psi_{i-1})} - \sum_{i=2}^{n} log(\omega + \alpha X_{i-1} + \beta\psi_{i-1}). \tag{4.3.2}$$

The MLE of the parameters are obtained by solving the following likelihood equations.

$$\frac{\partial logL}{\partial \theta} = 0 = \frac{n}{e^{\theta}(\theta-1)+1}\theta e^{\theta} - \frac{2ne^{\theta}}{e^{\theta}-1} + \sum_{i=2}^{n} \frac{x_i\left[e^{2\theta} - e^{\theta}(1+\theta)\right]}{(\omega+\alpha X_{i-1}+\beta\psi_{i-1})(e^{\theta}-1)^2}. \tag{4.3.3}$$

$$\frac{\partial logL}{\partial \omega} = 0 \Rightarrow \frac{-[e^{\theta}(\theta-1)+1]}{e^{\theta}-1}\sum_{i=2}^{n}\frac{x_i}{(\omega+\alpha X_{i-1}+\beta\psi_{i-1})^2} - \sum_{i=2}^{n}\frac{1}{(\omega+\alpha X_{i-1}+\beta\psi_{i-1})}. \tag{4.3.4}$$

$$\frac{\partial logL}{\partial \alpha} = 0 \Rightarrow \frac{-\left[e^{\theta}(\theta-1)+1\right]}{e^{\theta}-1}\sum_{i=2}^{n}\frac{X_iX_{i-1}}{(\omega+\alpha X_{i-1}+\beta\psi_{i-1})^2} - \sum_{i=2}^{n}\frac{X_{i-1}}{\omega+\alpha X_{i-1}+\beta\psi_{i-1}}. \tag{4.3.5}$$

$$\frac{\partial logL}{\partial \beta} = 0 \Rightarrow \frac{-\left[e^{\theta}(\theta-1)+1\right]}{e^{\theta}-1}\sum_{i=2}^{n}\frac{X_i\psi_{i-1}}{(\omega+\alpha X_{i-1}+\beta\psi_{i-1})^2} - \sum_{i=2}^{n}\frac{\psi_{i-1}}{\omega+\alpha X_{i-1}+\beta\psi_{i-1}}. \tag{4.3.6}$$

By solving these five equations numerically or maximising (4.3.1) we can find the estimates. For this purpose, we have used the GA function and maxLik package in R. However, we employed a heuristic least square approach using the sample values to ensure the correctness of the estimation procedure, which is described below. Here, a simple and direct strategy is employed. We first estimate the two model parameters α and β by equating the first two sample moments with the corresponding population values. These values are used only for initial value determination. Then the estimates of θ, α and β are obtained by maximising (4.3.1). The third model parameter ω is obtained using (4.2.7). This method reduces the bias and imprecision in the estimation procedure. We use GA in R for this, and dfsane and BBSolve in R for the initial value determination. The estimation methods are validated by simulating

samples of sizes 20, 50, and 100 and then repeating the simulation 100 times for the parameter values. Table 4.1 shows the various values of the parameters assumed for this purpose, as well as the corresponding estimated values. Table 4.1 also shows the computed MSE values and it is straightforward that the MSE decreases as the sample size increases.

Table 4.1: Parameter Estimates

True values are θ=3, α=0.1, β= 0.1							
Sample size	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\omega}$	MSE($\hat{\theta}$)	MSE($\hat{\alpha}$	MSE($\hat{\beta}$)
20	2.77	0.05	0.06	0.63	0.09	0.0200	0.004
50	2.84	0.07	0.08	0.60	0.064	0.0200	0.004
100	2.76	0.06	0.06	0.62	0.063	0.0003	0.003
True values are θ=3, α=0.1, β= 0.2							
20	2.91	0.06	0.06	0.63	0.063	0.0220	0.004
50	2.84	0.07	0.07	0.61	0.060	0.0210	0.004
100	2.79	0.06	0.06	0.62	0.058	0.0200	0.003
True values are θ=2, α=0.1, β= 0.3							
20	1.87	0.07	0.06	0.56	0.064	0.023	0.004
50	1.79	0.06	0.06	0.56	0.063	0.021	0.003
100	1.82	0.07	0.06	0.56	0.063	0.021	0.003
True values are θ=2, α=0.2, β= 0.1							
20	1.86	0.07	0.06	0.56	0.080	0.020	0.004
50	1.79	0.06	0.06	0.56	0.053	0.020	0.004
100	1.86	0.07	0.06	0.56	0.053	0.020	0.003
True values are θ=1, α=0.2, β= 0.3							
20	0.86	0.07	0.06	0.50	0.044	0.021	0.004
50	0.88	0.07	0.07	0.49	0.038	0.020	0.004
100	0.86	0.07	0.07	0.49	0.035	0.019	0.003
True values are θ=1, α=0.3, β= 0.4							
20	0.85	0.07	0.08	0.49	0.037	0.027	0.004
50	0.89	0.07	0.07	0.49	0.036	0.020	0.004
100	0.86	0.06	0.07	0.50	0.033	0.020	0.004

4.4 Real data analysis

Data set 1: The earthquake data for Japan, taken from the Japan Meteorological Agency website, is taken for analysis purposes. The data was accessed in November 2023. The data was from November 10, 2023, at 4.29 a.m. to November 17, 2023, at 10.55 p.m. First, we converted the duration to $[0, 1]$ scale using the min-max transformation. The transformed duration follows UTPD, which is clear from the p-value of 0.6882 of the K-S test. We got the parameter estimates as $\hat{\theta}$= -4.45, $\hat{\alpha}$=0.1 and $\hat{\beta}$=0.56, and $\hat{\omega} = 0.02$. The density of the estimated durations is superimposed on the histogram of the original durations in Figure 4.1. The corresponding time series plot is displayed in Figure 4.2. In Figure 4.3, errors are plotted.

Data set 2: To demonstrate the use of ACD models with UTPD, we have created a hypothetical data set in this example. The generated data set has a size of 100 and is on $(0, 1)$. The procedure is repeated as in the case of data set 1 with $\theta = 4$, $\alpha = 0.3$, $\beta = 0.4$, and $\omega = 0.22$. It is shown in Figures 4.4 and 4.5, where the estimated values are compared with original values showing the density plot and line plot, respectively.

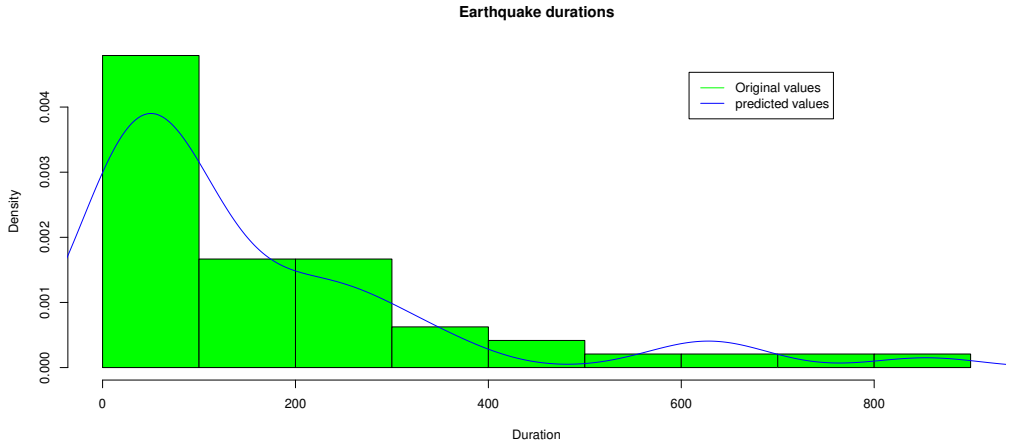

Figure 4.1: Fitted density plot of earthquake data

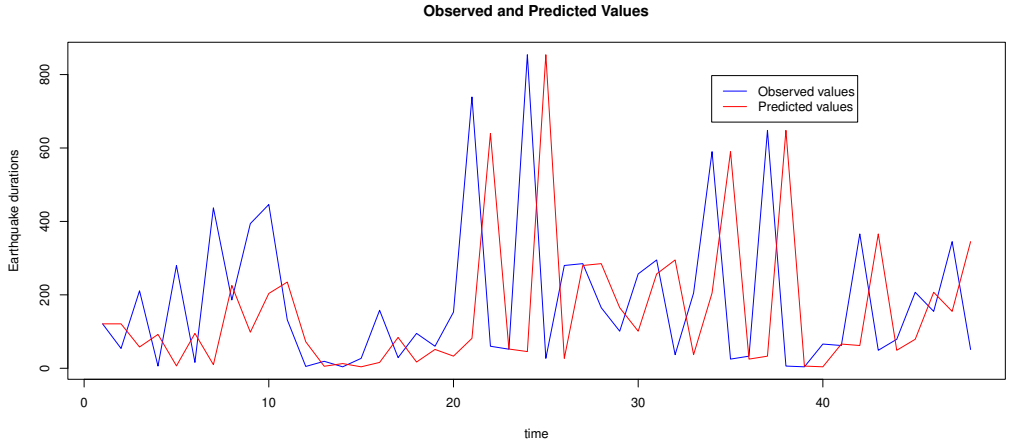

Figure 4.2: Observed and predicted duration values

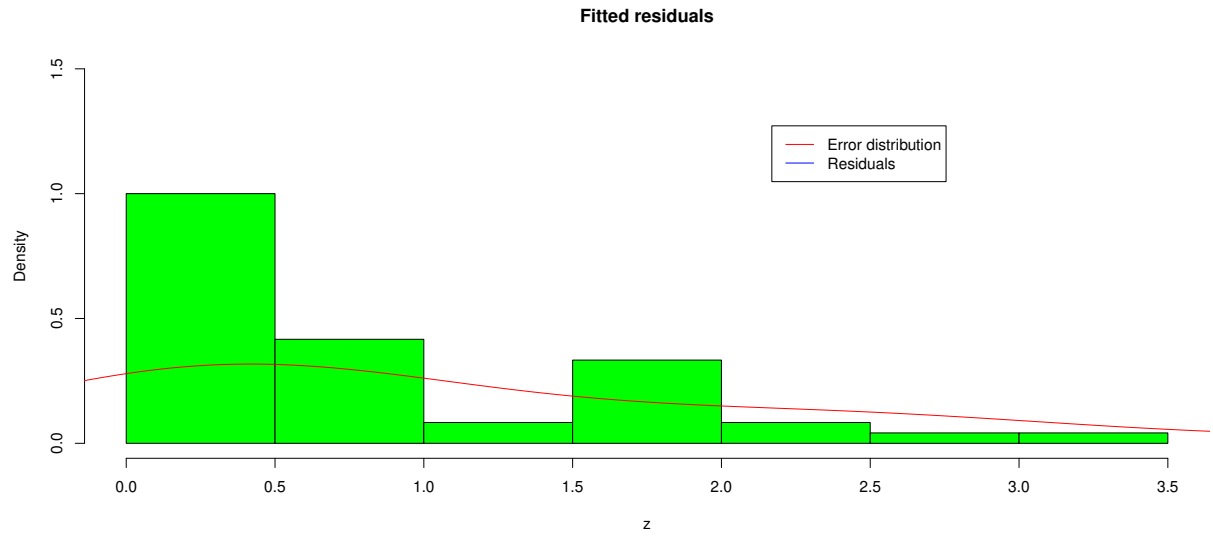

Figure 4.3: Residual plot with innovation series

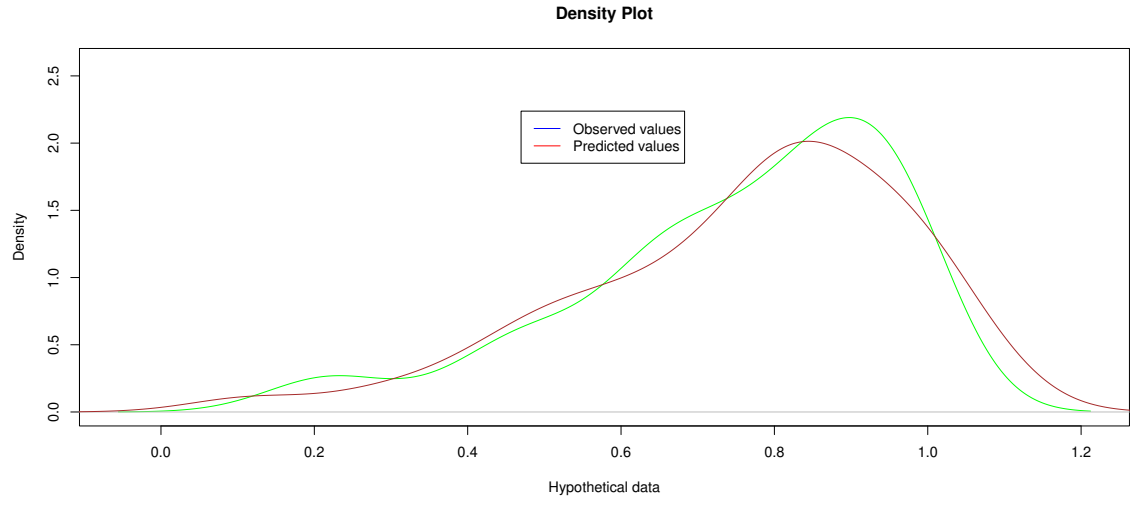

Figure 4.4: Fitted density plot of hypothetical data

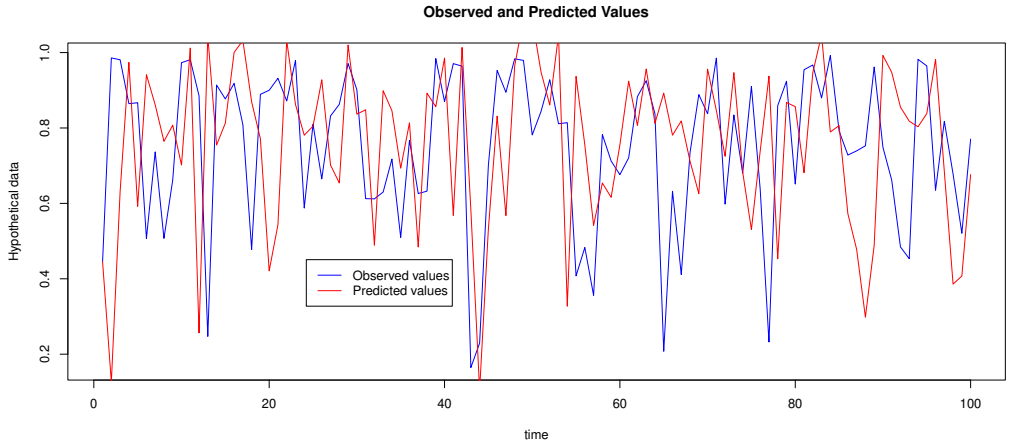

Figure 4.5: Observed and predicted values of hypothetical data

4.5 Conclusion

In this chapter, we constructed the autoregressive conditional duration process with UTPD as the distribution of durations . Distributional properties and classical procedures for the estimation of the associated parameters of the process are discussed. Simulation studies is done. Application of the UTPACD process are illustrated with real data.

CHAPTER 5

SPATIAL AUTOREGRESSIVE MODELS

5.1 Introduction

[1]

Spatial and temporal analysis have gained substantial prominence across various real-world scenarios in contemporary times. Through spatial analysis, we can recognize event patterns in relation to geographical locations, evaluate or assess them, construct event models, and subsequently forecast future values using these models. The credibility of these patterns can be verified by applying statistical tools and techniques skillfully. While spatial analysis and time series analysis are distinct fields, we can integrate them in certain contexts, such as when analyzing spatio-temporal

[1]This chapter is based on Krishnarani, Gautham Manoharan and Vidya (2021)

data where both spatial and temporal components are present. In such cases, the techniques of both may be used together to obtain a deeper understanding of the underlying patterns and relationships.

In this chapter we examine the influence of spatial factors in the healthcare sector, focusing specifically on child mortality rates. Child mortality serves as an indicator that mirrors the developmental metrics of a country, encompassing health, income, literacy, and other socio-economic elements. Child mortality serves as a significant gauge of the development of a country, particularly within the health sector. Child mortality, or under-five mortality, is defined as the number of deaths of children under age 5 per 1000 live births. Over time, it is evident from the records that the under-five mortality rate has had a discernible decline, with notable disparities evident across geographical regions. According to data from the World Health Organization (WHO), the African region struggles with the highest under-five mortality rate, at 74 per 1000 live births. This figure is approximately 9 times greater than that recorded in the European region, where the rate is recorded at 8 per 1000 live births, and such comparisons with several continents can be found on the WHO's official website. This highlights the importance of studying the spatial impact of child mortality rates.

Some studies on mortality with spatial impact can be seen in Li et al. (2019), Manuel et al. (2018), Singh and Masquelier (2018), Gupta et al. (2016), Pezzulo et al. (2016), Xiang and Song (2016), etc. Li et al. (2019) constructed a weighted logistic regression model for each place and time, and the Bayesian space-time smoothing

model is used to estimate trends for the mortality data obtained by conducting a survey of 35 countries in Africa. Manuel et al. (2018) used the spatial autoregressive model (SAR) and classical autoregressive error model to explore infant mortality by taking the influencing variables as the number of women in fertile age and the monthly income of women. This study was specifically concentrated in Alfenas City, Minas Gerais. Using ordinary least squares and geographically weighted regression, Singh and Masquelier (2018) studied the district-specific relationship between child mortality and a series of determinants. Gupta et al. (2016) analyzed infant mortality among several states of India using the SAR model with respect to household amenities, mother/child, and health facility variables. Pezzulo et al. (2016) described geospatial modelling of mortality data in sub-Saharan Africa. Xiang and Song (2016) modeled the perinatal mortality of China province using SAR models.

Apart from the spatial impact, there are numerous additional factors that influence child mortality. Some of them are income inequality, gross domestic product (GDP), the population of the country, and the facilities in the health sector. Low-income countries may possess high child mortality rates. According to the Organization for Economic Co-operation and Development (OECD) (2019) report, in the year 2018, OECD countries spent 8.8% of GDP on health, the USA 16.9%, Switzerland 12.2%, etc. The expenditure of each country, continent, and other clusters of countries in the health sectors is displayed on the website of the World Bank. These reports say Europe spent 9.85% of GDP, South Africa 8.25%, Sub-Saharan Africa 5.15%, Australia 9.28%, and East Asia and Pacific spent 6.68% of GDP on the health

sector. This indicates the necessity of studying the influence of GDP on child mortality. An analysis relating to the spatial and temporal elements of these variables will be useful to practitioners. It is explored in this study through spatio-temporal autoregressive models.

This chapter is organized as follows: In Section 5.2, we review the basic models that are already available in the literature, like regression, spatial lag, spatial error, and spatial Durbin models. Section 5.3 is devoted to the description of the data and the methodology used for our study. The preliminary model selection procedure is explained for the non-spatial scenario in Section 5.4, and the main theme of spatial modelling is transacted in Section 5.5. A concluding note is provided at the end.

5.2 Basic models

For analyzing the relationship between multiple continuous cross-sectional variables, the conventional approach involves employing regression analysis techniques to fit the best model based on the assumptions. These assumptions include the independence of sample observations and the independence and normality of errors. Furthermore, in scenarios where observations exhibit spatial correlation, these assumptions are not valid. Many datasets within geographical, health, and economic sectors show spatial correlation. As noted in the introduction, the popularity and application of spatial modelling have witnessed a notable increase in recent times due to its enhanced predictive accuracy. A comprehensive illustration of fundamental SAR models can

be found in the works of Anselin and Bera (1998), as well as Anselin (1988, 2001).

For a $n \times 1$ vector of random variables Y observed at n spatial points, a SAR process is defined as

$$Y - \mu = \rho W(Y - \mu) + \epsilon,$$

where ϵ is the random vector of dimension $n \times 1$ with $N_n(0, \sigma^2 I)$ distribution, μ is the mean vector of Y, and ρ is the spatial autoregressive parameter. Now the $n \times n$ matrix W is the spatial weight matrix that measures the nearest neighbors of the regions. Several types of weight matrices may be formed, and for details, one may refer to Anselin and Bera (1998), Anselin (2003), Getis and Aldstadt (2004), and Lu and Zhang (2010). Adaptive distance matrix, fixed distance matrix, matrix using the inverse distance weights, row standardized weights, spatial variogram function, etc. are some of the methods of formation of weight matrices seen in the literature. In the row-standardized method, the matrix W gives a weight or a scalar measure to the nearest regions or places. The elements of W are formed as below:

$$w_{i,j} = \begin{cases} 1 & \text{if i \& j are neighbors } (i \neq j); \\ 0 & \text{otherwise.} \end{cases}$$

The assignment of the value 0 serves to avoid the selection of a location itself as its closest neighbor. Subsequently, we perform row normalization on matrix W in such a way that the sum of each row equals one. Similar to the concept of chronological order in time series, here, a spatial order denoted by d implies that d places or areas

are regarded as neighbors or neighboring clusters.

The spatial lag model (SLM), spatial error model (SEM), and spatial Durbin model (SDM) are the extended versions of the SAR models. These three variants have been introduced in the literature as counterparts to the classical regression model. Within these models, the principles of regression are integrated to study the effects of explanatory variables. The SLM model by Anselin (1988) is defined as,

$$Y = \rho WY + X\beta + \epsilon, \tag{5.2.1}$$

where Y is a vector of the dependent variable with dimension $n \times 1$, the matrix X of order $n \times k$ contains the independent variables, β is a $k \times 1$ vector of the regression coefficients, ρ is a scalar spatial autoregressive coefficient, and ϵ is a vector of errors. The $n \times 1$ error vector ϵ follows $N_n(0, \sigma^2 I)$ (see Anselin 2001). In the SLM model, the spatial autocorrelation is in the data itself.

The SEM model is of the form

$$Y = X\beta + \eta, \tag{5.2.2}$$

where η is a $n \times 1$ random vector such that, $\eta = \lambda W\eta + \epsilon$ with λ as spatial coefficient and ϵ follows $N_n(0, \sigma^2 I)$. In this model, the error term follows the SAR model, and the spatial effect is on the error. The third model, SDM, exhibits spatial impact

both on the data and the covariates. SDM has the mathematical form,

$$Y = \rho WY + X\beta + WX\gamma + \epsilon, \tag{5.2.3}$$

where the coefficient γ is the spatial effect of the independent variables and ϵ follows $N_n(0, \sigma^2 I)$.

Our aim in this study is to analyze the spatial impact on the child mortality rate. In the subsequent section, we delineate the data and variables and modify models (5.2.1), (5.2.2), and (5.2.3) to accommodate non-linear relationship between the response variable and the explanatory factors.

5.3 Data and Modified Models

We have taken the child mortality and GDP per capita data from the website https://ourworldindata.org. The child mortality data is published by Gapminder, compiled and documented mainly based on www.mortality.org, the series of books called International Historical Statistics by Brian R Mitchell, and the GDP data by the Maddison Project Database. The total population data is also available in the same site published by Gapminder (v6), HYDE (v3.2), and United Nations Population Division (2019). We accessed the data in May 2021. The variables involved are the child mortality rate (share of newborns who die before reaching the age of 5), GDP (measured in dollars), and the total population (in millions) of the countries

from 2011 to 2016 worldwide. The choice of this range of dates is owing to the availability of the most recent period. The data consists of child mortality rate, GDP, and the total population of 241 countries. The child mortality rate is considered as the response variable and the other two variables as independent variables. If we do not take into account the spatial interaction, the regression equations using the ordinary least square method (OLS) is the usual method of finding the relationship between the response variable and the explanatory variables. The model fitted using this OLS method may be taken as an initial step in analysing the spatial effect on mortality. The scatter plot is used for the regression model fitting to identify the relationship. Linear or non-linear models may be fitted based on the standard procedures. We need to modify the existing SAR models using the particular regression curve used in the previous step.

The regression models under discussion in the next section are:
$Y = \beta_0 + \beta_1 ln(X)$, $Y = \beta_0 e^{\beta_1 X}$, $Y = \beta_0 + \beta_1 X^{\alpha}$, and the Makeham curve of the form $Y = \beta_0 + \beta_1 e^{\alpha X}$.

Then corresponding SLM, SEM and SDM are models are constructed with a model change in (5.2.1), (5.2.2), and (5.2.3) respectively.

In order to reduce the complexity as we moving ahead, we put $Z = c^X$ of the Makeham curve, where $c = e^{\alpha}$ and correspondingly the modified models formed are given below.

The SLM model under this case is ,

$$Y = \rho WY + Z\beta + \epsilon, \tag{5.3.1}$$

and SEM model is

$$Y = Z\beta + \eta, \tag{5.3.2}$$

with the error random vector η of order $n \times 1$, $\eta = \lambda W\eta + \epsilon$ and λ is a spatial coefficient.

The SDM model is formed as

$$Y = \rho WY + Z\beta + WZ\gamma + \epsilon. \tag{5.3.3}$$

See that here the vector of regression parameters β_0 and β_1 is β and the spatial coefficient γ may be considered as the effect of Z. The parameter α in the Makeham model is another parameter.

Hence, we have SAR models corresponding to a non-linear regression model. For the estimation of the parameters, we use the ML method (see Lu and Zhang (2010, 2011)) under the assumption that $\epsilon \sim N_n(0, \sigma^2 I)$, where the observations are taken from n spatial points.

For the SLM, we can write the log-likelihood as

$$\log L = \frac{-n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 + \ln|I-\rho W| - \frac{1}{2\sigma^2}(y-\rho Wy-z\beta)' (y-\rho Wy-z\beta), \quad (5.3.4)$$

where $|I-\rho W|$ is the Jacobin matrix of transformation of order $n \times n$. Here, the parameters β, ρ and σ can be obtained by maximising (5.3.4).

For the SEM, the log-likelihood obtained is

$$\ln L = \ln|I-\lambda W| - \frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(y-z\beta)'(I-\lambda W)' (I-\lambda W)(y-z\beta). \quad (5.3.5)$$

The estimates of the parameters β, λ, and σ are obtained by maximising (5.3.5). Now, we estimate the parameters β, ρ, γ and σ of the SDM by maximizing the logarithm of the likelihood function.

$$\ln L = \ln|I-\rho W| - \frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(y-\rho Wy-z\beta-Wz\gamma)' (y-\rho Wy-z\beta-Wz\gamma). \quad (5.3.6)$$

Analysis of the data is done using the R software of the 4.0.2 version. spdep, GA, rnaturalearth, rnaturalearthdata, sp, ggplot2, fmsb, basictrendline, spatialreg, and sf are the packages used in this work.

After being transformed to $[0, 1]$, the data child mortality rate likewise follows UTPD with parameter $\theta = -3.9$; the K-S test's p-value of 0.109 supports this. However, since the purpose of this chapter is to investigate the geographical impact, we will not be discussing it.

Now as a prelude to the main theme of the chapter we begin with the initial regression model selection procedure in the non-spatial context.

5.4 Preliminary modelling

The data of mortality, GDP and total population are from 241 countries around the globe. But some of the countries are having missing observations. To understand the effect of time on the mortality rate, in addition to the spatial effect, the data of the years from 2011 to 2016 is considered and analyzed separately. As the most recent year is 2016, we do concentrate on this year and the analysis of the previous years will be done in a similar manner. In the data of 2016, we could identify 8 outliers for GDP which is evident from the box plot given in Figure 5.1. So first data cleaning has been done to eliminate the missing observations and outliers from the data. Finally, 157 countries are filtered for the year 2016. Similar filtering has also been done for all the other years considered in the study.

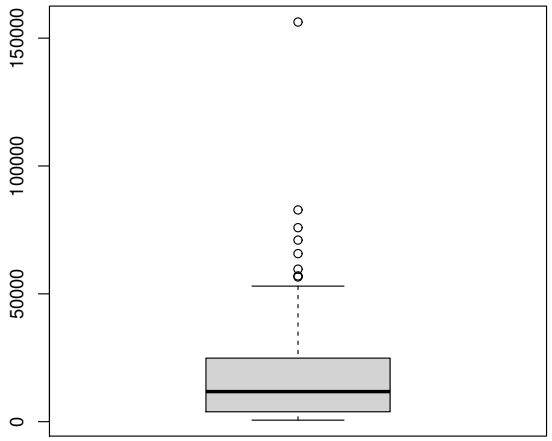

Figure 5.1: Box plot of GDP data of 2016

The summary statistics of the cleansed data for the relevant variables are given in Table 5.1. The number of observations after cleansing in different years is clearly evident from this Table.

We plotted the scatter diagram of the data in 2016 which is given in Figure 5.2. It gives an initial idea of the relationship between the variables considered. The scatter plot reveals an exponential decay in mortality for GDP but no significant relationship with the total population. The first hand information from the plot is that there is no linear relationship between the variables, but an exponential type relation exists. So we tried to fit different types of curves eliminating the total population size. The residuals obtained for the linear model shows non-randomness, and non-constant variance. Hence the regression models mentioned in Section 3 of this chapter have been tried including the Box-Cox transformed model, and the

parameters are estimated using the OLS method. As a tool for determining the best model, the selection criteria values used are adjusted R-squared and Akaikes' information criteria (AIC). The outputs obtained are given in Table 5.2. For the Makeham model, the adjusted R-squared values are the highest, but the AIC value is the lowest. So the Makeham model with the constant term (see Makeham (1860)) is found to be a suitable model for the child mortality rate with GDP as an explanatory variable. The fitted Makeham curve for child mortality rate against GDP is plotted in Figure 5.3. Based on the preliminary model, we shall assess the spatial impact in the next section.

Table 5.1: Summary measures

Year	**No. of observations**	**Mortality minimum**	**Mortality maximum**	**Mortality mean**	**Mortality median**
2011	152	0.25	14.98	3.880	2.205
2012	153	0.24	14.18	3.706	2.110
2013	153	0.23	13.81	3.564	2.040
2014	158	0.23	13.45	3.378	1.925
2015	158	0.22	13.09	3.250	1.845
2016	157	0.21	12.73	3.158	1.790
Year	**GDP minimum**	**GDP maximum**	**GDP mean**	**GDP median**	
2011	681	48980	13988	10014	
2012	710	50394	14484	10398	
2013	567	50863	14580	10760	
2014	561	52651	15094	11026	
2015	576	54278	15294	10982	
2016	589	53015	15153	11073	

5.5 Spatial modelling

We form the weight matrix to identify the impact of the mortality data. By the row-standardized method, a weight matrix W of order 157×157 is formed with elements 0 and 1 and a spatial order or count of 5, as explained in Section 5.2. Then the normalized weight matrix is computed such that for each neighboring country the same value $\frac{1}{no.of neighbours} = \frac{1}{2}$ is assigned and all other entries are zeros. We choose the nearest neighbors by spatial distance, considering latitude and longitude. Moran's I statistic (see Moran (1948), Cliff and Ord (1973), and Diniz-Filho et al. (2003)) is used for the calculation of the spatial correlation. Moran's I statistic is given by

$$I = \frac{\sum WY}{\sum W \sum Y}. \tag{5.5.1}$$

where Y is the child mortality rate for the year 2016. The value of this statistic is 0.72, with p value of 0. Hence, the null hypothesis of spatial independence $\rho = 0$ is rejected. The map of the edges of the spatial regions is given in Figure 5.4.

Since there is spatial correlation, we go for the construction of the models SLM, SEM, and SDM given in (5.3.1), (5.3.2) and (5.3.3). These models are constructed using the Makeham model we already obtained. Now we fix the coefficient c at the estimated value 0.9998 obtained by the OLS estimation method in order to keep the linear regression effect in the model. The fixing of the parameter is done to address the challenges in achieving convergence in the estimation process of multiple

parameters within spatial modelling. Despite the fact that the computed value of c closely approximates 1, it can be noted that a substantial GDP value could have a considerable impact on the child mortality rate within the Makeham model. The magnitude of β_1, which stands at 8.58, shows the pace at which the GDP undergoes geometric progression.

The estimation of the parameters of SLM, SEM, and SDM is obtained by maximizing the likelihood functions (5.3.4), (5.3.5), and (5.3.6) respectively. The best model is selected by considering the measures of AIC and Bayesian information criteria (BIC) mentioned in equations (1.4.1) and (1.4.2). The estimated values of the parameters and selection criteria values of the models are given in Table 5.3 and the standard errors (SE) and confidence intervals (CI) of the model parameters are given in Table 5.4. The selection criteria values AIC and BIC are the least important for the SLM model. Therefore, SLM is the best model for modelling the child mortality of the year 2016. The plots of predicted values of child mortality rate against the observed child mortality rate, the observed GDP values, and the residual plots against predicted values are given in Figures 5.5, 5.6, and 5.7 for SLM, SEM, and SDM, respectively.

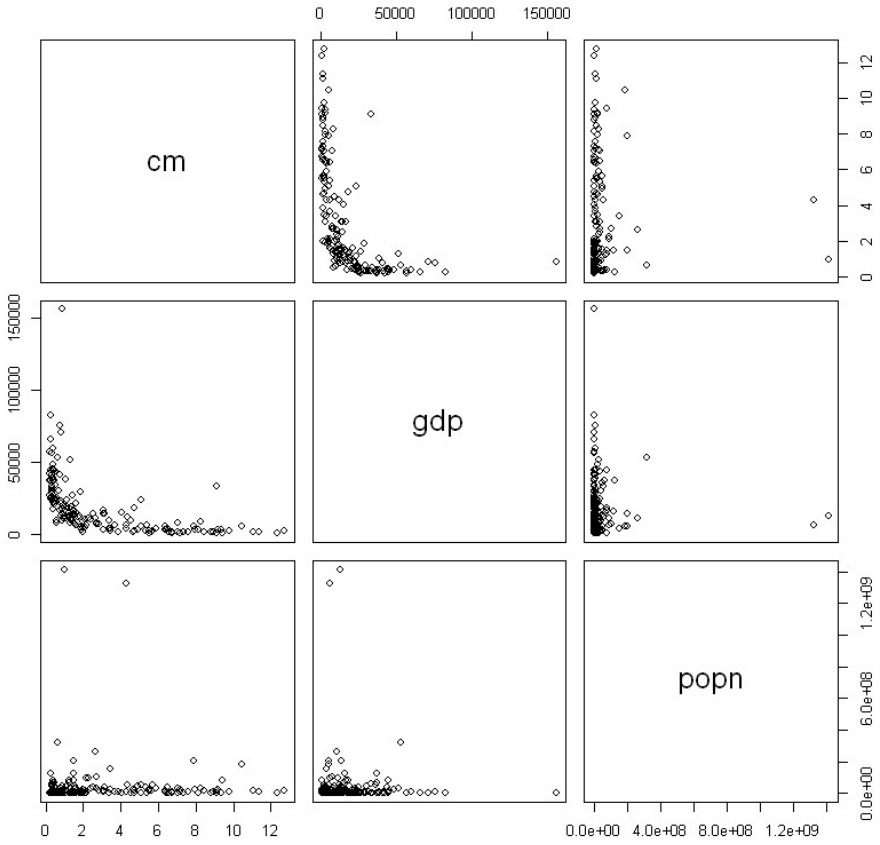

Figure 5.2: The scatter diagram

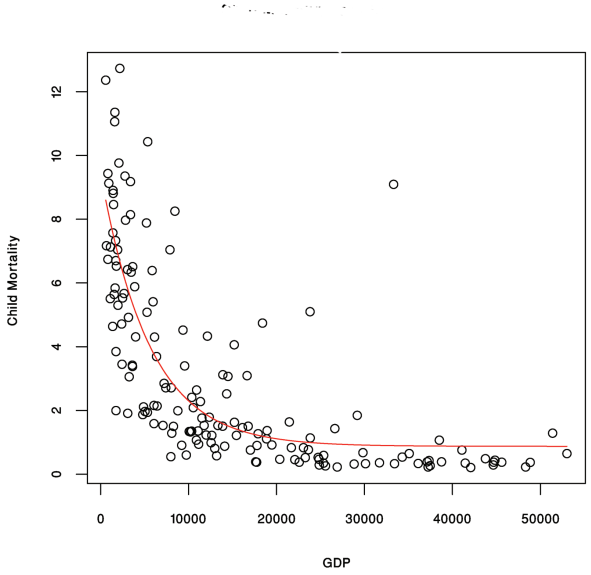

Figure 5.3: Fitted Makeham curve for child mortality rate

Table 5.2: Estimated parameter values and selection criteria values of the regression models for 2016 data

Models	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\lambda}$	Adjusted R-squared	AIC
$Y = \beta_0 + \beta_1 ln(X)$	22.53	-2.12	...	...	...	...
$Y = \beta_0 e^{\beta_1 X}$	8.72	-0.00012	...	...	0.609	656.42
$Y = \beta_0 + \beta_1 X^{\alpha}$	-5.48	70.95	-0.24	...	0.616	654.64
$Y = \beta_0 + \beta_1 e^{\alpha X}$	0.87	8.58	-0.0002	...	0.626	650.70
$\frac{Y^{\lambda}-1}{\lambda} = \beta_0 + \beta_1 ln(X)$	8.40	-0.85	...	0.1	0.620	...

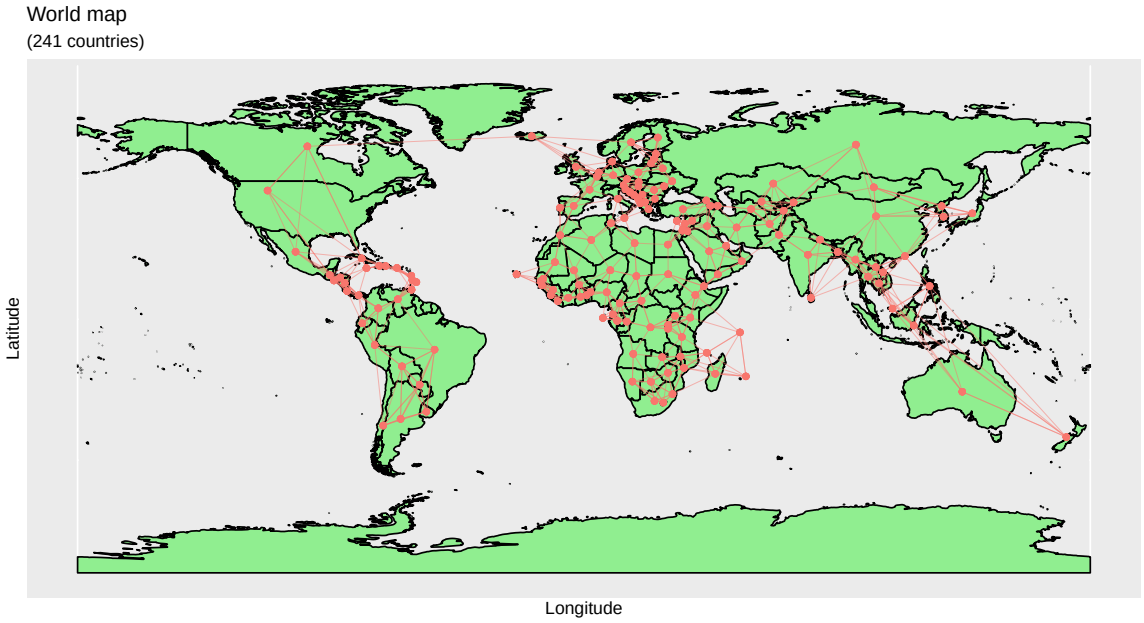

Figure 5.4: The map of the edges of the spatial regions

Table 5.3: Estimated parameter values and selection criteria values of spatial models of 2016 data

Models	β_0	β_1	c	ρ	λ	γ	σ	AIC	BIC
SLM	0.05	4.93	0.9998	0.53	...	...	1.48	587.24	599.47
SEM	1.86	5.06	0.9998	...	0.70	...	1.53	602.18	614.41
SDM	0.61	4.95	0.9998	0.55	...	-0.16	1.48	589.23	604.52

Table 5.4: Standard errors and confidence intervals of the model parameters of the 2016 data

Models	Parameters and values	SE	CI
SLM	β_0, 0.05	0.18	$(-0.30, 0.40)$
	β_1, 4.93	0.54	$(3.87, 5.98)$
	ρ, 0.53	0.06	$(0.41, 0.64)$
SEM	β_0, 1.86	0.42	$(1.04, 2.68)$
	β_1, 5.06	0.60	$(3.88, 6.23)$
	λ, 0.70	0.06	$(0.58, 0.81)$
SDM	β_0, 0.61	0.19	$(0.23, 0.98)$
	β_1, 4.95	0.61	$(3.75, 6.14)$
	ρ, 0.55	0.08	$(0.39, 0.71)$
	γ, -0.16	1.10	$(-2.31, 1.99)$

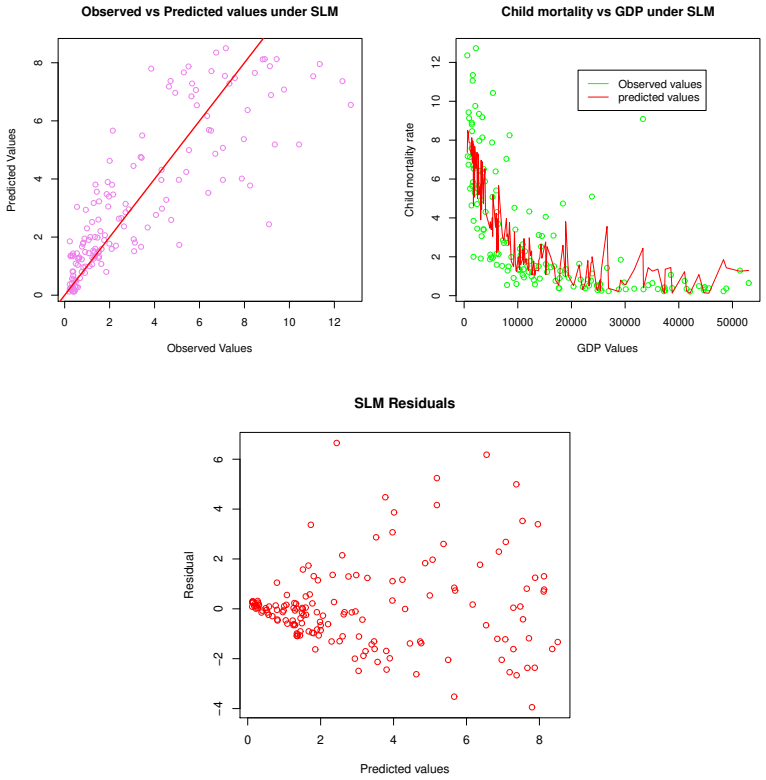

Figure 5.5: Plots of predicted values and residuals of child mortality rate of 2016 under SLM

To model the child mortality rate of the years 2011, 2012, 2013, 2014 and 2015, the procedure is repeated. The estimated parameter values and selection criteria values are computed and given in Table 5.5. Choosing the models with the least AIC and BIC values for the different years, we conclude that SLM is the best model for all the years. The estimates of the spatial auto regressive parameter ρ over the years in the SLM and SDM models are positive, indicating strong spatial dependence between the nearest neighbors with respect to the child mortality rate. But the estimates of the spatial parameter γ corresponding to the GDP in the SDM model is generally negative corroborating, an inverse relationship between the GDP and the child mortality rate in the spatial framework. The positive values of the estimates of λ designate the role of error variables affecting the child mortality rate. Hence there is a need for further studies by examining other possible factors contributing to the error term in the model.

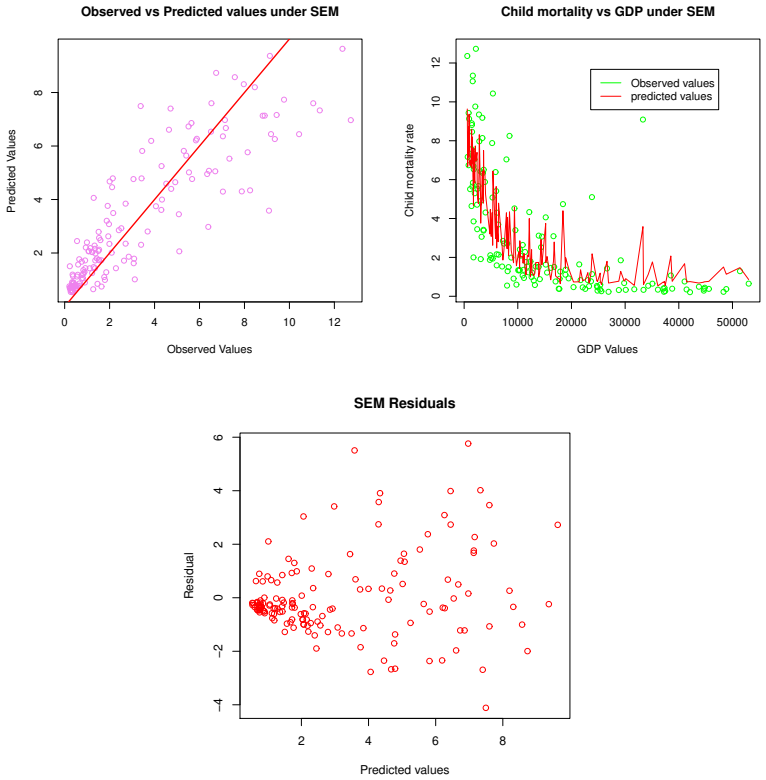

Figure 5.6: Plots of predicted values and residuals of child mortality rate of 2016 under SEM

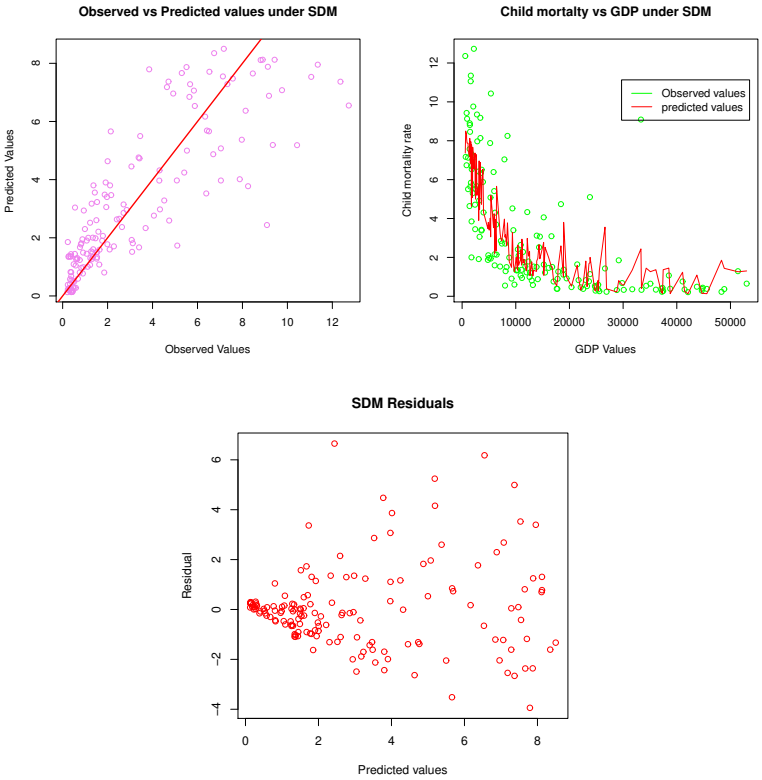

Figure 5.7: Plots of predicted values and residuals of child mortality rate under SDM

Table 5.5: Estimated parameter values and model selection criteria values from 2011 to 2015

Year	**Moran's I, p-value**	**c**	**Models**	β_0	β_1	ρ	λ	γ	σ	**AIC**	**BIC**
2011	0.75, 0	0.9997	Makeham	1.14	11.06	...	...	...	...	685.27	694.34
			SLM	0.11	6.20	0.54	...	...	1.70	609.61	621.70
			SEM	2.07	6.57	...	0.73	...	1.72	623.38	635.48
			SDM	0.30	6.48	0.57	...	-1.20	1.71	612.03	627.15
2012	0.75, 0	0.9997	Makeham	1.10	10.48	...	...	...	...	673.09	685.21
			SLM	0.09	5.88	0.55	...	...	1.65	604.97	617.09
			SEM	2.01	6.17	...	0.73	...	1.67	619.05	631.17
			SDM	0.13	5.95	0.57	...	-0.55	1.64	606.789	621.94
2013	0.74, 0	0.9998	Makeham	1.05	9.96	...	...	...	...	664.08	676.21
			SLM	0.08	5.54	0.55	...	...	1.61	597.79	609.91
			SEM	1.93	5.85	...	0.72	...	1.64	612.49	624.61
			SDM	0.10	5.62	0.56	...	-0.29	1.61	599.73	614.88
2014	0.73, 0	0.9998	Makeham	0.95	9.38	...	...	...	...	675.29	687.54
			SLM	0.06	5.31	0.54	...	...	1.57	608.80	621.05
			SEM	1.79	5.65	...	0.71	...	1.59	622.88	635.13
			SDM	0.14	5.36	0.53	...	-0.46	1.52	611.50	626.82
2015	0.72, 0	0.9998	Makeham	0.92	8.93	...	...	...	...	666.16	678.41
			SLM	0.06	5.07	0.54	...	...	1.53	600.01	612.26
			SEM	1.73	5.37	...	0.71	...	1.55	614.14	626.39
			SDM	0.08	5.13	0.55	...	-0.24	1.52	601.95	617.27

5.6 Conclusion

The spatial modelling of the child mortality rate is discussed with regard to the GDP of the worldwide data. While analyzing the data we could see that the total population size has no influence on the child mortality rate. The basic regression model approach is used for an initial mathematical model selection. Out of the several non-linear regression models, the best suitable one was the Makeham model. Using this curve we have formed SLM, SEM, and SDM models. Among the three SAR models, the most suitable one is seen to be SLM for all the years from 2011 to 2016. The study confirms the strong spatial or geographical dependence between the nearest neighbors in the child mortality rate. The GDP has an inverse spatial relationship in the child mortality. There are some unknown or random factors affecting the child mortality rate which has to be taken into consideration for further studies. This model can be further extended for studying spatial variation in heteroscedastic models.

CHAPTER 6

UNIFORM TRUNCATED POISSON MINIFICATION PROCESSESS

6.1 Introduction

The Laplace or characteristic functions are used to investigate a number of autoregressive models. However, there are a number of random variables with Laplace distribution that lack closed form, whereas survival function have. It is difficult to determine the distribution of the innovation random variable for modelling linear models with these kinds of distribution as marginal. To address this challenge, Tavares (1980) proposed an innovative approach involving a process of minification mentioned in (1.6.15). In this approach conventional addition is replaced with the operation of taking the minimum. This model have led to the development of several

non-Gaussian time series models. The minification models are developed and studied using the survival function of the underlying random variable. Tavares' introduction of minification models has motivated and driven the creation and advancement of various non-Gaussian time series models. In particular, the survival function, which characterizes the probability of a random variable exceeding a certain threshold, serves as a pivotal tool in the formulation and analysis of these models.

The first-order autoregressive exponential process was introduced in Tavares (1980). This process is a time-reversed version of the first-order exponential autoregressive (EAR(1)) process introduced by Graver and Lewis (1980). This result was proved by Chernick et al. (1988). Various aspects of Tavares model were investigated by different researchers. Yeh et al. (1988) studied Pareto random variables, and Pillai (1991) studied a model with a semi-Pareto marginal distribution. Arnold (1989, 1993) and Arnold and Robertson (1989) developed minification processes with logistic marginal distribution, and found utility in applications within biological contexts. One may refer to Arnold and Hallet (1989), Jayakumar and Pillai (2002), Kuttykrishnan and Jayakumar (2008), Krishnarani and Jayakumar (2008a, 2008b, 2013), Ristic (2008), for different elaborations of minification models. Extensions of both additive and minification models have been suggested through the substitution of fixed coefficients with random variables, resulting in the construction of random coefficient autoregressive models. Nicholls and Quinn(1982) introduced random coefficient model. We can see such models in Graver and Lewis (1980), Dewald and Lewis (1985), Lawrence and Lewis (1985), and Han et al. (2018). Many researchers stud-

ied bivariate minification models which are previously discussed in section (1.6.3). In the literature, we can see that no minification models were constructed in the range (0,1). These models may have practical applications in various fields, including finance, biology, and engineering, where relative proportions or ratios are more meaningful than absolute values. Finding a minification model with a marginal distribution in the range (0, 1) can be useful for modelling non-negative data that are bounded by 1, such as proportions, rates, or fractions. The daily rainfall rate of a particular geographical area, the monthly occupancy rate of a hotel, the weekly proportion of positive Covid-19 tests in a country, and the yearly fraction of renewable energy sources in a state's electricity generation are some examples. Motivated by these, we made an attempt to construct minification models with UTPD as the marginal distribution.

This chapter is organized systematically as follows: In Section 6.2, Type I uniform truncated Poisson minification process is introduced and it's properties are studied. The estimation of the parameter and the simulation study have been done in section 6.3. In Section 6.4, Type II uniform truncated Poisson autoregressive minification process is introduced and properties are investigated. Estimation of parameters and simulation studies are done in 6.5. Application of the process with real data is discussed in section 6.6, followed by a concluding section.

6.2 Uniform truncated Poisson minification processes

In this section, we construct minification models of different structures to illustrate the applications of UTPD in the time series analysis.

6.2.1 Type I uniform truncated Poisson minification process

The first model is introduced as follows.

$$X_n = \begin{cases} \varepsilon_n & \text{with probability } e^{-|\theta|} \\ min(X_{n-1}, \varepsilon_n) & \text{with probability } 1 - e^{-|\theta|} \end{cases} \tag{6.2.1}$$

where $\{\varepsilon_n, n \geq 1\}$ is the innovation series and ε_n is independent of X_i's $(i < n)$. Here we assume that X_0 follows UTPD with pdf mentioned in (2.2.2).

To identify the distribution of $\{\varepsilon_n, n \geq 1\}$, from (6.2.1),

$$\bar{F}_{X_n} = \bar{F}_{\varepsilon_n}(x)\left(e^{-|\theta|} + \left(1 - e^{-|\theta|}\right)\bar{F}_{X_{n-1}}(x)\right) \tag{6.2.2}$$

and after simplification, the survival function of the innovation variable is,

$$\bar{F}_{\varepsilon_n}(x) = \frac{e^{\theta} - e^{\theta x}}{e^{-|\theta|}\left(e^{\theta x} - 1\right) + e^{\theta} - e^{\theta x}}. \tag{6.2.3}$$

From (6.2.3), the pdf is derived, and it is

$$f_{\varepsilon_n}(x) = \frac{\theta e^{\theta x} e^{-|\theta|}\left(e^{\theta}-1\right)}{\left(e^{-|\theta|}\left(e^{\theta x}-1\right)+e^{\theta}-e^{\theta x}\right)^2} \tag{6.2.4}$$

The plot of the pdf of the innovation random variable is given in Figure 6.1.

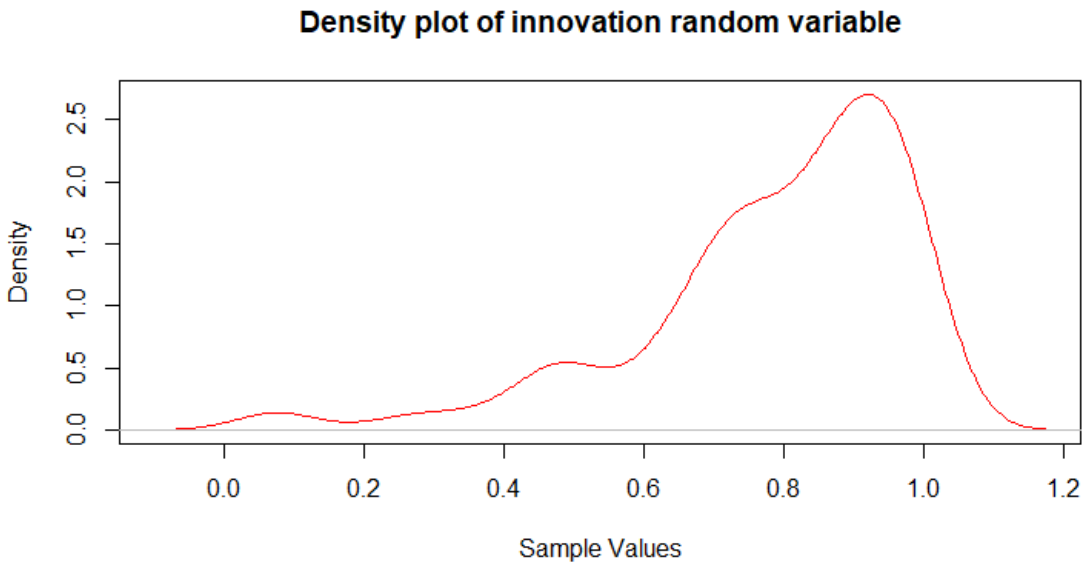

Figure 6.1: Density plot of Type I UTPM innovation random variable

The mean of the innovation variable is given by,

$$E\left(\varepsilon_n\right) = \frac{1}{e^{-|\theta|}-1} + \frac{e^{-|\theta|}\left(e^{\theta}-1\right)}{\left(e^{-|\theta|}-1\right)} \frac{\left(\theta+|\theta|\right)}{\left(e^{\theta}-e^{-|\theta|}\right)}$$

Now, an AR process of the form (6.2.3) with $\{X_n\}$ following UTPD as marginal and $\{\varepsilon_n\}$ follows distribution with pdf (6.2.4), where ε_n is independent of X_{n-1}, is called the Type I uniform truncated Poisson minification (UTPM) process. Note that if

$X_0 \sim$ UTPD with survival function given in (2.3.1), then the process X_n in (6.2.1) is a strictly stationary minification process if and only if $\varepsilon_n^{'}$s are iid with survival function in (6.2.3).

Theorem 6.2.1. *Let X_0 be distributed as UTPD with parameter θ. Define X_n as in (6.2.1). Then $\{X_n\}$ is a Type I UTPM process if and only if $\varepsilon_n{}'s$ are iid with pdf (6.2.4).*

Proof. Let $X_0 \sim$ UTPD(θ), and $\varepsilon_n{}'$s are iid following distribution with survival function (6.2.3).

We use mathematical induction procedure to prove this theorem. Substituting n=1 in (6.2.2), and using the survival function of UTPD,

$$\begin{aligned} \bar{F}_{X_1}(x) &= \bar{F}_{\varepsilon_1}(x)\left(e^{-|\theta|} + \left(1 - e^{-|\theta|}\right)\frac{e^{\theta} - e^{\theta x}}{e^{\theta} - 1}\right) \\ &= \frac{e^{\theta} - e^{\theta x}}{e^{\theta} - 1} \end{aligned}$$

Assuming $X_{n-1} \sim UTPD(\theta)$, following the same steps as in above, we get the result that $\{X_n\}$ is stationary $UTPD(\theta)$.

Conversely we assume $\{X_n\}$ is stationary and $X_0 \sim UTPD(\theta)$. From (6.2.2),

$$\bar{F}_X(x) = \bar{F}_{\varepsilon_n}(x)\left(e^{-|\theta|} + \left(1 - e^{-|\theta|}\right)\bar{F}_X(x)\right)$$

$$\begin{aligned}\bar{F}_{\varepsilon_n}(x) &= \frac{\bar{F}_X(x)}{e^{-|\theta|}+(1-e^{-|\theta|})\bar{F}_X(x)} \\ &= \frac{e^{\theta}-e^{\theta x}}{e^{-|\theta|}(e^{\theta x}-1)+e^{\theta}-e^{\theta x}}.\end{aligned}$$

which is same as (6.2.3). Hence the theorem. □

Now we will look through the conditional properties.

The conditional expectation of $X_n/X_{n-1}=x$ is given by,

$$\begin{aligned}E(X_n/X_{n-1}=x) = E(\varepsilon_n)e^{-|\theta|} &+ \left(1-e^{-|\theta|}\right)\frac{x\left(e^{\theta}-e^{\theta x}\right)}{e^{-|\theta|}(e^{\theta x}-1)+e^{\theta}-e^{\theta x}} \\ &- \frac{xe^{-|\theta|}\left(1-e^{-|\theta|}\right)\left(e^{\theta}-1\right)}{\left(e^{-|\theta|}-1\right)\left(e^{-|\theta|}(e^{\theta x}-1)+e^{\theta}-e^{\theta x}\right)^2} \\ &+ \frac{log e^{\theta x}\left(e^{\theta}-1\right)\left(1-e^{-|\theta|}\right)}{\left(e^{\theta}-e^{-|\theta|}\right)\left(e^{\theta x}\left(e^{-|\theta|}-1\right)+e^{-|\theta|}+e^{\theta}\right)}\end{aligned} \tag{6.2.5}$$

The conditional probability distribution of $X_n > y/X_{n-1}=x$ is given by

$$P(X_n>y/X_{n-1}=x) = \begin{cases} \frac{e^{-|\theta|}\left(e^{\theta}-1\right)}{e^{-|\theta|}-1}\left(\frac{1}{e^{-|\theta|}\left(e^{\theta y}-1\right)+e^{\theta}-e^{\theta y}}-\frac{1}{e^{-|\theta|}\left(e^{\theta}-1\right)}\right) & \text{if x } > y \\ \frac{e^{-2|\theta|}\left(e^{\theta}-1\right)}{e^{-|\theta|}-1}\left(\frac{1}{e^{-|\theta|}\left(e^{\theta y}-1\right)+e^{\theta}-e^{\theta y}}-\frac{1}{e^{-|\theta|}\left(e^{\theta}-1\right)}\right) & \text{if } x < y \end{cases}$$

6.3 Estimation of parameters

In this section we estimate the unknown parameters of the Type I UTPM process. Let X_0, X_1, ..., X_n be the realization of the Type I UTPM process. For the estimation

of the parameter θ, we will consider the stationary process $\{U_n\}$ given by

$$U_n = \begin{cases} 1 & \text{if } X_n \geq X_{n-1} \\ \\ 0 & \text{if } X_n < X_{n-1} \end{cases}$$

Now it can be seen that

$$\begin{aligned} E\left(U_n\right) = P\left(U_n = 1\right) \quad &= \frac{\left(e^{\theta}-1\right) log\left(e^{\theta}-1\right)}{\left(1-e^{-|\theta|}\right)^2} \\ &- \frac{\left(e^{\theta}-1\right)}{\left(1-e^{-|\theta|}\right)^2}\left(\left(1-e^{-|\theta|}\right) log\left(e^{\theta}-1\right) + \left(e^{-|\theta|}\left(1+|\theta|\right)\right)\right) \end{aligned}$$

Then the estimator of θ can be obtained by solving the equation

$$\begin{aligned} \bar{U_n} = &\frac{\left(e^{\hat{\theta}}-1\right) log\left(e^{\hat{\theta}}-1\right)}{\left(1-e^{-|\hat{\theta}|}\right)^2} - \frac{\left(e^{\hat{\theta}}-1\right)}{\left(1-e^{-|\hat{\theta}|}\right)^2}\left(\left(1-e^{-|\hat{\theta}|}\right) log\left(e^{\hat{\theta}}-1\right)\right. \\ &\left. + \left(e^{-|\hat{\theta}|}\left(1+\left|\hat{\theta}\right|\right)-1\right)\right) \end{aligned}$$

The value of $\hat{p} = P(X_n > X_{n-1})$ for different sample sizes and for different values of θ is given in Table 6.1 . From the table, it can be seen that as the parameter value θ increases, the value of $P(X_n > X_{n-1})$ decreases.

Table 6.1: Probabilities for different values of θ

Sample size(n)	Parameterθ	$(\hat{p})$
50	-1	0.2649
	1	0.2502
	2	0.0974
	3	0.0438
100	-1	0.2465
	1	0.2438
	2	0.1150
	3	0.0413
200	-1	0.2502
	1	0.2500
	2	0.0947
	3	0.0443

6.4 Type II uniform truncated Poisson minification process

Consider a minification process defined as,

$$X_n = \begin{cases} \varepsilon_n & with\ probability\ e^{-|\theta|} \\ min\left(e^{|\theta|}X_{n-1}, \varepsilon_n\right) & with\ probability\ 1 - e^{-|\theta|} \end{cases} \tag{6.4.1}$$

Here $\{\varepsilon_n, n \geq 1\}$ is the innovation sequence of iid random variables; and ε_n is independent of X_i, i= 0, 1, 2, ..., n-1. We assume that $\{X_n, n \geq 1\}$ is a stationary Markov process with marginal distribution, UTPD. Now (6.4.1) is referred to as the Type II uniform truncated Poisson minification (Type II UTPM) process.

The distribution of the innovation sequence can easily be derived from (6.4.1) by using the survival function.

The survival function of $\{X_n, n \geq 1\}$ can be written as

$$\begin{aligned}\bar{F}_{X_n}(x) &= e^{-|\theta|}\bar{F}_{\varepsilon_n}(x) + \left(1 - e^{-|\theta|}\right)\bar{F}_{X_{n-1}}\left(xe^{-|\theta|}\right)\bar{F}_{\varepsilon_n}(x) \\ &= \bar{F}_{\varepsilon_n}(x)\left(e^{-|\theta|} + \left(1 - e^{-|\theta|}\right)\bar{F}_{X_{n-1}}\left(xe^{-|\theta|}\right)\right) \qquad (6.4.2)\end{aligned}$$

Under stationarity the survival function of the innovation process is given by

$$\bar{F}_{\varepsilon_n}(x) = \frac{e^{\theta} - e^{\theta x}}{e^{\theta} - e^{-|\theta|} - e^{\theta e^{-|\theta|}x}\left(1 - e^{-|\theta|}\right)} \qquad (6.4.3)$$

The pdf of the innovation random variable is given by

$$f(\varepsilon_n) = \frac{\left(e^{\theta} - e^{-|\theta|}\right)\theta e^{\theta x} - \theta e^{\theta e^{-|\theta|}x}\left(1 - e^{-|\theta|}\right)\left(e^{\theta x}\left(1 - e^{-|\theta|}\right) + e^{\theta}e^{-|\theta|}\right)}{\left(e^{\theta} - e^{-|\theta|} - e^{\theta e^{-|\theta|}x}\left(1 - e^{-|\theta|}\right)\right)^2} \qquad (6.4.4)$$

The conditional probability distribution of $X_{n+1} > x / X_n = y$ is given by

$$P(X_{n+1} > x / X_n = y) = \begin{cases} P(\varepsilon_{n+1} > x) & \text{if } e^{|\theta|}y > \varepsilon_n \\ P(\varepsilon_{n+1} > x)\, e^{-|\theta|} + \left(1 - e^{-|\theta|}\right) P\left(e^{|\theta|}y > x\right) & \text{if } e^{|\theta|}y < \varepsilon_n \end{cases}$$

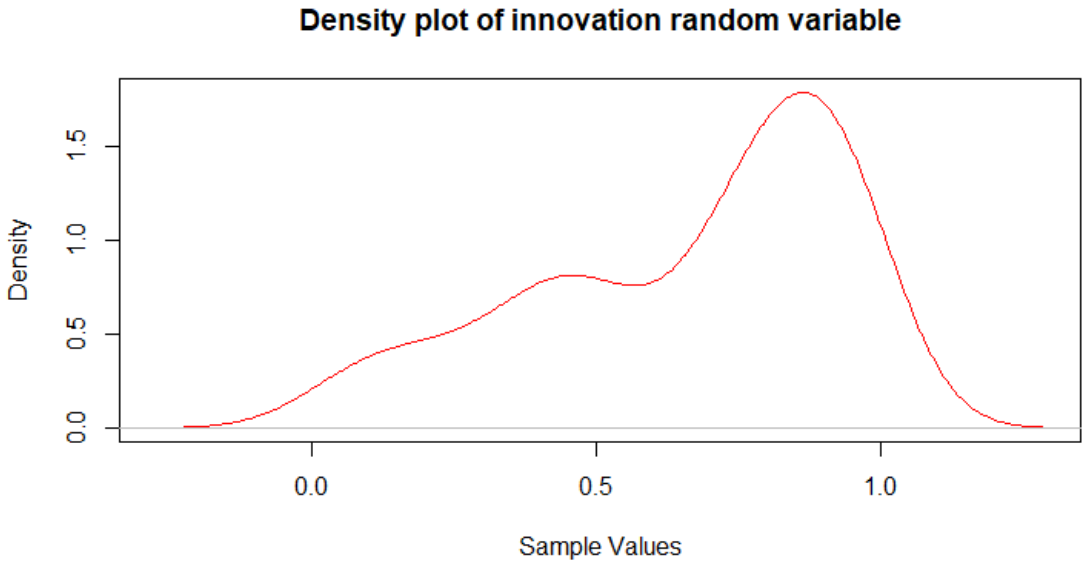

Figure 6.2: Density plot of Type II UTPM innovation random variable

Theorem 6.4.1. *Let X_0 be distributed as UTPD. Define X_n as in (6.4.1). Then $\{X_n\}$ is a Type II UTPM process if and only if ε_n's are iid with pdf (6.4.4).*

Proof. Let $X_0 \sim \text{UTPD}(\theta)$ and ε_n's are iid following distribution with survival function (6.4.3). We adopt mathematical induction procedure. Substituting n=1 in (6.4.2),

$$\begin{aligned} \bar{F}_{X_1}(x) &= \bar{F}_{\varepsilon_1}(x)\left(e^{-|\theta|} + \left(1 - e^{-|\theta|}\right)\frac{e^{\theta} - e^{\theta e^{-|\theta|}x}}{e^{\theta} - 1}\right) \\ &= \frac{e^{\theta} - e^{\theta x}}{e^{\theta} - 1} \end{aligned}$$

Assuming $X_{n-1} \sim UTPD(\theta)$, following the same steps, we get the result that $\{X_n\}$ is stationary $UTPD(\theta)$.

Conversely we assume $\{X_n\}$ is stationary and $X_0 \sim \text{UTPD}(\theta)$. From (6.4.2),

$$\bar{F}_X(x) = \bar{F}_{\varepsilon_n}(x)\left(e^{-|\theta|} + \left(1 - e^{-|\theta|}\right)\bar{F}_X(x)\right)$$

$$\begin{aligned}\bar{F}_{\varepsilon_n}(x) &= \frac{\bar{F}_X(x)}{e^{-|\theta|} + (1 - e^{-|\theta|})\,\bar{F}_X(e^{-|\theta|}x)} \\ &= \frac{e^{\theta} - e^{\theta x}}{e^{\theta} - e^{-|\theta|} - e^{\theta e^{-|\theta|}x}(1 - e^{-|\theta|})}\end{aligned}$$

which is same as (6.4.3). Hence the proof. □

The parameter estimation of the Type II UTPM process discussed in the following section. The simulations are also used to evaluate the precision of the methods of estimation.

6.5 Estimation and Simulation

The method of maximum likelihood is proposed here for the estimation of the parameters. The likelihood function is given by

$$L(x) = \begin{cases} \prod_{i=1}^{n} f_{\varepsilon_i}(x) & \text{w.p } e^{-|\theta|} \\ \prod_{i=1}^{n} f_{\varepsilon_i}(x) & \text{w.p } 1 - e^{-|\theta|}, \quad if\ e^{|\theta|}X_{i-1} > \varepsilon_i \\ \prod_{i=1}^{n} f_{y_i}(x) & \text{w.p} 1 - e^{-|\theta|}, \quad if\ e^{|\theta|}X_{i-1} < \varepsilon_i \end{cases} \tag{6.5.1}$$

where $y = e^{|\theta|}X_{n-1}$,

The p.d.f of y is given by

$$f(y) = \frac{\theta}{e^{|\theta|}(e^{\theta}-1)} e^{\theta e^{-|\theta|} y}, \quad 0 < y < e^{|\theta|} \tag{6.5.2}$$

$$logL = \begin{cases} \sum_{i=1}^{n} log f_{\varepsilon_i}(x) & \text{if } e^{|\theta|} X_{i-1} > \varepsilon_i \\ e^{|\theta|} \sum_{i=1}^{n} log f_{\varepsilon_i}(x) + \left(1 - e^{-|\theta|}\right) \sum_{i=1}^{n} log f_y(x) & \text{if } e^{|\theta|} X_{i-1} < \varepsilon_i \end{cases} \tag{6.5.3}$$

The maximum likelihood estimator θ is obtained by maximising (6.5.3). Since the parameter solution does not have an explicit expression, numerical techniques are used to identify it. In this case, the values for the parameter are obtained by applying the Nelder-Mead method to the R nlminb() function.

The estimation method is validated by simulating samples of sizes 20, 50, and 100. Table 6.2 shows the various values of the parameters assumed for this purpose and the corresponding estimated values. It is evident from the table that MSE decreases with increasing sample size. The density plot of innovation random variable is shown in Figure 6.2.

6.6 Real data analysis

We have considered the amount of carbon dioxide emissions per capita (metric tons) in Bahrain from 1990 to 2020 from the website https://data.worldbank.org. Using the min-max transformation, the observations are transformed. The ACF and PACF plots reveal that the data is stationary AR(1). Using the estimation procedure, $\hat{\theta}$

obtained is 2.9. The K-S test, with a p-value of 0.8211, confirms that the data follows UTPD. Then the values are predicted, assuming that the data follows the Type II UTPM process. The density and time series plots of observed and predicted values are shown in Figure 6.3 and Figure 6.4 respectively.

Table 6.2: Parameter estimates of θ

Sample size(n)	Parameterθ	Estimate($\hat{\theta}$)	MSE(θ)
20		0.1496	0.7211
50		0.2034	0.1573
100	0.5	0.3187	0.1454
20		1.3592	0.1290
50		1.3488	0.1217
100	1	1.3487	0.1216
20		1.1164	0.7807
50		1.3487	0.4242
100	2	2.1791	0.0321
20		2.8484	0.0221
50		2.7711	0.0124
100	3	2.9230	0.0059
20		5.8266	0.6832
50		4.2001	0.6385
100	5	4.8207	0.0321

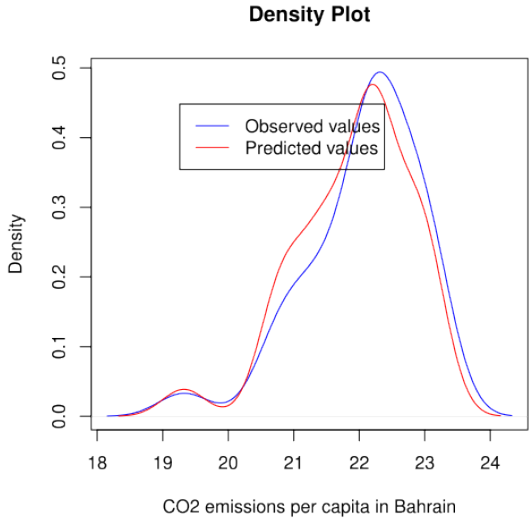

Figure 6.3: Density plot of observed and predicted values of Bahrain

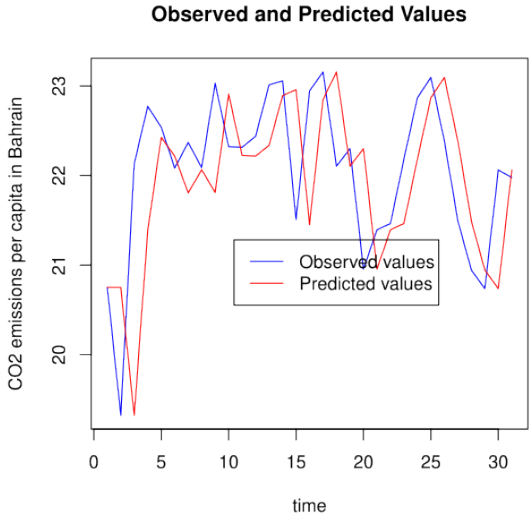

Figure 6.4: Time series plot of observed values and predicted values

6.7 Conclusion

In this chapter, two types of minification processes are constructed with UTPD as marginal. Analytical properties are derived. The estimation of the parameters of the process is discussed along with the simulation studies. Real-data application of the type II UTPM process is also done.

CHAPTER 7

RECOMMENDATIONS

Time series analysis is a pivotal statistical and data analysis method with wide-ranging applications across various domains. Its significance arises from its capacity to reveal valuable insights, facilitate predictive modelling, and guide informed decision-making through the examination of data that evolves over time. Recent studies on time series analysis have focused mostly on analyzing and researching the idea that linear time series models can accurately capture the structure of the series. However, there are situations when the subject, theory, or data indicate that linear models are unreliable. In those situations, it is desirable to take non-linear options into account. Most frequently, Gaussian distributions are used in linear time series models for errors or residuals, which may not adequately represent many real-world data sets. The ability to accommodate non-Gaussian and heavy-tailed distributions,

on the other hand, is a strength of non-linear models, allowing for more accurate modelling of extreme occurrences and outliers. Therefore, non-linear modelling using non-Gaussian distributions is necessary for the analysis of financial time series. This thesis deals with the construction of a new distribution, UTPD and its applications in different non-linear time series modelling. The first chapter is devoted to an introduction to time series data, linear and non-linear time series models, and their properties.

In the second chapter, we have studied in detail the uniform truncated Poisson distribution as the solution of a first-order differential equation and derived the same from the truncated uniform distribution. A comparison with a well-known distribution was done. The expressions for moments, distributions of the order statistics, etc. are further derived. Some transformed distributions were also studied. Some of the estimation procedures for the parameters, like maximum likelihood estimation and the method of moments, were discussed. The newly constructed distribution was applied to four real data sets. Generalized UTPD is derived and also studied it's properties.

The third chapter is devoted to the time series application of UTPD. We developed a flexible autoregressive process of order 1, with UTPD as marginal. The analytical characteristics of this process including the coditional properties have been studied in detail. The estimation of the parameter is done by using CLSE and Gaussian estimation method and its validity is established through simulation. Real-life applications are illustrated using the amount of carbon dioxide emissions per capita

(metric tons) in Japan.

Nowadays ACD models are used to analyze the dynamic behavior of financial duration data. So we developed an ACD model with UTPD as the distribution of duration in Chapter 4. Properties of the process, including mean, variance, and autocorrelation are studied. The estimation of the associated parameters are done by using the Gaussian estimation method. Simulation studies are carried out, and application is discussed with the help of two real data sets.

Mortality, especially child mortality, throws light into the health conditions of a country. The gross domestic product is an influential factor in child mortality. Spatial effects on child mortality rates through spatial autocorrelations are analyzed in chapter 5, and this study unifies the regression and spatial effects of gross domestic product. The analysis summarizes the results of recent years to compare the temporal effect. Spatial regression models are constructed and compared using several criteria to determine the best model. The regression models, spatial lag model, spatial error model, and spatial Durbin model are fitted at different time points.

Two different types of minification processes with the UTPD as stationary marginal distributions are developed in chapter 6. The properties of the process are derived. Estimation of the parameter and simulation studies are also done. Real data analysis are carried out with the minification processes

The following list contains some further research topics that could be used to expand this thesis work. We looked at a new distribution called UTPD in chapter 2. It is possible to model both left-skewed and right-skewed data with UTPD. As a re-

sult, it has numerous real-world applications. Many more time series models, as well as data sets from various fields of study, could be modeled using this distribution. Furthermore, a generalization of UTPD is presented, and its basic characteristics are investigated. Future research will focus on additional features and uses, among other things. It is possible to construct time series models linked to generalized UTPD. Multivariate generalizations of the UTPD can be considered, and vector autoregressive models may be investigated in the future. Some transformed distributions are introduced in chapter 2. Exploring the properties, estimation procedures, and applications will be another area of research.

We have not discussed the volatility models in this work. So ARCH, GARCH models, and stochastic volatility models could be addressed in the future using the UTPD or generalized UTPD.

Our focus has primarily been on autoregressive models of the first order in this thesis. In time series analysis, it's not unusual to encounter non-stationarity. So, it would be advantageous to investigate the creation of higher-order autoregressive models with the probability distribution presented in this study serving as the marginals or innovations.

BIBLIOGRAPHY

[1] Adamidis, K., and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics and Probability Letters*, **39**, 35–42.

[2] Adke, S. R., and Balakrishna, N. (1992). Renewal counting process induced by a discrete markov chain. *Australian Journal of Statistics*, **34(1)**, 115-121.

[3] Akdogan, Y., Kus, C., Asgharzadeh, A., Kinaci, I., and Sharafi, F. (2016). Uniform-Geometric distribution. *Journal of Statistical Computation and Simulation*, **86(9)**, 1754-1770.

[4] Alice, T., and Jose, K. K. (2004). Bivariate semi-Pareto minification processes. *Metrika*, **59**, 305-313.

[5] Al-Nachawati, H., Alwasel, I., and Alzaid, A. A. (1997). Estimating the parameters of the generalized Poisson AR(1) process. *Journal of Statistical Computation and Simulation*, **56(4)**, 337-352.

[6] Altawil, J. A. (2019). $[0, 1]$ Truncated Lomax – Uniform distribution with properties. *Journal of Interdisciplinary Mathematics*, **22(8)**, 1415-1431.

[7] Alwasel, I., Alzaid, A., and Al-Nachawati, H. (1998). Estimating the parameters of the binomial autoregressive process of order one. *Applied Mathematics and Computation*, **95**, 193–204.

[8] Anderson, T.W. (1971). The statistical analysis of time series. *John Wiley*, New York.

[9] Anselin, L. (1988). *Spatial econometrics: methods and models*, Kluwer Academic, Dordrecht.

[10] Anselin, L. (2001). Spatial effects in econometric practice in environmental and resource economics, *American Journal of Agricultural Economics*, **83(3)**, 705–710.

[11] Anselin, L. (2003). Spatial Econometrics, A companion to theoretical econometrics, *Blackwell Publishing Ltd.*

[12] Anselin, L. (2006). Spatial econometrics. *Handbook of Econometrics*: **1**, Econometric Theory, 901–969. Basingstoke.

[13] Anselin, L. (2021). Spatial econometrics. *Handbook of Spatial Analysis in the Social Sciences*. Edward Elgar Publishing, Cheltenham, UK. forthcoming.

[14] Anselin, L., and Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics, *Handbook of Applied Economic Statistics, New York: Marcel Dekker* **155**, 237-289.

[15] Anselin, L., Le Gallo, J., and Jayet, H. (2008). Spatial panel econometrics. *The Econometrics of Panel Data*, Advanced studies in theoretical and applied econometrics, **46**, springer-Verlag, Berlin.

[16] Anselin, L. and Rey, S. J. (2014). Modern spatial econometrics in Practice, *A Guide to GeoDa, GeoDaSpace and PySAL. GeoDa Press, Chicago, IL.*

[17] Arbia, G. (2014). A primer for spatial econometrics with applications in R. Palgrave Macmillan, Basingstoke, UK.

[18] Arnold, B.C. (1989). A logistic process constructed using geometric minimization. *Statistics and Probability Letters*, **7**, 253-257.

[19] Arnold, B. C. (1993). Logistic processes involving Markovian minimization. *Communications in Statistics-Theory and Methods*, **22**, 1699-1707.

[20] Arnold, B.C., and Hallett, T.J. (1989). A characterization of the Pareto process among stationary stochastic processes of the form $X_n = Cmin(X_{n-1}, Y_{n-1})$. *Statistics and Probability Letters* , **8**, 377-380.

[21] Arnold, B. C., and Robertson, C. A. (1989). Autoregressive logistic process, *Journal of AppIied Probability*, **26**, 524-531.

[22] Bai, J., Baltagi, B., and Pesaran, H. (2016). Cross-sectional dependence in panel data models: a special issue. *Journal of Applied Econometrics*, **31**,1–3.

[23] Bakouch, H. S., and Popvic, B. V. (2016). Lindley first- order autoregressive model with applications, *Communications in Statistics-Theory and Methods*, **45(17)**, 1-20.

[24] Balakrishna, N. (1998). Estimation for the semi- Pareto process. *Communications in Statistics-Theory and Methods*, **32**, 2307-2323.

[25] Balakrishna, N., and Jacob, T. M. (2003). Parameter estimation in minification processes. *Communications in Statistics- Theory and Methods,* **32**, 2139- 2152.

[26] Balakrishna, N., and Jayakumar, K. (1997). Bivariate semi-Pareto distributions and processes. *Statistical Papers*, **38**. 149- 165.

[27] Bauwens, L., and Giot, P. (2000). The logarithmic ACD-model: An application to the Bid-Ask quote process of three NYSE stocks. *Annales d'e'conomie et de statistiques*, **60**, 117-149.

[28] Bayer, F. M., Cintra, R. J., and Cribari-Neto, F. (2018). Beta seasonal autoregressive moving average models. *Journal of Statistical Computation and Simulation*, **88(15)**, 2961-2981.

[29] Benrhmach, G., Namir,K., Namir, A. and Bouyaghroumni, J. (2020), Nonlinear autoregressive neural network and extended Kalman filters for Prediction of financial time Series. *Hindawi Journal of Applied Mathematics.*

[30] Bhatti, and Chad, R. (2010). The Birnbaum-Sunders autoregressive conditional duration model. *Mathematics and Computers in Simulation(MATCOM)*, Elsevier, **80(10)**, 2062-2078.

[31] Billard, L. (2014). Exponential autoregressive models, *Wiley Stats Ref: Statistics Reference online, John WileySons Ltd 2014*.

[32] Blasques, F., Koopman, S. J. and Lucas, A. (2020). Non-linear autoregressive models with optimality properties. *Econometric Reviews*, **39(6)**, 559-578.

[33] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.

[34] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time series analysis: forecasting and control*. Englewood Cliffs, NJ: Prentice Hall.

[35] Box, G. E. P., and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, **65 (332)**, 1509—1526.

[36] Box, G. E. P., and Tiao, G. C. (1962). A further look at robustness via Bayes's theorem. *Biometrika*, **49 (3/4)**, 419–432.

[37] Breidt, F.J., Davis, R.A., Hsu, N.J., and Rosenblatt, M., (2006). Pile-up probabilities for the Laplace likelihood estimator of a non-invertible first order moving average. In: Ho, H.-C., Lai, T.L., (Eds.), Memory of Ching-Zong Wei, IMS Lecture Notes, pp. 1–19.

[38] Cancho, V., Louzada, F., and Barriga, G. (2011). The Poisson-exponential lifetime Distribution. *Computational Statistics and Data analysis*, **55**, 677-686.

[39] Carlin, Nicholas, G. Polson. and David S. Stoffer (1992). A Monte Carlo Approach to non normal and nonlinear state-space modelling. *Journal of the American Statistical Association*. **87(418)**.

[40] Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory*. **18**, 17-39.

[41] Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, **53**, 4433-4440.

[42] Chan, K.S. (1993). Consistency and limiting distribution of a least squares estimator of a threshold autoregressive model. *Annals of Statistics*. **21**, 520–533.

[43] Chan, K. S. (1991). Percentage points of likelihood ratio test for threshold autoregression. *Journal of the Royal Statistical Society*, series B, **53(3)**, 691-696

[44] Chan, K. S., and Cryer, J. D. (2008), *Time series Analysis with Applications in R, Springer series*

[45] Chan, K. S., and Tsay, R. S. (1998). Limitting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika*, **85**, 413-426.

[46] Chen, R., and Tsay, R. S. (1993a). Nonlinear additive ARX models. *Journal of American Statistical Association*, **88**, 955-67.

[47] Chen, R., and Tsay, R. S. (1993b). Functional-coefficient autoregressive models. *Jornal of American Statistical Association.***88**,298-308.

[48] Chernick, M. R. , Daley, D. J., and LittleJohn, R. P.(1988). A time-reversibility relationship between two Markov chains with exponential stationary distributions. *Journal of Applied Probability* , **25**, 418-422.

[49] Chi, G., and Zhu, J. (2020). Spatial regression models for the social sciences. Sage, Thousand Oaks, CA.

[50] Cifarelli, D.M., Gupta, R.P., and Jayakumar, K. (2010). On generalized semi-Pareto and semi-Burr distributions and random coefficient minification processes. *Statistical Papers*, **51**, 193–208.

[51] Cliff, A., and Ord, J. K. (1973). Spatial Autocorrelation, London: Pion.

[52] Cliff, A. , and Ord, J. K. (1981). Spatial Processes: Models and Applications. Pion, Londo.

[53] Crowder, M. (1985). Gaussian estimation for correlated binomial data. *Journal of the Royal Statistical Society*, Series B,**47(2)**, 229-237.

[54] Dewald, L.S., and Lewis, P.A.W. (1985). A new Laplace second-order autoregressive time series model- NLAR(2). *IEEE Transactions on Information Theory*, **31**, 645-651.

[55] Diniz-Filho, J. A. F., Bini, L. M., and Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology, *Global Ecology and Biogeography*, **12**, 53–64.

[56] Drost, F. C., and Werker, B.J. (2004). Semiparametric duration models. *Journal of Business & Economic Statistics. Taylor & Francis*, **22(1)**, 40-50.

[57] Dub´e, J., and Legros, D. (2014). Spatial Econometrics Using Microdata. *ISTE-Wiley, London* - Hoboken.

[58] Dufour, A., and Engle, R. F.(2000a). Time and the price impact of a trade. *The Journal of Finance, 2000 - Wiley Online Library.*

[59] Dufour, A., and Engle, R. F. (2000b). The ACD model: Predictability of the time between consecutive trades, in discussion papers in finance. ISMA Centre,**59**.

[60] Elhorst, J. P. (2012). Dynamic spatial panels: Models, methods, and inferences. *Journal of Geographical Systems*, **14**,5–28.

[61] Elhorst, J. P. (2014a). Spatial econometrics, from cross-sectional data to spatial panels. *Springer*, Heidelberg.

[62] Elhorst, J. P. (2014b). Spatial panel models. *Handbook of Regional Science Springer Verlag*, 1634-1652, Berlin, Heidelberg.

[63] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation. *Econometrica,***50**, 987-1007.

[64] Engle . R. F.(2000). The Econometrics of Ultra-high frequency data, *Econometrica* ,**68(1)**, 1—22.

[65] Engle, R. F.(2002). New frontiers for ARCH models. *Jornal of Applied Econometrics***17**, 425-446.

[66] Engle, R. F., and Russel, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* . **66(5)**, 1127–1162.

[67] Engle, R.F., and Russell, J. R. (2002), Analysis of High Frequency data, forthcoming in Handbook of Financial Econometrics, ed. by Y. Ait-Sahalia and L. P. Hansen, Elsevier Science: North-Holland.

[68] Feng, D., Jiang, G.J., and Song, P. X. K. (2004). Stochastic conditional duration models with "leverage effect" for financial trancaction data. *Journal of Financial Econometrics,*. **2(3)**, 390-421.

[69] Fernandes, M. (2004), Bounds for the probability distribution function of the linear ACD process. *Statistics & Probability Letters*, **68**, 169-176.

[70] Fernandes, M., and Grammig, J. (2005). Nonparametric specification tests for conditional duration models.*Journal of Econometrics*, **127(1)**, 35-68.

[71] Fernandes, M., and Grammig, J. (2006). A family of autoregressive conditional duration models,*Journal of Econometrics*, **130(1)**,1-23.

[72] Getis, A., and Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic, *Gegraphical Analysis*, **36**, 90–1104.

[73] Grammig, J., and Kai-Oliver Maurer (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal* , **3**, 6–38.

[74] Granger, C. W. J. , and Anderson, A. (1978), Non-linear Time series modelling. *Applied Time series Analysis*, Academic press, Newyork.

[75] Graver, D. P., and Lewis, P. A. (1980), First-order autoregressive gamma sequences and point processes. *Advances in Applied Probability*,**12**, 727-745.

[76] Gupta, A. K., Ladusingh, L., and Borkotoky, K. (2016). Spatial clustering and risk factors of infant mortality: district-level assessment of high-focus states in India. *Genus*, **72(2)**.

[77] Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357- 384.

[78] Hao, J., and Godbole, A. (2014). Distribution of the maximum and minimum of a random number of bounded random variables. *Open Journal of Statistics*, **06(02)**.

[79] Han, L. , Braun, W.J., and Jason Loeppky (2018). Random coefficient minification processes. *statistical papers*, **61**, 1741-1762.

[80] Hassan, A.S., Sabry, M. A. H., and Elsehetry, A. M. (2020) . A new family of upper-truncated distributions: Properties and estimation. *Thailand Statistician*, **18(2)**, 196-214.

[81] Hautsch, N. (2002), modelling intraday trading activity using Box-Cox ACD Models, Discussion Paper 02/05, CoFE, University of Konstanz.

[82] Heyde, C. C., and Scott, D. J. (1973). Invariance principles for the law of the iterated logarithm for martingales and processes with stationary increments. *The Annals of Probability*, **1(3)**,428–436.

[83] Higgins, M. L., and Bera, A.K. (1992). A Class of Nonlinear Arch Models. *International Economic Review*. **33(1)** 33, 137–58. https://doi.org/10.2307/2526988.

[84] Hu Y., Peng X., Li T., and Guo H. (2007). On the Poisson approximation to photon distribution for faint lasers. *Physics Letters A*, **367**, 173-176.

[85] Hunt, L. R., DeGroat, R. D., and Linebarger, D. A. (1995), Non-linear AR modelling. *Circuits, Systems and Signal Processing*, **14**, 689–705.

[86] Li, Z., Hsiao, Y., Godwin, J., Martin, B. D., Wakefield, J., and Clark, S. J. (2019). Changes in the spatial distribution of the under-five mortality rate: small-area analysis of 122 DHS surveys in 262 sub-regions of 35 countries in Africa, *PLoS ONE*, **14(1)**: e0210645. https://doi.org/10.1371/journal.pone.0210645.

[87] Jara A., Nieto-Barajas, L. E., and Quintana, F. (2013). A time series model for responses on the unit interval. *Bayesian Analysis*, **8(3)**, 723–740.

[88] Jayakumar, K., and Thomas. (2004). Semi logistic distributions and processes. *Stochastic modelling and Applications*, **7**, 20-32.

[89] Jayakumar, K., and Thomas. (2008). On generalisation to Marshall- Olken scheme and its application to Burr type XII distributions. *Statistical papers*, **49(3)**, 421-439.

[90] Jayakumar, K., and Pillai, R. N (2002). A class of stationary Markov processes. *Applied mathematics letters.***15**, 513-519.

[91] Kalamkar, V. A. (1995). Minification processes with discrete marginals.*Journal of Applied Probability*, **32(3)**, 692-706.

[92] Kale, B. K. (2007). *A first course on Parametric Inference*. Narosa, New Delhi.

[93] Kelejian, H. H, and Piras, G. (2017). *Spatial Econometrics*. Academic Press.

[94] Kieschnick, R. , and McCullough, B. D. (2003). Regression analysis of variates observed on (0,1); percentages, proportions and fractions. *Statistical Modelling*, **3**, 193-213.

[95] Klimko, L. A., and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics*, **Vol6(3)**, 629-642.

[96] Krishna, K. K., Jose, K.K., Alice, T., and Ristić, M.M. (2013). The Marshall-Olkin Fréchet Distribution. *Communications in Statistics - Theory and Methods*, **42(22)**.

[97] Krishnarani, S. D., and Jayakumar K. (2008a). On bivariate logistic distributions and processes,*Far East Journal of Theoretical Statistics*, **24(1)**, 109-124, P-ISSN: 0972–0863.

[98] Krishnarani S.D. , and Jayakumar, K. (2008b). A class of autoregressive processes, *Statistics and Probability Letters*, **78**. 1355-1361,ISSN: 0167-7152.

[99] Krishnarani, S. D., and Jayakumar, K. (2013). Generalized half semi-logistic distributions and Processes, *International Journal of Statistics and Analysis*, **3(2)**. 123-134 (Online ISSN: Print ISSN: 2248 - 9959).

[100] Kus, C. (2007). A new lifetime distribution. *Computational Statistics and Data Analysis*, **51**, 4497-4509.

[101] Kuttykrishnan, A.P., and Jayakumar, K. (2008). Bivariate semi a-Laplace distribution and processes. *Statistical Papers*, **49**,303-313.

[102] Lawrance, A . J (1992). Uniformly distributed first-order autoregressive time series models and multiplicative congruential random number generators, *J. Applied Probability.***29**, 896-903.

[103] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, New York, USA, 2nd edition.

[104] Lawrance, A.J, Lewis, P. A. W. (1985) . Modelling and residual analysis of non-linear autoregressive time series in exponential variables, *J.Roy. Statist.Soc.Ser.***B 47**, 165-202.

[105] Legendre, P. (1993). Spatial autocorrelation. Trouble or new paradigm?, Ecology. *Ecological Society of America.*, **74(6)**, 1659-1673.

[106] Lee, L. F., and Yu, J. (2010). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, **40**.255—271.

[107] Lee, L. F., and Yu, J. (2011). Estimation of spatial panels. *Foundations and Trends in Econometrics*, **4**:1—164.

[108] Lee, L. F., and Yu, J. (2015). Spatial panel data models. *The Oxford Handbook of Panel Data*, 363-401. Oxford University Press, Oxford, UK.

[109] LeSage, J. P. (2008). An introduction to spatial econometrics. *Revued'e'conomie industrielle*, 19-44.

[110] LeSage, J. P., and Pace, R. K. (2009). Introduction to Spatial Econometrics. *CRC Press*, Boca Raton, FL.

[111] Lewis, P. A. W., and Mckenzie, E. (1991). Minification processes and their transformations. *Journal of Applied Probability*, **28**, 45—57.

[112] Lewis, P. A. W , Mckenzie, E., and Hugus, D. K. (1989) Gamma processes, *Stochastic models* **5(1)**, 1-30.

[113] Lewis, P.A.W., and Stevens, J.G. (1991). Non-linear modelling of time series using multivariate adaptive regression splines (MARS). *Journal of the American Statistical Association*, **87**, 864—877

[114] Li, Z., Hsiao, Y., Godwin, J., Martin, B. D., Wakefield, J., and Clark, S. J. (2019). Changes in the spatial distribution of the under-five mortality rate: small-area analysis of 122 DHS surveys in 262 subregions of 35 countries in Africa, *PLoS ONE*, **14(1)**: e0210645. https://doi.org/10.1371/journal.pone.0210645

[115] Li, W. K., and Yu, Philip, L. H. (2003). On the residual autocorrelation of the autoregressive conditional duration model.*Economic Letters*, **79(2)**, 169–175.

[116] Lin, S. X., and Tamvakis, M. N. (2004). Effects of NYMEX trading on IPE Brent Crude futures markets: a duration analysis. *Energy Policy* , **32**, 77–82.

[117] Louzada, F., Roman, M., and Cancho, V. (2011). The Complementary Exponential Geometric Distribution: Model, Properties, and a Comparison with Its Counterpart. *Computational Statistics and Data analysis*, **55(8)**, 2516-2524.

[118] Lu, J., and Zhang, L. (2010). Evaluation of parameter estimation methods for fitting spatial regression models, *Forest Science*, **56(5)**, 505–514.

[119] Lu, J., and Zhang, L. (2011). Modelling and prediction of tree height–diameter relationships using spatial autoregressive models, *Forest Science*, **57(3)**, 252-264.

[120] Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. *Journal of Economic Literature*

[121] Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. *The Assurance Magazine, and Journal of the Institute of Actuaries*, **8(6)**, 301-310.

[122] Mališić, J. D. (1987). On exponential autoregressive time series models, *Mathematical Statistics and Probability Theory Vol.**B*** , 147-153.

[123] Manuel, L., Scalon, J. D., Nogueira, D. A., and Moreira, J. M. (2018). Spatial linear regression models in the infant mortality analysis, *Multi-Science Journal*, **1(13)**, 39–44.

[124] Mckenzie, E. (1985).An autoregressive process for beta random variables, *Management Science **31***, 988–997.

[125] McLeish, D. L. (1974). Dependent Central Limit Theorems and Invariance Principles. *The Annals of Probability*, **2**, 620–6.

[126] Meitz, M., and Terasvirta, T. (2006). Evaluating models of autoregressive conditional duration.*Journal of Business and Economic Statistics*, **24(1)**, 104-124.

[127] Moran, P. A. P. (1948). The interpretation of statistical maps, *Journal of the Royal Statistical Society. Series B (Methodological)*, **10(2)**, 243–251.

[128] Moran, P.A.P. (1953). The statistical analysis of the Canadian lynx cycle. *Australian Journal of Zoology.* **1**, 163-173.

[129] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *The Statistician*, **37**, 25-32.

[130] Morais, A. L., and Barreto-Souza, W. (2011). A compound class of Weibull and power series distributions. *Computational Statistics and Data Analysis*, Volume **55(3)**, 1410-1425.

[131] Nadarajah, S. (2008), Some Truncated Distributions, *Acta Applicandae Mathematicae*, **106**, 105–123.

[132] Nelson, D.B. (1991). Conditional heteroscedasticity in asset pricing: A new approach. *Econometrica*, **59**, 347–370.

[133] Nelson, H. L., and Granger, C. W. J (1979). Experiences with using the Box-Cox transformation when forecasting economic time series. *Journal Econometrics*, **10**, 57-69.

[134] Nicholls, D.F., and Quinn, B.C. (1982). Random coefficient autoregressive models- An introduction. *Lecture Notes in Statistics 11, Springer Verlag*, New York.

[135] OECD (2019). Health expenditure in relation to GDP, in Health at a Glance 2019: OECD Indicators, OECD Publishing, Paris. DOI: https://doi.org/10.1787/592ed0e4-en.

[136] Paelinck, J. H. P., and Klaassen, L. H., Ancot, J. P., Verster, A. C. P., and Wagenaar, S. (1979). *Spatial Econometrics.*

[137] Pacurar, M. (2006). Autoregressive conditional duration(ACD) models in finance, A survey of theoretical and empirical literature in SSRN. Dalhouse University. Canada.

[138] Pezzulo, C., Bird, T. J., Utazi, C. E., Sorichetta, A., Tatem, A., Yourkavitch, J., and Burgert, C. R. (2016). Geospatial modelling of child mortality across 27 countries in Sub- Saharan Africa, DHS Spatial Analysis Reports No. 13, Rockville, Maryland, USA: ICF International.

[139] Pillai, R. N. (1991). Semi-Pareto processes. *Journal of Applied Probability*, **28**, 461–465.

[140] Pillai, R. N., Jose, K. K., and Jayakumar, K. (1995), Autoregressive minification process and universal geometric minima, *Journal of the Indian Statistical Association*, **33**, 53-61.

[141] Popovi$\hat{c}$, B. V., Risti$\hat{c}$, M. M., and Balakrishna, N. (2016), A mixed stationary autoregressive model with exponential marginals. *Statistical Papers*, **58**:1125-1148.

[142] Quijano Xacur, O.(2019). The unifed distribution . *Journal of Statistical Distributions and Applications*, **6(13)**.

[143] Ristic, M. M. (2005). A Beta-Gamma autoregressive process of the second-order (BGAR(2)), *Statistics and Probability Letters*, **73**, 403-410.

[144] Ristic, M. M. (2006). Stationary bivariate minification process.*Statistics and Probability Letters*, **76**, 439-445.

[145] Ristic, M.M. (2008). A generalized Semi-Pareto minification process. *Statistical Papers* **49**, 343- 351.

[146] Ristic, M. M., and Popovic, B. (2000a). A new uniform AR(1) time series model (NUAR(1)). *Publications de l'Institut Mathématique*, **68(82)**, 145-152.

[147] Ristic, M. M., and Popovic, B. (2000b). Estimating of parameters: NUAR(1) process, *Publications de l'Institut Mathématique*,**70(84)**, 63-68.

[148] Ristic, M. M., and Popvic, B. C. (2002). The uniform autoregressive process of the second order (UAR(2)). *Statistics and Probability Letters*, **57**, 113-119.

[149] Rocha, A.V. , and Cribari-Neto, F. (2009). Beta autoregressive moving average models. *Test*, **18**, 529-49.

[150] Russell, J. R and Engle, R. F. (2005), A discrete- state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial - autoregressive conditional duration models. *Journal of Business & Economic statistics* **23(2)**, 166-180.

[151] Sim, C. H. (1986). Simulation of Weibull and gamma autoregressive stationary process, *Communications in Statistics- Simulation and Computation*, **15**, 1141-1146.

[152] Sim, C. H. (1990). Simulation of Weibull and gamma and exponential processes, *Journal Applied Probability* **27**, 325-332.

[153] Sim, C. H. (1993). First-order autoregressive logistic processes, *Journal of Probability*, **30**, 467-470.

[154] Singh, A., and Masquelier, B. (2018). Continuities and changes in spatial patterns of under-five mortality at the district level in India (1991 and 2011), *International Journal of Health Geographics*, **17(39)**. https://doi.org/10.1186/s12942-018-0159-3.

[155] Slutsky, E. (1927). The summation of random causes as the source of cycle processes (in Russian). In Problems of Economic Conditions, English translation(1937), *Econometrica*, **5**, 105-146.

[156] Sokal, R.R. , and Oden, N. L. (1978a). Spatial autocorrelation in biology, 1. Methodology. *Biological Journal of the Linnaean Society*, **10**, 199-228.

[157] Stout, W. F. (1974). Almost Sure Convergence. Academic press, New York.

[158] Tahmasbi, R., and Rezaei, S. (2008). A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics and Data Analysis*, **52**, 3889-3901.

[159] Tavares, L. V. (1977). The exact distribution of extremes of a non-Gaussian process. *Stochastic Processes and their Applications*, **5(2)**, 151–156.

[160] Tavares, L. V. (1980). An exponential Markovian stationary process. *Journal of Applied Probability*, **17**, 1117–1120.

[161] Taylor, S. J. (1986). Modelling financial time series. *John Wiley & Sons*.

[162] Tiao, G. C., and Tsay, R. S. (1994). Some advances in non-linear and adaptive modelling in time series. *Journal of Forecasting*, **13**, 109–131.

[163] Tjøstheim, D. (1994). Non-linear time series: A selective review. *Scandinavian Journal of Statistics*, **21**, 97–130.

[164] Tiku, M. L., Wong, W. K., Vaughan, D. C., and Bian, G. (2000). Time series models in non-normal situations: Symmetric innovations. *Journal of Time Series Analysis*, **21 (5)**, 571–596.

[165] Tong, H. (1983). Threshold models in non-linear time series analysis, Lecture Notes in Statistics, No. 21, Springer, Heidelberg.

[166] Tong, H.(1990), *Non-linear Time Series: A Dynamical System Approach*,Oxford: Clarendon Press.

[167] Tong, H.(2022) , Non-linear time series analysis since 1990: Some Personal Reflections. *Acta Mathematicae Applicatae Sinica, English Series*, **18(2)**, 177-184.

[168] Tsay, R. (2005). *Analysis of Financial Time Series*, John Wiley & sons.

[169] Van Wagner, C.E. (1987). Development and structure of the Canadian Forest Fire Weather Index system. Technical Report 35, Canadian Forestry Service, Ottawa.

[170] Weigend, A.S., and Gershenfeld, N.A. (Eds.) (1993). *'The future of time series', in: Time Series Prediction: Forecasting the Future and Understanding the Past.* Addison-Wesley, Reading, MA.

[171] Weiss, G. (1975). Time reversibility of linear stochastic processes. *Journal of Applied Probability*, **12**, 831-836.

[172] Whittle, P. (1962). Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute*, **39**, 105-129.

[173] WHO (2021). Child mortality and causes of death, World Health Organization. https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/child-mortality-and-causes-of-death.

[174] Wold, H. O. A. (1938). *A study of the Analysis of Stationary Time Series.* (2nd ed.1954). Almqristand and Witsett, Stockhol.

[175] Xiang, K., and Song, D. (2016). Spatial analysis of China province-level perinatal mortality, *Iranian Journal of Public Health*, **45(5)**, 614-622.

[176] Xu, D., Knight, J., & Wirjanto, T. S. (2010). Asymmetric stochastic conditional duration modela mixture-of-normal approach. *Journal of Financial Econometrics*, **9 (3)**, 469–488.

[177] Yakowitz, S. J (1973). A stochastic model for daily river flow time series. *Water resources research*, **9**, 1271-85.

[178] Yeh, B.C., Arnold, and Robertson, C. A. (1988). Pareto processes, *Journal of Applied Probability* **25**, 291-301.

[179] Yu, L., Wang, S., and Lai, K.K. (2006). An integrated data preparation scheme for neural network data analysis. *IEEE Transactions on Knowledge and Data Engineering*, **18**, 217–230.

[180] Yule, U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. Philos. *Transactions of the Royal Society, London, series A*, **226**, 267-298.

[181] Zhang, M. Y., Russel, J. R., and Tsay, R. S. (2001). A non-linear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics***104**, 179–207.

LIST OF PUBLISHED WORKS

(a) **Published:**

(1) Krishnarani, S. D., Gautham Manoharan, and Vidya, V. P. (2021). Spatial modelling of child mortality rate, *Journal of Indian Statistical Association*, **59(2)**, 1-23.

(2) Krishnarani, S. D., and Vidya, V. P. (2022). On uniform truncated Poisson distribution and its applications, *Statistics and Applications*, **20(1)**, 279–295. ISSN 2454-7395.

(b) **Presentations in Conferences/Seminars:**

(1) "On Exponential Uniform distribution and its associated Time series Models", National Seminar on Theoretical and Applied Statistics organized by the Post Graduate and Research Department of Statistics, Farook College(Autonomous), Kozhikode, January 28-29 , 2020.

(2) "Uniform Truncated Autoregressive Conditional Duration models", International conference on emerging trends in statistics and data science in conjunction with 40^{th} annual convention of Indian Society for Probability and Statistics (ISPS) jointly organized by the Departments of statistics of Cochin University of Science and technology, Cochin, M. D. University, Rohtak, University of Kerala, Trivandrum, Bharathiar Univerity, Coimbatore, The Madura College(Autonomous), Madurai, September 7-10, 2021.

(3) "Spatial Auto-regressive Analysis of Mortality Data", International Conference on Statistical Machine Learning and Its Applications to COVID-19, organised by Department of statistics, WMO Arts and Science College, Muttil , October 4-6, 2021.