MIXTURE INNOVATIONS' BASED AUTO-REGRESSIVE PROCESSES WITH APPLICATION TO SEA LEVEL RISE DATA

By

Fatimah Alshahrani

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Statistics - Doctor of Philosophy

2020

ABSTRACT

MIXTURE INNOVATIONS' BASED AUTO-REGRESSIVE PROCESSES WITH APPLICATION TO SEA LEVEL RISE DATA

By

Fatimah Alshahrani

Sea Level Rise (SLR) is one of the major challenges the world is facing in times of climate change. Using different approaches in obtaining, studying and modeling SLR data has often resulted in different conclusions for the same data set. This led to a debate between researchers some of whom reported accelerations whereas others advocated decelerations or linear trend. In this project, we model SLR as a linear mean reverting time series, also known as order one autoregressive processes (AR(1)), together with an in-depth study of the innovations associated with the AR process. In this direction, we define and evaluate moment ratio test to distinguish between innovations which come either from a Gaussian distribution, a Gumbel distribution or a mixture of the two. We were motivated by a claim from members of the Actuarial Climate Index (ACI) collaborative regarding the distribution of sea-level rise in the US Atlantic states. They claim that the monthly data, recorded over decades, have no inherent time structure and can be modeled as independent and identically observations from Gumbel distribution. Based on an exploratory data analysis, fitting a stationary, Markovian, mean-revering process to the data was essential, to understand its structure. Since the Ornstein–Uhlenbeck process (OU) is the high-frequency liming process for AR(1) processes, and our data has a fixed frequency of observations, we investigated fitting an AR(1) process where the innovations are either Gumbel noise, Gaussian noise, or a mixture of the two distributions. The result of these fittings, based on our moment-ratio tests and other diagnostic checks, demonstrates the preference for an AR(1) process with innovations from a mixture of Gaussian and Gumbel distributions.

ACKNOWLEDGMENTS

I am very grateful to my advisor and my committee Co-Chair Prof. Frederi Viens for his endless support, encouragement and patience. We had so many enlightened meetings and discussions. Prof. Viens introduced me like others to Malliavin Calculus, which is a big help in proving many theorems in Mathematical Statistics. He encouraged me to participate, attend and give talks in conferences and seminars. Prof. Viens was very supportive during my three years studying at MSU.

I also would like to thank my committee Co-Chair Prof. Shrijita Bhattacharya for her support and time. She and I had exchanged emails and held meetings together for two years while working on this project. Throughout our meetings, I learnt many techniques in Time Series Analysis and how to implement them.

I also would like to extend my gratitude to my committee members Prof. Alexandra Kravchenko and Prof. Lyudmila Sakhanenko for accepting to serve on my committee and their valuable discussion. In the beginning of this project we had had couple of meetings with Mr. Steve Kolk from The Actuarial Climate Index, and he had explained to us the data and some of their work. So thank you so much Mr. Kolk; the discussions with you made it easier for us to find the data and analyze it.

All of my study here at MSU was financially supported by Princess Nora bint Abdul Rahman University (PNU) through The Saudi Arabian Cultural Mission (SACM), so thank you so much PNU and SACM.

To my family, my mom and dad, to my husband and kids, I am very thankful for your support, love and believing in me.

TABLE OF CONTENTS

| LIST OF TABLES | vi |
|--|------|
| LIST OF FIGURES | viii |
| Introduction | 1 |
| autoregressive models | 5 |
| Studying and modeling the sea level data | 14 |
| The moments of the Gumbel random variable and the theoretical moments | |
| of the $AR(1)$ model | 21 |
| The Moments of the Gumbel Random Variable | 21 |
| The theoretical moments of the $AR(1)$ model \ldots \ldots \ldots \ldots \ldots \ldots \ldots The asymptotic variances and covariances of the empirical moments of the $AR(1)$ | 29 |
| model | 37 |
| The asymptotic distributions of $\widehat{m}_2, \ \widehat{m}_3$ and \widehat{m}_4 and the non-central chi- | |
| square S and K tests \ldots \ldots \ldots \ldots \ldots \ldots | 48 |
| The asymptotic normality of \widehat{m}_2 , \widehat{m}_3 and \widehat{m}_4 | 49 |
| The asymptotic normality of the empirical skewness and kurtosis | 66 |
| The non-central chi-square tests | 70 |
| The Simulation | 77 |
| The distributions of S_T and K_T | 77 |
| The non-central Chi-square test S and its empirical power $\ldots \ldots \ldots \ldots \ldots$ | 80 |
| The non-central chi-square test K and its empirical power $\ldots \ldots \ldots \ldots$ | 87 |
| The analysis of the sea level rise data | 93 |
| Data | 94 |
| Method | 95 |
| Results | 98 |
| Discussion | 100 |
| Conclusion | 107 |
| BIBLIOGRAPHY | 109 |

LIST OF TABLES

| Table 1: | Some non-Gaussian time series | 11 |
|-----------|---|----|
| Table 2: | The distribution of S_T using empirical moments and the moments of X_t when simulating from Gaussian AR(1) model | 78 |
| Table 3: | The distribution of S_T using empirical moments and using the moments of X_t when simulating from AR(1) model with Gumbel innovations \ldots | 78 |
| Table 4: | The distribution of K_T using empirical moments and the moments of X_t when simulating from Gaussian AR(1) model | 79 |
| Table 5: | The distribution of K_T using empirical moments and the moments of X_t when simulating from AR(1) model with Gumbel innovations $\ldots \ldots$ | 79 |
| Table 6: | Applying the non-central chi-square S test to reject the Gaussian innovations in favor of the Gumbel. | 81 |
| Table 7: | The empirical power of the non-central chi-square test S where $H_0: \xi_t \sim$ Gaussian vs $H_a: \xi_t \sim$ Gumbel | 82 |
| Table 8: | Applying the non-central chi-square S test to the hypothesis $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ Gaussian | 83 |
| Table 9: | The empirical power of the non-central chi-square test S where $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ Gaussian | 84 |
| Table 10: | Applying the non-central Chisquare test S when simulating from an AR(1) with mixture innovation to reject the Gumbel and the Gaussian innovations in favor of the mixture. | 85 |
| Table 11: | Applying the non-central Chisquare test S when simulating from an AR(1) with mixture innovation to reject the Gumbel and the Gaussian innovations in favor of the mixture innovations continuing. | 86 |
| Table 12: | The empirical power of the non-central chi-square test S where $H_0: \xi_t \sim$ Gumbel or Gaussian vs $H_a: \xi_t \sim$ mixture | 87 |
| Table 13: | Applying the non-central chi-square K test to reject the Gumbel innova- tions in favor of the Gaussian. | 89 |

| Table 14: | The empirical power of the non-central chi-square test K where $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ Gaussian | 90 |
|-----------|---|-----|
| Table 15: | Applying the non-central chi-square K test to reject the Gaussian innovations in favor of the Gumbel. | 91 |
| Table 16: | The empirical power of the non-central chi-square test K where $H_0: \xi_t \sim$ Gaussian vs $H_a: \xi_t \sim$ Gumbel | 92 |
| Table 17: | The estimators of $\hat{\phi}$, $\hat{\sigma}$ and \hat{p} when fitting AR(1) model with different innovations, to the SLR data, | 98 |
| Table 18: | The theoretical moments of AR(1) model when fitted with different innovations to the SLR data. | 98 |
| Table 19: | The asymptotic distribution of S_T and K_T when fitting the AR(1) model with different innovations to the SLR data | 99 |
| Table 20: | The confidence intervals of $m_3/m_2^{3/2}$ and m_4/m_2^2 applied to the SLR data | 99 |
| Table 21: | The results of applying the non-central chisquare test T to the SLR data under different null hypothesis scenarios | 100 |
| Table 22: | The p values of the S test applied to the SLR data with different scenarios. | 101 |
| Table 23: | The p value of KS test under the null hypothesis applied to the residuals of the SLR data with different scenarios. | 105 |

LIST OF FIGURES

| Figure 1: | Correlogram of the sea level data at the CEA region | 2 |
|-----------|--|-----|
| Figure 2: | The path of the de-trended data from 1990 to 2000 | 3 |
| Figure 3: | Map shows the 12 different regions according to ACI $\ldots \ldots \ldots$ | 94 |
| Figure 4: | The QQ-plot of the residuals of fitting a Gaussian AR(1) model to the STR data vs the Gaussian distribution | 103 |
| Figure 5: | The QQ-plot of the residuals of fitting an $AR(1)$ model with Gumbel innovations to the STR data vs the Gumbel distribution $\ldots \ldots \ldots$ | 104 |
| Figure 6: | The QQ-plot of the residuals of fitting an $AR(1)$ model with mixture innovations to the STR data vs the mixture distribution $\ldots \ldots \ldots$ | 104 |

Introduction

Sea Level Rise (SLR) is one of the major consequences the world is facing in times of climate change. Since the result of SLR involves social-economic response, regional and global SLR has been widely investigated and studied. Using different approaches in obtaining, studying and modeling SLR data resulted in different conclusions."This has led to an intensive debate on the existence of significant accelerations in regional and global sea level in recent years. Some researchers report accelerations where others find linear trends or even decelerations, sometimes based on the same data." [87]

Sea level data can be obtained either by climate models or measuring sea level, using methods such as tide gauge stations which is known as the historical observations or through satellite altimetry. According to sea level research group at university of Colorado, tide gauge data "offers the only source of historical, precise, long-term sea level data."

We obtained tide gauge data from The Actuarial Climate Index (ACI), and they claim that the monthly data, recorded over decades, of the Central East Atlantic (CEA) region follow the Gumbel distribution in the absence of any time structure. [5]. This claim is investigated here, considering the data as a random sample, and fitting the Gumbel distribution and applying the Kolmogorov–Smirnov test (KS test), we fail to reject the null hypothesis that the data as a random sample follows the Gumbel distribution. Looking at the auto-correlation function plot of the data in figure 1, we conclude that the assumption that the data is a "random sample" is not accurate.



Figure 1: Correlogram of the sea level data at the CEA region.

Figure 2shows the path of the data from 1990 to 2000, which looks like the path of the Ornstein–Uhlenbeck (OU) process. The OU process is a stationary, Gaussian, Markovian and mean reverting process. Since the Ornstein–Uhlenbeck process (OU) is the high-frequency limiting process for AR(1) processes, and our data has a fixed frequency of observations [30], we investigate fitting an AR(1) process where the innovations are Gumbel noise, Gaussian noise, or a mixture of both distributions.

Having three different models with the earlier debate in mind we defined a non-central chi square S and K tests based on the ratio of the empirical moments of the AR(1) model, such that these tests can reject the Gumbel innovations in favor of the Gaussian, the Gaussian innovations in favor of the Gumbel and the Gumbel or Gaussian innovations in favor of the mixture of both.

Employing the moments and the ratio of moments started in the 1930s, under the assumption of independent identically random variables. Many researchers were interested in checking the skewness, kurtosis and normality [23].



Figure 2: The path of the de-trended data from 1990 to 2000.

The result of fitting the three models to the SLR data, based on our non-central chi-square tests and other diagnostic checks that are based on the residuals, shows the preference for an AR(1) process where the innovations follow a mixture of Gaussian and Gumbel distributions.

This thesis is organized as follows: the related work in both parts the goodness of fit tests of the AR(1) model and modeling the sea level data in chapter 1. In chapter 2, the moments of the Gumbel random variable and the theoretical moments of the AR(1) model were calculated. The asymptotic distributions of the empirical second, third and fouth moments are provided in chapter 3, also the main results which are the defined non-central

chi-square S and K tests are in chapter 3. The simulations are in chapter 4, while the case study "the real data" in chapter 5.

The Literature review

The hypothesis tests that are related to the distribution of the innovations in the autoregressive models

The autoregressive process is one of the most successful statistical models for time series data due to its wide range of applications. An autoregressive process of order p(AR(p)), is written as:

$$X_{t} = \phi_{1}X_{t-1} + \phi_{2}X_{t-2} + \dots + \phi_{p}X_{t-p} + \epsilon_{t}$$
(1)

where ϵ_t is white noise (WN(., .)), i.e $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2$ and ϵ_t is uncorrelated with ϵ_s for each s < t. Autoregressive Processes are a sub-class of a more general linear class called Autoregressive Moving Average Process, which for orders p and q, (ARMA(p,q)), has the form:

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

$$\tag{2}$$

where $\epsilon_t \sim WN(0, \sigma^2)$ and the polynomial $(1-\phi_1 z - ... - \phi_p z^p)$ and $(1+\theta_1 z + ... + \theta_q z^q)$ have no common factors for identifiability. It is further assumed that X_t is stationary and invertible. Let *B* be the backward shift operator, $(B^j X_t = X_{t-j}, B^j \epsilon_t = \epsilon_{t-j}, j = 0, \pm 1, \pm 2, ...)$, then 2 can be written as

$$\phi(B)X_t = \theta(B)\epsilon_t \tag{3}$$

where $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + ... + \theta_q z^q$. X_t is an AR(p) if $\theta(z) \equiv 1$, and a moving average process of order q, (MA(q)) if $\phi(z) \equiv 1$. It is further assumed that X_t is a stationary and invertible. The necessary and sufficient condition for stationarity is that for the polynomial $\phi(z) \neq 0$ for all complex z with |z| = 1. Invertibility is equivalent to the condition $\theta(z) \neq 0$ for all $|z| \leq 1$.

In most applications ϵ_t is assumed to be Gaussian, hence if the length of realization is short, exact likelihood estimation can be used to estimate the parameters ϕ_i , i=1,2,...,p; θ_j , j=1,2,...,q. Alternatively, they can be asymptotically efficiently estimated by maximizing the conditional log-likelihood.

Statistical analysis of time series started in the 1920s. In 1927, Udny Yule fitted an AR(2) to Sunspots data [69, 84], making this the first application of autoregressive models to real data. Yule introduces the concept of corrolgram "serial correlation coefficient," though he did not invent the name according to [66]. Anderson [3] stated "Both Yule [94] and Bartlett [6] have shown that the ordinary test of significance are invalidated if successive observations are not independent of one another. The serial correlation coefficient has been introduced as a measure of the relationship between successive values of a variable ordered in time or space". This statement indicates a fundamental and standard approach in diagnostic checks in time series models.

In the early 1970s, with the computer progression and the so-called "Box–Jenkins method", which is a detailed strategy for time series forecasting, The ARMA process was studied and investigated widely [58]. The Box-Jenkins method is a method for analyzing discrete time series data, incorporating an iterative cycle of identification, estimation and verification [12, 2]. In the first step, based on the information on the sample path, a preliminary ARMA model is suggested. In the second step the estimation of a stationary ARMA model can be done using the exact or approximation methods. The third step is model diagnostic checking, which involves techniques like overfitting, residual plots and a group of tests that are preformed on the residuals.

A good model in time series data should be able to capture the dependence structure of the data appropriately. That is, checking the whiteness of the residuals, since a good time series model should produce residuals that are approximately white noise. Box and Jenkins [12] employ Cramer statistics [24] g_1 and g_2 for testing the normality of the residuals. In this frame, these are the residuals skewness and kurtosis defined respectively as:

$$K_3 = \left(\frac{1}{n}\sum_{t=1}^n \hat{\epsilon}_t^3\right) / \left(\frac{1}{n}\sum_{t=1}^n \hat{\epsilon}_t^2\right)^{3/2}$$

and

$$K_4 = \left(\frac{1}{n}\sum_{t=1}^n \widehat{\epsilon}_t^4\right) / \left(\frac{1}{n}\sum_{t=1}^n \widehat{\epsilon}_t^2\right)^2 - 3$$

Under the assumption of normality and if the model is correct, $K_3 \sim AN(0, 6/n)$ and $K_4 \sim AN(0, 24/n)$ [58, 39].

As mentioned by earlier statisticians, one important measurement of dependence is the autocorrelation function. This led to using the residual autocorrelation function \hat{r}_k as a test statistic, defined as follow:

$$\widehat{r}_k = \sum_{t=k+1}^n \widehat{\epsilon}_t \widehat{\epsilon}_{t-k} \Big/ \sum_{t=1}^n \widehat{\epsilon}_t^2, \tag{4}$$

k = 1, ..., m.

If the model is appropriate and $m \gg n$, then approximately $\hat{r}_k \cong 0$, for all k = 1, ..., m.

An informal analysis of these quantities is the core of the third step of the "Box–Jenkins method". For the ARMA model, Box and Pierce [14] derived the asymptotic distribution of \hat{r} , while for the AR model, Li [58] stated "As noted by Hosking [42], Walker [88] was the first to obtain the distribution under the autoregressive model".

In addition to the asymptotic distribution of \hat{r}_k , Box and Pierce [14] defined a several of statistics based on r_k and \hat{r}_k , where

$$r_k = \sum_{t=k+1}^n \epsilon_t \epsilon_{t-k} / \sum_{t=1}^n \epsilon_t^2.$$

Using the results of [3, 4], the limiting distribution of $r = (r_1, ..., r_m)'$ is multivariate with mean vector zero, $var(r_k) = (n-k)/n(n+2) \approx 1/n$ and $cov(r_k, r_l) = 0$ if $k \neq l$, they showed that, if the model is appropriate and the parameters are known then:

$$\widetilde{Q}(r) = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} r_k^2,$$
(5)

would asymptotically follow χ_m^2 . Furthermore, they suggested the following approximation

$$Q(r) = n \sum_{k=1}^{m} r_k^2 \sim \chi_m^2$$
(6)

based on the approximated variance. If the parameters are estimated and the model is appropriate, Box and Pierce [14] showed that,

$$Q(\hat{r}) = n \sum_{k=1}^{m} \hat{r}_k^2 \tag{7}$$

would be asymptotically χ^2_{m-p-q} yielding an approximate test for lack of fit. The statistics (7) is called the Box-Pierce statistic. Davies, Triggs and Newbold [25] and Prothero and Wallis [73] doubted the distribution of $Q(\hat{r})$ and its validity in practice for a moderated n. This led to the modified statistic:

$$\widetilde{Q}(\widehat{r}) = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} \widehat{r}_k^2.$$
(8)

which was recommended by Ljung and Box [61]. The statistic $\tilde{Q}(\hat{r})$ is called the Ljung-Box statistic, and it has a finite sample distribution that is closer to that of χ^2_{m-p-q} . In addition, they used the result by Anderson and Walker [4] to justify their belief in insensitivity of these tests to departures from normality of the ϵ_t 's.

Another modification on Box-Pierce statistic was suggested by Li and Mcleod [60] as follows:

$$Q_m^* = Q(\hat{r}) + \frac{k^2 m(m+1)}{2n}.$$
(9)

If the model is adequate then Q_m^* asymptotically χ^2 with $k^2(m-p-q)$ degree of freedom.

All the previous statistics to test weather or not a residual series is white noise, under the assumption of normality of the innovations, are based on the residuals' or innovations' autocorrelation functions. There are other approaches to check the whiteness of the residuals. For example, using nonparametric methods, Ducharme and Micheaux [29] employ a datadriven smooth test approach to define a goodness of fit test of normality for the innovations of an ARMA(p, q) model.

Since the sum and the conditional distribution of Gaussian random variables are Gaus-

sian, these facts encourage widespread interest in and investigation of the Gaussian AR(1) model. But in many areas of science, the observations of linear time series models are of non-Gaussian nature and are very common. For instance, one of the earliest applications of this type of model was in hydrology. O'Connell and Jones [71] consider linear time series models driven by lognormal white noise. The ARMA models driven by non-Gaussian innovations have been studied and investigated by many scholars, for example, Li and McLeod [71].

In particular, the simplicity of AR(1) structure, its usefulness and interpretation in a wide range of context, led to the building and studying of more than 30 non-Gaussian AR(1) models, according to [36].

One of the approaches in building the non-Gaussian AR(1) model is to specify the distribution of the innovations. In some cases the marginal distribution will be obtainable, for example the Cauchy AR(1) model by Brockwell and Davis [15]. Let X_t be an AR(1) model with $\epsilon_t \sim C(\lambda, 1)$ and $\lambda \in R$, then the marginal distribution of X_t is Cauchy with mean $\frac{\lambda}{1-\phi}$ and variance $\frac{1}{1-\phi^2}$. An example of non-Gaussian AR(1) with unknown marginal distribution is the Uniform AR(1) defined by Bell and Smith [7]. Let X_t be an AR(1) model with $\epsilon_t \sim U(0, 2\lambda)$, the X_t is a Uniform AR(1) model with unknown marginal distribution.

Another approach in constructing the non-Gaussian AR(1) model is starting with a specific marginal distribution of X_t , and then define a distribution for the innovations, even if it is not a known distribution. An example of this approach is the Logistic AR(1) [76]. Let X_t be an AR(1) model with ϵ_t has a distribution of the form $\frac{\sin(\phi\pi)}{2\phi\pi[\cosh(z-\lambda)+\cos(\phi\pi)]}$, where $-\infty < z < \infty$, then $X_t \sim Lg(\mu, 1)$, where $\mu = \frac{\lambda}{1-\phi}$. Table (1) contains some examples of non-Gaussian AR(1), note that the parameter μ is a function of the parameter λ as in the

Logistic AR(1).

| Name | Reference | Innovation distribution | Marginal distribution |
|---------|---------------------------------|---|-----------------------|
| GENTS | Lye and Martin [62] | $GT(\mu, \sigma^2)$ | Unknown |
| Laplace | Dewald and Lewis [27] | $\begin{cases} \lambda & \text{w.p.}\phi^2 \\ La(\lambda, 1) & \text{w.p.}1 - \phi^2 \end{cases}$ | $La(\mu, 1)$ |
| | Bell and Smith [7] Andĕl [1] | $Exp(1/\lambda)$ | Unknown |
| EAR | Gaver and Lewis [35] | $\begin{cases} 0 & \text{w.p.}\phi \\ Exp(1/\mu) & \text{w.p.}1 - \phi \end{cases}$ | $Exp(1/\mu)$ |

Table 1: Some non-Gaussian time series

Jacobs and Lewis [45, 46] describe steps toward modeling and estimating non-Gaussian time series, using the predesignated marginal distribution approach. When O'Connell and Jones [71] consider fitting an autoregressive model with log-normal innovations, the Yule-Walker equation is used to estimate the autoregressive parameters. Li and McLeod [59] study the asymptotic normality of the maximum likelihood estimator for non-Gaussian ARMA models under some general conditions. Other alternative approaches to modeling non-Gaussian time series have been considered including the Bayesian forecasting models of West and Migon [90], or state space models approach of Kashiwagi and Yanagimoto [52], and through the transformation as in Forecasting non-normal time series by Swift and Janacek [82].

As in the Gaussian ARMA model, Li and McLeod [59] prove the asymptotic normality of the residual autocorrelation function for the non-Gaussian ARMA model. Under the assumption in [59] it can be shown that:

$$\widehat{r}_k = \sum_{t=k+1}^n (\widehat{\epsilon}_t - \overline{\epsilon}) (\widehat{\epsilon}_{t-k} - \overline{\epsilon}) \Big/ \sum_{t=1}^n (\widehat{\epsilon}_t - \overline{\epsilon})^2,$$
(10)

where $\hat{\epsilon}_t$ are residuals from the fitted non-Gaussian ARMA(p,q) model, $\bar{\epsilon} = \sum_{t=1}^n \frac{\hat{\epsilon}_t}{n}$ and k = 1, ..., m, have an asymptotic multivariate normal distribution similar to that of (4) with a different information matrix.

The theoretical skewness and kurtosis of ARMA models with non-Gaussian innovations have been studied by Davies, Spedding, and Watson [26, 89]. The purpose of this study is "to deduce the appropriate skewness and kurtosis that need to be imposed on the noise in order to generate a series, X_t , with given skewness and kurtosis" [26]. In their paper, Davies et al., [26] obtained approximations of skewness and kurtosis of ARMA models with non-Gaussian innovations in terms of skewness and kurtosis of their innovations by making use of the Pearson family of distributions.

Let X_t be a series generated by ARMA(1,1) process, which according to (2), can be written as:

$$X_t - \phi X_{t-1} = \epsilon_t + \theta \epsilon_{t-1}$$

Davies et al., [26] claim that,

$$\sqrt{\beta_1(X)} = \frac{\{1 + \theta(3\phi^2 + 3\phi\theta + \theta^2)\}}{\{1 + \theta(2\phi + \theta)\}^{3/2}} \frac{(1 - \phi^2)^{3/2}}{(1 - \phi^3)} \sqrt{\beta_1(\epsilon)}$$
(11)

where

$$\sqrt{\beta_1(X)} = \frac{E[X_t^3]}{(E[X_t^2])^{3/2}}$$

is the skewness of X_t and

$$\sqrt{\beta_1(\epsilon)} = \frac{E[\epsilon_t^3]}{(E[\epsilon_t^2])^{3/2}}$$

is the skewness of the generating series $\{\epsilon_t\}$.

And

$$\beta_2(X) = \frac{\{(1-\phi^4) + (\phi+\theta)^4\}}{\{1-\phi^2 + (\phi+\theta)^2\}^2} \frac{(1-\phi^2)}{(1+\phi)^2} \beta_2(\epsilon) + 6\Big[\frac{(\phi+\theta)^2(1-\phi)^4 + \phi^2(\phi+\theta)^4}{\{(1-\phi^2) + (\phi+\theta)^2\}^2(1+\phi^2)}\Big]$$
(12)

where

$$\beta_2(X) = \frac{E[X_t^4]}{(E[X_t^2])^2}$$

is the kurtosis of X_t and

$$\beta_2(\epsilon) = \frac{E[\epsilon_t^4]}{(E[\epsilon_t^2])^2}$$

is the kurtosis of $\{\epsilon_t\}$. They showed " it is possible to simulate and match machined metal surface profiles having significantly non normal skewness and kurtosis" by using these techniques.

As in the diagnostic techniques of the Gaussian ARMA model, other approaches have been investigated to check the whiteness of the residuals of the non-Gaussian ARMA models. Using nonparametric approach, Fan and Zhang [34] "consider generalised likelihood ratio tests of whether or not the spectral density function of a stationary time series admits certain parametric forms," which can be applied to test the residuals' whiteness. Taking advantage of techniques from the empirical processes, many scholars propose goodness of fit tests for the ARMA models. For AR(1) when the innovations have unknown symmetric distribution, Koul [54] gives a class of minimum L_2 -distance estimators of the autoregression parameter, and based on that discussed a goodness of fit test for symmetry innovations distribution. In addition, Koul and Stute [56] investigate a class of goodness of fit test for autoregressive model. "These tests are based on a class of empirical processes marked by certain residuals", according to the authors. For heavy tailed innovations, Jureckova et al. [51] construct a class of nonparametric tests on the tail index of the innovation distribution in the linear autoregressive model.

Studying and modeling the sea level data

Studying and investigating sea level rise (SLR) has received attention since the late nineteenth century. According to Byrne [16] "Sea-level rise has been tracked since the late 1800s by tide level gauges, and since 1993 changes have been recorded with high precision from altimeter satellites."

Traditionally, sea level change has been estimated from tide gauge measurements of the 'still water level' collected over the last century. The 'still water level' (without waves) is averaged over some period of time, such as a month or a year yields Mean Sea Level (MSL) or "Sea Level" as in some other literature. This MSL is measured relative to fixed marks on the land known as 'benchmarks'. According to Sea Level Research Group at University of Colorado, "Although the global network of tide gauges comprises of a poorly distributed sea level measurement system, it offers the only source of historical, precise, long-term sea level data". The Permanent Service for Mean Sea Level (PSMSL) has been responsible for the collection, publication, analysis and interpretation of sea level data from the global network of tide gauges since its establishment in 1933.Alternatively, MSL in the context of satellite altimetry, are the measurements made in a geocentric reference frame (relative to the center of the Earth).

Due to the heavy density of population and infrastructure that are located at the coast, sea level rise presents an urgent topic in climate research, with direct socio-economic consequences [17, 77, 70]. "Impact and risk assessment, adaptation policies and long-term decision-making in coastal areas are crucially informed by projections of coastal mean sea level and extreme water level events [47]."

Projection sea level rise focuses on the physical climate system – atmosphere, land surface, ocean and sea ice- which serves the Fourth Asssessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC). The Program for Climate Model Diagnosis and Intercomprison (PCMDI) collects model output contributed by leading modeling centers around the world. This collection of recent model output is officially known as the "WCRP CMIP3 multi-model dataset," where WCRP is "World Climate Research Programme" and CMIP3 stands for phase 3 of the Coupled Model Intercomparison Project "archived data of climate models output" [63]. Such organizations and activities facilitate studying and investigating projection of sea level changes widely through different approaches. According to PCMDI "over 250 journal articles, based at least in part on the dataset, have been published or have been accepted for peer-reviewed publication".

Three different approaches are commonly considered to calculate the projections of future global sea level change based on climate models. The first approach is process-based projections by estimating the contribution of the primary processes that are causing sealevel change [8], "using specified forcing of the climate system" [9]. As stated by Jevrejeva et al. [47] "In a process-based approach, such as that used by Church et al. [20], various sources of information, including numerical simulations, are combined to model the evolution of different sea level components." Global mean sea level (GMSL) changes over time (t)are represented as:

$$GMSL(t) = T(t) + \sum_{i} GIC_{i}(t) + GrIS(t) + WAIS(t) + EAIS(t) + \sum_{j} LWS_{j}(t),$$
(13)

where T is thermostatic expansion, GIC_i ice mass loss from different glaciers, GrIS ice mass loss from the Greenland Ice Sheet, WAIS and EAIS and ice mass loss from the Antarctic Ice Sheet (West and East) and the contribution from land water storage LWS_j changes. Church et al. [20] estimate the median and 90% confidence interval (CI) of sea level change under specific scenarios (representative concentration pathway) as in [68], by the sum of (14) components with their uncertainties.

Regional sea level (RSL) projections using process-based approach combine multiple fac-

tors. The first is the dynamic sea level contribution (DSL), the contribution of sea level components, scaled by spatial "fingerprints" associated with gravitational, deformation and rotational effects (as in [67]). The next is the background (Bkgd) processes, that "unassociated with contemporary climate change that contribute to local sea level changes" [47], like localized processes such as sediment compaction.

RSL projections at locations x can be represented as:

$$RSL(x,t) = T(t) + DSL(x,t) + \sum_{i} F_{GIC_{i}}(x)GIC_{i}(t) + F_{GrIS}(x)GrIS(t)$$

+ $F_{WAIS}(x)WAIS(t) + F_{EAIS}(x)EAIS(t) + \sum_{j} F_{LWS_{j}}(x)LWS_{j}(t)$
+ $Bkgd(x,t),$ (14)

Where F_{GIC_i} , F_{GrIS} , F_{WAIS} , F_{EAIS} and F_{LWS_j} are the normalized gravitational, rotational and deformational fingerprints associated, with the different components.

As illustrated by Meehl et al. [64], this approach is both used to construct the past global sea level changes, and also to project the future. For projecting regional changes, Kopp et al. [53], Slangen et al. [78] and Spada et al. [79] build their studies based on model data from CMIP3, while the result of Perrette et al. [72] are based on model data from the latest suite of coupled climate model results to date CMIP5 [83]. An example of studies based on model data from combining both CMIP3 and CMIP3 was proposed by Slangen et al. [77]. Yin [93] applies GFDL CM2.1 climate model to model projections of rapid sea-level rise on the northeast coast of the United States. As stated earlier there is a huge amount of studies in this regard, the interested reader is advised to consult [47] for more references.

The second approach to project sea level changes is called the "semi-empirical" approach, this approach is based on the observed relationship between sea-level change and global temperature change [8]. Bolin et al. [9] "construct a time series regression model to predict global sea level from global temperature".

Other investigators use the current observations to extrapolate the future behavior of the cryosphere, and use this information in projecting sea level [8, 65].

As mentioned earlier, climate warming has contributed to global sea levels rise through the past century, and they are projected to rise throughout the 21st century [22]. Instead of using climate models or environmental factors as inputs, some authors have used historical sea level records to show sea level changes in mm/yr, that is, the estimation of long- term trends. "Trends in sea level have traditionally been calculated from long term data sets at a few locations" [44]. Several papers discussing evidence for acceleration from tide gauge records include Woodworth [91], Douglas [28], Holgate and Woodworth [41], Church and White [21], Jevrejeva et al. [48] and Woodworth et al. [92].

According to Visser [87], in general, there are three applications of trend estimation. Trends may be viewed in terms of prediction, highlight nonstationary behavior in a series, and trends are used for "calibration of recent attempts to model historic thermal expansion as well as glacier and ice sheet melting". There are five main methods for estimating trends in sea level records, see Chandler and Scott [18] for details on these methods, (i)Exploratory data analysis (or smoothed time series), (ii)Parametric trend estimation, (iii)Nonparametric trend estimation,(iv)Stochastic trend models (examples are ARIMA models and Structural Time series Models) and (v)Miscellaneous models. Applying trend models to sea level data has been done by many investigators. For example Vaziri [86] employs artificial neural network (ANN) and ARIMA models to predict the monthly Caspian Sea level using tide gauge records, whereas Imani et al. [43] applied the same technique to satellite altimetry records. In "Modeling regional sea level rise using local tide gauge data", Iz et al. [44] use the empirical trigonometric model to combine various tide gauge data, regardless of their time span, to estimate the representative of regional mean sea level trends. For a review of trend models applied to sea level data, the interested reader is referred to [87].

Using different methods to estimate sea level trends, several recent studies show that the rates of sea level rise (SLR) have been accelerating along the coastal mid-Atlantic region (the so-called "hotspot" of accelerated SLR, as referred to by Sallenger [74]) [31]. Boon et al. [11] use the least-square linear curve-fit method "simple linear regression" and quadratic regression to calculate the trend after filtering out seasonal and decadal variability, of sea level measurements at ten Chesapeake Bay water level stations (1975-2009).

Sallenger et al. [74] "present evidence of recently accelerated SLR in a unique 1,000km-long hotspot on the highly populated North American Atlantic coast north of Cape Hatteras" by analyzing tide-gauge records along the North American Atlantic coast. They suggest that changes in ocean dynamic contribute significantly to sea level rise. Monthly mean sea level records from 8 tide gauge stations in the Chesapeake Bay were used as an application of a new method of trend estimation by Ezer et al. [33]. The new method is based on Empirical Mode Decomposition (EMD) and Hilbert-Huang Transform (HHT). The same method (EMD/HHT) is applied by Ezer et al. [31] to study the relationship between sea level variations in Chesapeake Bay and the mid-Atlantic coast and the the Gulf Stream, using both monthly mean sea level records from 10 tide gauge stations and altimeter data from both regions.

In another study of measurements of sea level at tide stations along the Northeast Atlantic coastal region of the United States and Canada from 1969 to 2011, Boon [10] obtains results that are in agreement with the observational study by Sallenger et al. [74] and the pattern of modeled projections (climate model) by Yin et al. [93]. According to [10] "The observational data support the claim that sea level rise is now accelerating at locations from Norfolk, Virginia, to Halifax, Nova Scotia, at rates as high as $0.30mm/y^{2"}$.

In their paper, "Modeling of Coastal Inundation, Storm Surge, and Relative SeaLevel Rise at Naval Station Norfolk, Norfolk, Virginia, U.S.A.", Li et al. [57] applied the coastal modeling system to calculate the local water-surface elevation and storm-surge inundation in Norfolk, Virginia, for varying future relative SLR scenarios. They describe the consequences of potential relative SLR estimates and storm surges on inundation. In the context of low probability with high impact events "extreme event", Sweet and Park [81] studies the sea level data from different stations in the East and West coasts of the Unite States, including Norfolk (Sewells Point gauge). Among their results, data from Norfolk shows exponential growth when fitting growth model.

The moments of the Gumbel random variable and the theoretical moments of the AR(1) model

Since the defined hypothesis tests are based on the empirical skewness and kurtosis of the AR(1), then the theoretical moments for the AR(1) model and for the innovations are needed. In this chapter, we started by proving a general formula for the higher moments of Gumbel random variable, then obtained the first six moments. Then, the second, third, fourth, fifth, sixth and eighth moments for AR(1) model were calculated in section 2. In section 3 we calculate the variance of second, third and fourth empirical moments, and covariance of second and third empirical moments, and second and fourth empirical moments.

The Moments of the Gumbel Random Variable

In this section we start by proving a general formula for the higher moments of Gumbel random variable.

Theorem 0.0.1. Let X be a Gumbel random variable with location parameter μ and scale parameter β , "X ~ G(μ , β)". Then the higher moments around the origin τ_n , $E(X^n)$, are given by:

$$\tau_n = \sum_{k=0}^n \binom{n}{k} \Gamma^{(k)}(1) (-\beta)^k \mu^{n-k}, \qquad n \in N$$
(15)

where $\Gamma^{(k)}(1)$ is the k-th derivative of Gamma function evaluated at 1, and $\Gamma^{(0)}(1) = \Gamma(1)$.

 $\mathit{Proof.}$ We will use mathematical induction to prove :

$$\tau_n = m^{(n)}(0)$$

$$= m^{(t)} \Big|_{t=0}$$

$$= \frac{d^n}{dt^n} \Gamma(1 - \beta t) e^{\mu t} \Big|_{t=0}$$

$$= \sum_{k=0}^n \binom{n}{k} \Gamma^{(k)}(1) (-\beta)^k \mu^{n-k}$$

When n = 1, then

$$\tau_{1} = m^{(1)}(0)$$

$$= \frac{d}{dt}\Gamma(1-\beta t)e^{\mu t}\Big|_{t=0}$$

$$= \Gamma(1-\beta t)e^{\mu t}\mu - \Gamma'(1-\beta t)e^{\mu t}\beta\Big|_{t=0}$$

$$= \Gamma(1)\mu - \Gamma'(1)\beta$$

$$= \sum_{k=0}^{1} {\binom{1}{k}}\Gamma^{(k)}(1)(-\beta)^{k}\mu^{1-k}.$$

Now, assume the statement is correct for n = i, s.t.

$$\begin{split} \tau_{i} &= m^{(i)}(0) \\ &= \frac{d^{i}}{dt^{i}} \Gamma(1 - \beta t) e^{\mu t} \Big|_{t=0} \\ &= \binom{i}{0} \mu^{i} e^{\mu t} \Gamma(1 - \beta t) - \binom{i}{1} \mu^{i-1} \beta e^{\mu t} \Gamma'(1 - \beta t) + \dots + \\ &+ \binom{i}{i-1} \mu \beta^{i-1} e^{\mu t} \Gamma^{(i-1)}(1 - \beta t) - \binom{i}{i} \beta^{i} e^{\mu t} \Gamma^{(i)}(1 - \beta t) \Big|_{t=0} \\ &= \mu^{i} \Gamma(1) - \binom{i}{1} \mu^{i-1} \beta \Gamma'(1) + \dots + \binom{i}{i-1} \mu \beta^{i-1} \Gamma^{(i-1)}(1) - \beta^{i} \Gamma^{(i)}(1), \\ &= \sum_{k=0}^{i} \binom{i}{k} \Gamma^{(k)}(1)(-\beta)^{k} \mu^{i-k} \end{split}$$

since $\binom{i}{0} = \binom{i}{i} = 1$ we get the second last line. And to prove it for n = i + 1, note that

$$\begin{split} \tau_{i+1} &= \frac{d^{i+1}}{dt^{i+1}} m(t) \Big|_{t=0} \\ &= \frac{d}{dt} (m^{i}(t)) \Big|_{t=0} \\ &= \frac{d}{dt} \left[\binom{i}{0} \mu^{i} e^{\mu t} \Gamma(1 - \beta t) - \binom{i}{1} \mu^{i-1} \beta e^{\mu t} \Gamma'(1 - \beta t) + \dots + \right. \\ &+ \left(\binom{i}{i-1} \right) \mu \beta^{i-1} e^{\mu t} \Gamma^{(i-1)}(1 - \beta t) - \binom{i}{i} \beta^{i} e^{\mu t} \Gamma^{(i)}(1 - \beta t) \right] \Big|_{t=0} \\ &= \left[\binom{i}{0} \mu^{i+1} e^{\mu t} \Gamma(1 - \beta t) - \binom{i}{0} \mu^{i} e^{\mu t} \beta \Gamma'(1 - \beta t) \right. \\ &- \binom{i}{1} \mu^{i} \beta e^{\mu t} \Gamma'(1 - \beta t) + \binom{i}{1} \mu^{i-1} \beta^{2} e^{\mu t} \Gamma''(1 - \beta t) + \dots + \\ &+ \left(\binom{i}{i-1} \mu^{2} \beta^{i-1} e^{\mu t} \Gamma^{(i-1)}(1 - \beta t) - \binom{i}{i-1} \mu \beta^{i} e^{\mu t} \Gamma^{(i)}(1 - \beta t) \right. \\ &- \left. \binom{i}{i} \mu \beta^{i} e^{\mu t} \Gamma^{(i)}(1 - \beta t) + \binom{i}{i} \beta^{i+1} e^{\mu t} \Gamma^{(i+1)}(1 - \beta t) \right] \Big|_{t=0} \\ &= \left[\binom{i}{0} \mu^{i+1} \Gamma(1) - \binom{i}{0} \mu^{i} \beta \Gamma'(1) - \binom{i}{1} \mu^{i} \beta \Gamma'(1) + \binom{i}{1} \mu^{i-1} \beta^{2} \Gamma''(1) \right. \\ &+ \dots + \binom{i}{i-1} \mu^{2} \beta^{i-1} \Gamma^{(i-1)}(1) - \binom{i}{i-1} \mu \beta^{i} \Gamma^{(i)}(1) - \binom{i}{i} \mu \beta^{i} \Gamma^{(i)}(1) \right] \end{split}$$

notice that $\binom{i}{0} = \binom{i}{i} = 1$ implies:

$$\tau_{i+1} = \left[\mu^{i+1} \Gamma(1) - \left[\binom{i}{0} + \binom{i}{1} \right] \mu^{i} \beta \Gamma'(1) + \dots + \dots - \left[\binom{i}{i-1} + \binom{i}{i} \right] \mu \beta^{i} \Gamma^{(i)}(1) + \beta^{(i+1)} \Gamma^{(i+1)}(1) \right]$$

Now applying the identity $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ to the two terms inside the previous expression yields to:

$$\tau_{i+1} = \left[\mu^{i+1} \Gamma(1) - \binom{i+1}{1} \mu^i \beta \Gamma'(1) + \dots + \dots - \binom{i+1}{i} \mu \beta^i \Gamma^{(i)}(1) + \beta^{(i+1)} \Gamma^{(i+1)}(1) \right]$$
$$= \sum_{k=0}^{i+1} \binom{i+1}{k} \Gamma^{(k)}(1) (-\beta)^k \mu^{(i+1)-k}$$

Which is the required.

Since the formula contains the higher derivatives of Gamma function evaluated at 1, the following lemma states these values without prove. (see [19] for details)

Lemma 0.0.2. Let $\Gamma(z)$ be the Gamma function, defined on the complex plane as:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt , \text{ where } \quad R(z) > 0$$

and let $\zeta(s)$ be the Riemann Zeta function, defined as

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1}}{e^x - 1} dx$$

Euler evaluated $\zeta(2n), n \in N$, that is $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$, and $\zeta(6) = \frac{\pi^6}{945}$, then:

- 1. $\Gamma(1) = 1$.
- 2. $\Gamma'(1) = -\gamma$.

3. The second derivative

$$\Gamma''(1) = \gamma^2 + \zeta(2)$$
$$= \gamma^2 + \frac{\pi^2}{6}$$

4. The third derivative

$$\Gamma^{(3)}(1) = -\gamma^3 - 3\gamma\zeta(2) - 2\zeta(3)$$
$$= -\gamma^3 - \gamma\frac{\pi^2}{2} - 2\zeta(3)$$

5. The forth derivative

$$\Gamma^{(4)}(1) = \gamma^4 + 6\gamma^2 \zeta(2) + 8\gamma \zeta(3) + \frac{27}{4} \zeta(4)$$
$$= \gamma^4 + \gamma^2 \pi^2 + 8\gamma \zeta(3) + \frac{3}{20} \pi^4$$

6. The fifth derivative

$$\Gamma^{(5)}(1) = -\gamma^5 - 10\gamma^3\zeta(2) - 20\gamma^2\zeta(3) - \frac{135}{2}\zeta(4) - 20\zeta(2)\zeta(3) - 24\zeta(5)$$
$$= -\gamma^5 - \frac{5}{3}\gamma^3\pi^2 - 20\gamma^2\zeta(3) - \frac{3}{4}\gamma\pi^4 - \frac{10}{3}\pi^2\zeta(3) - 24\zeta(5)$$

7. The sixth derivative

$$\Gamma^{(6)}(1) = \gamma^{6} + 15\gamma^{4}\zeta(2) + 40\gamma^{3}\zeta(3) + \frac{404}{2}(\zeta(4))^{2} + 120\gamma\zeta(2)\zeta(3) + 144\gamma\zeta(5) + 40(\zeta(3))^{2} + \frac{2745}{8}\zeta(6) = \gamma^{6} + \frac{5}{2}\gamma^{4}\pi^{2} + 40\gamma^{3}\zeta(3) + \frac{45}{20}\gamma^{2}\pi^{4} + 20\gamma\pi^{2}\zeta(3) + 144\gamma\zeta(5) + 40(\zeta(3))^{2} + \frac{61}{168}\pi^{6}$$

Using these values with Theorem 0.0.1, the first six moments for Gumbel random variable are obtained in the following Lemma.

Lemma 0.0.3. Let X be a Gumbel random variable with location parameter μ and scale parameter β , "X ~ G(μ , β)", and define $E(X^n) = \kappa_n$ then:

1.
$$\tau_1 = \mu + \beta \gamma$$
.
2. $\tau_2 = \mu^2 + \beta^2 \gamma^2 + \frac{\pi^2}{6} \beta^2 + 2\beta \mu \gamma$
3. $\tau_3 = \mu^3 + 3\mu^2 \beta \gamma + 3\mu \beta^2 (\gamma^2 + \frac{\pi^2}{6}) + \beta^3 (\gamma^3 + \gamma \frac{\pi^2}{2} + 2\zeta(3))$.

$$\tau_4 = \mu^4 + 4\mu^3\beta\gamma + 6\beta^2\mu^2(\gamma^2 + \frac{\pi^2}{6}) + 4\beta^3\mu(\gamma^3 + \gamma\frac{\pi^2}{2} + 2\zeta(3)) + \beta^4(\gamma^4 + \gamma^2\pi^2 + 8\gamma\zeta(3) + \frac{3}{20}\pi^4)$$

$$\tau_{5} = \mu^{5} + 5\mu^{4}\beta\gamma + 10\mu^{3}\beta^{2}(\gamma^{2} + \frac{\pi^{2}}{6}) + 10\mu^{2}\beta^{3}(\gamma^{3} + 3\gamma\frac{\pi^{2}}{2} + 2\zeta(3))$$

+ $5\mu\beta^{4}(\gamma^{4} + \gamma^{2}\pi^{2} + 8\gamma\zeta(3) + \frac{3}{20}\pi^{4})$
+ $\beta^{5}(\gamma^{5} + \frac{5}{3}\gamma^{3}\pi^{2} + 20\gamma^{2}\zeta(3) + \frac{3}{4}\gamma\pi^{4} + \frac{10}{3}\pi^{2}\zeta(3) + 24\zeta(5))$

$$\begin{aligned} \tau_6 &= \mu^6 + 6\mu^5 \beta \gamma + 15\mu^4 \beta^2 (\gamma^2 + \frac{\pi^2}{6}) + 20\mu^3 \beta^3 (\gamma^3 + \gamma \frac{\pi^2}{2} + 2\zeta(3)) \\ &+ 15\mu^2 \beta^4 (\gamma^4 + \gamma^2 \pi^2 + 8\gamma\zeta(3) + \frac{3}{20}\pi^4) + 6\mu\beta^5 (\gamma^5 + \frac{5}{3}\gamma^3 \pi^2 + 20\gamma^2 \zeta(3)) \\ &+ \frac{3}{4}\gamma \pi^4 + \frac{10}{3}\pi^2 \zeta(3) + 24\zeta(5)) + \beta^6 (\gamma^6 + \frac{5}{2}\gamma^4 \pi^2 + 40\gamma^3 \zeta(3)) \\ &+ \frac{45}{20}\gamma^2 \pi^4 + 20\gamma \pi^2 \zeta(3) + 144\gamma \zeta(5) + 40(\zeta(3))^2 + \frac{61}{168}\pi^6) \end{aligned}$$

Proof. By applying Theorem 0.0.1 and Lemma0.0.2.

Now the moments of a centered "E(X) = 0 and $Var(X) = \sigma^2$ " Gumbel random variable, is stated in the following corollary.

Corollary 0.0.4. Suppose $X \sim G(-\frac{\sqrt{6}}{\pi}\gamma\sigma, \frac{\sqrt{6}}{\pi}\sigma)$, where $\sigma > 0$ and γ is Euler's constant, then

1. $\tau_1 = 0$ 2. $\tau_2 = \sigma^2$ 3. $\tau_3 = \frac{2 \cdot 6\sqrt{6}}{\pi^3} \sigma^3 \zeta(3)$
4.
$$\tau_4 = \frac{3 \cdot 9}{5} \sigma^4$$

5. $\tau_5 = \frac{36\sqrt{6}}{\pi^5} \sigma^5 [\frac{10}{3} \pi^2 \zeta(3) + 24\zeta(5)]$
6. $\tau_6 = \frac{36 \cdot 6}{\pi^6} \sigma^6 [40(\zeta(3))^2 + \frac{61}{168} \pi^6]$

Where $\zeta(.)$ is the Riemann Zeta function defined in Lemma 0.0.2.

The theoretical moments of the AR(1) model

In this section we calculate the theoretical moments of AR(1) model. Let $\xi_t = \text{IID}(0, \sigma^2)$ and

$$X_{t+1} = \phi X_t + \xi_{t+1}$$

which implies:

$$X_t = \sum_{i=0}^{\infty} \phi^i \xi_{t-i},$$

Let $E(X_t^r) = m_r$ and $E(\xi_t^r) = \tau_r$, then

Lemma 0.0.5.

$$m_2 = \frac{\tau_2}{1 - \phi^2}$$

Proof.

$$m_{2} = E[(\sum_{i=0}^{\infty} \phi^{i} \xi_{t-i})^{2}]$$

= $E\left[(\sum_{i=0}^{\infty} \phi^{i} \xi_{t-i})(\sum_{j=0}^{\infty} \phi^{j} \xi_{t-j})\right]$
= $\sum_{i=0}^{\infty} \phi^{2i} E[\xi_{t-i}^{2}] = \frac{\tau_{2}}{1-\phi^{2}}$

Since $E(\xi_t) = 0$ we keep only the case where i = j.

And the third moment is:

Lemma 0.0.6.

$$m_3 = \frac{\tau_3}{1 - \phi^3}$$

Proof.

$$m_{3} = E[(\sum_{i=0}^{\infty} \phi^{i} \xi_{t-i})^{3}]$$

= $E\Big[(\sum_{i=0}^{\infty} \phi^{i} \xi_{t-i})(\sum_{j=0}^{\infty} \phi^{j} \xi_{t-j})(\sum_{k=0}^{\infty} \phi^{k} \xi_{t-k})\Big]$
= $\sum_{i=0}^{\infty} \phi^{3i} E[\xi_{t-i}^{3}] = \frac{\tau_{3}}{1-\phi^{3}}$

Since $E(\xi_t) = 0$ we keep only the case where i = j = k.

The fourth moment is:

Lemma 0.0.7.

$$m_4 = \frac{\tau_4 - 3\tau_2^2}{(1 - \phi^4)} + \frac{3\tau_2^2}{(1 - \phi^2)^2}$$

Proof.

$$\begin{split} m_4 &= E[(\sum_{i=0}^{\infty} \phi^i \xi_{t-i})^4] \\ &= E\Big[(\sum_{i=0}^{\infty} \phi^i \xi_{t-i})(\sum_{j=0}^{\infty} \phi^j \xi_{t-j})(\sum_{k=0}^{\infty} \phi^k \xi_{t-k})(\sum_{l=0}^{\infty} \phi^l \xi_{t-l})\Big] \\ &= \sum_{i=0}^{\infty} \phi^{4i} E[\xi_{t-i}^4] + 3\sum_{i=0}^{\infty} \sum_{j=0, j \neq i}^{\infty} \phi^{2(i+j)} E[\xi_{t-i}^2] E[\xi_{t-j}^2] \\ &= \frac{\tau_4}{1 - \phi^4} + 3\tau_2^2 \Big(\frac{1}{(1 - \phi^2)^2} - \frac{1}{(1 - \phi^4)}\Big) \\ &= \frac{\tau_4 - 3\tau_2^2}{(1 - \phi^4)} + \frac{3\tau_2^2}{(1 - \phi^2)^2} \end{split}$$

in the second line we keep the cases where all the indices are equal and the indices when two are equal and the other two are equal, and this is the coefficient of the second term in the second line, from $\binom{4}{2}\binom{2}{2}/2!$. And the summation

$$\sum_{i=0}^{\infty} \sum_{j=0, j \neq i}^{\infty} \phi^{2(i+j)} E[\xi_{t-i}^2] E[\xi_{t-j}^2]$$

is done using the fact that

$$\sum_{i \neq k} = \sum_{i,k} - \sum_{i=k}$$

as follows

$$\begin{split} \sum_{i=0}^{\infty} \sum_{j=0, j \neq i}^{\infty} \phi^{2(i+j)} E[\xi_{t-i}^2] E[\xi_{t-j}^2] &= \tau_2^2 \sum_{i=0}^{\infty} \phi^{2i} \sum_{j=0, j \neq i}^{\infty} \phi^2 j \\ &= \tau_2^2 \sum_{i=0}^{\infty} \phi^{2i} \Big[\frac{1}{(1-\phi^2)} - \phi^{2i} \Big] \\ &= \tau_2^2 \Big[\frac{1}{(1-\phi^2)^2} - \frac{1}{(1-\phi^4)} \Big] \end{split}$$

And the fifth moment is:

Lemma 0.0.8.

$$m_5 = \frac{\tau_5 - 10\tau_2\tau_3}{(1 - \phi^5)} + \frac{10\tau_2\tau_3}{(1 - \phi^3)(1 - \phi^2)}$$

Proof.

$$\begin{split} m_5 &= E[(\sum_{i=0}^{\infty} \phi^i \xi_{t-i})^5] \\ &= E\Big[(\sum_{i=0}^{\infty} \phi^i \xi_{t-i})(\sum_{j=0}^{\infty} \phi^j \xi_{t-j})(\sum_{k=0}^{\infty} \phi^k \xi_{t-k})(\sum_{l=0}^{\infty} \phi^l \xi_{t-l})(\sum_{m=0}^{\infty} \phi^m \xi_{t-m})\Big] \\ &= \sum_{i=0}^{\infty} \phi^{5i} E[\xi_{t-i}^5] + {5 \choose 3} \sum_{i=0}^{\infty} \sum_{j=0, j \neq i}^{\infty} \phi^{3i+2j} E[\xi_{t-i}^3] E[\xi_{t-j}^2] \\ &= \frac{\tau_5}{1-\phi^5} + 10\tau_2\tau_3\Big(\frac{1}{(1-\phi^3)(1-\phi^2)} - \frac{1}{1-\phi^5}\Big) \\ &= \frac{\tau_5 - 10\tau_2\tau_3}{(1-\phi^5)} + \frac{10\tau_2\tau_3}{(1-\phi^3)(1-\phi^2)} \end{split}$$

in the second line we keep the cases where all the indices are equal and the indices when two are equal and the other three are equal, and this is the coefficient of the second term in the second line, from $\binom{5}{3}\binom{2}{2}$. Again the summation in the scond line is done using the fact that

$$\sum_{i \neq k} = \sum_{i,k} - \sum_{i=k}$$

Where as the sixth moment is:

Lemma 0.0.9.

•

$$m_6 = \frac{\tau_6 - 15\tau_2\tau_4 - 10\tau_3^2 + 30\tau_2^3}{(1 - \phi^6)} + \frac{15\tau_2\tau_4 - 45\tau_2^3}{(1 - \phi^4)(1 - \phi^2)} + \frac{10\tau_3^2}{(1 - \phi^3)^2} + \frac{15\tau_2^3}{(1 - \phi^2)^3}$$

Proof.

$$\begin{split} m_{6} &= E[(\sum_{i=0}^{\infty} \phi^{i} \xi_{t-i})^{6}] \\ &= \sum_{i=0}^{\infty} \phi^{6i} E[\xi_{t-i}^{6}] + \binom{6}{2} \sum_{i=0}^{\infty} \sum_{j=0, j \neq i}^{\infty} \phi^{4i+2j} E[\xi_{t-i}^{4}] E[\xi_{t-j}^{2}] \\ &+ \frac{\binom{6}{3}}{2} \sum_{i=0}^{\infty} \sum_{j=0, j \neq i}^{\infty} \phi^{3i+3j} E[\xi_{t-i}^{3}] E[\xi_{t-j}^{3}] \\ &+ \frac{\binom{6}{2}\binom{2}{2}}{6} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0, k \neq j \neq i}^{\infty} \phi^{2i+2j+2k} E[\xi_{t-i}^{2}] E[\xi_{t-j}^{2}] E[\xi_{t-k}^{2}] \\ &= \frac{\tau_{6}}{1-\phi^{6}} + 15\tau_{2}\tau_{4} \Big(\frac{1}{(1-\phi^{4})(1-\phi^{2})} - \frac{1}{1-\phi^{6}} \Big) + 10\tau_{3}^{2} \Big(\frac{1}{(1-\phi^{3})^{2}} - \frac{1}{1-\phi^{6}} \Big) \\ &+ 15\tau_{2}^{3} \Big(\frac{1}{(1-\phi^{2})^{3}} - \frac{3}{(1-\phi^{4})(1-\phi^{2})} + \frac{2}{(1-\phi^{6})} \Big) \\ &= \frac{\tau_{6} - 15\tau_{2}\tau_{4} - 10\tau_{3}^{2} + 30\tau_{2}^{3}}{(1-\phi^{6})} + \frac{15\tau_{2}\tau_{4} - 45\tau_{2}^{3}}{(1-\phi^{4})(1-\phi^{2})} + \frac{10\tau_{3}^{2}}{(1-\phi^{3})^{2}} + \frac{15\tau_{2}^{3}}{(1-\phi^{2})^{3}} \Big) \end{split}$$

Again, by repeating the same logic to count the indices, and using the fact

$$\sum_{i\neq k} = \sum_{i,k} - \sum_{i=k}$$

Lastly the eighth moment is:

•

Lemma 0.0.10.

$$m_8 = = \frac{\tau_8 - 28\tau_2\tau_6 - 56\tau_3\tau_5 - 35\tau_4^2 + 420\tau_2^2\tau_4 + 560\tau_2\tau_3^2 - 630\tau_2^4}{(1 - \phi^8)} + \frac{28\tau_2\tau_6 - 420\tau_2^2\tau_4 - 280\tau_2\tau_3^2 + 840\tau_2^4}{(1 - \phi^6)(1 - \phi^2)} + \frac{56\tau_3\tau_5 - 560\tau_2\tau_3^2}{(1 - \phi^5)(1 - \phi^3)} + \frac{35\tau_4^2 - 210\tau_2^2\tau_4 + 315\tau_2^4}{(1 - \phi^4)^2} + \frac{210\tau_2^2\tau_4 - 630\tau_2^4}{(1 - \phi^2)^2} + \frac{280\tau_2\tau_3^2}{(1 - \phi^3)^2(1 - \phi^2)} + \frac{105\tau_2^4}{(1 - \phi^2)^4}$$

Proof.

$$\begin{split} m_8 &= E[(\sum_{i=0}^{\infty} \phi^i \xi_{t-i})^8] \\ &= \sum_{i=0}^{\infty} \phi^{8i} E[\xi_{t-i}^8] + \binom{8}{6} \sum_{i=0}^{\infty} \sum_{j=0,j\neq i}^{\infty} \phi^{6i+2j} E[\xi_{t-i}^6] E[\xi_{t-j}^2] \\ &+ \binom{8}{5} \sum_{i=0}^{\infty} \sum_{j=0,j\neq i}^{\infty} \phi^{5i+3j} E[\xi_{t-i}^5] E[\xi_{t-j}^3] + \binom{4}{2} \sum_{i=0}^{\infty} \sum_{j=0,j\neq i}^{\infty} \phi^{4i+4j} E[\xi_{t-i}^4] E[\xi_{t-j}^4] \\ &+ \binom{8}{5} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0,k\neq j\neq i}^{\infty} \phi^{4i+2j+2k} E[\xi_{t-i}^4] E[\xi_{t-j}^2] E[\xi_{t-k}^2] \\ &+ \frac{\binom{8}{3} \binom{5}{2}}{2} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0,k\neq j\neq i}^{\infty} \phi^{3i+3j+2k} E[\xi_{t-i}^3] E[\xi_{t-j}^3] E[\xi_{t-k}^2] \\ &+ \frac{\binom{8}{2} \binom{6}{2} \binom{4}{2}}{24} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{k\neq j\neq i}^{\infty} \phi^{3i+3j+2k} E[\xi_{t-i}^3] E[\xi_{t-j}^2] E[\xi_{t-j}^2] E[\xi_{t-k}^2] E[\xi_{t-k}^2] \\ &+ \frac{\binom{8}{2} \binom{6}{2} \binom{4}{2}}{24} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0,i\neq j\neq k\neq l}^{\infty} \phi^{2i+2j+2k+2l} E[\xi_{t-i}^2] E[\xi_{t-j}^2] E[\xi_{t-k}^2] E[\xi_{t-k}^2] E[\xi_{t-k}^2] \\ &+ \frac{6}{(1-\phi)^8} + 28\tau_2\tau_6 \left(\frac{1}{(1-\phi^6)(1-\phi^2)} - \frac{1}{(1-\phi^8)}\right) \\ &+ 56\tau_3\tau_5 \left(\frac{1}{(1-\phi^4)(1-\phi^2)^2} - \frac{2}{(1-\phi^2)(1-\phi^6)} - \frac{1}{(1-\phi^4)(1-\phi^5)} + \frac{2}{(1-\phi^8)}\right) \\ &+ 280\tau_2\tau_3^2 \left(\frac{1}{(1-\phi^4)(1-\phi^2)^2} - \frac{6}{(1-\phi^4)(1-\phi^2)^2} + \frac{8}{(1-\phi^6)(1-\phi^2)} + \frac{3}{(1-\phi^4)^2} - \frac{6}{(1-\phi^8)}\right) \\ &= \frac{\tau_8 - 28\tau_2\tau_6 - 56\tau_3\tau_5 - 35\tau_4^2 + 420\tau_2^2\tau_4 + 560\tau_2\tau_3^2 - 630\tau_4^4}{(1-\phi^8)} \\ &+ \frac{28\tau_2\tau_6 - 420\tau_2^2\tau_4 - 280\tau_2\tau_3^2 + 840\tau_4^4}{(1-\phi^6)(1-\phi^2)} + \frac{105\tau_4^4}{(1-\phi^4)(1-\phi^2)^2} + \frac{280\tau_2\tau_3^2}{(1-\phi^3)^2(1-\phi^2)} + \frac{105\tau_4^4}{(1-\phi^4)(1-\phi^2)^2} + \frac{280\tau_2\tau_3^2}{(1-\phi^4)^2(1-\phi^4)^2} + \frac{280\tau_2\tau_3^2}{(1-\phi^4)^2} + \frac{280\tau_2\tau_3^2}{(1-\phi^4)^2} + \frac{28\tau_2\tau_6 - 56\tau_3\tau_5}{(1-\phi^5)(1-\phi^2)} + \frac{28\tau_2\tau_6 - 420\tau_2^2\tau_4}{(1-\phi^4)(1-\phi^2)^2} + \frac{280\tau_2\tau_3^2}{(1-\phi^4)^2} + \frac{28\tau_2\tau_6}{(1-\phi^4)(1-\phi^2)^2} + \frac{28\tau_2\tau_6}{(1-\phi^4)(1-\phi^2)^2} + \frac{28\tau_2\tau_6}{(1-\phi^4)^2(1-\phi^2)^2} + \frac{28\tau_2\tau_6}{(1-\phi^4)^2} + \frac{28\tau_2\tau_6}{(1-\phi^4)$$

The asymptotic variances and covariances of the empirical moments of the AR(1) model

In this section we calculate the variances and covariances of $\widehat{m}_i = \sum_{t=1}^T X_t^i / T$ where $X_t = \sum_{i=0}^\infty \phi^i \xi_{t-i}$, and $\xi_t = \text{IID}(0, \sigma^2)$ with $E(\xi^r) = \tau_r$ and $\tau_r < \infty$ for r = 2, 3, 4, 5, 6, 8.

We start by the asymptotic variance of the second empirical moment.

Theorem 0.0.11.

$$T \operatorname{Var}[\widehat{m}_2] = (m_4 - m_2^2) \left(\frac{2}{(1 - \phi^2)} - 1\right) + o(1)$$

Proof.

$$\operatorname{Var}[\widehat{m}_2] = \frac{(m_4 - m_2^2)}{T} + \frac{1}{T^2} \sum_{i=1}^T \sum_{j=1}^T \operatorname{Cov}(X_i^2, X_j^2)$$

Now verify there are exactly T - r terms whose covariance values are same and equal to $Cov(X_1^2, X_{1+r}^2)$. Thus,

$$\operatorname{Var}[\widehat{m}_{2}] = \frac{(m_{4} - m_{2}^{2})}{T} + \underbrace{\frac{2}{T^{2}} \sum_{r=1}^{T-1} (T - r) \operatorname{Cov}(X_{1}^{2}, X_{1+r}^{2})}_{C_{T,r}}$$

Thus, the problem boils down to the computation of $\operatorname{Cov}(X_1^2, X_{1+r}^2)$. Now note that

$$X_{1+r} = \phi^{r} X_{1} + \underbrace{\sum_{i=1}^{r} \phi^{r-i} \xi_{1+i}}_{R_{r}}$$

where R_r is independent of X_1 . Thus,

$$Cov(X_1^2, X_{1+r}^2) = Cov(X_1^2, \phi^{2r} X_1^2 + 2\phi^r X_1 R_r + R_r^2)$$

= $\phi^{2r} Var(X_1^2) + 2\phi^r Cov(X_1^2, X_1 R_r)$
= $\phi^{2r} (m_4 - m_3^2) + 2\phi^r (m_3 E[R_r] - m_2 m_1 E[R_r])$

First note that $m_1 = 0$ and $E[R_r] = 0$ which allows us to simplify a bit.

$$\operatorname{Cov}(X_1^2, X_{1+r}^2) = \phi^{2r}(m_4 - m_3^2)$$

This implies:

$$C_{T,r} = \frac{2}{T^2} \sum_{r=1}^{T-1} (T-r)\phi^{2r}(m_4 - m_3^2)$$

= $\frac{2}{T^2} \sum_{r=0}^{T-1} (T-r)\phi^{2r}(m_4 - m_3^2) - \frac{2(m_4 - m_3^2)}{T}$
= $\frac{2(m_4 - m_3^2)}{T} \left(\frac{1}{(1-\phi^2)} - 1\right) + o\left(\frac{1}{T}\right)$

| - | - | - | - | - | |
|---|---|---|---|---|--|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

The following theorem is for the asymptotic variance of the third empirical moment.

Theorem 0.0.12.

$$T \operatorname{Var}[\widehat{m}_3] = 2 \left[(m_6 - m_3^2) \left(\frac{1}{(1 - \phi^3)} - 1/2 \right) + 3m_4 \tau_2 \left(\frac{1}{(1 - \phi)(1 - \phi^2)} - \frac{1}{(1 - \phi^3)(1 - \phi^2)} \right) \right] + o(1)$$

Proof.

$$\operatorname{Var}[\widehat{m}_3] = \frac{(m_6 - m_3^2)}{T} + \frac{1}{T^2} \sum_{i=1}^T \sum_{j=1}^T \operatorname{Cov}(X_i^3, X_j^3)$$

Now verify there are exactly T - r terms whose covariance values are same and equal to $Cov(X_1^3, X_{1+r}^3)$. Thus,

$$\operatorname{Var}[\widehat{m}_3] = \frac{(m_6 - m_3^2)}{T} + \underbrace{\frac{2}{T^2} \sum_{r=1}^{T-1} (T - r) \operatorname{Cov}(X_1^3, X_{1+r}^3)}_{C_{T,r}}$$

Thus, the problem boils down to the computation of $\text{Cov}(X_1^3, X_{1+r}^3)$. Now note that

$$X_{1+r} = \phi^r X_1 + \underbrace{\sum_{i=1}^r \phi^{r-i} \xi_{1+i}}_{R_r}$$

where R_r is independent of X_1 . Thus,

$$Cov(X_1^3, X_{1+r}^3) = Cov(X_1^3, \phi^{3r} X_1^3 + 3\phi^{2r} X_1^2 R_r + 3\phi^r X_1 R_r^2 + R_r^3)$$

= $\phi^{3r} Var(X_1^3) + 3\phi^{2r} Cov(X_1^3, X_1^2 R_r) + 3\phi^r Cov(X_1^3, X_1 R_r^2)$
= $\phi^{3r} (m_6 - m_3^2) + 3\phi^{2r} (m_5 E[R_r] - m_3 m_2 E[R_r])$
+ $3\phi^r (m_4 E[R_r^2] - m_3 m_1 E[R_r^2])$

First note that $m_1 = 0$ and $E[R_r] = 0$ which allows us to simplify a bit.

$$\operatorname{Cov}(X_1^3, X_{1+r}^3) = \phi^{3r}(m_6 - m_3^2) + 3\phi^r m_4 E[R_r^2]$$

We simplify further by noting

$$E[R_r^2] = \frac{\tau_2(1-\phi^{2r})}{1-\phi^2}$$

$$C_{T,r} = \frac{2}{T^2} \sum_{r=1}^{T-1} (T-r)\phi^{3r} (m_6 - m_3^2) + 3\phi^r m_4 \frac{\tau_2(1-\phi^{2r})}{1-\phi^2}$$

= $\frac{2}{T^2} \sum_{r=0}^{T-1} (T-r) \left((m_6 - m_3^2)\phi^{3r} + 3m_4\tau_2 \frac{\phi^r (1-\phi^{2r})}{1-\phi^2} \right) - \frac{2(m_6 - m_3^2)}{T}$
= $\frac{2}{T} \left(\frac{m_6 - m_3^2}{(1-\phi^3)} - \frac{3m_4\tau_2}{(1-\phi^3)(1-\phi^2)} + \frac{3m_4\tau_2}{(1-\phi)(1-\phi^2)} \right) - \frac{2(m_6 - m_3^2)}{T} + o\left(\frac{1}{T}\right)$

| - 6 | | | | |
|-----|---|---|---|--|
| | | | | |
| | | | | |
| | | | | |
| 14 | _ | _ | _ | |

And the following theorem calculates the covariance between \widehat{m}_2 and \widehat{m}_3

Theorem 0.0.13.

$$T \operatorname{Cov}[\widehat{m}_3, \widehat{m}_2] = (m_5 - m_3 m_2) \left(\frac{1}{1 - \phi^2} + \frac{1}{1 - \phi^3} - 1 \right) + 3\tau_2 m_3 \left(\frac{1}{(1 - \phi)(1 - \phi^2)} - \frac{1}{(1 - \phi^3)(1 - \phi^2)} \right) + o(1)$$

Proof.

$$\operatorname{Cov}[\widehat{m}_{3}, \widehat{m}_{2}] = \frac{\operatorname{Cov}(X_{1}^{3}, X_{1}^{2})}{T} + \underbrace{\frac{1}{T^{2}} \sum_{r=1}^{T-1} (T-r) \operatorname{Cov}(X_{1}^{3}, X_{1+r}^{2})}_{C_{T,r}^{1}} + \underbrace{\frac{1}{T^{2}} \sum_{r=1}^{T-1} (T-r) \operatorname{Cov}(X_{1}^{2}, X_{1+r}^{3})}_{C_{T,r}^{2}}$$

By repeating logic of the previous proof:

$$Cov(X_1^3, X_{1+r}^2) = Cov(X_1^3, \phi^{2r} X_1^2 + 2\phi^r X_1 R_r + R_r^2)$$

= $\phi^{2r} Cov(X_1^3, X_1^2) + 2\phi^r Cov(X_1^3, X_1 R_r)$
= $\phi^{2r} (m_5 - m_3 m_2) + 2\phi^r (m_4 E[R_r] - m_3 m_1 E[R_r]) = \phi^{2r} (m_5 - m_3 m_2)$

since $E[R_r] = 0$. Also,

$$Cov(X_1^2, X_{1+r}^3) = Cov(X_1^2, \phi^{3r} X_1^3 + 3\phi^{2r} X_1^2 R_r + 3\phi^r X_1 R_r^2 + R_r^3)$$

$$= \phi^{3r} Cov(X_1^2, X_1^3) + 3\phi^{2r} Cov(X_1^2, X_1^2 R_r) + 3\phi^r Cov(X_1^2, X_1 R_r^2)$$

$$= \phi^{3r} (m_5 - m_3 m_2) + 3\phi^{2r} (m_4 E[R_r] - m_2^2 E[R_r])$$

$$+ 3\phi^r (m_3 E[R_r^2] - m_2 m_1 E[R_r^2])$$

$$= \phi^{3r} (m_5 - m_3 m_2) + 3\tau_2 m_3 \frac{\phi^r (1 - \phi^{2r})}{1 - \phi^2}$$

since $m_1 = 0$, $E[R_r] = 0$ and $E[R_r^2] = \tau_2(1 - \phi^{2r})/(1 - \phi^2)$. Substituting we get,

$$C_{T,r}^{1} = \frac{(m_{5} - m_{3}m_{2})}{T^{2}} \sum_{r=1}^{T-1} (T-r)\phi^{2r} = \frac{(m_{5} - m_{3}m_{2})}{T} \left(\frac{1}{1-\phi^{2}} - 1\right) + o\left(\frac{1}{T}\right)$$

$$C_{T,r}^{2} = \frac{(m_{5} - m_{3}m_{2})}{T^{2}} \sum_{r=1}^{T-1} (T-r)\phi^{3r} + 3\tau_{2}m_{3} \sum_{r=1}^{T-1} (T-r)\frac{\phi^{r}(1-\phi^{2r})}{1-\phi^{2}}$$
$$= \frac{(m_{5} - m_{3}m_{2})}{T} \left(\frac{1}{1-\phi^{3}} - 1\right) + 3\tau_{2}m_{3} \left(\frac{1}{(1-\phi)(1-\phi^{2})} - \frac{1}{(1-\phi^{3})(1-\phi^{2})}\right)$$
$$+ o\left(\frac{1}{T}\right)$$

| г | _ | - | _ | |
|---|---|---|---|--|
| | | | | |
| | | | | |
| | | | | |

Here is the asymptotic variance of \widehat{m}_4

Theorem 0.0.14.

$$T \operatorname{Var}[\widehat{m}_4] = 2 \left[(m_8 - m_4^2) \left(\frac{1}{(1 - \phi^4)} - 1/2 \right) + 6\tau_2 (m_6 - m_4 m_2) \left(\frac{1}{(1 - \phi^2)^2} - \frac{1}{(1 - \phi^2)(1 - \phi^4)} \right) + 4m_5 \tau_3 \left(\frac{1}{(1 - \phi)(1 - \phi^3)} - \frac{1}{(1 - \phi^3)(1 - \phi^4)} \right) \right] + o(1)$$

Proof.

$$\operatorname{Var}[\widehat{m}_4] = \frac{(m_8 - m_4^2)}{T} + \frac{1}{T^2} \sum_{i=1}^T \sum_{j=1}^T \operatorname{Cov}(X_i^4, X_j^4)$$

Now verify there are exactly T - r terms whose covariance values are same and equal to $Cov(X_1^4, X_{1+r}^4)$. Thus,

$$\operatorname{Var}[\widehat{m}_{4}] = \frac{(m_{8} - m_{4}^{2})}{T} + \underbrace{\frac{2}{T^{2}} \sum_{r=1}^{T-1} (T-r) \operatorname{Cov}(X_{1}^{4}, X_{1+r}^{4})}_{C_{T,r}}$$

Thus, the problem boils down to the computation of $Cov(X_1^4, X_{1+r}^4)$. Now note that

$$X_{1+r} = \phi^r X_1 + \underbrace{\sum_{i=1}^r \phi^{r-i} \xi_{1+i}}_{R_r}$$

where R_r is independent of X_1 . Thus,

$$Cov(X_1^4, X_{1+r}^4) = Cov(X_1^4, \phi^{4r} X_1^4 + 4\phi^{3r} X_1^3 R_r + 6\phi^{2r} X_1^2 R_r^2 + 4\phi^r X_1 R_r^3 + R_r^4)$$

$$= \phi^{4r} Var(X_1^4) + 4\phi^{3r} Cov(X_1^4, X_1^3 R_r) + 6\phi^{2r} Cov(X_1^4, X_1^2 R_r^2)$$

$$+ 4\phi^r Cov(X_1^4, X_1 R_r^3)$$

$$= \phi^{4r} (m_8 - m_4^2) + 4\phi^{3r} (m_7 E[R_r] - m_4 m_3 E[R_r])$$

$$+ 6\phi^{2r} (m_6 E[R_r^2] - m_4 m_2 E[R_r^2]) + 4\phi^r (m_5 E[R_r^3] - m_4 m_1 E[R_r^3])$$

First note that $m_1 = 0$ and $E[R_r] = 0$ which allows us to simplify a bit.

$$\operatorname{Cov}(X_1^4, X_{1+r}^4) = \phi^{4r}(m_8 - m_4^2) + 6\phi^{2r}(m_6 E[R_r^2] - m_4 m_2 E[R_r^2]) + 4\phi^r m_5 E[R_r^3]$$

Now note that

$$E[R_r^2] = \frac{\tau_2(1-\phi^{2r})}{1-\phi^2}$$
$$E[R_r^3] = \frac{\tau_3(1-\phi^{3r})}{1-\phi^3}$$

Substituting we get,

$$\begin{split} C_{T,r} &= \frac{2}{T^2} \sum_{r=1}^{T-1} (T-r) \left(\phi^{4r} (m_8 - m_4^2) + \frac{6\tau_2 (1 - \phi^{2r})}{1 - \phi^2} \phi^{2r} (m_6 - m_4 m_2) + \frac{4\phi m_5 \tau_3 (1 - \phi^{3r})}{1 - \phi^3} \right) \\ &= \frac{2}{T^2} \sum_{r=0}^{T-1} (T-r) \left(\phi^{4r} (m_8 - m_4^2) + \frac{6\tau_2 (1 - \phi^{2r})}{1 - \phi^2} \phi^{2r} (m_6 - m_4 m_2) + \frac{4\phi m_5 \tau_3 (1 - \phi^{3r})}{1 - \phi^3} \right) \\ &- \frac{2(m_8 - m_4^2)}{T} \\ &= \frac{2}{T} \left(\frac{m_8 - m_4^2}{(1 - \phi^4)} - \frac{6\tau_2 (m_6 - m_4 m_2)}{(1 - \phi^2) (1 - \phi^4)} - \frac{4m_5 \tau_3}{(1 - \phi^3) (1 - \phi^4)} + \frac{6\tau_2 (m_6 - m_4 m_2)}{(1 - \phi^2)^2} \right) \\ &+ \frac{8m_5 \tau_3}{T(1 - \phi) (1 - \phi^3)} - \frac{2(m_8 - m_4^2)}{T} + o\left(\frac{1}{T}\right) \end{split}$$

where o(1/T) involves terms like $T\alpha^T$.

And lastly the asymptotic covariance between \widehat{m}_2 and \widehat{m}_4

Theorem 0.0.15.

$$Cov[\widehat{m}_4, \widehat{m}_2] = (m_6 - m_4 m_2) \left(\frac{1}{1 - \phi^2} + \frac{1}{1 - \phi^4} - 1 \right) + 6\tau_2 (m_4 - m_2^2) \left(\frac{1}{(1 - \phi^2)^2} - \frac{1}{(1 - \phi^2)(1 - \phi^4)} \right) + 4m_3 \tau_3 \left(\frac{1}{(1 - \phi)(1 - \phi^3)} - \frac{1}{(1 - \phi^3)(1 - \phi^4)} \right) + o(1)$$

Proof.

$$\begin{aligned} \operatorname{Cov}[\widehat{m}_{4}, \widehat{m}_{2}] &= \frac{\operatorname{Cov}(X_{1}^{4}, X_{1}^{2})}{T} + \underbrace{\frac{1}{T^{2}} \sum_{r=1}^{T-1} (T-r) \operatorname{Cov}(X_{1}^{4}, X_{1+r}^{2})}_{C_{T,r}^{1}} \\ &+ \underbrace{\frac{1}{T^{2}} \sum_{r=1}^{T-1} (T-r) \operatorname{Cov}(X_{1}^{2}, X_{1+r}^{4})}_{C_{T,r}^{2}} \end{aligned}$$

By repeating logic of the previous proof:

$$Cov(X_1^4, X_{1+r}^2) = Cov(X_1^4, \phi^{2r} X_1^2 + 2\phi^{2r} X_1 R_r + R_r^2)$$

= $\phi^{2r} Cov(X_1^4, X_1^2) + 2\phi^r Cov(X_1^4, X_1 R_r)$
= $\phi^{2r} (m_6 - m_4 m_2) + 2\phi^r (m_5 E[R_r] - m_1 E[R_r]) = \phi^{2r} (m_6 - m_4 m_2)$

since $E[R_r] = 0$. Also,

$$\begin{aligned} \operatorname{Cov}(X_1^2, X_{1+r}^4) &= \operatorname{Cov}(X_1^2, \phi^{4r} X_1^4 + 4\phi^{3r} X_1^3 R_r + 6\phi^{2r} X_1^2 R_r^2 + 4\phi X_1 R_r^3 + R_r^4) \\ &= \phi^{4r} \operatorname{Cov}(X_1^2, X_1^4) + 4\phi^{3r} \operatorname{Cov}(X_1^2, X_1^3 R_r) + 6\phi^{2r} \operatorname{Cov}(X_1^2, X_1^2 R_r^2) \\ &+ 4\phi^r \operatorname{Cov}(X_1^2, X_1 R_r^3) \\ &= \phi^{4r} (m_6 - m_4 m_2) + 4\phi^{3r} (m_5 E[R_r] - m_2 m_3 E[R_r]) \\ &+ 6\phi^{2r} (m_4 E[R_r^2] - m_2^2 E[R_r^2]) + 4\phi^r (m_3 E[R_r^3] - m_2 m_1 E[R_r^3]) \\ &= \phi^{4r} (m_6 - m_4 m_2) + 6\tau_2 (m_4 - m_2^2) \frac{(1 - \phi^{2r})\phi^{2r}}{1 - \phi^2} + 4\tau_3 m_3 \frac{(1 - \phi^{3r})\phi^r}{1 - \phi^3} \end{aligned}$$

since $E[R_r] = 0$, $E[R_r^2] = \frac{\tau_2(1-\phi^{2r})}{1-\phi^2}$, $E[R_r^3] = \frac{\tau_3(1-\phi^{3r})}{1-\phi^3}$. Substituting we get,

$$C_{T,r}^{1} = \frac{1}{T^{2}} \sum_{r=1}^{T-1} (T-r)(m_{6} - m_{4}m_{2})\phi^{2r} = \frac{1}{T} \frac{(m_{6} - m_{4}m_{2})}{1 - \phi^{2}} - \frac{(m_{6} - m_{4}m_{2})}{T} + o\left(\frac{1}{T}\right)$$

$$C_{T,r}^{2} = \frac{1}{T(1-\phi^{4})} \left(m_{6} - m_{4}m_{2} - \frac{6\tau_{2}(m_{4} - m_{2}^{2})}{1-\phi^{2}} - \frac{4m_{3}\tau_{3}}{1-\phi^{3}} \right) + \frac{1}{T} \left(\frac{6\tau_{2}(m_{4} - m_{2}^{2})}{(1-\phi^{2})^{4}} + \frac{4m_{3}\tau_{3}}{(1-\phi)(1-\phi^{3})} \right) - \frac{(m_{6} - m_{4}m_{2})}{T} + o\left(\frac{1}{T}\right)$$

| - | - | _ | _ | - |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

The asymptotic distributions of \widehat{m}_2 , \widehat{m}_3 and \widehat{m}_4 and the non-central chi-square *S* and *K* tests

Asymptotic theory plays a very important role in time series, since the exact distributions of statistics of time series can be impossible and complicated to derive, even for Gaussian processes unless in limited known very special cases. Many classical theorems were first proved under the independent and identically distribution assumptions, and extended to the dependent case. One way to do it, is through m-dependence approximation, which first introduced by HOEFFDING AND ROBBINS (1948) [40]. In this chapter we proved the asymptotic normality of the empirical second, third and fourth moments of a class of stationary linear process, in the first section. Then apply the Delta method a conjecture to prove the asymptotic normality of the empirical skewness and kurtosis of the same class of processes, in section 2. In section 3, we define the non-central chisquare distribution goodness of fit tests S and K, to reject the null hypotheses of Gaussian innovations in favor of Gumbel, Gumbel innovations and Gussian in favor of mixture of both and Gumbel innovations in favor of the Gaussian.

The asymptotic normality of \widehat{m}_2 , \widehat{m}_3 and \widehat{m}_4

The sequence $X_1, X_2, ...$ is m-dependent for some m > 0 if the vector $(X_1, X_2, ..., X_i)$ is independent of $(X_{i+j}, X_{i+j+1}, ...)$ whenever j > m. The central limit theorem (CLT) for stationary m-dependent sequence is used to prove the asymptotic normality of the second, third and fourth empirical moments of stationary linear processes.

Let X_t be a linear process of the form:

$$X_t = \sum_{i=-\infty}^{\infty} \phi^i \xi_{t-i} \tag{16}$$

where $\xi_t = \text{IID}(0, \sigma^2)$, and the coefficients satisfy $\sum_{i=-\infty}^{\infty} |\phi^i| < \infty$. then the second empirical moment is:

$$\widehat{m}_2 = \sum_{t=1}^T X_t^2 / T,$$

and the third and fourth empirical moments respectively are:

$$\widehat{m}_3 = \sum_{t=1}^T X_t^3 / T$$
 and $\widehat{m}_3 = \sum_{t=1}^T X_t^3 / T.$

The CLT for stationary m-dependent sequence and "an approximation" theorem are the main tools used in the proof. Define

$$X_t^m = \sum_{i=-m}^m \phi_m^i \xi_{t-i} \tag{17}$$

where $\phi_m^i = \phi^i$ if $|i| \le m$ and 0 otherwise, then X_t^m is 2m-dependent and strictly stationary.

The idea of the proof is using truncation to construct an approximating sequence y_{nm} of random variables, that is similar to X_n , but their asymptotic distribution is obtainable. These theorems are stated as follow respectively without proofs, where the proofs can be found in [85, 75].

Theorem 0.0.16. If X_t is a strictly stationary m-dependent sequence of random variables with mean zero and autocovariance function $\gamma(\cdot)$ and if

$$V_m = \sum_{u=-m}^{u=m} \gamma(u)$$

where $V_m \neq 0$, then

$$\bar{X_n} \sim AN(0, V_m/n)$$

Theorem 0.0.17. Let x_n for n = 1, 2, ... and y_{mn} for m = 1, 2, ... be random $K \times 1$ vector such that:

- (i) $y_{mn} \xrightarrow{d} y_m$ as $n \to \infty$ for each m;
- (*ii*) $y_m \xrightarrow{d} y$ as $n \to \infty$;
- (iii) $\lim_{m \to \infty} \limsup_{n \to 0} \Pr\{|x_n y_{mn}| > \epsilon\} = 0 \text{ for every } \epsilon > 0. By Tchebycheff inequality, this condition is equivalent to E{|x_n y_{mn}|^2} \to 0, which is easier to verify.$

Then, $x_n \xrightarrow{d} y$

We begin by providing LLN for Stationary and Ergodic Time Series (Hamilton [38]).

Let X_t be a stationary process with mean $E[X_t] = \mu$ and autocovariance $\gamma_j = cov(X_t, X_{t-j})$. If

$$\sum_{j=0}^{\infty} |\gamma_j| < \infty$$

then X_t is ergodic for the mean. That is, $\overline{X} \xrightarrow{p} E[X_t] = \mu$. We state Ergodic Theorem as follows:

Theorem 0.0.18. Let X_t be stationary and ergodic with $E[X_i] = \mu$. Then

$$\bar{X} = \frac{1}{T} \sum_{t=1}^{T} X_t \xrightarrow{p} E[X_t] = \mu.$$

Remark 0.0.19. The ergodic theorem says that for a stationary and ergodic sequence $\{X_t\}$ the time average converges to the ensemble average as the sample size gets large.

Remark 0.0.20. Any transformation $g(\cdot)$ of a stationary and ergodic process $\{X_t\}$ is also stationary and ergodic. That is, $\{g(X_t)\}$ is stationary and ergodic. Therefore, if $E[g(X_t)]$ exists then the ergodic theorem gives

$$\bar{g} = \frac{1}{T} \sum_{t=1}^{T} g(X_t) \xrightarrow{p} E[g(X_t)].$$

as an application by take, $g(x) = x^i$, where i = 2, 3, 4 we get $\widehat{m_i} \xrightarrow{p} m_i$, where $\widehat{m_i} = \frac{1}{T} \sum_{t=1}^{T} X^i$ and $m_i = E[X^i]$.

Next we prove formulas for the variance and the asymptotic variance of \hat{m}_2 .

Theorem 0.0.21. Let X_t be a linear process of the form:

$$X_t = \sum_{i=-\infty}^{\infty} \phi^i \xi_{t-i}$$

where $\xi_t = \text{IID}(0, \sigma^2)$, and the coefficients satisfy $\sum_{i=-\infty}^{\infty} |\phi^i| < \infty$, and $\text{E}(\xi_t^4) < \infty$, then

$$T \operatorname{Var}(\widehat{m}_2) \to \sum_{r=-\infty}^{\infty} \gamma_2(r)$$

where

$$\gamma_2(r) = \operatorname{Cov}(X_t^2, X_{t+r}^2)$$
$$= \phi^{2r}(m_4 - m_2^2)$$

 $m_4 = \operatorname{E}(X_t^4)$ and $m_2 = \operatorname{E}(X_t^2)$

Proof.

$$\operatorname{Var}(\widehat{m_2}) = \frac{1}{T^2} \sum_{i=1}^{T} \sum_{j=1}^{n} \operatorname{Cov}(X_i^2, X_j^2),$$
$$= \frac{1}{T^2} \sum_{r=-(T-1)}^{T-1} (T-r) \operatorname{Cov}(X_1^2, X_{1+r}^2)$$

Now write X_t as:

$$X_{1+r} = \phi^r X_1 + \underbrace{\sum_{i=1}^r \phi^{r-i} \xi_{1+i}}_{R_r}$$

where R_r is independent of X_1 , with $E[R_r] = 0$ and $Var[R_r] = \frac{\sigma^2(1-\phi^{2r})}{1-\phi^2}$. Then the crosscovariance function look like the covariance function of X_1 . Thus,

$$Cov(X_1^2, X_{1+r}^2) = Cov(X_1^2, \phi^{2r} X_1^2 + 2\phi^r X_1 R_r + R_r^2)$$

= $\phi^{2r} Var(X_1^2) + 2\phi^r Cov(X_1^2, X_1 R_r)$
= $\phi^{2r} (m_4 - m_2^2) + 2\phi^r (m_3 E[R_r] - m_2 m_1 E[R_r])$
= $\phi^{2r} (m_4 - m_2^2)$

Let $\gamma_2(r) = \operatorname{Cov}(X_t^2, X_{t+r}^2)$, then

$$T \operatorname{Var}(\widehat{m_2}) = \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \gamma_2(r),$$

and since $\sum_{i} |\phi^{i}| < \infty$, this implies $\sum_{r} |\gamma_{2}(r)| < \infty$. Note that: $|(1 - \frac{|r|}{T})\gamma_{2}(r)| < |\gamma_{2}(r)|$ and $(1 - \frac{|r|}{T})\gamma_{2}(r) \rightarrow \gamma_{2}(r)$, and take $f_{n}(r) = (1 - \frac{|r|}{n})\gamma_{2}(r)$, for |r| < n and zero for |r| > n. Let $\mu(r) = 1$ for $r = \pm 1, \pm 2, \ldots$ be the counting measure then by the Dominated Convergence Theorem we have $T\operatorname{Var}(\widehat{m}_{2}) \rightarrow \sum_{r=-\infty}^{\infty} \gamma_{2}(r)$.

The above result for the variance enables proving the following CLT for the statistic \hat{m}_2 .

Theorem 0.0.22. If X_t is a stationary linear process and under the assumption of Theorem

0.0.21, then:

$$\widehat{m}_2 \sim AN(m_2, V/T)$$

Where $m_2 = E(X_t^2)$ and $V = \sum_{r=-\infty}^{\infty} \gamma_2(r)$.

Proof. First define the stationary 2m-dependent process

$$y_t^m = (X_t^m)^2,$$

where X_t^m is the approximation defined in (17) to X_t , then:

$$\bar{y}_{mT} = T^{-1} \sum_{t=1}^{T} y_t^m$$
$$= T^{-1} \left[(X_1^m)^2 + \dots + (X_T^m)^2 \right]$$

also $E(y_t^m) = E(X_t^m)^2$ is the second theoretical moment of the 2m dependent process X_t^m . Then consider

$$y_{mT} = T^{1/2} \left[\bar{y}_{mT} - E(\bar{y}_{mT}) \right]$$

as an approximation to

$$y_T = T^{1/2}(\hat{m}_2 - m_2),$$

where $E(\bar{y}_{mT})$ is the same as $E(y_t^m)$ given above. To obtain an asymptotic distribution for y_T , we apply Theorem 0.0.17, using y_{mT} as our approximation. Now verying (i),(ii) and (iii) of Theorem 0.0.17.

(i) Applying Theorem 0.0.16, (CLT for m-dependent sequence), to the 2m dependent series

 y_{mT} we obtain:

$$y_{mT} = T^{1/2} \left[\bar{y}_{mT} - \mathcal{E}(\bar{y}_{mT}) \right] \xrightarrow{d} y_m \sim N(0, V_m)$$

as $T \to \infty$ where $V_m = \sum_{r=-m}^{m} \gamma_2(r)$. (ii) Note that as $m \to \infty$, $V_m \to V$, using the Dominated Convergence Theorem, where V is defined in Theorem 0.0.22. Hence the characteristic function of y_m ,

$$\varphi_m(\lambda) = \exp\{-1/2\lambda^2 V_m\} \to \exp\{-1/2\lambda^2 V\}$$

as $m \to \infty$, which is the characteristic function of a random variable $y \sim N(0, V)$ (characterizing convergence of distribution of $F_n(.)$ in terms of convergence of sequence of characteristic functions $\varphi_n(.)$, i.e., $\varphi_n(\lambda) \to \varphi(\lambda) \Leftrightarrow F_n(x) \xrightarrow{d} F(x)$). (iii)To verify the last condition, note that:

$$E(y_T - y_{mn})^2 = TE \left[(\hat{m}_2 - E(X_t^2)) - (\bar{y}_{mT} - E(\bar{y}_{mT})) \right]^2$$

= $TE (\hat{m}_2 - E(X_t^2))^2 + TE (\bar{y}_{mT} - E(\bar{y}_{mT}))^2$
 $- 2TE ((\hat{m}_2 - E(X_t^2))(\bar{y}_{mT} - E(\bar{y}_{mT})))$
= $TVar(\hat{m}_2) + TVar(\bar{y}_{mT}) - 2TCov(\hat{m}_2, \bar{y}_{mT})$

using the calculations that led to the result in Theorem 0.0.21, and letting $m \to \infty$, we get

$$T\operatorname{Var}(\widehat{m}_2) + T\operatorname{Var}(\widehat{m}_2) - 2T\operatorname{Var}(\widehat{m}_2),$$

hence

$$(V+V-2V)=0$$

as $T \to \infty$.

In order to prove the limiting distribution of \hat{m}_3 we start by providing formulas for the variance and asymptotic variance of \hat{m}_3 .

Theorem 0.0.23. Let X_t be a linear process of the form:

$$X_t = \sum_{i=-\infty}^{\infty} \phi^i \xi_{t-i}$$

where $\xi_t = \text{IID}(0, \sigma^2)$, and the coefficients satisfy $\sum_{i=-\infty}^{\infty} |\phi^i| < \infty$, and $\text{E}(\xi_t^6) < \infty$, then

$$T \operatorname{Var}(\widehat{m}_3) \to \sum_{r=-\infty}^{\infty} \gamma_3(r)$$

where

$$\gamma_3(r) = \operatorname{Cov}(X_t^3, X_{t+r}^3)$$

= $\phi^{3r}(m_6 - m_3^2) + 3\phi^r m_4 \frac{\tau^2(1 - \phi^{2r})}{1 - \phi^2},$

 $m_6 = \mathcal{E}(X_t^6), \ m_3 = \mathcal{E}(X_t^3) \ and \ m_4 = \mathcal{E}(X_t^4).$

Proof.

$$\operatorname{Var}(\widehat{m_3}) = \frac{1}{T^2} \sum_{i=1}^{T} \sum_{j=1}^{T} \operatorname{Cov}(X_i^3, X_j^3),$$
$$= \frac{1}{T^2} \sum_{r=-(T-1)}^{T-1} (T-r) \operatorname{Cov}(X_1^3, X_{1+r}^3)$$

By repeating the same argument in proof of Theorem 0.0.21 we get:

$$\begin{aligned} \operatorname{Cov}(X_1^3, X_{1+r}^3) &= \operatorname{Cov}(X_1^3, \phi^{3r} X_1^3 + 3\phi^{2r} X_1^2 R_r + 3\phi^r X_1 R_r^2 + R_r^3) \\ &= \phi^{3r} \operatorname{Var}(X_1^3) + 3\phi^{2r} \operatorname{Cov}(X_1^3, X_1^2 R_r) + 3\phi^r \operatorname{Cov}(X_1^3, X_1 R_r^2) \\ &= \phi^{3r} (m_6 - m_3^2) + 3\phi^{2r} (m_5 E[R_r] - m_3 m_2 E[R_r]) \\ &+ 3\phi^r (m_4 E[R_r^2] - m_3 m_1 E[R_r^2]) \\ &= \phi^{3r} (m_6 - m_3^2) + 3\phi^r m_4 E[R_r^2] \\ &= \phi^{3r} (m_6 - m_3^2) + 3\phi^r m_4 \frac{\tau^2 (1 - \phi^{2r})}{1 - \phi^2} \end{aligned}$$

Let $\gamma_3(r) = \operatorname{Cov}(X_t^3, X_{t+r}^3)$, then

$$T \operatorname{Var}(\widehat{m_3}) = \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \gamma_3(r),$$

and since $\sum_{i} |\phi^{i}| < \infty$, this implies $\sum_{r} |\gamma_{3}(r)| < \infty$. Note that: $\left| \left(1 - \frac{|r|}{T} \right) \gamma_{3}(r) \right| < |\gamma_{3}(r)|$ and $\left(1 - \frac{|r|}{T} \right) \gamma_{3}(r) \rightarrow \gamma_{3}(r)$, and take $f_{n}(r) = \left(1 - \frac{|r|}{n} \right) \gamma_{3}(r)$, for |r| < n and zero for |r| > n. Let

 $\mu(r) = 1$ for $r = \pm 1, \pm 2, ...$ be the counting measure then by the Dominated Convergence Theorem we have $T\operatorname{Var}(\widehat{m}_3) \to \sum_{r=-\infty}^{\infty} \gamma_3(r)$.

Now the following theorem proves the asymptotic distribution for \hat{m}_3 . Note that the proof will be very similar to the proof of Theorem0.0.22, where the same method and theorems are used, but suitable approximations and estimators.

Theorem 0.0.24. If X_t is a stationary linear process and under the assumption of Theorem 0.0.23, then:

$$\widehat{m}_3 \sim AN(m_3, V/T)$$

Where $m_3 = \mathcal{E}(x_t^3)$ and $V = \sum_{r=-\infty}^{\infty} \gamma_3(r)$.

Proof. First define the stationary 2m-dependent process

$$y_t^m = (X_t^m)^3,$$

where X_t^m is the approximation defined in (17) to X_t , then:

$$\bar{y}_{mT} = T^{-1} \sum_{t=1}^{T} y_t^m$$
$$= T^{-1} \left[(X_1^m)^3 + \dots + (X_n^m)^3 \right]$$

also $\mathcal{E}(y_t^m) = \mathcal{E}(X_t^m)^3$ is the third theoretical moment of the 2m dependent process X_t^m .

Then consider

$$y_{mT} = T^{1/2} \left[\bar{y}_{mT} - E(\bar{y}_{mT}) \right]$$

as an approximation to

$$y_T = T^{1/2}(\widehat{m}_3 - m_3),$$

where $E(\bar{y}_{mT})$ is the same as $E(y_t^m)$ given above. To obtain an asymptotic distribution for y_T , we apply Theorem 0.0.17, using y_{mT} as our approximation. Now verifying (i),(ii) and (iii) of Theorem 0.0.17.

(i) Applying Theorem 0.0.16, (CLT for m-dependent sequence), to the 2m dependent series y_{mT} we obtain:

$$y_{mT} = T^{1/2} \left[\bar{y}_{mT} - \mathcal{E}(\bar{y}_{mT}) \right] \xrightarrow{d} y_m \sim N(0, V_m)$$

as $T \to \infty$ where $V_m = \sum_{r=-m}^m \gamma_3(r)$. (ii) Note that as $m \to \infty$, $V_m \to V$, using the Dominated Convergence Theorem, where V is defined in Theorem 0.0.24. Hence the characteristic function of y_m ,

$$\varphi_m(\lambda) = \exp\{-1/2\lambda^2 V_m\} \to \exp\{-1/2\lambda^2 V\}$$

as $m \to \infty$, which is the characteristic function of a random variable $y \sim N(0, V)$ (characterizing convergence of distribution of $F_n(.)$ in terms of convergence of sequence of characteristic functions $\varphi_n(.)$, i.e., $\varphi_n(\lambda) \to \varphi(\lambda) \Leftrightarrow F_n(x) \xrightarrow{d} F(x)$). (iii) To verify the last condition, note that:

$$E(y_T - y_{mT})^2 = TE\left[(\hat{m}_3 - E(X_t^3)) - (\bar{y}_{mT} - E(\bar{y}_{mT}))\right]^2$$

= $TE(\hat{m}_3 - E(X_t^3))^2 + TE(\bar{y}_{mT} - E(\bar{y}_{mT}))^2$
 $- 2TE((\hat{m}_3 - E(X_t^3))(\bar{y}_{mT} - E(\bar{y}_{mn})))$
= $TVar(\hat{m}_3) + TVar(\bar{y}_{mT}) - 2TCov(\hat{m}_3, \bar{y}_{mT})$

using the calculations that led to the result in Theorem 0.0.23, and letting $m \to \infty$, we get

$$T\operatorname{Var}(\widehat{m}_3) + T\operatorname{Var}(\widehat{m}_3) - 2T\operatorname{Var}(\widehat{m}_3),$$

hence

$$(V+V-2V)=0$$

as $T \to \infty$.

The same technique above applied to find the limiting distribution of \hat{m}_4 . First we start by providing formulas for the variance and asymptotic variance of \hat{m}_4 .

Theorem 0.0.25. Let X_t be a linear process of the form:

$$X_t = \sum_{i=-\infty}^{\infty} \phi^i \xi_{t-i}$$

where $\xi_t = \text{IID}(0, \sigma^2)$, the coefficients satisfy $\sum_{i=-\infty}^{\infty} |\phi^i| < \infty$ and $\text{E}(\xi_t^8) < \infty$, then

$$T\operatorname{Var}(\widehat{m}_4) \to \sum_{r=-\infty}^{\infty} \gamma_4(r)$$

where

$$\gamma_4(r) = \operatorname{Cov}(X_t^4, X_{t+r}^4)$$

= $\phi^{4r}(m_8 - m_4^2) + \frac{6\tau_2(1 - \phi^{2r})}{1 - \phi^2}\phi^{2r}(m_6 - m_4m_2) + \frac{4\phi m_5\tau_3(1 - \phi^{3r})}{1 - \phi^3}$

 $m_i = \mathcal{E}(X_t^i) \text{ for } i = 2, 4, 5, 6, 8, \text{ and } \tau_i = E(\xi^i) \text{ for } i = 2, 3, 4, 5, 6, 8.$

Proof.

$$\operatorname{Var}(\widehat{m_4}) = \frac{1}{T^2} \sum_{i=1}^{T} \sum_{j=1}^{T} \operatorname{Cov}(X_i^4, X_j^4),$$
$$= \frac{1}{T^2} \sum_{r=-(T-1)}^{T-1} (T-r) \operatorname{Cov}(X_1^4, X_{1+r}^4)$$

By repeating the same argument in proof of Theorem 0.0.21 we get:

$$\begin{aligned} \operatorname{Cov}(X_1^4, X_{1+r}^4) &= \operatorname{Cov}(X_1^4, \phi^{4r} X_1^4 + 4\phi^{3r} X_1^3 R_r + 6\phi^{2r} X_1^2 R_r^2 + 4\phi^r X_1 R_r^3 + R_r^4) \\ &= \phi^{4r} \operatorname{Var}(X_1^4) + 4\phi^{3r} \operatorname{Cov}(X_1^4, X_1^3 R_r) + 6\phi^{2r} \operatorname{Cov}(X_1^4, X_1^2 R_r^2) \\ &+ 4\phi^r \operatorname{Cov}(X_1^4, X_1 R_r^3) \\ &= \phi^{4r} (m_8 - m_4^2) + 4\phi^{3r} (m_7 E[R_r] - m_4 m_3 E[R_r]) \\ &+ 6\phi^{2r} (m_6 E[R_r^2] - m_4 m_2 E[R_r^2]) + 4\phi^r (m_5 E[R_r^3] - m_4 m_1 E[R_r^3]) \end{aligned}$$

Let $\gamma_4(r) = \operatorname{Cov}(X_t^4, X_{t+r}^4)$, then

$$T \operatorname{Var}(\widehat{m_4}) = \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \gamma_4(r),$$

and since $\sum_{i} |\phi^{i}| < \infty$, this implies $\sum_{r} |\gamma_{4}(r)| < \infty$. Note that: $\left| \left(1 - \frac{|r|}{T} \right) \gamma_{4}(r) \right| < |\gamma_{4}(r)|$ and $\left(1 - \frac{|r|}{T} \right) \gamma_{4}(r) \rightarrow \gamma_{4}(r)$, and take $f_{n}(r) = \left(1 - \frac{|r|}{n} \right) \gamma_{4}(r)$, for |r| < n and zero for |r| > n. Let $\mu(r) = 1$ for $r = \pm 1, \pm 2, \ldots$ be the counting measure then by the Dominated Convergence Theorem we have $T \operatorname{Var}(\widehat{m}_{4}) \rightarrow \sum_{r=-\infty}^{\infty} \gamma_{4}(r)$.

Now the following theorem proves the asymptotic distribution for \hat{m}_4 . Note that the proof will be very similar to the proof of Theorem0.0.22, where the same method and theorems are used, but suitable approximations and estimators.

Theorem 0.0.26. If X_t is a stationary linear process and under the assumption of Theorem

0.0.25, then:

$$\widehat{m}_4 \sim AN(m_4, V/T)$$

Where $m_4 = \mathcal{E}(x_t^4)$ and $V = \sum_{r=-\infty}^{\infty} \gamma_4(r)$.

Proof. First define the stationary 2m-dependent process

$$y_t^m = (X_t^m)^4,$$

where X_t^m is the approximation defined in (17) to X_t , then:

$$\bar{y}_{mT} = T^{-1} \sum_{t=1}^{T} y_t^m$$
$$= T^{-1} \left[(X_1^m)^4 + \dots + (X_n^m)^4 \right]$$

also $E(y_t^m) = E(X_t^m)^4$ is the fourth theoretical moment of the 2m dependent process X_t^m . Then consider

$$y_{mT} = T^{1/2} \left[\bar{y}_{mT} - E(\bar{y}_{mT}) \right]$$

as an approximation to

$$y_T = T^{1/2}(\widehat{m}_4 - m_4),$$

where $E(\bar{y}_{mT})$ is the same as $E(y_t^m)$ given above. To obtain an asymptotic distribution for y_T , we apply Theorem 0.0.17, using y_{mT} as our approximation. Now verifying (i),(ii) and (iii) of Theorem 0.0.17.

(i) Applying Theorem 0.0.16, (CLT for m-dependent sequence), to the 2m dependent series

 y_{mT} we obtain:

$$y_{mT} = T^{1/2} \left[\bar{y}_{mT} - \mathcal{E}(\bar{y}_{mT}) \right] \xrightarrow{d} y_m \sim N(0, V_m)$$

as $T \to \infty$ where $V_m = \sum_{r=-m}^{m} \gamma_4(r)$. (ii) Note that as $m \to \infty$, $V_m \to V$, using the Dominated Convergence Theorem, where V is defined in Theorem 0.0.25. Hence the characteristic function of y_m ,

$$\varphi_m(\lambda) = \exp\{-1/2\lambda^2 V_m\} \to \exp\{-1/2\lambda^2 V\}$$

as $m \to \infty$, which is the characteristic function of a random variable $y \sim N(0, V)$ (characterizing convergence of distribution of $F_n(.)$ in terms of convergence of sequence of characteristic functions $\varphi_n(.)$, i.e., $\varphi_n(\lambda) \to \varphi(\lambda) \Leftrightarrow F_n(x) \xrightarrow{d} F(x)$). (iii)To verify the last condition, note that:

$$E(y_T - y_{mT})^2 = TE\left[(\hat{m}_4 - E(X_t^4)) - (\bar{y}_{mT} - E(\bar{y}_{mT}))\right]^2$$

= $TE(\hat{m}_4 - E(X_t^4))^2 + TE(\bar{y}_{mT} - E(\bar{y}_{mT}))^2$
 $- 2TE((\hat{m}_4 - E(X_t^4))(\bar{y}_{mT} - E(\bar{y}_{mn})))$
= $TVar(\hat{m}_4) + TVar(\bar{y}_{mT}) - 2TCov(\hat{m}_4, \bar{y}_{mT})$

using the calculations that led to the result in Theorem 0.0.23, and letting $m \to \infty$, we get

$$T\operatorname{Var}(\widehat{m}_4) + T\operatorname{Var}(\widehat{m}_4) - 2T\operatorname{Var}(\widehat{m}_4),$$

hence

$$(V+V-2V)=0$$
as $T \to \infty$.

We state the following lemmas which contain the asymptotic covariances of (\hat{m}_2, \hat{m}_3) and (\hat{m}_2, \hat{m}_4) , without proof, since the idea of the proofs is the same as of the previous theorems 0.0.21, 0.0.23 and 0.0.25.

Lemma 0.0.27. Let X_t be a linear process of the form:

$$X_t = \sum_{i=-\infty}^{\infty} \phi^i \xi_{t-i}$$

where $\xi_t = \text{IID}(0, \sigma^2)$, the coefficients satisfy $\sum_{i=-\infty}^{\infty} |\phi^i| < \infty$, and $\text{E}(\xi_t^5) < \infty$, then

$$T \operatorname{Cov}[\widehat{m}_3, \widehat{m}_2] \to \sum_{r=-\infty}^{\infty} \gamma_{23}(r) + \sum_{r=-\infty}^{\infty} \gamma_{32}(r)$$

where

$$\gamma_{23}(r) = \operatorname{Cov}(X_t^2, X_{t+r}^3)$$

= $\phi^{3r}(m_5 - m_3m_2) + 3\tau_2m_3\frac{\phi^r(1 - \phi^{2r})}{1 - \phi^2}$

and

$$\gamma_{32}(r) = \text{Cov}(X_t^3, X_{t+r}^2)$$

= $\phi^{2r}(m_5 - m_3 m_2)$

65

Lemma 0.0.28. Let X_t be a linear process of the form:

$$X_t = \sum_{i=-\infty}^{\infty} \phi^i \xi_{t-i}$$

where $\xi_t = \text{IID}(0, \sigma^2)$, the coefficients satisfy $\sum_{i=-\infty}^{\infty} |\phi^i| < \infty$, and $\text{E}(\xi_t^6) < \infty$, then

$$T \operatorname{Cov}[\widehat{m}_4, \widehat{m}_2] \to \sum_{r=-\infty}^{\infty} \gamma_{24}(r) + \sum_{r=-\infty}^{\infty} \gamma_{42}(r)$$

where

$$\gamma_{24}(r) = \operatorname{Cov}(X_t^2, X_{t+r}^4)$$
$$= \phi^{4r}(m_6 - m_4 m_2) + 6\tau_2(m_4 - m_2^2) \frac{(1 - \phi^{2r})\phi^{2r}}{1 - \phi^2} + 4\tau_3 m_3 \frac{(1 - \phi^{3r})\phi^r}{1 - \phi^3}$$

and

$$\gamma_{42}(r) = \text{Cov}(X_t^4, X_{t+r}^2)$$

= $\phi^{2r}(m_6 - m_4 m_2)$

The asymptotic normality of the empirical skewness and kurtosis

In this section, the delta method was applied to get the asymptotic distribution of the empirical skewness and kurtosis of a stationary linear model to the asymptotic multivariate distribution of the vectors $M_T = (\hat{m}_2, \hat{m}_3)$ and $M_T = (\hat{m}_4, \hat{m}_2)$.

We start by a conjecture for the asymptotic multivariate distribution of the vector $M_T = (\hat{m}_2, \hat{m}_3)$.

Theorem 0.0.29. If X_t is a stationary linear process and under the assumptions of Theorem 0.0.23, then: For $M_T = (\hat{m}_2, \hat{m}_3)$, we have

$$\sqrt{T}(M_T - \mu) \stackrel{d}{\Longrightarrow} N(0, \Sigma)$$
 (18)

where

$$\mu = [m_2, m_3]^{\top}$$

$$\Sigma_{11} = \sum_{r=-\infty}^{\infty} \gamma_2(r)$$

$$\Sigma_{12} = \sum_{r=-\infty}^{\infty} \gamma_{23}(r) + \sum_{r=-\infty}^{\infty} \gamma_{32}(r)$$

$$\Sigma_{22} = \sum_{r=-\infty}^{\infty} \gamma_3(r)$$

Now applying the delta method we get the asymptotic distribution for the empirical skewness of the linear process X_t in the following theorem.

Theorem 0.0.30. For $S_T = \hat{m}_3 / \hat{m}_2^{3/2}$,

$$\sqrt{T}\left(S_T - \frac{m_3}{m_2^{3/2}}\right) \stackrel{d}{\Longrightarrow} N(0,\sigma^2)$$

where σ^2 is given by

$$\sigma^{2} = \left[-\frac{3m_{3}}{2m_{2}^{5/2}}, \frac{1}{m_{2}^{3/2}} \right] \Sigma \left[-\frac{3m_{3}}{2m_{2}^{5/2}}, \frac{1}{m_{2}^{3/2}} \right]^{\top}$$
$$= \frac{9m_{3}^{2}\Sigma_{11}}{4m_{2}^{5}} - \frac{6m_{3}\Sigma_{12}}{2m_{2}^{4}} + \frac{\Sigma_{22}}{m_{2}^{3}}$$

Proof. Applying theorem 0.0.29 and the multivariate delta method, where the function g is of the form $g(s_1, s_2) = \frac{s_2}{s_1^{3/2}}$, we get the required.

The same theory of the empirical skewness is applied to the empirical kurtosis. We start by the following conjecture for the asymptotic normality of the vector (\hat{m}_2, \hat{m}_4) .

Theorem 0.0.31. If X_t is a stationary linear process and under the assumptions of Theorem 0.0.25, then: For $M_T = (\hat{m}_2, \hat{m}_4)$, we have

$$\sqrt{T}(M_T - \mu) \stackrel{d}{\Longrightarrow} N(0, \Sigma) \tag{19}$$

where

$$\mu = [m_2, m_4]^{\top}$$
$$\Sigma_{11} = \sum_{r=-\infty}^{\infty} \gamma_2(r)$$
$$\Sigma_{12} = \sum_{r=-\infty}^{\infty} \gamma_{24}(r) + \sum_{r=-\infty}^{\infty} \gamma_{42}(r)$$
$$\Sigma_{22} = \sum_{r=-\infty}^{\infty} \gamma_4(r)$$

And lastly the distribution of ${\cal K}_T$

Theorem 0.0.32. For $K_T = \hat{m}_4 / \hat{m}_2^2$,

$$\sqrt{T}\left(K_T - \frac{m_4}{m_2^2}\right) \stackrel{d}{\Longrightarrow} N(0, \sigma^2)$$

where σ^2 is given by

$$\sigma^{2} = \left[-\frac{2m_{4}}{m_{2}^{3}}, \frac{1}{m_{2}^{2}} \right] \Sigma \left[-\frac{2m_{4}}{m_{2}^{3}}, \frac{1}{m_{2}^{2}} \right]^{\top} = \frac{4m_{4}^{2}\Sigma_{11}}{m_{2}^{6}} - \frac{4m_{4}\Sigma_{12}}{m_{2}^{5}} + \frac{\Sigma_{22}}{m_{2}^{4}}$$

Proof. Applying theorem 0.0.31 and the multivariate delta method, where the function g is of the form $g(s_1, s_2) = \frac{s_2}{s_1^2}$, we get the required.

The non-central chi-square tests

In this section, the empirical skewness and kurtosis are used to define a non-central chisquare distributions S and K, to reject the null hypotheses of Gaussian innovations in favor of Gumbel, Gumbel innovations and Gussian in favor of mixture of both and Gumbel innovations in favor of the Gaussian.

Theorem 0.0.33. Let X_t be an AR(1) process of the form:

$$X_{t+1} = \phi X_t + \xi_{t+1}$$

where $\xi_t \sim WN(0, \sigma^2)$, with $E(\xi_t)^8 < \infty$ and let $S = \hat{m}_3 / \hat{m}_2^{3/2}$ and $K = \hat{m}_4 / \hat{m}_2^2$, then for large T,

- the hypothesis test:
- $H_0: \{\xi_t\}$ follow Gaussian distribution $H_a: \{\xi_t\}$ follow Gumbel distribution

has a rejection region

$$\{x_t : (S+b_S)^2 \le c\} \quad if \quad (1/\sigma_{S_a}^2 - 1/\sigma_{S_0}^2) > 0 \tag{20}$$

and

$$\{x_t : (S+b_S)^2 > c\} \quad if \quad (1/\sigma_{S_a}^2 - 1/\sigma_{S_0}^2) < 0 \tag{21}$$

where $\sigma_{S_0}^2$ and $\sigma_{S_a}^2$ are the variances of S under the null and alternative hypotheses,

respectively, $b_S = \left(\frac{\mu_{S_0}\sigma_{S_a}^2 - \mu_{S_a}\sigma_{S_0}^2}{\sigma_{S_0}^2 - \sigma_{S_a}^2}\right)$ and c is a constant. Moreover, under the null hypothesis assumption, the test statistic

$$(S+b_S)^2/\sigma_{S_0}^2 \sim \chi_1^2 (\frac{\mu_{S_0}+b_S}{\sigma_{S_0}})^2.$$
(22)

• the hypothesis test:

 $H_0: \{\xi_t\}$ follow Gumbel distribution $H_a: \{\xi_t\}$ follow Gaussian distribution

has a rejection region

$$\{x_t : (K+b_K)^2 \le c\} \quad if \quad (1/\sigma_{Ka}^2 - 1/\sigma_{K0}^2) > 0 \tag{23}$$

and

$$\{x_t : (K+b_K)^2 > c\} \quad if \quad (1/\sigma_{K_a}^2 - 1/\sigma_{K_0}^2) < 0 \tag{24}$$

where $\sigma_{K_0}^2$ and $\sigma_{K_a}^2$ are the variances of K under the null and alternative hypotheses, respectively, $b_K = \left(\frac{\mu_{K_0}\sigma_{K_a}^2 - \mu_{K_a}\sigma_{K_0}^2}{\sigma_{K_0}^2 - \sigma_{K_a}^2}\right)$ and c is a constant. Moreover, under the null hypothesis assumption, the test statistic $(K + b_K)^2 / \sigma_{K_0}^2 \sim \chi_1^2 (\frac{\mu_{K_0} + b_K}{\sigma_{K_0}})^2$.

Proof. • Let $f_0(s)$ and $f_a(s)$ be the density functions of the statistic S under the null and alternative hypotheses, respectively. Applying the Neyman-Pearson Lemma to S,

we reject H_0 if

$$\frac{f_0(s)}{f_a(s)} \le c',$$

where c' is a constant determined by α (the probability of type I error).

Recall that, $S \sim AN(\mu_S, \sigma_S^2)$, which under the null hypothesis assumption (the innovations follow Gaussian distribution) implies $S \sim AN(\mu_{S_0}, \sigma_{S_0}^2)$, and $S \sim AN(\mu_{S_a}, \sigma_{S_a}^2)$ under the alternative (the innovations follow Gumbel distribution).

This implies

$$\left(\frac{1}{\sqrt{2\pi\sigma_{S_0}^2}}e^{-(s-\mu_{S_0})^2/2\sigma_{S_0}^2} \middle/ \frac{1}{\sqrt{2\pi\sigma_{S_a}^2}}e^{-(s-\mu_{S_a})^2/2\sigma_{S_a}^2}\right) \le c',$$

hence,

$$s^{2} \left(\frac{1}{\sigma_{S_{a}}^{2}} - \frac{1}{\sigma_{S_{0}}^{2}}\right) + 2s \left(\frac{\mu_{S_{0}}}{\sigma_{S_{0}}^{2}} - \frac{\mu_{S_{a}}}{\sigma_{S_{a}}^{2}}\right) \le c''$$
(25)

where c'' is a constant. If $\left(\frac{1}{\sigma_{S_a}^2} - \frac{1}{\sigma_{S_0}^2}\right) > 0$ then

$$\left(s + \left(\frac{\mu_{S_0}\sigma_{S_a}^2 - \mu_{S_a}\sigma_{S_0}^2}{\sigma_{S_0}^2 - \sigma_{S_a}^2}\right)\right)^2 \le c,$$

and c is a constant.

and c is a constant.
Let
$$b_S = \left(\frac{\mu S_0 \sigma_{S_a}^2 - \mu S_a \sigma_{S_0}^2}{\sigma_{S_0}^2 - \sigma_{S_a}^2}\right)$$
, then

$$\mathbb{P}\Big(\big(s+b_S\big)^2 \le c|H_0\Big) \le \alpha$$

where under the null hypothesis

$$(S+b_S)^2|H_0 \sim \chi_1^2(\frac{\mu_{S_0}+b_S}{\sigma_{S_0}})$$

And If $\left(\frac{1}{\sigma_{S_a}^2} - \frac{1}{\sigma_{S_0}^2}\right) < 0$, in (25), then the rejection region will be

$$\left(s + \left(\frac{\mu_{S_0}\sigma_{S_a}^2 - \mu_{S_a}\sigma_{S_0}^2}{\sigma_{S_0}^2 - \sigma_{S_a}^2}\right)\right)^2 > c,$$

which implies

$$\mathbb{P}\Big(\big(s+b_S\big)^2 > c|H_0\Big) \le \alpha.$$

• Repeating the same logic in the previous part applied on the statistic K. Let $f_0(k)$ and $f_a(k)$ be the density functions of the statistic K under the null and alternative hypotheses, respectively. Applying the Neyman-Pearson Lemma to K, we reject H_0 if

$$\frac{f_0(k)}{f_a(k)} \le c',$$

where c' is a constant determined by α (the probability of type I error).

Recall that, $K \sim AN(\mu_K, \sigma_K^2)$, which under the null hypothesis assumption (the innovations follow Gumbel distribution) implies $K \sim AN(\mu_{K_0}, \sigma_{K_0}^2)$, and $K \sim AN(\mu_{K_a}, \sigma_{K_a}^2)$ under the alternative (the innovations follow Gaussian distribution). This implies

$$\left(\frac{1}{\sqrt{2\pi\sigma_{K_0}^2}}e^{-(k-\mu_{K_0})^2/2\sigma_{K_0}^2} \middle/ \frac{1}{\sqrt{2\pi\sigma_{K_a}^2}}e^{-(k-\mu_{K_a})^2/2\sigma_{K_a}^2}\right) \le c',$$

hence,

$$k^{2} \left(\frac{1}{\sigma_{K_{a}}^{2}} - \frac{1}{\sigma_{K_{0}}^{2}}\right) + 2k \left(\frac{\mu_{K_{0}}}{\sigma_{K_{0}}^{2}} - \frac{\mu_{K_{a}}}{\sigma_{K_{a}}^{2}}\right) \le c''$$
(26)

where c'' is a constant. If $\left(\frac{1}{\sigma_{K_a}^2} - \frac{1}{\sigma_{K_0}^2}\right) > 0$ then

$$\left(k + \left(\frac{\mu_{S_0}\sigma_{K_a}^2 - \mu_{K_a}\sigma_{K_0}^2}{\sigma_{K_0}^2 - \sigma_{K_a}^2}\right)\right)^2 \le c,$$

and c is a constant.

Let
$$b_K = \left(\frac{\mu_{K_0}\sigma_{K_a}^2 - \mu_{K_a}\sigma_{K_0}^2}{\sigma_{K_0}^2 - \sigma_{K_a}^2}\right)$$
, then

$$\mathbb{P}\left(\left(k + b_K\right)^2 \le c|H_0\right) \le \alpha$$

where under the null hypothesis

$$(K+b_K)^2|H_0 \sim \chi_1^2(\frac{\mu_{K_0}+b_K}{\sigma_{K_0}}).$$

And If $\left(\frac{1}{\sigma_{K_a}^2} - \frac{1}{\sigma_{K_0}^2}\right) < 0$, in (26), then the rejection region will be

$$\left(k + \left(\frac{\mu_{K_0}\sigma_{K_a}^2 - \mu_{K_a}\sigma_{K_0}^2}{\sigma_{K_0}^2 - \sigma_{K_a}^2}\right)\right)^2 > c,$$

which implies

$$\mathbb{P}\Big(\big(k+b_K\big)^2 > c|H_0\Big) \le \alpha.$$

Here are some remarks regarding the non-central chisquare S and the non-central chisquare K tests:

Remark 0.0.34. The non-central chi-square S test can be used to reject the Gumbel innovations in favor of the Gaussian innovations, and it requires less data.

Remark 0.0.35. The non-central chi-square K test can be used to reject the Gaussian innovations, in favor of the Gumbel innovations, but it needs larger sample size than the non-central chisquare S test.

Remark 0.0.36. For the mixture of the Gaussian and Gumbel innovations, the non-central chisqure test S can be used to reject the Gumbel or the Gaussian innovations in favor of the mixture of both. Note that the S test can not be used to reject the mixture of both in favor of the Gaussian or in favor of the Gumbel, since any Gaussian AR(1) model or AR(1) model with Gumbel innovations can be written as a mixture with appropriate wights. While the non-central chisquare K test can not be used with the mixture.

Remark 0.0.37. In practice, starting with non-central chisquare S would be advised, that is

checking the normality first, and then checking the mixture assumption. That is because if there is a moderate portion of the observations are Gumbel (like the mixture), then applying the non-central chi-square S or K tests may not reject the Gumbel assumption in favor for the Gaussian since it is heavier tail than both of them.

Remark 0.0.38. The empirical power for these tests are computed, for more details see the simulation studies in chapter 4.

Remark 0.0.39. A very useful well known test is, the usual z-test, for example, by finding the confidence interval and use it to conclude the rejection of any of the null hypotheses the innovations are Gaussian, Gumbel or a mixture of both.

The Simulation

This chapter contains simulations for the main results in chapter 3. We start by simulations for the asymptotic distribution of the S_T and K_T . And then the non-central chi-square Sand K tests were illustrated by simulations for all the different null hypothesis scenarios. In addition, their empirical power was computed using simulations.

The distributions of S_T and K_T

In this section, we compare the asymptotic distributions of S_T and K_t using our calculations and using "the empirical moments".

In table 2 we calculate the asymptotic distribution of S_T where we simulate from a Gaussian AR(1) model. T is the length of the realization, "the sample size", and ϕ is the AR(1) parameter. In the third column, we simulate an AR(1) model with the same length as T and ϕ number of iterations (say 1000 times for example). each time we calculate the empirical moments, and then take their average over the number of iterations, and then we get the variances and covariances of the empirical second and third moments, and finally apply the delta method to the result which appear in column 3. In the last column, we first simulate one realization of the Gaussian AR(1), with the same T and ϕ and then obtain the asymptotic distributions of \hat{m}_2 and \hat{m}_3 using the results as in theorem0.0.29 and then applying the delta method.

| Т | ϕ | Distribution of S_T using empirical moments | Distribution of S_T using the moments of X_t |
|------|--------|---|--|
| 100 | 0.35 | $N(0, 0.50^2)$ | $N(0, 0.51^2)$ |
| 200 | 0.59 | $N(0, 0.47^2)$ | $N(0, 0.47^2)$ |
| 500 | 0.599 | $N(0, 0.30^2)$ | $N(0, 0.30^2)$ |
| 800 | 0.864 | $N(0, 0.44^2)$ | $N(0, 0.43^2)$ |
| 1600 | 0.850 | $N(0, 0.29^2)$ | $N(0, 0.29^2)$ |

Table 2: The distribution of S_T using empirical moments and the moments of X_t when simulating from Gaussian AR(1) model

As we can see from the table the distributions are almost the same. When the innovations are Gumbel, applying the same techniques above yield to the results in table 3. The results are satisfying in both cases.

| n | ϕ | Distribution of S_T using empirical moments | Distribution of S_T using the moments of X_t |
|------|--------|---|--|
| 100 | 0.31 | $N(0.966, 0.69^2)$ | $N(1.00, 0.69^2)$ |
| 200 | 0.662 | $N(0.667, 0.59^2)$ | $N(0.674, 0.60^2)$ |
| 800 | 0.856 | $N(0.419, 0.43^2)$ | $N(0.4216, 0.44^2)$ |
| 1000 | 0.863 | $N(0.408, 0.40^2)$ | $N(0.410, 0.40^2)$ |
| 1600 | 0.857 | $N(0.420, 0.31^2)$ | $N(0.420, 0.31^2)$ |

Table 3: The distribution of S_T using empirical moments and using the moments of X_t when simulating from AR(1) model with Gumbel innovations

The same method in comparing the asymptotic distribution of S_T using the results in chapter 3 with the empirical moments, are also applied to K_T . In this case the empirical second and fourth moments are used, and theorem 0.0.31. Table 4 shows the distribution of K_T where the innovations are Gaussian, while for the Gumbel innovations the results are illustrated in table 5. All of these distributions are in agreement with each other.

| Т | ϕ | Distribution of K_T using empirical moments | Distribution of K_T using the moments of X_t |
|-----|--------|--|--|
| 100 | 0.149 | $N(3.00, 0.48^2)$ | $N(2.99, 0.49^2)$ |
| 200 | 0.47 | $N(3.00, 0.363^2)$ | $N(2.99, 0.365^2)$ |
| 500 | 0.599 | $N(2.99, 0.25^2)$ | $N(2.99, 0.25^2)$ |
| 800 | 0.864 | $N(2.99, 0.272^2)$ | $N(2.99, 0.272^2)$ |
| 300 | 0.786 | $N(3.00, 0.412^2)$ | $N(2.99, 0.422^2)$ |
| 100 | 0.54 | $N(3.00, 0.532^2)$ | $N(2.99, 0.536^2)$ |

Table 4: The distribution of K_T using empirical moments and the moments of X_t when simulating from Gaussian AR(1) model

| Т | ϕ | Distribution of K_T using empirical moments | Distribution of K_T using the moments of X_t |
|-----|--------|--|--|
| 100 | 0.263 | $N(5.06, 3.68^2)$ | $N(5.10, 3.74^2)$ |
| 200 | 0.47 | $N(4.53, 2.12^2)$ | $N(4.47, 2.10^2)$ |
| 500 | 0.597 | $N(4.13, 1.12^2)$ | $N(4.14, 1.17^2)$ |
| 800 | 0.864 | $N(3.367, 0.679^2)$ | $N(3.36, 0.735^2)$ |
| 300 | 0.811 | $N(3.494, 1.149^2)$ | $N(3.496, 1.22^2)$ |
| 100 | 0.522 | $N(4.33, 2.61^2)$ | $N(4.37, 2.8^2)$ |

Table 5: The distribution of K_T using empirical moments and the moments of X_t when simulating from AR(1) model with Gumbel innovations

The non-central Chi-square test S and its empirical power

In this section we show the results of applying the non-central chi-square test S to reject the Gaussian innovations in favor of the Gumbel and the mixture of both innovations of AR(1) model, and to reject the Gumbel in favor of the Gaussian and the mixture of both mixture innovations of the Ar(1) model. Also we calculate the empirical power in all hypotheses.

We start with the hypothesis H_0 : $\xi_t \sim$ Gaussian vs H_a : $\xi_t \sim$ Gumbel. We simulate AR(1) with Gumbel innovations where T and ϕ as in the first and second column, and then we fit an AR(1) model with Gaussian innovations and obtain the estimators which are used to get the distribution of S_T as in theorem 0.0.30. And then on the same realization, we fit a Gaussian AR(1) model and get the estimators and get the distribution of S_T under the Gaussian assumption. Now the non-central chi square test S is applied as in theorem 0.0.33. As we can see the test works perfectly, we reject the null hypothesis in all of the simulations. Then when we simulate from Gaussian AR(1) with the same T and ϕ , and apply the same procedure above we fail to reject the null hypothesis, that is the innovations are Gaussian.

| | | simulation of $AR(1)$ | simulation of $AR(1)$ |
|------|--------|----------------------------|----------------------------|
| | | with Gumbel innova- | with Gaussian innova- |
| | | tions | tions |
| | | $H_0: \xi_t \sim Gaussian$ | $H_0: \xi_t \sim Gaussian$ |
| | | $H_a: \xi_t \sim Gumbel$ | $H_a: \xi_t \sim Gumbel$ |
| Т | ϕ | The result | The result |
| 250 | 0.168 | Reject H_0 | Fail to reject H_0 |
| 250 | 0.256 | Reject H_0 | Fail to reject H_0 |
| 300 | 0.351 | Reject H_0 | Fail to reject H_0 |
| 300 | 0.455 | Reject H_0 | Fail to reject H_0 |
| 400 | 0.519 | Reject H_0 | Fail to reject H_0 |
| 500 | 0.58 | Reject H_0 | Fail to reject H_0 |
| 700 | 0.659 | Reject H_0 | Fail to reject H_0 |
| 900 | 0.720 | Reject H_0 | Fail to reject H_0 |
| 900 | 0.783 | Reject H_0 | Fail to reject H_0 |
| 1000 | 0.801 | Reject H_0 | Fail to reject H_0 |

Table 6: Applying the non-central chi-square S test to reject the Gaussian innovations in favor of the Gumbel.

Table 7 shows the empirical power of the non-central chi-square S test of the previous hypothesis, that is the hypothesis $H_0: \xi_t \sim \text{Gaussian vs } H_a: \xi_t \sim \text{Gumbel}$. As we can see the power change T and ϕ . The bigger the $|\phi|$ the larger T is needed to get a good power.

| T | ϕ | The empirical power |
|------|--------|---------------------|
| 200 | 0.17 | 899/1000 |
| 200 | 0.30 | 790/1000 |
| 500 | 0,30 | 982/1000 |
| 200 | 0.50 | 557/1000 |
| 500 | 0.50 | 894/1000 |
| 300 | 0.60 | 550/1000 |
| 500 | 0.60 | 743/1000 |
| 800 | 0.60 | 887/1000 |
| 1000 | 0.60 | 925/1000 |
| 1200 | 0.70 | 846/1000 |

Table 7: The empirical power of the non-central chi-square test S where $H_0: \xi_t \sim \text{Gaussian}$ vs $H_a: \xi_t \sim \text{Gumbel}$

The second scenario the hypothesis H_0 : $\xi_t \sim$ Gumbel vs H_a : $\xi_t \sim$ Gaussian. We simulated from AR(1) with Gumbel innovations and fit first Gumbel innovations and get the parameters using EM algorithm and get the distribution of S_T . WE repeated the same techniques after fitting a Gaussian AR(1) model to the sample path. We applied the noncentral chi-square S test to the data.

| | | simulation of $AR(1)$ | simulation of $AR(1)$ |
|------|---------|----------------------------|----------------------------|
| | | with Gumbel innova- | with Gumbel innova- |
| | | tions | tions |
| | | $H_0: \xi_t \sim Gaussian$ | $H_0: \xi_t \sim Gumbel$ |
| | | $H_a: \xi_t \sim Gumbel$ | $H_a: \xi_t \sim Gaussian$ |
| Т | ϕ | The result | The result |
| 500 | 0.17 | Reject H_0 | Fail to reject H_0 |
| 500 | 0.0.291 | Reject H_0 | Fail to reject H_0 |
| 700 | 0.529 | Reject H_0 | Fail to reject H_0 |
| 1200 | 0.728 | Reject H_0 | Fail to reject H_0 |
| 1800 | 0.798 | Reject H_0 | Fail to reject H_0 |

And we get the results in table8, with the empirical power in table 9.

Table 8: Applying the non-central chi-square S test to the hypothesis $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ Gaussian

| T | ϕ | The empirical power |
|------|--------|---------------------|
| 500 | 0.17 | 999/1000 |
| 500 | 0.30 | 992/1000 |
| 700 | 0.50 | 952/1000 |
| 700 | 0.60 | 840/1000 |
| 1200 | 0.70 | 789/1000 |
| 2500 | 0.8 | 752/1000 |
| 3000 | 0.8 | 842/1000 |

And this is the empirical power. The last case where we apply the non-central chi square

Table 9: The empirical power of the non-central chi-square test S where $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ Gaussian

test S is to reject the Gumbel or Gaussian innovations in favor of the mixture. That is $H_0: \xi_t \sim$ Gaussian or $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ mixture. The same procedure above was applied here. We first simulate from AR(1) with mixture innovations, and then fit an AR(1) model with mixture, then find the estimators which enable us to calculate the distribution of S_T as in theorem 0.0.30. On the same realization we got from the simulation with mixture innovations, we fit an AR(1) with Gumbel innovations and then find the estimators which we use to find the distribution of S_T as in theorem 0.0.30 under the Gumbel assumption. Doing the same procedure on the same realization but with Gaussian innovations.

| | | | | | $H_0: \xi_t \sim Gaussian$ | $H_0: \xi_t \sim Gumbel$ |
|------|----------------------|--------------------|----------|-------|----------------------------|---------------------------|
| | | | | | $H_a: \xi_t \sim Mixture$ | $H_a: \xi_t \sim Mixture$ |
| T | $\widehat{\phi}_{g}$ | $\widehat{\phi}_n$ | ϕ_m | p | The result | The result |
| 700 | 0.109 | 0.075 | 0.096 | 0.711 | Reject H_0 | Reject H_0 |
| 700 | 0.150 | 0.137 | 0.133 | 0.657 | Fail to reject H_0 | Reject H_0 |
| 700 | 0.138 | 0.142 | 0.134 | 0.501 | Reject H_0 | Reject H_0 |
| 700 | 0.186 | 0.207 | 0.192 | 0.685 | Fail to reject H_0 | Reject H_0 |
| 700 | 0.209 | 0.193 | 0.197 | 0.685 | Reject H_0 | Reject H_0 |
| 700 | 0.138 | 0.159 | 0.174 | 0.557 | Reject H_0 | Reject H_0 |
| 700 | 0.235 | 0.229 | 0.229 | 0.260 | Reject H_0 | Reject H_0 |
| 2000 | 0.114 | 0.137 | 0.131 | 0.324 | Reject H_0 | Reject H_0 |
| 2000 | 0.185 | 0.175 | 0.176 | 0.255 | Reject H_0 | Reject H_0 |
| 2000 | 0.158 | 0.173 | 0.173 | 0.843 | Fail to reject H_0 | Reject H_0 |
| 2000 | 0.186 | 0.145 | 0.161 | 0.815 | Reject H_0 | Fail to reject H_0 |
| 2000 | 0.271 | 0.268 | 0.267 | 0.364 | Reject H_0 | Reject H_0 |
| | 0.209 | 0.197 | 0.205 | 0.546 | Reject H_0 | Reject H_0 |
| 2000 | | | | | | |
| 2000 | 0.276 | 0.266 | 0.273 | 0.754 | Fail to reject H_0 | Reject H_0 |
| | | | | | | |

Table 10: Applying the non-central Chisquare test S when simulating from an AR(1) with mixture innovation to reject the Gumbel and the Gaussian innovations in favor of the mixture.

After obtaining the distribution of S_T under the three assumptions where we simulate from a mixture we applied the non-central chi square test S as in theorem 0.0.33 and the results are in the following tables 10 and 11. The table is continued here:

| | | | | | $H_0: \xi_t \sim Gaussian$ | $H_0: \xi_t \sim Gumbel$ |
|------|--------------------|--------------------|----------|-------|----------------------------|---------------------------|
| | | | | | $H_a: \xi_t \sim Mixture$ | $H_a: \xi_t \sim Mixture$ |
| Т | $\widehat{\phi}_g$ | $\widehat{\phi}_n$ | ϕ_m | p | The result | The result |
| 3000 | 0.380 | 0.393 | 0.383 | 0.299 | Reject H_0 | Fail to reject H_0 |
| 3000 | 0.347 | 0.340 | 0.342 | 0.424 | Reject H_0 | Fail to reject H_0 |
| 3000 | 0.383 | 0.379 | 0.382 | 0.411 | Reject H_0 | Reject H_0 |
| 3000 | 0.308 | 0.321 | 0.327 | 0.592 | Reject H_0 | Reject H_0 |
| 3000 | 0.333 | 0.338 | 0.338 | 0.703 | Reject H_0 | Reject H_0 |
| 4000 | 0.445 | 0.439 | 0.442 | 0.254 | Reject H_0 | Reject H_0 |
| 4000 | 0.442 | 0.444 | 0.448 | 0.544 | Reject H_0 | Reject H_0 |
| 4000 | 0.540 | 0.542 | 0.539 | 0.508 | Reject H_0 | Reject H_0 |
| 4000 | 0.531 | 0.533 | 0.538 | 0.745 | Fail to reject H_0 | Reject H_0 |
| 9000 | 0.652 | 0.652 | 0.651 | 0.353 | Reject H_0 | Fail to reject H_0 |
| 9000 | 0.656 | 0.654 | 0.655 | 0.512 | Reject H_0 | Reject H_0 |
| 9000 | 0.740 | 0.737 | 0.739 | 0.690 | Fail to reject H_0 | Reject H_0 |
| 9000 | 0.663 | 0.662 | 0.665 | 0.769 | Reject H_0 | Reject H_0 |

Table 11: Applying the non-central Chisquare test S when simulating from an AR(1) with mixture innovation to reject the Gumbel and the Gaussian innovations in favor of the mixture innovations continuing.

As we see from table 10, the performance of the test depends on many factors, for example the ϕ , T and the p. Most of the cases where the test fails are when p is large or small and also related to the ϕ . But in general the test preforms better when we increase the sample size. As we see from the empirical power in table 12.

| T | ϕ | p | The empirical power |
|------|--------|------|---------------------|
| 700 | 0.17 | 0.53 | 61/100 |
| 1400 | 0.17 | 0.53 | 76/100 |
| 1800 | 0.17 | 0.53 | 80/100 |
| 2200 | 0.17 | 0.53 | 86/100 |
| 3000 | 0.17 | 0.53 | 91/100 |
| 1200 | 0.30 | 0.53 | 65/100 |
| 2500 | 0.30 | 0.53 | 80/100 |
| 3500 | 0.30 | 0.53 | 88/100 |
| 3500 | 0.50 | 0.53 | 71/100 |
| 4500 | 0.50 | 0.53 | 85/100 |
| 6000 | 0.70 | 0.53 | 63/100 |

Table 12: The empirical power of the non-central chi-square test S where $H_0: \xi_t \sim$ Gumbel or Gaussian vs $H_a: \xi_t \sim$ mixture

The non-central chi-square test K and its empirical power

In this section we show the results of applying the non-central chi-square K test. There are two cases where this test can be applied, to reject the Gumbel innovations in favor of Gaussian and the other way, to reject the Gaussian innovations in favor of the Gumbel.

We start by the hypothesis H_0 ; $\xi_t \sim$ Gumbel vs H_0 ; $\xi_t \sim$ Gaussian, where we simulate from a Gaussian AR(1), and then after obtaining the sample we fit a Gaussian AR(1) and use the estimators to get the distribution of K_t by applying theorem 0.0.32. Then next, we fit an AR(1) with Gumbel innovations on the same sample and use the estimators to get the distribution of K_T as in theorem 0.0.32. After having these distributions we apply theorem 0.0.33 to obtain the results of K test.

| | | simulation of $AR(1)$ with Gaussian innova- | simulation of $AR(1)$ with Gumbel innova- |
|------|--------|--|--|
| | | tions | tions |
| | | $H_0: \xi_t \sim Gumbel$ | $H_0: \xi_t \sim Gumbel$ |
| | | $H_a: \xi_t \sim Gaussian$ | $H_a: \xi_t \sim Gaussian$ |
| Т | ϕ | The result | The result |
| 700 | 0.171 | Reject H_0 | Fail to reject H_0 |
| 850 | 0.266 | Reject H_0 | Fail to reject H_0 |
| 900 | 0.33 | Reject H_0 | Fail to reject H_0 |
| 1000 | 0.483 | Reject H_0 | Fail to reject H_0 |
| 1500 | 0.56 | Reject H_0 | Fail to reject H_0 |
| 2000 | 0.508 | Reject H_0 | Fail to reject H_0 |
| 2500 | 0.60 | Reject H_0 | Fail to reject H_0 |
| 2500 | 0.762 | Reject H_0 | Fail to reject H_0 |
| 3000 | 0.77 | Reject H_0 | Fail to reject H_0 |
| 3000 | 0.84 | Reject H_0 | Fail to reject H_0 |

Table 13: Applying the non-central chi-square K test to reject the Gumbel innovations in favor of the Gaussian.

Lastly, we simulate from AR(1) with Gumbel innovations and do the same procedure above, and apply the same test under the same assumptions. As we can see in table 13 the test preforms well, where we reject the null hypothesis when we generate from Gaussian and we fail to reject the null hypothesis when the innovation is generated from the Gumbel. The following table 14 shows the empirical power of the test.

| T | ϕ | The empirical power |
|------|----------|---------------------|
| 500 | 0.17 | 770/1000 |
| 700 | 0.17 | 908/1000 |
| 700 | $0,\!30$ | 853/1000 |
| 1000 | 0.30 | 973/1000 |
| 1000 | 0.50 | 812/1000 |
| 1300 | 0.50 | 926/1000 |
| 1300 | 0.60 | 769/1000 |
| 1800 | 0.60 | 911/1000 |
| 1800 | 0.70 | 627/1000 |
| 2500 | 0.70 | 817/1000 |
| 3000 | 0.70 | 897/1000 |

Table 14: The empirical power of the non-central chi-square test K where $H_0: \xi_t \sim$ Gumbel vs $H_a: \xi_t \sim$ Gaussian

| | | simulation of $AR(1)$ with Gumbel innova- | simulation of $AR(1)$ with Gaussian innova- |
|------|--------|--|--|
| | | tions | tions |
| | | $H_0: \xi_t \sim Gaussian$ | $H_0: \xi_t \sim Gaussian$ |
| | | $H_a: \xi_t \sim Gumbel$ | $H_a: \xi_t \sim Gumbel$ |
| Т | ϕ | The result | The result |
| 700 | 0.171 | Reject H_0 | Fail to reject H_0 |
| 850 | 0.266 | Reject H_0 | Reject H_0 |
| 900 | 0.33 | Reject H_0 | Fail to reject H_0 |
| 1000 | 0.483 | Reject H_0 | Fail to reject H_0 |
| 1500 | 0.56 | Reject H_0 | Fail to reject H_0 |
| 2000 | 0.508 | Reject H_0 | Fail to reject H_0 |
| 2500 | 0.60 | Reject H_0 | Fail to reject H_0 |
| 2500 | 0.762 | Reject H_0 | Fail to reject H_0 |
| 3000 | 0.77 | Reject H_0 | Fail to reject H_0 |
| 3000 | 0.84 | Reject H_0 | Fail to reject H_0 |

Table 15: Applying the non-central chi-square K test to reject the Gaussian innovations in favor of the Gumbel.

Lastly, we applied the non-central chi-square test K under the hypothesis $H_0 : \xi \sim$ Gaussian vs $H_a : \xi_t \sim$ Gumbel to the same simulations above and we obtained the results in table15. And as we see the results of applying this test are satisfactory. And the empirical power of the of test K under the previous null hypothesis is shown in table 16.

| T | ϕ | The empirical power |
|------|----------|---------------------|
| 300 | 0.17 | 870/1000 |
| 350 | 0.17 | 917/1000 |
| 300 | $0,\!30$ | 836/1000 |
| 400 | 0.30 | 904/1000 |
| 400 | 0.50 | 757/1000 |
| 600 | 0.50 | 898/1000 |
| 700 | 0.5 | 920/1000 |
| 700 | 0.60 | 800/1000 |
| 900 | 0.60 | 885/1000 |
| 1000 | 0.60 | 908/1000 |
| 1000 | 0.70 | 725/1000 |
| 1400 | 0.70 | 851/1000 |
| 1900 | 0.70 | 930/1000 |

Table 16: The empirical power of the non-central chi-square test K where $H_0:\xi_t\sim \text{Gaussian}$ vs $H_a:\xi_t\sim \text{Gumbel}$

The analysis of the sea level rise data

It has been predicted that the global sea-level rise (SLR) will reach 0.5 to 2.00m by the end of this century (by 2100). The relative SLR (RSLR) can be greater than the global change because of local effects [57]. As mentioned in chapter one, several papers provide evidence of SLR using tide gauge data "historical observations", along the coastal mid-Atlantic region (a so-called "hotspot") [31, 10, 74].

This chapter analyzes sea level measurements obtained from the Actuarial climate index (ACI). These measurements are made available via tide gauge, which "measure sea level relative to the land below, but because the land may be moving, the ACI sea level component measures the combined effect on coastal shorelines of the generally rising seas and the rising or falling land," According to ACI.

The ACI sea level measurements are obtained from tide gague stations, which are located along 10 of the 12 regions as shown in Figure 3. These are for all regions except the CAR "Central Arctic" and this MID "Midwest." This study is concerned with sea level measurements of the CEA "Central East Atlantic" region, which includes the following states and districts: CT, DC, DE, MA, MD, ME, NH, NJ, NY, PA, RI, VT and WV. The population, infrastructure and economic activities in this region are dense. All this has led many authors to study this area as mentioned before, and our analysis shows new evidence of a rise in "hotspots" at tide stations in the CEA. There is more about the data in section one, and section two includes the methods, followed by the results and discussion in section three and four respectively.



Figure 3: Map shows the 12 different regions according to ACI

Data

Monthly sea level standardized anomaly records from tide gauge stations in the Central East Atlantic "CEA" were obtained from the ACI, recorded over a length of 56 years (1961-2017). The ACI acquired the raw data from the Permanent Service for Mean Sea Level. Let y_i be the monthly sea level measurements from 1961-2017, and let 1961-1990 be the reference period, with average $\mu_{ref}(y_i)$ and standard deviation $\sigma_{ref}(y_i)$ these measurements then convert to standardized anomaly according to the following equation:

$$y_{i_{std}} = \Delta y_i / \sigma_{ref}(y_i)$$

where $\Delta y_i = y_i - \mu_{ref}(y_i)$.

Method

The data (standardized anomaly records) were decomposed after converting them to time series. The path of the series $\{X_t\}$ (the random part after decomposition) suggests that an autoregressive model would be a good choice. Three centered AR(1) models with different innovations distributions were fitted to the data, Gaussian, Gumbel and a mixture of both. Recall AR(1) model is written in the form:

 $X_t = \phi X_{t-1} + \xi_t$ where $\xi_t \sim WN(0, \sigma)$.

For the centered Gaussian AR(1), $\xi \sim N(0.\sigma)$, $\sigma > 0$, the parameters ϕ and σ are estimated using the Expectation Maximization, where ξ_t is considered to be latent variable, and the log-likelihood of the Gaussian random variable ξ_t is:

$$\log L(0, \sigma^2 | \xi_t) = \frac{-T}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^T \xi_t^2.$$

And for the centered AR(1) model with Gumbel innovations the $\xi \sim G(\alpha, \beta)$ where $\alpha = -\gamma\beta$, $\beta > 0$ and γ is the Euler constant. As in the Gaussian case, ϕ and β are estimated using the Expectation Maximization where ξ_t is considered a latent variable, and the log-likelihood of the Gumbel random variable ξ_t is:

$$\log L(\alpha, \beta | \xi_t) = -T \log(\alpha) - \sum_{t=1}^T \frac{\xi_t - \beta}{\alpha} - \sum_{t=1}^T e^{-\left(\frac{\xi_t - \beta}{\alpha}\right)}.$$

And for the centered AR(1) model with mixture innovations, the innovations are the sum of the previous weighted innovations, $(\xi_t \sim p \cdot N(0, \sigma^2) + (1-p) \cdot G(-\gamma\beta, \beta))$, where $0 \le p \le 1$. The Expectation Maximization algorithm is used to estimate the parameters ϕ , σ^2 , p, and β , and the log likelihood function for the mixture innovations is:

$$\log L(\sigma, \beta, p, \xi_t) = \sum_t \log[(1-p) \cdot f(\xi_t) + p \cdot g(\xi_t)],$$

such that $f(\xi_t)$ and $g(\xi_t)$ are the densities for the above-centered Gaussian and Gumbel r.vs. respectively.

After obtaining all estimators, the theoretical AR(1) moments, $E[X_t^r]$ where r = 2, 3, 4, 5, 6, 8were calculated according to formulas Lemma 0.0.5, Lemma 0.0.6, Lemma 0.0.7, Lemma 0.0.8, Lemma 0.0.9 and Lemma 0.0.10, where under the Gaussian assumption, the estimated ϕ from fitting AR(1) with Gaussian innovations used in addition to moments of Gaussian random variable. The same applied under the Gumbel assumption, and for the mixture assumption, in addition to the estimated ϕ , \hat{p} , is used to calculate the weighted moments, and the sum of these weighted moments is the mixture moments.

Using these moments with $\hat{\phi}$, the means, variances and covariances of the empirical moments $\hat{m}_r = \sum_{t=1}^T X_t^r / T$, where r = 2, 3, 4 were obtained according to Theorem 0.0.18, Theorem 0.0.11, Theorem0.0.12, Theorem0.0.14, Theorem0.0.13 and Theorem0.0.15. Using theoremTheorem 0.0.29 and Theorem 0.0.31 imply the asymptotic multivariate normality of the vectors (\hat{m}_2, \hat{m}_3) and (\hat{m}_2, \hat{m}_4) of these statistics, and then the delta method in Theorem 0.0.30 and Theorem0.0.32 are applied to calculate the means and variances of the empirical skewness, $S_T = \hat{m}_3 / \hat{m}_2^{3/2}$ and empirical kurtosis $K_T = \hat{m}_4 / \hat{m}_2^2$, under the three different assumptions on the innovations.

Having these distributions enables us to apply the different tests (as in chapter three) to determine the best model. The confidence intervals were obtained, and then the z-test was applied to the means of the two statistics S_T and K_T under the three assumptions on ξ_t . To apply the non-central chi square S test as in Theorem 0.0.33, the value of the test statistic as in (22) was calculated as well as the coefficients (20) and (21) to determine the rejection regions. In addition, the p value of the non-central chisquare test S was calculated under different hypotheses (e.g. P(value of test statistic)).

Results

In this section, the results of applying the above method to the sea level rise are summarized. Table 18 shows the estimators of ϕ , σ , β and p when fitting AR(1) model with Gaussian, Gumbel and mixture innovations.

| ξ 's distribution | $\widehat{\phi}$ | $\widehat{\sigma}$ | $\widehat{\beta}$ | \widehat{p} |
|-----------------------|------------------|--------------------|-------------------|---------------|
| Gaussian | 0.172 | 0.863 | - | - |
| Gumbel | 0.190 | - | 0.7638 | - |
| Mixture | 0.170 | 0.8321 | 0.7063 | 0.534 |

Table 17: The estimators of $\hat{\phi}$, $\hat{\sigma}$ and \hat{p} when fitting AR(1) model with different innovations, to the SLR data,

The proceeding estimators are used to calculate the AR(1) theoretical moments, which are summarized in Table 18. These moments are essential in calculating the asymptotic distribution of the empirical skewness S_T and kurtosis K_T .

| ξ 's distribution | m_2 | m_3 | m_4 | m_5 | m_6 | <i>m</i> ₈ |
|-----------------------|--------|--------|--------|--------|---------|-----------------------|
| Gaussian | 1.032 | 0.00 | 3.198 | 0.0025 | 16.484 | 119.321 |
| Gumbel | 1.034 | 1.139 | 5.617 | 19.146 | 95.670 | 3218.199 |
| Mixture | 1.0300 | 0.5293 | 4.2999 | 8.8560 | 52.9747 | 1548.285 |

Table 18: The theoretical moments of AR(1) model when fitted with different innovations to the SLR data.

Using these moments with $\widehat{\phi}$ and the moments of the Gaussian, Gumbel and of the

| contains the a | symptotic distribut | ion of the empirical sl | kewness (S_T) and k | urtosis (K_T) . |
|----------------|---------------------|-------------------------|-----------------------|-------------------|
| | Ca distribution | The distribution of | The distribution | of |

mixture of both, we calculate the variances under the three different assumptions. Table 19

| ξ 's distribution | The distribution of | The distribution of |
|-----------------------|---------------------|---------------------|
| | S_T | K_T |
| Gaussian | $N(0.00, 0.167^2)$ | $N(2.99, 0.19^2)$ |
| Gumbel | $N(1.08, 0.26^2)$ | $N(5.24, 1.52^2)$ |
| Mixture | $N(0.50, 0.25^2)$ | $N(4.05, 1.12^2)$ |

Table 19: The asymptotic distribution of S_T and K_T when fitting the AR(1) model with different innovations to the SLR data

Since the distributions of K_T and S_T under the three assumptions are found, the confidence intervals for the means are calculated as shown in Table 20.

| ξ 's distribution | 95% C.I. of $E(S_T)$ | 64% C.I. of $E(K_T)$ |
|-----------------------|----------------------|----------------------|
| Gaussian | (0.173, 0.831) | (3.55, 3.90) |
| Gumbel | (-0.019, 1.024) | (2.30, 5.15) |
| Mixture | (0.004, 1.00) | (2.67, 4.78) |

Table 20: The confidence intervals of $m_3/m_2^{3/2}$ and m_4/m_2^2 applied to the SLR data

To apply the hypothesis tests as in Theorem 0.0.33, the values of test statistics are obtained under different assumptions. These values are summarized in Table 21 along with $(1/\sigma_{S_a}^2 - 1/\sigma_{S_0}^2)$ values and the C's values, where α is at a level 5%.

In table 22 we calculate the p value of the non-central chisquare test S under different null hypothesis scenarios.

| The hypothesis test | The Test Statistic Value | Value of $(1/\sigma_{Sa}^2 - 1/\sigma_{S0}^2)$ | The C Value |
|--|-----------------------------|--|---------------|
| $ \begin{array}{l} H_0: \xi_t \sim \text{Gumbel} \\ H_a: \xi_t \sim \text{mixture} \end{array} $ | 445.5 | 1.438 | 468.1 |
| $\begin{array}{l} H_0: \xi_t \sim \text{Gaussian} \\ H_a: \xi_t \sim \text{mixture} \end{array}$ | 28.49 | -19.94 | 15.932 |
| $\begin{array}{l} H_0: \xi_t \sim \text{mixture} \\ H_a: \xi_t \sim \text{Gumbel} \end{array}$ | 491.059 | -1.438 | 567.47 |
| $\begin{array}{l} H_0: \xi_t \sim \text{mixture} \\ H_a: \xi_t \sim \text{Gaussian} \end{array}$ | 12.46 | 19.94 | 3.620 |

Table 21: The results of applying the non-central chisquare test T to the SLR data under different null hypothesis scenarios.

Discussion

In this section we discuss the results of fitting an AR(1) model with three different innovations (Gaussian, Gumbel and mixture of both) to the SLR data, and we show the best model based on the non-central chi-square S test and compare the results to other diagnostic checks based on the residuals.

Table 19 contains the distributions of the empirical skewness S_T and the empirical kurtosis K_T , and their confidence intervals are given in table 20. When applying the z-test to $E(S_T)$ under the three different assumptions (ξ follow Gaussian, Gumbel and a mixture of both), using 95% confidence intervals, we reject the null hypothesis that ξ_t follow the Gumbel distribution, and the null hypothesis that ξ_t follow the Gaussian. This is because the means in both cases are not included in the confidence intervals. And since the mean of
| The hypothesis test | The P value | The result |
|--|---------------|----------------------|
| $\begin{array}{l} H_0: \xi_t \sim \text{Gumbel} \\ H_a: \xi_t \sim \text{mixture} \end{array}$ | 0.01489 | Reject H_0 |
| $\begin{array}{l} H_0: \xi_t \sim \text{Gaussian} \\ H_a: \xi_t \sim \text{mixture} \end{array}$ | 0.0013 | Reject H_0 |
| $\begin{array}{l} H_0: \xi_t \sim \text{mixture} \\ H_a: \xi_t \sim \text{Gumbel} \end{array}$ | 0.5067 | Fail to reject H_0 |
| $\begin{array}{l} H_0: \xi_t \sim \text{mixture} \\ H_a: \xi_t \sim \text{Gaussian} \end{array}$ | 0.493 | Fail to reject H_0 |

Table 22: The p values of the S test applied to the SLR data with different scenarios.

 S_T under the assumption that ξ_t follows a mixture of both the Gaussian and the Gumbel is included in the 95% confidence interval, we fail to reject the null hypothesis that the innovation follows a mixture of the Gaussian and Gumbel. We get the same conclusions when applying the z-test to the mean of K_T under the previous three assumptions. Note that the confidence level for the $E(K_T)$ is 65% because the sample size is small. In order to narrow the confidence interval, either increase the sample size, which leads to a decrease in the bias, or reduce the confidence level.

Table 21 includes the results of applying the S test to the SLR data under different hypotheses. The purpose of this test is to reject the innovations' hypotheses, whether Gaussian or Gumbel, in favor of the mixture. And it fails to reject the null hypothesis that ξ_t follows mixture in favor of the Gaussian or Gumbel innovations. As we see the value of $(1/\sigma_{S_a}^2 - 1/\sigma_{S_0}^2)$ under the assumption H_0 : ξ_t follows Gumbel vs H_a : ξ_t following the mixture, which is positive. This implies that we reject the null hypothesis since the value of the test statistic falls in the rejection region, (the value of the test statistic = 445.5 < the value of C= 468.1). For the hypothesis $H_0: \xi_t$ follow Gaussian vs $H_a: \xi_t$ follow mixture, the value of $(1/\sigma_{S_a}^2 - 1/\sigma_{S_0}^2)$ is negative, this implies we reject the null hypothesis since the value of the test statistic is greater than the value of C (28.49 > of C= 15.932), that is, the value of the test statistic falls in the rejection region.

Repeating the same logic above, we fail to reject the null hypothesis under the following assumption $H_0: \xi_t$ follows mixture vs $H_a: \xi_t$ follows Gumbel or Gaussian. That is because for the Gumbel alternative the value of the test statistic is 491.05 which is less than the value of C, (the negative value of $(1/\sigma_{S_a}^2 - 1/\sigma_{S_0}^2)$ implies the rejection region has to be greater than C). And for the Gaussian alternative, the value of the test statistic is greater than the value of C, hence it falls outside the rejection region.

The *p* value for the non-central chi square *S* test under the previous null hypothesis scenarios is included in table 22. From the table we see the values of *p* under the null hypotheses $H_0: \xi_t$ follows Gumbel or $H_0: \xi_t$ follows Gaussian are less than 5% (the p values are significant), this implies we reject these null hypotheses. Whereas we fail to reject the null hypothesis that $H_0: \xi_t$ follows a mixture of both, in favor of the Gaussian or the Gumbel. That is because the *p* values are greater than 5%.

All these results are in agreement with each others, that is the AR(1) model with mixture of both Gaussian and Gumbel fits the data more appropriately than the AR(1) with the Gumbel or with the Gaussian innovations.



Q-Q plot for Sea Level Rise data as AR(1) with Gaussian Innovations

Figure 4: The QQ-plot of the residuals of fitting a Gaussian AR(1) model to the STR data vs the Gaussian distribution

These results are also in agreement with other common diagnostic check based on the residuals $\hat{\xi}_t$. For example, the QQ plot of the residuals of fitting a Gaussian AR(1) model to the SLR data is shown in figure 4. The plot suggests a heaver tail than the Gaussian. And the QQ plot of the residuals of AR(1) with Gumbel innovations is shown in figure 5, which suggests a lighter tail than the Gumbel. Whereas the QQ plot of the residuals of the mixture A(1) as shown in 6 shows a good fit.

The QQ-plots of the Gumbel and mixture residuals.



Q-Q plot for Sea Level Rise data as AR(1) with Gumbel Innovations

Figure 5: The QQ-plot of the residuals of fitting an AR(1) model with Gumbel innovations to the STR data vs the Gumbel distribution





Figure 6: The QQ-plot of the residuals of fitting an AR(1) model with mixture innovations to the STR data vs the mixture distribution

Another diagnostic check based on the residuals is the Kolmogorov–Smirnov test (KS test). Table 23 shows the p value of the KS test applied to the residuals of the SLR data under different null hypothesis scenarios. For the Gaussian and Gumbel residuals', $H_0 : \hat{\xi}_t$ follows the Gaussian or Gumbel vs $H_0 : \hat{\xi}_t$ does not follow the Gaussian or the Gumbel. As we see from the table, we reject the null hypotheses that the residuals follow Gaussian or the Gumbel.

For the mixture residuals, we simulate a random sample y_t from a mixture with the same parameters we obtained from fitting a mixture AR(1), and apply the KS test to the hypothesis $H_0: \hat{\xi}_t$ and y_t come from the same distribution vs H_a the two samples are from different distributions. The p-value of the KS statistic indicates that we fail to reject the null hypothesis.

| The hypothesis test | The p value | |
|---|---------------|--|
| $H_0:\widehat{\xi_t}\sim 	ext{Gaussian}$ | 0.010 | |
| $H_0:\widehat{\xi_t}\sim \mathrm{Gumbel}$ | 2.2e - 16 | |
| H_0 : $\hat{\xi_t}$ and y_t fol- | 0.760 | |
| low the same distribution | | |

Table 23: The p value of KS test under the null hypothesis applied to the residuals of the SLR data with different scenarios.

All the results based on our methodology and based on the residuals diagnostic checks are in agreement. The advantage of our methodology is that the non-central chisqure S test enables us to reject the Gaussian and to reject the Gumbel in favor of the mixture. The limitation of this test is that it needs large sample.

Conclusion

A summary of this work as follows: we defined non-central chi-square S and K tests to reject the hypothesis that the innovations of AR(1) model follow Gumbel in favor of the Gaussian innovations, to reject the Gaussian innovations in favor of the Gumbel and to reject the Gumbel or Gaussian innovations in favor of the mixture of both of them.

we modeled the SLR data as an AR(1) process where the innovations are Gumbel, Gaussian, or a mixture of both. Based on our non-central chi square test, we reject the Gumbel and Gaussian innovations.

This work opens further problems in both theory and applications. In the theory part, proving the asymptotic multivariate normality of the vector (\hat{m}_2, \hat{m}_3) by the m-dependency or any other tools. Also in the theory, looking at the general formula of the higher moments of the Gumbel distribution, it is a function of the "Euler Mascheroni integral" which is the derivative (and the higher derivative) of the Gamma function evaluated at 1. And note that the general formula for the moments in the Gaussian case is a function of the Hermite polynomials. The question now does the "Euler Mascheroni integral" play other role for the Gumbel distribution? For the applied part, estimating the parameters of the centered AR(1) model with mixture innovations, without assuming the means of the mixture distributions equal to zero, this might be studied by using Bayesian analysis, in particular employing RUNJAGS.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] J Andel. On ar (1) processes with exponential white noise. Communications in Statistics-Theory and Methods, 17(5):1481–1495, 1988.
- [2] OD Anderson. The box-jenkins approach to time series analysis. RAIRO-Operations Research, 11(1):3–29, 1977.
- [3] Ronald L Anderson. Distribution of the serial correlation coefficient. The Annals of Mathematical Statistics, 13(1):1–13, 1942.
- [4] TW Anderson and AM Walker. On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process. *The Annals of Mathematical Statistics*, 35(3):1296–1303, 1964.
- [5] ANNMARIE GEDDES BARIBEAU. The slr factor: As sea levels rise, the flood risk equation changes. *Actuarial Review*, 2018.
- [6] MS Bartlett. Some aspects of the time-correlation problem in regard to tests of significance. Journal of the Royal Statistical Society, 98(3):536–543, 1935.
- [7] CB Bell and EP Smith. Infrence for non-negative autoregressive schemes. Communications in Statistics-Theory and Methods, 15(8):2267–2293, 1986.
- [8] Ocean Studies Board, National Research Council, et al. Sea-level rise for the coasts of California, Oregon, and Washington: past, present, and future. National Academies Press, 2012.
- [9] David Bolin, Peter Guttorp, Alex Januzzi, Daniel Jones, Marie Novak, Harry Podschwit, Lee Richardson, Aila Särkkä, Colin Sowder, and Aaron Zimmerman. Statistical prediction of global sea level from global temperature. *Statistica Sinica*, pages 351–367, 2015.
- [10] John D Boon. Evidence of sea level acceleration at us and canadian tide stations, atlantic coast, north america. *Journal of Coastal Research*, 28(6):1437–1445, 2012.
- [11] John D Boon, John M Brubaker, and David R Forrest. Chesapeake bay land subsidence and sea level change: An evaluation of past and present trends and future outlook. 2010.
- [12] George Edward Pelham Box. Time series analysis forecasting and control. 1970.

- [13] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time* series analysis: forecasting and control. John Wiley & Sons, 2015.
- [14] George EP Box and David A Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332):1509–1526, 1970.
- [15] Peter J Brockwell, Richard A Davis, and Stephen E Fienberg. Time series: theory and methods: theory and methods. Springer Science & Business Media, 1991.
- [16] J Peter Byrne. The cathedral engulfed: Sea-level rise, property rights, and time. La. L. Rev., 73:69, 2012.
- [17] Niamh Cahill, Andrew C Kemp, Benjamin P Horton, Andrew C Parnell, et al. Modeling sea-level change using errors-in-variables integrated gaussian processes. *The Annals of Applied Statistics*, 9(2):547–571, 2015.
- [18] Richard Chandler and Marian Scott. Statistical methods for trend detection and analysis in the environmental sciences. John Wiley & Sons, 2011.
- [19] Junesang Choi and HM Srivastava. Evaluation of higher-order derivatives of the gamma function. Publikacije Elektrotehničkog fakulteta. Serija Matematika, pages 9–18, 2000.
- [20] JA Church, PU Clark, A Cazenave, JM Gregory, S Jevrejeva, A Levermann, et al. Sea level change in: Stocker tf, qin d, plattner gk, tignor m, allen sk, boschung j, et al., editors. climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change. cambridge, united kingdom and new york, ny, 2013.
- [21] John A Church and Neil J White. A 20th century acceleration in global sea-level rise. Geophysical research letters, 33(1), 2006.
- [22] John A Church and Neil J White. Sea-level rise from the late 19th to the early 21st century. *Surveys in geophysics*, 32(4-5):585–602, 2011.
- [23] Denis Conniffe and John E Spencer. When moments of ratios are ratios of moments. Journal of the Royal Statistical Society: Series D (The Statistician), 50(2):161–168, 2001.
- [24] Harald Cramir. Mathematical methods of statistics. Princeton U. Press, Princeton, page 500, 1946.

- [25] Nelville Davies, CM Triggs, and Paul Newbold. Significance levels of the box-pierce portmanteau statistic in finite samples. *Biometrika*, 64(3):517–522, 1977.
- [26] Neville Davies, Trevor Spedding, and William Watson. Autoregressive moving average processes with non-normal residuals. *Journal of Time Series Analysis*, 1(2):103–109, 1980.
- [27] L Dewald and P Lewis. A new laplace second-order autoregressive time-series modelnlar (2). *IEEE Transactions on Information theory*, 31(5):645–651, 1985.
- [28] Bruce C Douglas. Global sea level acceleration. Journal of Geophysical Research: Oceans, 97(C8):12699–12706, 1992.
- [29] Gilles R Ducharme and Pierre Lafaye de Micheaux. Goodness-of-fit tests of normality for the innovations in arma models. *Journal of Time Series Analysis*, 25(3):373–395, 2004.
- [30] Philip A Ernst, Lawrence D Brown, Larry Shepp, and Robert L Wolpert. Stationary gaussian markov processes as limits of stationary autoregressive time series. *Journal of Multivariate Analysis*, 155:180–186, 2017.
- [31] Tal Ezer, Larry P Atkinson, William B Corlett, and Jose L Blanco. Gulf stream's induced sea level rise and variability along the us mid-atlantic coast. *Journal of Geophysical Research: Oceans*, 118(2):685–697, 2013.
- [32] Tal Ezer and William Bryce Corlett. Analysis of relative sea level variations and trends in the chesapeake bay: Is there evidence for acceleration in sea level rise? In 2012 Oceans, pages 1–5. IEEE, 2012.
- [33] Tal Ezer and William Bryce Corlett. Is sea level rise accelerating in the chesapeake bay? a demonstration of a novel new approach for analyzing sea level data. *Geophysical Research Letters*, 39(19), 2012.
- [34] Jianqing Fan and Wenyang Zhang. Generalised likelihood ratio tests for spectral density. Biometrika, 91(1):195–209, 2004.
- [35] Donald P Gaver and PAW Lewis. First-order autoregressive gamma sequences and point processes. Advances in Applied Probability, 12(3):727–745, 1980.
- [36] GK Grunwald, RJ Hyndman, and LM Tedesco. A unified view of linear ar (1) models. 1995.

- [37] John A Hall, Christopher P Weaver, Jayantha Obeysekera, Mark Crowell, Radley M Horton, Robert E Kopp, John Marburger, Douglas C Marcy, Adam Parris, William V Sweet, et al. Rising sea levels: Helping decision-makers confront the inevitable. *Coastal* management, 47(2):127–150, 2019.
- [38] James Hamilton. D.(1994), time series analysis, 1994.
- [39] Keith W Hipel and A Ian McLeod. *Time series modelling of water resources and environmental systems*. Elsevier, 1994.
- [40] Wassily Hoeffding, Herbert Robbins, et al. The central limit theorem for dependent random variables. *Duke Mathematical Journal*, 15(3):773–780, 1948.
- [41] SJ Holgate and PL Woodworth. Evidence for enhanced coastal sea level rise during the 1990s. Geophysical research letters, 31(7), 2004.
- [42] JRM Hosking. A unified derivation of the asymptotic distributions of goodness-of-fit statistics for autoregressive time-series models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 40(3):341–349, 1978.
- [43] M Imani, R-J You, and C-Y Kuo. Caspian sea level prediction using satellite altimetry by artificial neural networks. *International Journal of Environmental Science and Technology*, 11(4):1035–1042, 2014.
- [44] H Bâki Iz, L Berry, and M Koch. Modeling regional sea level rise using local tide gauge data. Journal of Geodetic Science, 2(3):188–199, 2012.
- [45] Patricia A Jacobs and Peter AW Lewis. Discrete time series generated by mixtures.
 i: Correlational and runs properties. Journal of the Royal Statistical Society: Series B (Methodological), 40(1):94–105, 1978.
- [46] Patricia A Jacobs and Peter AW Lewis. Discrete time series generated by mixtures ii: asymptotic properties. Journal of the Royal Statistical Society: Series B (Methodological), 40(2):222–228, 1978.
- [47] S Jevrejeva, T Frederikse, RE Kopp, Goneri Le Cozannet, LP Jackson, and RSW van de Wal. Probabilistic sea level projections at the coast by 2100. Surveys in Geophysics, 40(6):1673–1696, 2019.
- [48] S Jevrejeva, A Grinsted, JC Moore, and S Holgate. Nonlinear trends and multiyear cycles in sea level records. *Journal of Geophysical Research: Oceans*, 111(C9), 2006.

- [49] DN Joanes and CA Gill. Comparing measures of sample skewness and kurtosis. *Journal* of the Royal Statistical Society: Series D (The Statistician), 47(1):183–189, 1998.
- [50] Karl G Jöreskog, Ulf H Olsson, and Fan Y Wallentin. Multivariate analysis with LIS-REL. Springer, 2016.
- [51] Jana Jurečková, Hira L Koul, and Jan Picek. Testing the tail index in autoregressive models. Annals of the Institute of Statistical Mathematics, 61(3):579–598, 2009.
- [52] Nobuhisa Kashiwagi and Takemi Yanagimoto. Smoothing serial count data through a state-space model. *Biometrics*, pages 1187–1194, 1992.
- [53] Robert E Kopp, Jerry X Mitrovica, Stephen M Griffies, Jianjun Yin, Carling C Hay, and Ronald J Stouffer. The impact of greenland melt on local sea levels: a partially coupled analysis of dynamic and static equilibrium effects in idealized water-hosing experiments. *Climatic Change*, 103(3-4):619–625, 2010.
- [54] Hira L Koul. Minimum distance estimation and goodness-of-fit tests in first-order autoregression. The Annals of Statistics, pages 1194–1213, 1986.
- [55] Hira L Koul, Shiqing Ling, et al. Fitting an error distribution in some heteroscedastic time series models. *The Annals of Statistics*, 34(2):994–1012, 2006.
- [56] Hira L Koul, Winfried Stute, et al. Nonparametric model checks for time series. The Annals of Statistics, 27(1):204–236, 1999.
- [57] Honghai Li, Lihwa Lin, and Kelly A Burks-Copes. Modeling of coastal inundation, storm surge, and relative sea-level rise at naval station norfolk, norfolk, virginia, usa. *Journal of coastal research*, 29(1):18–30, 2013.
- [58] Wai Keung Li. *Diagnostic checks in time series*. CRC Press, 2003.
- [59] Wai Keung Li and A Ian McLeod. Arma modelling with non-gaussian innovations. Journal of Time Series Analysis, 9(2):155–168, 1988.
- [60] WK Li and AI McLeod. Distribution of the residual autocorrelations in multivariate arma time series models. *Journal of the Royal Statistical Society: Series B (Method*ological), 43(2):231–239, 1981.
- [61] Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. Biometrika, 65(2):297–303, 1978.

- [62] Jenny N Lye and Vance L Martin. Non-linear time series modelling and distributional flexibility. Journal of Time Series Analysis, 15(1):65–84, 1994.
- [63] Gerald A Meehl, Curt Covey, Thomas Delworth, Mojib Latif, Bryant McAvaney, John FB Mitchell, Ronald J Stouffer, and Karl E Taylor. The wcrp cmip3 multimodel dataset: A new era in climate change research. *Bulletin of the American meteorological* society, 88(9):1383–1394, 2007.
- [64] Gerard A Meehl, Thomas F Stocker, William D Collins, Pierre Friedlingstein, T Gaye, Jonathan M Gregory, Akio Kitoh, Reto Knutti, James M Murphy, Akira Noda, et al. Global climate projections. 2007.
- [65] Mark F Meier, Mark B Dyurgerov, Ursula K Rick, Shad O'neel, W Tad Pfeffer, Robert S Anderson, Suzanne P Anderson, and Andrey F Glazovsky. Glaciers dominate eustatic sea-level rise in the 21st century. *Science*, 317(5841):1064–1067, 2007.
- [66] Terence C Mills. A Statistical Biography of George Udny Yule: A Loafer of the World. Cambridge Scholars Publishing, 2017.
- [67] Jerry X Mitrovica, Mark E Tamisiea, James L Davis, and Glenn A Milne. Recent mass balance of polar ice sheets inferred from patterns of global sea-level change. *Nature*, 409(6823):1026–1029, 2001.
- [68] Richard H Moss, Jae A Edmonds, Kathy A Hibbard, Martin R Manning, Steven K Rose, Detlef P Van Vuuren, Timothy R Carter, Seita Emori, Mikiko Kainuma, Tom Kram, et al. The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756, 2010.
- [69] Manfred Mudelsee. Climate time series analysis. Springer, 2013.
- [70] Robert J Nicholls and Anny Cazenave. Sea-level rise and its impact on coastal zones. science, 328(5985):1517–1520, 2010.
- [71] PE O'Connell and DA Jones. Some experience with the development of models for the stochastic simulation of daily flows. *Inputs for Risk Analysis in Water Resources*, pages 287–314, 1979.
- [72] M Perrette, F Landerer, R Riva, K Frieler, and M Meinshausen. A scaling approach to project regional sea level rise and its uncertainties. 2013.
- [73] David L Prothero and Kenneth F Wallis. Modelling macroeconomic time series. Journal of the Royal Statistical Society: Series A (General), 139(4):468–486, 1976.

- [74] Asbury H Sallenger, Kara S Doran, and Peter A Howd. Hotspot of accelerated sea-level rise on the atlantic coast of north america. *Nature Climate Change*, 2(12):884–888, 2012.
- [75] Robert H Shumway and David S Stoffer. Time series analysis and its applications. Studies In Informatics And Control, 9(4):375–376, 2000.
- [76] Chiaw Hock Sim. First-order autoregressive logistic processes. Journal of applied probability, 30(2):467–470, 1993.
- [77] ABA Slangen, Mark Carson, CA Katsman, RSW Van de Wal, Armin Köhl, LLA Vermeersen, and Detlef Stammer. Projecting twenty-first century regional sea-level changes. *Climatic Change*, 124(1-2):317–332, 2014.
- [78] ABA Slangen, CA Katsman, RSW Van de Wal, LLA Vermeersen, and REM Riva. Towards regional projections of twenty-first century sea-level change based on ipcc sres scenarios. *Climate dynamics*, 38(5-6):1191–1209, 2012.
- [79] Giorgio Spada, JL Bamber, and RTWL Hurkmans. The gravitationally consistent sealevel fingerprint of future terrestrial ice loss. *Geophysical Research Letters*, 40(3):482– 486, 2013.
- [80] William V Sweet, Robert E Kopp, Christopher P Weaver, Jayantha Obeysekera, Radley M Horton, E Robert Thieler, and Chris Zervas. Global and regional sea level rise scenarios for the united states. 2017.
- [81] William V Sweet and Joseph Park. From the extreme to the mean: Acceleration and tipping points of coastal inundation from sea level rise. *Earth's Future*, 2(12):579–600, 2014.
- [82] A Louise Swift and Gareth J Janacek. Forecasting non-normal time series. Journal of Forecasting, 10(5):501–520, 1991.
- [83] Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. An overview of cmip5 and the experiment design. Bulletin of the American Meteorological Society, 93(4):485–498, 2012.
- [84] Ruey S Tsay. Time series and forecasting: Brief history and future research. *Journal* of the American Statistical Association, 95(450):638–643, 2000.
- [85] Aad W Van der Vaart. Time series. VU University Amsterdam, lecture notes, 2010.

- [86] Manouchehr Vaziri. Predicting caspian sea surface water level by ann and arima models. Journal of waterway, port, coastal, and ocean engineering, 123(4):158–162, 1997.
- [87] Hans Visser, Soenke Dangendorf, and Arthur C Petersen. A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate". *Journal of Geophysical Research: Oceans*, 120(6):3873–3895, 2015.
- [88] AM Walker. Some properties of the asymptotic power functions of goodness-of-fit tests for linear autoregressive schemes. Journal of the Royal Statistical Society: Series B (Methodological), 14(1):117–134, 1952.
- [89] W Watson, TG King, Trevor A Spedding, and KJ Stout. The machined surface—time series modelling. Wear, 57(1):195–205, 1979.
- [90] Mike West, P Jeff Harrison, and Helio S Migon. Dynamic generalized linear models and bayesian forecasting. *Journal of the American Statistical Association*, 80(389):73–83, 1985.
- [91] PL Woodworth. A search for accelerations in records of european mean sea level. International Journal of Climatology, 10(2):129–143, 1990.
- [92] PL Woodworth, Neil J White, S Jevrejeva, SJ Holgate, JA Church, and WR Gehrels. Evidence for the accelerations of sea level on multi-decade and century timescales. International Journal of Climatology: A Journal of the Royal Meteorological Society, 29(6):777–789, 2009.
- [93] Jianjun Yin, Michael E Schlesinger, and Ronald J Stouffer. Model projections of rapid sea-level rise on the northeast coast of the united states. *Nature Geoscience*, 2(4):262– 266, 2009.
- [94] G. Udny Yule. On the time-correlation problem, with especial reference to the variatedifference correlation method. *Journal of the Royal Statistical Society*, 84(4):497–537, 1921.