Specification tests for autoregressive conditional duration (ACD) models

DJ Kok

orcid.org 0000-0002-4686-5777

Dissertation accepted in fulfilment of the requirements for the degree *Master of Science in Mathematical Statistics* at the North-West University

Supervisor:	Dr M Smuts
Co-Supervisor:	Prof GL Grobler
Co-Supervisor:	Dr LM Ndwandwe

Graduation May 2024

Acknowledgements

First and foremost, I would like to thank our Heavenly Father, for the countless blessings, strength, and guidance in the preparation and completion of this dissertation.

I would like to express my profound and sincere gratitude to the following people who have supported me throughout this journey. Firstly, my supervisor, Dr. Marius Smuts, and co-supervisors, Prof. Gerrit Grobler and Dr. Lethani Ndwandwe, thank you for your patience, motivation, immense knowledge, guidance, and continued support throughout this process.

To my parents, Anton and Merinda, for their unconditional love, support, and constant encouragement. I would not be here if it weren't for you.

Lastly, I am deeply grateful to my sister, Zan-Mari, for words of encouragement when I needed it the most, and also for proofreading my dissertation.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Abstract

In this dissertation, we propose specification tests for the innovation distribution in autoregressive conditional duration (ACD) models, with specific emphasis on ACD(1,1) models with exponential and Lomax innovations. Literature on ACD models with Lomax innovations is relatively scarce, in particular specification tests for these models. As a result, we introduce a newly developed test for Lomax innovations, which is based on a characterisation of the Lomax law using Stein's method. For the hypothesis of exponential innovations, we compare the powers of classical tests as well as some modern tests using a simulation study. The classical tests include the Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling tests. Testing for Lomax distributed innovations, we compare the new test with classical tests, to analyse its finite sample performance. The numerical comparisons presented are conducted for various sample sizes and for a number of alternative distributions. Testing for exponential innovations, the results of the simulation study showed that the tests with the best overall performance are the score test by Cox and Oakes, as well as the Anderson-Darling test. The results for Lomax innovations showed that the new test outperformed all the classical tests. We illustrate some of the tests using a real-data example.

Key words: *ACD models, Specification test, High-frequency data, Goodness-of-fit, Warp speed bootstrap*

Table of Contents

Chapter 1

Introduction

Due to the rising availability of tick-by-tick data (or high-frequency data), the popularity of conditional duration (CD) models has recently increased tremendously (Meintanis et al., 2020). According to De Luca and Zuccolotto (2004), the main feature of high-frequency data is that the observations are no longer equally spaced in time; therefore, the duration between consecutive observations is treated as a random variable. As a result, standard time series models such as generalised autoregressive conditional heteroskedastic (GARCH) models can not be used when working with high-frequency data, since these models are based on observations equally spaced in time (Yan, 2021).

To circumvent the aforementioned inadequacy, Engle and Russell (1998) introduced the autoregressive conditional duration (ACD) model which is based on a point process and shares many features with the GARCH model. Since its introduction, the ACD model has been widely developed and has become a leading tool for modelling high-frequency financial data. An exhaustive review on the early development of ACD models was done by Pacurar (2008), whereas Bhogal and Thekke Variyam (2019) provides a comprehensive review of the theoretical and empirical work that has been done on conditional duration models since. More recently, Cavaliere et al. (2023) showed that the asymptotic theory of parameter estimators in ACD models, was based on the work done on GARCH models, which assumes nonrandom number of events. This contradicts the fact that the number of events in high-frequency data is random. Therefore, they introduced new limiting distribution results for likelihood estimators, which take the randomness of the events into account.

According to Meintanis et al. (2020), ACD models consist of two components: the first is duration as a function of the past, also called conditional duration, while the second is an independent and identical distributed (i.i.d.) component often called the "innovation". They also mentioned that there exist two specification problems, one concerning the functional form of conditional duration and the other concerning the correct specification of the innovation distribution. The main aim of this dissertation is to present specification tests for the innovation distribution in ACD models, with an emphasis on exponential as well as Lomax innovations. These tests are analogous to goodness-of-fit (GOF) tests.

The Lomax distribution, also known as the Pareto Type II distribution, was initially introduced to model business failure data (Nombebe et al., 2023). Literature on ACD models with Lomax innovations is relatively scarce, however, De Luca and Zuccolotto (2004) showed that the performance of these models is favourable compared to ACD models with other innovation distributions.

The primary objectives of this dissertation can be summarised as follows:

- Introduce ACD models and its theoretical framework.
- Provide an overview of the Lomax distribution and its use in the ACD context.
- Provide a discussion of existing GOF tests for the innovation distribution in ACD models.
- Provide a newly proposed test for testing Lomax distributed innovations in ACD models.
- Use a Monte Carlo study to test for exponential innovations, and obtain similar results as those in Meintanis et al. (2020).
- Use a Monte Carlo study to analyse the finite sample performance of the newly proposed test.
- Apply the newly proposed test with other existing tests to real-world data to assess the new test's competitiveness.

The dissertation is structured as follows: Chapter 2 introduces ACD models, including its theoretical framework, the standard ACD model, and one of its extensions (i.e., the ACD model with standard exponential innovations). Chapter 3 provides a brief overview of the Lomax distribution, followed by an examination of its use in ACD models. In Chapter 4, we discuss some existing GOF tests for the innovation distribution in ACD models. Additionally, we introduce a new test that can be used to test the hypothesis of Lomax distributed innovations. Chapter 5 contains Monte Carlo results of GOF tests for ACD models with exponential innovations, as well as an analysis of the newly proposed test for ACD(1,1) models with Lomax innovations. Chapter 6 provides a real-world example, on which some of the GOF tests considered in the simulation study, are applied. The dissertation concludes in Chapter 7 with some final remarks.

Chapter 2

ACD models

Since its introduction, the ACD model has been widely developed and has become a leading tool for modelling financial data irregularly spaced in time. As mentioned in Bhogal and Thekke Variyam (2019), Saart et al. (2015) noticed that ACD models can be categorised into three generations. The "*First generation*" ACD model, which was introduced by Engle and Russell (1998), assumes that the conditional durations are exponentially distributed and is often referred to as an exponential ACD or EACD model. The main feature of this model is that the standardised durations or innovations are i.i.d. with nonnegative support. Later, numerous studies have been dedicated to exploring various parametric extensions of the *first generation* model. These extensions, which are classified as the "*Second generation*" models, aim to enhance the dynamic specification of CD models, as well as the shape of the hazard function, offering increased flexibility. The Weibull ACD (WACD) and the gamma ACD (GACD) models were mentioned in Engle and Russell (1998) as possible extensions, while the Burr ACD model suggested by Grammig and Maurer (2000), the Log-ACD model of Bauwens and Giot (2000), the nonlinear ACD model of Zhang et al. (2001), the d-ACD model of Zuccolotto et al. (2002), the Asymmetric ACD model of Bauwens and Giot (2003), and the augmented ACD model of Fernandes and Grammig (2006) are also examples of second-generation models. More recently, researchers have started exploring semiparametric and nonparametric methods, aiming to establish the "*Third generation*" of models. The aim of these models is to provide a more useful generalisation to the ACD procedure. A notable instance of such research is the work presented by Drost and Werker (2004), who advocated for a semiparametric alternative for the i.i.d. assumption in *first* and *second generation* models, permitting the distribution function of the standardised durations to be conditional on the past. Meanwhile, Cosma and Galli (2006) addressed the nonlinearity in ACD models, by introducing a nonparametric regression procedure, which resulted in the Nonparametric ACD (N-ACD) model.

In the following section, we give the theoretical framework for ACD models. Section 2.2 introduces the ACD model, and Section 2.3 gives the standard ACD model. Finally, we state the EACD(1,1) model in Section 2.4, whereafter we provide a brief discussion on the randomness of the number of financial events and its effect on the asymptotics of parameter estimators of ACD models.

2.1 Theoretical framework

2.1.1 Point processes

As mentioned previously, the primary feature of high-frequency data (HFD) is that the observations are not equally spaced in time. However, since they are ordered in time, the time series of observed durations can be viewed as realisations of point processes. Bauwens and Giot (2001) defines a point process as "a unique kind of stochastic process that generates a random collection of points on the time axis". The simplest and most well-known point process is the Poisson process. Point processes are extensively applied in neuroscience, queueing theory, as well as in finance. Engle and Russell (1998) were the first to introduce them in a high-frequency finance context; however, the use thereof was detailed by Engle (2000), whereupon it was regarded as the starting point of the interest attracted.

Following Engle and Russell's framework (1998), let $\{t_0, t_1, \ldots, t_n, \ldots\}$ be a stochastic process which is a sequence of arrival times with $t_0 \leq t_1 \leq \cdots \leq t_n \leq \ldots$. The arrival times of the process are represented by the times of financial events, such as trades and quotes, contained in a high-frequency financial dataset. As can be seen above, simultaneous occurrence of events is possible; however, this possibility is usually eliminated in practice by ordinarily making use of thinning procedures. Let the counting function $N(t)$ be the number of events that have occurred by time $t \in [0, T]$, which is a step function that is left continuous with right limits. Clearly, $N(t)$ is a nondecreasing function of time with $N(t_0) = 0$. Furthermore, $t_{N(T)} = T$ denotes the last observed point of the sequence of arrival times, and $0 = t_0 \leq t_1 \leq \ldots \leq t_{N(T)} = T$ the observed point process. The characteristics associated with an event, such as the trading volume and price of a trade, are known as marks as they further describe the event that occurred. If there are marks associated with the arrival times, the process is called a marked point process. Following Snyder and Miller (1991), the two general characterisations of a point process, namely *conditionally orderly* and *evolve without after-effects*, which express the independence of the past and future of a point process, were introduced by Engle and Russell (1998):

> A point process on $[t_0, \infty)$ is said to *"evolve without after-effects"* if for any $t > t_0$, the realisation of points during $[t, \infty)$ does not depend in any way on the sequence of points during the interval $[t_0, t)$. A counting process is said to be *"conditionally orderly"* at time $t \geq t_0$ if, for a sufficiently short interval of time and conditional on any event P defined by the realisation of the process on $[t_0, t)$, the probability of two or more events occurring is infinitesimal relative to the probability of one event. (p. 1129)

The point processes used by Engle and Russell (1998) evolve with after-effects and are conditionally orderly. The abovementioned processes are defined in terms of the intensity, which is conditional on the available past information, which must at least include the count and the arrival times. Therefore, the conditional intensity process is defined as

$$\lambda\left(t \middle| N(t), t_0, \ldots, t_{N(t)}\right) = \lim_{\Delta t \to 0} \frac{P\left(N(t+\Delta t) > N(t) \middle| N(t), t_0, \ldots, t_{N(t)}\right)}{\Delta t}. \tag{2.1}$$

The conditional intensity in (2.1) is also known as the hazard function, particularly in survival analysis. Let f_i denote a family of conditional probability density functions for arrival time t_i, the log-likelihood, ℓ, can be expressed in terms of the conditional densities

$$\ell = \sum_{i=1}^{N(T)} \log\left(f_i\left(t_i | t_0, \ldots, t_{i-1}\right)\right), \tag{2.2}$$

or intensities

$$\ell = \sum_{i=1}^{N(T)} \log\left(\lambda\left(t_i | i-1, t_0, \ldots, t_{i-1}\right)\right) - \int_{t_0}^{T} \lambda\left(s \middle| N(s), t_0, \ldots, t_{N(s)}\right) ds. \tag{2.3}$$

The conditional intensity, the conditional density of the durations, and the conditional survival function completely describe a conditionally orderly process, as discussed in Lancaster (1990) and Snyder and Miller (1991). The parameterisation of the conditional intensity determines the success of using such processes.

2.1.2 General setup

Studying financial point processes usually consists of modelling the process of durations between consecutive points (Pacurar, 2008). Denote the ith duration, also known as the waiting time, between two events that occur at times t_{i-1} and t_i as $x_i = t_i - t_{i-1}$, and $\{z_0, z_1, \ldots, z_n, \ldots\}$ the sequence of marks associated with arrival times $\{t_0, t_1, \ldots, t_n, \ldots\}$. It is obvious that the sequence of durations $\{x_1, x_2, \ldots, x_{N(T)}\}$ has nonnegative elements, which impacts the choice of econometric models suitable for the durations. Following the framework of Engle (2000), we present the joint sequence of durations and marks as

$$\{(x_i, z_i), i = 1, \ldots, T\},$$

where $N(T) = T$ is known. Let F_{i-1} denote the filtration, which is the information set available at time t_{i-1}. Then, the joint density of the ith observation, conditional on F_{i-1}, is given by

$$(x_i, z_i)|F_{i-1} \sim p(x_i, z_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_x}, \boldsymbol{\theta_z}), \tag{2.4}$$

where $\breve{x}_{i-1} = \{x_{i-1}, x_{i-2}, \ldots, x_1\}$ and $\breve{z}_{i-1} = \{z_{i-1}, z_{i-2}, \ldots, z_1\}$ denote the past of x and z respectively and $\boldsymbol{\theta_x}$ and $\boldsymbol{\theta_z}$ are the set of parameters associated with x_i and z_i respectively. Without loss of generality, we can write the joint density in (2.4) as the product of the conditional density of the marks given the duration and the marginal density of the durations, where both are conditional upon the past marks and durations. That is,

$$p(x_i, z_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_x}, \boldsymbol{\theta_z}) = g(x_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_x})q(z_i|x_i, \breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_z}),$$

where $g\left(x_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_x}\right)$ is the marginal density of the duration x_i with parameter $\boldsymbol{\theta_x}$, conditional on past marks and durations, and $q\left(z_i|x_i, \breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_z}\right)$ is the conditional density of the mark z_i and parameter $\boldsymbol{\theta_z}$, conditional on the duration x_i as well as the past marks and durations. The log-likelihood can therefore be written as

$$\ell\left(\boldsymbol{\theta_x}, \boldsymbol{\theta_z}\right) = \sum_{i=1}^{T} \left[\log\left(g\left(x_i|\breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_x}\right)\right) + \log\left(q\left(z_i|x_i, \breve{x}_{i-1}, \breve{z}_{i-1}; \boldsymbol{\theta_z}\right)\right)\right]. \tag{2.5}$$

If we consider the durations to be weakly exogenous to the marks, then the two parts of the likelihood function in (2.5) are usually maximised separately, effectively simplifying the estimation (see, for example, Engle, 2000).

In this study we will not consider marks for the point process. Therefore, the log-likelihood in (2.5) can be written as

$$\ell\left(\boldsymbol{\theta_x}\right) = \sum_{i=1}^{T} \log\left(g\left(x_i|\breve{x}_{i-1}; \boldsymbol{\theta_x}\right)\right).$$

2.2 Models for the durations

The ACD model introduced by Engle (2000), which is a type of dependent Poisson process, is specified in terms of the conditional density of the durations x_i. They define ψ_i as the conditional expected duration given by

$$\psi_i \equiv \psi_i\left(\breve{x}_{i-1}\right) = \mathrm{E}\left(x_i|F_{i-1}\right). \tag{2.6}$$

The ACD model assumes that

$$x_i = \psi_i \varepsilon_i, \tag{2.7}$$

where the standardised durations, also called innovations, $\varepsilon_i = {}^{x_i}/_{\psi_i}$ are i.i.d. with $\mathrm{E}\left(\varepsilon_i\right) = 1$. Note that the unit mean assumption is without loss of generality. If $\mathrm{E}(\varepsilon_i) \neq 1$, we can express $x_i = \psi_i' \varepsilon_i'$, where $\varepsilon_i' = {}^{\varepsilon_i}/_{\mathrm{E}(\varepsilon_i)}$ and $\psi_i' = \psi_i \mathrm{E}(\varepsilon_i)$ such that $\mathrm{E}(\varepsilon_i') = 1$. The assumption in (2.7) requires that the temporal dependence in the durations to be captured by the conditional expected duration, which implies that $g\left(x_i|\breve{x}_{i-1}; \boldsymbol{\theta_x}\right) = g\left(x_i|\psi_i; \boldsymbol{\theta_x}\right)$. The density function of ε_i with parameters $\boldsymbol{\theta_\varepsilon}$ is given by $f_{\varepsilon_i}\left(u|\boldsymbol{\theta_\varepsilon}\right)$, which then has a nonnegative support. Subsequently, $g\left(x_i|F_{i-1}; \boldsymbol{\theta_x}\right) = \psi_i^{-1} f({}^{x_i}/_{\psi_i}|\boldsymbol{\theta_\varepsilon})$, where $\boldsymbol{\theta_x} = (\boldsymbol{\theta_\psi}, \boldsymbol{\theta_\varepsilon})$ is the set of parameters. Thus, the log-likelihood function is given by

$$\ell\left(\boldsymbol{\theta_x}\right) = \sum_{i=1}^{T} \log\left(g\left(x_i|F_{i-1}; \boldsymbol{\theta_x}\right)\right) = \sum_{i=1}^{T} \left[\log\left(f\left(\frac{x_i}{\psi_i}\middle|\boldsymbol{\theta_\varepsilon}\right)\right) - \log\left(\psi_i\right)\right]. \tag{2.8}$$

After specifying a parametric distribution for ε_i, estimates of $\boldsymbol{\theta_x}$ can be obtained by the use of several numerical optimisation algorithms. Since the setup in (2.6) and (2.7) is very general, a variety of models can be obtained by using different distributions for ε_i, as well as different specifications for the expected duration ψ_i. In the following section, we will discuss the standard ACD model.

2.3 Standard ACD model

The standard ACD model of Engle and Russell (1998) consists of a linear parameterisation of (2.6), in which the expected duration ψ_i relies on p past durations and q past expected durations, and is given by

$$\psi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}, \tag{2.9}$$

where the conditions $\omega > 0, \alpha_j \geq 0, \beta_j \geq 0$ are required to ensure nonnegative conditional durations. The abovementioned model is referred to as the ACD(p,q) model. When examining the ACD model alongside Bollerslev's (1986) generalised autoregressive conditional heteroskedasticity (GARCH) model, it becomes evident that they exhibit several shared characteristics. The ACD model is often regarded as the analogous counterpart of the GARCH model for duration data. The AR part of equation (2.9) captures duration clustering observed in HFD, where we see a pattern of large (small) durations being followed by other large (small) durations, analogous to how the GARCH model accounts for volatility clustering.

A convenient feature of the ACD model is that it can be formulated as an ARMA(max(p,q),q) model for the durations x_i. Specifically, let $\eta_i \equiv x_i - \psi_i$ be a martingale difference sequence (i.e., $\mathrm{E}(\eta_i|F_{i-1}) = 0$), the ACD($p$,$q$) model in (2.9) after rearranging terms becomes

$$x_i = \omega + \sum_{j=1}^{\max(p,q)} (\alpha_j + \beta_j)\, x_{i-j} - \sum_{j=1}^{q} \beta_j \eta_{i-j} + \eta_i. \tag{2.10}$$

Pacurar (2008) deduced from this ARMA representation, that "all the coefficients within the infinite-order AR representation implied by inverting the MA component must be nonnegative". This will establish a well-defined process for durations. Furthermore, for x_i to be weakly stationary, the following condition is sufficient

$$\sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1.$$

Hence, for an ACD(1,1) model it is necessary for the roots of polynomials $1 - (\alpha_1 + \beta_1)z$ and $(1 - \beta_1)z$ to lie outside the unit circle, in order to satisfy the stationarity and invertibility conditions. Equation (2.10) can also be used to derive the autocorrelation function (ACF) of the ACD model (see Bauwens and Giot (2000), for a detailed discussion).

The conditional mean of x_i is equal to ψ_i, resulting from definition (2.6), whereas the unconditional mean is given by

$$\mathrm{E}\left(x_i\right) = \frac{\omega}{1 - \sum_{j=1}^{p} \alpha_j - \sum_{j=1}^{q} \beta_j}.$$

From (2.7), the conditional variance of x_i is given by

$$\mathrm{Var}\left(x_i|F_{i-1}\right) = \psi_i^2 \mathrm{Var}\left(\varepsilon_i\right).$$

The unconditional variance of x_i is based on (2.7) and (2.9), however to obtain a general expression for the ACD(m,q) model is rather cumbersome. For the ACD(1,1) model, however, the unconditional variance can be derived as follows

$$\mathrm{E}\left(x_i^2\right) = \mathrm{E}\left(\psi_i^2\right) - \mathrm{E}\left(\varepsilon_i^2\right) = \frac{\omega^2 \mathrm{E}\left(\varepsilon_i^2\right)}{1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2 \mathrm{E}\left(\varepsilon_i^2\right)} + \frac{\left(\alpha_1 + \beta_1\right) 2\omega^2 \mathrm{E}\left(\varepsilon_i^2\right)}{\left(1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2 \mathrm{E}\left(\varepsilon_i^2\right)\right)\left(1 - \alpha_1 - \beta_1\right)}$$

Therefore,

$$\mathrm{Var}\left(x_i\right) = \mathrm{E}\left(x_i^2\right) - \mathrm{E}\left(x_i\right)^2 = \mathrm{E}\left(x_i\right)^2 \left[\frac{\mathrm{E}\left(\varepsilon_i^2\right)\left(1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2\right) - \left(1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2 \mathrm{E}\left(\varepsilon_i^2\right)\right)}{1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2 \mathrm{E}\left(\varepsilon_i^2\right)}\right].$$

If $\alpha_1 > 0$, the unconditional standard deviation of x_i will be larger than the unconditional mean. This phenomenon is usually referred to as overdispersion/excess dispersion and is often noticed in duration data (see Engle and Russell (1998)). Similar to the GARCH(1,1) model which is regarded as a good starting point, an ACD(1,1) model is seen as a natural starting point. The ACD(1,1) model's statistical properties are also extensively studied by Engle and Russell (1998), who derived the first two moments while Bauwens and Giot (2000) computed the ACF.

Throughout this dissertation, we will focus on two ACD models: the ACD(1,1) with exponential errors, also known as EACD(1,1), and the ACD(1,1) model with Lomax errors, also known as LACD(1,1). We will introduce the LACD(1,1) model in Chapter 3.

2.4 EACD(1,1) model

In this section we will introduce the simplest member of the ACD family, which assumes standard exponential (i.e. the shape parameter is equal to one) innovations, named the EACD(1,1) model.

The model can be written as

$$x_i = \psi_i \varepsilon_i,$$

where

$$\psi_i = \omega + \alpha_1 x_{i-1} + \beta_1 \psi_{i-1}$$

and

$$\varepsilon_i \sim \mathrm{Exp}(1),$$

with expected value $\mathrm{E}(\varepsilon_i) = 1$ and variance $\mathrm{Var}(\varepsilon_i) = \mathrm{E}\left(\varepsilon_i^2\right) - \mathrm{E}\left(\varepsilon_i\right)^2 = 2 - 1 = 1$.

The unconditional mean and variance of the ith duration are given as

$$\mathrm{E}(x_i) = \frac{\omega}{1 - \alpha_1 - \beta_1}$$

and

$$\mathrm{Var}\left(x_i\right) = \mathrm{E}\left(x_i^2\right) - \mathrm{E}\left(x_i\right)^2 = \mathrm{E}\left(x_i\right)^2 \left(\frac{1 - \beta_1^2 - 2\alpha_1\beta_1}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 2\alpha_1^2}\right)$$

respectively.

One of the main advantages of assuming that the innovations follow a standard exponential distribution is that it provides quasi-maximum likelihood estimators (QMLE) for the ACD parameters of a general ACD(1,1) model, as can be seen in Engle and Russell (1998) and Engle (2000).

In this specific case the log-likelihood function of the EACD(1,1) model takes the form

$$\ell(\boldsymbol{\vartheta}) = -\sum_{i=1}^{T} \left[\left(\frac{x_i}{\psi_i} \right) + \log(\psi_i) \right], \tag{2.11}$$

where $\boldsymbol{\vartheta}$ are the unknown model parameters to be estimated.

2.5 Random number of events

Engle and Russell (1998) used ACD models to model the durations between financial events, which are observed over a given period of time, resulting in the number of durations being random. Until now, the asymptotic theory for likelihood-based estimators for these models treats the number of events, $N(T)$, as deterministic, hence not random. Cavaliere et al. (2022) showed that the number of events being random has a significant impact on asymptotics and inference in duration models. As a result, they provide asymptotic theory for likelihood-based estimators, which takes the randomness of the number of events into account.

Engle and Russell (1998) observed that the log-likelihood in (2.11) is similar to the log-likelihood function of the GARCH(1,1) model with Gaussian innovations, and as a result, quoted the asymptotic theory of GARCH models in Lee and Hansen (1994). They showed that the asymptotic normality (AN) of the QMLEs holds if the durations are stationary and ergodic for the ACD(1,1) model. Additionally, these conditions allow $\alpha_1 + \beta_1 \geq 1$ and hence, for the durations to have infinite mean. However, as mentioned previously, these results are based on the number of events $N(T)$ being nonrandom. Cavaliere et al. (2023) provided an additional, more restrictive, condition for the ACD(1,1) model when $N(T)$ is random, stating that the durations need to have finite mean or equivalently, finite unconditional expectation (i.e. $\alpha_1 + \beta_1 < 1$).

Cavaliere et al. (2022) show that taking the randomness of the events into account, the asymptotic theory for QMLEs significantly depends on the tail behaviour, characterised by a tail index of the marginal distribution of the durations. Furthermore, they showed that for an ACD(1,0) model, where $\beta = 0$, that the AN breaks down if the tail index is less than one. Specifically, for the EACD(1,0) model, the AN breaks down if $\alpha \notin [0, a_u)$, with $a_u = e^{\lambda} \simeq 1.8$, where λ is Euler's constant. A wide range of tail indices were observed in their empirical applications on a wide range of duration data.

These new developments can have an impact on our study since one needs to take the randomness of the events into account when applying the bootstrap procedure to obtain estimates of the model parameters; which complicates matters. In this dissertation, we will treat the number of events as deterministic. In the next chapter, the Lomax distribution in an ACD setup will be presented.

Chapter 3

Lomax distribution

In this chapter, the Lomax distribution and the ACD(1,1) model with Lomax innovations will be discussed. The Pareto distribution, also referred to as the Pareto type I distribution, is a heavy-tailed distribution that was first introduced by Pareto (1897) to model the distribution of wealth among individuals. Over the years, the popularity of the Pareto distribution has led to various modifications and adaptions, giving rise to several variants known as the Pareto types II; III; and IV; as well as the so-called generalised Pareto distribution (GPD).

Lomax (1954) proposed the Lomax distribution, which is a special case of the Pareto type II distribution with the location parameter set to zero. This distribution has been used as an alternative to the exponential, gamma, and Weibull distributions for heavy-tailed data by Bryson (1974). The Lomax distribution was originally used to model business failures and has subsequently been used in a variety of fields such as actuarial science, medical and biological sciences and engineering. Furthermore, it has been applied to model data obtained from income and wealth by Harris (1968), and Atkinson and Harrison (1978), sizes of files on a computer server by Holland et al. (2006), reliability and life testing by Hassan and Al-Ghamdi (2009), and firm size by Corbellini et al. (2010).

3.1 Characteristics of the Lomax distribution

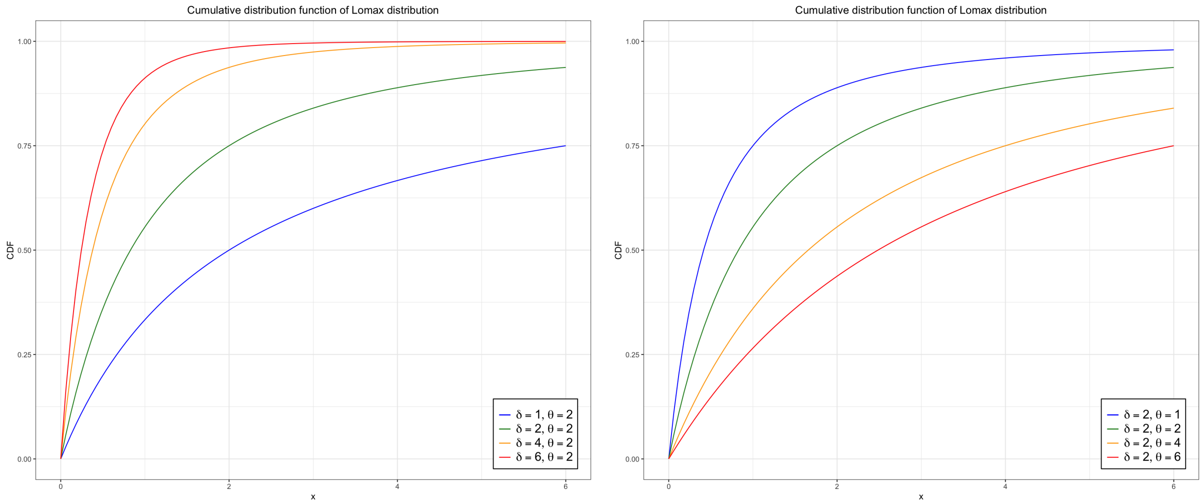
Let X be a nonnegative random variable, then X is Lomax distributed with parameters δ and θ ($X \sim \text{Lomax}(\delta, \theta)$) if X has the distribution function

$$F(x|\delta, \theta) = 1 - \left[1 + \frac{x}{\theta}\right]^{-\delta}, \qquad x \geq 0,$$

where $\delta, \theta > 0$ are shape and scale parameters respectively. The density function of X is given by

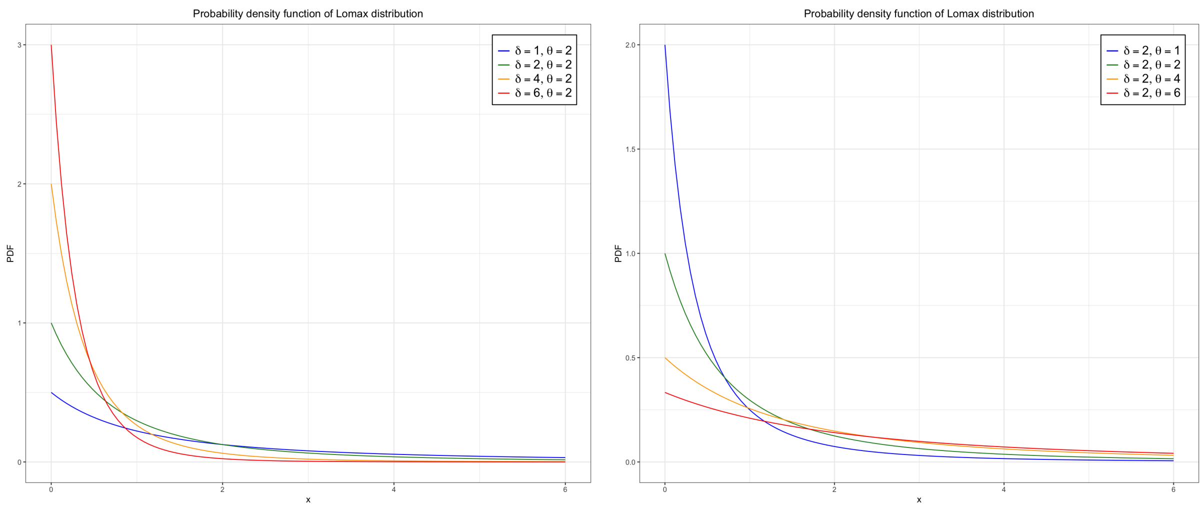
$$f(x|\delta, \theta) = \theta^{\delta}\delta(\theta + x)^{-(\delta+1)}, \qquad x \geq 0.$$

Some possible shapes of the probability density function (pdf) and the cumulative distribution function (cdf) of the Lomax distribution are given in Figures 3.1 and 3.2.



(a) Different δ with $\theta = 2$ (b) Different θ with $\delta = 2$

Fig. 3.1: Distribution function of Lomax distribution with different parameter combinations.



(a) Different δ with $\theta = 2$ (b) Different θ with $\delta = 2$

Fig. 3.2: Density function of Lomax distribution with different parameter combinations.

Properties of the Lomax distribution will now be discussed. The mean and the variance of the Lomax distribution are

$$E[X] = \frac{\theta}{\delta - 1}, \qquad \delta > 1$$

and

$$\mathrm{Var}[X] = \frac{\theta^2 \delta}{(\delta - 1)^2(\delta - 2)}, \qquad \delta > 2.$$

We also state the skewness and kurtosis below, since with duration models leptokurtic skewed distributions are popular:

$$\text{Skew}[X] = \frac{2(1+\delta)}{\delta-3}\sqrt{\frac{\delta-2}{\delta}}, \qquad \delta > 3$$

and

$$\text{Kurt}[X] = 3 + \frac{6(\delta^3+\delta^2-6\delta-2)}{\delta(\delta-3)(\delta-4)}, \qquad \delta > 4.$$

We notice that if $\delta = 1$, then $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ does not exist. Also as δ approaches one from the right side, $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ increases, while if δ goes to infinity, $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ goes to zero. This can be seen in the left planes of Figures 3.1 and 3.2. On the other hand, $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ has a linear and quadratic relationship with θ, where both the mean and variance increase as θ increases.

There exists a relationship between the Lomax distribution and a few other well-known distributions. These include, but are not limited to, distributions such as the Pareto Type I and Type IV, the beta prime distribution (also called beta distribution of the second kind), the GPD, and the log-logistic distribution. These relationships are given below:

- The Lomax distribution is a Pareto Type I distribution shifted so that its support begins at zero. Therefore, if $Y \sim \mathit{Pareto}(\delta, \theta)$, then $Y - \theta \sim \text{Lomax}(\delta, \theta)$.
- The Lomax distribution is a Pareto Type IV with location parameter $\mu = 0$ and inequality parameter $\gamma = 1$.
- If $X \sim \text{Lomax}(\delta, \theta)$ and $Y = \frac{X}{\theta}$, then the distribution function of Y is

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P\left(\frac{X}{\theta} \leq y\right) \\
&= P(X \leq \theta y) \\
&= F_X(\theta y) \\
&= 1 - \left(1 + \frac{\theta y}{\theta}\right)^{-\delta} \\
&= 1 - (1+y)^{-\delta}, \qquad y \geq 0.
\end{aligned}$$

Therefore, the density is given by

$$f_Y(y) = \delta(1+y)^{-\delta-1}, \qquad y \geq 0,$$

which is the density of a special case of the beta prime distribution with one of its shape parameters equal to one. Specifically $Y \sim \beta'(1, \delta)$.

- The Lomax distribution is a generalised Pareto distribution with location parameter $\mu = 0$, shape parameter $\xi = \frac{1}{\delta}$, and scale parameter $\beta = \frac{\theta}{\delta}$.
- If $X \sim \text{Lomax}(1, \theta)$ and $Y = \log(X)$, the distribution function of Y is given as

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(\log(X) \leq y) \\
&= P(X \leq e^y) \\
&= F_X(e^y) \\
&= 1 - \left(1 + \frac{e^y}{\theta}\right)^{-1} \\
&= \frac{1 + e^y/\theta - 1}{1 + e^y/\theta} \\
&= \frac{e^y}{\theta}\left(1 + \frac{e^y}{\theta}\right)^{-1} \\
&= \left(1 + e^{-(y - \log\theta)}\right)^{-1}, \qquad -\infty < y < \infty.
\end{aligned}$$

Therefore, Y has a logistic distribution with location $\log(\theta)$ and scale 1. This implies that X equals a log-logistic distribution with shape 1 and scale θ.

Chahkandi and Ganjali (2009) mentions that the Lomax distribution is also considered important in lifetime modelling, since it belongs to the family of decreasing failure rates. Several generalisations for the Lomax distribution have been developed to enhance the flexibility of the new distribution to model more diverse types of data, which mainly consists of adding scale/shape parameters accordingly. Some worth mentioning include the exponentiated Lomax by Abdul-Moniem (2012), Poisson-Lomax by Al-Jarallah et al. (2014), Weibull-Lomax by Tahir et al. (2015), power-Lomax by Rady et al. (2016), Rayleigh-Lomax distribution by Fatima et al. (2018), Marshall–Olkin length biased Lomax by Mathew and Chesneau (2020), and the Kumaraswamy generalised power-Lomax by Nagarjuna et al. (2021) (see Nagarjuna et al. (2022), for more examples).

Now that we have introduced the Lomax distribution, the ACD(1,1) model with Lomax innovations will be discussed.

3.2 LACD(1,1) model

In this section the ACD model with Lomax innovations is introduced, which we will refer to as the LACD(1,1) model. The LACD(1,1) model can be written as

$$x_i = \psi_i \varepsilon_i,$$

where

$$\psi_i = \omega + \alpha_1 x_{i-1} + \beta_1 \psi_{i-1}$$

and

$$\varepsilon_i \sim \text{Lomax}(\delta, \theta).$$

We need to constrain one of the parameters due to the unit mean restriction. Let $\delta = \theta + 1$, then $\varepsilon_i \sim \text{Lomax}(\theta + 1, \theta)$. The density function now only depends on θ and can be written as

$$f_{\varepsilon_i}(u|\theta) = \theta^{(\theta+1)}(\theta+1)(\theta+u)^{-(\theta+2)}, \qquad u \geq 0.$$

Therefore, the expected value and variance of ε_i are given by

$$\text{E}(\varepsilon_i) = 1$$

and

$$\text{Var}(\varepsilon_i) = \frac{\theta+1}{\theta-1},$$

respectively. From Section 2.2, we can obtain the density of the conditional duration $x_i|\psi_i$

$$\begin{aligned} g(x_i|\psi_i, \boldsymbol{\vartheta}) &= \frac{1}{\psi_i} f\left(\frac{x_i}{\psi_i}\middle|\theta\right) \\ &= \frac{\theta^{(\theta+1)}}{\psi_i}(\theta+1)\left(\theta + \frac{x_i}{\psi_i}\right)^{-(\theta+2)}, \qquad x_i \geq 0, \end{aligned}$$

where $\boldsymbol{\vartheta} = \{\omega, \alpha_1, \beta_1, \theta\}$ is the model parameters. The unconditional mean and variance of the ith duration are given as

$$\mathrm{E}(x_i) = \frac{\omega}{1-\alpha_1-\beta_1}$$

and

$$\mathrm{Var}\left(x_i\right) = \mathrm{E}\left(x_i^2\right) - \mathrm{E}\left(x_i\right)^2 = \mathrm{E}\left(x_i\right)^2 \left[\frac{(2\theta)\left(1-\beta_1^2-2\alpha_1\beta_1-\alpha_1^2\right)}{(\theta-1)\left(1-\beta_1^2-2\alpha_1\beta_1-\frac{2\alpha_1^2\theta}{\theta-1}\right)} - 1\right],$$

respectively. From (2.8), the parameters of the ACD model and θ can be jointly estimated by maximising the log-likelihood function

$$\begin{aligned}
\ell(\boldsymbol{\vartheta}) &= \sum_{i=1}^{T}\left[\log\left(f\left(\frac{x_i}{\psi_i}\right)\middle|\theta\right) - \log(\psi_i)\right] \\
&= \sum_{i=1}^{T}\left[(\theta+1)\log(\theta) - \log(\psi_i) + \log(\theta+1) - (\theta+2)\log\left(\theta+\frac{x_i}{\psi_i}\right)\right],
\end{aligned}$$

where $\boldsymbol{\vartheta}$ are the unknown model parameters to be estimated. In the next chapter, we will consider goodness-of-fit tests for assessing the fit of ACD models assuming different innovation distributions.

Chapter 4

Goodness-of-fit tests

From the earliest days of statistics, statisticians have commenced their analysis by proposing a distribution for their observations, whereafter, perhaps with somewhat less enthusiasm, they would then investigate whether the aforementioned distribution was true. Therefore, a vast number of test procedures have appeared over the years, whereupon the study of these procedures has come to be known as goodness-of-fit (D'Agostino, 2017). In this chapter, we will discuss several GOF tests for the innovation distribution in ACD models. The focus will specifically be on tests that have proven powerful for testing GOF in the i.i.d. setting, to the current context of ACD, as well as newly developed tests.

Let $\boldsymbol{x}_T = (x_1, x_2, \ldots, x_T)$ be observations from model (2.7) and (2.9). The composite hypothesis to be tested is

$$H_0 : F \in \mathcal{F}_{\boldsymbol{\theta}},$$

where $\mathcal{F}_{\boldsymbol{\theta}}$ is a specified family of distributions with parameter $\boldsymbol{\theta}$, against general alternatives.

All the test statistics that will be considered are based on the corresponding residuals

$$\widehat{\varepsilon}_t = \frac{x_t}{\widehat{\psi}_t}, \qquad t = 1, 2, ..., T,$$

where $\widehat{\psi}_t$ is an appropriate estimator of the conditional expected duration ψ_t under the ACD specification. Let $\widehat{\varepsilon}_{(1)} \leq \widehat{\varepsilon}_{(2)} \leq \cdots \leq \widehat{\varepsilon}_{(T)}$ denote the order statistics corresponding to $\widehat{\varepsilon}_1, \widehat{\varepsilon}_2, \ldots, \widehat{\varepsilon}_T$.

For the hypothesis of exponentially distributed innovations, we consider classical tests, as well as modern tests discussed in Meintanis et al. (2020). However, when testing for Lomax distributed innovations, we only consider classical tests and a newly proposed test.

4.1 Classical tests based on the empirical distribution function

The classical GOF tests are based on the empirical distribution function (edf)

$$\widehat{F}_T(x) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{I}(\widehat{\varepsilon}_t \leq x), \qquad 0 < x < \infty,$$

where $\boldsymbol{I}(\cdot)$ denotes the indicator function. Specifically, we consider the Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling tests. Each of these tests is based on the comparison between the hypothetical distribution function $F_0(\cdot)$ and the edf $\widehat{F}_T(\cdot)$, using a specific distance measure.

4.1.1 Kolmogorov-Smirnov statistic (KS_T)

The Kolmogorov-Smirnov test statistic, proposed by Kolmogorov (1933), is

$$KS_T = \sup_{x\geq 0} \left| \widehat{F}_T(x) - F_0(x) \right|.$$

This test statistic can be simplified to

$$KS_T = \max\{KS_T^+, KS_T^-\},$$

where

$$KS_T^+ = \max_{1\leq t\leq T} \left[\frac{t}{T} - F_0(\widehat{\varepsilon}_{(t)})\right],$$

$$KS_T^- = \max_{1\leq t\leq T} \left[F_0(\widehat{\varepsilon}_{(t)}) - \frac{t-1}{T}\right].$$

Large values of KS_T will lead to the rejection of the null hypothesis.

4.1.2 Cramér-von Mises statistic (CM_T)

The Cramér-von Mises test statistic is given by

$$CM_T = T\int_0^\infty \left(\widehat{F}_T(x) - F_0(x)\right)^2 dF_0(x).$$

This test statistic can be simplified (see Smit (2018) for the derivation) to

$$CM_T = \frac{1}{12T} + \sum_{t=1}^{T}\left(F_0(\widehat{\varepsilon}_{(t)}) - \frac{2t-1}{2T}\right)^2.$$

This test rejects the null hypothesis for large values of CM_T.

4.1.3 Anderson-Darling statistic (AD_T)

Anderson and Darling (1954) introduced a reweighted version of the Cramér-von Mises test statistic, known as the Anderson-Darling statistic which is given by

$$AD_T = T \int_0^\infty \frac{\left(\widehat{F}_T(x) - F_0(x)\right)^2}{F_0(x)(1 - F_0(x))} dF_0(x).$$

This test statistic can be simplified (see Appendix A.1 for the derivation) to

$$AD_T = -T - \frac{1}{T} \sum_{t=1}^{T} \left((2t-1)\log\left(F_0(\widehat{\varepsilon}_{(t)})\right) + (2T+1-2t)\log\left(1 - F_0(\widehat{\varepsilon}_{(t)})\right)\right).$$

This test also rejects the null hypothesis for large values of AD_T.

The tests discussed above can be applied to test for any distribution and are referred to as classical tests in the literature. We will now discuss tests specifically testing for the exponential distribution.

4.2 A test based on a score function

The score function is the first derivative (or gradient) of the log-likelihood function, which is a powerful tool that can be used to test statistical hypotheses. The test that we consider which employs this score function, is the test by Cox and Oakes (1984).

Cox-Oakes test (CO_T)

The score test in Cox and Oakes (1984) is used in a survival analysis setting where there is censoring present in the data. Since we are working with observations that are all observable, the test statistic has the following form

$$CO_T = T + \sum_{t=1}^{T} (1 - \widehat{\varepsilon}_t) \log(\widehat{\varepsilon}_t).$$

The test rejects the null hypothesis for both small and large values of CO_T.

4.3 A test based on the Puri-Rubin characterisation

The test we consider, which is developed by Milošević and Obradović (2016), is based on the characterisation by Puri and Rubin (1970). This characterisation says the following

Let X_1 and X_2 denote two independent copies of a random variable X with pdf $f(x)$. Then X and $|X_1 - X_2|$ have the same distribution if and only if for some $\lambda > 0$, $\quad f(x) = \lambda e^{-\lambda x}$, for $x \geq 0$.

Milošević-Obradović test statistic (MO_T)

The Milošević-Obradović test statistic (see Milošević and Obradović (2016)) is given by

$$MO_T = \frac{1}{T(T-1)} \sum_{1 \leq t < s \leq T} \left(\frac{2}{a + |\widehat{\varepsilon}_t - \widehat{\varepsilon}_s|} - \frac{1}{a + \widehat{\varepsilon}_t} - \frac{1}{a + \widehat{\varepsilon}_s} \right),$$

where $a > 0$ is a constant tuning parameter. Large values of MO_T will lead to the rejection of H_0.

4.4 A test based on the empirical characteristic function

The empirical characteristic function (ECF) has a long history as a tool for statistical inference. For GOF problems in particular, early work was done by Heathcote (1972) and Feigin and Heathcote (1976). More recent research includes, among others, Henze et al. (2003), Gürtler and Henze (2000), Koutrouvelis and Meintanis (1999) and Kankainen and Ushakov (1998). A large part of the literature on the ECF is covered by Ushakov (1999).

In these tests the characteristic function (cf) of a random variable X, given by

$$\phi(t) = \mathrm{E}[e^{itX}],$$

is estimated by the ECF of the data $x_1, x_2, \ldots, x_T$, defined as

$$\widehat{\phi}_T(t) = \frac{1}{T} \sum_{j=1}^{T} e^{itx_j}.$$

We will only focus on the Epps-Pulley test proposed by Epps and Pulley (1986) in this dissertation.

Epps-Pulley test statistic (EP_T)

The test of Epps and Pulley (1986) is based on the difference between the ECF $\widehat{\phi}_T(t)$ defined above, and the cf of the exponential distribution,

$$\phi_0(t, \lambda) = \frac{\lambda}{\lambda - it},$$

where λ is the parameter of the exponential distribution.

The Epps-Pulley test statistic is given by

$$EP_T = \sqrt{48T}\left(\frac{1}{T}\sum_{t=1}^{T} e^{-\widehat{\varepsilon}_t} - \frac{1}{2}\right).$$

This test rejects H_0 for large values of $|EP_T|$. In the next section we will introduce a newly proposed test to test for Lomax innovations.

4.5 A test based on a characterisation of the Lomax law using Stein's method

The test statistic, newly proposed by Ebner et al. (2023) in the i.i.d. setting, is based on the characterisation of the Lomax law using the method of Stein et al. (2004). This approach states that, under suitable conditions, a real-valued random variable X follows a distribution with differentiable density f in $[a, b]$, $-\infty \leq a \leq b \leq \infty$, if and only if

$$\begin{aligned}
\mathrm{E}\left[p'(X) + \frac{f'(X)}{f(X)}p(X)\right] &= \int_a^b \left(p'(x) + \frac{f'(x)}{f(x)}p(x)\right) f(x)dx \\
&= \int_a^b \left(p'(x)f(x) + f'(x)p(x)\right) dx \\
&= \int_a^b \frac{d}{dx}\left(p(x)f(x)\right) dx \\
&= \lim_{x\to b^-} p(x)f(x) - \lim_{x\to a^+} p(x)f(x) \\
&= p(b^-)f(b^-) - p(a^+)f(a^+),
\end{aligned}$$

where $p(\cdot)$ is a function from a sufficiently large class of test functions. For example, in our case, if we have $a = 0$ and $b = \infty$, then $\lim_{x\to\infty} p(x)$ and $\lim_{x\to 0} p(x)$ should exist. $p(x)$ should also be smooth on the interval $[0, \infty)$. One example of a function, that was used in Ebner et al. (2023) is $p(x) = e^{-tx}$. This was also used similarly in Ebner et al. (2022). We apply the above on the innovations $\varepsilon_i, i = 1, 2, \ldots, T$. If

$\varepsilon_i \sim \text{Lomax}(\theta+1,\theta)$, recall that the density function of ε_i is given by

$$f_{\varepsilon_i}(u;\theta) = \theta^{(\theta+1)}(\theta+1)(\theta+u)^{-(\theta+2)}, \qquad u \geq 0.$$

Using the above choice of $p(\cdot)$ leads to the following characterisation of the Lomax distribution: $\varepsilon_i \sim \text{Lomax}(\theta+1,\theta)$ if, and only if

$$\begin{aligned}
\mathrm{E}\left[-e^{-t\varepsilon_i}\left\{t+(\theta+2)(\theta+\varepsilon_i)^{-1}\right\}\right] &= p(\infty)f(\infty) - p(0)f(0) \\
&= \lim_{\varepsilon_i\to\infty} e^{-t\varepsilon_i}\frac{\theta^{(\theta+1)}(\theta+1)}{(\theta+\varepsilon_i)^{(\theta+2)}} - \lim_{\varepsilon_i\to 0} e^{-t\varepsilon_i}\frac{\theta^{(\theta+1)}(\theta+1)}{(\theta+\varepsilon_i)^{(\theta+2)}} \\
&= 0 - \frac{\theta^{(\theta+1)}(\theta+1)}{\theta^{(\theta+2)}} \\
&= -\frac{\theta+1}{\theta}.
\end{aligned}$$

Therefore

$$\mathrm{E}\left[e^{-t\varepsilon_i}\left\{t+(\theta+2)(\theta+\varepsilon_i)^{-1}\right\}\right] = \frac{\theta+1}{\theta}.$$

The test statistic is given by

$$S_{T,a} = T\int_0^\infty \left\{\frac{1}{T}\sum_{j=1}^{T}\left[t+(\widehat{\theta}+2)((\widehat{\theta}+\widehat{\varepsilon}_j)^{-1}\right]e^{-t\widehat{\varepsilon}_j} - \frac{\widehat{\theta}+1}{\widehat{\theta}}\right\}^2 w(t)dt,$$

where $w(t) = e^{-at}$ is a weight function with $a > 0$ a constant tuning parameter. For this choice of weight function, the test statistic has the easily calculable form (see Appendix A.2 for the derivation)

$$\begin{aligned}
S_{T,a} = \frac{1}{T}\sum_{i,j=1}^{T}&\left\{\frac{2}{(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)^3} + \frac{2(\widehat{\theta}+2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)^2} + \frac{(\widehat{\theta}+2)^2}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\theta}+\widehat{\varepsilon}_j)(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)}\right\} \\
&- \frac{2(\widehat{\theta}+1)}{\widehat{\theta}}\sum_{i=1}^{T}\left\{\frac{1}{(\widehat{\varepsilon}_i+a)^2} + \frac{(\widehat{\theta}+2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\varepsilon}_i+a)}\right\} + \frac{T(\widehat{\theta}+1)^2}{\widehat{\theta}^2 a},
\end{aligned}$$

where $\widehat{\theta}$ is an estimator of θ.

In the next chapter, we will employ a power study to test the efficiency of the various GOF tests presented in this chapter.

Chapter 5

Specification tests and Simulation study

In this chapter, the Monte Carlo setup to evaluate the power of the various tests presented in Chapter 4, is discussed. First, we describe the estimation of the model parameters $\boldsymbol{\vartheta}$, followed by a discussion of the resampling procedure used to approximate the asymptotic null distribution of the test statistics. The empirical powers for an ACD(1,1) model with exponential and Lomax innovation distributions are presented thereafter. For the ACD(1,1) model with Lomax innovations, an analysis of the sensitivity of maximum likelihood estimation on the optimisation procedure's starting values is presented. Furthermore, we discuss the choice of values used for the tuning parameter, a, to obtain the approximated powers of the newly proposed test, $S_{T,a}$.

A significance level of 5% is used throughout the simulation study. All calculations are done in R (R Core Team, 2023). Power estimates are calculated for sample sizes $T = \{100, 200, 400\}$ using 10 000 Monte Carlo replications for various alternative distributions.

5.1 Estimation of model parameters

For an ACD(1,1) model, consider the observations $\boldsymbol{x}_T = (x_1, x_2, \ldots, x_T)$. In this dissertation, conditional maximum likelihood estimation (CMLE) is used, which incorporates the density of the standard durations under the null hypothesis H_0 (Tsay, 2010). The likelihood function of the durations $\boldsymbol{x}_T$ is

$$L(\boldsymbol{x}_T|\boldsymbol{\vartheta}) = g(x_1; \boldsymbol{\vartheta}) \prod_{t=2}^{T} g(x_t|F_{t-1}, \boldsymbol{\vartheta}), \tag{5.1}$$

where $\boldsymbol{\vartheta} = (\boldsymbol{\theta}_x, \boldsymbol{\theta}_\varepsilon)$ denotes the vector of model parameters. Since the marginal probability density function $g(x_1; \boldsymbol{\vartheta})$ in (5.1) is too complicated for a general ACD model, and its impact on the likelihood function is very small for sufficiently large T, we can ignore the factor and work with the conditional likelihood function

$$L_C(\boldsymbol{x}_T|\boldsymbol{\vartheta}) = \prod_{t=2}^{T} g(x_t|F_{t-1}, \boldsymbol{\vartheta})$$
$$= \prod_{t=2}^{T} \left(\frac{1}{\psi_t} f_{\varepsilon_t} \left(\frac{x_t}{\psi_t} \right) \right),$$

where f_{ε_t} is the density of the innovations ε_t under the null hypothesis H_0. The maximum likelihood estimates can be derived by maximising the conditional log-likelihood function

$$\ell_C(\boldsymbol{x}_T|\boldsymbol{\vartheta}, x_1) = \sum_{t=2}^{T} \left[\log \left(f_{\varepsilon_t} \left(\frac{x_t}{\psi_t} \right) \right) - \log(\psi_t) \right].$$

Therefore, the conditional log-likelihood function of an ACD(1,1) model with exponential and Lomax innovation distributions is given by

$$\ell_C(\boldsymbol{x}_T|\boldsymbol{\vartheta}, x_1) = -\sum_{t=2}^{T} \left[\left(\frac{x_t}{\psi_t} \right) + \log(\psi_t) \right],$$

and

$$\ell_C(\boldsymbol{x}_T|\boldsymbol{\vartheta}, x_1) = \sum_{t=2}^{T} \left[(\theta+1)\log(\theta) + \log(\theta+1) - (\theta+2)\log\left(\theta + \frac{x_t}{\psi_t}\right) - \log(\psi_t) \right], \tag{5.2}$$

respectively. Note that θ in (5.2) represents the Lomax parameter. Another popular method used is the quasi-maximum likelihood estimation (QMLE), where estimates can be obtained by choosing an appropriate distribution for ε_t and maximising the conditional maximum likelihood function using the chosen distribution; however, only CMLE, where x_1 is known, is considered in this dissertation.

5.2 Resampling procedure

Due to the complexity of the asymptotic null distribution of GOF test statistics in an autoregressive-like scenario, as can be seen in Lee and Oh (2015), a parametric bootstrap methodology is used to calculate critical values, which are then used to obtain empirical powers. However, to reduce the computational cost of calculating these powers, we employ the warp-speed bootstrap of Giacomini et al. (2013), which essentially involves using a single bootstrap replication for each generated Monte Carlo sample. The warp-speed bootstrap, implemented in the context of ACD models can be summarised as follows

1. Simulate a sample $x_1, ..., x_T$ from a hypothesised model with true parameter-vector $\boldsymbol{\vartheta}_0 = (\omega, \alpha_1, \beta_1, \theta_\varepsilon)$.

2. Obtain the CMLE $\widehat{\boldsymbol{\vartheta}}_T = \vartheta_T(x_1, \ldots, x_T)$ of the parameter-vector $\boldsymbol{\vartheta}_0$.

3. Compute the test statistic $S = S(\widehat{\varepsilon}_1, ..., \widehat{\varepsilon}_T)$, where $\widehat{\varepsilon}_t = \frac{x_t}{\widehat{\psi}_t}$ and $\widehat{\psi}_t = \widehat{\omega} + \widehat{\alpha}_1 x_{t-1} + \widehat{\beta}_1 \widehat{\psi}_{t-1}$.

4. Simulate an i.i.d. sample $\varepsilon_t^*, \quad t = 1, ..., T$, from the distribution under the null hypothesis F_0.

5. Using $\varepsilon_t^*,\ t = 1, ..., T$, and $\widehat{\boldsymbol{\vartheta}}_T$ construct pseudo-observations $x_1^*, ..., x_T^*$, from $x_t^* = \widehat{\psi}_t^* \varepsilon_t^*$, where $\widehat{\psi}_t^* = \widehat{\omega} + \widehat{\alpha}_1 x_{t-1}^* + \widehat{\beta}_1 \widehat{\psi}_{t-1}^*$; with initial value $x_1^* = \widehat{\psi}_1^* = \bar{x}$, where $\bar{x} = T^{-1} \sum_{t=1}^{T} x_t$.

6. Compute the CMLE $\widehat{\boldsymbol{\vartheta}}_T^* = \vartheta_T(x_1^*, \ldots, x_T^*)$ and $\widehat{\psi}_t^* = \widehat{\omega}^* + \widehat{\alpha}_1^* x_{t-1}^* + \widehat{\beta}_1^* \widehat{\psi}_{t-1}^*$.

7. Compute the test statistic $S^* = S(\widehat{\varepsilon}_1^*, ..., \widehat{\varepsilon}_T^*)$, where $\widehat{\varepsilon}_t^* = \frac{x_t^*}{\widehat{\psi}_t^*}, \quad t = 1, ..., T$.

8. Repeat steps 1-7 a number of times, say MC, and obtain the sequence of test statistics S_m, $m = 1, ..., MC$ and bootstrap test statistics $S_m^*,\ m = 1, ..., MC$.

9. Use the MC bootstrap test statistics to estimate the sample $(1 - \alpha)$ quantile, $c_{T,\alpha}^* := S_{(\mathrm{MC}(1-\alpha))}^*$, where $S_{(1)}^* \leq \ldots \leq S_{(T)}^*$, denote the corresponding order statistics.

10. The power estimate is then given by $\frac{1}{MC} \sum_{m=1}^{MC} \mathrm{I}\left(S_m > c_{T,\alpha}^*\right)$.

5.3 Exponential innovations

In this section, the results for the tests for exponential innovations are presented, with the aim to obtain similar results as those in Meintanis et al. (2020). The null hypothesis takes the form

$$H_0 : F(x) = 1 - e^{-x}, \qquad x > 0$$

where $F(\cdot)$ denotes the distribution function of the exponential distribution.

5.3.1 Accuracy of MLEs

Before conducting the power study, we need to test whether the estimation procedure (i.e. maximum likelihood estimation (MLE)) used in the bootstrap procedure, is accurate. This will ensure consistent estimators of the model parameters, and in turn, the reliability of the bootstrap procedure. We will do this by means of an accuracy study, focusing on the following metrics applied to the estimated parameters: Average of the MLEs (AE), standard errors (SE), biases, and roots of the mean squared errors (RMSE). To calculate the MLEs we simulate from an EACD(1,1) model with the same true model parameters as those used by Meintanis et al. (2020), i.e. $\boldsymbol{\vartheta_0} := (\omega, \alpha_1, \beta_1) = (0.3, 0.3, 0.4)$. The results of the accuracy study are given in Table 5.1, which reveal that the average of the maximum likelihood estimates are close to the true values even for the sample size of 100, and converge to the true parameter values as the sample size increases. We also see that the SEs, bias, and RMSEs decrease as the sample size increases, indicating that the underlying estimation procedure is accurate and producing consistent estimators.

Table 5.1: AE, SE, bias and RMSE for ACD(1,1) model with exponential innovations.

	AE	**SE**	**Bias**	**RMSE**
		$T = 100$		
ω	0.371	0.193	0.071	0.193
α_1	0.288	0.133	-0.012	0.133
β_1	0.332	0.233	-0.068	0.233
		$T = 200$		
ω	0.345	0.149	0.045	0.149
α_1	0.293	0.095	-0.007	0.095
β_1	0.357	0.189	-0.043	0.188
		$T = 400$		
ω	0.323	0.107	0.023	0.107
α_1	0.297	0.067	-0.003	0.067
β_1	0.377	0.137	-0.023	0.137

In the next section, we present the results of the power study. The simulated model we will focus on in the power study is the ACD(1,1) model where we again employ the same true model parameters as those used by Meintanis et al. (2020), i.e. $\boldsymbol{\vartheta_0} = (0.3, 0.3, 0.4)$.

5.3.2 Power study

The alternative distributions used in testing for the ACD(1,1) model with exponential innovations are given in Table 5.2. These alternative distributions were chosen since they are frequently used alternatives when testing for exponential innovations.

Table 5.2: Alternative distributions

Alternative	**f(x)**	**Notation**	**Support**
Weibull	$f(x) = \lambda^{\theta}\theta x^{\theta-1}e^{-(x\lambda)^{\theta}}, \quad \lambda = \Gamma(\frac{\theta+1}{\theta})$	$W(\theta)$	$(0,\infty)$
Gamma	$f(x) = \frac{\theta^{\theta}}{\Gamma(\theta)}x^{\theta-1}e^{-\theta x}$	$G(\theta)$	$(0,\infty)$
Lognormal	$f(x) = \frac{1}{\sqrt{2\pi}\theta x}e^{-\frac{\left(\log(x)+\frac{\theta^2}{2}\right)^2}{2\theta^2}}$	$LN(\theta)$	$(0,\infty)$

Tables 5.3 – 5.5 show the approximated sizes and powers for sample sizes $T = \{100, 200, 400\}$ against the alternative distributions in Table 5.2. The entries in these tables are the percentages of 10 000 Monte Carlo samples that resulted in the rejection of the null hypothesis. These approximations are rounded to the nearest integer, and the highest power for each distribution is highlighted.

Table 5.3: Tests for exponential innovations: rejection rate for $T = 100$

	$\boldsymbol{KS_T}$	$\boldsymbol{CM_T}$	$\boldsymbol{AD_T}$	$\boldsymbol{CO_T}$	$\boldsymbol{MO_T}$	$\boldsymbol{EP_T}$
E(1)	**5**	**5**	**5**	6	**5**	**5**
W(1.1)	17	20	19	**32**	29	30
W(1.2)	46	54	55	**73**	69	66
W(1.3)	76	84	85	**95**	93	91
W(1.4)	93	98	98	**100**	99	98
G(1.2)	19	23	24	**37**	35	33
G(1.3)	35	43	45	**64**	57	51
G(1.4)	54	64	66	**82**	75	72
G(1.5)	69	79	83	**93**	88	83
LN(0.8)	97	98	**100**	93	92	91
LN(0.9)	68	73	**90**	57	46	60
LN(1)	38	48	**65**	16	10	20
LN(1.1)	48	54	**58**	2	1	4

From Tables 5.3 - 5.5 we see that the empirical sizes are satisfactory and that the powers generally increase as the sample size increases. For Weibull and gamma alternatives, we notice that the powers increase with the shape parameter.

The powers of the KS_T test do not compare favourably to those of the other tests, while we find that the CO_T test generally outperforms the other tests, having the highest approximated power for these alternatives. Although the CO_T test is the most powerful, it is only marginally more powerful compared to MO_T and EP_T, specifically when the sample size increases. Considering the lognormal alternatives, the test that performs unfavourably in comparison to the other tests is the MO_T test. On the other hand, the KS_T and CM_T tests generally perform well, while the AD_T test is the most powerful test, as it outperforms the other tests.

Table 5.4: Tests for exponential innovations: rejection rate for $T = 200$

	$\boldsymbol{KS_T}$	$\boldsymbol{CM_T}$	$\boldsymbol{AD_T}$	$\boldsymbol{CO_T}$	$\boldsymbol{MO_T}$	$\boldsymbol{EP_T}$
E(1)	**5**	**5**	**5**	**5**	**5**	**5**
W(1.1)	28	35	37	**51**	48	47
W(1.2)	76	85	87	**96**	93	91
W(1.3)	97	99	99	**100**	**100**	**100**
W(1.4)	**100**	**100**	**100**	**100**	**100**	**100**
G(1.2)	35	42	45	**62**	55	53
G(1.3)	60	70	75	**88**	82	79
G(1.4)	84	90	94	**98**	95	93
G(1.5)	95	98	99	**100**	99	98
LN(0.8)	**100**	**100**	**100**	99	**100**	99
LN(0.9)	98	99	**100**	74	68	84
LN(1)	79	88	**97**	17	10	31
LN(1.1)	81	89	**94**	1	0	3

Overall, the difference between the results in this study and that of Meintanis et al. (2020) can be attributed to Monte Carlo errors. There are, however, some substantial differences with the powers obtained against the LN distribution for CO_T, MO_T and EP_T tests. For example, weak powers of $2, 1$ and 4 are obtained in the last row of Table 5.3 above, while the powers reported in Meintanis et al. (2020) are $23, 10$ and 42.

Table 5.5: Tests for exponential innovations: rejection rate for $T = 400$

	$\boldsymbol{KS_T}$	$\boldsymbol{CM_T}$	$\boldsymbol{AD_T}$	$\boldsymbol{CO_T}$	$\boldsymbol{MO_T}$	$\boldsymbol{EP_T}$
E(1)	**5**	**5**	**5**	**5**	**5**	4
W(1.1)	48	58	61	**76**	73	74
W(1.2)	97	99	99	**100**	**100**	99
W(1.3)	**100**	**100**	**100**	**100**	**100**	**100**
W(1.4)	**100**	**100**	**100**	**100**	**100**	**100**
G(1.2)	60	70	75	**88**	81	78
G(1.3)	90	94	97	**99**	98	96
G(1.4)	99	**100**	**100**	**100**	**100**	**100**
G(1.5)	**100**	**100**	**100**	**100**	**100**	**100**
LN(0.8)	**100**	**100**	**100**	**100**	**100**	**100**
LN(0.9)	**100**	**100**	**100**	89	92	97
LN(1)	**100**	**100**	**100**	17	12	45
LN(1.1)	99	**100**	**100**	0	0	2

5.4 Lomax innovations

We will now present results for the hypothesis of Lomax distributed innovations. In this case, the null hypothesis is

$$H_0 : F(x) = 1 - \left[1 + \frac{x}{\theta}\right]^{-(\theta+1)}, \qquad x > 0$$

where $F(\cdot)$ denotes the distribution function of the Lomax innovations. Firsty, we will test the convexity of the estimation procedure, followed by an accuracy study. Thereafter, we present the results of the power study, where we will also consider some local power approximations. The choice of tuning parameter, a, used to obtain the powers of the newly proposed test will be briefly discussed at the end of the section.

5.4.1 Sensitivity of MLEs: Local vs Global extreme points

Before executing the power study, we will investigate whether the estimation procedure used in Section 5.2 is concave, hence a concave log-likelihood function. If the concavity of the log-likelihood function holds, the function has a global extreme point (i.e. global maximum), resulting in consistent MLEs. A nonconcave estimation procedure is one where the log-likelihood function has local extreme points, which will produce inaccurate estimators. To test the concavity of the estimation procedure, we will analyse whether the procedure is sensitive to the choice of starting values, by making use of a simulation study. If the estimation procedure is sensitive, then the starting values will have an impact on the estimates, as the estimates will be inconsistent, which is an indication of nonconcave estimation.

As usual in such a study, we choose the Lomax distribution as the true distribution for the innovations. That is, in the estimation procedure, we need to maximise the log-likelihood function in (5.2). In the power study, we report approximated powers for various different alternative distributions. Therefore, we also test the sensitivity of the MLEs for some of the alternative distributions. Here we consider an exponential, a Weibull, and a gamma distribution. The steps to test the sensitivity are as follows:

1. Simulate a sample $x_1, ..., x_T$ from an ACD(1,1) model with specific distributed innovations and actual model parameters $\boldsymbol{\vartheta_0} = (\omega, \alpha_1, \beta_1, \theta)$, where θ is the parameter of the specific distribution.
2. Obtain the CMLE $\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T}} = (\widehat{\omega}, \widehat{\alpha}_1, \widehat{\beta}_1, \widehat{\theta})$ of $\boldsymbol{\vartheta_0}$, by using $\widetilde{\boldsymbol{\vartheta}}_{\boldsymbol{0}} := (\widetilde{\omega}, \widetilde{\alpha}_1, \widetilde{\beta}_1, \widetilde{\theta})$ as the starting values in the estimation procedure, where $\widetilde{\theta}$ is the starting value of the Lomax parameter and $\widehat{\theta}$ the estimate of the Lomax parameter.

If we choose the Lomax distribution as the specific distribution of the innovations, the estimates need to be close to the actual model parameters for the estimation procedure not to be sensitive. When one of the alternative distributions is used as the specific distribution of the innovations, the estimates do not necessarily need to be close to the actual model parameters; however, the values of the estimates should be similar and within the same range for the estimation procedure not to be sensitive.

Table 5.6: Actual model parameter values

	ω	α_1	β_1	θ
$\boldsymbol{\vartheta_{0_1}}$	0.5	0.49	0.30	4.0
$\boldsymbol{\vartheta_{0_2}}$	1.2	0.25	0.65	7.0
$\boldsymbol{\vartheta_{0_3}}$	4.0	0.69	0.20	12.0
$\boldsymbol{\vartheta_{0_4}}$	2.5	0.10	0.70	5.5
$\boldsymbol{\vartheta_{0_5}}$	4.2	0.60	0.20	2.0

The different combinations of actual model parameters we will use in the study are given in Table 5.6. In Table 5.7 we report several combinations of starting values for the estimation procedure that will be used in this simulation study. A sample size of $T = 400$ is used in this simulation study.

Table 5.7: Starting values

	$\widetilde{\omega}$	$\widetilde{\alpha}_1$	$\widetilde{\beta}_1$	$\widetilde{\theta}$
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_1}}$	0.15	0.12	0.80	1.1
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_2}}$	0.60	0.20	0.55	2.3
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_3}}$	1.30	0.31	0.60	3.7
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_4}}$	1.85	0.80	0.15	4.5
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_5}}$	2.10	0.65	0.20	5.1
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_6}}$	2.70	0.40	0.40	6.4
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_7}}$	3.40	0.52	0.35	7.7
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_8}}$	3.95	0.77	0.10	8.9
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_9}}$	4.50	0.45	0.29	9.2
$\widetilde{\boldsymbol{\vartheta}}_{\mathbf{0_{10}}}$	5.30	0.85	0.08	10.0

Lomax distribution

In Table 5.8 the MLEs are shown when Lomax distributed innovations are used for the simulated model. Here, the estimates need to be close to the actual model parameters $\boldsymbol{\vartheta}_{\mathbf{0_2}}$, for the estimation procedure not to be sensitive. From the results, we see that β_1 is underestimated, while the other parameters are overestimated. Although the estimates are not the same as the actual values, they are relatively close to the actual values, and each parameter's estimates are in the same range. The results for the other actual values are shown in Appendix B. We can therefore say that for an ACD(1,1) model with Lomax innovations, the MLE is not sensitive to the choice of starting values.

Exponential and Weibull distribution

The estimates of the ACD(1,1) model with exponential innovations are presented in Table 5.9, while the estimates of the ACD(1,1) model with Weibull innovations can be observed in Table 5.10. Here the estimates need to be in the same range for the estimation procedure to be concave. We notice from the abovementioned tables, that for both distributions, the estimates of the parameters are all in the same range. Therefore, We find that for an ACD(1,1) model with Exponential and Weibull innovations, the MLE is not sensitive to the choice of starting values. The results of the other choices of actual model parameters are given in Appendix B for exponential and Weibull innovations.

Table 5.8: Estimates of LACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_2}} = (1.2, 0.25, 0.65, 7)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	1.380	0.279	0.629	7.462
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	1.378	0.279	0.629	7.473
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	1.379	0.279	0.629	7.465
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	1.380	0.279	0.629	7.467
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	1.381	0.279	0.629	7.474
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	1.378	0.279	0.629	7.466
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	1.379	0.279	0.629	7.461
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	1.380	0.279	0.629	7.469
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	1.379	0.279	0.629	7.466
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	1.379	0.279	0.629	7.473

Table 5.9: Estimates of EACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_2}} = (1.2, 0.25, 0.65, 7)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.172	0.061	0.619	18.264
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.172	0.060	0.619	18.277
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.172	0.061	0.619	18.257
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.172	0.061	0.619	18.262
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.172	0.061	0.619	18.266
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.172	0.061	0.619	18.260
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.172	0.061	0.619	18.249
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.172	0.061	0.619	18.246
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.172	0.061	0.619	18.259
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.172	0.061	0.619	18.253

Gamma distribution

In Tables 5.11 and 5.12 we observe the MLEs when gamma distributed innovations are used for the simulated model. For the estimation procedure not to be sensitive to the choice of starting value, the estimates need to be in the same range. Table 5.11 shows that the estimates of the parameters are not all in the same range. We further see that $\widehat{\alpha}_1 + \widehat{\beta}_1 = 1$. For an ACD model to be stationary, $\alpha_1 + \beta_1 < 1$ needs to hold. Therefore, the model is not stationary. Similar results can be obtained from Table 5.12, where the model is also not stationary and the estimates of the parameters are also not in the same range.

Table 5.10: Estimates of WACD(1,1), with true parameters $\boldsymbol{\vartheta}_{\mathbf{0_1}} = (0.5, 0.49, 0.3, 4)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.495	0.531	0.206	18384.394
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.495	0.531	0.206	18640.321
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.496	0.531	0.206	18711.024
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.495	0.532	0.206	18428.544
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.495	0.532	0.206	18610.052
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.496	0.532	0.205	18607.159
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.495	0.531	0.206	18342.575
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.495	0.531	0.206	18609.604
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.496	0.531	0.206	18531.927
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.496	0.532	0.205	18588.616

Table 5.11: Estimates of GACD(1,1), with actual parameters $\boldsymbol{\vartheta}_{\mathbf{0_1}} = (0.5, 0.49, 0.3, 4)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.629	1.000	0.000	20.193
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	170.330	0.999	0.001	118.216
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.124	0.993	0.007	19.065
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.127	0.997	0.003	18.997
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.309	0.981	0.019	193.976
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.149	0.981	0.019	19.189
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	115.701	0.986	0.014	985.907
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.207	0.900	0.100	808.151
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.472	0.994	0.006	65.702
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.522	0.994	0.006	19.985

To see how the results will change when a lower value of the gamma parameter is used, we will now consider three different choices for the gamma parameter used in the actual model parameters $\boldsymbol{\vartheta}_{\mathbf{0_1}}$, $\theta = 2, 1.9$ and 1.8 respectively. From Tables 5.13 to 5.15, we observe that as the value of θ becomes smaller, the model becomes stationary ($\alpha_1 + \beta_1 < 1$), and the estimates of each parameter are in the same range.

Table 5.12: Estimates of GACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_4}} = (2.5, 0.1, 0.7, 5.5)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.228	1.000	0.000	98.915
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	114.039	0.644	0.356	2036.676
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	208.173	0.750	0.250	1734.948
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	422.166	0.816	0.184	1591.640
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	70.565	0.635	0.365	1933.013
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	41.227	0.815	0.185	1861.384
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	66.896	0.808	0.192	1852.607
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	141.695	0.795	0.205	1787.164
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	90.880	0.815	0.185	1813.745
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	49.658	0.813	0.187	1872.208

Table 5.13: Estimates of GACD(1,1), with adjusted actual parameters $\boldsymbol{\vartheta_{0_1}} = (0.5, 0.49, 0.3, 2)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.003	0.699	0.301	1505.676
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.002	0.695	0.305	972.028
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.003	0.695	0.305	1449.146
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.003	0.675	0.325	496.597
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.003	0.696	0.304	1593.187
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.002	0.696	0.304	907.030
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.002	0.703	0.297	1092.331
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.002	0.665	0.335	1638.144
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.001	0.695	0.305	1511.850
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.001	0.690	0.310	1644.434

We can therefore conclude that for an ACD(1,1) model with gamma innovations to be stationary and for the MLE not to be sensitive to the choice of starting values, we should choose $\theta \leq 1.8$. Similar results for other choices of true values can be found in Appendix B.

Table 5.14: Estimates of GACD(1,1), with adjusted actual parameters $\boldsymbol{\vartheta_{0_1}} = (0.5, 0.49, 0.3, 1.9)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.033	0.708	0.292	1507.534
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.014	0.685	0.315	324.750
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.015	0.686	0.314	389.590
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.009	0.671	0.329	644.574
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.016	0.683	0.317	438.324
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.006	0.680	0.320	487.966
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.021	0.687	0.313	1084.377
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.021	0.687	0.313	786.056
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.012	0.676	0.324	576.949
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.031	0.693	0.307	309.352

Table 5.15: Estimates of GACD(1,1), with adjusted actual parameters $\boldsymbol{\vartheta_{0_1}} = (0.5, 0.49, 0.3, 1.8)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	2.853	0.687	0.290	13495.974
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	2.850	0.687	0.290	13057.813
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	2.852	0.687	0.290	13506.593
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	2.849	0.688	0.289	13467.114
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	2.849	0.687	0.290	13480.956
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	2.850	0.687	0.290	13479.535
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	2.849	0.688	0.289	13499.822
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	2.852	0.687	0.290	13502.228
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	2.847	0.687	0.290	13444.576
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	2.851	0.687	0.290	13502.876

5.4.2 Accuracy of MLEs

Moving our attention to the accuracy of the estimation procedure, we again will use an accuracy study, by focusing on the AE, SE, biases, and RMSE of the etimates. The design of this study is the same as in Section 5.3.1. From the results of the accuracy study presented in Table 5.16, we observe that the average of maximum likelihood estimates is close to the majority of true values, and gets closer as the sample size increases. For the Lomax parameter, the average is far from the true value when $T = 100$,

Table 5.16: Average estimates (AE), SE, bias and RMSE for ACD(1,1) model with Lomax innovations.

	AE	**SE**	**Bias**	**RMSE**
		$T = 100$		
ω	0.361	0.206	0.061	0.206
α_1	0.292	0.162	-0.008	0.162
β_1	0.341	0.252	-0.059	0.252
θ	60.065	405.064	58.065	405.044
		$T = 200$		
ω	0.339	0.156	0.039	0.156
α_1	0.294	0.118	-0.006	0.118
β_1	0.362	0.204	-0.038	0.204
θ	4.928	88.844	2.928	88.840
		$T = 400$		
ω	0.322	0.110	0.022	0.110
α_1	0.298	0.085	-0.002	0.085
β_1	0.378	0.149	-0.022	0.148
θ	2.253	0.863	0.253	0.863

however when $T = 400$ the average is close to the true value. Additionally, we observe a decrease in SEs, bias, and RMSEs as the sample size increases. This indicates the accuracy of the underlying estimation approach, ensuring the production of consistent estimates.

5.4.3 Power study

In this section, the empirical powers for the ACD(1,1) model with Lomax innovations are presented. The alternative distributions given in Table 5.2 are used, as well as two additional distributions given in Table 5.17. First, we consider some local power estimates, specifically focusing on two mixture distributions. The first mixture is where we sample with probability w from a $W(1.2)$ distribution and with probability $(1 - w)$ from a Lomax(2) distribution. The second mixture is obtained by sampling with probability w from a $G(1.3)$ distribution and with probability $(1 - w)$ from a Lomax(2) distribution. The approximated powers of these two mixture distributions are given in Tables 5.18 and 5.19, respectively.

Table 5.17: Alternative distributions

Alternative	f(x)	Notation	Support
Power	$f(x) = \left(\frac{\theta+1}{\theta}\right)^{-\theta}\theta x^{\theta-1}$	$P(\theta)$	$(0,1)$
Generalized gamma	$f(x) = \frac{\theta x^{\theta\lambda-1}}{\gamma^{\theta\lambda}\Gamma(\lambda)}e^{-\left(\frac{x}{\gamma}\right)^{\theta}}, \quad \gamma = \frac{\Gamma(\lambda)}{\Gamma(\lambda+\frac{1}{\theta})}$	$GG(\theta,\lambda)$	$(0,\infty)$

The empirical sizes and powers for sample sizes $T = \{100, 200, 400\}$ are given in Tables 5.20 - 5.22. The entries in these tables are the percentages of 10 000 Monte Carlo samples that resulted in the rejection of the null hypothesis, which are rounded to the nearest integer. The highest power for each distribution in the results is again highlighted for the reader's comfort. Finally, we provide a brief discussion of the choice of tuning parameter, a, used to obtain the powers of the newly proposed test.

Some general conclusions regarding the results obtained in the above-mentioned tables are now presented. First, consider the approximated local powers, presented in Tables 5.18 and 5.19.

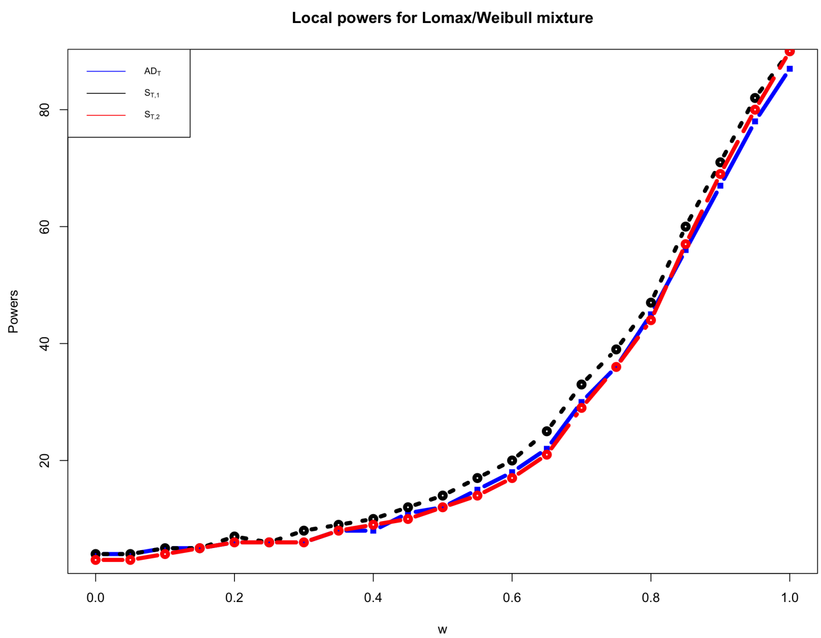


Fig. 5.1: Local powers for some of the tests over the entire range of mixture probabilities of the Lomax-Weibull mixture distribution for $T = 200$.

The results of the mixture distributions consisting of the Lomax distribution with parameter $\theta = 2$ and the Weibull distribution with parameter $\theta = 1.2$ are presented in Table 5.18. The performance of the $S_{T,0.1}$ test is unfavourable; having the lowest powers for the majority of the choices of the mixture probability,

w. The $S_{T,1}$ test outperforms all other tests as the approximated powers are generally the highest for the majority values of w considered. The AD_T, $S_{T,0.5}$ and $S_{T,2}$ tests also perform quite well, as their powers are similar to those of the $S_{T,1}$ test. Even though the $S_{T,2}$ test is undersized, it outperforms the other classical tests at low values of w. This shows higher levels of local powers compared to the other classical tests. These results can be observed in Figure 5.1, which contains two of the most favourable newly proposed tests ($S_{T,1}$ and $S_{T,2}$) and the classical test with favourable performance (AD_T).

Table 5.18: Approximated local powers for the mixture of the Lomax(2) and Weibull(1.2) distributions for various choices of the mixture parameter, w. ($T = 200$)

w	KS_T	CM_T	AD_T	$S_{T,0.1}$	$S_{T,0.5}$	$S_{T,1}$	$S_{T,2}$
0	4	**5**	4	**5**	**5**	4	3
0.05	4	4	4	**5**	4	4	3
0.1	5	5	5	5	**6**	5	4
0.15	**5**	**5**	**5**	**5**	**5**	**5**	**5**
0.2	6	6	6	5	6	**7**	6
0.25	**6**	**6**	**6**	5	**6**	**6**	**6**
0.3	6	6	6	6	7	**8**	6
0.35	7	8	8	7	8	**9**	8
0.4	7	8	8	7	9	**10**	9
0.45	9	11	11	8	11	**12**	10
0.5	10	12	12	8	13	**14**	12
0.55	12	15	15	11	16	**17**	14
0.6	15	18	18	12	**20**	**20**	17
0.65	18	22	22	15	24	**25**	21
0.7	23	29	30	20	32	**33**	29
0.75	28	35	36	25	38	**39**	36
0.8	36	44	45	30	**47**	**47**	44
0.85	45	54	56	38	58	**60**	57
0.9	54	65	67	49	69	**71**	69
0.95	65	75	78	61	80	**82**	80
1	75	85	87	70	88	**90**	**90**

The approximated local powers in Table 5.19 will now be discussed. These results are obtained using the mixture distribution consisting of the Lomax distribution with parameter $\theta = 2$ and the gamma distribution with parameter $\theta = 1.3$. Here, the KS_T, CM_T, $S_{T,0.1}$ and $S_{T,2}$ tests exhibit poor performance, as their powers are lower than the rest of the tests for the majority of mixture probabilities. We notice that the $S_{T,0.5}$ test performs the best, having the highest approximated local power for almost all values of w considered. Since the $S_{T,1}$ is the best test for a handful of mixture probabilities, we can consider this test the closest competitor for $S_{T,0.5}$. Figure 5.2 presents the local powers for some of the tests that exhibit good performance.

Table 5.19: Approximated local powers for the mixture of the Lomax(2) and gamma(1.3) distributions for various choices of the mixture parameter, w. ($T = 200$)

$\boldsymbol{w}$	$\boldsymbol{KS_T}$	$\boldsymbol{CM_T}$	$\boldsymbol{AD_T}$	$\boldsymbol{S_{T,0.1}}$	$\boldsymbol{S_{T,0.5}}$	$\boldsymbol{S_{T,1}}$	$\boldsymbol{S_{T,2}}$
0	4	**5**	4	**5**	**5**	**5**	4
0.05	4	**5**	**5**	**5**	**5**	**5**	4
0.1	**5**	**5**	**5**	**5**	**5**	**5**	4
0.15	**5**	**5**	**5**	**5**	**5**	**5**	4
0.2	5	5	5	**6**	**6**	**6**	5
0.25	5	6	6	6	**7**	**7**	6
0.3	6	6	6	6	7	**8**	6
0.35	7	8	8	7	**9**	**9**	8
0.4	7	8	8	7	**10**	**10**	8
0.45	9	10	10	8	**12**	**12**	10
0.5	9	10	11	9	**13**	12	9
0.55	11	13	13	12	**16**	15	12
0.6	12	15	16	14	**19**	17	14
0.65	15	17	20	18	**24**	21	16
0.7	17	21	24	22	**29**	28	21
0.75	23	28	31	26	**35**	34	27
0.8	27	34	37	34	**44**	41	35
0.85	36	42	46	41	**52**	49	42
0.9	43	50	55	51	**61**	59	53
0.95	52	62	66	60	**72**	70	65
1	62	70	75	71	**81**	80	76

Next, we consider the performance of the tests in general against all distributions listed in Tables 5.2 and 5.17. From Tables 5.20 - 5.22, we observe that the empirical sizes are not satisfactory (i.e., undersized) for all of the tests except the $S_{T,0.1}$ test when we use a sample size $T = 100$; however as the sample size increases the empirical sizes become adequate. We also see that we do not have similar problems as we had in Section 5.3.2 with the EACD(1,1) model, where the powers obtained were very low in some instances. In general, the powers of the KS_T test are lower for the majority of the alternatives and perform poorly in comparison to other tests, for all sample sizes. In turn, the $S_{T,0.1}$ test generally performs well for all sample sizes, while the $S_{T,0.5}$ test has the highest approximated powers for the majority of the alternatives, and outperforms the other tests.

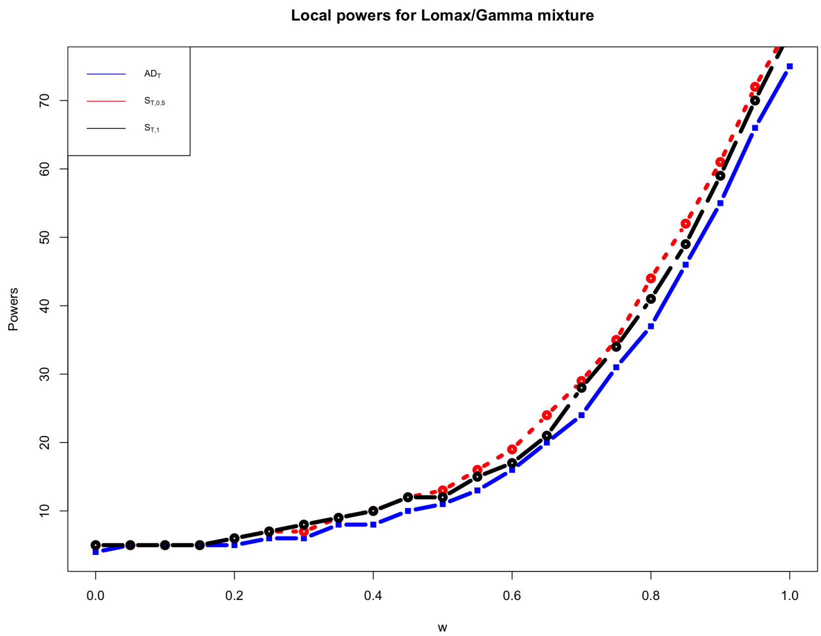


Fig. 5.2: Local powers for some of the tests over the entire range of mixture probabilities of the Lomax-gamma mixture distribution for $T = 200$.

Shifting the focus to a more in-depth analysis of the results obtained for the different alternatives, we find for the Weibull and gamma alternatives the powers grow with the shape parameter. For these alternatives, the KS_T and $S_{T,0.1}$ tests achieved relatively low powers compared to the remaining tests. $S_{T,0.5}$ and $S_{T,1}$ outperform the other tests here, achieving the highest power.

Considering the lognormal alternatives, we see that the $S_{T,2}$ test exhibits the lowest powers, while $S_{T,0.1}$ has the highest power. Furthermore, the powers decrease as the scale parameter increases.

For the power alternatives, in general, the power increases with the shape parameter for the classical tests and higher values of the tuning parameter used in the newly proposed test (i.e., $a = 1$ and $a = 2$), whereas it decreases with the shape parameter for lower values of the tuning parameter (i.e., $a = 0.1$ and

$a = 0.5$). As the sample size increases, $S_{T,0.1}$ has the lowest power for the majority of the alternatives. Another test that generally does not achieve good results is $S_{T,1}$. On the other hand, the CM_T test performs well, while the AD_T test is considered the most powerful test, as it outperforms the other tests.

Table 5.20: Tests for Lomax innovations: rejection rate for $T = 100$

	$\boldsymbol{KS_T}$	$\boldsymbol{CM_T}$	$\boldsymbol{AD_T}$	$\boldsymbol{S_{T,0.1}}$	$\boldsymbol{S_{T,0.5}}$	$\boldsymbol{S_{T,1}}$	$\boldsymbol{S_{T,2}}$
L(2)	3	3	3	**5**	3	2	3
L(3)	3	3	3	**5**	4	3	3
L(4)	3	3	3	**5**	3	2	2
W(1.1)	18	19	20	12	21	**23**	22
W(1.2)	44	53	54	36	56	**59**	57
W(1.3)	74	84	86	67	87	**89**	86
G(1.2)	21	23	24	19	**29**	28	24
G(1.3)	35	41	44	38	**50**	49	43
G(1.4)	53	61	64	57	**72**	70	63
LN(0.8)	98	98	**100**	**100**	**100**	98	89
LN(0.9)	81	82	94	**100**	96	84	53
LN(1.1)	32	35	51	**81**	55	32	9
P(0.6)	59	75	**88**	60	19	9	13
P(0.7)	79	90	**93**	18	17	36	57
P(0.8)	94	98	**99**	11	55	79	90
GG(0.8,1.5)	5	5	5	**8**	**8**	7	5
GG(0.9,1.5)	27	33	36	35	**44**	41	34
GG(0.8,1.6)	9	10	11	**15**	**15**	13	10
GG(0.9,1.6)	39	46	50	51	**60**	56	49

The results of the generalised gamma alternatives show that the power increases with both shape parameters. The powers achieved by the KS_T test are substantially lower than those of the remaining tests. Other tests that generally do not produce good results are CM_T, AD_T, and $S_{T,2}$. For the sample size $T = 100$, we find that the $S_{T,0.1}$ achieves the highest powers of all alternatives; however, as the sample increases the $S_{T,0.1}$ test also performs favourable.

Table 5.21: Tests for Lomax innovations: rejection rate for $T = 200$

	KS_T	CM_T	AD_T	$S_{T,0.1}$	$S_{T,0.5}$	$S_{T,1}$	$S_{T,2}$
L(2)	**5**	**5**	**5**	**5**	**5**	**5**	4
L(3)	**5**	**5**	**5**	6	6	**5**	4
L(4)	4	4	4	**5**	4	4	3
W(1.1)	30	33	35	25	38	**39**	38
W(1.2)	76	84	84	69	87	**89**	88
W(1.3)	97	99	99	95	99	**100**	**100**
G(1.2)	35	42	46	42	**53**	52	47
G(1.3)	61	70	75	71	**81**	80	74
G(1.4)	83	90	94	90	**95**	**95**	92
LN(0.8)	**100**	**100**	**100**	**100**	**100**	**100**	**100**
LN(0.9)	**100**	**100**	**100**	**100**	**100**	**100**	92
LN(1.1)	79	81	95	**99**	92	77	47
P(0.6)	95	99	**100**	84	35	14	26
P(0.7)	99	**100**	**100**	26	30	66	89
P(0.8)	**100**	**100**	**100**	16	86	98	**100**
GG(0.8,1.5)	6	7	8	**13**	11	9	7
GG(0.9,1.5)	46	54	62	68	**70**	66	58
GG(0.8,1.6)	13	15	19	**28**	25	20	15
GG(0.9,1.6)	70	79	85	86	**90**	87	82

Finally, we demonstrate how the selection of the tuning parameter, a, impacts the power of the recent proposed test, $S_{T,a}$. To provide a visual representation of how the power varies with different values of a, Figure 5.3 shows the powers for the $S_{200,a}$ test over a grid of a values for five alternative distributions, which also aid as a motivation for the choice of a values used in the study.

From Figure 5.3 it is clear that there is no trend for the various alternative distributions, and as a result, we chose $a = 0.1, 0.5, 1$ and 2 for the tuning parameter. Another way to obtain the value of the tuning parameter is by making use of a data-dependent choice of the parameter (See for example Allison and Santana, 2015); however, we do not implement this method.

Table 5.22: Tests for Lomax innovations: rejection rate for $T = 400$

	KS_T	CM_T	AD_T	$S_{T,0.1}$	$S_{T,0.5}$	$S_{T,1}$	$S_{T,2}$
L(2)	**5**	**5**	**5**	**5**	**5**	**5**	4
L(3)	**5**	**5**	**5**	**5**	**5**	**5**	**5**
L(4)	**5**	**5**	**5**	**5**	**5**	**5**	4
W(1.1)	50	60	63	47	67	**69**	68
W(1.2)	97	**99**	**99**	95	**99**	**99**	**99**
W(1.3)	**100**	**100**	**100**	**100**	**100**	**100**	**100**
G(1.2)	58	68	74	73	**80**	78	75
G(1.3)	89	94	97	96	**98**	**98**	97
G(1.4)	99	**100**	**100**	**100**	**100**	**100**	**100**
LN(0.8)	**100**	**100**	**100**	**100**	**100**	**100**	**100**
LN(0.9)	**100**	**100**	**100**	**100**	**100**	**100**	**100**
LN(1.1)	99	99	**100**	**100**	**100**	99	89
P(0.6)	**100**	**100**	**100**	98	71	33	53
P(0.7)	**100**	**100**	**100**	43	68	96	**100**
P(0.8)	**100**	**100**	**100**	31	**100**	**100**	**100**
GG(0.8,1.5)	10	11	15	**26**	18	14	9
GG(0.9,1.5)	80	86	93	94	**96**	94	89
GG(0.8,1.6)	23	24	36	**57**	47	35	23
GG(0.9,1.6)	96	97	99	99	**100**	99	98

In this chapter, we tested two hypotheses where the innovations are exponential and Lomax distributed respectively by means of a power study. For the hypothesis of exponential innovations, we found that the estimation procedure is accurate and the results of the power study are similar to those in Meintanis et al. (2020) except for three tests (CO_T, MP_T, EP_T) against the LN(1.1) alternative.

For the hypothesis of Lomax innovations, we observed that the estimation procedure is not sensitive to the choice of starting values, hence producing global maxima, and conclude the procedure is accurate. Focusing on the results of the power study, it is clear that the selection of the tuning parameter is important and that the recently proposed test has higher local powers compared to the classical tests. It is also possible to obtain adequate size with the new test when the classical tests are undersized for small sample sizes. Generally, the new test outperforms classical tests, except against the Power distribution, which is interestingly the only distribution in our study with a closed interval support, and therefore not a heavy-tailed distribution.

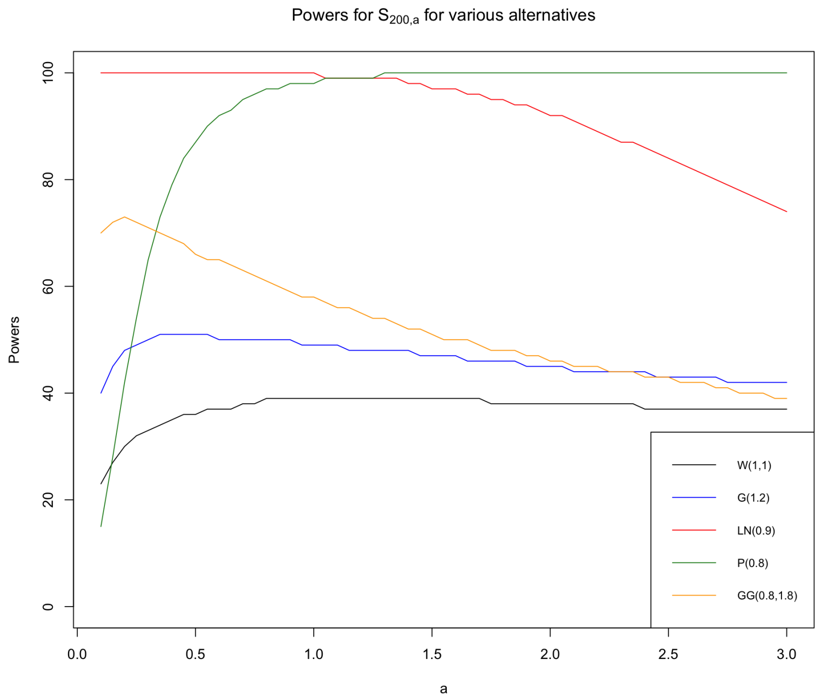


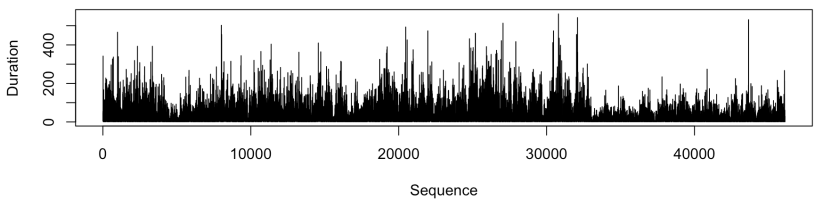
Fig. 5.3: Powers for $S_{200,a}$ for various alternatives.

In the next chapter, we will apply some of the tests considered in Chapter 4 to a real-world data set.

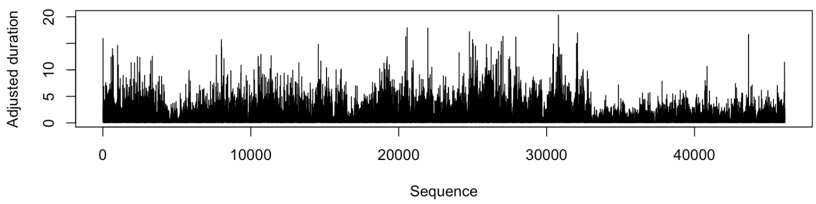
Chapter 6

Practical application

In this chapter, some of the tests considered in Chapter 4 are applied to a real-world dataset: the International Business Machines Corporation (IBM) transactions data from November 1, 1990, to January 31, 1991. The data, which was downloaded from the home page of Ruey S. Tsay, are from the Trades, Orders Reports, and Quotes (TORQ) data set constructed by Joel Hasbrouck and the New York Stock Exchange (NYSE); see Engle and Russell (1998) for a detailed description of the data. There are a total of 60 328 transactions and 63 trading days. As discussed by Engle and Russell (1998), we make two adjustments to the data. Two days were deleted from the 63 trading days (i.e. 23 November and 27 December) and we only focus on trades that occur from 10:00 a.m. to 4:00 p.m.. After these adjustments, there remain 51 365 observations of the original 60 328 transactions.



(a) Durations



(b) Adjusted durations

Fig. 6.1: IBM transactions data from 01/11/1990 to 31/1/1991.

We used the ACDm package of Belfrage (2022) in R (R Core Team, 2023) to obtain the 46 120 positive trade durations. As mentioned in Engle and Russell (1998), since intraday transactions exhibit some diurnal pattern, where the transactions are higher near the open and the close of the market, we need to focus on the diurnally adjusted durations (also called adjusted durations), by removing the daily seasonal component. If x_i is the trade duration between two consecutive events, the corresponding adjusted duration is given as

$$\tilde{x}_i = \frac{x_i}{\phi(t_i)},$$

where $\phi(t_i)$ is the daily seasonal component. Figure 6.1 presents the durations and adjusted durations, and Figure 6.2 the estimated diurnal pattern of the IBM data.

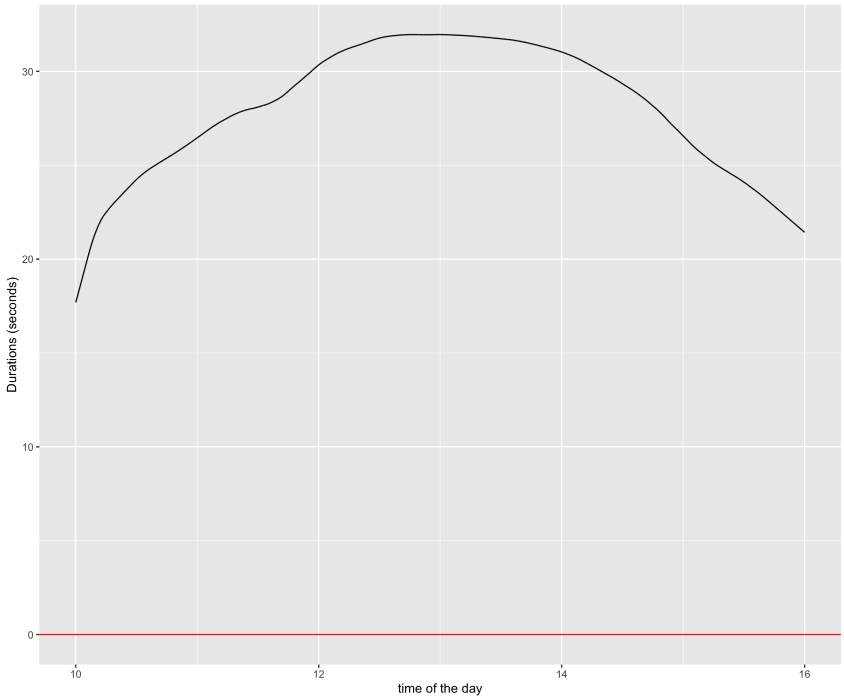
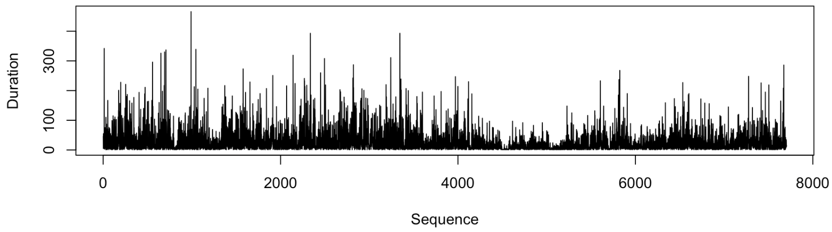
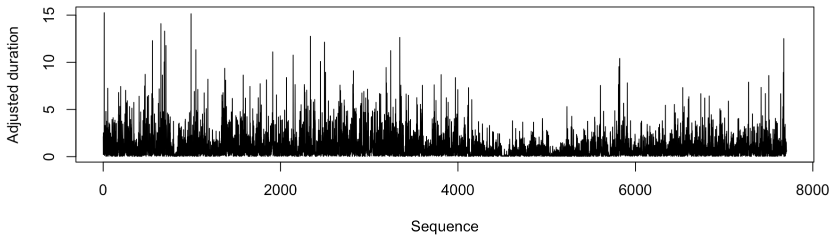


Fig. 6.2: Estimate of diurnal pattern for IBM transaction durations.

In this dissertation, we make use of a subsample of the data, considering only the positive transaction durations of IBM stock on the first 10 consecutive trading days from November 1 to November 14, 1990. Therefore we focus on 7 704 observations, where the durations and adjusted durations are given in Figure 6.3.



(a) Durations



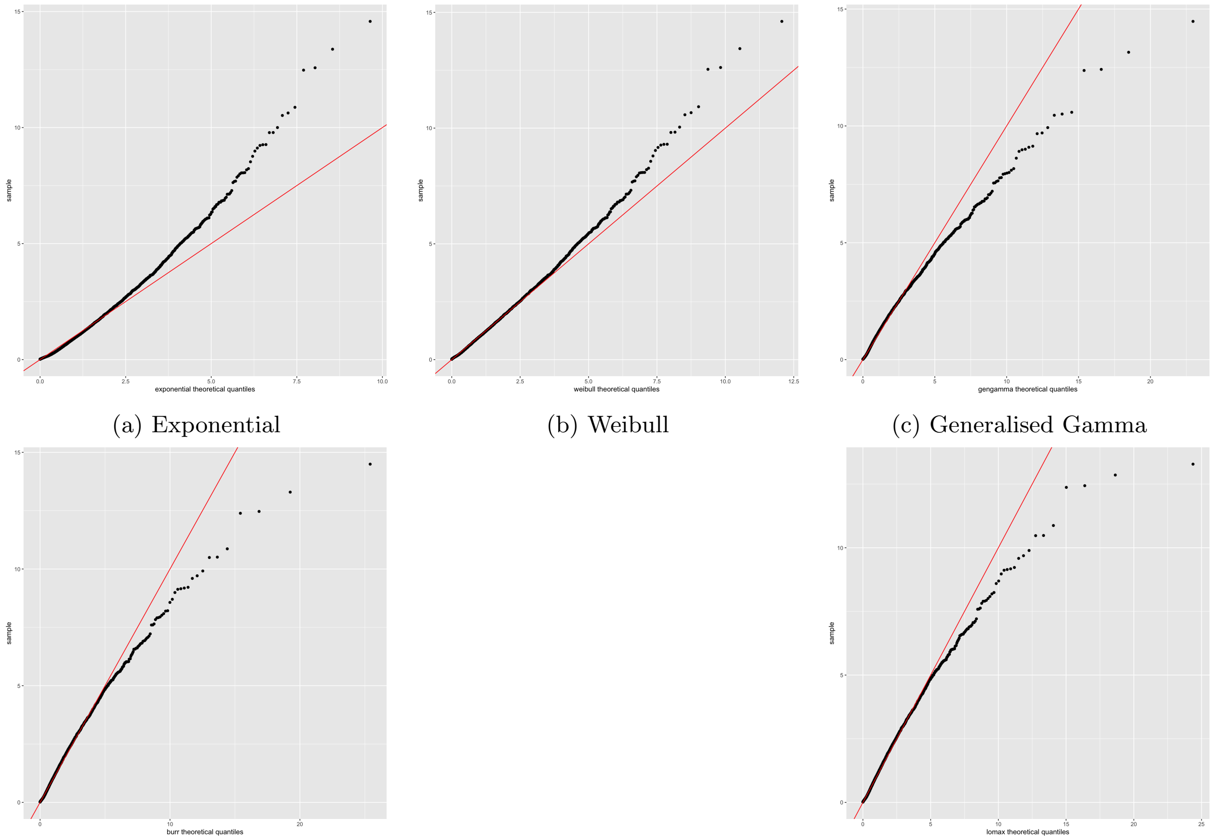
(b) Adjusted durations

Fig. 6.3: Time plots of 7 704 nonzero durations for IBM stock traded in first ten trading days of November 1990.

First, a graphical approach is used to analyse the fit of the model, by looking at the extent the fitted and the empirical distributions of the residuals differ. Quantile-quantile (QQ) plots will allow us to graphically observe whether the residuals follow a specified distribution. Figure 6.4 shows the QQ-plots for evaluating the GOF of five distributions (exponential, Weibull, generalised gamma, Burr, and Lomax) to the residuals, after fitting the ACD(1, 1) model to the IBM duration data. If a specific distribution fits the residuals well, then the QQ-plot of the residuals should be close to a linear line. The QQ-plots in all the panels (6.4a)–(6.4e) in Figure 6.4 show systematic departures from the straight line, where the exponential distribution in Figure 6.4a shows the biggest departure. The QQ-plots indicate that not one of the distributions we considered, fits the data.

Table 6.1: The p-values for GOF tests of different families of parametric distributions for the error term of the ACD(1,1) model for IBM trade durations.

	KS	**CM**	**AD**	$\mathbf{S_{T,0.5}}$
Exponential	0	0	0	–
Weibull	0	0	0	–
Generalized gamma	0.807	0.8855	0	–
Burr	0	0	0	–
Lomax	0	0	0	0



(a) Exponential (b) Weibull (c) Generalised Gamma

(d) Burr (e) Lomax

Fig. 6.4: QQ-plots for evaluating the GOF after fitting the ACD(1,1) model with five different innovations to IBM transaction data.

In Table 6.1 we present the p-values for the classical GOF tests and the newly proposed test we considered in Chapter 4, for testing the hypothesis that the innovations follow a specific distribution (i.e. exponential, Weibull, generalised gamma, Burr or Lomax). These p-values were obtained using the same bootstrap technique described in Chapter 5. All the tests rejected ACD(1,1) models with exponential, Weibull, Burr and Lomax innovations, respectively. These results are consistent with the results from the QQ-plots. For a generalised gamma distribution, we see that the KS_T and CM_T tests have a p-value of 0.807 and 0.8855 respectively, which contradicts our findings from the QQ-plots. However, for the AD_T test, the p-value is 0 which is again consistent with the findings from the QQ-plot. A possible reason why the AD_T test is the only test among the classical tests that correctly reject the hypothesis of generalised gamma innovations, is that this test places more weight on observations in the tails of a distribution compared to the other classical tests.

Chapter 7

Conclusion

In this dissertation, we presented specification tests for the innovation distribution in ACD models, with specific emphasis on ACD(1,1) models with exponential and Lomax innovations. The primary objectives of this dissertation, with their corresponding outcomes, can be summarised as follows:

- Introduce ACD models and its theoretical framework.

  The ACD model together with its theoretical framework, was presented in Chapter 2.

- Provide an overview of the Lomax distribution and its use in the ACD context.

  The overview of the Lomax distribution was done in Chapter 3 where we also introduced the ACD(1,1) model with Lomax innovations, also known as LACD(1,1).

- Provide a discussion of existing GOF tests for the innovation distribution in ACD models.

  Some existing GOF tests were given in Chapter 4.

- Provide a newly proposed test for testing Lomax distributed innovations in ACD models.

  This was done at the end of Chapter 4. This new test is based on a characterisation of the Lomax law using Stein's method

- Use a Monte Carlo study to test for exponential innovations, and obtain similar results as those in Meintanis et al. (2020).

  Testing the hypothesis of exponential innovations, the results of the study in 5 is similar as those in Meintanis et al. (2020), except for the $\text{LN}(1,1)$ alternative.

- Use a Monte Carlo study to analyse the finite sample performance of the newly proposed test.

  The results of the Monte Carlo study in Chapter 5 show that the newly proposed test is competitive, outperforming the classical tests against the majority of the alternatives considered. It is also possible to obtain adequate size with the new test when the classical tests are undersized for small sample sizes.

- Apply the newly proposed test with other existing tests to real-world data to assess the new test's competitiveness.

In Chapter 6 the QQ-plot of the residuals from an ACD(1,1) model with Lomax innovations, showed that the model did not fit the data well, whereupon the newly proposed test accurately confirmed this.

For possible future research, the new asymptotic theory, introduced by Cavaliere et al. (2023), for the parameter estimation of ACD models can be included in the bootstrap procedure. From Chapter 5 we saw that the choice of tuning parameter in the newly proposed test is important, therefore a data-dependent choice to obtain the tuning parameter can be considered. Finally, more recent data can be used to test the performance of the newly proposed test.

Bibliography

Abdul-Moniem, I. B. (2012). Recurrence relations for moments of lower generalized order statistics from exponentiated Lomax distribution and its characterization. *J. Math. Comput. Sci.*, 2(4):999–1011.

Al-Jarallah, R., Ghitany, M., and Gupta, R. C. (2014). A proportional hazard Marshall–Olkin extended family of distributions and its application to Gompertz distribution. *Communications in Statistics-Theory and Methods*, 43(20):4428–4443.

Allison, J. and Santana, L. (2015). On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, 85(16):3276–3288.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American statistical association*, 49(268):765–769.

Atkinson, A. B. and Harrison, A. J. (1978). Distribution of personal wealth in Britain. *(Cambridge University Press).*

Bauwens, L. and Giot, P. (2000). The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et de Statistique*, pages 117–149.

Bauwens, L. and Giot, P. (2001). *Econometric modelling of stock market intraday activity*, volume 38. Springer Science & Business Media.

Bauwens, L. and Giot, P. (2003). Asymmetric ACD models: introducing price information in ACD models. *Empirical Economics*, 28:709–731.

Belfrage, M. (2022). *ACDm: Tools for Autoregressive Conditional Duration Models.* R package version 1.0.4.2.

Bhogal, S. K. and Thekke Variyam, R. (2019). Conditional duration models for high-frequency data: a review on recent developments. *Journal of Economic Surveys*, 33(1):252–273.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Bryson, M. C. (1974). Heavy-tailed distributions: properties and tests. *Technometrics*, 16(1):61–68.

Cavaliere, G., Mikosch, T., Rahbek, A., and Vilandt, F. (2022). The Econometrics of Financial Duration Modeling. *J.Econometrics, forthcoming.*

Cavaliere, G., Mikosch, T., Rahbek, A., and Vilandt, F. (2023). Asymptotics for the Generalized Autoregressive Conditional Duration Model. *arXiv preprint arXiv:2307.01779.*

Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, 53(12):4433–4440.

Corbellini, A., Crosato, L., Ganugi, P., and Mazzoli, M. (2010). Fitting Pareto II distributions on firm size: Statistical methodology and economic puzzles. *Advances in Data Analysis: Theory and Applications to Reliability and Inference, Data Mining, Bioinformatics, Lifetime Data, and Neural Networks*, pages 321–328.

Cosma, A. and Galli, F. (2006). A nonparametric ACD model. CORE Discussion Paper No. 2006/67.

Cox, D. R. and Oakes, D. (1984). *Analysis of survival data*, volume 21. CRC press.

D'Agostino, R. (2017). *Goodness-of-fit-techniques.* Routledge.

De Luca, G. and Zuccolotto, P. (2004). Detecting features of different financial durations through the Pareto distribution. *Quaderni di Statistica*, 6:71–86.

Drost, F. C. and Werker, B. J. M. (2004). Semiparametric duration models. *Journal of Business & Economic Statistics*, 22(1):40–50.

Ebner, B., Fischer, A., Henze, N., and Mayer, C. (2022). Weibull or not Weibull? *arXiv preprint arXiv:2206.06643.*

Ebner, B., Grobler, G. L., and Smuts, M. (2023). A goodness-of-fit test for the Lomax distribution. Manuscript in preparation.

Engle, R. F. (2000). The econometrics of ultra-high-frequency data. *Econometrica*, 68(1):1–22.

Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, pages 1127–1162.

Epps, T. and Pulley, L. (1986). A test of exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(2):206–213.

Fatima, K., Jan, U., and Ahmad, S. (2018). Statistical Properties of Rayleigh Lomax distribution with applications in Survival Analysis. *Journal of Data Science*, 16(3).

Feigin, P. and Heathcote, C. (1976). The empirical characteristic function and the cramer-von mises statistic. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 309–325.

Fernandes, M. and Grammig, J. (2006). A family of autoregressive conditional duration models. *Journal of Econometrics*, 130(1):1–23.

Giacomini, R., Politis, D. N., and White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric theory*, 29(3):567–589.

Grammig, J. and Maurer, K.-O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *The Econometrics Journal*, 3(1):16–38.

Gürtler, N. and Henze, N. (2000). Goodness-of-fit tests for the Cauchy distribution based on the empirical characteristic function. *Annals of the Institute of Statistical Mathematics*, 52:267–286.

Harris, C. M. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, 16(2):307–313.

Hassan, A. and Al-Ghamdi, A. (2009). Optimum Step Stress Accelerated Life Testing for Lomax Distribution. *Journal of Applied Sciences Research*, 5.

Heathcote, C. (1972). A test of goodness of fit for symmetric random variables. *Australian Journal of Statistics*, 14(2):172–181.

Henze, N., Klar, B., and Meintanis, S. G. (2003). Invariant tests for symmetry about an unspecified point based on the empirical characteristic function. *Journal of Multivariate Analysis*, 87(2):275–297.

Holland, O., Golaup, A., and Aghvami, A. (2006). Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. *IEE Proceedings-communications*, 153(5):683–690.

Kankainen, A. and Ushakov, N. G. (1998). A consistent modification of a test for independence based on the empirical characteristic function. *Journal of Mathematical Sciences*, 89:1486–1494.

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att*, 4:89–91.

Koutrouvelis, I. A. and Meintanis, S. G. (1999). Testing for stability based on the empirical characteristic function with applications to financial data. *Journal of Statistical Computation and Simulation*, 64(4):275–300.

Lee, S. and Oh, H. (2015). Entropy test and residual empirical process for autoregressive conditional duration models. *Computational Statistics & Data Analysis*, 86:1–12.

Lee, S.-W. and Hansen, B. E. (1994). Asymptotic theory for the GARCH(1, 1) quasi-maximum likelihood estimator. *Econometric theory*, 10(1):29–52.

Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American statistical association*, 49(268):847–852.

Mathew, J. and Chesneau, C. (2020). Some new contributions on the Marshall–Olkin length biased Lomax distribution: Theory, modelling and data analysis. *Mathematical and Computational Applications*, 25(4):79.

Meintanis, S. G., Milošević, B., and Obradović, M. (2020). Goodness-of-fit tests in conditional duration models. *Statistical Papers*, 61:123–140.

Milošević, B. and Obradović, M. (2016). New class of exponentiality tests based on U-empirical Laplace transform. *Statistical Papers*, 57(4):977–990.

Nagarjuna, V. B., Vardhan, R. V., and Chesneau, C. (2021). Kumaraswamy generalized power Lomax distribution and its applications. *Stats*, 4(1):28–45.

Nagarjuna, V. B., Vardhan, R. V., and Chesneau, C. (2022). Nadarajah–haghighi lomax distribution and its applications. *Mathematical and Computational Applications*, 27(2):30.

Nombebe, T., Allison, J., Santana, L., and Visagie, J. (2023). On fitting the Lomax distribution: a comparison between minimum distance estimators and other estimation techniques. *Computation*, 11(3):44.

Pacurar, M. (2008). Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Surveys*, 22(4):711–751.

Pareto, V. (1897). The new theories of economics. *Journal of political economy*, 5(4):485–502.

Puri, P. S. and Rubin, H. (1970). A characterization based on the absolute difference of two iid random variables. *The Annals of Mathematical Statistics*, 41(6):2113–2122.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rady, E.-H. A., Hassanein, W., and Elhaddad, T. (2016). The power Lomax distribution with an application to bladder cancer data. *SpringerPlus*, 5:1–22.

Saart, P. W., Gao, J., and Allen, D. E. (2015). Semiparametric autoregressive conditional duration model: Theory and practice. *Econometric Reviews*, 34(6-10):849–881.

Smit, N. (2018). An objective comparison between various goodness-of-fit tests for exponentiality. Master's thesis, North-West University.

Snyder, D. L. and Miller, M. I. (1991). Random point processes in time and space. *Springer Texts in Electrical Engineering*.

Stein, C., Diaconis, P., Holmes, S., and Reinert, G. (2004). Use of exchangeable pairs in the analysis of simulations. *Lecture Notes-Monograph Series*, pages 1–26.

Tahir, M. H., Cordeiro, G. M., Mansoor, M., and Zubair, M. (2015). The Weibull-Lomax distribution: properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 44(2):455–474.

Tsay, R. S. (2010). Analysis of Financial Time Series. *Wiley Series in Probability and Statistics*.

Ushakov, N. (1999). Selected topics in characteristic functions. modern probability and statistics. *VSP, Utrecht*.

Yan, X. (2021). Autoregressive conditional duration modelling of high frequency data. *arXiv preprint arXiv:2111.02300*.

Zhang, M. Y., Russell, J. R., and Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, 104(1):179–207.

Zuccolotto, P. et al. (2002). Modelling the impact of open volume on inter-trade autoregressive durations. *Metron*, 60(3-4):51–65.

Appendix A

Closed-form expression derivations

A.1 Anderson-Darling test

We prove the closed-form expression of the Anderson-Darling test for the uniform distribution case. The test statistic can be written as follows:

$$\begin{aligned}
AD_T &= T \int_0^1 \frac{\left(\widehat{F}_T(x) - x\right)^2}{x(1-x)} dx \\
&= T \sum_{t=0}^{T} \int_{X_{(t)}}^{X_{(t+1)}} \frac{\left(\frac{t}{T} - x\right)^2}{x(1-x)} dx \qquad \left(X_{(0)} = 0, X_{(t+1)} = 1\right) \\
&= T \sum_{t=0}^{T} \int_{X_{(t)}}^{X_{(t+1)}} \frac{\left(\left(\frac{t}{T}\right)^2 - 2\left(\frac{t}{T}\right)x + x^2\right)}{x(1-x)} dx \\
&= T \sum_{t=0}^{T} \left[\int_{X_{(t)}}^{X_{(t+1)}} \frac{(\frac{t}{T})^2}{x(1-x)} dx + \int_{X_{(t)}}^{X_{(t+1)}} \frac{-2(\frac{t}{T})x}{x(1-x)} dx + \int_{X_{(t)}}^{X_{(t+1)}} \frac{x^2}{x(1-x)} dx \right].
\end{aligned}$$

But one can show that

$$\begin{aligned}
\int_{X_{(t)}}^{X_{(t+1)}} \frac{(\frac{t}{T})^2}{x(1-x)} dx &= \left(\frac{t}{T}\right)^2 \int_{X_{(t)}}^{X_{(t+1)}} \frac{1}{x(1-x)} dx \\
&= \left(\frac{t}{T}\right)^2 \int_{X_{(t)}}^{X_{(t+1)}} \left(\frac{1}{x} + \frac{1}{1-x}\right) dx \\
&= \left(\frac{t}{T}\right)^2 \left[[\ln(x)]_{X_{(t)}}^{X_{(t+1)}} - [1 - \ln(x)]_{X_{(t)}}^{X_{(t+1)}} \right] \\
&= \left(\frac{t}{T}\right)^2 \left[\ln\left(X_{(t+1)}\right) - \ln\left(X_{(t)}\right) - \ln\left(1 - X_{(t+1)}\right) + \ln\left(1 - X_{(t)}\right)\right],
\end{aligned}$$

$$\begin{aligned}
\int_{X_{(t)}}^{X_{(t+1)}} \frac{-2(\frac{t}{T})x}{x(1-x)} dx &= 2\left(\frac{t}{T}\right) \int_{X_{(t)}}^{X_{(t+1)}} \frac{-1}{(1-x)} dx \\
&= 2\left(\frac{t}{T}\right) \left[[1 - \ln(x)]_{X_{(t)}}^{X_{(t+1)}} \right] \\
&= 2\left(\frac{t}{T}\right) \left[\ln\left(1 - X_{(t+1)}\right) - \ln\left(1 - X_{(t)}\right)\right],
\end{aligned}$$

and

$$\int_{X_{(t)}}^{X_{(t+1)}} \frac{x^2}{x(1-x)} dx = \int_{X_{(t)}}^{X_{(t+1)}} \frac{x}{(1-x)} dx$$

$$\text{Let } u = 1-x \implies du = -dx$$

$$\begin{aligned}
\therefore \int_{X_{(t)}}^{X_{(t+1)}} \frac{x}{(1-x)} dx &= \int_{1-X_{(t)}}^{1-X_{(t+1)}} \frac{u-1}{u} du \\
&= \int_{1-X_{(t)}}^{1-X_{(t+1)}} 1 du - \int_{1-X_{(t)}}^{1-X_{(t+1)}} \frac{1}{u} du \\
&= [u]_{1-X_{(t)}}^{1-X_{(t+1)}} - [\ln(u)]_{1-X_{(t)}}^{1-X_{(t+1)}} \\
&= \left[1 - X_{(t+1)} - 1 + X_{(t)}\right] - \left[\ln\left(1 - X_{(t+1)}\right) - \ln\left(1 - X_{(t)}\right)\right] \\
&= X_{(t)} - X_{(t+1)} + \ln\left(1 - X_{(t)}\right) - \ln\left(1 - X_{(t+1)}\right).
\end{aligned}$$

Substituting back into AD_T we have

$$\begin{aligned}
AD_T = T \sum_{t=0}^{T} \Bigg[&\left(\frac{t}{T}\right)^2 \left[\ln\left(X_{(t+1)}\right) - \ln\left(X_{(t)}\right) - \ln\left(1 - X_{(t+1)}\right) + \ln\left(1 - X_{(t)}\right)\right] \\
&+ 2\left(\frac{t}{T}\right)\left[\ln\left(1 - X_{(t+1)}\right) - \ln\left(1 - X_{(t)}\right)\right] \\
&+ X_{(t)} - X_{(t+1)} + \ln\left(1 - X_{(t)}\right) - \ln\left(1 - X_{(t+1)}\right)\Bigg] \\
= T \sum_{t=0}^{T} \Bigg[&\left(\frac{t^2}{T^2}\right)\ln\left(X_{(t+1)}\right) - \left(\frac{t^2}{T^2}\right)\ln\left(X_{(t)}\right) - \left(\frac{t^2}{T^2}\right)\ln\left(1 - X_{(t+1)}\right) + \left(\frac{t^2}{T^2}\right)\ln\left(1 - X_{(t)}\right) \\
&+ \left(\frac{2t}{T}\right)\ln\left(1 - X_{(t+1)}\right) - \left(\frac{2t}{T}\right)\ln\left(1 - X_{(t)}\right) \\
&+ X_{(t)} - X_{(t+1)} + \ln\left(1 - X_{(t)}\right) - \ln\left(1 - X_{(t+1)}\right)\Bigg] \\
= T \sum_{t=0}^{T} \Bigg[&X_{(t)} - \frac{t^2}{T^2}\ln\left(X_{(t)}\right) + \left(1 + \frac{t^2}{T^2} - \frac{2t}{T}\right)\ln\left(1 - X_{(t)}\right) \\
&- X_{(t+1)} + \frac{t^2}{T^2}\ln\left(X_{(t+1)}\right) + \left(\frac{2t}{T} - \frac{t^2}{T^2} - 1\right)\ln\left(1 - X_{(t+1)}\right)\Bigg]
\end{aligned}$$

$$
\begin{aligned}
= T\Bigg[& X_{(0)} - \frac{0^2}{T^2}\ln\left(X_{(0)}\right) + \left(1 + \frac{0^2}{T^2} - \frac{2(0)}{T}\right)\ln\left(1 - X_{(0)}\right) \\
& - X_{(1)} + \frac{0^2}{T^2}\ln\left(X_{(1)}\right) + \left(\frac{2(0)}{T} - \frac{0^2}{T^2} - 1\right)\ln\left(1 - X_{(1)}\right) \\
& + X_{(1)} - \frac{1^2}{T^2}\ln\left(X_{(1)}\right) + \left(1 + \frac{1^2}{T^2} - \frac{2(1)}{T}\right)\ln\left(1 - X_{(1)}\right) \\
& - X_{(2)} + \frac{1^2}{T^2}\ln\left(X_{(2)}\right) + \left(\frac{2(1)}{T} - \frac{1^2}{T^2} - 1\right)\ln\left(1 - X_{(2)}\right) \\
& + X_{(2)} - \frac{2^2}{T^2}\ln\left(X_{(2)}\right) + \left(1 + \frac{2^2}{T^2} - \frac{2(2)}{T}\right)\ln\left(1 - X_{(2)}\right) \\
& - X_{(3)} + \frac{2^2}{T^2}\ln\left(X_{(3)}\right) + \left(\frac{2(2)}{T} - \frac{2^2}{T^2} - 1\right)\ln\left(1 - X_{(3)}\right) \\
& \vdots \\
& + X_{(t)} - \frac{t^2}{T^2}\ln\left(X_{(t)}\right) + \left(1 + \frac{t^2}{T^2} - \frac{2(t)}{T}\right)\ln\left(1 - X_{(t)}\right) \\
& - X_{(t+1)} + \frac{t^2}{T^2}\ln\left(X_{(t+1)}\right) + \left(\frac{2(t)}{T} - \frac{t^2}{T^2} - 1\right)\ln\left(1 - X_{(t+1)}\right)\Bigg] \\
= T & \sum_{t=1}^{T}\left[\left(\frac{1-2t}{T^2}\right)\ln\left(X_{(t)}\right) + \left(\frac{-2T-1+2t}{T^2}\right)\ln\left(1 - X_{(t)}\right)\right] + T\left[-1+0+0\right] \\
= & \sum_{t=1}^{T}\left[\left(\frac{1-2t}{T}\right)\ln\left(X_{(t)}\right) + \left(\frac{-2T-1+2t}{T}\right)\ln\left(1 - X_{(t)}\right)\right] - T \\
= & -T - \frac{1}{T}\sum_{t=1}^{T}\left[(2t-1)\ln\left(X_{(t)}\right) + (2T+1-2t)\ln\left(1 - X_{(t)}\right)\right].
\end{aligned}
$$

A.2 A test based on a characterisation of the Lomax law using Stein's method

$$
\begin{aligned}
S_{T,a} &= T\int_0^\infty \left\{\frac{1}{T}\sum_{j=1}^{T}\left[t+(\widehat{\theta}+2)((\widehat{\theta}+\widehat{\varepsilon}_j)^{-1}\right]e^{-t\widehat{\varepsilon}_j}-\frac{\widehat{\theta}+1}{\widehat{\theta}}\right\}^2 e^{-at}dt\\
&= T\int_0^\infty \frac{1}{T^2}\sum_{i,j=1}^{T} t^2 e^{-(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)t}dt\\
&\quad + T\int_0^\infty \frac{2}{T^2}\sum_{i,j=1}^{T}(\widehat{\theta}+2)(\widehat{\theta}+\widehat{\varepsilon}_i)^{-1}te^{-(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)t}dt\\
&\quad + T\int_0^\infty \frac{1}{T^2}\sum_{i,j=1}^{T}(\widehat{\theta}+2)^2(\widehat{\theta}+\widehat{\varepsilon}_i)^{-1}(\widehat{\theta}+\widehat{\varepsilon}_j)^{-1}e^{-(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)t}dt\\
&\quad - T\int_0^\infty \frac{2}{T}\left(\frac{\widehat{\theta}+1}{\widehat{\theta}}\right)\sum_{i=1}^{T}\left[te^{-(\widehat{\varepsilon}_i+a)t}+(\widehat{\theta}+2)(\widehat{\theta}+\widehat{\varepsilon}_i)^{-1}te^{-(\widehat{\varepsilon}_i+a)t}\right]dt\\
&\quad + T\int_0^\infty \frac{(\widehat{\theta}+1)^2}{\widehat{\theta}^2}e^{-at}dt\\
&= \frac{1}{T}\sum_{i,j=1}^{T}\left\{\int_0^\infty t^2e^{-(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)t}dt+\frac{2(\widehat{\theta}+2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)}\int_0^\infty te^{-(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)t}dt+\frac{(\widehat{\theta}+2)^2}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\theta}+\widehat{\varepsilon}_j)}\int_0^\infty e^{-(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)t}dt\right\}\\
&\quad - 2\left(\frac{\widehat{\theta}+1}{\widehat{\theta}}\right)\sum_{i=1}^{T}\left\{\int_0^\infty te^{-(\widehat{\varepsilon}_i+a)t}dt+\frac{(\widehat{\theta}+2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)}\int_0^\infty e^{-(\widehat{\varepsilon}_i+a)t}dt\right\}+\frac{T(\widehat{\theta}+1)^2}{\widehat{\theta}^2}\int_0^\infty e^{-at}dt.
\end{aligned}
$$

But recall that if $X \sim Gamma(\alpha,\beta)$ then

$$
\begin{aligned}
\int_0^\infty f(x)dx = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx &= 1, \qquad x\geq 0\\
\therefore \int_0^\infty x^{\alpha-1}e^{-\beta x}dx &= \frac{\Gamma(\alpha)}{\beta^\alpha},
\end{aligned}
$$

where $\Gamma(\alpha)=(\alpha-1)!$.

Using the above, the test statistic becomes

$$
\begin{aligned}
S_{T,a} &= \frac{1}{T}\sum_{i,j=1}^{T}\left\{\frac{\Gamma(3)}{(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)^3}+\frac{2(\widehat{\theta}+2)\Gamma(2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)^2}+\frac{(\widehat{\theta}+2)^2\Gamma(1)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\theta}+\widehat{\varepsilon}_j)(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)}\right\}\\
&\quad - \frac{2(\widehat{\theta}+1)}{\widehat{\theta}}\sum_{i=1}^{T}\left\{\frac{\Gamma(2)}{(\widehat{\varepsilon}_i+a)^2}+\frac{(\widehat{\theta}+2)\Gamma(1)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\varepsilon}_i+a)}\right\}+\frac{T(\widehat{\theta}+1)^2\Gamma(1)}{\widehat{\theta}^2 a}\\
&= \frac{1}{T}\sum_{i,j=1}^{T}\left\{\frac{2}{(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)^3}+\frac{2(\widehat{\theta}+2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)^2}+\frac{(\widehat{\theta}+2)^2}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\theta}+\widehat{\varepsilon}_j)(\widehat{\varepsilon}_i+\widehat{\varepsilon}_j+a)}\right\}\\
&\quad - \frac{2(\widehat{\theta}+1)}{\widehat{\theta}}\sum_{i=1}^{T}\left\{\frac{1}{(\widehat{\varepsilon}_i+a)^2}+\frac{(\widehat{\theta}+2)}{(\widehat{\theta}+\widehat{\varepsilon}_i)(\widehat{\varepsilon}_i+a)}\right\}+\frac{T(\widehat{\theta}+1)^2}{\widehat{\theta}^2 a}.
\end{aligned}
$$

Appendix B

Sensitivity of MLEs: Extra results for true and alternative distributions

Lomax distribution

Table B.1: Estimates of LACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_1}} = (0.5, 0.49, 0.3, 4)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.660	0.417	0.290	4.932
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.660	0.417	0.290	4.931
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.660	0.417	0.290	4.934
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.660	0.417	0.290	4.930
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.660	0.417	0.290	4.933
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.660	0.417	0.290	4.929
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.660	0.417	0.290	4.929
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.660	0.417	0.290	4.932
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.660	0.417	0.290	4.927
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.660	0.417	0.290	4.931

Table B.2: Estimates of LACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_3}} = (4, 0.69, 0.2, 12)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	5.525	0.530	0.174	22.779
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	5.529	0.530	0.174	22.903
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	5.532	0.531	0.174	22.903
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	5.529	0.531	0.174	22.859
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	5.529	0.530	0.174	22.785
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	5.527	0.530	0.174	22.843
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	5.529	0.530	0.174	22.768
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	5.532	0.530	0.174	22.853
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	5.531	0.530	0.174	22.963
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	5.529	0.530	0.174	22.924

Table B.3: Estimates of LACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_4}} = (2.5, 0.1, 0.7, 5.5)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.505	0.050	0.908	4.596
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.506	0.050	0.908	4.592
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.504	0.050	0.908	4.593
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.506	0.051	0.908	4.592
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.505	0.051	0.908	4.589
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.506	0.050	0.908	4.591
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.504	0.050	0.908	4.598
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.506	0.051	0.908	4.596
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.506	0.051	0.908	4.587
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	0.506	0.050	0.908	4.590

Table B.4: Estimates of LACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_5}} = (4.2, 0.6, 0.2, 2)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	4.150	0.686	0.233	2.364
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	4.148	0.686	0.233	2.366
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	4.149	0.686	0.233	2.364
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	4.149	0.686	0.233	2.365
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	4.149	0.686	0.233	2.365
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	4.151	0.686	0.233	2.366
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	4.150	0.685	0.233	2.364
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	4.148	0.686	0.233	2.367
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	4.148	0.686	0.233	2.364
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_{10}}}$	4.150	0.686	0.233	2.364

Exponential distribution

Table B.5: Estimates of EACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_1}} = (0.5, 0.49, 0.3, 4)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.208	0.093	4.88E-08	6371.646
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.208	0.093	1.15E-07	6383.410
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.208	0.093	4.87E-07	6362.403
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.208	0.093	1.03E-07	6397.058
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.208	0.093	3.53E-07	6377.489
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.208	0.092	1.72E-07	6375.056
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.208	0.093	2.12E-07	6289.900
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.208	0.093	2.12E-07	6381.182
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.208	0.093	3.48E-07	6382.055
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	0.208	0.093	4.67E-07	6394.475

Table B.6: Estimates of EACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_3}} = (4, 0.69, 0.2, 12)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.748	0.037	3.89E-07	5138.186
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.748	0.037	3.05E-06	5138.532
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.252	0.031	0.645	4903.198
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.748	0.037	2.49E-06	5104.707
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.252	0.031	0.645	4952.122
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.251	0.031	0.645	4920.514
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.252	0.031	0.645	4951.758
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.748	0.037	2.88E-06	5136.798
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.252	0.031	0.644	4942.791
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	0.251	0.031	0.645	4950.164

Table B.7: Estimates of EACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_4}} = (2.5, 0.1, 0.7, 5.5)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	1.572	4.02E-08	0.031	4380.490
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.024	1.06E-08	0.985	4194.443
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.000	1.39E-10	1.000	3099.117
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	1.590	4.76E-08	0.020	4403.278
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.009	3.41E-10	0.995	4086.511
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.264	3.80E-08	0.837	4303.519
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.062	2.55E-08	0.962	4265.667
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	1.537	9.24E-09	0.052	4365.450
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.162	8.03E-09	0.900	4322.297
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	1.618	4.97E-08	0.003	4411.531

Table B.8: Estimates of EACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_5}} = (4.2, 0.6, 0.2, 2)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	2.992	0.240	0.030	2547.554
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	2.991	0.240	0.030	2546.937
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	2.991	0.240	0.030	2560.675
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	2.990	0.240	0.030	2561.819
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	2.991	0.240	0.030	2558.812
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	2.991	0.240	0.030	2580.895
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	2.990	0.240	0.030	2554.836
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	2.990	0.240	0.030	2553.236
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	2.991	0.240	0.030	2547.683
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	2.992	0.240	0.030	2567.526

Weibull distribution

Table B.9: Estimates of WACD(1,1), with true parameters $\boldsymbol{\vartheta_{0_2}} = (1.2, 0.25, 0.65, 7)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.872	0.264	0.646	18789.475
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.877	0.264	0.646	18545.680
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.875	0.264	0.646	18863.485
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.000	2.45E-06	1.000	12773.928
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.871	0.263	0.647	18766.549
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.875	0.264	0.646	18661.424
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.874	0.264	0.647	18692.660
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.872	0.264	0.647	18793.853
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.874	0.264	0.646	18885.532
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	0.872	0.263	0.647	18808.819

Table B.10: Estimates of WACD(1,1), with true parameters $\boldsymbol{\vartheta_{0_3}} = (4, 0.69, 0.2, 12)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	7.938	0.677	0.258	18282.136
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	7.947	0.676	0.259	18965.005
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	7.926	0.678	0.257	18845.414
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	7.960	0.678	0.256	18957.954
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	7.997	0.675	0.259	19229.858
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	7.880	0.676	0.260	18723.594
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	7.914	0.677	0.258	19000.702
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	7.893	0.676	0.259	19014.303
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	7.844	0.676	0.259	18919.921
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	7.878	0.676	0.260	18977.934

Table B.11: Estimates of WACD(1,1), with true parameters $\boldsymbol{\vartheta_{0_4}} = (2.5, 0.1, 0.7, 5.5)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	1.173	0.100	0.796	18602.109
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	1.156	0.099	0.798	18542.002
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	1.155	0.099	0.798	18766.296
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	1.161	0.100	0.798	19151.008
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	1.164	0.100	0.797	18105.027
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	1.182	0.100	0.796	18848.344
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	1.161	0.099	0.798	18422.263
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	1.162	0.099	0.798	18773.937
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	1.166	0.099	0.797	18587.193
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	3.00E-06	9.92E-08	1.000	12780.560

Table B.12: Estimates of WACD(1,1), with true parameters $\boldsymbol{\vartheta_{0_5}} = (4.2, 0.6, 0.2, 2)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	3.575	0.637	0.128	16717.542
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	3.576	0.637	0.128	16334.960
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	3.573	0.637	0.128	16403.213
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	3.573	0.637	0.128	16457.138
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	3.572	0.637	0.128	16720.669
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	3.575	0.637	0.128	16806.302
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	3.574	0.637	0.128	16394.458
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	3.571	0.637	0.128	16803.965
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	3.574	0.637	0.128	16443.961
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	3.572	0.637	0.128	16455.998

Gamma distribution

Table B.13: Estimates of GACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_2}} = (1.2, 0.25, 0.65, 7)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	8.075	1.000	0.000	7.820
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	7.474	0.994	0.006	7.618
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	7.770	0.997	0.003	7.717
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	1688.921	0.996	0.004	1744.177
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	7.432	0.992	0.008	7.584
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	7.935	0.999	0.001	7.775
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	7.323	0.985	0.015	7.534
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	182.159	0.966	0.034	48.614
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	9.018	0.992	0.008	8.293
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	159.696	0.992	0.008	46.822

Table B.14: Estimates of GACD(1,1), with adjusted actual parameters $\boldsymbol{\vartheta_{0_2}} = (1.2, 0.25, 0.65, 1.8)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.001	0.336	0.664	14402.569
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.002	0.342	0.658	14594.812
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.001	0.338	0.662	14433.812
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.001	0.339	0.661	14408.425
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.002	0.334	0.666	14623.665
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.001	0.338	0.662	14631.49
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.001	0.331	0.669	14405.285
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.001	0.33	0.67	673.616
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.002	0.324	0.676	14462.569
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	0.002	0.324	0.676	14326.454

Table B.15: Estimates of GACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_3}} = (4, 0.69, 0.2, 12)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.512	0.993	0.007	1.225
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	1.462	0.999	0.001	1.262
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	3.949	0.995	0.005	1.278
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	4.440	0.988	0.012	1.279
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	3.267	0.998	0.002	4.566
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	8.312	0.994	0.006	1.288
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	4.996	0.993	0.007	7.092
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	5.765	0.999	0.001	8.492
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	6.244	0.999	0.001	8.619
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	8.281	0.997	0.003	9.799

Table B.16: Estimates of GACD(1,1), with adjusted actual parameters $\boldsymbol{\vartheta_{0_3}} = (4, 0.69, 0.2, 1.8)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.001	0.726	0.274	1419.925
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.001	0.701	0.299	1368.914
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.000	0.711	0.289	1353.209
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.000	0.707	0.293	1005.202
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.000	0.704	0.296	1282.371
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.001	0.704	0.296	979.731
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.001	0.700	0.300	1064.627
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.003	0.704	0.296	165.753
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.000	0.692	0.308	1363.875
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	0.002	0.699	0.301	1350.780

Table B.17: Estimates of GACD(1,1), with actual parameters $\boldsymbol{\vartheta_{0_5}} = (4.2, 0.6, 0.2, 2)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	0.098	0.903	0.097	292.817
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	0.011	0.803	0.197	637.953
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	0.010	0.815	0.185	433.472
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	0.009	0.815	0.185	752.812
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	0.008	0.816	0.184	615.170
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	0.010	0.811	0.189	760.125
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	0.717	0.815	0.185	437.804
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	0.015	0.810	0.190	1174.777
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	0.025	0.812	0.188	562.488
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	0.003	0.813	0.187	1176.254

Table B.18: Estimates of GACD(1,1), with adjusted actual parameters $\boldsymbol{\vartheta_{0_5}} = (4.2, 0.6, 0.2, 1.8)$

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\theta}$
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_1}}$	11.326	0.740	0.260	9705.712
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_2}}$	14.150	0.839	0.161	9518.480
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_3}}$	10.385	0.702	0.298	9505.665
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_4}}$	13.678	0.824	0.176	9876.777
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_5}}$	10.803	0.719	0.281	9636.859
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_6}}$	11.021	0.728	0.272	9659.121
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_7}}$	11.279	0.739	0.261	9672.383
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_8}}$	11.617	0.752	0.248	9671.827
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_9}}$	11.646	0.753	0.247	9711.045
$\widehat{\boldsymbol{\vartheta}}_{\boldsymbol{T_10}}$	11.837	0.760	0.240	9726.526