Modeling Multivariate Time Series

By
Tsamtsakiri Panagiota

Submitted to the Department of Statistics
of the Athens University of Economics and Business
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Athens, Greece
2023
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Διακριτές Πολυμεταβλητές Χρονοσειρές
Τσαμτσακίρη Παναγιώτα

Διδακτορική Διατριβή
Που υποβλήθηκε στο Τμήμα Στατιστικής
tου Οικονομικού Πανεπιστημίου Αθηνών
για την απόκτηση
Διδακτορικού Τίτλου στη Στατιστική

Αθήνα
2023
DEDICATION

To my siblings Kiki and Iraklis
ACKNOWLEDGEMENTS

It is even more difficult and strange to me this because words are not enough to express my feelings and to thank all people who were close to me and they supported me with my thesis.

First of all i would like to express wholeheartedly my gratitude to my supervisor Prof. Dimitris Karlis who was the leader and he taught me on how to think, to search and to learn. I would like to thank him for his support, his aid, his countless revisions contributing by this way at the improvement of the thesis. I feel proud to be one of his students. Many thanks to Prof. Ioannis Ntzoufras for his support whenever i asked him for.

I owe a debt of gratitude to my parents Theodoros and Georgia for their whole-life support, valuable trust to this step of my life and infinite love. Words cannot describe how grateful i am to my siblings Kiki and Iraklis for always being there to encourage me for everything in my life. You are the best! I am much appreciated to my husband Bill for his patient, fostering and for being in my life.

Finally i wish to acknowledge State Scholarship Foundation in Greece for the scholarship that offered to me.
«This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY)»
ABSTRACT

Modeling Multivariate Time Series

by Tsamtsakiri Panagiota

The study of time series has become a subject of interest last years especially for count data with heteroskedasticity. While there are numerous research studies with univariate models, in multivariate case the literature is limited. The reason for the above is that there are countless difficulties at the construction and analysis of multivariate counting processes. There are two general categories of time series models with different constructions leading to different ways of their theoretical properties analysis. While considering heteroskedastic time series models from the category ergodic properties can be studied more easier than the the second one, methods for estimation are difficult to established.

This thesis deals with the construction of multivariate Integer Generalized Autoregressive Conditional Heteroskedastic(INGARCH) and Conditional Autoregressive Range(CARR) time series processes. The INGARCH(1,1) model has been also studied in multivariate case and implementations in 2 dimensions have been also presented recent years. The drawback was found at the restricted values of correlation coefficient boundaries depending on the way of model's construction. Based on a family of copulas a multivariate INGARCH(1,1) model and a CARR(1,1) model are studied considering interdependencies and self-dependencies respectively. According to model’s complexity, its appropriateness and capability to study data with small sample sizes are examined and provided with simulations. H-steps ahead forecasting is considered by taking conditional expectation on volatilities and calculating marginal probability mass function.

Firstly, considering univariate INGARCH models where volatilities are linearly or log-linearly expressed offering more flexibility on conditions of stationarity respectively, a Bayesian Trans-dimensional Markov Chain Monte Carlo is provided. At a second stage a new alternative family of Sarmanov distribution is also presented in order to ameliorate boundaries of correlation coefficient and comparison with the known Sarmanov families are graphically discussed. A multivariate INGARCH(1,1) process is studied based on this alternative Sarmanov distribution and an implementation with daily crash counts on three towns in Netherlands is presented. A multivariate Conditional Autoregressive Range(CARR(1,1)) model assuming an exponential distribution and reconstructing correlation coefficients boundaries is discussed. The proposed model is illustrated on bivariate series from the fields of renewable sources.
ΠΕΡΙΛΗΨΗ

Διακριτές Πολυμεταβλητές Χρονοσειρές
Τσαμτσακίρη Παναγιώτα

Η μελέτη των χρονοσειρών αποτελεί αντικείμενο έρευνας τα τελευταία χρόνια ειδικά των διακριτών μοντέλων με ετεροσκεδαστικότητα. Ενώ υπάρχουν πολλές μελέτες στην μονοδιάστατη περίπτωση, στην πολυδιάστατη η βιβλιογραφία είναι περιορισμένη. Ο λόγος για το παραπάνω είναι ότι υπάρχουν αρκετές δυσκολίες στην δομή και την ανάλυση πολυμεταβλητών μοντέλων χρονοσειρών. Συμπεριλαμβάνοντας δύο γενικές κατηγορίες χρονοσειρών τα εμπόδια της πρώτης κατηγορίας σχετικά με την δομή και την ανάλυση, δεν είναι σημαντικές δυσκολίες της δεύτερης και το αντίστροφο. Λαμβάνοντας υπόψη τα παραπάνω ενώ στα ετεροσκεδαστικά μοντέλα η εργοδικότητα είναι εύκολο να μελετηθεί, η εφαρμογή μεθόδων για την εκτίμηση των παραμέτρων είναι δύσκολη. Επιπλέον όταν η σχετική συσχέτιση πρέπει να ληφθεί υπόψη, η εκτίμηση του μοντέλου γίνεται περισσότερο πολύπλοκη.

Στην διατριβή αυτή η δομή ενός πολυμεταβλητού Integer Generalized Autoregressive Conditional Heteroskedastic (INGARCH) και ενός πολυμεταβλητών Conditional Autoregressive Range (CARR) μοντέλου χρονοσειρών μελετώνται. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια.
Contents

1 Introduction
   1.1 Current models for time series ...................................................... 1
   1.2 Continuous-valued time series models ............................................. 2
      1.2.1 Autoregressive time series models of order 1 .............................. 2
      1.2.2 Heteroskedastic continuous-valued time series models ................. 6
   1.3 Integer-valued time series models ............................................... 11
      1.3.1 Integer-valued heteroskedastic time series models ...................... 11
   1.4 Stationarity and Ergodicity of OD models ........................................ 20
      1.4.1 Ergodicity ................................................................................. 20
      1.4.2 Consistency issue of Observation Driven(OD) Models ................... 24
   1.5 Multivariate heteroskedastic OD time series models ......................... 24
      1.5.1 Bivariate processes for continuous data .................................... 24
      1.5.2 Bivariate processes for discrete data ........................................ 27
   1.6 Objective & Structure of the thesis ................................................. 31

2 Bayesian models selection for INGARCH models .................................. 35
   2.1 Stationarity of linear and log-linear INGARCH models ....................... 37
   2.2 Parameters estimation ....................................................................... 39
   2.3 Parameters estimation based on Bayesian statistics methods ................ 41
   2.4 Methods for Bayesian model selection .............................................. 43
      2.4.1 Simulations ................................................................................. 48
      2.4.2 Applications ............................................................................... 50

3 A novel multivariate Sarmanov distribution ......................................... 57
   3.1 Sarmanov distribution & FGM copula ................................................. 57
   3.2 Families of copulas ............................................................................ 61
   3.3 Alternative construction of FGM copula ............................................ 62
   3.4 A proposed alternative FGM copula .................................................. 64
      3.4.1 Copulas in higher dimensions ..................................................... 78

4 A multivariate INGARCH(1,1) time series model ................................ 81
   4.1 Model ............................................................................................... 81
   4.2 Ergodic Properties ............................................................................. 82
      4.2.1 Weak dependence criteria ......................................................... 82
   4.3 Estimation ......................................................................................... 84
4.4 Simulations .................................. 87
4.5 Application .................................. 95
  4.5.1 Goodness of fit ................................ 98
  4.5.2 Predictive marginals ......................... 99

5 A bivariate CARR(1,1) model .................. 105
  5.1 The model .................................. 105
    5.1.1 Ergodic properties .......................... 107
  5.2 Ergodic properties ............................ 109
  5.3 Estimation .................................. 110
  5.4 Goodness of fit test ........................... 111
  5.5 Simulations .................................. 112
  5.6 Implementation ............................... 115

6 Main contribution & Future challenges ........ 119
  6.1 Main contribution ............................. 119
  6.2 Future challenges ............................. 121

7 Appendix ..................................... 125
  7.1 Appendix A: Bivariate distribution based on copulas ............. 125
  7.2 Appendix B: Partial derivatives of INGARCH process ............. 126
  7.3 Appendix C: Partial derivatives of CARR process ............... 127
List of Tables

2.1 Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size $n=100$, $M_i$ is the class of linear and $LM_i$ for log-linear INGARCH models accordingly. 49

2.2 Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size $n=200$, $M_i$ is the class of linear and $LM_i$ for log-linear INGARCH models accordingly. 50

2.3 Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size $n=500$, $M_i$ is the class of linear and $LM_i$ for log-linear INGARCH models accordingly. 50

2.4 Estimated Posterior Means and standard deviations (sd) for the competing models. 51

2.5 Geweke’s convergence z-scores for the parameters of five linear and five log-linear INGARCH models 52

2.6 Posterior probabilities and Bayes factor for the competing models 52

2.7 Estimated Posterior Means and standard deviations (sd) for the competing models. 53

2.8 Geweke’s convergence z-scores for the parameters of five linear and five log-linear INGARCH models 54

2.9 Posterior probabilities and Bayes factor for the competing models 54

4.1 Descriptive statistics for the time series of accidents 96

4.2 ML estimates for trivariate Poisson model parameters based on volatility construction defined by (1.14; 4.3). Numbers in parentheses are standard errors. 98

5.1 Mean, sd and MADE values after 200 replications of simulation experiment. 113

5.2 Mean, sd and MADE values after 200 replications of simulation experiment. 113

5.3 Descriptive statistics for the time series of electricity price ranges 117

5.4 ML estimates and se for model parameters. 117
List of Figures

1.1 Series of polio dataset (left) and its ACF (right) .......................... 12
2.1 Series of polio dataset (left) and its ACF (right) .......................... 51
2.2 Series of campylobacterosis dataset (left) and its ACF (right) ........ 53
2.3 Posterior density for \( \lambda_{T+1} \) for polio data ....................... 55
2.4 Posterior density for \( \lambda_{T+1} \) for Campylobacterosis data .......... 56
3.1 Values for correlation with equal values at both \( \lambda_i, k_i, i = 1, 2 \). Lower boundary(first graph) and upper boundary(second graph) .................. 72
3.2 Values for correlation with equal values at both \( \lambda_i, k_i, i = 1, 2 \). Lower boundary(first graph) and upper boundary(second graph) .................. 72
3.3 Correlation coefficient values considering numerous values for bases \( c_i \) and powers \( k_i \), lower boundary(first graph) and upper boundary(second graph) .................. 73
3.4 Correlation coefficient values for numerous equal values of \( \lambda_i \) and \( k_i \) ............................ 76
3.5 Correlation coefficient values when \( \lambda_i = k_j, i \neq j, i, j = 1, 2 \) .......... 77
4.1 Boxplot with results from the simulation experiment with initial parameter values \( (a_{ij}, b_{ij}, d_{ij}, \omega_{ij}) = (0.12, 0.1, 0.1, 0.22, 0.15, 0.18, 0.15, 0.18, 0.22, 0.16, 0.18, 0.18, 0.1, 0.1, 0.1, 0.15, 0.15, 0.12, 1, 1.5, 4, 5, 6) \) ................. 90
4.2 Boxplot with results from the simulation experiment with initial parameter values \( (a_{ij}, b_{ij}, d_{ij}, \omega_{ij}) = (0.35, 0.12, 0.05, 0.18, 0.12, 0.15, 0.1, 0.4, 0.1, 0.15, 0.15, 0.18, 0.1, 0.1, 0.1, 0.15, 0.15, 0.1, 1.5, 3, 2, 8, 4, 3, 1.5) \) .................. 92
4.3 Boxplot with results from the simulation experiment with initial parameter values \( (a_{ij}, b_{ij}, d_{ij}, \omega_{ij}) = (0.2, 0.2, 0.1, 0.25, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.2, 0.1, 0.1, 1.1, 1.5, 1, 8, 2, 2, 3) \) .................. 94
4.4 Time series plots of the series of accidents in Schiphol, De Bilt and Soesterberg .............................................. 96
4.5 ACF for the three time series ................................................. 97
4.6 Cross-correlation function .................................................... 97
4.7 Cumulative periodogram plots of the Pearson residuals for daily crash counts for Shiphol(left upper), De Bilt(right upper) and Soesterberg. .................. 99
4.8 One step ahead bivariate predictive probability for each one of the three places in Netherlands. .............................................. 100
4.9 Number of transactions per 15 seconds for IBM (top) and Coca-Cola (bottom). .............................................. 101
4.10 Auto- and cross-correlation function of the transaction data. . . . . . . 102
4.11 Boxplot of residuals $\hat{e}_{1,t}$ for IBM(left) and $\hat{e}_{2,t}$ for KO(right). . . . . . . 102

5.1 Boxplot with results from the simulation experiment with initial parameter values $(d_1,d_2,a_{11},a_{22},b_{11},b_{22},\omega)=(2,1.5,0.3,0.45,0.2,0.15,1.8)$ . . 114
5.2 Boxplot with results from the simulation experiment with initial parameter values $(d_1,d_2,a_{11},a_{22},b_{11},b_{22},\omega)=(3,0.5,0.4,0.6,0.5,0.3,0.2)$ . . . 114
5.3 Time series plots of the series of daily ranges between maximum and minimum electricity prices for Greece (first plot) and for Italy(second plot) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116
List of abbreviations

ACP . . . . . . Autoregressive Conditional Poisson
AR . . . . . . AutoRegressive
ARCH . . . . . . Autoregressive Conditional Heteroskedastic
ARMA . . . . . . Autoregressive Moving Average
BAR . . . . . . Bivariate AutoRegressive
BCP . . . . . . . Bivariate Conditional Poisson
BIN . . . . . . Binomial
BNB . . . . . . Beta Negative Binomial
CARR . . . . . Conditional AutoRegressive Range
CCC . . . . . . . Conditional Constant Correlation
cdf . . . . . . . Cumulative distribution function
CGINAR . . . Combined Geometric INteger AutoRegressive
CPINAR . . . Compound Poisson INteger AutoRegressive
DACP . . . . . Double AutoRegressive Conditional Poisson
DCC . . . . . . Dynamic Conditional Correlation
DINGARCH . Dynamic INteger Generalized AutoRegressive Conditional Heteroskedastic
DSTCC . . . Double Smooth Transition Conditional Correlation
EAR . . . . . . Exponential AutoRegressive
ECARR . . . Exponential Conditional AutoRegressive Range
ECCC . . . . . Extended Conditional Constant Correlation
EGB2 . . . . . Exponential Generalized Beta of second order
FCGARCH . . Flexible Coefficient Generalized AutoRegressive Heteroskedastic
FGM . . . . . . Farlie-Gumbel-Morgenstern
FINCARCH . Functional INGARCH
GARCH . . . . Generalized AutoRegressive Conditional Heteroskedastic
GEAR . . . . . Generalized Exponential AutoRegressive
GED . . . . . . Generalized Error Distribution
GEV . . . . . . Gumbel Extreme Value
GFDCC . . . . Quadratic Flexible Dynamic Conditional
GJR . . . . . . Glosten-Jagannathan-Runkle
GLM . . . . . . Generalized Linear Models
GO-GARCH . . Generalized Orthogonal GARCH
GOF-GARCH . . Generalized Orthogonal Factor GARCH
IGARCH . . . . Integrated Generalized AutoRegressive Conditional Heteroskedastic
i.i.d . . . . independent identically distributed
INAGARCH . . INteger Assymmetric eneralized AutoRegressive Conditional Heteroskedastic
INBL . . . . . . INteger BiLinear
MA . . . . . . . . Moving Average
MCMC . . . . Markov Chain Monte Carlo
MINARCH . . . Mixture INteger e AutoRegressive Conditional Heteroskedastic
MLE . . . . . . Maximum Likelihood Estimation
MSM . . . . . . Marginal Specific Modeld
NB . . . . . . . . Negative Binomial
NEAR . . . . . . New Exponential AutoRegressive
NIG . . . . . . . Normal Inverse Gamma
NLAR . . . . . . New Laplace AutoRegressive
NMEAR . . . . New Multivariate Exponential AutoRegressive
OD . . . . . . . . Observation Driven
O-GARCH . . Orthogonal Generalized AutoRegressive Conditional Heteroskedastic
PAGARCH . . . Periodic Asymmetric Generalized AutoRegressive Conditional Heteroskedastic
PDM . . . . . . Probability Distributed Model
pdf . . . . . . Probability Density Function
pgf . . . . . . Probability Generating Function
PIG . . . . . . Poisson Inverse Gaussian
pmf . . . . . . Probability Mass Function
PSINAR . . . . p-order Signed INteger AutoRegressive
PTARCH . . . . Power Tranformation Threshold AutoRegressive Conditional Heteroskedastic
QARCH . . . . Quadratic AutoRegressive Conditional Heteroskedastic
QMLE  . . . . .  Quasi Maximum Likelihood Estimation
RCAR  . . . . .  Random Coefficient AutoRegressive
RCINAR . . .  Random Coefficient INteger AutoRegressive
RGS  . . . . .  Regime Switching
RJMCMC . . .  Reversible Jump Markov Chain Monte Carlo
RRINGARCH .  Random Rounded INteger AutoRegressive Conditional Heteroskedastic
RrNGINGARCH  Random Environment New
RSDC . . . . .  Regime Switching Dynamic Correlation
SINAR . . . .  Signed INteger AutoRegressive
STCC . . . . .  Smoothed Transition Conditional Correlation
STGARCH . .  Smooth Transition Generalized AutoRegressive Conditional Heteroskedastic
STINAR . . .  Skew True INteger AutoRegressive
TLAR . . . . .  Transposed Laplace AutoRegressive
TVCC . . . . .  Time Varying Conditional
VC-GARCH . . Varying Correlation Generalized AutoRegressive Conditional Heteroskedastic
ZI . . . . . .  Zero Inflated
ZICP . . . . .  Zero Inflated Compound Poisson
Chapter 1

Introduction

1.1 Current models for time series

A series of observations \( \{X_t\}_{t=0}^T \) is defined as a count series if \( \{X_t\} \in \mathbb{N}_0, \forall \ t \) and as continuous series if \( \{X_t\} \in \mathbb{R}, \forall \ t \). Cox et al. (1981) classifies time-varying parameter models into two categories: observation driven (OD) models and parameter driven (PD) models. At the first one parameters depend on functions of lagged dependent values and past observations. An example of an observation driven model is:

\[
X_t \mid x_{t-1}, y_t \sim Poisson(\exp(y_t^T \beta + f(x_{t-1})))
\]

where \( y_t \) is a p-dimensional row vector and \( \beta \) is a vector of regression coefficients and \( f(\cdot) \) can be any function. Sometimes exogenous variables are needed on parameters expressions offering flexibility on the order of the model. Autoregressive (AR) and Integer valued AR (INAR) models considering different thinning operators proposed by Al-Osh and Alzaid (1987); McKenzie (1985), continuous ARMA, discrete ARMA (DARMA) and Integer valued ARMA (INARMA) models of Al-Osh and Alzaid (1988); Alzaid and Al-Osh (1990); Jacobs and Lewis (1978), Integer Autoregressive Conditional Heteroskedastic (INARCH) and Integer-valued Autoregressive Conditional Heteroskedastic models of Ferland et al. (2006); Fokianos et al. (2009); Fokianos and Tjøstheim (2011) are representative examples of OD models with applications to various disciplines such as finance Bollerslev (1986), epidemiology and public health Fokianos et al. (2009); Ferland et al. (2006); Fokianos and Tjøstheim (2011) On the other hand at parameter driven models parameters are stochastic processes that are subject to their own error. Furthermore at this category of models observed data are assumed to be independent given latent variables:

\[
X_t \mid \epsilon_t, y_t \sim Poisson(\exp(y_t^T \beta + \epsilon_t))
\]
where \( y_t \) are defined as previously whilst \( \epsilon_t \) is a latent process usually of a standard form as a classical time series model e.g. \( \epsilon_t = \alpha \epsilon_{t-1} + r_t \) where \( r_t \) are i.i.d. \( N(0, \sigma^2) \).

Models studied by Zeger and Qaqish (1988) as an extension to the generalized linear model (GLM), stochastic volatility models proposed by Shephard (1995) and hidden Markov models MacDonald and Zucchini (1997) are examples at the class of PD models. On the other hand in a PD autocorrelation expressed by using latent variables. On the other hand at the category of OD models AR, OD and PD models have crucial different properties and usages. Parameter driven models are hard to estimate because likelihood function cannot be always written analytically and parameters estimation seems to be more challenging than these in OD models but stability properties including stationarity and mixing conditions are easier to established.

1.2 Continuous-valued time series models

1.2.1 Autoregressive time series models of order 1

The first stochastic model that was useful on the representation and study of series is the autoregressive model (AR) in which the current value of the process is expressed as a finite, linear aggregate of its previous values and a variable called innovation \( \epsilon_t \). The model is defined as:

\[
X_t = \phi_1 X_{t-1} + \epsilon_t
\]

or equivalently

\[
\phi(B)X_t = \epsilon_t
\]

where \( \phi(B) = 1 - \phi_1 B \), B is a back shift operator defined by \( BX_t = X_{t-1} \), \( \epsilon_t \) is a white noise, a sequence of uncorrelated random variable with 0 mean and variance \( \sigma^2_\epsilon \). The above model is stationary under certain conditions. Another model expressed as a finite linear combination of white noise variables, proposed by George (1970) and defined as:

\[
X_t = \epsilon_t - v_1 \epsilon_{t-1}
\]

or equivalently

\[
X_t = v(B)\epsilon_t
\]

where \( v(B) = 1 - v_1(B) \) is a polynomial associated with the moving average (MA) model. Due to the expression of construction, the model is always stationary and it is capable to identify long-trends in data and smooth out short-term fluctuations. Numerous models have been studied and distinguished in two categories. By the
mean of construction where a marginal distribution is firstly specified and then the error distribution is obtained, the first class called class of marginal specific models (MSM). In contrast in the second class of error specific models, the analysis of model properties is done by specifying a suitable distribution for the i.i.d. error variable $\epsilon_t$ without bothering on the exact form of the marginal distribution of the stationary time series $X_t$. A variety of the presented models is also considered to be part of random coefficient autoregressive models where the distribution of $X_t$ derived from the distribution of error term $\epsilon_t$.

The first non-Gaussian AR(1) model was the one with stationary exponential marginal distribution introduced by Gaver and Lewis (1980). Considering that $\epsilon_t$ is a sequence of i.i.d. non-negative random variables, chosen in such a way that $X_t$ follows a specified stationary marginal distribution. That is, one has to find the form of the distribution of $\epsilon_t$ for a specified distribution of $X_t$ considering properties of generating functions (g.f.) or characteristic functions (c.f.). This means that if the function of $X_t$ is self-decomposable then there exists a well-defined distribution for the innovation $\epsilon_t$ as a consequence of $\phi_{X_t}(s) = \phi_{X_{t-1}}(s)\phi_{\epsilon_t}s$ where $\phi$ is the characteristic function. Based on the above, Exponential Autoregressive EAR(1) model expressed as:

$$X_t = \epsilon_t + \begin{cases} \phi X_{t-1} & \text{w.p. } \alpha \\ 0 & \text{w.p. } 1-\alpha \end{cases}$$

The zero-defect of the EAR(1) model means that the relationship between the current observation and past observations decays exponentially over time and this was regarded as one of the drawbacks of that model. To overcome the problem Lawrance and Lewis (1985) constructed a transposed exponential autoregressive (TEAR) model keeping as marginals distribution exponential distributions. Moreover Lawrance and Lewis (1981) generalized the EAR(1) model by introducing other parameter $\phi$ known as new exponential autoregressive model (NEAR(1)). The benefit of this model is that it has an expression of simple random similar combinations of exponential independent and thus it is easy to estimate. A generalized version of the NEAR(1) model proposed by Sim (1990) in order to define a stationary sequence of exponential random variables. Efforts for constructions of Gamma AR models based on thinning operators have been made by Gaver and Lewis (1980); Sim (1990). The first Gamma AR model based on marginal specific models where marginal distributions are $\text{Gamma}(p, \lambda) = \frac{x^{p-1}\exp(-\frac{x}{\lambda})}{\lambda^p \Gamma(p)}$, where $\text{Gamma}(p)$ is a Gamma function with $p, \lambda > 0$, shape and scale parameters respectively while at the second one the
process defined as:

\[ X_t = \lambda * X_{t-1} + \epsilon_t, \ t = 1, 2, \ldots \]

where \( \lambda * X_t = \sum_{i=1}^{N(X)} E_i \), \( E_i \) are exponential random variables with parameter \( \beta \) and for each fixed value of \( X_t \), \( N(X) \) is a positive random variable with parameter \( \lambda \beta x \). Pitt et al. (2002); Pitt and Walker (2005); Pitt and Walker (2006) suggested alternative methods of constructing the distribution of \( \epsilon_t \) by specifying suitable conditional and marginal distributions. They introduced some latent variables to get the specified marginal distributions for the AR(1) sequences. A mixed exponential AR(1) model defined by Lawrance (1980) and a random coefficient AR(1) model with marginal mixed exponential distribution defined by Lawrance and Lewis (1985) where \( X_t \) follows an exponential distribution if the distribution of innovations is a mixture of exponential distributions. Similar to the case of EAR(1) model a Constant coefficient Laplace and a Random coefficient Laplace AR model referred as transposed Laplace AR(1)(TLAR) models also defined by Sim (1990). To bypass the problem of zero-defect in the LAR(1) process Sim (1993) studied an autoregressive process with symmetric Laplace marginal distributions with location 0 and scale parameter 1 as:

\[ f(x) = \frac{1}{2} \exp(-|x|), \ x \in \mathbb{R}. \]

The construction of an autoregressive time series model with stationary marginal Gumbel Extreme Value(GEV) distribution given by:

\[ f(x; \mu, \sigma) = \exp \left\{ - \exp \left( -\frac{x - \mu}{\sigma} \right) \right\}, \ -\infty < x, \mu < \infty, \sigma > 0 \]

was constructed to model extreme values of a dataset. This distribution is widely used in many areas such as engineering to model maximum wind speed, in finance to model daily loss or gain in stock prices. Based on the above distribution Toulemonde et al. (2010) and Balakrishna and Shiji (2014) constructed Gumbel’s extreme value AR model summarizing the properties of a first order model. Many other stationary autoregressive time series processes also constructed considering each time different innovation’s distribution. At the bibliography there are references of Cauchy AR process discussed by Tauchen and Pitts (1983), Logistic AR process constructed by Sim (1993) and Hyperbolic Secant autoregressive processes Rao and Johnson (1988). They showed that the autocorrelation function of stationary processes mentioned previously can decay slowly, which implies that the processes exhibits long-range dependence. Sometimes, suitable mixtures of more than one standard distribution
may be a better choice compared to any single distribution. Normal-Laplace is one such distribution introduced by Reed and Jorgensen (2004), which is a convolution of normal and Laplace distributions. Jose et al. (2008) and Tomy and Jose (2009) discussed first-order autoregressive processes with normal-Laplace and generalized normal-Laplace stationary marginal distributions, where such processes provide the combined effect of Gaussian and non-Gaussian time series models. There are also other time series models with similar stationary marginal distributions in the literature. For example, Marshall–Olkin Esscher transformed Laplace distribution (George and George, 2013), mixed asymmetric Laplace distribution (Krishnan and George, 2017) and the references therein.

By combining moving average and autoregressive models, one gets what is known as autoregressive moving average (ARMA) model defined as:

\[ X_t = \phi_1 X_{t-1} + \theta_1 \epsilon_{t-1}, \epsilon_t \sim N(0, \sigma^2) \]

where \( \mathcal{F}_{t-1} \) is the \( \sigma \)-field algebra generated by \( X_0, X_t, t \in \mathbb{N} \), \( \sigma_t = \text{var}(X_t | \mathcal{F}_{t-1}) = \text{Var}(\epsilon_t | \mathcal{F}_{t-1}) \). The advantage of the above model is that it gives a more parsimonious model with relatively few unknown parameters. ARMA models with exponential stationary marginal distributions were introduced to study the sequences of non-negative random variables. For example, Jacobs (1978); Jacobs (1980) analysed the queuing networks using exponential ARMA(1, 1) model. Balakrishna and Ranganath (2015) analysed Bombay Stock Exchange data using ARMA(1, 1) model with generalized error distributed innovations. While the range of marginal distributions becomes too restricted then the innovations have non-standard distribution with discontinuities. Lewis et al. (1989) introduced a class of random coefficient models to generate a dependent sequence of Gamma\((p, \lambda)\) random variables where, the coefficients are i.i.d. Beta random variables.

The logical extension of AR(1) model to AR(2) has been introduced by George (1970) where an Autoregressive Integrated Moving Average (ARIMA) model was introduced and an AR(2) was included as a special case. Of course \((X_{t-1}, X_{t-2})\) follows a bivariate normal distribution for a Gaussian AR model and hence it is easy to solve this problem. In order to avoid such complications in non-Gaussian cases, Lawrance and Lewis (1980) proposed an AR(2) model based on the work of Gaver and Lewis (1980).

Lawrance and Lewis (1985) studied simultaneously Exponential autoregressive models (EAR) of order 1 and 2. The NEAR(1) has been generalized to an NEAR(2) model identified as Random Lag Autoregressive RLAR(2) model incorporating ran-
domness of the selection in terms. Numerous models have been studied in order to capture complex relationships between past and current values of time series providing flexibility at models forecasting. Such models have been studied by by

Jayakumar and Kuttykrishnan (2007) and Kim and Lynne (1997). Several models where the authors concentrate on the definition, the existence of the innovation distribution for specified marginal distributions and their second-order properties have been proposed in bibliography by Box et al. (2015), Trindade et al. (2010), Polasek and Pai (2004). Many other models were studied such autoregressive models with normal inverse Gaussian innovations(NIG), finite mixture autoregressive models Le et al. (1996);Wong and Li (2000);Wong et al. (2009);Nguyen et al. (2016) autoregressive models with slowly varying innovations,Davis and McCormick (1989);Feigin and Resnick (1992);Abraham and Balakrishna (1999);Ing and Yang (2014).

1.2.2 Heteroskedastic continuous-valued time series models

Extreme values in return series with financial data have been captured with heteroskedastic time series models. The idea for the construction of the first heteroskedastic model has been generated considering a non-standard ARMA model for $\epsilon^2_t$, introduced by Engle (1982) well known as Autoregressive Conditional Heteroskedastic(ARCH) model and has the following form

$$
\epsilon_t = \sigma_t \epsilon_{t-1}, \quad \sigma^2_t = \omega + \sum_{i=1}^{m} a_i \epsilon^2_{t-i}
$$

where $\epsilon_t$ is the error term, $\sigma^2_t$ is the conditional variance of the error term at time $t$, $\omega > 0$ and $a_i \geq 0$ for $i \geq 1$. One of the weaknesses of the above model is that both positive and negative values of $\epsilon^2_{t-i}$ provide similar effect on volatility as it depends on their previous values. Moreover based on the construction of volatility parameters $a_i, \quad i = 1, 2, ...$ must be positive and considering values of series moments, values of $a_i$‘s are too restrictive. Bollerslev (1986) proposed a Generalised Autoregressive Conditional Heteroskedastic (GARCH) model in order to gives flexibility for the model of time-varying conditional variance assuming that the variance of the error term expressed not only of squared error term values at previous lags but also of the conditional variance at previous lags.

$$
\sigma_t = \epsilon_t Z_t,
$$
where
\[ \sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \sigma_{t-j}^2. \]

After comparison with the ARCH model, GARCH model is more parsimonious because its volatility has terms of variance in past. One drawback of the model is that it takes into account only the movement of \( z_t \). Conditions of stationarity are some of the properties that have been discussed at the study of Bollerslev (1986). A potential disturbance was that the unconditional variance of the process to be modeled did not exist. Nelson (1991) also showed that when an Integrated Generalized Conditional Heteroskedastic (IGARCH) process starts at some infinite time point, its behaviour depends on the intercept \( \omega \) and if the intercept is positive then the unconditional variance of the process grows linearly with time. Considering the fact that a GARCH model could be expressed as an ARMA model equivalently an IGARCH model could be expressed as an ARIMA model. An attempt to find which parameters are capable to decrease slowly autocorrelation function, was done by Baillie et al. (1996) and the models is known as first fractionally integrated GARCH(FIGARCH) model. In these models, the underlying volatility of the data is not directly observable and instead is modeled as a hidden variable that evolves over time according to some stochastic process. The purpose is to consider parameters using and modeling data at a high frequency within a single day named intra-daily data. Engle (2002) proposed a GARCH-X model including exogenous variables \( r_{t-l} \) as measures of intraday volatility. While the model was more flexible allowing modeling intraday data, the shortcoming is that \( r_t \) is independent of \( \sigma_t \). An attempt allowing flexibility on the range of parameters’ values introduced by Nelson (1991) where volatility expressed by a log-linear form. Exponential Generalized Conditional Heteroskedastic(EGARCH) and Exponential Generalized Conditional Heteroskedastic with exogenous variables(EGARCH-X) are extensions of the ARCH/GARCH models and allows for asymmetric responses of volatility.

Another class of models assuming additional latent volatility processes for each exogenous variable considered at the volatility expression for modeling high frequency data within a single day, has been composed considering the models proposed by {Engle and Gallo (2006), Shephard and Sheppard (2010)}. Assuming numerous functions as density functions for error term, Engle and Gallo (2006) studied estimators consistency and asymptotic normality are consistent and asymptotically normal. The shortcoming of poor forecasting performance convey structural changes of the ARCH process leading to studies of other model proposed by Hamilton (1989),
Diebold and Inoue (2001), Lamoureux and Lastrapes (1990), Hamilton and Susmel (1994), Engle and Lee (1999), Granger and Hyung (2004) and Medeiros and Veiga (2004). An important issue at this class of models was made by Hansen et al. (2012) who introduced models with endogenous variables to model intradaily data linked to the return variance by one more equation in the model. Furthermore based on the work of Nelson (1991) a Realized EGARCH(1,1) model had been also discussed. A Flexible Coefficient Generalized Autoregressive Conditional Heteroskedastic (FCGARCH) model with many features observed in empirical financial distributions including time-varying conditional variance, asymmetry, thick tails, and high peakedness. At the above models conditional distribution of $y_t$ is an exponential generalized beta distribution of the second kind (EGBD2) introduced by McDonald (1991); McDonald (2008) and studied further from Wang et al. (2001).

Definition 1. The exponential generalized beta distribution defined as

$$EGBD2(z; \delta, \sigma, p, q) = \frac{\exp \left( \frac{p(z-\delta)}{\sigma} \right)}{|\sigma| B(p, q) \left( 1 + \exp \left( \frac{z-\delta}{\sigma} \right)^{p+q} \right)}$$

where $\delta$ is the location parameter that affects the mean of the distribution, $\sigma$ reflects the scale of the density function, and $p$ and $q$ are shape parameters that together determine the skewness and kurtosis of the distribution.

An other class of GARCH models capable to accommodate asymmetries in the response have been studied by Glosten et al. (1993); Engle and Ng (1993); Sentana (1995). Construction, study and modeling of heteroskedastic models where not only variance but also standard deviation given by a volatility expression proposed by Ding et al. (1993); Taylor (2008); Schweer and Weiß (2014). Moreover considering special cases of GARCH model Duan (1997) constructed an augmented GARCH model. The first non-linear model named smooth transition GARCH model (STGARCH) proposed by Hagerud (1997); González-Rivera (1998); Anderson et al. (1999). Lanne and Saikkonen (2005) Zakoian (1994) and Rabemananjara and Zakoian (1993) proposed two other threshold models without a wide acceptance due to complex construction and behavior of standard deviation. Li and Li (1996) studied a double threshold ARCH(DTARCH) model where both conditional mean and conditional variance have a threshold-type structure as defined from Tong (1990). The observed problem at the DTARCH and STGARCH model is that when $k \rightarrow \infty$ then the models are not identifiable. To overcome the above problem Hagerud (1997) and González-Rivera (1998) proposed some criteria. Engle III and Sheppard
1.2. CONTINUOUS-VALUED TIME SERIES MODELS

(2001) referred that at applications the assumption of GARCH models having constants parameters could not be appropriate when the series are too long. To bypass the problem, the assumption that parameters change at specific points of time, divide the series into subseries provided a time-varying GARCH(TV-GARCH) model. An other one way to model data containing breaks is the construction of Hidden Markov models or Markov switching models. Hamilton and Susmel (1994) studying a Markov-switching model in order to model very large shocks of stocks in 1987. Cai (1994) proposed an alternative construction of the above model. Furthermore Rydén et al. (1998) showed that a simplified model of this proposed by Cai (1994) was capable to model daily data with high frequency. Numerous other Markov-switching models have been also proposed by Granger and Ding (1995);Gray (1996);Klaassen (2002);Haas et al. (2004);Liu (2006);Lange and Rahbek (2009). An overview of ARCH/GARCH models had been given by Teräsvirta (2009). The increased automatization of financial markets and intradaily data in exchanges lead to study and modeling time between events. Moreover in the field of finance daily or monthly differences of $\max(P_t) - \min(P_t)$ where $P_t$ is the logarithmic price of high-return assets given by a Geometric Brownian motion leads to the importance of a new heteroscedastic time series process. Considering the above, two models have been also discussed. The first namely Autoregressive Conditional Duration(ACD) model studied by Engle and Russell (1998) and Dufour and Engle (2000). Assuming that $\{t_0, t_1, \ldots, t_n\}$ be a sequence of arrival times and $x_i = t_i - t_{i-1}$ denotes the duration between events occurred at $t_{i-1}$ and $t_i$. The model for durations $X_i$ defined as:

$$X_i = \lambda_t \epsilon_t$$

where $\lambda_t = E(X_t \mid \mathcal{F}_{t-1})$ linked linearly with its previous values and past observations of $x_t$ representing the time since the last event at time $t$ and $\epsilon_t$ is a sequence of i.i.d. random variables. The second model known as Conditional Autoregressive Range(CARR) model given by:

$$R_t = \lambda_t \epsilon_t$$  \hspace{1cm} (1.1)

where $R_t$ is the range of the financial asset at time $t$, defined as the difference between the high and low prices over a given time period, $\lambda_t$ is equal to conditional expectation of $R_t$ up to time $t$ representing the volatility and expressed linearly with range values of previous time periods and $\epsilon_t$ is a sequence of positive independent and identical random variables of the normalized range. The main difference between those two models is that while the time interval in the CARR model is constant in the ACD model is not. For both models numerous distributions for the error terms
have been discussed. The standard ACD model is the Exponential ACD(1,1) which is the counterpart of the GARCH(1,1) stochastic process. Firstly it has been studied by Engle and Russell (1998) and then autocorrelation function and some other properties have been discussed by Bauwens and Giot (2000). Furthermore an alternative model allows for flexibility at parameter values of volatility construction discussed by Bauwens et al. (2001) where logarithmic scale of volatility defined similar to this of Fokianos and Tjøstheim (2011). Properties of ergodicity and stationarity of first order linear and log-linear Exponential ACD processes studied by Carrasco and Chen (2002) and Meitz and Teräsvirta (2006). studied a Generalized Gamma ACD model observing that the data fit the model better than other models. Zhang et al. (2001) introduced a Threshold Autoregressive Conditional duration model(TACD) based on the threshold GARCH model studied by Zakoian (1994); Rabemananjara and Zakoian (1993) in order to capture asymmetries in volatility. To model and forecast the dynamics of the price range, Chou (2005) studied the first Conditional Autoregressive Range(CARR) model that is very similar to the GARCH model of Bollerslev (1986) and the autoregressive conditional duration (ACD) model of Engle and Russell (1998). Then Xie and Wu (2017) proposed a Gamma CARR model and Chou (2006) studied an asymmetric CARR model. Chou (2005) and Chou et al. (2010) show that the range-based CARR model produces more accurate volatility Fernandes et al. (2005) and Chou et al. (2015) studied a CARR model when the range expressed from the difference max(Price)-min(Price) where prices are daily or monthly. At the second stochastic process studied the time between events. The main difference is that at the first model the time interval is fixed but at the second is not. Our study is concentrated at a bivariate CARR model where volatilities expressed linearly by their past values and past volatilities. Assume that \( R_t = \lambda_t \epsilon_t \) where \( \epsilon_t \) is a sequence of positive independent and identical random variables with mean equal to 1 and \( \lambda_t \) is equal to conditional expectation of \( R_t \) up to time \( t \). A drawback with models in financial field is that autocorrelation function is decreased up to time 5 and the order of the model becomes large. A way to simplify the model is to incorporate exogenous variables in the volatility. Chou (2005) made a brief discussion by proposing the CARRX model where \( R_t \) given again by (1.1) and \( \lambda_t = d_0 + \sum_{i=1}^{p} a_i \lambda_{t-i} + \sum_{j=1}^{q} b_j R_{t-j} + \sum_{l=1}^{L} \gamma_{t-l} X_{t-l} \). Yatigammana et al. (2016) studied a model where errors follow a Weibull distribution. Similar to GARCH models Chiang et al. (2016) and Bauwens and Giot (2000) proposed lognormal log-CARR and log-ACD model respectively providing flexibility at the range of parameters values considering stationarity conditions. combine a CARR model with a DCC model
1.3 Integer-valued time series models

1.3.1 Integer-valued heteroskedastic time series models

The construction of integer-valued time series models was conducted from the necessity of modeling data as the number of incidences from a certain disease or the number of people waiting at a cashier in discrete time points. Numerous models for studying data of various disciplines such as sports, epidemiology, ecology and marketing proposed by Grunwald et al. (2000). The first integer-valued time series model was firstly introduced by McKenzie (1985); Al-Osh and Alzaid (1987).

Integer-valued heteroskedastic time series models have a vital role based on their construction where parameters or low order conditional moments expressed via past observations and past values assuming that those conditional moments change over time. Data with extreme values over time could be study by using heteroskedastic models. The first Integer-valued Autoregressive Conditional Heteroskedastic (INARCH) model given by:

\[ X_t | \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \]  

(1.2)

where

\[ \lambda_t = d_1 + a_1 \lambda_{t-1}. \]

\( \mathcal{F}_{t-1} \) is the \( \sigma \)-algebra defined as \( \{ X_s, s \leq t \} \) and parameter \( a_1 \in [0,1] \). At the INARCH model dispersion of data derived from the construction of conditional moment \( \lambda_t \). Coefficient \( (a_i) \) values affect moment \( \lambda_t \) at time \( t \). Those values at volatility lead to data with extreme values. Furthermore a wider range at volatility permitted considering not only past values of its own but also past observations. The above model is the almost known Integer-valued Generalized Autoregressive Conditional Heteroskedastic(INGARCH) model and defined as:

\[ X_t | \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \]  

(1.3)

where \( \mathcal{F}_{t-1} \) is the \( \sigma \)-algebra defined as \( \{ X_s, s \leq t \} \) and

\[ \lambda_t = d_1 + a_1 \lambda_{t-1} + b_1 X_{t-1}. \]  

(1.4)
At the graphs below we present data studied from an integer valued autoregressive time series (INAR(1)) plot with Poisson innovations defined as:

\[ X_t = a_1 \circ X_{t-1} + R_t \]

where

\[ R_t \sim \text{Poisson}(\lambda) \quad \text{and} \quad a_1 \circ X = \sum_{i=1}^{X} Y_i \]

and data studied by an integer autoregressive heteroskedastic model (INGARCH(1,1)):

\[ X_t \mid \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \quad (1.5) \]

where \( \mathcal{F}_{t-1} \) is the \( \sigma \)-algebra defined as \( \{X_s, s \leq t\} \) and

\[ \lambda_t = d_1 + a_1 \lambda_{t-1} + b_1 X_{t-1} \]

Furthermore OD models are a way to study data where autocorrelation function decays slowly. Rydberg and Shephard (2000) introduced an heteroskedastic time series process modeling the price of an asset at time \( t \) \( p(t) \) given by

\[ p(t) = p(0) + \sum_{i=1}^{N(t)} Z_i \]

where \( N(t) \) is the number of trades recorded up to time \( t \) and \( Z_i \) is the change of price associated with the \( i^{th} \) transaction. An alternative expression for the price \( p(t) \) given by:

\[ p_t = p((t+1)\Delta) - p(t\Delta) = [N(t+1)\Delta] - \sum_{i=1}^{N(t)\Delta} Z_i = \sum_{i=N(t)\Delta+1}^{N((t+1)\Delta)} Z_i \]

**Figure 1.1:** Series of polio dataset (left) and its ACF (right)
1.3. INTEGER-VALUED TIME SERIES MODELS

Considering the idea to "bin" out the time interval and count the number of transactions, the first integer valued conditional autoregressive heteroskedastic model was constructed for the number of transactions that occurs in the interval \([n\Delta, (n + 1)\Delta]\) and expressed as: \(N_t = N[(t + 1)\Delta] - N[t\Delta]\). Streett (2000) discussed the BIN(1,0) and BIN(1,1) models considering the first autoregressive heteroskedastic (ARCH(1)=GARCH(1,0)) process and the generalized heteroskedastic process (GARCH(1,1)) introduced by Bollerslev (1986) where data come from a normal distribution and variance conditional the past changes over time. Numerous applications have been also developed for several economic phenomena. An attempt for modeling count data with over or underdispersion Heinen (2003) studied two models by replacing the Poisson distribution with double Poisson distribution proposed by Efron (1986) and facing with the problem of normalizing constant’s intractability and moments. At the first one, double autoregressive conditional Poisson model(DACP\(1,1)\)) variance is proportional to the mean and at the second one DACP\(2,1)\)) variance expressed as a quadratic function. Furthermore Heinen (2003) considered only the expression for the number of transactions rename the model as autoregressive conditional Poisson(ACP) and studied some theoretical properties. Ferland et al. (2006) discriminated that at the ACP model when \(a_1, b_1 > 0\) then the model could be expressed as an autoregressive moving average model of order 1(ARMA(1,1)) where volatility is expressed by (1.4). Additionally if \(a_1 = b_1 = 0\) then the model of i.i.d. data has a mean \(\lambda_t = b_0\) and when \(a_1 = 0\) then the model reduced to an autoregressive model of order 1. A brief discussion for generalised double autoregressive conditional Poisson models has also been done from Heinen (2003). Zhu (2012a) examined properties of an INGARCH model assuming a Generalised Poisson distribution.

Definition 2. Let \(\{X_t\}\) be a time series of counts. Assuming that conditional on the past information the random variables \(X_i, \ text {for} i = 1, \ldots, n\) are independent and the conditional distribution of \(X_t\) is specified by a generalised Poisson (GP) distribution:

\[
X_t | \mathcal{F}_{t-1} \sim \mathcal{GP}(\lambda_t^*, k), \quad \frac{\lambda_t^*}{1 - k} = \lambda_t = d_1 + a_1 \lambda_{t-1} + b_1 X_{t-1}
\]

where \(b_0 > 0, a_1, b_1 \geq 0, \max(-1,-\lambda_t^*/4) < k < 1, \mathcal{F}_{t-1}\) is the \(\sigma\)-field generated by \(\{X_s : s < t - 1\}\). When \(a_1 = 0\) then the model is called GP-INARCH(1) model and when \(k = 0\) then the model reduced to the INGARCH(1,1) model defined by (1.3)

On the other hand to model underdispersion a binomial \(f(x = k) = \binom{n}{k} p^k (1-p)^{n-k}, n \in \mathbb{N}, p \in [0,1]\) and a negative binomial \(f(x = r) = \binom{n+r-1}{r} (1-p)^{n} p^r, n \in \mathbb{N}, p \in [0,1]\)
\( r, n \in \mathbb{N}, \ p \in [0,1] \) distributions have been proposed but differentiation with respect to parameter \( n \) is problematic and as a result difficulties with estimation have been provoked. Zhu (2011) constructed a univariate INGARCH Negative Binomial process considering that \( n \) is a known parameter. To estimate the parameter vector, maximum likelihood approach has been provided but for the choice of best initial values Yule-Walker and Conditional Least Squares methods have been also discussed. Chen et al. (2016) studied a Negative Binomial OD model with a time-varying conditional autoregressive mean where shape and scale parameters are considered unknown and Bayesian MCMC sampling used for parameters estimation. A Neyman type A OD model also discussed by Gonçalves et al. (2015) assuming a Compound Poisson law where its compounding distribution is also a Poisson law. Further study and analysis of an CP-INGARCH(1,1) time series process has been done from Silva (2016). Furthermore the same year Gonçalves et al. (2016) studied a zero-inflated integer-valued GARCH with general Compound Poisson deviates(ZICP-INGARCH(1,1)). A Zero-Inflated Compound Poisson Integer-valued GARCH model with orders 1 and 1, is calculated based on characteristic function of \( X_t \mid \mathcal{F}_{t-1}, \forall t \in \mathbb{Z} \):

\[
\Phi_{x_t | X_{t-1}}(u | \mathcal{F}_{t-1}) = \omega + (1 - \omega) \exp \left\{ \left( \frac{\lambda_t | \phi_t(u)}{\phi_t(0)} (\phi_t(u) - 1) \right) i \right\}, \ u \in \mathbb{R},
\]

\( \lambda_t \) defined by (1.4) for some constants \( 0 \leq \omega < 1, b_0 > 1 \) and where \( (\phi(t), t \in \mathbb{Z}) \) is a family of characteristic functions on \( \mathbb{R} \) and \( i \) denotes the imaginary unit.

The ZICP-INGARCH(1,1) model when \( a_1 = 0 \) is denoted as ZICP-INARCH(1) model. While \( \omega = 0 \) the above model reduced to a \( CP - INGARCH(1,1) \) model. Moreover the model has exceptional forms of zero-inflated Poisson(ZIP-INGARCH(1,1)) and negative binomial (NBINGARCH(1,1)) in case of \( 0 \leq a_1 < 1 \)

In case of numerous datasets with an excessive number of zeros, heteroskedastic models have been discussed by Zhu (2012c) based on the almost proposed zero-inflated Poisson distribution by Xie et al. (2001); Yang et al. (2011) and on weighted Poisson distribution proposed by Del Castillo and Pérez-Casany (1998); del Castillo and Pérez-Casany (2005) respectively. Based on the construction of volatility, parameter values are too restricted. For the above reason, Fokianos and Tjøstheim (2011) studied a new model where considering volatility’s construction negative and positive parameters values are allowed. To overcome this problem another model based on theory of GLM has been proposed by Fokianos and Tjøstheim (2011)
assuming

\[ X_t \mid \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \]  \hspace{1cm} (1.6)

where

\[ v_t = d_1 + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j \log(X_{t-j} + 1), \quad v_t = \log(\lambda_t) \]  \hspace{1cm} (1.7)

In this model, parameters \(d_1, a_i, b_j\) for any \(i = 1, \ldots, p, j = 1, \ldots, q\) take values in \(\mathbb{R}\) and both negative and positive correlations can occur. Certainly other specifications are possible. In particular, one may consider a model for the log-mean process by introducing \(\log(X_{t-j} + u)\), where \(u\) is a constant. Fokianos and Tjøstheim (2011) presented results of the data analysis that do not indicate any gross deviations in terms of the mean square error of residuals for values of \(u\) varying from 1 to 10 with a step equal to 0.5. Bentarzi and Bentarzi (2017) studied a periodic INGARCH(1,1) model with the requirement of establishment of first and second order conditional moments. An alternative construction proposed by Xu et al. (2012) where not only conditional mean expressed similar to this of initial INGARCH model given by (1.3) but also conditional variance is assumed to have a constant ratio to the conditional mean. The new dispersed integer-valued autoregressive heteroskedastic (DINARCH) model is capable to study data with overdispersion or underdispersion and this flexibility can be reflected by the value of the defined ratio. General DINARCH(p) model, where the process \(X_t \mid \mathcal{F}_{t-1} \sim P(\lambda_t)\), \(\lambda_t = b_0 + \sum_{i=1}^{p} a_i X_{t-i}\), \(E[X_t \mid \mathcal{F}_{t-1}] = \lambda_t\), \(\text{Var}[X_t \mid \mathcal{F}_{t-1}] = \alpha \lambda_t\) where \(\lambda_t\) satisfies (1.3) and \(\alpha\) is assumed constant. In case of \(a < 1\) model is appropriate for data with overdispersion but datasets with underdispersion are used at the proposed DINGARCH(p,q) model. A negative binomial(NB), a double Poisson (DP) and a generalized Poisson(GP) DINARCH processes have been proposed based on Kotz et al. (2004);Efron (1986) and Gelfand and Dalal (1990);Consul and Jain (1973) respectively. Additionally considering exceptional cases such as this of a zero-inflated Poisson INGARCH or a weighted Poisson in case of numerous zeros have been also studied by Xie et al. (2001);Yang et al. (2011) and Zhu (2012a) respectively. The NB-DINARCH model is defined assuming a NB distribution \(\text{NB}(r_t, \rho_t)\) with a pmf:

\[ f(x_t \mid \mathcal{F}_{t-1}) = \frac{\Gamma(r_t + x_t)}{\Gamma(x_t + 1) \Gamma(r_t)} \rho_t^{x_t} (1 - \rho_t)^{r_t} \]

where \(\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt\). First and second order moments defined as \(E_{\text{NB}}(X_t \mid \mathcal{F}_{t-1}) = r_t \rho_t/(1 - \rho_t)\) and \(\text{Var}_{\text{NB}}(x_t \mid \mathcal{F}_{t-1}) = r_t \rho_t/(1 - \rho_t)^2\) but considering the general equations from DINGARCH model \(r_t = \frac{\lambda_t}{1 - \alpha}\) and \(\rho_t = 1 - \frac{1}{\alpha}\) obtaining that conditional variance change over time. The DP-DINARCH model has been also
studied by Heinen (2003) and was called DACP. At the GP-DINARCH model where $GP(\lambda^*, k)$ referred previously after matching conditional moments with those at the general case $\lambda_t^* = \frac{\lambda_t}{\sqrt{\alpha}}$ and $k = 1 - \frac{1}{\alpha}$. An initiative idea for a mix construction of a beta-negative binomial integer-valued heteroskedastic (BNB-INGARCH) model that reduces to a Poisson -INGARCH or a NB-INGARCH model considering some restrictions and based on the study of Wang (2011) discussed by Gorgi (2020). A Beta Negative Binomial ($BNB(\lambda, r, \alpha)$) distribution is defined as:

$$f(x) = \frac{\Gamma(x + r)}{\Gamma(x + 1)\Gamma(r)} B\{a + r, (a - 1)\lambda/r + x\} B\{a, (a - 1)\lambda/r\}$$

with dispersion parameter $r > 0$, mean $\lambda > 0$ and tail parameter $a > 1$. When $a \to \infty$ the BNB converges to a NB distribution with success probability $\frac{\lambda}{r + \lambda}$ and in case that $r \to \infty$ the BNB converges to a Poisson distribution with mean $\lambda$. A hysteretic Poisson INGARCH process with advantages of modeling data with overdispersion and asymmetry studied by Truong et al. (2017). At the hysteretic Poisson INGARCH model a hysteretic variables responsible for dynamic switching modeling and and it can be an exogenous variable. Chen and Lee (2016) introduced generalized threshold Poisson INGARCH model with the main difference of relaxing the sharp transition in the parameter changes. An integer-valued process with positive and negative observations has been studied by Alomani et al. (2018) using a symmetric Skellam distribution and trying to model data with overdispersion with time-varying variance. An integer-valued asymmetric generalized autoregressive conditional heteroskedastic (INAGARCH) model was proposed by Hu and Andrews (2021) and defined as:

$$X_t = Z_t Y_t, \quad Y_t | F_{t-1} \sim Poisson(\lambda_t)$$

$$\lambda_t = \left(\sqrt{1 + 4\eta_t} - 1\right)/2$$

$$\eta_t = a_{00} + \sum_{i=1}^{p} a_{0i} (|X_{t-i}| - \gamma_0 X_{t-i})^2 + \sum_{j=1}^{q} \beta_{0j} \eta_{t-j},$$

(1.8)

where $F_{t-1}$ is the $\sigma$-field, $a_{00} > 0, a_{1i} > 0, b_{1j} > 0, i = 1, \ldots, p$ and $j = 1, \ldots, q, \ p \geq 1, \ q \geq 0$. Furthermore, following Straumann (2005) parameter values $a_{00}, a_{1i}, b_{1j}$ are all unique and the polynomials $a_0(x) = \sum_{i=1}^{p} a_{0i} x^i$ and $\beta_0(x) = \sum_{j=1}^{q} \beta_{0j} x^j$ have no common roots. The idea for the above construction based on the asymmetric generalized autoregressive conditional heteroskedastic (AGARCH) model where variables are continuously distributed in $\mathbb{R}$ studied by Ding et al. (1993). This model is appropriate to model data with positive and negative values while those observations
can have asymmetric impact on conditional variance. The pmf of \( X_t \) is:

\[
P(X_t = k \mid \mathcal{F}_{t-1}) = \begin{cases} 
\frac{1}{2} e^{-\lambda t} \frac{\lambda^{|k|}}{|k|!}, & k \neq 0 \\
e^{-\lambda t}, & k = 0
\end{cases}
\]

which is symmetric about zero and exhibits exponential decay when \( k \to \pm \infty \).

An heteroskedastic time series process with no assumptions on the relationship between conditional mean and variance included in the family of semiparametric models introduced by Liu and Yuan (2013) and called random rounded integer-valued autoregressive conditional heteroskedastic (RRINARCH) model. Other advantage except separateness between mean and variance is that it allows negative values at the series and at the autocorrelation function. Additionally autocorrelation of the RRINARCH(p,q) model is similar to this of an AR(p) model. The construction of the above model based on two random rounding operators where the first one is a special case of probabilistic rounding introduced by M’Raïhi et al. (2000) defined as:

\[
\odot_1(x, U) = \Delta(x) + \mathbb{I}(U \geq 1 + \Delta(x) - x), \quad x \in \mathbb{R}
\]

\[
\odot_2(x, U) = \Delta(x^{1/2}) + \mathbb{I}(U \geq B(x)), \quad x \in \mathbb{R}_+
\]

where

1. \( \mathbb{I}(A) \) is an indicator function of \( A \)

2. \( \Delta(x) = \max\{z \in \mathbb{Z} : z \leq x\} \)

3. \( B(x) = \frac{(\Delta(x^{1/2} + 1)^2 - x)}{(\Delta(x^{1/2} + 1)^2 - (\Delta(x^{1/2}))^2)} \)

4. \( U \) is a uniform random variable defined on the interval [0,1].

Considering that

\[
\mathcal{F}(c) = \{X : E(X) = c, \text{Var}(X) < \infty\}
\]

and

\[
R(c) = (\Delta(c) + 1 - c)(\Delta(c) - c).
\]

Then the random rounding operator lies on the above collection \( \mathcal{F}(c) \) and \( \text{Var}(\odot_1(c, U)) = R(c) \) and RRINARCH(p,q) model has mean given by:

\[
E(X_t \mid \mathcal{F}_{t-1}) = c + \sum_{j=1}^{p} \phi(j)X_{t-j}
\]
and variance defined as:

\[ \text{Var}(X_t \mid \mathcal{F}_{t-1}) = R(E(X_t \mid \mathcal{F}_{t-1})) + h_t \]

where \( h_t \) is a non-negative \( \mathcal{F}_{t-1} \)-measurable function. A generalised regime-switching integer-value (RGS-INGARCH(1,1)) GARCH(1,1) model defined by Lee and Hwang (2018) in order to model processes with time-varying dependent transition probabilities. A threshold integer-valued time series given by:

\[ X_t \mid \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \]

where \( \lambda_t \) given by (1.3) and

\[
X_{t-1}^{(r)} = \begin{cases} 
X_{t-1} & \text{if } X_{t-1} > m_t \\
0 & \text{if } X_{t-1} \leq m_t
\end{cases}
\]

\[
X_{t-1}^{(l)} = \begin{cases} 
X_{t-1} & \text{if } X_{t-1} \leq m_t \\
0 & \text{if } X_{t-1} > m_t
\end{cases}
\]

Parameter \( m_t \) is a time-varying threshold variable that determines the dynamic switching mechanism of the model. Another threshold model has been proposed by Liu (2012). The main difference with the previous model is that the threshold parameter at this model lies on the set of natural numbers without any dependence on time. Furthermore, an attempt for the construction of a mixture integer-valued autoregressive conditional heteroskedastic (MINARCH) model was given by Zhu et al. (2010) based on the idea of the general mixture model of Saikkonen (2007). The usefulness of the above model obtained at the modeling of data from bimodal marginal distributions as for example Canadian Lynx data studied from Tong (1990). A novel structure of an integer-valued heteroskedastic model to overcome the problem of equidispersion proposed by Barreto-Souza et al. (2022). Time-varying dispersion parameter was considered at this new structure in order to control conditional variance and to provide alternative and more flexible structure of volatility. The time varying dispersion integer-valued generalized autoregressive conditional heteroskedastic (tv-DINGARCH) model studied in case of Negative binomial distribution while a log-linear tv-INGARCH model with covariates has been also discussed.

**Definition 3.** A \( tv-DINGARCH(p_1,p_2,q_1,q_2) \) process \( \{X_t\}_{t \in \mathbb{N}} \) is defined by \( X_t \mid \mathcal{F}_{t-1} \sim MP(\lambda_t, \phi_t) \) with

\[
\lambda_t = f(X_{t-1}, \ldots, X_{t-p_1}, \lambda_{t-1}, \ldots, \lambda_{t-q_1}), \quad \phi_t = f(X_{t-1}, \ldots, X_{t-p_2}, \phi_{t-1}, \ldots, \phi_{t-q_2})
\]

where \( \mathcal{F}_{t-1} \) and \( \lambda_0, \phi_0 \) denoting some starting value, \( f: \mathbb{N}^{p_1} \times (0, \infty)^{q_1} \to (0, \infty) \) and \( g: \mathbb{N}^{p_2} \times (0, \infty)^{q_2} \to (0, \infty) \)
A variable follows a mixed Poisson (MP) distribution if $X \mid Z = z \sim \text{Poisson}(\lambda z)$ and $Z$ follows a non-negative distribution with $E(z) = 1$ and $\text{Var}(z) = \phi$, for $\lambda, \phi > 0$. In case of $X \sim MP(\lambda_t, \phi_t)$ then $E(X_t) = \lambda_t$ and $\text{Var}(X_t) = \lambda_t + \phi_t \lambda_t^2$.

The random variable $Z$ follows a Gamma distribution $G(a, b)$. Fokianos and Tjøstheim (2011) studied two models: the first one of a Negative binomial distribution and the second one of a Poisson inverse Gaussian (PIG) distribution. At the first case when $Z \sim G(\phi^{-1}, \phi^{-1})$ with $E(z) = 1$ and $\text{Var}(Z) = \phi$ then $X \sim \text{NB}(\lambda, \phi)$ given by:

$$f(x \mid F_{t-1}) = \frac{\Gamma(x + \phi^{-1}t)}{x!\Gamma(\phi^{-1}t)} \left( \frac{1}{\lambda_t \phi_t + 1} \right)^x \left( \frac{1}{\lambda_t \phi_t + 1} \right)^{1/\phi_t}, \quad x \in \mathbb{N}_0$$

At the second model when $Z \sim IG(1, \phi)$ then $X \sim PIG(\lambda, \phi)$. An exceptional case of a log-linear tv-DINGARCH model also proposed without study of theoretical properties. While the construction based on this proposed before many years from Fokianos and Tjøstheim (2011), the main difference is obtained on the consideration of covariates at the expression of volatilities $\log(\lambda_t)$ and $\log(\phi_t)$. This necessity appeared after the idea of modeling Covid-19 data about deaths in Ireland. A variable with the name $\log - hosp$ which gives the number of hospitalized patients due to the disease, was included as covariate in order to model and study the effect to the mean and dispersion. The structure of the NB model is similar to this given previously in linear case but $\mu_t = \log(\lambda_t)$ and $v_t = \log(\phi_t)$ are defined as

$$\left\{ \begin{array}{l}
\mu_t = b_0 + b_1 \log(Y_{t-1} + 1) + b_2 \mu_{t-1} + \delta_1 (\log - hosp_t) \\
v_t = a_0 + a_1 \log(Y_{t-1} + 1) + a_2 v_{t-1} + \gamma_1 (\log - hosp_t)
\end{array} \right.$$  

An INGARCH(1,1) model with Markovian covariates constructed by Liu (2012). Assuming a time homogeneous Markov chain $Z_t$ given as a $p$-dimensional vector $Z_t = (Z_{1,t}, \ldots, Z_{p,t})$ that contains the covariates information at time $t$, the model defined as

$$X_t \mid F_{t-1}^{X,Z}, Z_t \sim \text{Poisson}(e^{Z_t^T \gamma \mu_t}),$$

where

$$\mu_t = \delta + a \mu_{t-1} + \beta X_{t-1} e^{-Z_{t-1}^T \gamma},$$

$F_{t-1}^{X,Z} = \sigma\{\mu_1, Y_s, Z_s, s \leq t\}$, $L(Z_t \mid Z_{t-1})$ represents a distribution of the random variable $Z_t$. By this way a Negative Binomial INGARCH model with Markovian covariates has been also discussed. Clark and Dixon (2021) studied a spatially correlated self-excited INGARCH model assuming data model dependence and capturing spatial variation. A new model based on Poisson-Gamma conjugacy with a
dynamic specification proposed by Gouriéroux and Lu (2019). The main feature of
the above model is the flexibility offered by the construction providing homoskedas-
tic or heteroskedastic models. At this study a Negative Binomial INAR(1) model
and a Negative Binomial INARCH(1,1) model have been discussed. An overview
about integer-valued generalized conditional heteroskedastic time series models given
by Fokianos (2012) and recently a methodological review proposed by Davis et al.
(2021).

1.4 Stationarity and Ergodicity of OD models

1.4.1 Ergodicity

At observation-driven process ergodicity can be examined by using Tweedie criteria
under the assumption that the process is irreducible Markov chain. First idea, pro-
posed by Rosenblatt (1956), is based on mixing condition satisfied by a sequence of
random variables. Even though Tweedie (1976) has proposed irreducibility criteria,
assumptions are not always satisfied. Meyn and Tweedie (1992);Meyn and Tweedie
(1993);Meyn and Tweedie (1994) studied conditions for stability of discrete-time
stochastic processes, of continuous-time stochastic processes respectively. At dis-
crete time processes criteria of Harris recurrence, positive Harris recurrence and
geometric recurrence and ergodicity were discussed. The chain is probabilistically
stable if \(\Phi\) returns to sets of positive measure (‘Harris recurrence’) or if there is a
unique invariant probability measure for \(\Phi\) (‘positive Harris recurrence’). Consid-
ering the Markov chain \(\Phi = \{\Phi_0, \Phi_1, \ldots, \}\) evolving on a locally compact separable
metric space \(X\), whose Borel \(\sigma\)-algebra shall be denoted \(\mathcal{B}\). With \(P_\mu\) and \(E_\mu\) denoted
the probability and expectation conditional on \(\Phi_0\) having distribution \(\mu\) and \(P_x\) and
\(E_x\) when \(\mu\) is concentrated at \(x\). Criteria of \(\phi\)-irreducibility are the most commonly
used in cases of heteroskedastic time series models. If \(\Phi\) is positive Harris recurrent
and \(t\) with invariant probability measure \(\pi\) and period \(m \geq 1\), and if \(\pi(f) < \infty\),
then for any initial condition \(x\) satisfying

\[
E_X \left[ \sum_{k=1}^{\tau_A} f(\Phi_k) \right] < \infty
\]

the distributions converge in \(f\)-norm:

\[
\left\| \frac{1}{m} \sum_{i=1}^{m} P^{i+l}(x, \circ) - \pi \right\|_f \rightarrow 0
\]
where Frobenius norm defined as $||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{q} |a_{ij}|^2}$ and $A$ is an $n \times m$ matrix. Furthermore proved that if a Markov chain is aperiodic and geometrically recurrent then it is geometrically ergodic. The main problems are when Markov chains are not irreducible or in general do not satisfy any mixing condition discussed by Doukhan and Doukhan (1994). Fonseca and Tweedie (2002) proposed stability condition and ergodic theorems for non-irreducible chains. Lu (1996) studied geometric ergodicity of an ARCH process while Rahbek (2003) studied stochastic properties and stationary solutions of multivariate ARCH processes. Francq and Zakoïan (2006) proposed mixing properties of a general class of GARCH (1, 1) models and before this paper Ling and McAleer (2002) discussed properties for stationarity and existence of moments for a class of GARCH models. Bibi and Aknouche (2008) examined stationarity and geometric ergodicity of a periodic GARCH processes. More generally Boussama et al. (2011) studied criteria for stationarity and geometric ergodicity of Baba, Engle, Kraft, and Kroner (BEKK) multivariate GARCH models by expressing GARCH processes as semi-polynomial Markov processes. Lescheb (2015) proposed necessary and sufficient conditions ensuring the existence of stationary solutions basing on the top Lyapunov exponent considering a periodic Asymmetric Generalized Autoregressive Conditional Heteroskedastic model (PAGARCH). Furthermore limit theorems and moment inequalities are difficult to obtain in many time series processes. Taking into account the above obstacle criteria of weak stationarity have been proposed and study. Meyn and Tweedie (1994) studied under which criteria Markov chains converge. Doukhan et al. (2012) introduced the more adapted weak dependence condition. The idea based on the asymptotic independence between "past" and "future". Considering a time series $X = (X_n)_{n \in \mathbb{N}}$ with values in a local topological space $E (E = \mathbb{R}^d)$. Assuming a variable $(X_{i_1}, X_{i_2}, \ldots, X_{i_u})$ for the past and one $(X_{j_1}, X_{j_2}, \ldots, X_{j_v})$ for the future where $i_1 \leq i_2 \leq \cdots \leq i_u \leq j_1 \leq j_2 \leq \cdots \leq j_v$, $u, v \in \mathbb{N}^*$, independence of time series is calculated via independence between $X_i$ and $X_j$ given by:

$$
\epsilon(r) = \sup_{D(X_i, X_j) \geq r} \sup_{(f, g) \in \mathcal{F} \times \mathcal{G}} \left\| \text{Cov}(f(x_i), g(x_j)) \right\|
$$

where $D(X_i, X_j) = j_1 - i_u$. Many dependence criteria have been constructed considering mixing condition introduced by Rosenblatt (1956) or on martingales approximations (mixingales).

**Definition 4.** Let $p \geq 1$ and let $(\mathcal{F}_n)_{n \in \mathbb{Z}}$ be an increasing sequence of $\sigma$-algebras. The sequence $(X_n, \mathcal{F}_n)_{n \in \mathbb{Z}}$ called an $L^p$ mixingale if there exist non-negative sequences $(\epsilon_n)_{n \in \mathbb{Z}}$ and $(\phi(n))_{n \in \mathbb{Z}}$ such that $\psi(n) \rightarrow 0$ as $n \rightarrow \infty$ and for all integers
\[ n \in \mathbb{Z}, \ k \geq 0 \]
\[
\left\| X_n - E(X_n \mid \mathcal{F}_{n+k}) \right\|_p \leq c_n \psi(k+1) \\
\left\| E(X_n \mid \mathcal{F}_{n-k}) \right\| \leq c_n \psi(k)
\]

However many classes of time series do not satisfy any mixing condition or limit theorems and moment inequalities are difficult to obtain. Bickel and Bühlmann (1999) and simultaneously Doukhan and Louhichi (1999) introduced a new idea of weak dependence. Let \( \Lambda(\delta) \) be the set of Lipschitz functions from \( \mathcal{X} \) to \( \mathbb{R} \) with respect to the distance \( \delta \). For \( f \in \Lambda(\delta) \), we denote this by \( \text{Lip}(f) \)

\[
\Lambda^{(1)}(\delta) = \left\{ f \in \Lambda(\delta) \mid \text{Lip}(f) \leq 1 \right\}.
\]

Furthermore considering a probability space \( (\Omega, \mathcal{A}, P) \) and a Polish space \( \mathcal{X} \), two classes of functions \( \mathcal{F}_u, \mathcal{G}_u : \mathcal{X}_u \rightarrow \mathbb{R} \) have been also defined and \( \mathcal{F} = \bigcup \mathcal{F}_u \) and \( \mathcal{G} = \bigcup \mathcal{G}_u \). Assuming two random variables \( X \) and \( Y \) with values in \( \mathcal{X}_u \) and \( \mathcal{X}_v \) respectively and \( \Psi \) some function from \( \mathcal{F} \times \mathcal{G} \) to \( \mathbb{R}_+ \), the \((\mathcal{F}, \mathcal{G}, \mathbb{R})\)-dependence coefficient \( \epsilon(X, Y) \) is given by:

\[
\epsilon(X, Y) = \sup_{f \in \mathcal{F}_u, g \in \mathcal{G}_u} \frac{\text{Cov}(f(X), g(Y))}{\Psi(f, g)}.
\]

If \( (X_n)_{n \in \mathbb{Z}} \) is a sequence of \( \mathcal{X} \)-valued random variables and \( \Gamma(u, v, k) \) a set of \((i, j) \in \mathbb{Z}^u \times \mathbb{Z}^v \) such that \( i_1 \leq \cdots \leq i_u \leq i_u + k \leq j_1 \leq \cdots \leq j_v \). The dependence coefficient \( \epsilon(k) \) is defined by:

\[
\epsilon(k) = \sup_{u,v} \sup_{(i,j) \in \Gamma(u,v,k)} \left( \epsilon \left( X_{i_1}, \ldots, X_{i_u} \right), \left( X_{j_1}, \ldots, X_{j_v} \right) \right).
\]

The sequence \( (X_n)_{n \in \mathbb{N}} \) is \((\mathcal{F}, \mathcal{G}, \Psi)\)-dependent if the sequence \( \epsilon(k) \) tends to zero. Doukhan et al. (2012) proposed numerous criteria of weak dependence like \( \eta, k, \lambda, \tilde{\alpha}, \tilde{\beta}, \phi, \theta \) and \( \tau \) where a part of those constructed based on Rosenblatt (1956)'s criteria. The coefficient \( \theta \) corresponds to

\[
\Psi(f, g) = d_g \|f\|_\infty \text{Lip}(g),
\]

where \( f \) belongs to the set of bounded functions \( \mathcal{F}_u \), \( \|f\|_\infty = \max f_i \). Moreover \( (X_n)_{n \in \mathbb{Z}} \) be a sequence of \( R^k \)-valued random variables. We define the sequence \( (Y_n)_{n \in \mathbb{Z}} \) by \( Y_n = h(X_n) \). If \( X \) is \( \theta \)-weak dependent then \( (Y_n)_{n \in \mathbb{Z}} \) is also \( \theta \) weak dependent and

\[
\theta_Y(n) = O(\theta(n)^{\frac{p-a}{p-1}}).
\]
1.4. STATIONARITY AND ERGODICITY OF OD MODELS

**Definition 5.** Let $(\Omega, A, \mathbb{P})$ be a probability space and $\mathcal{F}$ be a $\sigma$-algebra on $A$. Let $X$ be a random variable on $(\Omega, A, \mathbb{P})$ with values on $\mathbb{R}^m$ and $||X||_1 < \infty$. Let $\mathcal{X}$ be a polish space and $\delta$ be a metric on $\mathcal{X}$. For any $L^p$-integrable function $X$ on $\mathcal{X}$:

$$
\theta_p(\mathcal{F}, X) = \sup\{||E(f(X) \mid \mathcal{F}) - E(f(X))||_p / f \in \Lambda^{(1)}(\mathbb{R}^m)\}
$$

and

$$
\tau_p(\mathcal{F}, X) = \{||\sup\{E(f(X) \mid \mathcal{F}) - E(f(X))\} / f \in \Lambda^{(1)}(\mathbb{R}^m)||\}. 
$$

The main difference between Tweedie and weak dependence criteria is that for the satisfaction of the second only the first order moments of the process must be infinite. Simpler inequalities for $\beta$, $\alpha$ and $\tau$ weak dependence coefficients have been also provided. Considering the probability space $(\Omega, A, \mathbb{P})$, where $\Omega$ is rich enough and $X$ be another variable distributed as $Y$ and independent of $Y$ we can bound the coefficient $\tau$ as: $\tau(\mathcal{F}, Y) = ||Y - X||$. The above implies that there exists a variable $Y^* \text{ measurable with respect } \max\{\mathcal{F}, \sigma(Y), \sigma(\delta)\}$ independent of $\mathcal{F}$ distributed as $Y$

$$
||Y - Y^*||_1 = \tau_1(\mathcal{F}, Y).
$$

It is easy to see that

$$
\theta_p(\mathcal{F}, Y) \leq \tau_p(\mathcal{F}, Y).
$$

The drawback is that weak dependence criteria provide stationarity of only first order moments. The usefulness of each one of those depend on the closed form of stochastic process. Fokianos et al. (2009) studied ergodicity of univariate INGARCH process considering a perturbed model. {Doukhan and Louhichi (1999), Dedecker et al. (2007)} overcome those difficulties by proposing weak dependence criteria for different types of processes. In case where recursive equation is expressed linearly weak dependence criteria can be used more easier than other cases. Furthermore the problem of ergodicity become more intensive when there are interdependencies between two or more stochastic processes. For an extended study of weak dependence, someone could see the work of Dedecker et al. (2007).

To sum up there are numerous criteria to provide ergodicity. The study of ergodicity is very important because via ergodic criteria, stationarity also provided. Even though there are too many weak dependence criteria at some cases, criteria proposed by Tweedie (1976) could be also used. There are studies as Fokianos et al. (2009) where stationarity proved using weak dependence criteria but a few years later at the same process stationarity has been also provided using Tweedie criteria by Tjøstheim (2012). The reason is that in some way Tweedie criteria are more difficult and complicated to use based on more theoretical aspect.
CHAPTER 1. BACKGROUND

1.4.2 Consistency issue of Observation Driven (OD) Models

In the thesis we consider multivariate discrete and continuous OD models. Considering
\[ X_t | \mathcal{F}_{t-1} \sim G^0(\lambda_t), \]
where \((\lambda, d_\lambda)\) be a locally compact and separable metric space and \(\lambda_t = \psi_{X_t}(\lambda_t)\) is a measurable function. Examination for the consistency of MLE based on theorem developed by Lee et al. (2016), providing stationarity and ergodicity of the process \(\{X_t, t \geq 0\}\) and considering that second order partial derivatives are continuous and bounded. In the thesis ergodicity provided considering criteria proposed by Dedecker et al. (2007). We study numerical simulations considering a trivariate INGARCH(1,1) and a bivariate ECARR(1,1) model. Fokianos et al. (2009) provided ergodic properties assuming classical approaches as criteria proposed by Tweedie (1976) and \(\beta\)-mixing properties. The obstacle is the irreducibility of the Markov chain and for this reason Fokianos et al. (2009) considered a perturbed model. On the contrary in the continuous case of the ECARR model Fernandes et al. (2005) considering a trivariate model provided ergodicity via Tweedie criteria of the model without any disturbance. The drawback with the criteria of weak dependence is that only the first two conditional moments the process are stationary.

1.5 Multivariate heteroskedastic OD time series models

1.5.1 Bivariate processes for continuous data

In real world the construction of models taking into account properties of data expressed by two or more variables and examined for the correlation was proved necessary. The first bivariate autoregressive process is given by:
\[ X_t = A_1 X_{t-1} + \epsilon_t, \]
where \(A_1\) is a \(2 \times 2\) matrix and the error process is a 2-dimensional zero mean white noise process with covariance matrix \(E(\epsilon_t\epsilon'_t) = \Sigma_e\) such that \(\epsilon_t \sim N(0, \Sigma_e)\). In similar way the expansion of the model at more than two dimensions has been also studied without any difficulty. The two-dimensional Vector Autoregressive (VAR(p)) model could be written as a 2p-dimensional VAR(1) model by stacking p consecutive \(X_t\) variables in a 2p-dimensional vector as:
\[ X_t = AX_{t-1} + \epsilon_t \]
where $A$ referred as the companion matrix and $A = \begin{bmatrix} A_1 & A_2 & \ldots & A_{p-1} & A_p \\ I_2 & 0 & \ldots & 0 & 0 \\ 0 & I_2 & \ldots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \ldots & 0 & I_2 \end{bmatrix}$ and

$e_t = \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ Similarly to the univariate case the stable VAR process can be expressed as the weighted sum of past and present innovations. Considering the Wold decomposition theorem and under suitable conditions every $K$-dimensional nondeterministic zero mean stationary process $y_t$ has an MA representation. A vector autoregressive moving average (VARMA) defined as:

$$X_t = A_1 x_{t-1} + \epsilon_t + B_1 \epsilon_t$$

where $A_1$ and $B_1$ are $2 \times 2$ matrices, $x_t$ is a 2-dimensional vector of observations and $\epsilon_t$ is a 2-dimensional vector of white noise error terms. At this model if the determinant of the MA operator of the VARMA process has all its roots outside the unit circle, the process has an equivalent pure VAR representation of possibly infinite order. A detailed introductory exposition of VARMA processes is provided by Lütkepohl (2005) and a more advanced treatment can be found in Deistler and Hannan (1988), Hsiao (1979); Hsiao (1981). Keating (2000) introduces the closely related idea of asymmetric VAR (AVAR) models, in which the lag length differs across variables such that the lag length is the same for each variable in all equations, but may differ across variables. The use of exogenous variables to a VAR model which include the same number of lags as the other equations but only lags of the exogenous variables have been also studied. The appearance of exogenous variables was not commonly used in VAR models but after few years financial market showed that those models were necessary. A better example is a small open economy that faces exogenous variation in world interest rates or in its terms of trade (see, e.g., Cushman and Zha (1997)). To represent the dynamics of the conditional variances and covariances, the construction of multivariate generalized autoregressive conditional heteroskedastic models (MGARCH) became useful. The drawback was that as the dimension of the model becomes high then the number of parameters increases and computational methods for parameters estimation are needed. On the other hand, the above fact is not a drawback because models with only a few parameters may not be able to capture the relevant dynamics in the covariance structure. An obligatory assumption is that the conditional covariance matrices implied by the model
must be positive definite, but this is often infeasible in practice except the formula construction in way that positive definiteness is implied by the model structure.

The first GARCH model for the conditional covariance matrices was the so-called VEC model of Bollerslev et al. (1988) and an ARCH model of Engle et al. (1984). The lack at properties for consideration of positive definiteness in the construction of the model provide at the construction of more parsimonious models. Bollerslev et al. (1988) presented a simplified version assuming that the matrices are diagonal. A model that can be viewed as a restricted version of the VEC model is the Baba-EngleKraft-Kroner (BEKK) defined in Engle and Kroner (1995). It has the attractive property that the conditional covariance matrices are positive definite by construction. Engle and Kroner (1995) provide sufficient conditions for the two models, BEKK and VEC, to be equivalent. They also give a representation theorem that establishes the equivalence of diagonal VEC models and general diagonal BEKK models. A generalization of the univariate exponential GARCH model of Nelson (1991),where parameter restrictions to ensure positive definiteness are not needed proposed by Kawakatsu (2006). Engle et al. (1990) introduced the first factor GARCH model, assuming that the observations are generated by underlying factors that are conditionally heteroskedastic and possessing by this way a GARCH-type structure. Alexander and Chibumba (1997) proposed an Orthogonal GARCH (O-GARCH) model and Van der Weide (2002) extended the above model to a Generalized Orthogonal GARCH model (GO-GARCH).

Diebold and Nerlove (1989) propose a model similar to the previous one and Sentana (1998) made a comparison between the above models. Full Factor (FF–) GARCH model and Generalized Orthogonal Factor GARCH (GOF-GARCH) model have been studied by Vrontos et al. (2003) and Lanne and Saikkonen (2007) respectively. Correlation models are based on the decomposition of the conditional covariance matrix into conditional standard deviations and correlations. The simplest multivariate correlation model that is nested in the other conditional correlation models, is the Constant Conditional Correlation (CCC–) GARCH model of Bollerslev (1990). At varying-correlation GARCH(VC-GARCH) model proposed by Tse and Tsui (2002) conditional correlation is function of previous periods and a set of estimated correlations. Other multivariate GARCH models with transformations at Conditional correlation have been also studied by Kwan et al. (2005), Engle (2002), Billio et al. (2006), Cappiello et al. (2006), Silvennoinen and Teräsvirta (2005), Silvennoinen and Teräsvirta (2009a), Berben et al. (2005), Pelletier (2006), Silvennoinen and Teräsvirta (2009b), Silvennoinen and Teräsvirta (2009b), Fernandes
1.5. MULTIVARIATE HETEROSKEDASTIC OD TIME SERIES MODELS

et al. (2005).

1.5.2 Bivariate processes for discrete data

Definition for a joint conditional distribution is required due to necessity of modeling data in more than one dimension. In discrete case, the construction of a multivariate distribution is more complex than the continuous based on the study of theoretical properties like ergodicity and consistency of estimators. Furthermore while studies in case of homoskedastic models, at heteroskedastic models literature is restricted. For example while there is a multivariate negative binomial distribution studied by Andreassen (2013), there are no studies for multivariate Negative Binomial Integer Generalized Autoregressive Conditional Heteroskedastic (NB-INGARCH). A way in order to construct a multivariate model is the method of trivariate reduction introduced by Mardia (1970). Based on the above model a bivariate Poisson distribution had been also proposed by Liu (2012). Considering three mutually independent Poisson distributed random variables $X_{1t} \sim \text{Poisson}(\lambda_{1t} - \phi)$, $X_{2t} \sim \text{Poisson}(\lambda_{2t} - \phi)$ and $X_{3,t} \sim \text{Poisson}(\phi)$ then $Y_{1t} = X_{1t} + X_{3,t}$ and $Y_{2t} = X_{2t} + X_{3,t}$ have a joint bivariate Poisson distribution given by:

$$f(Y_{1t}, Y_{2t} \mid \mathcal{F}_{t-1}) = \min(m,l) \sum_{i=0}^{\min(m,l)} P(X_1 = m - i)P(X_2 = l - i)P(X_3 = i)$$

$$= e^{-(\lambda_{1t} + \lambda_{2t} - \phi)} \frac{(\lambda_{1t} - \phi)^m (\lambda_{2t} - \phi)^l}{m! l!} \sum_{i=0}^{\min(m,l)} \left( \frac{\phi}{(\lambda_{1t} - \phi)(\lambda_{2t} - \phi)} \right)^i$$

Andreassen (2013) presented three simulation studies based on three methods the first one of setting initial values at $\lambda_t$ and $X_1 = (X_{1,1}, X_{2,1})$ and then using general structure for $\lambda_t$ simulations from $BP(\lambda_{1,t}, \lambda_{2,t}, \phi)$ were provided. The second one based on generation of $X_{1,t} \sim \text{Poisson}(\lambda_{1,t} - \phi)$, $X_{2,t} \sim \text{Poisson}(\lambda_{2,t} - \phi)$ and $X_{3,t} \sim \text{Poisson}(\phi)$ and setting $Y = (X_{1,t} + X_{3,t}, X_{2,t} + X_{3,t})^T$ and the third one by using a closed form of a copula as usually an Archimedean copula or a Clayton copula. Based on method of trivariate reduction, the first bivariate heteroskedastic integer valued conditional heteroskedastic process constructed by . The effort was concentrated at the extension of the autoregressive conditional Poisson model of a vector autoregressive(AR)-type model for the conditional mean given as:

$$X_t \mid \mathcal{F}_{t-1}^X \sim BP(\lambda_t) \quad (1.13)$$
where
\[
\lambda_t = d + A\lambda_{t-1} + BX_{t-1}
\] (1.14)

\(\lambda_{it}, i = 1, \ldots, n\) is the number of variables and \(t \in \mathbb{N}\) is volatility lag. Each \(\lambda_{it}\) does not depend only on its own past and this indicates that \(A\) and \(B\) are squared matrices with non zero elements. Providing that \(||A||_1 + ||B||_1 < 1\) where \(||A||_1 = \max_{1 \leq j \leq p} \sum_{i=1}^q |a_{ij}|\), then there exists a unique solution \((X_t, \lambda_t)\) to model (4.1) which is stationary and ergodic. The elements of \(d, A, B\) are assumed to be positive considering positivity of \(\lambda_{i,t}\). The construction of the model has been done using the trivariate reduction method where the bivariate Poisson distribution given by:

\[
f(x_1, x_2) = f(x_1) f(x_2) \left[ 1 + \omega (g_1(x_1) g_2(x_2) - E(g_1(x_1)) E(g_2(x_2)) \right]
\]

Cui and Zhu (2018) considered as marginal \(f(x_i), i = 1, 2\) Poisson(\(\lambda_i\)) and \(q(x_i) = e^{-y_i} - e^{-c\lambda_i}\), where \(c = 1 - e^{-1}\)

\[
f(y_1, y_2) = \prod_{i=1}^{2} \frac{\lambda_i^{y_i}}{y_i!} e^{-\lambda_i} \left( 1 + \delta \prod_{i=1}^{2} (e^{-y_i} - e^{-c\lambda_i}) \right),
\]

where \(c = 1 - e^{-1}\).

The proposed model was not well defined as mentioned from other researchers who proposed alternative constructions of multivariate distributions because while the structure of correlation coefficient is based on the equivalent study of Lakshminarayana et al. (1999), the way on how boundaries values have been also calculated missing or it is not clearly defined. Another attempt taking into advantage three families of copulas the family of Frank, the family of Gaussian and the family of FGM copula had been discussed by Cui et al. (2020) where bivariate Poisson distribution defined by:

\[
f(x_1, x_2) = \frac{1}{Z(\lambda_1, \lambda_2, \omega)} \prod_{i=1}^{2} \frac{\lambda_i^{x_i}}{x_i!} e^{-\lambda_i} c(F_1(x_1), F_2(x_2)),
\]

where \(c(F_1(x_1), F_2(x_2))\) is one of the above mentioned copulas, \(\omega\) be the parameter of dependence and normalizing constant defined as :

\[
Z(\lambda_1, \lambda_2, \omega) = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \frac{\lambda_1^{x_1} \lambda_2^{x_2}}{x_1! x_2!} e^{-(\lambda_1 + \lambda_2)} c(F_1(x_1), F_2(x_2)).
\]
Based on Shmueli et al. (2005), series $Z$ converges and both $Z, Z^{-1}$ are bounded. Those motivations for the construction derived from the idea of Lakshminarayana et al. (1999) and copulas. At the above cases researchers deal with the problem of negative correlation. Although the difficulty is focused not only on the construction of a multivariate Poisson distribution that provides negative correlation but on the dependence structure which does not offer flexibility. Crucial attention is needed at parameters’ estimation where classical approaches as maximum likelihood approximation seem to be time-consuming depending each time on the copulas’ form. Cui and Zhu (2018) and Cui et al. (2020) decomposed the log-likelihood function based on maximization by parts and a modified version of this as alternative approaches with less complexity but with a bit high computational cost. Another approximation was presented by Fokianos et al. (2020) whose aim is an introduction of a copula via a vector of continuous random variables.

Considering the fact that copulas have a wide range of applications providing a general way for construction of multivariate distributions with continuous distributions, Fokianos et al. (2020) constructed a bivariate Poisson by introducing a copula form via a vector of continuous distributions and by utilizing fundamental properties of Poisson process. The method is flexible because it implies that each $Y_i$ can be a Poisson process but the joint distribution is not mandatory a Poisson distribution.

1. Define initial values for parameter vector $\lambda_0$ and generate values of $U_\ell = (U_{1,\ell}, U_{2,\ell}), \ell = 1, 2, ..., k$ from a 2-dimensional copula. Then $U_{i,\ell}$ follow marginally the uniform distribution on $(0,1)$ for $i=1,2$.

2. Consider the transformation $X_{i,\ell} = -\log U_{i,\ell}/\lambda_{i,0}, i = 1, 2$. Then the marginal distribution of $X_{i,\ell}, \ell = 1, 2, \ldots, k$ is exponential with parameter $\lambda_{i,0}, i = 1, 2$.

3. Define now taking $K$ large enough $Y_{i,0} = \max \left\{ \sum_{l=1}^{K} X_{i,l} \leq 1 \right\}, i = 1, 2$. Then $Y_0$ is marginally a set of first values of a Poisson process with parameter $\lambda_0$.

Fokianos et al. (2020) with the same way study a bivariate log-linear Poisson INGARCH model given by:

$$Y_t = N_t(v_t), \ v_t = b_0 + A v_{t-1} + B \log(Y_{t-1} + 1)$$

where $v_t = \log \lambda_t$ and $1_2$ is a 2-dimensional vector which consists of ones. Similarly to the linear model, the process $N_t$ denotes a sequence of independent 2-
CHAPTER 1. BACKGROUND

variate copula–Poisson processes which counts the number of events in $[0, \exp(v_{1,t})] \times [0, \exp(v_{2,t})]$.

At this model the author dealt with the problem of equidispersion. For this reason another model where data comes from a double Poisson distribution, was constructed by Heinen and Rengifo (2007) based on the definition given by:

$$X_t | F_{t-1} \sim DP(\lambda_t, \phi), \ t = 1, 2, \ldots$$

where $E(X_t | F_{t-1}) = \lambda_t$ and $Var(X_t | F_{t-1}) = \frac{\lambda_t}{\phi}$. A brief discussion with distinction on the rank of matrices A and B has been also presented. The above models were uncorrelated and a way with copulas of multivariate time series processes also discussed. The main problem was the lack of uniqueness of multivariate distribution in discrete case. To bypass the problem Heinen (2003) proposed a way for variables continuousation by adding at a discrete variable Z a randomly generated variable U from a Uniform distribution in (0,1) expressed as:

$$Z^* = Z + (U - 1)$$

The idea was based on Denuit and Lambert (2005) considering that this way of continuousation does not alter the concordance between the pairs of variables. A few years after the above study Heinen and Rengifo (2007) proposed an other one bivariate DPINGARCH model via the way of continuousation and using as tool for the construction of the bivariate distribution an archimedean copula. Assuming as previously a $DP(\lambda_t, \phi)$ where $\lambda_t$ given by (1.4), the joint density of the counts in the double Poisson distribution is:

$$h(X_{1,t}, X_{2,t}, \theta, \Sigma) = \prod_{i=1}^{2} f_{DP}(X_i, \lambda_{i,t}, \phi_i)c(q_i; \Sigma).$$

$f_{DP}$ denotes the double Poisson density as a function of the observation $Y_{i,t}$, $c(\cdot; \cdot)$ denotes the copula density of a multivariate normal and $\theta = (b_0, A, B)$. The $q_{i,t}$, gathered in the vector $q(t)$ are the normal quantiles of the $z_{i,t}$ and $z_{i,t}$ are the PIT of the continued extension of the count data under the marginal densities

$$z_{i,t} = F_{i,t}^*(X_{i,t}^*) = F_{i,t}(X_{i,t} - 1) + f_{i,t}(X_{i,t})U_{i,t},$$

where $F_{i,t}^*, F_{i,t}, f_{i,t}$ are the conditional cdf of the continued extension of the data, the conditional cdf and the conditional pdf respectively and $X_{i,t}^*$ are the continued extension of the original count data.
A new development of a multi-factor dynamic copula model with time-varying loadings and observation driven dynamics proposed recently by Opschoor et al. (2021). The flexibility of the above structure and two other features as scalability to high dimensions and positivity of covariance matrix provide a model that allow simple methods of estimation and easy study of many theoretical properties. A new bivariate conditional Poisson(BCP) INGARCH model defined by based on the BCP distribution introduced from Piancastelli et al. (2020).

**Definition 6.** A vector \((X_1, X_2)\) follows a BCP distribution with parameters \(\lambda_1, \lambda_2 > 0\) and \(\phi \in \mathbb{R}\) if \(Z_1 \sim \text{Poisson}(\lambda_1)\) and \(X_2 \mid X_1 = x_1 \sim \text{Poisson}(\mu_2 e^{\phi x_1})\) where \(\mu_2 = \lambda_2 \exp\{-\lambda_1 (e^\phi - 1)\}\) and the marginal of the Poisson distribution is a mixture Poisson distribution.

The main difference is that at the study of Piancastelli et al. (2020) \(\lambda_2\) replaced by the marginal mean of \(X_2\) and the joint BCP given by:

\[
f(x_1, x_2) = \prod_{i=1}^{2} \frac{\lambda_i^{x_i}}{x_i!} \exp\left\{-\lambda_1 (1 + x_2 (e^\phi - 1)) - \lambda_2 \exp\left(-\lambda_1 (e^\phi - 1) + \phi x_1\right) + \phi x_1 x_2\right\}.
\]

Parameter \(\phi\) is appropriate for model’s dependence considering as covariance and correlation respectively

\[
\text{cov}(X_1, X_2) = \lambda_1 \lambda_2 (e^\phi - 1),
\]

\[
\text{corr}(X_1, X_2) = (e^\phi - 1) \sqrt{\frac{\lambda_1 \lambda_2}{1 + \lambda_2 (e^{\lambda_1 (e^\phi - 1)}^2) - 1}}.
\]

It is worth mentioned that this model is mathematically tractable and considering a wide range of values for \(\phi\) a broad interval for correlation coefficient is also provided. At this paper a brief discussion on how unclear correlation coefficient’s range had been constructed at the work of Cui and Zhu (2018) also presented. Piancastelli et al. (2020) studied a bivariate INGARCH(1,1) model and using a parametrization for the bivariate conditional Poisson distribution different from the almost proposed by Berkhout and Plug (2004) provide a wider range of contemporaneous correlation.

1.6 Objective & Structure of the thesis

While the above models have been extensively studied, main obstacles considering the lack of closed forms for construction of multivariate Poisson at more than two
dimensions have been observed. Cui and Zhu (2018); Cui et al. (2020) studied bivariate INGARCH models with ideas based on the work of Lakshminarayana et al. (1999) and copulas. At the first study of Cui and Zhu (2018) distributions expressed based on the namely Farlie-Gumbel-Morgenstern copula while at the second one of Cui et al. (2020) at Frank and Gaussian copulas. At the above cases researchers deal with the problem of negative correlation. Although the difficulty is focused not only on the construction of a multivariate Poisson distribution that provides negative correlation but on the dependence structure which does not offer flexibility. Crucial attention is needed at parameters’ estimation where classical approaches as maximum likelihood approximation seem to be time-consuming depending each time on the copulas’ form. Cui and Zhu (2018) and Cui et al. (2020) decomposed the log-likelihood function based on maximization by parts and a modified version of this as alternative approaches with less complexity but with a bit high computational cost. Another approximation was presented by Fokianos et al. (2020) whose aim is an introduction of a copula via a vector of continuous random variables. Instead of quasi likelihood approximation at this study a bootstrap procedure also provided for copulas parameters. The limitation is that correlation boundaries at those multivariate models are restricted and this derives from the choice of copula which provide complex models. At the bivariate cases considering measures proposed by Meyn and Tweedie (2012), some theoretical aspects such as properties of stationarity and ergodicity are also provided. At the multivariate case theoretical aspects have been proved based on the similar way to this of univariate linear INGARCH Poisson model by Fokianos et al. (2009). Theoretical aspects are more easier to prove in case where volatilities at heteroskedastic models are linearly expressed than those of exponential volatilities expressions. It is worth mentioned the study of theoretical properties of a multivariate log-linear INGARCH Poisson model presented by Fokianos et al. (2020). The need of study is derived considering the vital lack of flexible Poisson distribution and regarding the flourishing literature of OD models with a limited research and progress of multivariate flexible heteroskedastic Poisson models.

The objective of the thesis is to develop and study fundamentally time series models leading to study and modeling multivariate time series models for count and continuous data. Apart from the construction of multivariate stochastic processes, we examine multivariate distributions that offer not only positive and negative correlation but also provide wider range for correlation coefficient boundaries. Furthermore specific cases of correlation coefficient in case of Poisson and Exponential
distributions that arise from the construction of multivariate distributions have been
discussed in detail. Methods for variable and model selection have been developed
in order firstly to simplify a model considering less variables and than the proposed
and to select among few models not necessarily from the same family. Those meth-
ods are not based on classical approaches but on bayesian methodologies. This is
an important issue when there are studies with data that have been modelled using
different processes but we need to examine which process is the best to fit the data.
For this reason a bayesian model selection method will be developed to the next
chapter among linear and log-linear INGARCH models.

In particular in chapter 2 using GLM models such as those proposed by Ferland
et al. (2006) where observations are Poisson distributed and mean expressed linearly
or log-linearly via its past values and past-observations, a bayesian model selection
method is provided. Based on posterior models probabilities, posterior predictive
distributions for volatilities have been also calculated. Efficiency and consistency
of prior distribution are examined taking into account two criteria and providing a
simulation study. Model selection method with the proposed priors is demonstrated
through applications to two different datasets.

In chapter 3 a very interesting issue is discussed on the identification of multivari-
ate distributions. The contribution is focused on the development of multivariate
distributions providing negative correlation coefficient values. The construction of
multivariate distribution is providing using a multiplicative factor responsible for
variables dependence and allowing flexibility at correlation coefficient. Based on
this structure, two case studies where random variables are Poisson and Exponen-
tially distributed are discussed. Considering that out multivariate distribution is an
alternative construction to this proposed by Le et al. (1996), a graphical inspection
is represented yielding correlation coefficient boundaries in comparison with others
using different values for exponents in multiplicative factor.

In chapter 4 a multivariate INGARCH(1,1) model is studied based on multi-
ivariate distribution proposed at the previous chapter. Two models have been also
discussed by Fokianos et al. (2020) and Piancastelli et al. (2020). The objective of
the new model is to allow flexibility on the way of study theoretical properties such
as ergodicity and stationarity of the model and parameter estimation. In addition
considering that in trivariate case number of parameters becomes large because mul-
tiplicative factor contains all the combinations of two out of three variables, other
ways of estimations could be usefull for parameter estimation. For that reason de-
composition of score function tends to be necessary and a likelihood by parts method
briefly discussed.

In chapter 5 based on the work of Chou (2005); Chan et al. (2019); Miralles-Marcelo et al. (2013) we study a CARR(1,1) model where variable is defined as the difference between daily maximum and minimum stock prices. Study for ergodicity does not differ from this of INGARCH(1,1) model but the main difference in bivariate model is that we do not consider interactions between the processes. The drawback is that no implementation has been provided due to lack of intradaily data that are not needed for model construction. The contribution is focused on the the structure of exponential CARR(1,1) model with strong correlation lying in the interval (-0.95,0.95).

In the last chapter some challenges for future work are summarized.
Chapter 2

Bayesian models selection for INGARCH models

This chapter focuses on the theory and model selection for a class of INGARCH processes. Assuming that observations $Y_t, t = 0, 1, \ldots$ are Poisson distributed, we proposed two models where the difference obtained on the definition of the mean process. Based on Poisson distribution the first and second order conditional moments are equal. To overcome the problem of equidispersion and deal with observations with extreme values, further extensions with other conditional distributions of $Y_t$ such as a negative binomial, a generalized Poisson, a COM-Poisson distribution and a zero-inflated Poisson by Zhu (2011, 2012a,b,c) have been studied respectively. We focus on the model selection among a few linear and log-linear Poisson INGARCH models. The reason to do this is that at log-linear INGARCH models parameter values lie on less restrictive intervals offering flexibility at volatility. More specifically we consider observations $\{X_t\} \in 0, 1, \ldots$ and $\lambda_t, t \in \mathbb{Z}$ be a sequence of mean process. In the first Poisson heteroskedastic model, we consider that observations $X_t$ given the past values are Poisson distributed and mean $\lambda_t$ is linked linearly with its past values and past observations. Namely we assume that

$$X_t \mid \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t)$$  \hspace{1cm} (2.1)

where $\mathcal{F}_{t-1}$ is the $\sigma$-field generated by $\{X_s : s \leq t - 1\}$ and

$$\lambda_t = d_1 + \sum_{i=1}^{p} a_i \lambda_{t-i} + \sum_{j=1}^{q} b_j X_{t-j}$$  \hspace{1cm} (2.2)

Assuming that $\lambda_t$ is the mean of Poisson distribution, parameters $d_1, a_i, b_j$ are positive for all $i = 1, \ldots, p, j = 1, \ldots, q$. $p$ and $q$ determine the order of the model and initial values $X_0$ and $\lambda_0$ are fixed. Furthermore considering Poisson process
\( E(X_t \mid \mathcal{F}_{t-1}) = \operatorname{Var}(X_t \mid \mathcal{F}_{t-1}) = \lambda_t \). This model proposed by Ferland et al. (2006) based on the idea of GARCH model proposed by Bollerslev (1986). Heinen (2003) studied unconditional moments of an INGARCH(1,1) model based on martingale properties and Ferland et al. (2006) study properties and moments of the same model based on theory of ARMA processes. Assuming that \( \{X_t\} \) and \( \{Y_t\} \) are zero mean stationary processes with autocovariance function \( \gamma(\cdot) \) and \( \{X_t\} \) is an ARMA(1,1) process defined as

\[
Y_t - \phi Y_{t-1} = \epsilon_t + \theta \epsilon_{t-1} \tag{2.3}
\]

where \( \epsilon_t \) is an i.i.d. sequence \( N(0, \sigma^2) \), \( \phi_1 = a_1 + b_1 \), \( \theta_1 = -a_1 \) and \( \sigma^2 = E(Y_t) \) then \( \{Y_t\} \) is also an ARMA(1,1) process. According to stationarity conditions

\[
E(X_t) = E(E(X_t \mid \mathcal{F}_{t-1})) = E(\lambda_t) = d_1 + a_1 E(\lambda_{t-1}) + b_1 E(X_{t-1})
\]

\[
= d_1 + a_1 E(\lambda_t) + b_1 E(\lambda_{t-1} | \mathcal{F}_{t-1})
\]

\[
= d_1 + a_1 E(\lambda_t) + b_1 E(\lambda_t) \tag{2.4}
\]

where

\[
\mu = E(\lambda_t) = d_1 (1 - a_1 - b_1)^{-1} \tag{2.5}
\]

Based on stationarity condition of ARMA(1,1) model we have that \( a_1 + b_1 < 1 \) and \( d_1 > 0 \) Considering martingale property where the deviation between dependent variable and conditional mean are independent from the information set at time \( t \) and \( \mathcal{I}_t \) is the \( \sigma \)-algebra of \( \{\lambda_s, s \leq t-1\} \) we have:

\[
\operatorname{Cov}(X_t - \lambda_t, \lambda_{t-k}) = E[(X_t - \lambda_t)(\lambda_{t-k} - \lambda_t)] = 0 \tag{2.6}
\]

\[
\operatorname{Cov}(X_t, X_{t-k} - \lambda_{t-k}) = E[(X_t - \mu)(X_{t-k} - \lambda_{t-k})] = 0 \tag{2.7}
\]

\[
\operatorname{Cov}(\lambda_t, \lambda_{t-k}) = a_1 \operatorname{Cov}(X_{t-1}, \lambda_{t-k}) + b_1 \operatorname{Cov}(\lambda_t, \lambda_{t-k}) \tag{2.8}
\]

and by iterating (2.8) we obtain:

\[
\operatorname{Cov}(\lambda_t, \lambda_{t-k}) = (a_1 + b_1)^k \gamma_\lambda(0)
\]
2.1. STATIONARITY OF LINEAR AND LOG-LINEAR INGARCH MODELS

and

\[ \text{Cov}(X_t, X_{t-k}) = \text{Cov}(X_t, \lambda_{t-k}) = \text{Cov}(\lambda_t, X_{t-k}) \]

In order to prove the second order unconditional moment we could consider that

\[ \text{Var}(X_t) = E(\text{Var}(X_t \mid F_{t-1})) + \text{Var}(E(X_t \mid F_{t-1})) \]

or based on ARMA(1,1) model we have that \( \text{Var}(X_t) = \frac{1 - (a_1 + b_1)^2 + a_1^2}{1 - (a_1 + b_1)^2} E(X_t) \) where \( E(X_t) \) given by (2.4). The drawback with this model is that parameters take only positive values according to the expression of volatility. To overcome this problem another model based on theory of GLM has been proposed by Fokianos and Tjøstheim (2011) assuming

\[ X_t \mid F_{t-1} \sim \text{Poisson}(v_t = \log(\lambda_t)) \]  (2.9)

where

\[ v_t = d_1 + \sum_{i=1}^{p} a_i v_{t-i} + \sum_{j=1}^{q} b_j \log(X_{t-j} + 1) \]  (2.10)

In this model, parameters \( d_1, a_i, b_j \) for any \( i = 1, \ldots, p, j = 1, \ldots, q \) take values in \( \mathbb{R} \) and both negative and positive correlations can occur. Certainly other specifications are possible. In particular, one may consider a model for the log-mean process by introducing \( \log(X_{t-j} + u) \), where \( u \) is a constant. Fokianos and Tjøstheim (2011) presented results of the data analysis that do not indicate any gross deviations in terms of the mean square error of residuals for values of \( u \) varying from 1 to 10 with a step equal to 0.5. We consider that \( u = 1 \) throughout the thesis and parameter \( \lambda_t \) is defined in (2.1), (2.2), (2.9), (2.10). Zeger and Qaqish (1988) studied parameters estimation approach without discussion for conditions of ergodicity and stationarity. Furthermore Shephard (1995) studied a similar model called Integer moving average OD model, where the main difference is that mean expressed log-linearly only from past values of observations. Fokianos and Kedem (2004) studied moments of the model given by (2.9), (2.10) and a different approach of estimation has been discussed.

2.1 Stationarity of linear and log-linear INGARCH models

Ferland et al. (2006) discussed first and second order stationarity based on backshift operator for the general INGARCH(p,q) process. Based on \( B(d) \lambda_t = d_1 + A(d)X_t \) where \( B(d) = 1 - B^d \) and \( A(d) = 1 - A^d \) are lag-operators used to produce previous elements at volatilities and time series. Then \( \lambda_t = B^{-1}(d)d_1 + B^{-1}dA(d) \). There
are numerous criteria in order to prove process ergodicity. Fokianos et al. (2009) study conditions of ergodicity considering a perturbed model and based on criteria of \( \phi \)-irreducibility, geometric ergodicity has been provided. Assuming that \( \lambda_t \) is a Markov chain, where \( \lambda_t = d_1 + a_1 \lambda_{t-1} \), Fokianos et al. (2009) tried to prove geometric ergodicity considering \( \phi \)-irreducibility. This is equivalent to show that \( \sum P^m(\lambda, A) > 0 \) or \( \phi(A) > 0 \), where \( P^m(\lambda, A) = P\{\lambda_n \in A \mid \lambda_0 = \lambda\} \). Streett (2000) proved that under certain conditions there exist a stationary initial distribution for \( \lambda_t \) by showing that a point \( \lambda^* \) is reachable. The problem is that some sets \( A \) of positive Lebesque measure are not open sets and for this reason Fokianos et al. (2009) used a perturbed model defined by

\[
\lambda_t^m = d_1 + a_1 \lambda_{t-1}^m + b_1 X_{t-1}^m + \epsilon_t^m
\]

where \( \epsilon_{t,m} = c_m I(X_{t-1}^m = 1) U_t \), \( I(\cdot) \) is an indicator function, \( U_t \) is a sequence of iid uniform random variables appropriate to establish irreducibility. Furthermore \( c_m \rightarrow 0 \) when \( m \rightarrow \infty \). The above indicate small differences between the initial and the perturbed model offering to provide infinite partial derivatives and estimator convergence for both of two models. Several specific cases and examples have been also proposed by Dedecker et al. (2007). Franke (2010) based on \( \theta \) and \( \tau \) weak dependence criteria proved stationarity of the univariate INGARCH process.

The main idea for the ergodicity is that considering two convenient functions \( f, g \), \( \text{Cov}(f(\text{past}), g(\text{future})) \) tends to 0 as the distance between past and future becomes large, where ‘past’ and ‘future’ are elementary events given through finite dimensional marginals. For those criteria we study functions which satisfy Lipschitz condition. Considering Euclidean space \( \mathbb{R}^m \) and \( h : \mathbb{R}^m \rightarrow \mathbb{R}^n \)

\[
Liph = \sup \left\{ \frac{|h(\lambda_t, X_t) - h(\tilde{\lambda}_t, \tilde{X}_t)|}{|\lambda_t - \tilde{\lambda}_t| + |X_t - \tilde{X}_t|} \right\}
\]

Let \( A_1(\mathbb{R}^m) \) be the set of functions \( h \) where \( Lip(h) \leq 1 \).

The above conditions indicate that model (2.1); (2.2) is ergodic and has a strictly stationary solution. Doukhan et al. (2012) studied weak dependence criteria for two INGARCH time series models studied also by Fokianos et al. (2009). In contrast for the model defined by (2.10) criteria of weak dependence cannot be proved easily because logarithmic function does not possess the contraction property. Fokianos and Tjøstheim (2011) provided conditions of geometric ergodicity based on the idea of Fokianos et al. (2009) and considering a perturbed model given by:

\[
v_t^m = d_1 + a_1 v_{t-1}^m + b_1 \log(X_{t-1}^m + 1) + \epsilon_{t,m}
\]
where similar to the linear model $\epsilon_{t,m} = c_m I(X_{t-1}^m = 1)U_t$, $U_t$ are iid uniformly distributed random variables, $I()$ is the indicator function and $c_m > 0$, $c_m \to \infty$, $m \to \infty$. Considering the above study the chain $v_t$ is $\phi$-irreducible and geometric ergodic when $|a_1| < 1$ and $|a_1 + b_1| < 1$. Liboschik et al. (2015) proposed alternative and more flexible conditions for log-linear model’s stationarity given by:

$$\max\{|a_1|, |b_1|, |a_1 + b_1|\} < 1$$

Fokianos and Tjøstheim (2012) studied geometric ergodicity and moments for a class of non-linear models considering perturbed models.

Similar to linear model considering conditions of stationarity, for the log-linear INGARCH(1,1) we obtain that:

$$E(X_t) = E(E(X_t \mid \mathcal{F}_{t-1})) = E(E(v_t)) = E(E(\log(\lambda_t)))$$
$$= E(d_1 + a_1 E(\log(\lambda_{t-1})) + b_1 \log(Y_{t-1} + 1))$$
$$= d_1 + a_1 E(\log(\lambda_{t-1})) + b_1 E(\log(\lambda_{t-1}))$$

and

$$E(\log(\lambda_t)) = \frac{d_1}{1 - a_1 - b_1}.$$ 

### 2.2 Parameters estimation

Estimation of parameters is necessary before applying one of the methods for bayesian model selection. Problem of estimation via maximum likelihood is treated firstly by Ferland et al. (2006). The conditional likelihood function for the parameters of linear INGARCH model given by:

$$L(x \mid \theta) = \prod_{t=1}^{n} \frac{e^{-\lambda_t^x t} \lambda_t^{x_t}}{x_t!}$$ (2.11)

where $\lambda_t$ is expressed via (2.2) and $\theta = (\theta_1, \theta_2, \ldots, \theta_{p+q+1}) = (\alpha_1, \ldots, \alpha_p, d_1, b_1, \ldots, b_q)$ is the vector of the unknown parameters and $y_t$ the observed value at time $t$. The score function defined as:

$$S(\theta) = \frac{\partial l}{\partial \theta} = \sum_{t=1}^{n} \left( \frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta}$$ (2.12)

while the second derivatives are:

$$\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 \lambda_t}{\partial \theta_i \partial \theta_j} \left( \frac{X_t}{\lambda_t} - 1 \right) - \frac{X_t}{\lambda_t^2} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\lambda_t}{\partial \theta_j}.$$ (2.13)
where \( \ell \) is the log-likelihood given by:

\[
l = \sum_{t=1}^{n} \lambda_t + x_t \log(\lambda_t - \log(x_t!))
\]

First and second order derivatives are calculated and considering conditions of stationarity it is easily to prove that \( E\|\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}\| \leq \infty \).

\[
\frac{\partial^2 l}{\partial a_1^2} = -\frac{X_t \partial^2 \lambda_t}{\lambda_t^2} \frac{\partial a_1^2}{\partial a_1^2}
\]

\[
E\left( \frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \mid F_{t-1} \right) = -\frac{1}{X_t} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j}
\]

Considering Taylor series expansion for \( \frac{\partial l}{\partial \theta} \) and considering Lemma 6 of Lee et al. (2016)

\[
-\frac{1}{T} \frac{\partial l(\theta_0)}{\theta_i \theta_j} \to I(\theta_0)
\]

we have that

\[
\sqrt{T}(\hat{\theta} - \theta_0) \to N(0, I(\theta_0)^{-1}).
\]

Likelihood function defined similarly to (2.11) for the linear model and considering construction for the volatility

\[
l(\theta) = \sum_{i=1}^{T} (X_t v_t - e^{v_t} - \log(X_t!))
\]

On the other hand the log-likelihood function for the perturbed model given by:

\[
l(\theta) = \sum_{t=1}^{n} \left( X_t^m v_t^m - e^{v_t^m} - \log(X_t^m!) + \log(f_u(U_t)) \right)
\]

where \( f_u(\cdot) \) denotes the uniform density respectively, second order derivatives are given by:

\[
\frac{\partial^2 v_t(\theta)}{\partial d_1^2} = a_1 \frac{\partial^2 v_{t-1}}{\partial b_0^2}
\]

\[
\frac{\partial^2 v_t(\theta)}{\partial a_1^2} = 2 \frac{\partial v_{t-1}}{\partial a_1} + a_1 \frac{\partial^2 v_{t-1}}{\partial a_1^2} + a_1 \frac{\partial^2 v_{t-1}}{\partial b_1^2}
\]

\[
\frac{\partial^2 v_t(\theta)}{\partial b_1^2} = a_1 \frac{\partial^2 v_{t-1}}{\partial b_1^2}
\]
Providing that difference between the perturbed and the unperturbed model can be eliminated as \( m \to \infty \). Fokianos and Tjøstheim (2011) studied asymptotic normality and consistency of estimators by using the fact that the likelihood function of the perturbed model is three times differentiable. By using proposition 6.3.9 of Brockwell and Davis proved that the score function, information matrix and third derivatives tend to the quantities of the unperturbed model.

\[
E\left(\frac{1}{\lambda_t^m} \frac{\partial \lambda_t^m}{\partial \theta} \frac{\partial \lambda_t^m}{\partial \theta}^T\right) \to E\left(\frac{1}{\lambda_t} \frac{\partial \lambda_t}{\partial \theta} \frac{\partial \lambda_t}{\partial \theta}^T\right)
\]

\[
\frac{1}{\sqrt{n}} \frac{\partial^2 l_m(\theta_0)}{\partial \theta_i \partial \theta_j} \to I_m(\theta_0) \text{ as } T \to \infty
\]

\[
I_m(\theta_0) \to I(\theta_0) \text{ as } m \to \infty
\]  

(2.18)

### 2.3 Parameters estimation based on Bayesian statistics methods

Using Bayes theorem posterior density \( p(\theta \mid y) \) for the parameter vector \( \theta = (\theta_1, ..., \theta_{p+q+1}) = (a_1, a_2, ..., a_p, d_1, b_1, ..., b_q) \) can be derived. We have as usual that

\[
p(\theta \mid x) \propto L(x \mid \theta)p(\theta) \tag{2.19}
\]

\[
p(\theta \mid x)p(\alpha_1) \ldots p(\alpha_p)p(d_1) \ldots p(b_q)
\]

where \( L(x \mid \theta) \) is the likelihood function and \( p(\theta) \) is the prior distribution. Here we assume independent priors for all the parameters. We use MCMC method to obtain samples from the posterior distribution. Based on Bayes theorem and under the assumption of stationarity we choose as prior distributions truncated normals for the parameters \( \alpha_i \) and \( b_j \) where \( 0 < \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} b_j < 1 \) or \( -1 < \alpha_i, b_j < 1 \) and \( -1 \leq \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} b_j < 1 \) in linear and log-linear case accordingly and a Gamma(1,2) prior distribution for the parameter \( d_1 \) and after Metropolis-Hasting steps we reject values that they do not satisfy stationarity conditions.

Given that \( X_t \mid \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t) \), where \( \lambda_t \) is expressed via equations (2.2) or (2.10), posterior density of parameter vector \( \theta = (\alpha_1, ..., \alpha_p, d_1, b_1, ..., b_q) \) can be expressed by equation
\[ p(\theta \mid \mathbf{x}) \propto \prod_{t=1}^{n} e^{-\lambda_{t} x_{t}^{2}} \times \frac{1}{e^{2\pi^{2}}} \left(\frac{\alpha_{1} - \mu_{1}}{\sigma_{\alpha_{1}}} \right)^{2} \times \ldots \times \frac{1}{e^{2\pi^{2}}} \left(\frac{\alpha_{p} - \mu_{p}}{\sigma_{\alpha_{p}}} \right)^{2} \times \frac{1}{e^{2\pi^{2}}} \left(\frac{\beta_{1} - \mu_{\beta_{1}}}{\sigma_{\beta_{1}}} \right)^{2} \times \ldots \times \frac{1}{e^{2\pi^{2}}} \left(\frac{\beta_{q} - \mu_{\beta_{q}}}{\sigma_{\beta_{q}}} \right)^{2} \times \Phi \left(\frac{1}{\sigma_{\alpha_{1}}} - \mu_{\alpha_{1}} \left| \theta_{1} \right. \right) - \Phi \left(\frac{1}{\sigma_{\alpha_{p}}} \mu_{\alpha_{p}} \right) \times \ldots \times \Phi \left(\frac{1}{\sigma_{\beta_{1}}} - \mu_{\beta_{1}} \left| \theta_{1} \right. \right) - \Phi \left(\frac{1}{\sigma_{\beta_{q}}} \mu_{\beta_{q}} \right)
\]

where

\[ \mathbb{S} = \{\alpha_{1}, \ldots, \alpha_{p}, \beta_{1}, \ldots, \beta_{q}\} = \begin{cases} \sum_{i=1}^{p} a_{i} + \sum_{j=1}^{q} b_{j} \in (0, 1) \\
|\alpha_{i}|, |b_{j}| < 1, \ |\sum_{i=1}^{p} \alpha_{i} + \sum_{j=1}^{q} b_{j}| < 1, \end{cases} \]

\[ u = \begin{cases} 0, \text{if } \sum_{i=1}^{p} a_{i} + \sum_{j=1}^{q} b_{j} \in (0, 1) \\
-1, \text{if } |\sum_{i=1}^{p} \alpha_{i} + \sum_{j=1}^{q} b_{j}| < 1. \end{cases} \]

and \( \gamma \left(\gamma_{1}, \frac{d_{1}}{\beta} \right) \) is the lower incomplete gamma function. We use \( p+q+1 \) independent Metropolis steps (Tierney (1994); Chib and Greenberg (1995)) for all parameters \( \theta_{i} \) and as proposal distribution \( q(\theta^{(can)} \mid \theta_{1}, \theta_{2}, \ldots, \theta_{p+q+1}) \) we use \( N(\theta, \sigma^{2}) \) distributions. The choice for this variance \( \sigma^{2} \) is important for the convergence of the chain. Considering the conditions of stationarity, the interval for each parameter our work is \( (0, 1) \) and \( (-1, 1) \) at the linear and the log-linear cases respectively but for the parameter \( d_{1} \) the positivity is the only restriction in both cases. In the work of Fokianos and Tjøstheim (2011) the restrictions of the parameters have been discussed, because they provide different correlation structure and they are deterministic for the calculation of the maximum likelihood estimators. Acceptance probability at the \( j \)-th iteration is calculated by

\[ \alpha = \min \left\{ 1, \frac{\prod_{t=1}^{T} e^{-\lambda_{t}^{(can)} \left(\lambda_{t}^{(can)} \left| \theta_{t}^{(can)} \right. \right) y_{t} p(\theta_{t}^{(can)})}}{\prod_{t=1}^{T} e^{-\lambda_{t}^{(j)} \left(\lambda_{t}^{(j)} \left| \theta_{t}^{(j)} \right. \right) y_{t} p(\theta_{t}^{(j)})}} \right\} , \]

where \( \lambda^{(can)} \) implies that it has been calculated based on the candidate value for the parameter. Marginal posterior means and posterior variances are used on trans-dimensional MCMC method for the definition of pseudopriors.
2.4. METHODS FOR BAYESIAN MODEL SELECTION

Denote as \( \theta = (\alpha_1, \ldots, \alpha_p, d_1, b_1, \ldots, b_{p+q}) \) the vector with all the parameters. We describe the MCMC algorithm in a general form applicable to both linear and log-linear models. The key distinction is the conditions that need to be satisfied for each model to ensure stationarity.

We use \( p+q+1 \) independent Metropolis steps for all parameters \( \theta_i, i = 1, \ldots, p+q+1 \). The algorithm runs as follows

- Initialize \( \theta^{(0)} = (\alpha_1^{(0)}, \ldots, \alpha_p^{(0)}, b_1^{(0)}, \ldots, b_q^{(0)}, d_1^{(0)}) \) from truncated normal distributions for \( \alpha \)'s and \( b_1, \ldots, b_q \) and a Gamma for \( d_1 \) such as to satisfy the stationarity conditions of the particular model.
- For the \( r+1 \)-th iteration of the algorithm
  - Generate values from proposal densities for \( \theta^{(can)}_i \sim N(\theta_i, \sigma_i^2) \) where the tuning parameter \( \sigma_i^2 \) was selected such as to achieve 25% acceptance rate.
  - Update \( \theta^{(r)}_i \) to \( \theta^{(r+1)}_i \) with probability given by (2.21) and if the candidate value does not violate the stationarity conditions of the model.
  - Iterate this updating procedure.
- Go to the next iteration or stop if the chain has converged.

2.4 Methods for Bayesian model selection

While the literature on discrete valued time series is increasing very fast, there are very few papers that consider the problem of model selection, namely that of selecting the order of the model. In simple INAR\((p)\) models this is merely the order of autoregressive terms, in our INGARCH setting this relates to the order of \( p \) and \( q \) to be used. An efficient reversible jump Markov chain Monte Carlo (RJ-MCMC) algorithm is constructed by Enciso-Mora et al. (2009) for moving between INARMA processes of different orders. See also work for model selection in INAR\((p)\) models in Bu and McCabe (2008).

In addition, Alzahrani et al. (2018) studied the model selection problem among INAR and linear INGARCH models by using particle filtering MCMC method. Wang et al. (2020) used a penalized conditional maximum likelihood to estimate the parameters of an INGARCH model. Note that this approach while can help to identify useful terms in the model it does not select between competing models. Assume that we have a countable set of models denoted by \( \mathcal{M} \) for a given set of data \( x \). Let consider a model indexed by \( m \in \mathcal{M} \), we denote the vector of
unknown parameters for model $m$ as $\theta_m \in \Theta_m$. Two models, say $m_1$ and $m_2$, can have different dimensions. Let $p(m)$ the prior distribution for each model $m$. Then posterior probability for model $m \in \mathcal{M}$ is given by

$$p(m \mid x) = \frac{p(m) L(x \mid m)}{\sum_{m' \in \mathcal{M}} p(m') L(x \mid m')}.$$  

(2.22)

where $L(x \mid m)$ is the marginal likelihood for model $m$ calculated by the integral

$$\int L(x \mid \theta_m)p(\theta_m \mid m)d\theta_m$$

and $p(\theta_m \mid m)$ is the prior distribution of the parameter vector $\theta_m$. There are many approaches for model selection problem using Bayes factor of model $m_i$ against model $m_j$. Basically there are two interpretations for Bayes factor. Firstly Bayes factor is the ratio of two marginal likelihoods, representing how well the data are fitted under each model. More analytically

$$BF = \frac{p(m_i \mid x)p(m_j)}{p(m_j \mid x)p(m_i)} = \frac{\int_{\Theta_m} p(x \mid m_i, \theta_m)p(\theta_m \mid m_i)d\theta_m}{\int_{\Theta_m} p(x \mid m_j, \theta_m)p(\theta_m \mid m_j)d\theta_m}.$$  

(2.23)

Secondly, we could say that posterior model odds can be represented as

$$\frac{p(m_i \mid x)}{p(m_j \mid x)} = B_{ij} \times \frac{p(m_i)}{p(m_j)}.$$  

(2.24)

The integrals required for Bayes factor are analytically intractable. Consequently many methods have been proposed to approximate Bayes factor. Green (1995) proposed the so called reversible jump MCMC method allowing jumps between models with low computational cost. The idea of Dellaportas et al. (2002) based on the above methods. Reversible jump MCMC is more flexible because proposals only for a subset of parameters is required. Assuming that we want to make a jump from model $m_i$ with parameter vector $\theta_{m_i} = (\theta_1, \theta_2)$ to model $m_j$ with parameter vector $\theta_{m_j} = (\theta_1, \theta_2, \theta'_3)$ where those two models are nested. For this reason we generate $u$ from a specified proposal density $q(u \mid \theta_{m_i}, m_i, m_j)$. When we are in a model where those 2 models are nested, a proposal only for the third parameter is needed. In the general case $(\theta_{m_j}, u') = g_{m,m_j}$ where $g_{m,m_j}$ is a specified function where $g_{m,m_j} = g_{m,m_j}^{-1}$. Then the acceptance probability is calculated by:

$$\frac{p(x \mid m_j, \theta_{m_j})p(\theta_{m_j} \mid m_j)}{p(x \mid m_i, \theta_{m_i})p(\theta_{m_i} \mid m_i)q(u \mid \theta_{m_i}, m_i, m_j)}.$$  

Methods, introduced by Green (1995), Carlin and Chib (1995), Dellaportas et al. (2002), generate observations from the joint posterior distribution from $(m, \theta_m)$, for estimating $p(\theta_m \mid m)$. A method introduced by Carlin and Chib (1995) proposes
jumps between all models. The full conditional posterior density for each model is given by
\[
p(m \mid \theta_m, x) = \frac{A_m}{\sum_{m' \in M} A_{m'}},
\]
where numerator is given by equation
\[
A_m = p(x \mid \theta_m, m) \prod_{m' \in \$} p(\theta_{m'} \mid x)p(m').
\]
CHAPTER 2. BAYESIAN MODELS SELECTION FOR INGARCH MODELS

In this method and in the method introduced by Dellaportas et al. (2002) there is no linked densities and also the calculation of the Jacobian is not needed. Those methods require proposal densities of higher dimensions, called pseudopriors.

The crucial problem here is the matching of dimensions of the three models. For doing this consider that if model $m_1$ is visited then the corresponding parameter vector $\theta_{m_1}$ is connected to the model likelihood. So this parameter vector can be updated from posterior distributions. If $m_1$ is not visited then the parameter vector is disconnected from the likelihood and is generated from pseudopriors. A similar reasoning can be followed for the models $m_2$ and $m_3$. For example, for model $m_1$ the joint distribution is defined by

$$p(x, \theta | m_1) = p(x | \theta, m_1)p(\theta | m_1)$$

$$= p(x | \theta, m_1)p(\theta_{m_1} | m_1)p(\theta_{m_2} | m_1)p(\theta_{m_3} | m_1),$$

where $p(\theta_{m_2} | m_1)$, $p(\theta_{m_3} | m_1)$ are proper distributions integrating to 1. Then for model $m_i$, $i = 1, 2, 3$ we sample from

$$p(m_i | \theta, x) \propto \begin{cases} \frac{L(x | \theta, m_1)p(\theta_{m_1} | m_1)p(\theta_{m_2} | m_1)p(\theta_{m_3} | m_1)p(m_1), \text{ for } i = 1} \frac{L(x | \theta, m_2)p(\theta_{m_2} | m_2)p(\theta_{m_3} | m_2)p(m_2), \text{ for } i = 2} \frac{L(x | \theta, m_3)p(\theta_{m_3} | m_3)p(m_3), \text{ for } i = 3} \end{cases}$$

where

$$p(\theta_{m_1} | m_i) \propto \begin{cases} p(x | \theta_{m_1}, m_1)p(\theta_{m_1} | m_i) \text{ if } i = 1 \frac{p(\theta_{m_1} | m_i)}{p(\theta_{m_1} | m_i)} \text{ if } i = 2 \frac{p(\theta_{m_1} | m_i)}{p(\theta_{m_1} | m_i)} \text{ if } i = 3 \end{cases}$$

and $p(m_1), p(m_2), p(m_3)$ are prior model probabilities. Those methods are deterministic for the time that the algorithm visit and stay at each model. Finally posterior model probabilities for model $m \in M$ are estimated by

$$\hat{\rho}(m | x) = \frac{\sum_{i=1}^{B} I(m^{(i)} = m)}{B},$$

where $B$ is the total number of iterations and $m^{(i)}$ denotes the model we are in the $i$-th iteration. We will apply this trans-dimensional approach to jump between different INGARCH models. A detailed description of the trans-dimensional MCMC can be found in the Appendix B. Predictions in time varying volatility in linear and log-linear INGARCH models can be obtained from the output of the MCMC. McCabe and Martin (2005) studied methods for prediction of INAR(1) model. Raftery
et al. (1997) discussed model averaging and the choice of models for Bayesian model averaging. A \( p \)-step ahead predictive pmf is defined as

\[
p(x_{T+p} \mid x) = \sum_{m \in M} p(x_{T+p} \mid x, m)p(m \mid x),
\]

(2.31)

where \( M \) denotes the countable set of candidate models and \( P(x_{T+p} \mid x, m) \) is the \( p \)-step ahead predictive pmf from model \( m \) and \( P(m \mid x) \) is the posterior probability of model \( m \) calculated in trans-dimensional MCMC. So in our work

\[
p(\lambda_{T+1} \mid x) = \sum_{m \in M} p(\lambda_{T+1} \mid x, m)p(m \mid x),
\]

(2.32)

which is an average of the posterior predictive distribution under each model weighted by their posterior model probabilities. For each parameter vector we calculate at each iteration \( \lambda_{T+1} \). Each sampled point should be taken with probability \( p(m \mid x) \).

Then we obtain the sample of \( p(\lambda_{T+1} \mid x) \) by weighting all samples of \( p(\lambda_{T+1} \mid m, x) \) by the corresponding \( p(m \mid x) \).

Consider \( s \) candidate models, namely \( m_1, \ldots, m_s \). We denote the parameter vector for model \( m \) as \( \theta_m \). Due to the different number of parameters for each model they may have different dimension. Also denote the set of indices not having \( j \) as \( S_{-j} = \{1, 2, \ldots, j-1, j+1, \ldots, s\} \). We have obtained from the MCMC from all the models the mean \( \tilde{\mu}_{\theta_{ij}} \) and the variance \( \tilde{\sigma}_{\theta_{ij}}^2 \) of parameter \( \theta_{ij} \) which is the \( i \)-th parameters of the \( j \)-th model. Then the algorithm runs as follows.

For a number of iterations say \( B \) we work as follows

1. For model \( m_k \) we generate the parameter vector for this model, say \( \theta_{mk} \) from the conditional posterior distribution \( L(x \mid \theta_{mk}, m_k)p(\theta_{mk} \mid m_k) \) given by (2.20) and the other parameters \( \theta_{mj} \), from proper pseudopriors \( p(\theta_{mj} \mid m_j) = \prod_{i=1}^{p_j} N(\tilde{\mu}_{\theta_{ij}}, \tilde{\sigma}_{\theta_{ij}}^2) \) for \( j \in S_{-k} \) where \( p_k \) denotes the number of parameters for model \( k \).

2. We repeat this for each candidate model so for all \( k = 1, \ldots, s \).

3. Calculate posterior probability for model \( k \) defined by

\[
Pr_k = p(m_k \mid \theta, x) = \frac{A}{B},
\]

where

\[
A = L(x \mid \theta_{mk}, m_k)p(\theta_{mk} \mid m_k)p(m_k) \times \prod_{j \in S_{-k}} \frac{1}{\sqrt{2\pi \tilde{\sigma}_{\theta_{ij}}}} \exp \left( -\frac{(\theta_{ij} - \tilde{\mu}_{\theta_{ij}})^2}{2\tilde{\sigma}_{\theta_{ij}}^2} \right).
\]
and

\[ B = \sum_{w=1}^{s} L(x \mid \theta_{m_w}, m_w)p(\theta_{m_w} \mid m_w)p(m_w) \times \prod_{j \in S_{-w}} \prod_{i=1}^{p_i} \frac{1}{\tilde{\sigma}_{ij} \sqrt{2\pi}} \exp \left( -\frac{(\theta_{ij} - \tilde{\mu}_{ij})^2}{2\tilde{\sigma}_{ij}^2} \right). \]

4. We select the model with larger value \( Pr_k \). this is the model selected at the \( i \)-th iteration,

5. Repeat steps 1-3. for \( B \) times.

6. At the end calculate s posterior model probabilities given by (2.30).

2.4.1 Simulations

In order to examine if our approach can identify the correct structure of the time series we study a small simulation study considering two criteria proposed by Liang et al. (2008) for model selection consistency. These are:

CRITERION 1: Each conditional prior must be proper (integrating to one) and cannot be arbitrarily vague in the sense of almost all of its mass being outside any believable compact set.

CRITERION 2: (Model selection consistency) If data \( y \) have been generated from model \( M_i \) then posterior model probability of model \( M_i \) should converge to 1 as the sample size \( n \to \infty \).

If criterion 2 is not accomplished may be a different choice of prior distributions is necessary. In our case for the accomplishment of criterion 1 we suggest as pseudopriors normal densities by taking means and variances after "pilot runs" of MCMC for each model and considering that those pseudopriors are proper. More specifically we generate 100 datasets of size \( n = 200 \) with randomly chosen parameter values for each model, and we run a trans-dimensional Markov chain with length 10000. According to criterion 2 posterior model probability of model from which we generate the data, must be close to 1 while the probabilities must be small for the other models. Furthermore the choice of prior model probabilities is deterministic for the definition of the posterior model probabilities. Lodewyckx et al. (2011) proposed bisection algorithm to choose prior model probabilities which have a vital role at
the algorithm. We present the results of averages of 100 replications for each model. The columns correspond to the frequency of selecting each one of 10 models, while the rows are the true models that generated the data. Note that in addition to the 10 INGARCH models we have two additional models different from INGARCH type. The first model was an INAR(2) model of the form $X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + R_t$, where $\circ$ is the binomial thinning operator and $R_t$ are innovations terms that follow a Poisson distribution, while the second model was an INAR(1) model of the form $X_t = \alpha_1 \circ X_{t-1} + R_t$, $t = 1, \ldots, n$ where now the innovation term depends on some covariate, namely $R_t \sim \text{Poisson}(\exp(\gamma_1 z_t))$ with $Z_t \sim \text{Ber} (\zeta)$ where $\gamma_1 = 2$ and $\zeta = 0.2$. So, reading from table (2.1) 91.5\% of the times we have generated the data from model $M_1$ we found this model as the chosen one.

Looking the results from Table (2.1) we can see that even for small sample size ($n = 100$) we can identify the correct structure with great success for most of the models. Furthermore based on the criteria mentioned above and considering results on Tables (2.2) and (2.3), when sample sizes are $n = 200$ and $500$ respectively, we have increasing probability to find the correct model.

Overall, the experiment shows that the method can identify the underlying model with good success. Our approach correctly detects the model with high probability.

<table>
<thead>
<tr>
<th>Model</th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$M_3$</th>
<th>$M_4$</th>
<th>$M_5$</th>
<th>$LM_1$</th>
<th>$LM_2$</th>
<th>$LM_3$</th>
<th>$LM_4$</th>
<th>$LM_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1 = \text{INGARCH}(1)$</td>
<td>90.57</td>
<td>1.74</td>
<td>1.13</td>
<td>3.48</td>
<td>0.09</td>
<td>2.55</td>
<td>0.24</td>
<td>0.10</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td>$M_2 = \text{INGARCH}(1,1)$</td>
<td>0.04</td>
<td>81.43</td>
<td>0.59</td>
<td>16.12</td>
<td>0.06</td>
<td>0.86</td>
<td>0.63</td>
<td>0.17</td>
<td>0.10</td>
<td>0.00</td>
</tr>
<tr>
<td>$M_3 = \text{INGARCH}(1,2)$</td>
<td>0.04</td>
<td>1.23</td>
<td>96.76</td>
<td>0.17</td>
<td>0.31</td>
<td>0.09</td>
<td>0.10</td>
<td>0.42</td>
<td>0.10</td>
<td>0.38</td>
</tr>
<tr>
<td>$M_4 = \text{INGARCH}(2,1)$</td>
<td>0.00</td>
<td>3.01</td>
<td>0.42</td>
<td>93.72</td>
<td>0.38</td>
<td>2.07</td>
<td>0.36</td>
<td>0.01</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>$M_5 = \text{INGARCH}(2,2)$</td>
<td>0.00</td>
<td>1.25</td>
<td>2.84</td>
<td>0.09</td>
<td>87.34</td>
<td>0.02</td>
<td>7.32</td>
<td>1.02</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td>$LM_1 = \text{INGARCH}(1,1)$</td>
<td>0.14</td>
<td>0.18</td>
<td>0.16</td>
<td>0.09</td>
<td>0.07</td>
<td>97.38</td>
<td>0.94</td>
<td>0.43</td>
<td>0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>$LM_2 = \text{INGARCH}(1,2)$</td>
<td>0.08</td>
<td>0.37</td>
<td>0.26</td>
<td>0.14</td>
<td>0.01</td>
<td>3.01</td>
<td>90.72</td>
<td>1.72</td>
<td>2.34</td>
<td>1.35</td>
</tr>
<tr>
<td>$LM_3 = \text{INGARCH}(2,1)$</td>
<td>0.02</td>
<td>0.10</td>
<td>0.08</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>2.35</td>
<td>87.76</td>
<td>9.32</td>
<td>0.37</td>
</tr>
<tr>
<td>$LM_4 = \text{INGARCH}(2,2)$</td>
<td>0.10</td>
<td>0.04</td>
<td>0.29</td>
<td>0.20</td>
<td>0.00</td>
<td>0.28</td>
<td>8.72</td>
<td>0.41</td>
<td>84.64</td>
<td>5.32</td>
</tr>
<tr>
<td>$LM_5 = \text{INGARCH}(2,3)$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.12</td>
<td>0.05</td>
<td>0.42</td>
<td>1.07</td>
<td>8.28</td>
<td>89.02</td>
</tr>
<tr>
<td>IN2=INAR(2)</td>
<td>6.28</td>
<td>37.48</td>
<td>4.72</td>
<td>12.29</td>
<td>1.24</td>
<td>22.12</td>
<td>12.28</td>
<td>2.71</td>
<td>2.98</td>
<td>0.10</td>
</tr>
<tr>
<td>INexp=Exponential INAR(1)</td>
<td>0.35</td>
<td>40.16</td>
<td>3.82</td>
<td>11.96</td>
<td>2.53</td>
<td>7.60</td>
<td>24.72</td>
<td>2.91</td>
<td>5.21</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Table 2.1: Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size $n = 100$, $M_i$ is the class of linear and $LM_i$ for log-linear INGARCH models accordingly.
CHAPTER 2. BAYESIAN MODELS SELECTION FOR INGARCH MODELS

<table>
<thead>
<tr>
<th>Model</th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$M_3$</th>
<th>$M_4$</th>
<th>$M_5$</th>
<th>$LM_1$</th>
<th>$LM_2$</th>
<th>$LM_3$</th>
<th>$LM_4$</th>
<th>$LM_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$ = INGARCH(1)</td>
<td>91.50</td>
<td>4.18</td>
<td>0.19</td>
<td>2.16</td>
<td>0.03</td>
<td>1.75</td>
<td>0.10</td>
<td>0.06</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$M_2$ = INGARCH(1,1)</td>
<td>0.10</td>
<td>78.73</td>
<td>0.39</td>
<td>18.91</td>
<td>0.15</td>
<td>0.79</td>
<td>0.45</td>
<td>0.21</td>
<td>0.21</td>
<td>0.07</td>
</tr>
<tr>
<td>$M_3$ = INGARCH(1,2)</td>
<td>0.01</td>
<td>0.36</td>
<td>98.27</td>
<td>0.09</td>
<td>0.22</td>
<td>0.01</td>
<td>0.08</td>
<td>0.36</td>
<td>0.05</td>
<td>0.36</td>
</tr>
<tr>
<td>$M_4$ = INGARCH(2,1)</td>
<td>0.00</td>
<td>2.82</td>
<td>0.60</td>
<td>92.80</td>
<td>0.47</td>
<td>2.50</td>
<td>0.42</td>
<td>0.07</td>
<td>0.26</td>
<td>0.05</td>
</tr>
<tr>
<td>$M_5$ = INGARCH(2,2)</td>
<td>0.00</td>
<td>1.19</td>
<td>3.22</td>
<td>82.98</td>
<td>0.26</td>
<td>10.24</td>
<td>1.08</td>
<td>0.18</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>$LM_1$ = INGARCH(1)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>99.05</td>
<td>0.41</td>
<td>0.28</td>
<td>0.15</td>
<td>0.11</td>
</tr>
<tr>
<td>$LM_2$ = INGARCH(1,1)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>2.81</td>
<td>91.23</td>
<td>2.19</td>
<td>2.77</td>
<td>0.97</td>
</tr>
<tr>
<td>$LM_3$ = INGARCH(1,2)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.92</td>
<td>89.13</td>
<td>7.63</td>
<td>1.29</td>
</tr>
<tr>
<td>$LM_4$ = INGARCH(2,1)</td>
<td>0.03</td>
<td>0.09</td>
<td>0.01</td>
<td>0.04</td>
<td>0.01</td>
<td>0.05</td>
<td>9.46</td>
<td>8.17</td>
<td>72.88</td>
<td>9.27</td>
</tr>
<tr>
<td>$LM_5$ = INGARCH(2,2)</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.77</td>
<td>1.67</td>
<td>9.95</td>
<td>87.54</td>
</tr>
<tr>
<td>$IN_2$ = INAR(2)</td>
<td>4.67</td>
<td>31.96</td>
<td>4.20</td>
<td>13.19</td>
<td>1.42</td>
<td>25.24</td>
<td>14.30</td>
<td>1.71</td>
<td>3.12</td>
<td>0.19</td>
</tr>
<tr>
<td>$IN_{exp}$ = Exponential INAR(1)</td>
<td>0.15</td>
<td>41.50</td>
<td>3.10</td>
<td>13.01</td>
<td>1.32</td>
<td>8.05</td>
<td>26.81</td>
<td>1.33</td>
<td>4.54</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Table 2.2: Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size n = 200, $M_i$ is the class of linear and $LM_i$ for log-linear INGARCH models accordingly.

<table>
<thead>
<tr>
<th>Model</th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$M_3$</th>
<th>$M_4$</th>
<th>$M_5$</th>
<th>$LM_1$</th>
<th>$LM_2$</th>
<th>$LM_3$</th>
<th>$LM_4$</th>
<th>$LM_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$ = INGARCH(1)</td>
<td>95.83</td>
<td>1.87</td>
<td>0.21</td>
<td>1.12</td>
<td>0.03</td>
<td>0.80</td>
<td>0.08</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$M_2$ = INGARCH(1,1)</td>
<td>0.07</td>
<td>90.79</td>
<td>0.32</td>
<td>5.03</td>
<td>0.09</td>
<td>2.72</td>
<td>0.55</td>
<td>0.24</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td>$M_3$ = INGARCH(1,2)</td>
<td>0.01</td>
<td>0.32</td>
<td>99.03</td>
<td>0.08</td>
<td>0.11</td>
<td>0.06</td>
<td>0.05</td>
<td>0.18</td>
<td>0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>$M_4$ = INGARCH(2,1)</td>
<td>0.03</td>
<td>2.13</td>
<td>0.42</td>
<td>95.62</td>
<td>0.27</td>
<td>1.04</td>
<td>0.79</td>
<td>0.13</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>$M_5$ = INGARCH(2,2)</td>
<td>0.06</td>
<td>0.89</td>
<td>1.72</td>
<td>94.07</td>
<td>0.10</td>
<td>2.32</td>
<td>0.43</td>
<td>0.09</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>$LM_1$ = INGARCH(1)</td>
<td>0.01</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>99.36</td>
<td>0.25</td>
<td>0.12</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>$LM_2$ = INGARCH(1,1)</td>
<td>0.01</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>1.73</td>
<td>96.46</td>
<td>0.34</td>
<td>1.12</td>
<td>0.24</td>
</tr>
<tr>
<td>$LM_3$ = INGARCH(1,2)</td>
<td>0.00</td>
<td>0.13</td>
<td>0.02</td>
<td>0.09</td>
<td>0.01</td>
<td>0.04</td>
<td>95.72</td>
<td>3.42</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>$LM_4$ = INGARCH(2,1)</td>
<td>0.08</td>
<td>0.97</td>
<td>0.11</td>
<td>0.17</td>
<td>0.14</td>
<td>0.24</td>
<td>3.98</td>
<td>1.14</td>
<td>91.34</td>
<td>1.83</td>
</tr>
<tr>
<td>$LM_5$ = INGARCH(2,2)</td>
<td>0.02</td>
<td>0.04</td>
<td>0.11</td>
<td>0.08</td>
<td>0.28</td>
<td>0.13</td>
<td>0.52</td>
<td>1.07</td>
<td>4.32</td>
<td>93.43</td>
</tr>
<tr>
<td>$IN_2$ = INAR(2)</td>
<td>4.67</td>
<td>31.96</td>
<td>4.20</td>
<td>13.19</td>
<td>1.42</td>
<td>25.24</td>
<td>14.30</td>
<td>1.71</td>
<td>3.12</td>
<td>0.19</td>
</tr>
<tr>
<td>$IN_{exp}$ = Exponential INAR(1)</td>
<td>0.15</td>
<td>41.50</td>
<td>3.10</td>
<td>13.01</td>
<td>1.32</td>
<td>8.05</td>
<td>26.81</td>
<td>1.33</td>
<td>4.54</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Table 2.3: Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size n = 500, $M_i$ is the class of linear and $LM_i$ for log-linear INGARCH models accordingly.

2.4.2 Applications

Polio data

The data consist of monthly counts of poliomyelitis cases in the United States from 1970 to 1983 (168 observations) reported by the Centres for Disease Control and discussed in Zeger (1988) among others. Figure (2.1) presents the original series and the autocorrelation function of the series. In trans-dimensional method we consider jumps between five linear and five log-linear INGARCH model.

Table (2.4) presents estimators for each of five linear and log-linear INGARCH($p, q$) models. The proposal densities are normal distributions where the value for the variance is very small and was selected to achieve acceptance rate between 25% and 30%.
namely it has a value such as $0.01 < \sigma^2 < 0.025$ for the parameters $\alpha_1, \ldots, \alpha_p, b_1, \ldots, b_q$ and 0.04 for the parameter $d_1$ in linear case and 0.01 in log-linear case accordingly. The chain was run for 10000 to 20000 iterations depending on the model. A burn-in part of 2500 iterations was discarded and the resulting samples were checked for convergence by using the test proposed by Geweke (1992). Z-scores are presented in Table (2.5) and indicate that the convergence has been achieved.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\textbf{Model} & \textbf{parameters} & \textbf{linear} & & & & & \textbf{log-linear} & & & & & \\
\hline
\textbf{Parameters} & $d_1$ & $\alpha_1$ & $b_1$ & $\alpha_2$ & $b_2$ & & $d_1$ & $\alpha_1$ & $b_1$ & $\alpha_2$ & $b_2$ & \\
\hline
INGARCH(0,1) & mean & 0.86 & - & 0.37 & - & - & 0.04 & - & 0.47 & - & - & \\
& sd & 0.09 & - & 0.07 & - & - & 0.03 & - & 0.07 & - & - & \\
\hline
INGARCH(1,1) & mean & 0.62 & 0.20 & 0.35 & - & - & 0.04 & 0.04 & 0.44 & - & - & \\
& sd & 0.15 & 0.12 & 0.07 & - & - & 0.03 & 0.15 & 0.08 & - & - & \\
\hline
INGARCH(1,2) & mean & 0.67 & 0.11 & 0.34 & - & 0.08 & 0.05 & -0.47 & 0.39 & - & 0.35 & \\
& sd & 0.12 & 0.09 & 0.07 & - & 0.06 & 0.05 & 0.32 & 0.08 & - & 0.19 & \\
\hline
INGARCH(2,1) & mean & 0.54 & 0.19 & 0.33 & 0.09 & - & 0.07 & 0.06 & 0.54 & -0.52 & - & \\
& sd & 0.15 & 0.11 & 0.07 & 0.08 & - & 0.07 & 0.14 & 0.08 & 0.11 & - & \\
\hline
INGARCH(2,2) & mean & 0.56 & 0.11 & 0.31 & 0.09 & 0.08 & 0.08 & -0.49 & 0.47 & -0.54 & 0.42 & \\
& sd & 0.14 & 0.09 & 0.07 & 0.05 & 0.08 & 0.08 & 0.22 & 0.09 & 0.12 & 0.16 & \\
\hline
\end{tabular}
\caption{Estimated Posterior Means and standard deviations (sd) for the competing models.}
\end{table}

For the models where parameters satisfy stationarity conditions we apply trans-dimensional MCMC method of Lodewyckx et al. (2011) and posterior probabilities are presented in Table (2.6). The INGARCH(1,1) model is the one mostly visited which indicates that this is the selected model. Note that the 4 best models are of the linear type while the log-linear models have much smaller posterior probabilities. Bayes factors are also reported in Table (2.6). They have been calculated with

Figure 2.1: Series of polio dataset (left) and its ACF (right)
CHAPTER 2. BAYESIAN MODELS SELECTION FOR INGARCH MODELS

Table 2.5: Geweke’s convergence z-scores for the parameters of five linear and five log-linear INGARCH models

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Linear</th>
<th>Log-linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>INGARCH(0,1)</td>
<td>$d_1$</td>
<td>0.454</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>0.128</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>INGARCH(1,1)</td>
<td>$d_1$</td>
<td>-0.677</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-</td>
<td>-0.842</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>INGARCH(1,2)</td>
<td>$d_1$</td>
<td>0.182</td>
<td>-1.197</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-</td>
<td>-0.844</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>-1.178</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-</td>
<td>-0.844</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-</td>
<td>-0.844</td>
</tr>
<tr>
<td>INGARCH(2,1)</td>
<td>$d_1$</td>
<td>1.644</td>
<td>0.642</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-</td>
<td>-1.281</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>-</td>
<td>-0.848</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-</td>
<td>-0.848</td>
</tr>
<tr>
<td>INGARCH(2,2)</td>
<td>$d_1$</td>
<td>1.724</td>
<td>-0.849</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-</td>
<td>-0.404</td>
</tr>
<tr>
<td></td>
<td>$b_1$</td>
<td>-</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-</td>
<td>-0.733</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>-</td>
<td>-0.825</td>
</tr>
</tbody>
</table>

respect the last (and more complicated) model. They also support simple models.

Table 2.6: Posterior probabilities and Bayes factor for the competing models

<table>
<thead>
<tr>
<th>Model</th>
<th>Posterior probability</th>
<th>Bayes Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>INGARCH(1,1)</td>
<td>0.4564</td>
<td>3245.066</td>
</tr>
<tr>
<td>INGARCH(0,1)</td>
<td>0.3741</td>
<td>2659.801</td>
</tr>
<tr>
<td>INGARCH(1,2)</td>
<td>0.0641</td>
<td>455.838</td>
</tr>
<tr>
<td>INGARCH(2,1)</td>
<td>0.0588</td>
<td>418.188</td>
</tr>
<tr>
<td>LINGARCH(0,1)</td>
<td>0.0341</td>
<td>242.644</td>
</tr>
<tr>
<td>INGARCH(2,2)</td>
<td>0.0085</td>
<td>60.462</td>
</tr>
<tr>
<td>INGARCH(1,2)</td>
<td>0.0023</td>
<td>16.199</td>
</tr>
<tr>
<td>LINGARCH(2,1)</td>
<td>0.0016</td>
<td>11.527</td>
</tr>
<tr>
<td>LINGARCH(2,2)</td>
<td>0.0001</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Finally, considering that we have posterior model probabilities for all models, we apply a Bayesian model-averaging (BMA) procedure. Based on posterior model probabilities, as estimated via trans-dimensional MCMC, we derive the density of predictive volatilities. Consequently we calculate $p(\lambda_{T+1} | m, y)$, i.e. the mean of the next unseen observation, weighted at its posterior model probability. We present densities for the four linear models with the higher posterior model probabilities in Figure (2.3). As expected the BMA estimate summarizes the individual densities.

**Campylobacterosis data**

The data (e.g. Liboschik et al., 2015) consist of number of campylobacterosis cases in the North of Québec in Canada. Figure (2.2) presents the original series and the autocorrelation function of the series. Similarly, as for the polio dataset, we present parameters estimation and Z-scores concerning and checking the convergence. Considering same conditions of stationarity as for the polio dataset we present in Table 7 results after the trans-dimensional MCMC algorithm. In this case we observe that
the 4 best models are not only linear but also they have a complicated structure. Furthermore the simple model INGARCH(1,0) is the less preferable.

Finally, we have again calculated \( p(\lambda_{T+1} \mid y) \) for the next unseen observation, weighted at its posterior model probability, depicted in Figure (2.4). As the sample size becomes large, the probability where we generate the data converge to 1. This is evident that our prior distributions are appropriate for parameters. As we can see from the first implementation the most visited model is the \( \text{INGARCH}(1,1) \) model instead of the simplest \( \text{INGARCH}(0,1) \) model. On the contrary at the second application the greatest probability is gotten in \( \text{INGARCH}(2,1) \) model that is the expected considering the results presented at table 2.9. Larger data values

![Figure 2.2: Series of campylobacterosis dataset (left) and its ACF (right)](image-url)
CHAPTER 2. BAYESIAN MODELS SELECTION FOR INGARCH MODELS

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>Linear</th>
<th>Log-linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>INGARCH(0,1)</td>
<td>0.037</td>
<td>-</td>
<td>-0.099</td>
</tr>
<tr>
<td>INGARCH(1,1)</td>
<td>0.592</td>
<td>-0.990</td>
<td>1.318</td>
</tr>
<tr>
<td>INGARCH(1,2)</td>
<td>-0.657</td>
<td>0.544</td>
<td>-0.335</td>
</tr>
<tr>
<td>INGARCH(2,1)</td>
<td>-0.373</td>
<td>-1.256</td>
<td>-1.607</td>
</tr>
<tr>
<td>INGARCH(2,2)</td>
<td>-0.823</td>
<td>0.431</td>
<td>0.745</td>
</tr>
</tbody>
</table>

Table 2.8: Geweke’s convergence z-scores for the parameters of five linear and five log-linear INGARCH models

<table>
<thead>
<tr>
<th>Model</th>
<th>Posterior probability</th>
<th>Bayes Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>INGARCH(2,1)</td>
<td>0.5595</td>
<td>576.0052</td>
</tr>
<tr>
<td>INGARCH(1,1)</td>
<td>0.2784</td>
<td>286.5797</td>
</tr>
<tr>
<td>LINGARCH(1,1)</td>
<td>0.0515</td>
<td>52.9751</td>
</tr>
<tr>
<td>INGARCH(2,2)</td>
<td>0.0339</td>
<td>34.8693</td>
</tr>
<tr>
<td>LINGARCH(1,2)</td>
<td>0.0222</td>
<td>22.8036</td>
</tr>
<tr>
<td>LINGARCH(0,1)</td>
<td>0.0177</td>
<td>18.2094</td>
</tr>
<tr>
<td>INGARCH(1,2)</td>
<td>0.0170</td>
<td>17.5422</td>
</tr>
<tr>
<td>LINGARCH(2,1)</td>
<td>0.0125</td>
<td>12.9080</td>
</tr>
<tr>
<td>LINGARCH(2,2)</td>
<td>0.0064</td>
<td>6.5652</td>
</tr>
<tr>
<td>INGARCH(0,1)</td>
<td>0.0009</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Table 2.9: Posterior probabilities and Bayes factor for the competing models

generally provide more information and reduce the uncertainty associated with parameter estimation, leading to greater precision. As a result, the posterior model probabilities become more precise and reliable, reflecting the increased confidence in the model’s fit to the data.

In this section a bayesian model selection method has been discussed for the selection of the appropriate model among linear and log-linear INGARCH models. The more complicated method, the more flexible to use. Reversible Jump MCMC is difficult to construct assuming that appropriate proposal distributions must be defined for the variables of the new model. In contrary while the other two methods proposed by Carlin and Chib (1995) and Dellaportas et al. (2002) have a tolerant computational cost, they are less complicated at construction. Based on two criteria for model consistency, we observed that the average of posterior model probability is closed to one for the model that data are generated without any previous reference to this. Moreover to our simulation study two models different from INGARCH type have been also considered in order to examine algorithm’s efficacy.

Two datasets are taken into account to examine which of the proposed five linear
2.4. METHODS FOR BAYESIAN MODEL SELECTION

and five log-linear INGARCH models better fits the data. In both cases linear INGARCH models have been selected, indicating that less complicated models are preferable for studies. Finally Bayesian Model Averaging procedure has been also applied to get densities of predictive volatilities. The method of Carlin and Chib (1995) provides flexibility with the construction of parameter space and the way that method decide to jump from one model to another. For those reasons the above method is very useful both in cases of nested and non-nested models and it can be generalized when we have processes of different types such as homoskedastic and heteroskedastic.

Figure 2.3: Posterior density for $\lambda_{T+1}$ for polio data
Figure 2.4: Posterior density for $\lambda_{T+1}$ for Campylobacterosis data
Chapter 3

A novel multivariate Sarmanov distribution

3.1 Sarmanov distribution & FGM copula

The study of copulas in the domain of statistics plays an important role for the construction of multivariate distributions. Bivariate distributions could be a starting point for the generalization in more than three dimensions. Considering different structures, copula is a way for the calculation and modeling of the dependence. Sklar in 1959 proposed a theorem which plays an important role between multivariate distributions and their univariate margins. Considering inverse transformation method:

Theorem 1. Let $H$ be a joint distribution function with margins $F, G$. Then there exist a copula $C$ such for all $x_1, x_2$ in $\mathbb{R}$

$$H(x_1, x_2) = C(F(x_1), G(x_2); \omega).$$ (3.1)

If $F,G$ are continuous, then $C$ is unique otherwise $C$ is uniquely determined on Range$F \times$ Range$G$. Conversely if $C$ is a copula and $F, G$ are distribution functions, then the function $H$ defined by (3.1) is a joint distribution function with margins $F,G$.

The main assumption is that the marginals must be monotone functions. Some further properties for copulas proposed by Nelsen (2007):

1. For every $u, v \in [0, 1]$ $C(u, 0) = C(0, v) = 0$ and $C(u, 1) = u, C(1, v) = v$
2. For every $u_1, u_2, v_1, v_2 \in [0, 1]$ such that $u_1 \leq u_2$ $v_1 \leq v_2$

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0.$$
The first condition ensures that $U, V$ are uniform marginals and the second referred as rectangular inequality $P(u_1 \leq U \leq u_2, v_1 \leq V \leq v_2) \geq 0$. Based on the above theorem there are many studies provided bivariate and multivariate parametric continuous and discrete distributions. Furthermore copulas is a way to study scale-free measures of dependence. There is flourishing literature about different named structures of copulas. Archimedean copulas firstly appeared in study of probabilistic metric spaces Schweizer (1991). Methods for constructions of copulas and tail dependence presented on a review of Nelsen (2007). There are numerous studies for bivariate distributions with some restrictive assumptions such as the range of dependence or the form of the cdf. At many studies alternative constructions of copulas have been proposed in order to provide a wide range at values of parameters of dependence. Furthermore those values depend on copulas function each time. The bounds for dependence based on Fréchet-Hoeffding inequality

$$\max\{u + v - 1, 0\} \leq C(u, v) \leq \min\{u, v\}.$$  

The left side of inequality proves negative dependence and the right strong positive dependence. Cambanis (1977) studied bounds depending on the above inequality for a specified family of copula.

Since we have a form of bivariate distribution via a copula, it will be reasonable to measure dependence and association between the variables. Moreover the opportunity to capture tail dependence is offered via copulas. Two measures of dependence namely Spearman’s $\rho$ and Kendall’s tau also used from many researchers for the above reason. In each case we examine the probability of concordance minus the probability of discordance for the pairs $(x_i, y_i)$ and $(x_j, y_j)$, where those two pairs are concordant when $(X_i - X_j)(Y_i - Y_j) \geq 0$ and discordant when $(X_i - X_j)(Y_i - Y_j) \leq 0$. Under the assumption that these are two vectors of independent variables $(X_1, Y_1)$ and $(X_2, Y_2)$ with common margins $F$ and $G$, Kendall’s tau defined as

$$\tau_C = 4 \int_1 \int_1 C(u, v)dC(u, v) - 1.$$  

For Spearman’s correlation examination of concordance and discordance between independent pairs of variables $(X_i, Y_i), i = 1, 2, 3$ defined as

$$\rho_C = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0])$$

$$= 12 \int_0^1 \int_0^1 uvdC(u, v) - 3,$$  

(3.2)
where $X_2, Y_3$ are independent with marginals $F$ and $G$ respectively. Pearson correlation is often used as linear dependence measure and for this reason it is not appropriate for copula dependence measure. Although the above two measures of association take different values, depending on the family of copula each time, there is a relationship between them and measures of dependence.

**Definition 7.** Let $X$ and $Y$ be two random variables. $X$ and $Y$ are positive quadrant dependent (PQD) if for all $(x_1, x_2) \in \mathbb{R}^2$

$$P[X_1 \leq x_1, X_2 \leq x_2] \geq P[X_1 \leq x_1]P[X_2 \leq x_2],$$

or equivalently

$$P[X_1 > x_1, X_2 > x_2] \geq P[X_1 > x_1]P[X_2 > x_2].$$

One important consequence for the above measures of association for continuous positively quadrant dependent variables is the following theorem

**Theorem 2.** Let $X$ and $Y$ are two continuous random variables with joint distribution function $H$ and marginals $F$ and $G$ respectively and copula $C$. If $X$ and $Y$ are PQD, then

$$3\tau_C \geq \rho_C \geq 0.$$

Furthermore with copulas the tail monotonicity is deterministic for positive quadrant dependence.

**Definition 8.** Let $X$ and $Y$ be two random variables,

1. $X_2$ is left tail decreasing in $X_1$ if $P[X_2 \leq x_2, X_1 \leq x_1]$ is a non-increasing function of $x_1$ for all $x_2$.

2. $X_1$ is left tail decreasing in $X_2$ if $P[X_1 \leq x_1, X_2 \leq x_2]$ is a non-increasing function of $x_2$ for all $x_1$.

3. $X_2$ is right tail increasing in $X_1$ if $P[X_2 > x_2, X_1 > x_1]$ is a non-decreasing function of $x_1$ for all $x_2$.

4. $X_1$ is right tail increasing in $X_2$ if $P[X_1 > x_1, X_2 > x_2]$ is a non-decreasing function of $x_2$ for all $x_1$. 
Chapter 5 in Nelsen (2007) provides more details about tail monotonicity of continuous and discrete random variables. There are numerous families of copulas constructed with different methods. Some of those methods are geometric methods, algebraic methods and the method of inversion. Nelsen (2007) mentioned as geometric methods ordinal sums and singular copulas with prescribed support. We briefly discuss about those methods.

**Definition 9.** Assuming a finite index set $K$ and $J_K$ a partition of $\mathbb{I}$ with $J_K \subseteq [a_k, b_k]$ for $k \in K$ and $\{C_k\}_{k \in K}$ a collection of copulas the ordinal sum defined by:

$$ C(u, v) = \begin{cases} 
  a_k + (b_k - a_k)C \left( \frac{u-a_k}{b_k-a_k}, \frac{v-a_k}{b_k-a_k} \right), & \text{if } (u, v) \in J_k^2, \\
  M(u, v), & \text{if } (u, v) \notin J_k^2. 
\end{cases} \quad (3.3) $$

Furthermore three examples have been also proposed in order to construct singular copulas with support consists of line segments or arcs of circles. As algebraic methods Nelsen (2007) referred Plackett distributions and Ali-Mikhail-Haq distributions. The idea for the construction of the first family based on contigency tables. Considering two continuous random variables $X, Y$ with marginals $F, G$ respectively and their joint distribution $H$, column variables correspond to the events "$X \leq x$" and "$X > x$" similarly for the row variable $Y$. Thus we obtain

<table>
<thead>
<tr>
<th>$X_1 \leq x_1$</th>
<th>$X_2 \leq x_2$</th>
<th>$X_1 &gt; x_1$</th>
<th>$X_2 &gt; x_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(x_1, x_2)$</td>
<td>$F(x_1) - H(x_1, x_2)$</td>
<td>$G(x_2) - H(x_1, x_2)$</td>
<td>$1 - F(x_1) - G(x_2) + H(x_1, x_2)$</td>
</tr>
</tbody>
</table>

$$ \theta = \frac{H(x_1, x_2) \left[ 1 - F(x_1) - G(x_2) + H(x_1, x_2) \right]}{[F(x_1) - H(x_1, x_2)] [G(x_2) - H(x_1, x_2)]} \quad (3.4) $$

and using probability transforms $F(x_1) = u, G(x_2) = v$ and Sklar’s theorem

$$ \theta = \frac{C(u, v) \left[ 1 - u - v + C(u, v) \right]}{[u - C(u, v)] [v - C(u, v)]}. $$

Second family of distributions constructed based on Poisson process. In the univariate case the odds for survival defined as: $\frac{F(x)}{F(x_1)} = \frac{1 - F(x_1)}{F(x_1)}$. Analogously in bivariate case survival odds ratio given by:

$$ \frac{1 - H(x_1, x_2)}{H(x_1, x_2)}. $$

For example Gumbel’s bivariate logistic distribution defined considering two exponential distributions and the above function. In this case the odds ratio given as
the sum of the survival odds ratio for variables $X$ and $Y$. Furthermore considering two independent random variables $X$, $Y$ and expressing each one as $F(x) = \left(1 + \left[\frac{(1-F(x_1))}{F(x_1)}\right]\right)^{-1}$ then

$$
\frac{1 - H(x_1, x_2)}{H(x_1, x_2)} = \frac{1 - F(x_1)}{F(x_1)} + \frac{1 - G(x_2)}{G(x_2)} + \frac{1 - F(x_1)}{F(x_1)} \frac{1 - G(x_2)}{G(x_2)}. \tag{3.5}
$$

Ali-Mikhail-Haq based on (3.5) proposed a bivariate distribution given by:

$$
\frac{1 - H(x_1, x_2)}{H(x_1, x_2)} = \frac{1 - F(x_1)}{F(x_1)} + \frac{1 - G(x_2)}{G(x_2)} + (1 - \theta) \frac{1 - F(x_1)}{F(x_1)} \frac{1 - G(x_2)}{G(x_2)},
$$

where $\theta$ lies in the interval $[-1, 1]$. The third and mostly used method is this of inversion assuming Sklar’s theorem and considering two marginal distributions $F, G$ then $C(u, v) = H(F^{-1}(u), G^{-1}(v))$.

### 3.2 Families of Copulas

In this section we provide a brief overview of the most commonly used families of parametric copulas. Based on Sklar’s theorem one can construct bivariate one parameter or two parameter families of copulas. The first prevalent family is the family of Archimedean copulas with a construction based on Laplace transform appeared firstly in study of probabilistic metric spaces Schweizer (1991). A CDF $F$ then it could be expressed via the laplace transform of an other one CDF $G$ as

$$
F(X) = \int_0^\infty G^a(x) dM(a) = \phi(-\log G(x)).
$$

Respectively in bivariate case

$$
C(F(x_1), G(x_2)) = \phi(-\log G_1 - \log G_2) = \phi(\phi^{-1}(F_1) + \phi^{-1}(F_2)),
$$

or after transformation

$$
C(u, v) = \phi \left(\phi^{-1}(u) + \phi^{-1}(v)\right),
$$

where $\phi : [0, 1] \rightarrow [0, 1]$ are continuous, strictly decreasing and convex functions with $\phi(1) = 0$. This family has an appeal at many different scientific fields because it does not have many difficulties. Obviously this copula has a closed form if both $\phi$ and $\phi^{-1}$ have closed forms and parameter of dependence appears in Laplace transform. Numerous copulas have been constructed with different Laplace transforms as generators. For example Gumbel (1960b) proposed a new copula using as generator
Laplace transform of Exponential distribution where parameter of dependence lies in the interval $[-1, 1]$, namely the cfd is given by:

$$C(u, v) = e^{-(((-\log u)^\theta + (-\log v)^\theta)^{1/\theta})}, \ \phi_\theta(x) = (-\log x)^\theta.$$  

Furthermore another discussed family is this of Frank (1979) which first appeared in a non-statistical context. The reason is that this family is the only one which satisfies radial symmetry. Details and statistical properties have been also provided by Nelsen (1986); Genest (1987). The Frank copula has cdf given by:

$$C(u, v) = -\frac{1}{\theta} \log \left( 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right), \ \phi_\theta(x) = -\log \frac{e^{-\theta x} - 1}{e^{-\theta} - 1}.$$  

At Nelsen (2007) a variety of families of Archimedean copulas is concentrated with a little discussion about some of them and many scatterplots are also provided in order to highlight the importance of parameter $\theta$.

### 3.3 Alternative construction of FGM copula

Another family of copulas with discernible effects is Sarmanov family of bivariate distributions proposed firstly by Sarmanov (1966). Assume two univariate probability mass functions (p.m.f.) or probability density functions (p.d.f.) with supports defined on $A_1 \subseteq \mathbb{R}$, $A_2 \subseteq \mathbb{R}$ respectively. Additionally assume $q_i(x_i), i = 1, 2$ two bounded non constant functions where

$$\sum_{x_i = -\infty}^{\infty} q_i(x_i)f_i(x_i) = 0 \text{ or } \int_{x_i = -\infty}^{\infty} q_i(x_i)f_i(x_i) = 0.$$  

The joint probability mass function (p.m.f.) (or probability density function (p.d.f.)) for the discrete and the continuous case is:

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)[1 + \omega q_1(x_1)q_2(x_2)], \ (3.6)$$  

where the multiplicative factor $[1 + \omega q_1(x_1)q_2(x_2)]$ indicates the deviance of two variables $X_1, X_2$ from independence and parameter $\omega$ is a real number that satisfies the condition

$$1 + \omega q_1(x_1)q_2(x_2) \geq 0. \ (3.7)$$  

In the case where $\omega = 0$, the variables $X_1, X_2$ are independent.

Values of parameter $\omega$ are deterministic for values of correlation coefficient and depend on approximate values of functions $q_i(x_i), \ i = 1, 2$. Ting Lee (1996) proposed countless constructions of mixing functions $q_i(x_i)$ considering each time support of $f_i(x_i)$. When the support of $f_i(x_i)$ is contained in $[0,1]$, then $q_i(x_i) =$
where \( x_i - \mu_i, \ i = 1, 2 \) where \( \mu_i = \sum_{i=1}^{2} \sum_{z=0}^{1} z_i f_i(z_i) \left( \int_{z_i=0}^{1} z_i f_i(z_i) dz_i \right) \). The same construction could be applied when variable’s support is any finite interval. In the case where \( x_i \in \mathbb{R} \) then the proposed mixing functions are \( q_i(x_i) = f_i(x_i) - \sum_{z_i=0}^{\infty} f_i^2(z_i) \left( \int_{0}^{\infty} f_i^2(t_i) dt_i \right) \) and in case where \( x_i \in \mathbb{R}^+ \) then \( q_i(x_i) = \exp(-x_i) - L_i(t) \) or \( q_i(x_i) = 1 - F_i(x_i) \). Ting Lee (1996) expressed \( q_i(x_i) \) by the function below where \(|q_i(x_i)| < 1\) and variable’s support is \([0, 1]\):

\[
q_i(x_i) = 1 - 2F_i(x_i), \ i = 1, 2
\] (3.8)

Based on (3.6), (3.7), considering Sklar’s theorem and inversion method Farlie, Gumbel and Morgerstern at the same period constructed the FGM family given by:

\[
c(u, v) = 1 + \omega(1 - 2u)(1 - 2v), \ u, v \in [0, 1], \quad (3.9)
\]

where \( F(x_1) = u \) and \( F(x_2) = v \) The overall constraint on \( \omega \) is given by:

\[
c(u, v) \geq 0,
\] (3.10)

or equivalently:

\[
1 + \omega (1 - 2u)(1 - 2v) \geq 0,
\]

where finally \( \omega \in [-0.33, 0.33] \) Product moment and correlation coefficient are given by:

\[
E(X_1X_2) = \mu_1\mu_2 + \omega v_1v_2, \quad (3.11a)
\]

\[
\rho = \frac{\omega v_1v_2}{\sigma_1\sigma_2}, \quad (3.11b)
\]

where

\[
v_i = E(f_i(x_i)\phi_i(x_i)), \ i = 1, 2.
\] (3.12)

Based on (3.11b), it is prevalent that values of correlation coefficient depend on values of \( \omega \). Considering (3.11b) and (3.12) values of \( \omega \) lie in the interval \([-1, 1]\). Considering FGM, Farlie (1960) extended the idea of Morgerstern and Gumbel by proposing the following

\[
C(u, v) = uv (1 + \omega A(u)B(v)) ,
\]

where \( A(u), B(v) \to 0 \) as \( u, v \to 1 \) According to the above general family, Farlie (1960) and then Huang and Kotz (1999) proposed alternative construction embed-
ding mixtures of powers, namely

\[ C(u, v) = uv (1 + \omega (1 - u)\gamma (1 - v)\gamma), \quad \gamma \geq 1 \]
\[ C(u, v) = uv (1 + \omega (1 - u^\gamma) (1 - v^\gamma)), \quad \gamma \geq 1/2 \]
\[ C(u, v) = uv (1 + \omega (1 - u)^{\gamma_1} (1 - v)^{\gamma_2}), \quad \gamma_1 \neq \gamma_2 \]
\[ C(u, v) = uv (1 + \omega (1 - u^\gamma)(1 - v^\gamma)), \quad \gamma \geq 0 \]

The first two cases are focused in symmetry. For the third there are unclear properties because it is about a future work. The fourth has been studied by Huang and Kotz (1999) considering a division of the unit square into four quadrants to obtain lower and upper bound for parameter of association \( \omega \). In this case the inequality that provide bounds for parameter \( \omega \) is given by:

\[ 1 + \omega (1 - (1 + \gamma) u^\gamma) (1 - (1 + \gamma) v^\gamma), \]

where finally the range for \( \omega \) is \([-\max\{1,p\}^{-2}, p^{-1}]\)

Bekrizadeh et al. (2012) probed the workings of amendment of correlation values constructing an alternative model as

\[ C(u, v) = uv [1 + \theta (1 - u^a)(1 - v^a)]^n, \]

where \( a > 0 \) and \( n \in \mathbb{N} \). Taking advantage of the fact that \( c(u, v) \geq 0 \), the interval for \( \omega \) is defined as \(-\min\{1, \frac{1}{na}\} \leq \omega \leq \frac{1}{na}\). The above studies have been concentrated(focused) in the case where a bivariate distribution function can be interpreted as a bivariate copula or in other words when \( u, v \) are uniformly distributed.

### 3.4 A proposed alternative FGM copula

In the whole study about association parameter’s interval, numerous applications have been mentioned. Gumbel (1960a) proposed a new bivariate Exponential distribution based on the first construction of FGM copula and Bekrizadeh et al. (2012) studied his proposed FGM copula considering as marginals Beta and Weibull distributions. Assuming the properties studied by Huang and Kotz (1999), Ting Lee (1996) proposed Gamma, Beta, Exponential and Normal bivariate distributions considering each time different construction for functions encompassed in multiplicative factor. The main problems are: (a) there is not always a closed form for the bivariate distribution leading to infeasibility of studying other properties of copulas e.g. parameters estimation and (b) restrictive correlation coefficients values are provided.
3.4. A PROPOSED ALTERNATIVE FGM COPULA

At the case of discrete variables there are many studies focusing not only on the interval of parameter $\omega$ but also on methods for parameters’ estimation. Ting Lee (1996) proposed two distributions considering Binomial and Poisson marginals respectively. Lakshminarayana et al. (1999) and after many years Cui and Zhu (2018) proposed bivariate Poisson distributions. The main drawback is the violation of Frechet-Hoeffding boundaries in order to get better values for dependence parameter range. Piperigou (2009) studied a Poisson and a Geometric bivariate distributions considering probability generating functions in order to increase values for $\omega$ and correlation coefficient $\rho$. For the construction of boundaries for dependence parameter Piperigou (2009) used the theory introduced by Cambanis (1977). Miravete (2009a) based on Ting Lee (1996) built two multivariate count data regression models considering as marginals Double Poisson distributions. Moreover Ting Lee (1996) proposed an extended Sarmanov distribution and Miravete (2009a) based on this extension proposed trivariate Double Poisson and Gamma distributions without studying theoretical properties. Lakshminarayana et al. (1999) provided different general construction for the mixing functions

$$q_i(x_i) = g_i(x_i) - E(g_i(x_i)), \ i = 1, 2,$$

and discussed about the vital problem of the choice of $g_i(x_i)$. Assuming that mixing functions must be bounded functions, damped polynomials have been proposed where as damped polynomials Lakshminarayana et al. (1999) considered any polynomial multiplied by a dampening function $w(x)$ that tends to zero faster than any polynomial tends to infinity as $x$ increases. A prevalent option is $e^{-x}$ where $g_i(x_i) = e^{-x_i}P(x_i)$. The idea of Lakshminarayana et al. (1999) seems to be similar to this of Ting Lee (1996) where under the assumption that $f_i(x_i)$ are defined in the interval $[0, \infty)$. Ting Lee (1996) proposed as mixing functions

$$q_i(x_i) = e^{-x_i} - LT_i(1), \ i = 1, 2, \ \forall \ x_i \geq 0.$$

At this study $g_i(x_i) = e^{-x_i}$ and $LT_i(t) = \sum_{x_i=0}^{\infty} e^{-tx_i}f_i(x_i)$ is the Laplace transform of $f_i$, for $i = 1, 2$. Integrating $x_1x_2$ with respect to (3.6) the product moment and correlation coefficient are given by (3.11a; 3.11b) where

$$v_i = -LT_i'(1) - LT_i(1)\mu_i, \ i = 1, 2.$$

Under the restriction (3.7) $\omega \in [b_1, b_2]$, where

$$b_1 = \frac{-1}{\max \{LT_1(1)LT_2(1), (1 - LT_1(1))(1 - LT_2(1))\}}$$

(3.15)
and

\[ b_2 = \frac{1}{\max\{LT_1(1)(1 - LT_2(1)), (1 - LT_1(1))(LT_2(1))\}} \]  

(3.16)

and based on (3.6) product moment and correlation coefficient are defined as:

\[ E(X_1, X_2) = \mu_1\mu_2 + \omega v_1 v_2, \]

(3.17)

\[ \rho = \frac{\omega v_1 v_2}{\sigma_1 \sigma_2}, \]

(3.18)

where \( v_i, \ i = 1, 2 \) given by (3.14). The drawback with all of those constructions is that the range of correlation is too restrictive. Vernic (2020) and a few years later Piancastelli et al. (2023), based on the previous definition and on the idea that \( q_i(x_i) \) could be expressed as \( u_i(x_i) - E[u_i(X_i)] \), proposed a new function \( u_i(x_i) \) defined as \( u_i(x_i) = e^{-ax_i} \) where \( a \neq 1 \). The proposed function is called Exponential kernel yielding a greater range for correlation coefficient boundaries. In our study based on Sarmanov distribution given by (3.6), on the work of Ting Lee (1996) and generalized the idea of Vernic (2020), we focus firstly on bivariate case, defining new mixing functions as alternative approaches of Laplace transforms given by:

\[ q_i(x_i) = c_i^{-k_i} x_i - \sum_{x_i=0}^{\infty} c_i^{-k_i} x_i f(x_i), \quad i = 1, 2 \]

where

\[ c_i > 1, \ i = 1, 2 \] and \( k_i > 0, \ i = 1, 2 \).

Based on the above structure, functions \( q_i(x_i), \ i = 1, 2 \) are bounded, non constant functions. In discrete case the definition of bivariate distribution has a drawback at the use of finite sums. For the above reason many studies have been done based on Sarmanov distribution and it is alternative construction. Moreover the crucial problem is tracked for the variables with support \( \mathbb{R}^+ \) because there is a restricted number of bounded functions which allow flexibility at values of correlation coefficient. Considering now two discrete variables and by definition of joint distribution a new construction has been developed where \( q_i(x_i) = f_i(x_i) + 2F_x(x_i) - 1 \):
\[ p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2) = P(X_1 \leq x_1, X_2 = x_2) - P(X_1 < x_1, X_2 = x_2) \]
\[ = P(X_1 \leq x_1, X_2 \leq x_2) - P(X_1 \leq x_1, X_2 \leq x_2 - 1) \]
\[ - P(X_1 \leq x_1 - 1, X_2 \leq x_2) + P(X_1 \leq x_1 - 1, X_2 \leq x_2 - 1) \]
\[ = P(x_1, x_2) - P(x_1, x_2 - 1) - P(x_1 - 1, x_2) + P(x_1 - 1, x_2 - 1) \]
\[ = P(x_1)P(x_2)[1 + \omega (1 - P(x_1))(1 - P(x_2))] \]
\[ - P(x_1)P(x_2 - 1)[1 + \omega (1 - P(x_1))(1 - P(x_2 - 1))] \]
\[ - P(x_1 - 1)P(x_2)[1 + \omega (1 - P(x_1 - 1))(1 - P(x_2))] \]
\[ + P(x_1 - 1)P(x_2 - 1)[1 + \omega (1 - P(x_1 - 1))(1 - P(x_2 - 1))] \]
\[ = (P(x_1) - P(x_1 - 1))(P(x_2) - P(x_2 - 1)) \]
\[ [(P(x_1) + P(x_1 - 1) - 1)(P(x_2) + P(x_2 - 1) - 1)] \quad (3.19) \]

and considering that

\[ P(x_1) + P(x_1 - 1) - 1 = P(x_1) - P(x_1 - 1) + 2P(x_1 - 1) - 1 \]
\[ = -[P(x_1) - P(x_1 - 1)] + 2P(x_1) - 1 \]
\[ = -p(x_1) + 2P(x_1) - 1, \]

we obtain finally another construction of multiplicative factor different from those that have been discussed previously. This formula provide an improvement at bounds of correlation coefficient. Barbiero (2019) studied values of correlation coefficient based on the above structure for a discrete Weibull distribution.

\[ p(x_1, x_2) = \prod_{i=1}^{2}[P(x_i) - P(x_i - 1)] \left( 1 + \omega \prod_{i=1}^{2}[p(x_i) - 2P(x_i) + 1] \right). \]

For the construction of bivariate Poisson distribution two general closed forms have been also proposed from Kocherlakota and Kocherlakota (2001) and Lakshminarayana et al. (1999) respectively. The focus is concentrated on the flexibility of the closed form of bivariate distribution and on the construction which allows high values for correlation. At the first general construction, researchers tried to define a bivariate model paying attention on methods for estimation caused of the non-closed
form in conjunction with the second where different functions have been proposed to increase correlation coefficient bounds. Kocherlakota and Kocherlakota (2001) studied a bivariate Poisson distribution based on trivariate reduction method. Assuming that variables $X_i$ follow independent Poisson distributions $P(\lambda_i)$ with parameters $\lambda_i, i = 1, 2, 3$ and the variables $X = X_1 + X_3, Y = X_2 + X_3$ follow jointly a bivariate distribution given by

$$p(x = r, y = s) = \sum_{k=0}^{\min(r,s)} f_1(r - k) f_2(s - k) f_3(k)$$

$$= e^{-\sum_{i=1}^{3} \lambda_i} \sum_{k=0}^{\min(r,s)} \frac{\lambda_1^r \lambda_2^s \lambda_3^k}{(\prod_{i=1}^{2} \lambda_i)^k} (r-k)! (s-k)! k!$$

$$r, s = 0, 1, \ldots, \lambda_i > 0, i = 1, 2, 3.$$ 

Marginally each one of those variables follows a Poisson distribution with mean $E(X) = \lambda_1 + \lambda_3$ and $E(Y) = \lambda_2 + \lambda_3$. Parameter $\lambda_3$ is appropriate for dependence between two variables $X, Y$ while $\text{cov}(X, Y) = \lambda_3$. In case where $\lambda_3 = 0$ two variables are independent and the bivariate Poisson distribution is the product of two independent Poisson distributions. Vernic (1997) studied properties for the above bivariate Poisson distribution and parameters’ estimation using method of moments while Kocherlakota and Kocherlakota (2001) proposed method of maximum likelihood for parameters’ estimation. Karlis and Ntzoufras (2003) proposed an EM algorithm considering the bivariate Poisson model proposed by Kocherlakota and Kocherlakota (2001) and assuming log link regression functions for the mean assuming

$$\log \lambda_{ik} = w_{ik} \beta_k, \quad k = 1, 2, 3, \quad i = 1, \ldots, n$$

where $w_{ik}$ denotes the vector of explanatory variables and $\beta_k$ is the vector of regression coefficients.

Lakshminarayana et al. (1999) studied a bivariate Poisson model based on Sarmanov distribution as we mentioned at the beginning of this chapter. Correlation coefficient bounds in this study violated the Frechet-Hoeffding bounds because additional restriction is necessary. Similar to the above model another model was also proposed by Cui and Zhu (2018). The main difference with the model construct by Lakshminarayana et al. (1999) is the non deterministic and inappropriate way of correlation coefficient bounds’ definition. While Lakshminarayana et al. (1999) proposed that:

$$|\rho| \leq \sqrt{\frac{\lambda_1 \lambda_2 e^{-(1-e^{-1}) (\lambda_1 + \lambda_2)} (1 - e^{-1})^2}{(1 - e^{-(1-e^{-1}) \lambda_1}) (1 - e^{-(1-e^{-1}) \lambda_2})}}.$$
Cui and Zhu (2018) restrict $\rho$ as:

$$|\rho| \leq \min \left\{ \sqrt{\lambda_1 \lambda_2} e^{-(1-e^{-1})(\lambda_1 + \lambda_2)} (1 - e^{-1})^2 \right\}.$$ 

In our study based on (3.4) we define a new bivariate Poisson distribution as:

$$f(x_1, x_2) = \prod_{i=1}^{2} e^{-\lambda_i} \frac{x_i}{x_i!} \left[ 1 + \omega \prod_{i=1}^{2} q_i(x_i) \right],$$

where

$$q_i(x_i) = c_i^{-k_i x_i} - e^{\lambda_i (c_i^{-k_i x_i} - 1)}, \quad c_i \geq 0, k_i \geq 0.$$

Moreover considering (3.19), we can conclude to the same bivariate Poisson distribution proposed above.

$$F(x_1, x_2) = \prod_{i=1}^{2} \left( e^{-\lambda_i} \sum_{t=0}^{X_i} \frac{(\lambda_i c_i^{-k_i})^t}{t!} - e^{\lambda_i (c_i^{-k_i} - 1)} F(x_i) \right),$$

$$F(x_1 - 1, x_2) = \left( e^{-\lambda_1} \sum_{t=0}^{X_1-1} \frac{(\lambda_1 c_1^{-k_1})^t}{t!} - e^{\lambda_1 (c_1^{-k_1} - 1)} F(x_1 - 1) \right)$$

$$\prod_{i=1}^{2} \left( e^{-\lambda_i} \sum_{t=0}^{X_i} \frac{(\lambda_i c_i^{-k_i})^t}{t!} - e^{\lambda_i (c_i^{-k_i} - 1)} F(x_i) \right),$$

$$F(x_1, x_2 - 1) = \left( e^{-\lambda_1} \sum_{t=0}^{X_1} \frac{(\lambda_1 c_1^{-k_1})^t}{t!} - e^{\lambda_1 (c_1^{-k_1} - 1)} F(x_1) \right)$$

$$\prod_{i=1}^{2} \left( e^{-\lambda_i} \sum_{t=0}^{X_i-1} \frac{(\lambda_i c_i^{-k_i})^t}{t!} - e^{\lambda_i (c_i^{-k_i} - 1)} F(x_i) \right),$$

$$F(x_1 - 1, x_2 - 1) = \prod_{i=1}^{2} \left( e^{-\lambda_i} \sum_{t=0}^{X_i-1} \frac{(\lambda_i c_i^{-k_i})^t}{t!} - e^{\lambda_i (c_i^{-k_i} - 1)} F(x_i - 1) \right).$$

Assuming that $f(x_1, x_2) = F(x_1, x_2) - F(x_1 - 1, x_2) - F(x_1, x_2 - 1) + F(x_1 -$
where \( x_1, x_2 \leq 1 \) and substituting the equations above we have:

\[
f(x_1, x_2) = \prod_{i=1}^{2} \left( F(x_i) - F(x_i - 1) \right) e^{\sum_{i=1}^{2} \lambda_i (c_i^{k_i-1})} \left( F(x_2) - F(x_2 - 1) \right)
- e^{\lambda_1 (c_1^{k_1-1}) - \lambda_2} \left( F(x_1) - F(x_1 - 1) \right) \sum_{t=0}^{x_2} \frac{\lambda_2 c_2^{k_2}}{t!} \\
+ e^{\lambda_1 (c_1^{k_1-1}) - \lambda_2} \left( F(x_1) - F(x_1 - 1) \right) \sum_{t=0}^{x_2-1} \frac{\lambda_2 c_2^{k_2}}{t!} \\
- e^{-\lambda_1 + \lambda_2 (c_2^{k_2}-1)} \left( \sum_{t=0}^{X_1} \frac{(\lambda_1 c_1^{k_1})}{t!} - \sum_{t=0}^{X_1-1} \frac{(\lambda_1 c_1^{k_1})}{t!} \right) \left( F(x_2) - F(x_2 - 1) \right) \\
+ e^{-(\lambda_1 + \lambda_2)} \prod_{j=1}^{2} \left( \sum_{t=0}^{X_1} \frac{(\lambda_j c_j^{k_j})}{t!} - \sum_{t=0}^{X_1-1} \frac{(\lambda_j c_j^{k_j})}{t!} \right) .
\]

Taking into account that in discrete case \( f(x_i) = F(x_i) - F(x_i - 1) \), \( i = 1, 2 \) and

\[
f(x_i) = \left( \sum_{t=0}^{X_i} \frac{(\lambda_i c_i^{k_i})}{t!} - \sum_{t=0}^{X_i-1} \frac{(\lambda_i c_i^{k_i})}{t!} \right) , \quad i = 1, 2
\]

for variables \( X_i, \ i = 1, 2 \) respectively we take that

\[
p(x_1, x_2) = \prod_{i=1}^{2} f(x_i) \left( 1 + \omega \prod_{i=1}^{2} \left( c_i^{k_i x_i} - e^{\lambda_i (c_i^{k_i-1})} \right) \right) .
\]

Another way is the variable transformation to change the domain and to bypass the problem especially at the case of \([0, \infty)\). Miravete (2009b) based on the model defined by Lakshminarayana et al. (1999) constructed two bivariate models the first when marginals are distributed according to a double Poisson distribution and the second one when marginals follow a Gamma distribution. At those models there is a closed form for the bivariate cdf. On the contrary in case on Poisson distribution this is not efficient cause of cdf’s construction. For this reason a closed form for a copula is not always possible and methods for parameters estimation are more extensive.

Our new model is flexible not only in case of a discrete variable but also for continuous variables. Calculating cdf of a bivariate distribution and considering \( g_i(x_i) = c_i^{k_i x_i} \) in order to construct a new family of copulas with flexible correlation bounds depend each time on the choice of mixing functions, we have:

\[
C(u, v) = uv \left[ 1 + \omega \left( \sum_{l_1=0}^{F^{-1}(u)} c_1^{-k_1} - uLT_1(f(x_1)) \right) \left( \sum_{l_2=0}^{F^{-1}(v)} c_2^{-k_2} - vLT_2(f(x_2)) \right) \right].
\]
In general this family has the properties of a copula. Considering structure for correlation coefficient and assuming that $q_i(x_i), \ i = 1, 2$ are defined by the mixing functions given above we have that:

$$E(X_1X_2) = \prod_{i=1}^{2} \lambda_i + \omega \prod_{i=1}^{2} \left[ -\frac{1}{\log c_i} q_i'(k_i) - \lambda_i e^{\lambda_i(c_i^{-k_i}-1)} \right]$$

$$= \prod_{i=1}^{2} \lambda_i + \omega \prod_{i=1}^{2} \lambda_i e^{\lambda_i(c_i^{-k_i}-1)(c_i^{-k_i} - 1)}, \ i = 1, 2 \quad (3.20)$$

Under the assumption that $q_i'(k_i)$ are the derivatives of $\exp(\lambda_i(c_i^{-k_i}) - 1), \ i=1,2$ with respect to $k_i, \ i = 1, 2$ or

$$\frac{\partial}{\partial k_i} \sum_{x_i=0}^{\infty} f_i(x_i) c_i^{-k_i x_i} = \sum_{x_i=0}^{\infty} f_i(x_i) \frac{\partial}{\partial k_i} c_i^{-k_i x_i} = -\log c_i \sum_{x_i=0}^{\infty} x_i f_i(x_i) c_i^{-k_i x_i}.$$ 

Combining (3.20) into (3.18) correlation coefficient given by:

$$\rho = \omega \prod_{i=1}^{2} \sqrt{\lambda_i e^{\lambda_i(c_i^{-k_i}-1)(c_i^{-k_i} - 1)}}, \ i = 1, 2. \quad (3.21)$$

The above structure of (3.20) has been calculated based on (3.14). According to the above mixing functions new boundaries are calculated considering restriction given by (3.6) and calculating the limits as $x_i, \ i = 1, 2$ approaches to 0 and $\infty$ and given by:

$$b_1 = \frac{-1}{\max \left\{ (1 - e^{\lambda_1(c_1^{-k_1}-1)}) (1 - e^{\lambda_2(c_2^{-k_2}-1)}), e^{\lambda_1(c_1^{-k_1}-1)} e^{\lambda_2(c_2^{-k_2}-1)} \right\}}, \quad (3.22)$$

$$b_2 = \frac{1}{\max \left\{ e^{\lambda_1(c_1^{-k_1}-1)} (1 - e^{\lambda_2(c_2^{-k_2}-1)}), e^{\lambda_2(c_2^{-k_2}-1)} (1 - e^{\lambda_1(c_1^{-k_1}-1)}) \right\}}, \quad (3.23)$$

Based on the Cambanis (1977) this construction does not violate Frechet-Hoeffding inequality and it offers flexible boundaries for correlation coefficient. Firstly we examine $\rho$’s values taking 3 possible values for $\lambda_i, \ i = 1, 2$ and numerous values for $k_i, \ i = 1, 2$. From the graphs below we observe that the optimal value for $\rho$ is obtained when $\lambda_i = 0.5, \ i = 1, 2$. 


Figure 3.1: Values for correlation with equal values at both $\lambda_i, k_i, i = 1, 2$. Lower boundary (first graph) and upper boundary (second graph).

Then keeping the value for $\lambda_i, i = 1, 2$ and considering different combinations between two parameters $k_1, k_2$ boundaries for correlation coefficient are also presented. From those graphs and by using $c_i = \exp(1), \ i = 1, 2$ we observe that upper boundary for correlation is 0.77 and lower boundary is -0.5. Furthermore approximately when $k_1, k_2$ are equal to 10 correlation coefficient value converge to upper and lower boundaries respectively and it is closed to the same value when $k_1 = k_2, k_1 = k_2/2$ and $k_1, k_2 > 8$.

Figure 3.2: Values for correlation with equal values at both $\lambda_i, k_i, i = 1, 2$. Lower boundary (first graph) and upper boundary (second graph).

Finally a comparison among the models that have been also proposed with the new model is presented via graphs with the appropriate values for $\rho$ lie in different interval each time based on values for parameters $c_i, \lambda_i$ and $k_i, i = 1, 2$. 
3.4. A PROPOSED ALTERNATIVE FGM COPULA

We obtain values for $\rho$ for countless values of $c_i$, $i = 1, 2$. At the graphs below we present correlation coefficient values taking numerous values for parameters $\lambda_i$ and for $k_1 = k_2 = 1$. Considering models proposed by Lakshminarayana et al. (1999), Ting Lee (1996) and only assuming conditions for inequality proposed by Cambanis (1977) the optimal values are given when $c_i = 3.5$.

![Graphs showing correlation coefficient values](image)

**Figure 3.3:** Correlation coefficient values considering numerous values for bases $c_i$ and powers $k_i$, lower boundary(first graph) and upper boundary(second graph)

Under the assumption that correlation coefficient given by (3.18) and assuming Sarmanov distribution, the proposed bivariate Poisson distribution can be written equivalently as:

$$f(x_1, x_2) = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} \frac{[f_i(x'_i) - f_i(x_i)]}{\sqrt{\lambda_i (c_i^{-k_i} - 1)}}, \quad (3.24)$$

where

$$f_i(x'_i) \sim Poisson(\lambda_i \alpha^{-k_i}), \quad i = 1, 2.$$
\[
\begin{align*}
    f(x_1, x_2) &= \prod_{i=1}^{2} f_i(x_i) [1 + \rho \prod_{i=1}^{2} \frac{\sigma_i}{v_i} q_i(x_i)] \\
    &= f_1(x_1) f_2(x_2) + \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1) f_2(x_2) c_1^{-k_1 x_1} c_2^{-k_2 x_2} \\
    &\quad - \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1) f_2(x_2) c_1^{-k_1 x_1} e^{\lambda_2 (c_2^{-k_2} - 1)} \\
    &\quad - \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1) f_2(x_2) e^{\lambda_1 (c_1^{-k_1} - 1)} (c_2^{-k_2} - 1) \\
    &\quad + \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1) f_2(x_2) e^{\lambda_1 (c_1^{-k_1} - 1)} e^{\lambda_2 (c_2^{-k_2} - 1)} \\
    &= f_1(x_1) f_2(x_2) + \rho \frac{f_1(x_1')}{\sqrt{\lambda_1 (c_1^{-k_1} - 1)}} \frac{f_2(x_2')}{\sqrt{\lambda_2 (c_2^{-k_2} - 1)}} \\
    &\quad - \rho \frac{f_1(x_1')}{\sqrt{\lambda_1 (c_1^{-k_1} - 1)}} \frac{f_2(x_2)}{\sqrt{\lambda_2 (c_2^{-k_2} - 1)}} + \rho \frac{f_1(x_1)}{\sqrt{\lambda_1 (c_1^{-k_1} - 1)}} \frac{f_2(x_2)}{\sqrt{\lambda_2 (c_2^{-k_2} - 1)}} \\
    &= f_1(x_1) f_2(x_2) + \rho \frac{[f_1(x_1') - f_1(x_1)][f_2(x_2') - f_2(x_2)]}{\sqrt{\lambda_1 \lambda_2 (c_1^{-k_1} - 1)(c_2^{-k_2} - 1)}}. 
\end{align*}
\]

Based on (3.6) and considering (3.7) the construction of continuous distributions is possible. Miravete (2009b) based on Sarmanov distribution studied not only a Double Poisson distribution in discrete case but also a bivariate Gamma distribution. Assuming that exponential distribution gain a dominant role because it holds for distances in time and more precisely the time between events, it has served as a first approach to a model for life testing. Many researchers developed bivariate exponential distributions without considering Sarmanov distribution for the construction. Gumbel (1960a) and Marshall and Olkin (1967) study constructions for new bivariate exponential distribution. On the work of Gumbel (1960a) bounds for correlation have been proved to be between -0.25 to +0.25 in one model and on the second model correlation takes only negative values between -0.4 to 0. Gumbel (1960a) based on work of Fréchet (1951) constructed a new family of copula named Gumbel copula. Farlie after his investigation to correlation coefficient studied a generalization of bivariate Farlie-Gumbel-Morgerstern distribution proposed by Morgerstern and Gumbel. For a review of various exponential distributions someone can study Balakrishnan (2014). Shubina and Lee (2004) study an alternative construction of bivariate distributions by using a new upper boundary for parameter \( \omega \). This idea could be correct in general case or in case where we have a variable in a bounded interval but not in case of \([0, \infty)\). Ting Lee (1996) proposed different construction of multiplicative factor depend on the variable’s domain. Considering that \( X \) is
continuous and taking values in \([0, \infty)\), we choose as marginal distributions exponential with parameters \(\lambda_1\) and \(\lambda_2\) accordingly. This choice of continuous variable provide that the construction of the proposed mixing function \(q_i(x_i), i = 1, 2\) can be used either with discrete or with continuous variables. At exponential distribution the choice for the bounds for dependence parameter \(\omega\) differ from those that have been already proposed at discrete case in previous chapter. Considering that \(X_i \sim \text{Exp}(\lambda_i) = \lambda_i \exp(-\lambda_i x_i), \lambda_i, x_i > 0\) and based on (3.6), (3.7) and alternative definition for mixing functions given at section 3 we have:

\[
q_i(x_i) = c_i^{-k_i x_i} - \frac{\lambda_i}{\log(e^{\lambda_i c_i})}, \ i = 1, 2.
\] (3.26)

By using (3.18) correlation given by \(\rho = \frac{\omega v_1 v_2}{\sigma_1 \sigma_2}\) where \(v_i = E(q_i(x_i)f_i(x_i))\) or equivalently by

\[
\rho = \omega \prod_{i=1}^{2} \frac{\lambda_i k_i \log c_i}{(\lambda_i + k_i \log c_i)^2}.
\] (3.27)

Bounds for \(\omega\) are calculated assuming that \(1 + \omega q_1(x_1)q_2(x_2)\) must be positive. The construction of boundaries based on Cambanis (1977) and the boundaries defined in section 4 could be given by:

\[
b_1 = \min \left\{ \frac{-1}{\prod_{i=1}^{2} (1 - E_{f_i} c_i^{-k_i x_i})}, \prod_{i=1}^{2} E_{f_i} c_i^{-k_i x_i} \right\}
\]

and

\[
b_2 = \min \left\{ \frac{1}{(1 - E_{f_1} c_1^{-k_1 x_1}) E_{f_2} c_2^{-k_2 x_2}}, \frac{1}{(1 - E_{f_2} c_2^{-k_2 x_2}) E_{f_1} c_1^{-k_1 x_1}} \right\}
\] (3.28)

where \(E_{f_i} c_i^{-k_i x_i} = \int_{0}^{\infty} c_i^{-k_i x_i} \lambda_i \exp(-\lambda_i x_i)\).

In case of exponential distribution equations (3.18, 3.27) can be written as

\[
\rho = \omega \prod_{i=1}^{2} (1 - E_{f_i} c_i^{-k_i x_i}) E_{f_i} c_i^{-k_i x_i}.
\] (3.29)

Taking all possible cases for \(\omega\) based on the above boundaries give by (3.28) and correlation coefficient given by (3.29), we can see that \(\rho\) is always the other choice of the boundary and is always smaller than 1. For this reason we can reconstruct boundaries for \(\omega\) as :

\[
b_1 = \min \left\{ \prod_{i=1}^{2} (1 - E_{f_i} c_i^{-k_i x_i}), \prod_{i=1}^{2} E_{f_i} c_i^{-k_i x_i} \right\}
\] (3.30)
and

\[ b_2 = \frac{1}{\min \left\{ \left(1 - E_{f_1}c_1^{-k_1x_1} \right) E_{f_1}c_1^{-k_1x_1}, E_{f_1}c_1^{-k_1x_1} \left(1 - E_{f_2}c_2^{-k_2x_2} \right) \right\}}. \] (3.31)

Calculation of those boundaries in general case based on the study of Cambanis (1977). The main property that must be satisfied is the inequality of Frechet-Hoeffding bounds for joint distribution \( F(x_1, x_2) \) where \( F(x_1) \) and \( G(x_2) \) are Exponential.

Based on the above equations we observe that \(-0.95 \leq \rho \leq 0.95\) when \( c_i = \exp(1), \ i = 1, 2\). As we can see at the first two graphs in case of \( k_i = \lambda_i \in [0.5, 9] \) and \( \alpha_i \in [0.5, 10] \) as values of parameters \( k_i, \lambda_i \) increase and \( \alpha_i \) decrease \( \rho \) is in \([-0.249, 0.506]\). The color of line change when equation for the calculation of correlation change. As we can see for the lower bound in the case where parameters take the same values simultaneously equation does not change.

At the graphs below we observe the higher upper and lower boundaries for correlation of exponential distribution. As we can see when \( \lambda_1 = \lambda_2 \) and \( k_1 = k_2 \) we observe the higher value for the upper boundary and the opposite when \( \lambda_i = k_j, \ i, j = 1, 2, \ i \neq j \).

![Figure 3.4: Correlation coefficient values for numerous equal values of \( \lambda_i \) and \( k_i \)](image-url)
3.4. A PROPOSED ALTERNATIVE FGM COPULA

The fact that Exponential distribution has a cdf in a closed form provides a bivariate copula in a closed form. Bivariate cdf given by:

\[
F(x_1, x_2) = \prod_{i=1}^{2} (1 - e^{-\lambda_i x_i}) + \omega \prod_{i=1}^{2} \left[ \frac{\lambda_i}{\lambda_i + k_i \log(c_i)} \left(1 - e^{-(\lambda_i + k_i \log(c_i))x_i} \right) - \frac{\lambda_i}{\lambda_i + k_i \log(c_i)} F_i(x_i) \right]
\]

and copula defined as

\[
C(u, v) = uv + \omega \left( \frac{\lambda_1}{\lambda_1 + k_1 \log(c_1)} \left(1 - e^{-(\lambda_1 + k_1 \log(c_1))F_1^{-1}(u)} - u \right) \right) \times \left( \frac{\lambda_2}{\lambda_2 + k_2 \log(c_2)} \left(1 - e^{-(\lambda_2 + k_2 \log(c_2))F_1^{-1}(v)} - v \right) \right).
\]

Bivariate distribution is given by:

\[
f(x_1, x_2) = \prod_{i=1}^{2} f_i(x_i) + \prod_{i=1}^{2} \log(e^{\lambda_i k_i}) \left[f_i(x_i) - f_i(x_i) \right], \tag{3.32}
\]

where

\[
f_i(x_i) = \text{Exp}(\lambda_i + k_i \log c_i), \ i = 1, 2.
\]
where

\[ f(x_1, x_2) = \prod_{i=1}^{2} f_i(x_i)[1 + \rho \prod_{i=1}^{2} \sigma_i q_i(x_i)] \]

\[ = f_1(x_1)f_2(x_2) + \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1)f_2(x_2) c_1^{-k_1 x_1} c_2^{-k_2 x_2} \]

\[ - \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1)f_2(x_2) c_1^{-k_1 x_1} \lambda_2 \log(e^{\lambda_2 c_2^2}) \]

\[ - \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1)f_2(x_2) \lambda_1 \log(e^{\lambda_1 c_1^2}) c_2^{-k_2 x_2} \]

\[ + \rho \frac{\sigma_1 \sigma_2}{v_1 v_2} f_1(x_1)f_2(x_2) \lambda_1 \lambda_2 \log(e^{\lambda_1 c_1^2} e^{\lambda_2 c_2^2}) \]

\[ = f_1(x_1)f_2(x_2) + \rho \frac{f_1(x_1)^{''}}{\sqrt{\lambda_1 (c_1^{-k_1} - 1)}} \frac{f_2(x_2)^{''}}{\sqrt{\lambda_2 (c_2^{-k_2} - 1)}} \]

\[ - \rho \frac{f_1(x_1)^{''}}{-\lambda_1 - 1 - \lambda_2 - 1} f_2(x_2)^{''} \]

\[ - \rho \frac{f_1(x_1)^{''}}{-\lambda_1 - 1 - \lambda_2 - 1} f_2(x_2)^{''} + \rho \frac{f_1(x_1)}{-\lambda_1 - 1 - \lambda_2 - 1} \frac{f_2(x_2)}{-\lambda_1 - 1 - \lambda_2 - 1} \]

\[ = f_1(x_1)f_2(x_2) + \rho \frac{[f_1(x_1)^{''} - f_1(x_1)] [f_2(x_2)^{''} - f_2(x_2)]}{(\lambda_1 + 1)(\lambda_2 + 1)}. \quad (3.33) \]

### 3.4.1 Copulas in higher dimensions

Ting Lee (1996) proposed the multivariate Sarmanov distribution extended (3.6) as:

\[ f_{12...n}(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} f_i(x_i)[1 + R_{q_1q_2...q_n}(x_1, x_2, \ldots, x_n)], \]

where

\[ R_{12...n} = \sum_{i<n}^{n-1} \sum_{j=1}^{n} \omega_{ij} q_i(x_i) q_j(x_j) + \]

\[ \sum_{i<n}^{n-2} \sum_{j<k=3}^{n} \sum_{k=3}^{n} \omega_{ijk} q_i(x_i) q_j(x_j) q_k(x_k) + \cdots + \omega_{ijk...m} \prod_{i=1}^{m} q_i(x_i) \quad (3.34) \]

and \( \Omega_n = \{ \omega_{ijk}, \ldots, \omega_{ijkm} \} \) is the set of real numbers chosen such that

\[ 1 + R_{q_1q_2...q_n}(x_1x_2 \ldots x_n) \geq 0. \]

The main drawback is the high computational cost needed to calculate all the combinations mentioned to the above equation. Miravete (2009b) based on the given
3.4. A PROPOSED ALTERNATIVE FGM COPULA

multivariate extension, proposed the extended Compound Poisson and Gamma distributions in three dimensions. Another multivariate distribution is proposed considering dependence only between two variables.

\[ R_{q_1q_2\ldots q_n}(x_1, x_2, \ldots, x_n) = \sum_{i<j}^{n-1} \sum_{j=2}^{n} \omega_{ij} q_i(x_i) q_j(x_j). \]  

(3.35)

Correlation coefficient is given by:

\[ \rho_{ij} = \frac{\omega_{ij} v_i v_j \mu_k}{\sigma_i \sigma_j}, \quad i, j, k = 1, 2, 3, \ i \neq j \neq k, \]

where \( v_i \), \( i = 1, 2, 3 \) given by (3.14) and \( \mu_i = Ef_i(x_i), \ i = 1, 2, 3 \).

A trivariate Poisson distribution similar to this proposed previously given by:

\[
p(x_1, x_2, x_3) = P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3) \]

\[ - P(X_1 \leq x_1, X_2 \leq x_2 - 1, X_3 \leq x_3) - P(X_1 \leq x_1 - 1, X_2 \leq x_2, X_3 \leq x_3) \]

\[ - [P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3 - 1) - P(X_1 \leq x_1 - 1, X_2 \leq x_2 - 1, X_3 \leq x_3)] \]

\[ - [P(X_1 \leq x_1 - 1, X_2 \leq x_2, X_3 \leq x_3 - 1) - P(X_1 \leq x_1 - 1, X_2 \leq x_2 - 1, X_3 \leq x_3 - 1)], \]

where trivariate Sarmanov distribution can be defined more simpler as:

\[
F(x_1, x_2, x_3) = \prod_{i=1}^{3} F(x_i) [1 + \omega_{12} F(x_1) F(x_2) (1 - F(x_1)) (1 - F(x_2)) F(x_3) \\
+ \omega_{13} F(x_1) F(x_3) (1 - F(x_1)) (1 - F(x_3)) F(x_2) \\
+ \omega_{23} F(x_2) F(x_3) (1 - F(x_2)) (1 - F(x_3)) F(x_1)].
\]

At this chapter we discussed a way to construct multivariate distributions. Based on Sklar’s theorem and Sarmanov distribution, many alternative approaches have been also studied. The novelty in our study is not only that we try to construct a less complicated multivariate distribution at the same period with Piancastelli et al. (2023) but we provide a wider range at correlation coefficient considering another function in the term of multiplicative factor. The importance is that we could imply this function in both discrete and continuous case taking each time different construction for \( \rho \)'s boundaries. Furthermore when we have a continuous variable, calculations are less restrictive and complicated as proved in case of exponential distribution. The construction of our new copula has been done not only to be useful for construction of any bivariate distribution considering appropriate marginals each time, but also to construct a new Poisson INGARCH(1,1) and a new CARR(1,1) time series models with studies provided at the next two chapters.
Chapter 4

A multivariate INGARCH(1,1) time series model

4.1 Model

Fokianos et al. (2020) studied linear and log-linear models for multivariate count time series models with Poisson marginals. The model was structured taking into account fundamental properties of Poisson process via a vector of continuous distributions and considering a copula form. The dependence among marginal Poisson components based on the properties of the Poisson process. The construction of the multivariate Poisson INGARCH model based on the construction of Freeland and McCabe (2004) A brief discussion about this model was done in the introduction of the thesis. In what follows we assume that $X_t = (X_{i,t}), \ i = 1,...,p, \ t = 1,2,...,$ denotes a $p$–dimensional count time series. Let $\lambda_t = (\lambda_{i,t}), \ i = 1,...,p, \ t = 1,2,...,T$ be the corresponding $p$-dimensional intensity process and $\mathcal{F}_t^{X_{i}}$ the $\sigma$–field generated by $X_0, X_t, \lambda_0$ with $\lambda_0$ being a $p$-dimensional vector denoting the starting value of $\lambda_t$. With this notation, the intensity process is given by $\lambda_t = E[X_t|\mathcal{F}_t^{X_{i}}]$. The new trivariate model constructed assuming that the linear dependence of $\lambda_t$ on $\lambda_{i,t-1}$ and $X_{j,t-1}$ in the presence of $\lambda_{i,t-1}$ and $X_{j,t-1}, i \neq j, \ i, j = 1,2,3$

$$X_t \mid \mathcal{F}_{t-1}^{X_{i}} \sim TP(\lambda_t)$$

where

$$\lambda_t = d + A\lambda_{t-1} + BX_{t-1}, \quad (4.1)$$

assuming that for each $i = 1, 2, 3$

$$X_{i,t} \mid \mathcal{F}_{t-1}^{X_{i}} \sim P(\lambda_{i,t}), \ i = 1,2,3, \quad (4.2)$$

81
where
\[ \lambda_{i,t} = d_i + \sum_{j=1}^{3} a_{i,j} \lambda_{j,t-1} + \sum_{j=1}^{3} b_{i,j} X_{j,t-1}, \quad i, j = 1, 2, 3. \] (4.3)

Each \( \lambda_{i,t}, \ i = 1, \ldots, p \) depends not only on its own past but also to past observations of the other series and this indicates that \( A \) and \( B \) are matrices of dimension \( p \times p \) (in multivariate case). Those matrices are responsible for heteroskedasticity because in case of constant conditional mean and variance \( (A = B = 0) \) the model becomes an integer AR Poisson model. Furthermore if elements \( a_{ij}, b_{ij}, i \neq j, \ i, j = 1, 2, 3 \) are zero then matrices \( A, B \) are diagonal and conditional mean depends only on its own past and its past values. In the case where \( A = 0 \) then the model reduced to an INARCH process. Fokianos et al. (2020) proved that consider model (1.14) and suppose that \( \|A\|_1 + \|B\|_1 < 1 \), \( \|A\|_1 = \max_{1 \leq j \leq p} \sum_{i=1}^{q} |a_{ij}| \), then there exists a unique solution \( (Y_t, \lambda_t) \) to model (4.1) which is stationary and ergodic. The elements of \( d,A,B \) are assumed to be positive considering positivity of \( \lambda_{i,t} \).

### 4.2 Ergodic Properties

#### 4.2.1 Weak dependence criteria

Fokianos et al. (2020) studied stationarity and ergodicity of the linear model by using Tweedie criteria. Based on the way with which Fokianos et al. (2009) studied ergodicity of univariate linear INGARCH model using a perturbed model, on a similar way Fokianos et al. (2020) study ergodic properties for a bivariate INGARCH model. At the proposed trivariate Poisson INGARCH model we study ergodicity of the model by using weak dependence criteria based on the work of Dedecker et al. (2007) without using a perturbation model. There are numerous criteria of weak dependence but we study ergodicity based on \( \theta \) and \( \tau \) weak dependence criteria and based on (Dedecker et al. (2007), pg. 115) \( \tau_1(\mathcal{F}_{t-1}^X, X_t) = \|X_t - \hat{X}_t\|_1 \). Under the assumption that \( h : \mathbb{R}^m \to \mathbb{R}^n \) is a Lipschitz function, where \( m \) is and \( n \) is , , . We prove that \( \{X_t\} \) is ergodic. Proof is given in the Appendix.

**Theorem 3.** If \( h(\lambda_t, X_t) \) satisfies the Lipschitz condition, then \( \{X_t\}_{t \in \mathbb{Z}} \) is stationary and \( \theta \)-weak dependent.

Considering that \( \tilde{\mathcal{F}}_{t-1} \) is a \( \sigma \)-algebra generated by \( \lambda_t, X_t, \tilde{\lambda}_t, \tilde{X}_t \), i=0,...,t-1 where \( \tilde{\lambda}_t, \tilde{X}_t \) are perturbed volatilities, then
\[
\|X - \tilde{X}\|_1 = E|X - \tilde{X}| = E[E \left( |X - \tilde{X}| \mid \tilde{\mathcal{F}}_{t-1} \right)] = E|\lambda_t - \tilde{\lambda}_t| = ||\lambda - \tilde{\lambda}||_1.
\]
Based on the above, under the assumption that multivariate function \( h: \mathbb{R}^m \rightarrow \mathbb{R}^n \) satisfy the Lipschitz condition \( |h(x) - h(y)| \leq L|x - y| \) where \( L \) is constant less than 1 and by using Minkowski's inequality where \( A_i, B_j \) are squared, invertible matrices in general case that we have a multivariate INGARCH(p,q) model we have that volatility given by:

\[
\lambda_t = d + A \lambda_{t-i} + BX_{t-i},
\]

where

\[
\lambda_t = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad d = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}, \quad A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix}
\]

and

\[
\lambda_t = d + \sum_{i=1}^{p} A_i \lambda_{t-i} + \sum_{j=1}^{q} B_j Y_{t-j}.
\]

\[
||X_t - \tilde{X}_t||_1 = ||\lambda_t - \tilde{\lambda}_t||_1 = ||h(\lambda_{t-i}, X_{t-i}) - h(\lambda_{t-i}, \tilde{X}_{t-i})||_1
\]

\[
\leq \left( \sum_{i=1}^{p} ||A_i||_1 + \sum_{j=1}^{q} ||B_j||_1 \right) \max_{1 \leq k_i \leq m_1} ||X_{k_i, t-k_i} - \tilde{X}_{k_i, t-k_i}||_1.
\]

By iterating the procedure we have that:

\[
||X_t - \tilde{X}_t||_1 \leq \left( \sum_{i=1}^{p} ||A_i||_1 + \sum_{j=1}^{q} ||B_j||_1 \right) \sum_{i=1}^{*} k_i \max ||X_{k, t-\sum_{i=1}^{*} k_i} - \tilde{X}_{k, t-\sum_{j=1}^{*} k_i}||_1,
\]

where \( k_i \in [1, \max(p, q)] \).

In our study we consider a trivariate INGARCH time series model with Poisson marginals. More specifically by iterating (1.14) we have an equation which decrease
CHAPTER 4. A MULTIVARIATE INGARCH(1,1) TIME SERIES MODEL

computational cost:

\[
\lambda_t = a + Ax_{t-1} + BX_{t-1}
\]

\[
= d + Ad + A^2x_{t-2} + ABX_{t-2} + BX_{t-1}
\]

\[
= \ldots
\]

\[
= \sum_{i=0}^{k-1} A^id + A^kx_{t-k} + B \sum_{i=0}^{k-1} A^iX_{t-i-1}, \quad (4.4)
\]

where \(d\) is the i-th element of \(d\) and \(a_{ij}\) is the (i,j)-element of \(A\) (B respectively).

Assuming first order stationarity, taking conditional expectation and from passage to the limit as \(k \to \infty\) the above equation can be written as:

\[
\lambda_t = (I - A)^{-1}d + \sum_{i=0}^{\infty} A^iBX_{t-i-1}, \quad (4.5)
\]

where \(I\) is the identical matrix of 3-dimensions.

In trivariate INGARCH model according to construction of trivariate Poisson distribution given by (3.34) both negative and positive correlation can be calculated. In trivariate Poisson distribution from the restriction \(\omega_{12}q_1(x_1)q_2(x_2) + \omega_{13}q_1(x_1)q_3(x_3) + \omega_{23}q_2(x_2)q_3(x_3)\) parameters \(\omega_{ij}, \ i, j = 1, 2, 3, i \neq j\) are bounded based on (3.22 , 3.23) but the main difference is that at the numerators \(l_k, k = 1, 2, 3\) are chosen randomly such that \(l_1 + l_2 + l_3 = 1\).

4.3 Estimation

The vital problem in multivariate distribution is the expression of the joint distribution. At those models that have been already proposed the multivariate pmf is quite complicated and maximum likelihood based inference is challenging. Considering that \(\{X_t, t = 1, 2, \ldots, n\}\) is a sample from the count time series model proposed above and the vector of the parameters denoted by \(\theta\) where \(\theta = (d^T, A^T, B^T, \omega^T)\). Using the general distribution defined by (3.36), (3.34) we can estimate parameters of each multivariate model. Fokianos et al. (2020) maximized the Quasi log-likelihood function. Based on lemmas (4.1) and (4.2), the examination of asymptotic normality has been provided based on the unperturbed model. The log-likelihood function can be written as:

\[
\log L(\theta) = \sum_{t=1}^{n} \sum_{i=1}^{3} \log p_i(x_{it}) + \sum_{t=1}^{n} \log[1 + \sum_{i<2}^{3} \sum_{j=2}^{3} \omega_{ij}q_i(x_{it})q_j(x_{jt})].
\]
4.3. ESTIMATION

Considering (4.3) log-likelihood function for trivariate Poisson INGARCH model is given by:

\[ l(\theta) = n \sum_{t=1}^{n} \left( \sum_{i=1}^{3} [-\lambda_i(\theta) + X_{i,t} \log(\lambda_i(\theta)) - \log(X_{i,t}!)] + \log \phi_t(\theta) \right), \]  

(4.6)

where

\[ \phi_t(\theta) = 1 + \sum_{i<j}^{3} 2 \omega_{ij} \left( c_i^{(-k_i-1)} - e^{\lambda_i(c_i^{-k_i-1})} \right) \left( c_j^{(-k_j-1)} - e^{\lambda_j(c_j^{-k_j-1})} \right). \]  

(4.7)

Score function is given by:

\[ S_n(\theta) = \frac{\partial l(\theta)}{\partial \theta} = \sum_{t=2}^{n} \frac{\partial l_t(\theta)}{\partial \theta} = \sum_{t=1}^{n} \left( \sum_{i=1}^{3} \left( \frac{X_{i,t}}{\lambda_i,t} - 1 \right) \frac{\partial \lambda_i,t(\theta)}{\partial \theta} + \frac{1}{\phi_t(\theta)} \frac{\partial \phi_t(\theta)}{\partial \theta} \right), \]

where \( \lambda_{i,t} \) defined by (1.14) and \( \theta = (d_i, a_{ij}, b_{ij}, \omega_{ij}), \ i,j = 1,2,3 \). The score function satisfies:

\[ \frac{1}{\sqrt{n}} S_n(\theta) \to N(0,I(\theta_0)) \]  

(4.8)

**Theorem 4.** For the INGARCH model as \( n \to \infty \)

\[ \sqrt{n}(\hat{\theta} - \theta_0) \to N(0,I(\theta_0)^{-1}) \]

where

\[ I(\theta_0) = E \left( \frac{\partial l(\theta_0)}{\partial \theta} \frac{\partial l(\theta_0)}{\partial \theta^T} \right) = -E \left( \frac{\partial^2 l(\theta_0)}{\partial \theta \partial \theta^T} \right). \]

Based on the work of Lee et al. (2016) conditions for strong consistency of MLE estimator in INGARCH models are accomplished in our case. The true parameter value \( \theta_0 \) is an interior point on \( \Theta \), with \( \Theta \) being a compact set. Furthermore considering conditions of stationarity of the model \( \mathbf{d,A,B} \) do not have negative entries and \( \text{Cov}(X_{i,t},X_{j,t}) < \min(a_1,a_2) \) where \( (a_1,a_2)^T = (I - \mathbf{A})^{-1}\mathbf{d} \) for all \( \theta \in \Theta \). Considering (3.22, 3.23), \( \phi_L \leq \phi_t(\theta) \leq \phi_U \) and based on condition of stationarity \( \lambda_t \) is bounded \( \lambda_L \leq \lambda_t \leq \lambda_U \). Assuming parameter vectors \( \theta_i = (a_{ij}, b_{ij}, d_i), i,j = 1,2,3 \) and \( \theta_4 = (\omega_{12}, \omega_{13}, \omega_{23}) \) first and second order partial derivatives are presented in appendix B. Moreover based on equation (1.14) derivatives for \( \lambda_{i,t}(\theta) \) are:

\[ \frac{\partial \lambda_{i,t}(\theta)}{\partial d_i} = d_i + b_{ij} \frac{\partial \lambda_{i,t-1}}{\partial d_i}, \]

\[ \frac{\partial \lambda_{i,t}(\theta)}{\partial b_{ij}} = \lambda_{j,t-1} + b_{ij} \frac{\partial \lambda_{i,t-1}}{\partial b_{ij}}, \]

\[ \frac{\partial \lambda_{i,t}(\theta)}{\partial a_{ij}} = \lambda_{j,t-1} + a_{ij} \frac{\partial \lambda_{i,t-1}}{\partial a_{ij}}. \]
CHAPTER 4. A MULTIVARIATE INGARCH(1,1) TIME SERIES MODEL

and they are $\mathcal{F}_{t-1}$-measurable. Assuming second order partial derivatives we have:

$$E \left( \frac{X_{it}}{\lambda_i(\theta)} \mid \mathcal{F}_{t-1} \right) = -\frac{1}{\lambda_i} E \left( \left( c_i^{-k_i x_{it}} - e^{\lambda_i(c_i^{k_i-1})} \right) \mid \mathcal{F}_{t-1} \right) = 0$$

Under above conditions and considering Minkowski’s inequality we have:

$$E \left( -\frac{\partial^2 l}{\partial \theta^2_{ij}} \mid \mathcal{F}_{t-1} \right) \leq \frac{1}{\lambda_U} + \frac{c_i^{-k_i} - 1}{\phi_U} + \frac{(c_i^{-k_i} - 1)^2}{\phi_U^2} + 2 \left( \frac{c_i^{-k_i} - 1}{\phi_U} \right)^2.$$ (4.9)

Based on the above $E \sup_{\theta \in \Theta} \left| \frac{\partial^2 l}{\partial \theta^2} \right|_1 < \infty$. Proofs are given in Appendix A. Using Taylor series expansion for $\frac{\partial l}{\partial \theta}$ we have

$$0 = \frac{\partial l}{\partial \theta} = \frac{\partial l}{\partial \theta_0} + (\hat{\theta} - \theta_0) \frac{\partial^2 l}{\partial \theta^2} \text{ or } \frac{1}{\sqrt{n}} \frac{\partial l}{\partial \theta_0} = \sqrt{n}(\hat{\theta} - \theta_0) \left( -\frac{1}{\sqrt{n}} \frac{\partial^2 l}{\partial \theta^2} \right).$$

Considering Lemma 6 of Lee et al. (2016) $\frac{1}{n} \frac{\partial^2 l}{\partial \theta^2} \rightarrow I(\theta_0)$ and by using martingale central limit theorem $\frac{1}{\sqrt{n}} \frac{\partial l}{\partial \theta_0} \rightarrow N(0, I(\theta_0))$. Based on the above

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N(0, I(\theta_0)^{-1}).$$

Considering that in trivariate case there are numerous parameter for estimations other ways for estimations could be provided. Here we briefly discussed the method of likelihood by parts based on score function’s decomposition. Assuming that model given by (4.2), (4.3) likelihood by parts defined as:

$$l(\theta) = l(\theta_1) + l(\theta_1, \theta_2),$$ (4.10)

where

$$l(\theta_1) = \sum_{t=1}^{n} \left( \sum_{i=1}^{3} \left[ -\lambda_i(\theta) + X_{i,t} \log(\lambda_i(\theta)) - \log(X_{i,t}) \right] \right),$$

$$l(\theta_1, \theta_2) = \sum_{t=1}^{n} \log \phi_i(\theta)$$

and

$$\theta_1 = (d_i, a_{ij}, b_{ij}), \theta_2 = (\omega_{ij}), \text{ i, j = 1, 2, 3.}$$

Song et al. (2005) proposed repetition of steps 1-k to solve the above equations under the assumption that as $k \rightarrow \infty$, $\theta_n^k$ will converge to $\theta$. 
4.4. SIMULATIONS

- **Step 1:** \( \frac{\partial l(\theta_1)}{\partial \theta_1} = 0 \), for \( \theta_1 \) and \( \frac{\partial l(\theta_1, \theta_2)}{\partial \theta_2} = 0 \), for \( \theta_2 \).

- **Step k:** \( \frac{\partial l(\theta_1)}{\partial \theta_1} = -\frac{\partial l(\theta_{k-1}, \theta_{k-1})}{\partial \theta_1} \), for \( \theta_1 \) and \( \frac{\partial l(\theta_{k-1}, \theta_2)}{\partial \theta_2} = 0 \), for \( \theta_2 \).

In our case this method is very useful in comparison with the maximum likelihood approach of the full bivariate distribution. This means that calculating the partial derivative of \( l(\theta_1) \) and then substituting at \( l(\theta_{k-1}, \theta_2) \) the number of parameters for estimation of \( l(\theta_{k-1}, \theta_2) \) is decreased and the algorithm becomes less intensive.

4.4 Simulations

We run a simulation study in order to check the performance of estimators. In our simulation study we use sample size of \( n=50, n=200 \) and \( n=500 \) with \( m=200 \) replications. Based on inverse transformation method for the data generation we have:

- **Step 1:** Choose initial values for the parameter vector \( \theta = (d_i, a_{ij}, b_{ij}, \omega_{ij}) \), \( i, j = 1, 2 \).

- **Step 2:** Generate an observation \( x_1 \) from a \( P(\lambda_1) \) where \( \lambda_{1t} = d_1 + a_{11} \lambda_{1,t-1} + a_{12} \lambda_{2,t-1} + b_{11} X_{1,t-1} + b_{12} X_{2,t-1} \).

- **Step 3:** Generate a random number \( u_1 \sim U(0, 1) \).

- **Step 4:** If there exist a \( k_1 \) such that \( P_1(k_1 - 1 \mid x_1) \leq u_1 \leq P_1(k_1 \mid x_1) \) set \( x_2 = k_1 \), where \( P_1(k_1 - 1 \mid x_1) \) and \( P(k_1 \mid x_1) \) are conditional distributions.

- **Step 5:** Generate a number \( u_2 \sim U(0, 1) \).

- **Step 6:** If there exist a \( k_2 \) such that \( P_2(k_2 - 1 \mid x_1, x_2) \leq u_2 \leq P_2(k_2 \mid x_1, x_2) \) then set \( x_3 = k_2 \).

where

\[
P_1(x_2 \mid x_1) = P(\lambda_2) (1 + A(\lambda_1, \lambda_2)),
\]

\[
P_2(x_3 \mid x_1, x_2) = P(\lambda_3^*) \frac{1 + \sum_{i<j=1}^3 A(\lambda_i^*, \lambda_j^*)}{A(\lambda_1^*, \lambda_2^*)},
\]
A(\lambda_i, \lambda_j) = \omega_{ij} \left( e^{-k_i x_i} - e^{-\lambda_i^s(e^{-k_i-1})} \right) \left( e^{-k_j x_j} - e^{-\lambda_j^s(e^{-k_j-1})} \right)

and \lambda_{it} defined in step 2 and \lambda_{it}^\ast defined as

\lambda_{it}^\ast = d_i + a_{i1} \lambda_{1,t-1}^\ast + a_{i2} \lambda_{2,t-1}^\ast + a_{i3} \lambda_{3,t-1}^\ast + b_{i1} X_{1,t-1} + b_{i2} X_{2,t-1} + b_{i3} X_{3,t-1}.

Furthermore at the formulas above \( P(\lambda_i) \) and \( P(\lambda_i^\ast) \) are Poisson distributions where \( \lambda_{it} \) and \( \lambda_{it}^\ast \) defined at steps above. For the trivariate Poisson model we estimate parameter vector \( \theta = (d_1, d_2, d_3, a_{ij}, b_{ij}, \omega_{ij}) \) with both negative and positive values for parameters \( \omega_{ij}, i, j = 1, 2, 3 \). In order to take optimal values for parameters \( \omega_{ij}, i, j = 1, 2, 3 \) we take into account (3.22; 3.23) and substituting optimal values of the other parameters by the way described above, we get corresponding \( \omega_{ij}' \)'s values.

At the case in which the joint distribution does not have a closed form the choice for initial value of \( \omega \) is demanded. Fokianos et al. (2020) at the proposed model overcome the problem considering the method studied by Tjøstheim and Hufthammer (2013) for local Gaussian correlation. Considering a two-dimensional random variable the bivariate density can be approximated by a Gaussian bivariate density in a neighbourhood at each point \( x = (x_1, x_2) \) where correlation coefficient defined as \( \rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \). Tjøstheim and Hufthammer (2013) provided that the parameter vector of Gaussian distribution is defined by using a likelihood related penalty function \( q \) given by:

\[
q = \int K_b(v - x) \left[ \psi(v, \theta(x)) - \log \psi(v, \theta(x)) f(v) \right] dv, \tag{4.11}
\]

where \( K_b(v - x) = \frac{K_b^{-1}(v_1 - x_1) K_b^{-1}(v_2 - x_2)}{b_1 b_2} \) and \( b = (b_1, b_2) \) is a bivariate bandwidth. An other one interpretation for the above equation is a locally weighted Kullback-Leibler distance from \( f \) to \( \psi(\cdot, \theta(x)) \).

The biases and standard deviations are presented in (4.1; 4.2; 4.3) in Appendix. Simulations were carried out in R using constrOptim considering constraints defined at the previous sections. All estimators perform well as small biases are obtained.
4.4. SIMULATIONS
Figure 4.1: Boxplot with results from the simulation experiment with initial parameter values $(a_{ij}, b_{ij}, d_{ij}, \omega_{ij}) = (0.12, 0.1, 0.1, 0.22, 0.15, 0.18, 0.15, 0.18, 0.22, 0.16, 0.18, 0.18, 0.1, 0.1, 0.1, 0.1, 0.15, 0.1, 0.1, 2, 1, 1.5, 4, 5, 6)$
4.4. SIMULATIONS
**Figure 4.2:** Boxplot with results from the simulation experiment with initial parameter values $(a_{ij}, b_{ij}, d_{ij}, \omega_{ij}) = (0.35, 0.12, 0.05, 0.18, 0.12, 0.15, 0.1, 0.4, 0.1, 0.15, 0.18, 0.1, 0.05, 0.1, 0.1, 0.1, 1.5, 1.5, 3, 2.8, 4, 3, 1.5)$
4.4. SIMULATIONS
Figure 4.3: Boxplot with results from the simulation experiment with initial parameter values \((a_{ij}, b_{ij}, d_{ij}, \omega_{ij}) = (0.2, 0.2, 0.1, 0.25, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.2, 0.1, 0.1, 0.1, 0.2, 0.1, 0.1, 1, 1.5, 1.8, 2, 2, 3)\)
4.5 Application

Recent years there is a flourishing interest about data for road accident statistics. A number of approaches have been suggested to model time-series crash count data, most of them do not explicitly take into account the significant autocorrelation that traditionally is present in the data. The ignorance of autocorrelation lead to statistical inconsistency of the coefficient estimates producing standard estimates that are too optimistic. Brijs et al. (2008) proposed a Poisson INAR(1) regression model for modelling daily car crash data with time interdependencies and examine the risk impact of weather conditions on the observed counts. At this application, we use daily crash counts in three areas of Netherlands that is, Shiphol, De Bilt and Soesterberg. The data share some similar environment, especially with respect to weather conditions, road characteristics and traffic exposure. From time series plots presented in (4.4) we could see that the large number of accidents is obtained in weekdays.
Pedeli and Karlis (2013) calculated cross-correlation between the three time series providing the emergency study of a multivariate time series process. Autocorrelation for the three series and cross-correlation also presented at (4.5, 4.5). The first two series are auto-correlated at lag=1. Some descriptive statistics are also presented at the table 5.3 . Calculating means and variances we take values for variances larger than those for means with an indication to overdispersion. This implies that Poisson distribution is not appropriate some other distribution such as a Negative binomial distribution is more appropriate for those data. Considering the study of Zhu (2011) about a univariate negative binomial integer-valued GARCH model, a multivariate model need to discussed for modeling those data. Fokianos et al. (2020) based on study of Latour (1997) approaches INGARCH(1,1) model with an ARMA(1,1) model as:

\[(X_t - \mu) - (B + D)(X_{t-1} - \mu) = e_t - De_{t-1}, \tag{4.12}\]

where \(\mu = E(Y_t) = (I - B - A)^{-1}d\). Based on the general ARMA(1,1) model \(X_t - \phi X_{t-1} = e_t + \theta e_{t-1}\) compared with (4.12) we get that \(\phi = A + B\), \(\theta = -A\) and \(e_t \sim \text{Normal}(0, \Sigma)\).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shiphol</td>
<td>8.781</td>
<td>24.600</td>
</tr>
<tr>
<td>Weekdays/Weekends</td>
<td>9.770/6.298</td>
<td>2.39/3.06</td>
</tr>
<tr>
<td>De Bilt</td>
<td>2.584</td>
<td>3.681</td>
</tr>
<tr>
<td>Weekdays/Weekends</td>
<td>2.893/1.740</td>
<td>1.28/1.55</td>
</tr>
<tr>
<td>Soesterberg</td>
<td>3.370</td>
<td>5.673</td>
</tr>
<tr>
<td>Weekdays/Weekends</td>
<td>3.667/2.625</td>
<td>1.53/1.95</td>
</tr>
</tbody>
</table>

**Table 4.1:** Descriptive statistics for the time series of accidents
We have indication of dependence between the number of crashes at three cities.
One can observe that cross correlation between Schiphol and De Bilt decline faster than the others. On the other hand positive cross correlation is observed between daily crashes at Soesterberg and De Bilt indicating that when there is an increase at number of daily crashes at Soesterberg, there is also an increase at those of De Bilt h days later, where h stands for the time lag.

In our study the parameters of the multivariate Poisson model are calculated by ML method. Furthermore the parameters of the trivariate INGARCH Poisson time series model are calculated by using constrOptim in R and considering parameter vector $\theta$ as presented in simulation. As constraints we consider conditions of stationarity for parameters $a_{ij}, b_{ij}, \ i, j = 1, 2, 3$ and under the assumption that $k_i = 1, \ i = 1, 2, 3$, we restrict parameters $\omega^{12}, \omega^{13}, \omega^{23}$ according to (3.22), (3.23).

<table>
<thead>
<tr>
<th></th>
<th>$d_1$</th>
<th>1.002(0.016)</th>
<th>$b_{11}$</th>
<th>0.124(0.024)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d_2$</td>
<td>0.810(0.010)</td>
<td>$b_{12}$</td>
<td>0.123(0.023)</td>
</tr>
<tr>
<td></td>
<td>$d_3$</td>
<td>0.814(0.019)</td>
<td>$b_{13}$</td>
<td>0.113(0.013)</td>
</tr>
<tr>
<td>$a_{11}$</td>
<td>0.215(0.015)</td>
<td>$b_{21}$</td>
<td>0.084(0.016)</td>
<td></td>
</tr>
<tr>
<td>$a_{12}$</td>
<td>0.192(0.007)</td>
<td>$b_{22}$</td>
<td>0.0115(0.015)</td>
<td></td>
</tr>
<tr>
<td>$a_{13}$</td>
<td>0.113(0.013)</td>
<td>$b_{23}$</td>
<td>0.105(0.014)</td>
<td></td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>0.128(0.028)</td>
<td>$b_{31}$</td>
<td>0.243(0.043)</td>
<td></td>
</tr>
<tr>
<td>$a_{22}$</td>
<td>0.117(0.017)</td>
<td>$b_{32}$</td>
<td>0.084(0.015)</td>
<td></td>
</tr>
<tr>
<td>$a_{23}$</td>
<td>0.188(0.012)</td>
<td>$b_{33}$</td>
<td>0.122(0.022)</td>
<td></td>
</tr>
<tr>
<td>$a_{31}$</td>
<td>0.299(0.001)</td>
<td>$\omega^{12}$</td>
<td>-2.465(0.035)</td>
<td></td>
</tr>
<tr>
<td>$a_{32}$</td>
<td>0.134(0.034)</td>
<td>$\omega^{13}$</td>
<td>-1.448(0.051)</td>
<td></td>
</tr>
<tr>
<td>$a_{33}$</td>
<td>0.039(0.061)</td>
<td>$\omega^{23}$</td>
<td>-1.961(0.039)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: ML estimates for trivariate Poisson model parameters based on volatility construction defined by (1.14; 4.3). Numbers in parentheses are standard errors.

### 4.5.1 Goodness of fit

We consider Pearson residuals defined by $e_{i,t} = \frac{Y_{i,t} - \lambda_{i,t}}{\sqrt{\lambda_{i,t}}} , \ i = 1, 2$ in order to examine model fit. Assuming that those residuals could be expressed as:

$$
e_t = \sum_{j=1}^{n/2} \beta_1 \left( \frac{j}{n} \right) \cos(2\pi \omega_j t) + \beta_2 \left( \frac{j}{n} \right) \sin(2\pi \omega_j t),$$

and calculating periodograms observations with high-valued frequencies were not observed. This means that there are not extreme values at Pearson residuals and
4.5. APPLICATION

equivalently there are no divergences between estimators and real values. Based on
theory under the correct model residuals come from a white noise with constant
variance and cumulative periodograms presented at 4.5.1 are smooth. In our study
for calculating \( \hat{e}_{i,t} \), we consider \( \lambda_{i,t} \) as \( \lambda_{i,t}(\hat{\theta}) \) where \( \hat{\theta} = (\hat{\omega}_{ij}, \hat{a}_{ij}, \hat{b}_{ij}) \). In graphs below
we present cumulative periodograms for trivariate INGARCH model.

![Cumulative periodogram plots](image)

**Figure 4.7:** Cumulative periodogram plots of the Pearson residuals for daily crash
counts for Schiphol(left upper), De Bilt(right upper) and Soesterberg.

4.5.2 Predictive marginals

We make prediction based on method of conditional expectation. Based on eqn (4.4)
an one step ahead forecasting is:

\[
\hat{\lambda}_{t+1} = E(\lambda_{t+1} | \mathcal{F}) = d + A\lambda_t + BX_t
\]  

(4.13)

and using recursion we have:

\[
\hat{\lambda}_{t+1} = \sum_{i=0}^{k-1} (A^i d + B^i X_{t-i}) + A^k \lambda_{t-k-1}.
\]
Considering recursive formula a p-steps ahead forecasting can be calculated by:

\[ \hat{\lambda}_{t+p} = E \left( \lambda_{t+p} \mid F^X \right) = \sum_{i=0}^{k-1} A^i d + \sum_{i=0}^{k-1} \left( BA^i X_{t+p-i} \right) + A^k E \left( \lambda_{t+p-k} \mid F \right). \]

Consecutively calculating \( \lambda_{t+p}, \ p = 1, 2, ..., n \) we choose every time this \( \lambda_t \) with the greatest probability. Based on (3.4) we predict \( \lambda \) at time \( t = 366 \) and then we calculate bivariate marginal predictive pmfs given by:

\[
p(x_i, x_j) = p(x_i) p(x_j) \left[ 1 + \omega_{ij} \left( c_i^{(-k_i-1)} - e^{\lambda_i(c_i^{-k_i-1})} \right) \left( c_j^{(-k_j-1)} - e^{\lambda_j(c_j^{-k_j-1})} \right) \right].
\]

Figure 4.8: One step ahead bivariate predictive probability for each one of the three places in Netherlands.
Moreover we compare our model with this studied by Fokianos et al. (2020). At this illustration working with the same data consists of number of transactions per 15 seconds for the stocks Coca-Cola (KO) and IBM on September 19th 2005, a likelihood ratio test is provided and graphical inspection for residuals for both models is discussed. On the same way we keep 1440 where transactions before 9:30 and after 16:30 are not included. Figure (4.9) shows time series plots and (4.10) depicts the autocorrelation and cross-correlation functions. Autocorrelation functions plot indicates high correlated transactions data within and between the individual transaction series.

There is evidence that our proposed model fit better the data providing a likelihood ratio test where log-likelihoods are -2519.479 and -6884.62 respectively. Moreover we calculate Pearson residuals as \( \hat{e}_{i,t} = \frac{X_{i,t} - \hat{X}_{i,t}}{\sqrt{\hat{X}_{i,t}}} \), where \( \hat{X}_{i,t} = \lambda_{i,t}(\hat{\theta}) \) and \( \hat{\theta} \) is parameters estimations vector. Based on the graph below (4.11) in model 2, which is the INGARCH model constructed based on Sarmanov distribution discussed in chapter 3, we have a more consistent spread implying a less substantial dispersion.
At this chapter, considering multivariate distribution discussed at chapter 3, a new multivariate INGARCH time series model has been provided. The keypoint is the construction of a flexible Poisson distribution which provide many ways for parameters estimation. The contribution to the almost existed models is that we provide a flexible INGARCH(1,1) model with contemporaneous correlation assuming...
that conditional expectation depends on past values. In our study volatilities are calculated considering interdependencies. Parameters estimation was provided by Maximum Likelihood approach but many other methods like likelihood by parts could be applied in order to simplify and to decrease computational cost. A brief discussion of likelihood by parts method is also provided. A dataset with daily crash counts has been studied without being totally appropriate to study with a Poisson INGARCH model. Moreover in order to provide the efficacy of the copula proposed in chapter 3 in case of continuous variable and using the new range of correlation coefficient, a CARR model is also studied in chapter 5.
Chapter 5

A bivariate CARR(1,1) model

5.1 The model

The study of interactions between two or more assets in finance seems to be of particular importance. Assuming that $X_t$ is a variable for daily stock prices, Fernandes et al. (2005) studied a CARR model where $R_t$ is Exponentially distributed range $R_t$ between daily maximum and minimum stock prices. At the proposed trivariate CARR model volatilities expressed linearly with previous volatilities and ranges of days up to $t - 1$. The above structure is capable to capture covariates and spill-over effects but many drawbacks lead to more extensively study of multivariate CARR models. The first one is the restrictive interval of correlation coefficient provided by the structure of multivariate exponential distribution. The second one is that no implementations have been provided yet considering the lack of intradaily data and for that reason any parameter estimation method does not have been discussed. Furthermore ergodic properties have been studied using criteria developed by Tweedie (1976) and $\beta$-mixing property but other methods assuming weak dependence criteria discussed by Doukhan et al. (2012) could be also yielded for simplicity. The main shortcoming is that while ergodic properties and estimation approaches are not difficult to study theoretically, it is hard to implement data under the assumption of Exponential distribution. In our study we contribute at CARR models by constructing a bivariate conditional autoregressive range process based on bivariate exponential distribution defined by (5.10). A maximum likelihood approach is discussed and convergence of maximum likelihood estimations is also provided. Numerous efforts for strengthening goodness of fit criteria at heteroskedastic models have been done by Meintanis et al. (2020) and Jiménez-Gamero et al. (2020). Based on those, a brief discussion presented in our study and results of simulation are . A review for range volatilities of univariate and multivariate conditional stochastic
CHAPTER 5. A BIVARIATE CARR(1,1) MODEL

processes given by Chou et al. (2015).
Assume that \( R_t = \lambda_t \epsilon_t \) where \( \epsilon_t \) is a sequence of positive independent and identical exponentially distributed random variables with mean equal to 1 and \( \lambda_t \) equal to conditional expectation of \( R_t = \max(X_t) - \min(X_t) \) up to time \( t \), where \( X_t \) is a continuous variable usually express the price. The proposed bivariate exponential process is defined as:

\[
R_t \mid F_{t-1} \sim BExp(\lambda_t).
\]

In our study we assume that each conditional expectation expressed only on its past ignoring the interactions between the two processes and volatility spill-over effects. In other words:

\[
\begin{bmatrix}
R_{1t} \\
R_{2t}
\end{bmatrix} =
\begin{bmatrix}
\lambda_{1t} & 0 \\
0 & \lambda_{2t}
\end{bmatrix}
\begin{bmatrix}
\epsilon_{1t} \\
\epsilon_{2t}
\end{bmatrix},
\]

(5.1)

where

\[
\lambda_{jt} = d_j + a_{j1} \lambda_{j-1} + b_{j1} R_{jt-1}, \quad j = 1, 2.
\]

(5.2)

Furthermore based on (5.2) and under the assumption of exponential distribution on residuals \( \epsilon_t \) conditional first and second order moments are given by:

\[
E[R_{jt} \mid F_{t-1}] = \lambda_{jt}, \quad \text{Var}[R_{jt} \mid F_{t-1}] = \lambda_{jt}^2.
\]

Unconditional moments of our process are defined based on the idea of Engle and Russell (1998) for martingale difference. Considering a martingale difference \( \eta_{jt} = R_{jt} - \lambda_{jt} \), \( j = 1, 2 \), where \( \lambda_{jt} \) defined by (5.2), a new expression for the calculation of unconditional moments is given by:

\[
R_{jt} = d_j + (a_{j1} + b_{j1}) R_{jt-1} - a_{j1} \eta_{jt-1} + \eta_{jt},
\]

where

\[
E(\eta_{jt}) = 0 \quad \text{and} \quad \text{Var}(\eta_{jt}) = E(\eta_{jt}^2).
\]

For the calculation of first and second order moments we assume conditions of weak stationarity. Assuming that the process is first order stationarity we have that

\[
E(\lambda_{jt} \mid F_{t-1}) = E(\lambda_{jt-1} \mid F_{t-1}) \quad \text{and} \quad E[R_{jt}] = \frac{d_j}{1 - \sum_{i=0}^{\max(p,q)} (a_{ji} + b_{ji})}
\]

and

\[
\text{Var}[R_{jt}] = \frac{E[R_{jt}^2(1-(a_j+b_j)^2+a_j^2)]}{1-(a_j+b_j)^2},
\]
5.1. THE MODEL

in case of an ECARR(p,q) model. When we have a ECARR(1,1) model then
\[ \mu_j = E(R_{jt}) = \frac{d_j}{1 - (a_{1j} + b_{1j})}. \]
Based on the above equation for the range \( R_{jt}, j = 1, 2 \) we have that:

\[
(R_{jt} - \mu_j) - (a_{j1} + b_{j1}) (R_{jt-1} - \mu_j) = \eta_{jt} - a_{j1}\eta_{jt-1}
\]
\[
(R_{jt} - \mu_j) R_{jt-1} - (a_{j1} + b_{j1}) (R_{jt-1} - \mu_j) R_{jt-1} = \eta_{jt} R_{jt-1} - a_{j1}\eta_{jt-1} R_{jt-1}
\]
\[
(R_{jt} - \mu_j) R_{jt} - (a_{j1} + b_{j1}) (R_{jt-1} - \mu_j) R_{jt} = \eta_{jt} R_{jt} - a_{j1}\eta_{jt-1} R_{jt}
\]

(5.3)

It is easy to verify that

\[
E(\eta_{jt-1} R_{jt}) = E(\eta_{jt-1} (d_j + (a_{j1} + b_{j1}) R_{jt-1} + \eta_{jt} - a_{j1}\eta_{jt-1})) = b_{j1}\sigma_{\eta_j}^2,
\]
\[
E(\eta_{jt} R_{jt}) = E(\eta_{jt} (d_j + (a_{j1} + b_{j1}) R_{jt-1} + \eta_{jt} - a_{j1}\eta_{jt-1})) = \sigma_{\eta_j}^2,
\]

(5.4)

Combining (5.3) and (5.4) we have:

\[
Var(R_{jt}) = (a_{j1} + b_{j1}) \gamma_{R_{jt}}(1) + \sigma_{\eta_j}^2 - a_{j1}b_{j1}\sigma_{\eta_j}^2,
\]
\[
\gamma_{R_{jt}}(1) = (a_{j1} + b_{j1}) Var(R_{jt}) - a_{jt}\sigma_{jt}^2,
\]

(5.5)

or

\[
Var(R_{jt}) = \frac{E[R_{jt}](1 - (a_{j1} + b_{j1})^2 + a_{j1}^2)}{1 - (a_{j1} + b_{j1})^2},
\]
\[
Cov(R_{jt}R_{jt-1}) = \frac{b_{j1} (1 - a_{j1}(a_{j1} + b_{j1})) E(R_{jt})}{(1 - (a_{j1} + b_{j1})^2)},
\]

and

\[
Cor(R_{jt}R_{jt-1}) = \frac{b_{j1} (1 - a_{j1}(a_{j1} + b_{j1}))}{1 - 2a_{j1}b_{j1} - a_{j1}^2}.
\]

5.1.1 Ergodic properties

Fernandes et al. (2005) proposed a trivariate model where each expected value expressed not only on its previous values but on the previous volatilities of other models in order to capture the interactions between the variance. In other words the model given by:

\[ R_t = \text{diag}(\lambda_t) \epsilon_t, \]
where \( R_t = (R_{1t}, R_{2t}, R_{3t}) \), \( \lambda_t = (\lambda_{1t}, \lambda_{2t}, \lambda_{3t}) \) and \( \epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t}) \).

Carrasco and Chen (2002) studied stationarity and properties of ergodicity based on Tweedie criteria. Fernandes et al. (2005) studied under which conditions the process is strictly stationary and \( \beta \)-mixing with exponential decay after expressing the model defined above as a trivariate autoregressive model of order 1 (MAR(1)) given by:

\[
X_{t+1} = C(\epsilon_t)X_t + D(\epsilon_t),
\]

where \( X_t = (R'_{t}, \lambda'_{t}) \), \( D(\epsilon_t) = (d_1\epsilon_{1t}, d_2\epsilon_{2t}, d_3\epsilon_{3t}, d_1, d_2, d_3) \) and

\[
C(\epsilon_t) = \begin{bmatrix}
b_{11}\epsilon_{1t} & b_{12}\epsilon_{2t} & b_{13}\epsilon_{3t} & a_{11}\epsilon_{1t} & a_{12}\epsilon_{2t} & a_{13}\epsilon_{3t} \\
b_{21}\epsilon_{1t} & b_{22}\epsilon_{2t} & b_{23}\epsilon_{3t} & a_{21}\epsilon_{1t} & a_{22}\epsilon_{2t} & a_{23}\epsilon_{3t} \\
b_{31}\epsilon_{1t} & b_{32}\epsilon_{2t} & b_{33}\epsilon_{3t} & a_{31}\epsilon_{1t} & a_{32}\epsilon_{2t} & a_{33}\epsilon_{3t} \\
b_{11} & b_{12} & b_{13} & a_{11} & a_{12} & a_{13} \\
b_{21} & b_{22} & b_{23} & a_{21} & a_{22} & a_{23} \\
b_{31} & b_{32} & b_{33} & a_{31} & a_{32} & a_{33}
\end{bmatrix}
\]

Dedecker et al. (2007) proposed numerous weak dependence criteria appropriate for the examination processes ergodicity. For more information someone can study the book of Dedecker et al. (2007). The main difference between Tweedie criteria and weak dependence criteria is that in the first case we have strong conditions of stationarity for all moments in contrast with the second case of weak stationarity of the first two moments. On the same way to previous chapter based on the work that we have also provided for INGARCH(1,1) model and considering study of Franke (2010), ergodic properties of a CARR(1,1) model are also provided. Assuming a Lipschitz function \( h : \mathbb{R}^m \rightarrow \mathbb{R}^n \), we prove that the process is ergodic. At those models where the conditional expectation expressed linearly, criteria of weak dependence for the proof of ergodicity are less complicated in the study. On the other hand when the expression is not linear other methods must be used for example see Fokianos et al. (2009).

**Theorem 5.** If \( \|A\|_1 + \|B\|_1 \leq 1 \) then the process \( \{R_t\} \) defined by (5.1) is stationary and \( \theta \)-weak dependent.

Based on Lemma 5.2(p 115) Dedecker et al. (2007) we have that \( \|R - R^*\|_1 = \tau_1(P_{t-1}^R, R) \), where \( \tau \)-weak dependence criterion has been discussed in chapter 2. Furthermore if the process \( R_t \) is \( \tau \)-weak dependent then satisfies the criterion of \( \theta \)-weak
dependence. The proof is similar to the proposed in trivariate INGARCH Poisson
process. In addition considering that parameter \( \lambda_t \) is the conditional expectation of
the Exponential process all parameters \( d_j, a_{ij}, b_{ij}, j = 1, 2 \) are positive.

5.2 Ergodic properties

Conditions of stationarity embedded on conditions of ergodicity. Carrasco and Chen
(2002) studied stationarity and properties of ergodicity based on Tweedie criteria.
Fernandes et al. (2005) studied appropriate conditions that the process is strictly
stationary and \( \beta \)-mixing with exponential decay.

Assuming the model defined by (5.1) and considering that mixing functions defined as (3.26)
for two exponentially distributed random variables with parameters
\( \frac{1}{\lambda_{it}}, i = 1, 2 \) we have:

\[
f(r_{1t}, r_{2t}) = \left[ \prod_{i=1}^{2} \frac{1}{\lambda_{it}} e^{-\frac{r_{it}}{\lambda_{it}}} \right] q(r_{1t}, r_{2t}),
\]

where

\[
q(r_{1t}, r_{2t}) = 1 + \omega \prod_{i=1}^{2} \left( c_i^{k_t r_{it}} - \frac{1}{1 + \lambda_{it} k_i \log c_i} \right)
\]

and

\[
\lambda_{jt} = d_j + a_j \lambda_{j,t-1} + b_j R_{j,t-1}, \quad j = 1, 2.
\]

Log-likelihood function is given by:

\[
l(\theta; r) = - \sum_{i=1}^{2} \sum_{t=1}^{n} \log \lambda_{it} - \sum_{i=1}^{2} \sum_{t=1}^{n} \frac{R_{it}}{\lambda_{it}} + \sum_{t=1}^{n} \log(q(r_{1t}, r_{2t})).
\]

Score function is

\[
S_n(\theta) = \frac{\partial l}{\partial \theta},
\]

\[
\frac{1}{\sqrt{n}} S_n(\theta) \rightarrow N(0, I(\theta_0)).
\]

**Theorem 6.** For the CARR model as \( n \rightarrow \infty \)

\[
\sqrt{n}(\hat{\theta} - I(\theta_0)^{-1}),
\]

where \( I(\theta_0)^{-1} = E \left( \frac{\partial^2 l(\theta_0)}{\partial \theta \partial \theta^T} \right) \).
First and second order partial derivatives are calculated in order to prove consistency of MLE. Assuming parameter vector \( \theta = (d_j, b_j, \omega) \), \( j = 1, 2 \) and based on condition of stationarity and boundedness of parameter \( \omega \), \( E_{sup_{\omega \in \Theta}} \| \frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} \|_1 \leq \infty \).

Proof are given in Appendix. Moreover using Taylor series expansion for \( \frac{\partial l(\theta)}{\partial \theta} \) we have:

\[
0 = \frac{\partial l(\hat{\theta})}{\partial \theta} = \frac{\partial l(\theta_0)}{\partial \theta} + (\hat{\theta} - \theta_0) \frac{\partial^2 l(\hat{\theta})}{\partial \theta \partial \theta^T} \text{ or } \frac{1}{\sqrt{n}} \frac{\partial l(\theta_0)}{\partial \theta} = \frac{1}{\sqrt{n}} (\hat{\theta} - \theta_0) \left( -\frac{1}{n} \frac{\partial^2 l(\theta)}{\partial \theta \partial \theta^T} \right).
\]

Considering Lemma 6 of Ting Lee (1996) \( -\frac{1}{n} \frac{\partial^2 l(\theta_0)}{\partial \theta \partial \theta^T} \rightarrow I(\theta_0) \) and by using martingale central limit theorem \( \frac{1}{\sqrt{n}} \frac{\partial l(\theta_0)}{\partial \theta} \rightarrow N(0, I(\theta_0)) \). Based on the above \( \sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N(0, I(\theta_0)^{-1}) \)

### 5.3 Estimation

Taking into account (3.7) and considering that \( q_i(r_i) \) defined as (3.26) for two exponentially distributed random variables with parameters \( \frac{1}{\lambda_{it}}, j = 1, 2 \), bivariate Exponential distribution defined as:

\[
f(r_{1t}, r_{2t}) = \prod_{i=1}^{2} \frac{1}{\lambda_{it}} e^{-\frac{R_{it}}{\lambda_{it}}} q(r_{1t}, r_{2t}),
\]

where

\[
q(r_{1t}, r_{2t}) = 1 + \omega \prod_{i=1}^{2} \left( c_i^{-k_{i}r_{it}} - \frac{1}{1 + \lambda_{it} k_{i} \log c_i} \right),
\]

The score function is given by:

\[
l(\theta; r) = -\sum_{i=1}^{2} \sum_{t=1}^{n} \log \lambda_{it} - \sum_{i=1}^{2} \sum_{t=1}^{n} \frac{R_{it}}{\lambda_{it}} + \sum_{t=1}^{n} \log(q(r_{1t}, r_{2t})).
\]

Partial derivatives are given in the Appendix. Considering conditions of stationarity given by Theorem 1 and bounds for parameter \( \omega \) given by (3.30; 3.31) assumptions for estimators consistency are satisfied. The score function is given by

\[
S_n(\theta) = \sum_{i=1}^{2} \sum_{t=1}^{n} \left( \frac{R_{it}}{\lambda_{it}^2} - \frac{1}{\lambda_{it}} \right) \frac{\partial \lambda_{it}}{\partial \theta} + \sum_{t=1}^{n} \frac{1}{q(r_{1t}, r_{2t})} \frac{\partial q(r_{1t}, r_{2t})}{\partial \theta},
\]

where \( \hat{\theta} = \arg \max_\theta l(\theta) \).
5.4 Goodness of fit test

Consider observations $r_{it}$, $i = 1, 2$, $t = 1, \ldots, n$ and under the assumption that residuals $\epsilon_{it}$ come from a unit exponential distribution we test the null hypothesis that bivariate distribution is a bivariate exponential distribution defined by (5.10) and (5.11). Jiménez-Gamero et al. (2020) proposed a GOF test for residuals considering the classical GARCH model. The idea is based on characterization that $E(X | Y) = 0$ holds if and only if $E(X e^{iuy}) = 0$. Meintanis et al. (2020) proposed GOF tests for the class of Autoregressive Conditional Duration (ACD) models in univariate case. Considering that residuals $\epsilon_t = x_t / E(x_t | F_t)$ are unobservable a test for the null hypothesis that residuals come from a unit exponential distribution has been examined. At this study there is a brief discussion about a Goodness of fit test for a bivariate Stochastic Conditional duration (SCD) model proposed by Bauwens and Veredas (2004). Meintanis et al. (2020) for the bivariate SCD model considered the characteristic function of the bivariate distribution and defined the distance statistic measure as:

$$K(u_1, u_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\hat{K}_T(u_1, u_2) - K_F(u_1, u_2)|^2 w(u_1, u_2) du_1 du_2.$$

Furthermore in a similar way Hudecová et al. (2021) discussed goodness of fit tests for a class of bivariate time series of counts considering the L2-type distance between two estimators of the probability generating function one nonparametric and one semiparametric under the same null hypothesis as an extension of the study in univariate case proposed by Meintanis and Karlis (2014). At all of the above test statistics the importance is concentrated in the asymptotic behavior of the statistic. At the proposed model where bivariate distribution defined by (5.10) the statistic based on the moment generating function defined as:

$$S = \int_{0}^{\infty} \int_{0}^{\infty} |\hat{M}_T(u_1, u_2) - M_F(u_1, u_2)|^2 w(u_1, u_2) du_1 du_2,$$

where

- $M_F(u_1, u_2) = \prod_{i=1}^{2} \frac{1}{\lambda_i(u_i-1/\lambda_i)} + \omega \left( \frac{1}{1-\lambda_i u_i + \lambda_i k_i \log c_i} - \frac{1}{\lambda_i + k_i \log c_i} \frac{1}{\lambda_i u_i - 1} \right)$.

- $\hat{M}_T(u_1, u_2) = \frac{1}{n} \sum_{t=1}^{n} e^{u_1 \epsilon_{1t} + u_2 \epsilon_{2t}}$.

- $w(u_1, u_2)$ is a weight function satisfying $0 \leq \int \int w(u_1, u_2) du_1 du_2 \leq 1$. 

The asymptotic distribution of the test statistic $S$ depends on several unknown quantities. For this reason we use a parametric bootstrap approach.

1. Generate $\epsilon^*_it$, $i = 1, 2$, where $\epsilon_it$ are iid and follow Exp(1).

2. Compute pseudo-observations $r^*_it, i = 1, 2, t = 1, \ldots, n$ from (5.1) where $\epsilon_t$ replaced by $\epsilon^*$ and $\lambda_t$ by $\hat{\lambda}_t$ where all parameters $a_{ii}$ and $b_{ii}$ replaced by their estimators.

3. Fit the model (5.1) using $r^*_it, i = 1, 2, t = 1, \ldots, n$ and compute the bootstrap estimator $\hat{\lambda}^*_t$ of $\lambda_t$.

4. Compute the corresponding test statistic $S^*$.

5. Repeat steps 1-4 several times, say B, and obtain the sequence of test statistics, $S^*_1, S^*_2, \ldots, S^*_B$.

6. Compute the p-value as $p = \frac{I(S^* - S)}{B}$ where $I =$ \begin{cases} 1, & S^* > S \\ 0, & Otherwise \end{cases}.

Asymptotic behavior of the test statistic implies asymptotic behavior of the MLE estimator as $Y_t, t \geq 0$ is strictly stationary, estimators converge to a Normal distribution given by theorem 4 and parametric space $\Theta$ is a compact set. Partial derivatives with respect to all parameters are presented below.

\[
\begin{align*}
\frac{\partial l_1(\theta; r)}{\partial d_j} &= \sum_{t=1}^{n} \left( -\lambda^{-1}_{jt} + R_{jt} \lambda^{-2}_{jt} \right) \frac{\partial \lambda_{jt}}{\partial d_j} \\
\frac{\partial l_2(\theta; r)}{\partial k_2} &= \sum_{t=1}^{n} \frac{1}{q(r_{1t}r_{2t})} \omega \left( c_1^{k_1r_{1t}} - c_2^{k_2r_{2t}} \log c_2 \right) \left( c_1^{k_1r_{1t}} - c_2^{k_2r_{2t}} \log c_2 \right) \\
\frac{\partial l_2(\theta; r)}{\partial k_2} &= -\sum_{t=1}^{n} \frac{1}{q(r_{1t}r_{2t})} \left( a_1^{k_1r_{1t}} - c_1 \log(c_1) \right) \left( a_2^{k_2r_{2t}} \log c_2 \right) \\
\frac{\partial \lambda_{jt}}{\partial a_j} &= \sum_{t=1}^{n} a_{jt}^{t-1}d_j + Td_j^{n-1} \lambda_{jt-t} + b_j \sum_{t=1}^{n-1} a_{jt-t-1}^{t-1} R_{jt-1-t-1} \\
\frac{\partial \lambda_{jt}}{\partial b_j} &= \sum_{t=1}^{n} d_j^{t-1} R_{jt-1} \\
\frac{\partial \lambda_{jt}}{\partial d_j} &= \sum_{t=0}^{n-1} a_j^{t}
\end{align*}
\]

5.5 Simulations

We obtain starting values of $l(\theta; r)$ by considering a bivariate ARMA(1,1) model similar to this defined in chapter 2 for the calculation of the unconditional moments see Ferland et al. (2006). Considering (3.27), we define initial value for parameter
5.5. SIMULATIONS

ω after obtaining initial values of other parameters and calculating volatilities. For data generation:

1. Choose initial values for the parameter vector \( \theta = c(a_{ij}, b_{ij}, \omega) \),

2. Generate a value for the univariate \( \text{Exp}(\lambda t) \), where \( \lambda t \) defined by (5.2),

3. Generate a \( u \sim U(0, 1) \),

4. If this \( u \leq f(r_2 \mid r_1) \) then let \( r_2 = u \).

<table>
<thead>
<tr>
<th>sample size(n)</th>
<th>50</th>
<th>200</th>
<th>500</th>
<th>50</th>
<th>200</th>
<th>500</th>
<th>50</th>
<th>200</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial value</td>
<td>mean</td>
<td>sd</td>
<td>MADE</td>
<td>mean</td>
<td>sd</td>
<td>MADE</td>
<td>mean</td>
<td>sd</td>
<td>MADE</td>
</tr>
<tr>
<td>( d_1 = 2.000 )</td>
<td>1.933</td>
<td>1.981</td>
<td>1.958</td>
<td>0.545</td>
<td>0.234</td>
<td>0.195</td>
<td>0.711</td>
<td>0.349</td>
<td>0.223</td>
</tr>
<tr>
<td>( d_2 = 1.500 )</td>
<td>1.578</td>
<td>1.479</td>
<td>1.516</td>
<td>0.394</td>
<td>0.189</td>
<td>0.123</td>
<td>0.357</td>
<td>0.243</td>
<td>0.146</td>
</tr>
<tr>
<td>( a_{11} = 0.300 )</td>
<td>0.378</td>
<td>0.319</td>
<td>0.294</td>
<td>0.209</td>
<td>0.118</td>
<td>0.087</td>
<td>0.256</td>
<td>0.157</td>
<td>0.102</td>
</tr>
<tr>
<td>( a_{22} = 0.450 )</td>
<td>0.386</td>
<td>0.391</td>
<td>0.398</td>
<td>0.134</td>
<td>0.062</td>
<td>0.048</td>
<td>0.137</td>
<td>0.081</td>
<td>0.062</td>
</tr>
<tr>
<td>( b_{11} = 0.200 )</td>
<td>0.204</td>
<td>0.199</td>
<td>0.201</td>
<td>0.119</td>
<td>0.056</td>
<td>0.045</td>
<td>0.144</td>
<td>0.073</td>
<td>0.063</td>
</tr>
<tr>
<td>( b_{22} = 0.150 )</td>
<td>0.181</td>
<td>0.166</td>
<td>0.152</td>
<td>0.135</td>
<td>0.088</td>
<td>0.066</td>
<td>0.118</td>
<td>0.086</td>
<td>0.075</td>
</tr>
<tr>
<td>( \omega = -0.200 )</td>
<td>1.992</td>
<td>1.767</td>
<td>1.827</td>
<td>0.919</td>
<td>0.553</td>
<td>0.317</td>
<td>1.054</td>
<td>0.537</td>
<td>0.402</td>
</tr>
</tbody>
</table>

Table 5.1: Mean, sd and MADE values after 200 replications of simulation experiment.

<table>
<thead>
<tr>
<th>sample size(n)</th>
<th>50</th>
<th>200</th>
<th>500</th>
<th>50</th>
<th>200</th>
<th>500</th>
<th>50</th>
<th>200</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial value</td>
<td>mean</td>
<td>sd</td>
<td>MADE</td>
<td>mean</td>
<td>sd</td>
<td>MADE</td>
<td>mean</td>
<td>sd</td>
<td>MADE</td>
</tr>
<tr>
<td>( d_1 = 3.000 )</td>
<td>3.035</td>
<td>2.989</td>
<td>2.995</td>
<td>0.450</td>
<td>0.274</td>
<td>0.239</td>
<td>0.518</td>
<td>0.313</td>
<td>0.299</td>
</tr>
<tr>
<td>( d_2 = 0.500 )</td>
<td>0.461</td>
<td>0.503</td>
<td>0.509</td>
<td>0.211</td>
<td>0.103</td>
<td>0.079</td>
<td>0.196</td>
<td>0.123</td>
<td>0.100</td>
</tr>
<tr>
<td>( a_{11} = 0.400 )</td>
<td>0.468</td>
<td>0.414</td>
<td>0.404</td>
<td>0.205</td>
<td>0.094</td>
<td>0.051</td>
<td>0.202</td>
<td>0.120</td>
<td>0.060</td>
</tr>
<tr>
<td>( a_{22} = 0.600 )</td>
<td>0.538</td>
<td>0.608</td>
<td>0.617</td>
<td>0.248</td>
<td>0.100</td>
<td>0.077</td>
<td>0.232</td>
<td>0.126</td>
<td>0.102</td>
</tr>
<tr>
<td>( b_{11} = 0.500 )</td>
<td>0.483</td>
<td>0.476</td>
<td>0.490</td>
<td>0.223</td>
<td>0.093</td>
<td>0.069</td>
<td>0.252</td>
<td>0.112</td>
<td>0.087</td>
</tr>
<tr>
<td>( b_{22} = 0.300 )</td>
<td>0.362</td>
<td>0.286</td>
<td>0.272</td>
<td>0.225</td>
<td>0.115</td>
<td>0.081</td>
<td>0.277</td>
<td>0.131</td>
<td>0.087</td>
</tr>
<tr>
<td>( \omega = 0.200 )</td>
<td>-0.100</td>
<td>-0.154</td>
<td>-0.172</td>
<td>0.518</td>
<td>0.313</td>
<td>0.257</td>
<td>0.655</td>
<td>0.316</td>
<td>0.260</td>
</tr>
</tbody>
</table>

Table 5.2: Mean, sd and MADE values after 200 replications of simulation experiment.
Figure 5.1: Boxplot with results from the simulation experiment with initial parameter values $(d_1, d_2, a_{11}, a_{22}, b_{11}, b_{22}, \omega) = (2, 1.5, 0.3, 0.45, 0.2, 0.15, 1.8)$

Figure 5.2: Boxplot with results from the simulation experiment with initial parameter values $(d_1, d_2, a_{11}, a_{22}, b_{11}, b_{22}, \omega) = (3, 0.5, 0.4, 0.6, 0.5, 0.3, 0.2)$
\[
\hat{\lambda}_{jt+1} = E(\lambda_{jt+1} | \mathcal{F}_{jt}) = \sum_{i=0}^{T} (a_j^i d_j + b_j a_j^i R_{jt-i}) + a_j^T \lambda_1.
\] (5.14)

5.6 Implementation

Many efforts to protect the climate provide creation of a variety of regulations. One way for the environment protection is the introduction and usefulness of Renewable sources. With renewable sources we mean wind and photovoltaic electricity generators which have an impact on energy market. This impact depend on weather conditions because when weather is rainy or solar energy is not enough there is scarcity of renewable sources and an augmentation of prices of electricity market is obtained. Firstly German policy makers introduced regulations to promote renewable sources but the issue of how Germany’s green energy policy has affected electricity market volatility (i.e., the question of whether such mechanisms are effective) has not been analysed so far. To avoid unstable increased market volatility Energy Exchange Law (Energiewirtschaftsgesetz, EnWG) and Renewable Energy Law (Erneuerbare-Energien-Gesetz, EEG), seek to promote renewables in a way that maintains electricity market stability. Auer (2016) are the first who proposed a pure electricity volatility model. GARCH models are inappropriate because they cannot capture the extraordinary characteristics of electricity prices and volatilities which have their origins in the non-storability of electricity requiring a constant balance between production and consumption and (ii) the dependence of electricity demand on the weather and the intensity of business and everyday activities. Moreover CARR models is designed to accommodate extreme values. It is worth mentioned that in Nordic European countries where climate is cold during winter hydroelectric power plants are activated for low prices in electricity . In contrast at Pennsylvania-New Jersey-Maryland(PJM) high and very volatile electricity prices are obtained during the summer because during the winter gas-fired plants are used to generate electricity. used weekly data about German electricity prices from Energy Exchange Market(EXX) for period of June 2000 to August 2015. In our study we use data from Hellenic Energy exchange group and Gestore dei Mercati Energetici SpA in Italy. Theoretically there is a correlation based on similarities of climate conditions(daytime maximum temperature and daily low temperature), humidity, yearly shiny hours, yearly rainy days and hours of daily light. At both countries high electricity values are observed at periods with maximum temperature at summer because electrical devices are very useful to decrease temperature. At
countless studies with energy economics trends and periodicity both at heteroskedastic and homoskedastic time series models are frequently observed. To eliminate trend models considering exogenous variables have been proposed. Narayan and Liu (2015) studied trend-based GARCH unit root model that outperforms a GARCH model without trend. Sin (2013) proposed a conditional autoregressive range model (CARRX) with exogenous variables in volatilities. At this application only data from September to April with size 264 are used in order to avoid any trend. The drawback obtained at forecasting because intradaily data are not used only a range of maximum and minimum daily values. For that reason we cannot predict specific values of energy. At the table below descriptive statistics are presented considering skewness and kurtosis. Based on the results the ranges at both countries are exponentially distributed. To estimate parameters $d_i, a_i, b_i$, $i = 1, 2$, we choose initial
5.6. IMPLEMENTATION

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Sd</th>
<th>Skewness</th>
<th>Kyrtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greece</td>
<td>76.499</td>
<td>56.774</td>
<td>1.627</td>
<td>5.531</td>
</tr>
<tr>
<td>Italy</td>
<td>59.989</td>
<td>41.585</td>
<td>1.534</td>
<td>4.629</td>
</tr>
</tbody>
</table>

Table 5.3: Descriptive statistics for the time series of electricity price ranges

<table>
<thead>
<tr>
<th></th>
<th>(d_1)</th>
<th>(d_2)</th>
<th>(a_{11})</th>
<th>(a_{22})</th>
<th>(b_{11})</th>
<th>(b_{22})</th>
<th>(\omega)</th>
</tr>
</thead>
<tbody>
<tr>
<td>estimate</td>
<td>6.810</td>
<td>9.760</td>
<td>0.678</td>
<td>0.572</td>
<td>0.181</td>
<td>0.211</td>
<td>1.687</td>
</tr>
<tr>
<td>se</td>
<td>0.022</td>
<td>0.024</td>
<td>0.019</td>
<td>0.018</td>
<td>0.029</td>
<td>0.024</td>
<td>0.037</td>
</tr>
</tbody>
</table>

| Log-Lik= | -3702.26 |

Table 5.4: ML estimates and se for model parameters.

values based on ARIMA model described at the beginning of the chapter and for dependence parameter we take into account boundaries given by (3.30; 3.31).

To conclude with this chapter, a new bivariate CARR time series model has been studied based on copulas defined by (5.10). While there are few studies were bivariate and trivariate CARR processes have been provided, difficulties with other properties like ergodicity or parameters estimation can be found. In our study the structure of bivariate exponential distribution offers the opportunity to study those properties for this model. The main drawback with all of those studies is the way of prediction. To predict a value at this model data observations during a time period are very useful. The fact that variable \(R_t\) is a range between maximum and minimum values without considering values in between this interval leads to appropriateness for forecasting. An implementation using daily electricity prices has been studied including data from Greece and Italy. This aforementioned model addressed to researchers who want to study financial data and the shortcomings are that financial data usually have periodicity and trends. For those reasons time series processes at finance could be used to model data for certain time periods dropping out in this way the above problems.
Chapter 6

Main contribution & Future challenges

6.1 Main contribution

We have developed a trans-dimensional MCMC approach for model selection between a family of Poisson INGARCH type models including linear and log-linear models. This allows to fit the models and select the one that best fits the data. Model selection for discrete valued time series is a topic of less research and we think that we contribute towards this. Our approach is based on satisfying the stationarity conditions, i.e. assuming that the series are stationary. Allowing for combinations of parameters beyond stationarity is possible but the interpretation of the models is more difficult afterwards.

The method that has been discussed by Carlin and Chib (1995) is the most flexible method in our case when we have a number of nested models. This is an important issue because since now there are limited studies about trans-dimensional methods for count data.

Furthermore we note that, although we only analyzed methodology by using linear and log-linear INGARCH models when all parameters are positive, our analysis can be extended when parameters are negative in case of log-linear INGARCH models. In addition the use of other functions for the mean process $\lambda_t$ presented by Fokianos et al. (2009); Fokianos and Tjøstheim (2011) could be examined for computational efficacy.

Additionally we treated only Poisson INGARCH models. However the method can be easily extended to the case of other models of the same type with different conditional distributions that have appeared in the literature.

We contributed at the construction of multivariate discrete and continuous dis-
tributions that offer strong correlation based on the new structure of Sarmanov distribution. We only discussed this family assuming properties of p.m.f and p.d.f. and we proposed the closed formula of a copula without taking into account any study about copula’s properties. We presented boundaries of correlation coefficient at the case of Poisson and exponential distributions providing the flexible way of use in both cases. While Miravete (2009b) proposed a multivariate Sarmanov distribution, in our study an alternative construction of this distribution is provided. The advantage is that this multiplicative factor consider only products between all combinations of appropriate functions each time. This decreases computational cost and theoretical properties are more easier to study.

While a multivariate Poisson and a multivariate Exponential time series models have been structured and theoretical properties have been discussed, the proposed Sarmanov could be used for the study of the above multivariate distributions. Moreover computational cost for parameters estimation could be decreased if maximization by parts method proposed by Song et al. (2005) implemented. Poisson INGARCH model dealt with the problem of equidispersion and for this reason less datasets can be studied with this model. In univariate case INGARCH models of Binomial and Negative Binomial distributions have been studied. Mimic those models and based on the new structure of copula, other multivariate INGARCH models where processes $X_t$ are simulated for other discrete distributions could be provided.

In case of exponential distribution homoskedastic time series processes are difficult to study. Assuming that innovations simulated from exponential distributions and using characteristic functions, one can proved that process $X_t$ can not simulated from an exponential distribution and this leads to more complicated structures. In contrary in case of data with heteroskedasticity, models with exponential distribution can be constructed without obstacles. Based on the idea of Bollerslev et al. (1994), an exponential CARR model has been developed and an application with data of electricity energy sources has been provided.

The model becomes complicated when interdependencies can be considered at volatilities expressions. One can observe difficulties not only with the study of theoretical properties like ergodicity but also with the intensive computational way of parameters estimation. At both studies weak dependence criteria have been discussed for studying ergodic properties but other theoretical tools could be also used to provide stationarity at $p^{th}$-order moments where $p \geq 2$ as future work.
6.2 Future challenges

Several extensions of Bivariate INGARCH or Multivariate INGARCH models can enhance the research. Studies of multivariate INGARCH models where volatilities will be expressed by another nonlinear functions could be produced. Log-linear INGARCH model has been also constructed in one dimension but the construction in more dimensions is a topic of interest. One can study the model when

\[ X_t \sim BP(\lambda_t), \]

where \( \lambda_t = d + A\lambda_{t-1} + BY_{t-1}. \)

BP is a bivariate Poisson distribution, \( A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \) and \( B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}. \) Assuming the constructions above, positive and negative correlations and cross-correlations can occur and interdependencies can be captured to the model. Furthermore INGARCH\((p,q)\) models with incorporation of other discrete distributions can be account. For example Zhu (2011), Zhu (2012c); Zhu (2012b) studied univariate Negative Binomial and Compound Poisson INGARCH models and Miravete (2009b) studied bivariate Negative Binomial distribution without considering time series processes. Based on the above study one can develop bivariate Negative Binomial and Compound Poisson INGARCH processes to model data with over or under-dispersion. In our study we propose a trivariate INGARCH\((1,1)\) model defined by (2.1), (5.2) but one may extend this model in \( n \) dimensions \( n \geq 3 \) where volatilities could be expressed by:

\[ \lambda_t = d + A \sum_{i=1}^p \lambda_{t-i} + B \sum_{j=1}^q Y_{t-j}, \]

where

\( d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \) \( A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}, \)

and \( B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \ddots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} \end{bmatrix}. \)

Stationarity could be provided using other theoretical approaches except weak dependence criteria but taking into account how many weak dependence criteria there
are, stationarity of low order moments could be also provided in cases where more theoretical approaches are difficult to proceed.

Moreover the contribution to the new bivariate distribution given by (5.1) is the definition of another new where one variable will be discrete and the other will be continuous. At this case we will have:

\[ f(x_1, x_2) = f_1(x_1)f_2(x_2)[1 + \omega q_1(x_1)q_2(x_2)], \]

where:

\[ q_1(x_1) = c_1^{-k_1x_1} - \sum_{x_1=0}^{\infty} c_1^{-k_1x_1} f_1(x_1) \]

and

\[ q_2(x_2) = c_2^{-k_2x_2} - \int_{x_2=0}^{\infty} c_2^{-k_2x_2} f_2(x_2).dx_2. \]

This is a challenging problem because there are not many studies with distributions of this form to account for. The problem becomes more intricated when a multivariate distribution is going to constructed and multivariate time series models could be defined without assuming interdependencies. The transformation of this bivariate distribution to a copula is:

\[ C(u, v) = uv \left[ 1 + \omega \left( \sum_{t_1=0}^{F^{-1}(u)} c_1^{-k_1t_1} - uLT_1(f(x_1)) \right) \left( \sum_{t_2=0}^{F^{-1}(v)} c_2^{-k_2t_2} - vLT_2(f(x_2)) \right) \right]. \]

Considering the appropriate marginals, the construction is easy to derived calculating the corresponding Laplace transforms. While the formula has been given and at chapter 3 a copula with exponential marginals was proposed, properties to be a copula is an aforementioned problem.

At CARR time series model, studies with data from financial and econometrics fields have been implied. The crucial problem with this family of models is to find a way to forecast values considering not only the maximum and minimum values of a period but also values in between those periods. The difficulty stems from volatility’s expression. Moreover Chou and Wang (2006) compared the performance of an Exponential CARR model and a GARCH model. An exponential CARR model with the incorporation of exogenous variables has been also appeared at this study. A model similar to this could be studied in order to decrease model’s order and in general to improve it’s accuracy.

\[ R_t \sim BExp(\lambda_t), \]
where

$$\lambda_t = d_1 + A\lambda_{t-1} + BR_{t-1} + \Gamma Y_{t-1}$$

and $Y_t$ could be simulated from any continuous distribution.
Chapter 7
Appendix

7.1 Appendix A: Bivariate distribution based on copulas

Proof of bivariate Poisson distribution:

\[ f(x_1, x_2) = \prod_{i=1}^{2} f_i(x_i) \left[ 1 + \rho \prod_{i=1}^{2} \frac{\sigma_i q_i(x_i)}{v_i} \right] = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} \frac{\sigma_i f_i(x_i) c_i^{-k_i x_i}}{v_i} \]

\[ - \rho \prod_{i=1}^{2} \frac{\sigma_i f_i(x_i)}{v_i} \left[ e^{-k_1 x_1} e^{\lambda_2 (c_2^{-k_2} - 1)} + c_2^{-k_2 x_2} e^{\lambda_1 (c_1^{-k_1} - 1)} \right] \]

\[ + \rho \prod_{i=1}^{2} P_i(x_i) - \rho D_1(x_1) P_2(x_2) - \rho P_1(x_1) D_2(x_2) + \rho \prod_{i=1}^{2} D_i(x_i) \]

\[ = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} \frac{[f_i(x'_i) - f_i(x_i)]}{\sqrt{\lambda_i (c_i^{-k_i} - 1)}}, \quad (7.1) \]

where

\[ D_i(x_i) = \frac{f_i(x_i)}{\sqrt{\lambda_i (c_i^{-k_i} - 1)}}, \quad P_i(x_i) = \frac{f_i(x'_i)}{\sqrt{\lambda_i (c_i^{-k_i} - 1)}}, \quad f_i(x'_i) \sim \text{Poisson}(\lambda_i c_i^{-k_i}). \]

Proof of bivariate Exponential distribution:

\[ f(x_1, x_2) = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} \frac{\sigma_i f_i(x_i) c_i^{-k_i x_i}}{v_i} + \rho \prod_{i=1}^{2} \frac{\sigma_i f_i(x_i) Q(\theta)}{v_i} + \rho \prod_{i=1}^{2} \frac{\sigma_i f_i(x_i)}{v_i} \frac{\lambda_i}{\log(e^{\lambda_i c_i^{k_i}})} \]

\[ = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} P_i(x_i) - \rho D_1(x_1) P_2(x_2) - \rho \prod_{i=1}^{2} P_1(x_1) D_2(x_2) + \rho \prod_{i=1}^{2} D_i(x_i) \]

\[ = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} \frac{\log(e^{\lambda_i c_i^{k_i}})}{k_i \log c_i} [f_i(x'_i) - f_i(x_i)], \quad 125 \]
where

\[ Q(\theta) = \left[ c_1^{-k_1 x_1} \frac{\lambda_2}{\lambda_2 + k_2 \log(c_2)} + \frac{\lambda_1}{\lambda_1 + k_1 \log c_1} c_2^{-k_2 x_2} \right], \]

\[ P_i(x_{i,t}) = \log \left( \frac{(e^{\lambda_i c_i^k})}{k_i \log c_i} f_i(x_i) \right), \]

\[ D_i(x_{i,t}) = \log \left( \frac{(e^{\lambda_i c_i^k})}{k_i \log c_i} f_i(x_i) \right), \]

\[ f_i(x_i) \sim \text{Exp}(\lambda_i + k_i \log c_i). \]

### 7.2 Appendix B: Partial derivatives of INGARCH process

Assuming parameter vector \( \theta_i = (\alpha_{ij}, b_{ij}, d_i, \omega_{ij}) \), \( i, j = 1, 2, 3 \) first order partial derivatives with respect to parameter vector \( \theta_i \) and second order partial derivatives with respect to \( b_{ij}, \omega_{ij} \) are given accordingly by:

\[
\frac{\partial l(\theta_i)}{\partial d_i} = \sum_{t=1}^{n} (Q_t(\theta)) \frac{\partial \lambda_i(\theta_i)}{\partial d_i},
\]

\[
\frac{\partial l(\theta_i)}{\partial b_{ij}} = \sum_{t=1}^{n} (Q_t(\theta)) \frac{\partial \lambda_i(\theta)}{\partial b_{ij}},
\]

\[
\frac{\partial l(\theta_i)}{\partial a_{ij}} = \sum_{t=1}^{n} (Q_t(\theta)) \frac{\partial \lambda_i(\theta)}{\partial a_{ij}},
\]

\[
\frac{\partial l(\theta_i)}{\partial \omega_{ij}} = \sum_{t=1}^{n} \left( \frac{1}{\phi_t(\theta)} R_i(x_{i,t}) R_j(x_{j,t}) \right) \frac{\partial \lambda_i(\theta)}{\partial \omega_{ij}},
\]

\[
\frac{\partial^2 l(\theta)}{\partial \omega_{ij}^2} = -\sum_{t=1}^{n} \frac{R_i^2(x_{i,t}) R_j^2(x_{j,t}) \partial \lambda_i^2(\theta)}{\phi_t^2(\theta)} \frac{\partial \lambda_i(\theta)}{\partial \omega_{ij}}.
\]
7.3. APPENDIX C: PARTIAL DERIVATIVES OF CARR PROCESS

\[
\frac{\partial^2 l(\theta)}{\partial b_{ij}^2} = \sum_{t=1}^{n} \left( \sum_{i=1}^{n} \left( \frac{-X_{i,t}}{\lambda_i^2(\theta)} \left( \frac{\partial \lambda_i(\theta)}{\partial b_{ij}} \right)^2 \right) + \sum_{i=1}^{n} \left( \frac{X_{i,t}}{\lambda_i(\theta)} - 1 \right) \left( \frac{\partial^2 \lambda_i(\theta)}{\partial b_{ij}^2} \right) \right)
\]
\[+ \sum_{t=1}^{n} \frac{1}{\phi_t(\theta)} \left( \sum_{i<j} \omega_{ij} \left( c_{ij} - 1 \right) D_i(x_{i,t}) R_j(x_{j,t}) \left( \frac{\partial \lambda_i(\theta)}{\partial b_{ij}} \right)^2 \right) - \left( \omega_{ij} \left( c_{ij} - 1 \right) D_j(x_{j,t}) R_i(x_{i,t}) \left( \frac{\partial \lambda_j(\theta)}{\partial b_{ij}} \right)^2 \right) + 2D_i(x_{i,t}) D_j(x_{j,t}) \left( \frac{\partial \lambda_i(\theta)}{\partial b_{ij}} \right) \left( \frac{\partial \lambda_j(\theta)}{\partial b_{ij}} \right)
\]
\[+ \frac{1}{\phi_t^2(\theta)} \left( \sum_{i<j} \omega_{ij}^2 D_i^2(x_{i,t}) R_j^2(x_{j,t}) \left( \frac{\partial \lambda_i(\theta)}{\partial b_{ij}} \right)^2 \right)
\]
\[+ \left( \omega_{ij}^2 D_j^2(x_{j,t}) R_i^2(x_{i,t}) \left( \frac{\partial \lambda_j(\theta)}{\partial b_{ij}} \right)^2 \right) + 2\omega_{ij} D_i(x_{i,t}) D_j(x_{j,t}) R_i(x_{i,t}) R_j(x_{j,t}) \right),
\]

where

\[
R_t(x_{i,t}) = c_{i}^{-k_i x_{i,t}} - e^{\lambda_i(c_{i}^{-k_i} - 1)},
\]
\[
D_t(x_{i,t}) = e^{\lambda_i(c_{i}^{-k_i} - 1)} \left( c_{i}^{-k_i} - 1 \right),
\]
\[
Q_t(\theta) = \sum_{i=1}^{n} \left( \frac{X_{i,t}}{\lambda_i(\theta)} - 1 \right) - \frac{1}{\phi_t(\theta)} \sum_{i<j} \left( \omega_{ij} e^{\lambda_i(c_{i}^{-k_i} - 1)} R_t(x_{i,t}) + \omega_{ij} R_j(x_{j,t}) e^{\lambda_j(c_{j}^{-k_j} - 1)} \right).
\]

Based on equation (1.14) derivatives of $\lambda_{it}$, $i = 1, 2$ with respect to parameter vector $\theta_{it}$, $i = 1, 2$ defined previously are given by:

\[
\frac{\partial \lambda_i(\theta)}{\partial d_i} = d_i + b_{ij} \frac{\partial \lambda_{i,t-1}}{\partial d_i},
\]
\[
\frac{\partial \lambda_i(\theta)}{\partial b_{ij}} = \lambda_{j,t-1} + b_{ij} \frac{\partial \lambda_{j,t-1}}{\partial b_{ij}},
\]
\[
\frac{\partial \lambda_i(\theta)}{\partial a_{ij}} = \lambda_{j,t-1} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial a_{ij}}.
\]

7.3 Appendix C: Partial derivatives of CARR process

Considering bivariate CARR model defined by (5.1) and volatilities defined by (5.2) first and second order partial derivatives with respect to parameters $a_i$ and $\omega$ are
given by:

\[
\frac{\partial l(\theta)}{\partial c_i} = \sum_{t=1}^{n} \omega \left( c_j^{-k_j r_{jt}} - \frac{1}{1 + \lambda_{it} k_j \log c_j} \right) \frac{k_i \log c_i}{(1 + \lambda_{it} k_i \log c_i)^2} \frac{1}{q(r_{1t}, r_{2t})} \frac{\partial \lambda_{it}}{\partial c_i} \\
+ \sum_{t=1}^{n} \left( \frac{\lambda_{it} + R_{it}}{\lambda_{it}^3} \right) \frac{\partial \lambda_{it}}{\partial c_i},
\]

\[
\frac{\partial l(\theta)}{\partial \omega} = \sum_{t=1}^{n} \prod_{i=1}^{2} \left( \frac{c_i^{-k_i r_{it}} - \frac{1}{1 + \lambda_{it} k_i \log c_i}}{q(r_{1t}, r_{2t})} \right),
\]

\[
\frac{\partial^2 l(\theta)}{\partial c_i^2} = \sum_{t=1}^{n} \left( \frac{2 \omega k_i \log c_i q(r_{1t}, r_{2t})}{(1 + \lambda_{it} k_i \log c_i)^{\frac{3}{2}}} - \frac{\omega c_j^{-k_j r_{jt}}}{(1 + \lambda_{it} k_j \log c_j)^{\frac{3}{2}}} \right) \frac{\partial \lambda_{it}}{\partial c_i} \\
+ \sum_{t=1}^{n} \left( \frac{2 \omega k_i \log c_i q(r_{1t}, r_{2t})}{(1 + \lambda_{it} k_j \log c_j)^4} - \frac{\omega c_j^{-k_j r_{jt}}}{(1 + \lambda_{it} k_j \log c_j)^{\frac{5}{2}}} \right) \frac{\partial \lambda_{it}}{\partial c_i} + \left( \frac{\lambda_{it} + R_{it}}{\lambda_{it}^3} \right) \frac{\partial^2 \lambda_{it}}{\partial c_i^2} \\
+ \sum_{t=1}^{n} \omega k_i \log c_i \left( c_j^{-k_j r_{jt}} q(r_{1t}, r_{2t}) - \frac{1}{1 + \lambda_{jt} k_j \log c_j} \right) \frac{\partial^2 \lambda_{it}}{\partial c_i^2},
\]

\[
\frac{\partial^2 l(\theta)}{\partial \omega^2} = -\sum_{t=1}^{n} \prod_{i=1}^{2} \left( \frac{c_i^{-k_i r_{it}} - \frac{1}{1 + \lambda_{it} k_i \log c_i}}{q^2(r_{1t}, r_{2t})} \right)^2.
\]

where \( q(r_{1t}, r_{2t}) \) given by (5.11) and partial derivatives with respect to \( d_i \) and \( b_i \) are calculated in a similar way.


Engle III, R. F. and K. Sheppard (2001). Theoretical and empirical properties of
dynamic conditional correlation multivariate GARCH.

Farlie, D. J. (1960). The performance of some correlation coefficients for a general

positive innovations. *Communications in Statistics. Stochastic Models* 8(3), 685–
717.


197.

of the American Statistical Association* 104(488), 1430–1439.


of Multivariate Analysis* 102(3), 563–578.

Fokianos, K. and D. Tjøstheim (2012). Nonlinear poisson autoregression. *Annals of
the Institute of Statistical Mathematics* 64(6), 1205–1225.

Fonseca, G. and R. Tweedie (2002). Stationary measures for non-irreducible non-

(1, 1) models without moment assumptions on the observed process. *Econometric
Theory* 22(5), 815–834.

Frank, M. J. (1979). On the simultaneous associativity off (x, y) andx + y- f (x, y).
*Aequationes mathematicae* 19(1), 194–226.


