BOOTSTRAP UNIT ROOT TESTS
FOR HEAVY-TAILED OBSERVATIONS

by

Andrejus Parfionovas

A thesis submitted in partial fulfillment
of the requirements for the degree
of
MASTER OF SCIENCE
in
Statistics

UTAH STATE UNIVERSITY
Logan, Utah
2003
ABSTRACT

Bootstrap Unit Root Tests for
Heavy-Tailed Observations

by

Andrejus Parfionovas, Master of Science
Utah State University, 2003

Major Professor: Dr. Piotr Kokoszka
Department: Mathematics and Statistics

We explore the application of the bootstrap unit root test to time series with heavy-tailed errors. The size and power of the tests are estimated for two different autoregressive models (AR(1)) using computer simulated data. Real-data examples are also presented. Two different bootstrap methods and the subsampling approach are compared. Conclusions on the optimal bootstrap parameters, the range of applicability, and the performance of the tests are made.

(80 pages)
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 UNIT ROOT TESTS</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Brownian Motion as a Limit of a Random Walk Process</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Convergence of Random Functions</td>
<td>5</td>
</tr>
<tr>
<td>2.3 Functional Central Limit Theorem</td>
<td>6</td>
</tr>
<tr>
<td>2.4 Applications to Unit Root Processes</td>
<td>7</td>
</tr>
<tr>
<td>2.4.1 Model 1: AR(1) with No Constant Term or Time Trend</td>
<td>8</td>
</tr>
<tr>
<td>2.4.2 Model 2: AR(1) with Constant Term but No Time Trend</td>
<td>9</td>
</tr>
<tr>
<td>2.4.3 Model 3: Random Walk with a Constant Term</td>
<td>10</td>
</tr>
<tr>
<td>2.4.4 Model 4: Random Walk with a Constant Term and Time Trend</td>
<td>10</td>
</tr>
<tr>
<td>3 STABLE RANDOM VARIABLES</td>
<td>12</td>
</tr>
<tr>
<td>3.1 Definition of a Stable Distribution</td>
<td>12</td>
</tr>
<tr>
<td>3.2 Relation to Well-Known Distributions</td>
<td>14</td>
</tr>
<tr>
<td>3.3 Properties of the Stable Random Variables</td>
<td>14</td>
</tr>
<tr>
<td>3.4 Extension of the Functional Central Limit Theorem</td>
<td>17</td>
</tr>
<tr>
<td>3.5 Asymptotic Theory for Unit Root Tests for Heavy-Tailed Observations</td>
<td>17</td>
</tr>
<tr>
<td>3.5.1 Model 1: AR(1) with No Constant Term</td>
<td>18</td>
</tr>
<tr>
<td>3.5.2 Model 2: AR(1) with a Constant Term</td>
<td>18</td>
</tr>
<tr>
<td>4 BOOTSTRAP FOR TIME SERIES</td>
<td>20</td>
</tr>
<tr>
<td>4.1 Block Bootstrap</td>
<td>20</td>
</tr>
<tr>
<td>4.1.1 Description of the Bootstrapping Procedure</td>
<td>20</td>
</tr>
<tr>
<td>4.1.2 Specific Features of the Method</td>
<td>21</td>
</tr>
<tr>
<td>4.1.3 Choosing the Optimal Parameters</td>
<td>22</td>
</tr>
<tr>
<td>4.2 AR-sieve Bootstrap</td>
<td>22</td>
</tr>
<tr>
<td>4.2.1 Description of the Bootstrapping Procedure</td>
<td>22</td>
</tr>
<tr>
<td>4.2.2 Specific Features of the Method</td>
<td>23</td>
</tr>
<tr>
<td>4.2.3 Choosing the Optimal Parameter</td>
<td>23</td>
</tr>
<tr>
<td>4.3 Other Methods</td>
<td>24</td>
</tr>
<tr>
<td>4.3.1 Variable Length Markov Chain (VLMC) Sieve Bootstrap</td>
<td>24</td>
</tr>
</tbody>
</table>
### 4.3.2 Local Bootstrap for Conditional Mean Estimates

4.4 Residual Subsampling for Unit Root Test

4.5 Comparison and Conclusions

---

5 BOOTSTRAP UNIT ROOT TESTS FOR HEAVY TAILED OBSERVATIONS

5.1 Model 1: AR(1) Model Without a Constant Term or Time Trend

5.1.1 Residual-based AR-Sieve Bootstrap Algorithm

5.1.2 The Rate of Convergence of the Test Statistic

5.1.3 Size and Power of the Test

5.1.4 Performance of the Test for Different Stability Indexes α

5.1.5 Changing the Bootstrap Sample Size

5.1.6 CI-based Bootstrap Algorithm

5.1.7 The Performance of the CI-based Bootstrap

5.2 Model 2: AR(1) Model with a Constant Term but No Time Trend

5.2.1 Bootstrap Algorithm

5.2.2 Size of the Test

5.2.3 Power of the Test

5.3 Application to Yield Curves on Corporate Bonds

---

6 CONCLUSIONS

6.1 REFERENCES

APPENDIX
LIST OF TABLES

Table | Page
--- | ---
5.1 Series (8.1.1) with $\alpha = 1.75$ were used to estimate the percentage of rejections of $H_0 : \rho = 1$ for different values of the real parameter $\rho$ (nominal size = 5%). 10000 of experiments were performed with 2000 of bootstrap replications each. | 34
5.2 Type-I errors for the model (8.1.1)(percentage of the rejection the null-hypothesis $H_0 : \rho = 1$) for the series of different sizes $n$ and different values of the index of stability $\alpha$. Based on 10000 experiments. Nominal size = 5% | 37
5.3 Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) with the reduced bootstrap samples for model (8.1.1) with $\alpha = 1.75$. Based on 10000 experiments with 2000 bootstrap simulations. | 39
5.4 Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) with the increased bootstrap samples for the model (8.1.1) with $\alpha = 1.75$. | 40
5.5 Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) with the increased bootstrap samples for the model (8.1.1) with $\alpha = 1.5$. | 41
5.6 CI-bootstrap estimates of the type I error (reject true $H_0$) for the AR(1) series of different length, assuming the model (8.1.1) with the $\alpha = 1.75$. The nominal size of the error = 5%. | 43
5.7 Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) of the CI-based bootstrap tests with the reduced bootstrap samples. Based on 10000 experiments with 1000 bootstrap simulations for the model (8.1.1) with $\alpha = 1.75$. The nominal error rate is 5%. | 44
5.8 Type-I errors (percentage of the rejection the null-hypothesis $H_0 : \rho = 1$ and $\tau = 0$) for the series of different sizes $n$ and different values of the index of stability $\alpha$. Based on 10000 experiments. Model 8.2.1. | 46
5.9 Number of rejections of $H_0 : \rho = 1$ and $\tau = 0$ out of 10000 experiments for series of different lengths with $\alpha = 1.75$. We compare the performance of the test for $\tau = 0$ and different values of the real $\rho$, based on 2000 bootstrap replications. 

5.10 Number of rejections of $H_0 : \rho = 1$ and $\tau = 0$ out of 10000 experiments for series of different lengths with $\alpha = 1.75$. We compare the performance of the test for $\rho = 1$ and different values of the parameter $\tau$, based on 2000 bootstrap replications.

5.11 The LSE estimates $\hat{\rho}_{500}$ for the yield curves on corporate bonds.

5.12 The p-values estimated for testing the hypothesis $H_0 : \rho = 1$ on the yield curves (length = 500) using 1000 bootstrap replications.

5.13 The results of the bootstrap tests for the yield curves on corporate bonds: LS estimates $\hat{\rho}_{500}$ and $\hat{\tau}_{500}$, statistic of interest $\Phi_{500}$ and the p-value. The results are based on 1000 bootstrap replications.
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Random samples of size 1000 of (a) standard normal random variable $N(0,1)$, and stable distribution with different values of $\alpha$: (b) $S_{1.5}(1.0.0)$, (c) $S_{1.5}(1.0.0)$, (d) $S_{1.2}(1,0.0)$.</td>
</tr>
<tr>
<td>5.1</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{100}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.2</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{200}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.3</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{300}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.4</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{400}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.5</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{500}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.6</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{600}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.7</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{700}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.8</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{800}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.9</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{900}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.10</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{1000}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.11</td>
<td>Comparison of density estimators $\hat{f}<em>{T</em>{1100}}(t)$ and $\hat{f}<em>{T</em>{500}}(t)$</td>
</tr>
<tr>
<td>5.12</td>
<td>Bold solid curve represents the density estimator of the statistic $T_{500}$ calculated from 10000 of different series (5.1.1). The dashed curves represent ten density estimates of the bootstrapped statistic $T_{500}^*(t)$.</td>
</tr>
<tr>
<td>5.13</td>
<td>Graphical comparison of the power estimates for testing the hypothesis $H_0: \rho = 1$, for the series of different lengths with $\alpha = 1.75$. Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.</td>
</tr>
</tbody>
</table>
Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$, for the series following model (5.1.1) with $\alpha = 1.5$, by using AR-sieve bootstrap (solid line) and the subsampling (dotted line). Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

Type-I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ for the model 5.1.1) for the series of different lengths ($n = 500, 375, 250, 125, 60$) and different values of the stability index ($\alpha = 1.75, 1.5, 1.25, 1.1$) used for the stable distribution. Based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation.

Comparison of the type I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ for the model (5.1.1) with $\alpha = 1.5$) for the AR-sieve bootstrap and the subsampling methods. Bootstrap results are based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation. Subsampling results are based on $R = 5000$ replications.

Type-I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ for the model (5.1.1) with $\alpha = 1.75$) for the series of different lengths ($n = 500, 375, 250, 125, 60$) versus the percentage the bootstrap sample compares to the original series length. Results are based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation. Nominal level = 5%.

Type-I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ for the model (5.1.1) with $\alpha = 1.5$) for the series of different lengths ($n = 500, 375, 250$) versus the percentage the bootstrap sample compares to the original series length. Results are based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation. Nominal level = 5%.
5.19 Type-I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ and $\tau = 0$) for the series of different length ($n = 500, 375, 250, 125, 60$) with different values of the index of stability ($\alpha = 1.75, 1.5, 1.25, 1.1$) used for the stable distribution. Based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation.

5.20 Comparison of the type-I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ and $\tau = 0$) for the series of different length ($n = 500, 375, 250, 125, 60$) with different the index of stability ($\alpha = 1.5$) used for the stable distribution. Based on 10000 experiments for every type of series and 2000 replications for every bootstrap estimation and 5000 subsampling replications.

5.21 Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$ and $\tau = 0$, for the series of different lengths. Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

5.22 Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$ and $\tau = 0$, for the series with $\alpha = 1.75$. Values for the parameter $\tau$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

5.23 Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$ and $\tau = 0$, for the series following the model (5.2.1) with $\alpha = 1.5$ (lengths of series $n = 500$) by using AR-sieve bootstrap (solid line) and the subsampling (dotted line). Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

5.24 Yield curves on corporate bonds and corresponding returns. The bottom panels represent the first differences of the top panels.

5.25 Yield curves on corporate bonds and corresponding returns. The bottom panels represent the first differences of the top panels.

5.26 Yield curves on corporate bonds and corresponding returns. The bottom panels represent the first differences of the top panels.
CHAPTER 1

INTRODUCTION

One of the major problems in econometrics and time series analysis is the forecasting of the random processes, when the task is to predict the value of the process \( \{X_t\} \) at some future moment \( t + k \) by using its previous realizations \((X_1, X_2, \ldots, X_t)\). To see how the forecasting relates to the unit root testing consider the following situation. Suppose we have a simple AR(1) (autoregressive of order 1) model, i.e., the data are generated by the process

\[
X_t = \tau (t + \rho(1 - t)) + \rho X_{t-1} + \varepsilon_t,
\]

where \( \varepsilon_t \) are some zero-mean errors, \( \tau \in \mathbb{R} \), \( 0 < \rho \leq 1 \). It turns out that in this case \( E(X_{t+k} | X_t) = \tau(t + k) + \rho^k (X_t - \tau t) \). One may see that for \( \rho < 1 \) the conditional expectation of \( X_{t+k} \) will depend less on the previous values of \( X_t \) as the prediction step \( k \) increases. Moreover, the decay of this dependence is exponential, which practically means one needs only “a few” previous observations to predict future values of the process. For \( \rho = 1 \), the process turns into

\[
X_t = T + X_{t-1} + \varepsilon_t,
\]

and the conditional expectation is \( E(X_{t+k} | X_t) = \tau k + X_t \), i.e., future values of the process depend on all previous realizations of \( X_t \), which makes the prediction procedure more complicated. Thus, it is important to have proper tools for testing the hypothesis \( H_0 : \rho = 1 \).

The unit root problem is not new. Many unit root tests have been developed for the model described above. However, most of the tests assume the random errors \( \varepsilon_t \) to be independent and identically distributed (i.i.d.) with zero mean and finite variance, i.e., \( E(\varepsilon_t) = 0 \) and \( \text{Var}(\varepsilon_t) < \infty \). On the other hand, in many econometric applications we encounter processes that cannot be described using these assumptions. In particular, the behavior of some financial and macroeconomic time series (like the stock market prices or the financial asset returns) suggests that the errors have an infinite variance \( \text{Var}(\varepsilon_t) = \infty \). Although there are theoretical methods developed for the exploration of such models, the practical application is not well developed yet. Fortunately, the level of the computer power nowadays allows us to use intensive computational methods to overcome these
difficulties.

In this thesis we explore the application of the bootstrap methods for the unit root testing in infinite variance environments. We study their effectiveness and validity using computer simulated data.

The thesis is organized in the following way: in Section 2 we introduce the general theory of unit root tests. Necessary concepts and proofs are provided. The applicability to the most common data models is described. Section 3 provides the theoretical background on stable distributions which are often used to model infinite variance data. Definitions and the main properties of stable random variables are established. Also, the extension of the asymptotic theory of the unit root tests to some of the heavy-tailed models is considered. The bootstrap methods used in time series analysis are examined in Section 4. Different bootstrap approaches are compared and comments on their applicability to certain models are provided. In Section 5, we discuss the results of the bootstrap implementation to the unit root tests for heavy-tailed series. Two different bootstrap methods as well as two different data generating processes with numerous values of the parameters are considered. The comparison with a different numerical approach (subsampling) is performed. In addition to the exploration of the properties of the method using the computer simulated data we demonstrate its performance with several real data examples. Finally, we address some practical considerations on setting the optimal bootstrap parameters and finish with our conclusions in Section 6.
CHAPTER 2
UNIT ROOT TESTS

Consider a Gaussian AR(1) model (Markov process):

\[ X_t = \rho X_{t-1} + \epsilon_t, \tag{2.0.1} \]

where \( \epsilon_t \sim \text{i.i.d. } N(0, \sigma^2) \), \( \rho \leq 1 \), and \( X_0 = 0 \).

Properties of this model (confidence intervals, forecasting, etc.) can be explored by using classical methods of time series analysis [15]. However, most of these methods were created for the weakly stationary time series, i.e., for such \( \{X_t\} \) that \( \text{E}(X_t^2) < \infty \), for all \( t \). In case of the Markov process (2.0.1) the variance is

\[ \text{Var}(X_t) = \frac{\sigma^2}{1-\rho^2}. \]

We can see that \( \text{Var}(X_t) \to \infty \), when \( \rho \to 1 \). This means that the process becomes non-stationary and requires completely different methods of exploration to be used.

So our primary objective can be formulated as testing the hypothesis \( H_0 : \rho = 1 \). Such tests are called “unit root tests.” Many unit root tests were investigated (see [22] and [11]). We will focus on the method described in [15]. It was shown that for \( |\rho| < 1 \) the Ordinary Least Square (OLS) estimator

\[ \hat{\rho}_n = \hat{\rho}_{LS} = \frac{\sum_{j=2}^{n} X_{j-1}X_j}{\sum_{j=2}^{n} X_{j-1}^2} \tag{2.0.2} \]

can be used to construct a statistic which asymptotically has a Normal distribution, i.e.,

\[ \sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{L} N(0, (1-\rho^2)) \]

However, in case of the unit root process (\( \rho = 1 \)) this distribution turns out to be degenerated \( (1-\rho^2 = 0) \). This happens because of the fast convergence rate of the estimator of the unit root coefficient when \( \rho = 1 \).

To overcome this difficulty and obtain a reasonable asymptotic distribution, the statistic

\[ n(\hat{\rho}_n - 1) \tag{2.0.3} \]
should be used instead. To find the asymptotic distribution of (2.0.3), the Brownian motion process must be used.

2.1 Brownian Motion as a Limit of a Random Walk Process

**Definition 2.1.1**: Standard Brownian motion $W(\cdot)$ is a continuous-time stochastic process, associating each $t \in [0, 1]$ with the scalar random variable $W(t) \in \mathbb{R}$, such that:

1) $W(0) = 0$;
2) For any $0 \leq t_1 < t_2 < \ldots < t_k \leq 1$, the increments $[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \ldots,$ $[W(t_k) - W(t_{k-1})]$ are independent and Gaussian with

$$[W(s) - W(t)] \sim N(0, s - t)$$

3) $W(t)$ is continuous in $t$ with probability 1.

To see how the Brownian motion relates to the random walk process, consider the model

$$y_t = y_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim N(0, 1)$. We assume that $y_0 = 0$. This corresponds to the first statement of the definition 2.1.1.

Because of that we can rewrite (2.1.1) as a sum of $\varepsilon_t$:

$$y_t = \sum_{i=1}^{t} \varepsilon_i \sim N(0, t).$$

Since the $\varepsilon_i$ are i.i.d. $N(0, 1)$, it follows that $y_t \sim N(0, t)$. This means that for any integers $0 < t < s$ the difference $(y_s - y_t)$ also has a Normal distribution, i.e.,

$$y_s - y_t = \sum_{i=t+1}^{s} \varepsilon_i \sim N(0, (s - t)).$$

Moreover, $(y_s - y_t)$ is independent of any $(y_q - y_r)$ for $0 < t < s < r < q$. This actually is a special case of the second statement of the definition 2.1.1.

However, a random walk is a discrete process. A continuous-time Brownian motion can be obtained in the following way. Each $\varepsilon_t$ can be treated as a sum of i.i.d. random variables:

$$\varepsilon_t = \sum_{i=1}^{n} \varepsilon_i,$$ where $\varepsilon_i \sim N(0, 1/n)$. 
Then, as we consider a limit as \( n \to \infty \). The process \( y_t \) will become continuous, i.e., the difference 
\[(y_t - y_{t+\Delta}) \to 0, \quad \text{when} \quad \Delta \to 0.
\]
Finally, we want to restrict the domain of such process to \( t \in [0,1) \) so all conditions of the definition 2.1.1 will be satisfied.

2.2 Convergence of Random Functions

In order to explore asymptotic properties of functions of stochastic processes, the convergence for continuous random functions should be defined. Here we basically extend the definitions of convergence for random variables. A more formal definition which also covers discontinuous functions can be found in [4].

**Definition 2.2.1:** Consider a continuous-time stochastic process \( S(\cdot) \), such that any given realization \( S(r) \) is a continuous function of \( r \in [0,1] \) with probability 1. Let a sequence of such continuous random functions \( S_T(\cdot), T = 1,2, \ldots \), satisfy the following conditions:

1) For any finite collection of \( k \) dates \( 0 \leq r_1 < r_2 < \ldots < r_k \leq 1 \) the sequence of \( k \)-dimensional vectors 
\[
[S_T(r_1), S_T(r_2), \ldots, S_T(r_k)], \quad T = 1,2, \ldots \text{ converges in distribution to the vector } [S(r_1), S(r_2), \ldots, S(r_k)].
\]

2) For each \( \varepsilon > 0 \) the probability \( P\{|S_T(r_1) - S_T(r_2)| > \varepsilon\} \to 0 \) uniformly in \( T \), as \( |r_1 - r_2| = \delta \to 0 \).

3) The probability \( P\{|S_T(0)| > \lambda\} \to 0 \) uniformly in \( T \), as \( \lambda \to \infty \).

Then we say that the sequence of random functions \( S_T(\cdot) \) converges in law to \( S(\cdot) \), i.e., \( S_T(\cdot) \xrightarrow{L} S(\cdot) \).

**Definition 2.2.2:** Let \( S_T(\cdot) \) and \( V_T(\cdot), T = 1,2, \ldots \), be sequences of random continuous functions, such that \( S_T : r \in [0,1] \to \mathbb{R}^1 \) and \( V_T : r \in [0,1] \to \mathbb{R}^1 \). Let the sequence of the random variables \( Y_T, T = 1,2, \ldots \), be defined as follows:

\[
Y_T = \sup_{r \in [0,1]} |S_T(r) - V_T(r)|.
\]

If this \( Y_T \) converges in probability to zero, then we say that the sequence \( S_T(\cdot) \) converges in probability to \( V_T(\cdot) \), i.e., \( S_T(\cdot) \xrightarrow{P} V_T(\cdot) \).

The following continuous mapping theorem will be used widely in further derivations.
THEOREM 2.2.1 (CONTINUOUS MAPPING THEOREM): Consider a continuous functional \( g(\cdot) : g(S(\cdot)) \to \mathbb{R}^1 \), where \( S(\cdot) \) is a random continuous function. Then for any converging sequence of random functions such that \( S_T(\cdot) \xrightarrow{L} S(\cdot) \), convergence \( g(S_T(\cdot)) \xrightarrow{L} g(S(\cdot)) \) also holds.

The proof of this theorem can be found in [15], Chapter 17, page 482.

2.3 Functional Central Limit Theorem

Here we describe the Functional Central Limit Theorem, which will be used to determine the asymptotic properties of the random walk statistics. The standard Central Limit Theorem (CLT) says that if \( \epsilon_t \) is a sequence of i.i.d. random variables with mean 0 and variance \( \sigma^2 \), then

\[
\sqrt{T} \frac{\epsilon_t}{\sigma} \xrightarrow{L} N(0,1),
\]

where \( \epsilon_t = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \) is the sample mean.

Similarly to the sample mean, we construct a random variable

\[
X_T(r) = \frac{1}{T} \sum_{t=1}^{[Tr]} \epsilon_t,
\]

where \( r \in [0,1] \), and \([Tr]\) denotes the largest integer that is less or equal to \( Tr \). To obtain a reasonable distribution, this statistic must be re-scaled as:

\[
\sqrt{T} \frac{X_T(r)}{\sigma} = \sqrt{\frac{[Tr]}{T}} \frac{\sum_{t=1}^{[Tr]} \epsilon_t}{\sigma \sqrt{[Tr]}},
\]

where the first multiplier approaches \( \sqrt{r} \) when \( T \to \infty \), and the second multiplier follows the standard CLT, hence

\[
\sqrt{T} \frac{X_T(r)}{\sigma} \xrightarrow{L} N(0,r).
\]

It can be easily shown that for any \( 0 < r_1 < r_2 < 1 \) the difference \( X_T(r_2) - X_T(r_1) \) is also asymptotically Normal:

\[
\sqrt{T} \frac{X_T(r_2) - X_T(r_1)}{\sigma} = \sqrt{T} \frac{\frac{1}{T} \sum_{t=1}^{[Tr_2]} \epsilon_t}{\sigma} \xrightarrow{L} N(0, r_2 - r_1).
\]

Moreover, this difference is independent of the statistic (2.3.2) assuming that \( r < r_1 \).
The above discussion motivates the Functional CLT which states that for any i.i.d. random variables \( \xi_t \) with mean 0 and variance \( \sigma^2 \) the sequence of stochastic functions defined in (2.3.2) converges to the standard Brownian Motion process, i.e.,

\[
\sqrt{T} \frac{X_T(t)}{\sigma} \xrightarrow{L} W(t).
\]

We can see that the standard CLT follows from (2.3.3), because when setting \( r = 1 \) we have:

\[
\sqrt{T} \frac{X_T(1)}{\sigma} = \sum_{i=1}^{T} \frac{\xi_i}{\sigma \sqrt{T}} \xrightarrow{L} W(1) \sim N(0,1).
\]

2.4 Applications to Unit Root Processes

Consider a random walk process

\[ y_t = y_{t-1} + \xi_t, \]

where \( \xi_t \) are i.i.d. with \( E(\xi_t) = 0 \) and \( Var(\xi_t) = \sigma^2 \). We also assume that \( y_0 = 0 \), hence \( y_t = \sum_{i=1}^{t} \xi_i \).

Before describing any particular models that can be applied for this process, we introduce some useful asymptotic distributions of its statistics, namely: \( T^{-3/2} \sum_{i=1}^{T} y_{t-1}, T^{-2} \sum_{i=1}^{T} y_{i-1}^2, T^{-1} \sum_{i=1}^{T} y_{t-1} \xi_t \) and \( T^{-3/2} \sum_{i=1}^{T} \xi_t^2 \).

First of all we construct a stochastic function \( X_T(t) \) according to the equation (2.3.1), i.e., for any realization we have

\[ X_T(r) = \frac{1}{T} \sum_{t=1}^{[Tr]} \xi_t = \frac{y_T}{T}. \]

Notice that

\[ \int_{0}^{1} X_T(r) dr = \sum_{t=1}^{T} \frac{1}{T} \frac{y_{t-1}}{T} = \frac{1}{T^2} \sum_{t=1}^{T} y_{t-1}. \]

Now, by using Theorem 2.2.1 and (2.3.4) we obtain

\[
T^{-3/2} \sum_{t=1}^{T} y_{t-1} = \sqrt{T} \int_{0}^{1} X_T(r) dr \xrightarrow{L} \sigma \int_{0}^{1} W(r) dr.
\]

Similarly, by using the stochastic function \( S_T(r) = T (X_T(r))^2 \), it can be shown that

\[
T^{-2} \sum_{i=1}^{T} y_{i-1}^2 = \frac{L}{\sigma^2} \int_{0}^{1} [W(r)]^2 dr,
\]
and

\[(2.4.3) \quad T^{-1} \sum_{t=1}^{T} y_{t-1} \varepsilon_t = \frac{L}{2} \sigma^2 ([W(1)]^2 - 1),\]

See [15], Section 17.3, page 485, for the detailed proof.

Finally, let us find the distribution of the statistic \( T^{-3/2} \sum_{t=1}^{T} t \varepsilon_t \). By using the fact that \( y_t = \sum_{i=1}^{t} u_i \), one can easily show that \( T^{-3/2} \sum_{t=1}^{T} y_{t-1} = T^{-1/2} \sum_{t=1}^{T} \varepsilon_t - T^{-3/2} \sum_{t=1}^{T} \varepsilon_t \), hence

\[
T^{-3/2} \sum_{t=1}^{T} \varepsilon_t = T^{-1/2} \sum_{t=1}^{T} \varepsilon_t - T^{-3/2} \sum_{t=1}^{T} y_{t-1}.
\]

Then, according to (2.3.5) and (2.4.1) it follows that

\[(2.4.4) \quad T^{-3/2} \sum_{t=1}^{T} \varepsilon_t \sim \sigma \left( W(t) - \int_{0}^{1} W(r) dr \right).\]

Below we discuss several models that are commonly used for unit root testing.

2.4.1 Model 1: AR(1) with No Constant Term or Time Trend

Suppose we want to fit an AR(1) regression model with only one parameter \( \rho \):

\[ y_t = \rho y_{t-1} + \varepsilon_t, \]

where \( \varepsilon_t \) are i.i.d. with mean zero and variance \( \sigma^2 \).

Our goal is to test the hypothesis \( H_0 : \rho = 1 \) using the OLS estimator of the unknown parameter:

\[ \hat{\rho}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2}. \]

To obtain an asymptotic distribution of this statistic when \( \rho = 1 \), the estimator must be re-scaled as it follows from (2.0.3):

\[(2.4.5) \quad T(\hat{\rho}_T - 1) = \frac{T^{-1} \sum_{t=1}^{T} y_{t-1} \varepsilon_t}{T^{-2} \sum_{t=1}^{T} y_{t-1}^2}. \]

The asymptotic distributions of the numerator and denominator are given in (2.4.4) and (2.4.2), respectively. Now, since both the numerator and the denominator converge in law, it follows according to Theorem 2.2.1 that under the null hypothesis \( H_0 : \rho = 1 \) the statistic (2.4.5) has the following asymptotic distribution:

\[(2.4.6) \quad T(\hat{\rho}_T - 1) \xrightarrow{L} \frac{[W(1)]^2 - 1}{2 \int_{0}^{1} [W(r)]^2 dr}.\]
This limiting distribution cannot be transformed to any simple form involving some standard distribu­tions. However, it can be shown (see [15], Section 17.4, page 488) that it is skewed to the left.

2.4.2 Model 2: AR(1) with Constant Term but No Time Trend

Here we want to fit a model with a constant term

\[ y_t = \tau + \rho y_{t-1} + \varepsilon_t, \]

where \( \varepsilon_t \) are i.i.d. with mean zero and variance \( \sigma^2 \), while the data generating process remains the same as in (2.0.1).

As before, we want to check whether our data is a random walk, i.e., we want to test the hypothesis \( H_0 : \tau = 0 \text{ and } \rho = 1 \). For this purpose we need to know the asymptotic distribution of the OLS estimates of \( \tau \) and \( \rho \), which are obtained as:

\[
\begin{bmatrix}
\hat{\tau}_T \\
\hat{\rho}_T
\end{bmatrix} = \left[ \sum_{y_{t-1}}^T \sum_{y_{t-1}} \right]^{-1} \left[ \sum_{y_{t-1}} \sum_{y_{t-1}} \right].
\]

Here the indexes of summation go over \( t = 1, 2, \ldots, T \).

The difference between the vector estimate and the true value is:

\[
(2.4.7) \quad \begin{bmatrix} \hat{\tau}_T \\
\hat{\rho}_T - 1 \end{bmatrix} = \begin{bmatrix} \hat{\tau}_T \\
\hat{\rho}_T - 1 \end{bmatrix} = \left[ \sum_{y_{t-1}}^T \sum_{y_{t-1}} \right]^{-1} \left[ \sum_{y_{t-1}} \sum_{y_{t-1}} \right].
\]

It turns out that the sums presented in this formula have different convergence rates: \( \sum y_{t-1} = O_p(T^{3/2}) \), \( \sum y_{t-1} \varepsilon_t = O_p(T) \), \( \sum y_{t-1} = O_p(T^2) \), and \( \sum \varepsilon_t = O_p(T^{1/2}) \). Because of these different convergence rates, to get an asymptotic distribution, the difference (2.4.7) needs to be re-scaled by the matrix

\[
Y_T = \begin{bmatrix} T^{1/2} & 0 \\ 0 & T \end{bmatrix}.
\]

It can be shown (see [15], Section 17.4, page 491) that

\[
Y_T \begin{bmatrix} \hat{\tau}_T \\
\hat{\rho}_T - 1 \end{bmatrix} \overset{L}{\to} \begin{bmatrix} \sigma & 0 \\ 0 & 1 \end{bmatrix} \left[ \int_0^1 W(r)dr \int_0^1 W(r)^2 dr \right]^{-1} \left[ \int_0^1 (W(1))^2 dr - 1/2 \right],
\]

which means that

\[
(2.4.8) \quad T(\hat{\rho}_T - 1) \overset{L}{\to} \frac{1}{2} \left[ \int_0^1 (W(1))^2 dr - 1 \right] W(1) \cdot \int_0^1 W(r) dr - \left[ \int_0^1 W(r) dr \right]^2.
\]
As we can see the distribution of the statistic becomes even more complicated than (2.4.6). It is also more strongly skewed to the left, what means that null hypothesis $H_0: \rho = 1$ will be rejected more frequently.

2.4.3 Model 3: Random Walk with a Constant Term

Now suppose that the data was obtained from a random walk process with a drift:

$$y_t = \tau + y_{t-1} + \varepsilon_t, \tau \neq 0.$$ 

Say we want to fit the following model:

$$y_t = \tau + \rho y_{t-1} + \varepsilon_t,$$

and test the hypothesis $H_0: \rho = 0$.

One might expect that the asymptotic distribution of $\hat{\tau}$ and $\hat{\rho}$ should be similar to (2.4.8), but actually it can be shown (see [15], Section 17.4, page 495) that both estimates are asymptotically Gaussian:

$$T^{1/2}(\hat{\tau} - \tau) \sim N(0, \sigma^2 Q^{-1}),$$

where $Q = \begin{bmatrix} 1 & \tau^2 / 3 \\ \tau / 2 & \tau^2 / 3 \end{bmatrix}$.

2.4.4 Model 4: Random Walk with a Constant Term and Time Trend

Finally we want to fit a model with a constant term and time trend:

$$y_t = \tau + \delta t + \rho y_{t-1} + \varepsilon_t, \tau \neq 0,$$

for the data which was generated by the process:

$$y_t = \tau + y_{t-1} + \varepsilon_t.$$

It is useful to rewrite (2.4.10) in the form:

$$y_t = (1 - \rho)\tau + \rho [y_{t-1} - \tau(t - 1)] + (\delta + \rho \tau) t + \varepsilon_t = \tau^* \rho^* \varepsilon_{t-1} + \delta^* t + \varepsilon_t.$$
Then the hypothesis \( H_0 : \tau = \tau_0, \rho = 1, \) and \( \delta = 0 \) is equivalent to \( \tau^* = 0, \rho^* = 1, \) and \( \delta^* = \tau_0. \)

In [15] (Section 17.4, page 498) it is shown that

\[
\begin{bmatrix}
T^{1/2} \tau_* \\
T(\delta_*^T - 1) \\
T^{2} (\delta_*^T - \tau_0)
\end{bmatrix}
\sim
\begin{bmatrix}
\frac{1}{2} \int W(r) dr \\
\int [W(r)]^2 dr \\
\int W(r) dr
\end{bmatrix}
\begin{bmatrix}
1 & \int W(r) dr & 1/2 \\
\int [W(r)]^2 dr & \int W(r) dr & 1/3 \\
W(1) & \frac{1}{2} \{[W(1)]^2 - 1\} & W(1) - \int W(r) dr
\end{bmatrix}^{-1} \Bigg[ \frac{W(1)}{2} \{[W(1)]^2 - 1\} \Bigg].
\]
CHAPTER 3

STABLE RANDOM VARIABLES

In this section we describe the family of univariate $\alpha$-stable distributions. The distribution has four parameters which makes it very flexible and applicable in many situations. Econometrics and financial mathematics use stable models for exploration of financial asset returns, stock market prices, etc. Stable distributions are also extensively used in hydrology and in modelling network traffic.

3.1 Definition of a Stable Distribution

There are several ways to define a stable distribution see [23], Chapter 1, Section 1.1. The following definition explains the name of the distribution:

**DEFINITION 3.1.1:** A real-valued random variable $X$ is said to have a stable distribution if for any $n \geq 2$ there are $a_n > 0$ and $b_n \in \mathbb{R}$ such that:

\[
    a_n \sum_{i=1}^{n} X_i + b_n \overset{d}{=} X,
\]

where $X_i$ are i.i.d. copies of $X$.

$X$ is said to be strictly stable if $b_n = 0$.

This definition reflects the key-feature of this distribution family and also explains its name. In other words a random variable is called stable, if a certain linear combination of its copies has the same distribution. For example a Normal random variable belongs to the stable distribution family, because if $X_i \sim N(\mu, \sigma^2), i = 1, 2, \ldots, n$ then

\[
    \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2) \iff \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i - \sqrt{n}\mu \sim N(\mu, \sigma^2).
\]

As a counter-example consider a set of exponential random variables $X_i \sim Exp(\lambda), i = 1, 2, \ldots, n$.

In this case

\[
    \sum_{i=1}^{n} X_i \sim Gamma(\lambda, n),
\]

which is not an exponential distribution.
Though Definition 3.1.1 does not give any specific information about the stable distribution itself, it can be shown that to satisfy equation 3.1.1 the distribution of a random variable must be of a certain form. In general case it depends on four parameters and neither the cumulative distribution, nor the density function can be written in a closed form. However, properties of this distribution can be explored by using its characteristic function:

\[
E e^{iXt} = \begin{cases} 
\exp\left\{ -\sigma^\alpha |t|^\alpha \left( 1 - i\beta \text{sign}(t) \tan(\pi\alpha/2) \right) + i\mu t \right\}, & \text{if } \alpha \neq 1 \\
\exp\left\{ -|t| \left( 1 + i\beta^2 \text{sign}(t) \ln |t| \right) + i\mu t \right\}, & \text{if } \alpha = 1
\end{cases}
\]

where \( \alpha \in (0, 2], \sigma \geq 0, \beta \in [-1, 1], \mu \in \mathbb{R}^1 \), and

\[
\text{sign}(t) = \begin{cases} 
-1, & \text{if } t < 0 \\
0, & \text{if } t = 0 \\
1, & \text{if } t > 0
\end{cases}
\]

To denote that \( X \) follows the stable distribution we use notation: \( X \sim S_\alpha(\sigma, \beta, \mu) \).

The parameter \( \alpha \) is called the index of stability. When \( \alpha = 2 \), equation (3.1.1) becomes the characteristic function of a Gaussian random variable with mean \( \mu \) and variance \( 2\sigma^2 \). Another special case is Cauchy distribution which can be obtained by setting \( \alpha = 1 \) and \( \beta = 0 \) (see 3.2.2). However if \( \alpha = 1 \) and \( \beta \neq 0 \), then properties of the distribution become difficult to explore and require separate exploration because of the factor \( \ln |\theta| \).

The parameter \( \beta \) is known as a skewness parameter. Density function of \( S_\alpha(\sigma, \beta, \mu) \) is symmetric with respect to \( \mu \) if and only if \( \beta = 0 \). The distribution is skewed to the right if \( \beta > 0 \) or to the left if \( \beta < 0 \). When \( |\beta| = 1 \), it is said to be totaly skewed to the right or to the left, respectively.

The parameter \( \mu \) is a location (shift) parameter, because if \( X \sim S_\alpha(\sigma, \beta, \mu) \) and \( a \in \mathbb{R}^1 \), then \( X + a \sim S_\alpha(a \sigma, \beta, \mu + a) \). However it can be shown that \( EX = \mu \) if and only if \( \alpha \in (1, 2] \) ([23], chapter 1, section 1.2). This is because \( E[X] = \infty \) if \( \alpha \leq 1 \), see Property 3.3.11.

The parameter \( \sigma \) is often called the scale parameter, because the product of a stable \( X \sim S_\alpha(\sigma, \beta, \mu) \) and any real \( a \neq 0 \) can be written as:

\[
aX \sim \begin{cases} 
S_\alpha([a]\sigma, \text{sign}(a)\beta, a\mu), & \alpha \neq 1 \\
S_1([a]\sigma, \text{sign}(a)\beta, a\mu - \frac{2}{\alpha} \ln(|a|)\sigma \beta), & \alpha = 1
\end{cases}
\]

This name however is a misnomer when \( \alpha = 1 \) and \( \beta \neq 0 \).
3.2 Relation to Well-Known Distributions

With certain values of the parameters α-stable density can be written in a closed form. Moreover it can take form of a well-known distribution.

When \( \alpha = 2 \), α-stable distribution becomes Gaussian with density function

\[
\frac{1}{2\sigma\sqrt{\pi}} \exp\left\{ -\frac{(x - \mu)^2}{4\sigma^2} \right\}.
\]

Moreover, parameter \( \beta \) becomes irrelevant, i.e., can take any value without affecting the distribution.

For the simplicity we assume \( \beta = 0 \) so that \( S_2(\sigma, 0, \mu) = N(\mu, 2\sigma^2) \).

When \( \alpha = 1 \) and \( \beta = 0 \) the distribution turns into Cauchy(\( \mu, \sigma \)) with the density function

\[
\frac{\sigma}{\pi ((x - \mu)^2 + \sigma^2)}.
\]

When \( \alpha = 1/2 \) and \( \beta = 1 \) we have Lévy distribution with the density function

\[
\frac{1}{(x - \mu)^{3/2}} \sqrt{\frac{\sigma}{2\pi}} \exp\left\{ -\frac{\sigma}{2(x - \mu)} \right\}.
\]

When both \( \sigma = \beta = 0 \) then the distribution degenerates to the constant \( \mu \) for any \( \alpha \in (0,2] \).

This case has no practical use and we exclude it to maintain simplicity.

3.3 Properties of the Stable Random Variables

Here we describe some properties of the α-stable distribution that are important for our further exploration. Other properties and their proofs can be found in [23] (Chapter 1, Section 1.2) and [21] (Chapter 2, Section 2.1). Below we assume \( X \sim S_\alpha(\sigma, \beta, \mu) \) if not specified otherwise.

**PROPERTY 3.3.1:** For any \( a \in \mathbb{R}^1 \)

\( X + a \sim S_\alpha(\sigma, \beta, \mu + a) \).

**PROPERTY 3.3.2:** If \( X \sim S_\alpha(\sigma, \beta, \mu) \) and \( 0 \neq a \in \mathbb{R}^1 \), then

\[
aX \sim \left\{ \begin{array}{ll}
S_\alpha(|a|\sigma, \text{sign}(a)\beta, a\mu), & \alpha \neq 1 \\
S_1(|a|\sigma, \text{sign}(a)\beta, a\mu - \frac{2}{\sigma} \ln(|a|) \sigma \beta), & \alpha = 1.
\end{array} \right.
\]

**PROPERTY 3.3.3:** If \( \mu = 0 \), then for any \( 0 < a < 2 \)

\( X \sim S_\alpha(\sigma, \beta, 0) \iff -X \sim S_\alpha(\sigma, -\beta, 0) \).
PROPERTY 3.3.4: \( X \sim S_\alpha(\sigma, \beta, \mu) \) is symmetric if and only if \( \beta = 0 \) and \( \mu = 0 \). It is symmetric about \( \mu \) if and only if \( \beta = 0 \).

It also follows that a symmetric stable random variable is strictly stable. Converse statement in general is not true.

PROPERTY 3.3.5: If \( X \sim S_\alpha(\sigma, \beta, \mu) \) with \( \alpha \neq 1 \), then \( X \) is strictly stable if and only if \( \mu = 0 \).

It follows from properties 3.3.1 and 3.3.5 that any non-stable random variable \( X \sim S_\alpha(\sigma, \beta, \mu) \), \( \alpha \neq 1 \) can be made strictly stable by shifting: \( X - \mu \). Condition \( \alpha \neq 1 \) is vital, because of the following property:

PROPERTY 3.3.6: If \( \alpha = 1 \), i.e., \( X \sim S_1(\sigma, \beta, \mu) \), then \( X \) is strictly stable if and only if \( \beta = 0 \).

In other words non-strictly 1-stable random variable cannot be made strictly stable by shifting. However shifting can make it symmetric about 0 (Property 3.3.4).

PROPERTY 3.3.7: If \( X_i, i = 1, 2, \ldots, n \) are i.i.d. \( S_\alpha(\sigma, \beta, \mu) \), then
\[
\sum_{i=1}^{n} X_i \stackrel{d}{=} \begin{cases} 
    n^{1/\alpha} X_1 + \mu \left( n - n^{1/\alpha} \right), & \alpha \neq 1 \\
    n X_1 + \frac{\sigma \beta}{n} \ln(n), & \alpha = 1.
\end{cases}
\]

PROPERTY 3.3.8: For \( \alpha < 1 \) and \( \mu = 0 \) distribution \( S_\alpha(\sigma, \beta, 0) \) has support on the whole real line, i.e., random variable \( X \) takes any real values from \((-\infty, \infty)\)

PROPERTY 3.3.9: If \( \alpha < 1, \mu = 0 \), and any fixed \( \sigma > 0 \) the distribution \( S_\alpha(\sigma, \beta_2, 0) \) is stochastically greater than \( S_\alpha(\sigma, \beta_1, 0) \) for any \(-1 < \beta_1 < \beta_2 < 1\), i.e., if \( X_i \sim S_\alpha(\sigma, \beta_i, 0) \), then \( P\{X_1 \geq x\} \leq P\{X_2 \geq x\} \).

The following property describes the asymptotic behavior of the tail probabilities of non-Gaussian \( \alpha \)-stable random variables (\( \alpha \neq 2 \)).

PROPERTY 3.3.10: If \( X \sim S_\alpha(\sigma, \beta, \mu) \) with \( 0 < \alpha < 2 \), then
\[
\begin{align*}
\lim_{\lambda \to \infty} \lambda^\alpha P\{X > \lambda\} &= C_\alpha \frac{1+\beta}{2} \sigma^\alpha, \\
\lim_{\lambda \to \infty} \lambda^\alpha P\{X < -\lambda\} &= C_\alpha \frac{1-\beta}{2} \sigma^\alpha,
\end{align*}
\]

where
\[
C_\alpha = \left( \int_0^{\infty} x^{-\alpha} \sin x dx \right)^{-1} = \begin{cases} 
    (1 - \alpha) [\Gamma(2 - \alpha) \cos(\pi \alpha / 2)]^{-1}, & \text{if } \alpha \neq 1 \\
    2/\pi, & \text{if } \alpha = 1.
\end{cases}
\]
In other words, equation 3.3.1 states that the tail probabilities behave like \( \lambda^{-\alpha} \) in contrast with normal whose decay is exponential. To illustrate it compare the graphs of \( \alpha \)-stable samples of size 1000 with \( \alpha = 1.9, 1.5, 1.2 \) and Gaussian with \( \mu = 0 \) and \( \sigma = 1, 5 \):

![Graphs of \( \alpha \)-stable samples with \( \alpha = 1.9, 1.5, 1.2 \) and Gaussian distributions with \( \mu = 0 \) and \( \sigma = 1, 5 \).]

Such behavior leads to very unpleasant properties of this distribution. One of them is given below:

**Property 3.3.11:** For \( X \sim S_\alpha(\sigma, \beta, \mu) \) with \( 0 < \alpha < 2 \)

\[
E|X|^p < \infty, \text{ for any } 0 < p < \alpha, \\
E|X|^p = \infty, \text{ for any } p \geq \alpha.
\] 

(3.3.2)

To overcome this difficulty when dealing with the moments of stable random variable the following property is used:

**Property 3.3.12:** Suppose \( X \sim S_\alpha(\sigma, \beta, \mu) \) with \( 0 < \alpha < 2 \) (\( \beta = 0 \) when \( \alpha = 1 \)). Then for every
$0 < p < \alpha$ there is a constant $c_{\alpha, \beta}(p)$ such that

\[(E|X|^p)^{1/p} = \sigma c_{\alpha, \beta}(p) = \sigma (E|X_0|^p)^{1/p},\]

where $X_0 \sim \mathcal{S}_\alpha(1, \beta, 0)$

### 3.4 Extension of the Functional Central Limit Theorem

In Section 1.3 we have established the Functional CLT in terms of the standard Brownian Motion process, see (2.3.4).

A similar theorem can be proved for the random variables that follow a strictly stable distribution, namely:

\[\frac{1}{n^{1/\alpha} I(n)} \sum_{i=1}^{[nt]} \xi_i \overset{L}{\rightarrow} L_\alpha(t),\]

where $I(n)$ is a certain function slowly varying as $n \to \infty$, and $L_\alpha(t)$ is a Lévy process with strictly stable increments.

Note that when the stability index $\alpha = 2$, the Lévy process turns into a Brownian motion: $L_2(t) = W(t)$.

It was also shown in [12] that for $n \to \infty$ the following asymptotic distribution is valid:

\[\left(\frac{1}{n^{1/\alpha} I(n)} \sum_{i=1}^{[nt]} u_i, \frac{1}{n^{2/\alpha^2} I(n)} \sum_{i=1}^{[nt]} u_i^2\right) \overset{L}{\rightarrow} (L_\alpha(t), [L_\alpha(t)]),\]

where $[L_\alpha](t) = L_\alpha^2(t) - 2 \int_0^t L_\alpha(s-d)L_\alpha(s), t \geq 0$.

### 3.5 Asymptotic Theory For Unit Root Tests for Heavy-Tailed Observations

Now we can review the unit root process (2.0.1), by setting the random components $\xi_i$ to be i.i.d. $\mathcal{S}_\alpha(1, 0, 0), \alpha \in (1, 2]$ (most of the properties are also valid for $\alpha \in (0, 1]$). In this section we describe (without giving a proof) the asymptotic properties of the statistics of interest for the fitted models with (2.4.2) and without constant term (2.4.1). More detailed description can be found in [15].
3.5.1 Model 1: AR(1) with No Constant Term

Say we want to fit one-parametric AR(1) regression model (2.4.1) to the data generated by the process:

\[ y_t = \rho y_{t-1} + \varepsilon_t, \]

where \( \varepsilon_t \sim \text{i.i.d. } S_\alpha(1,0,0). \)

It can be shown that under the null-hypothesis \( H_0: \rho = 1 \), the OLS estimator \( \hat{\rho} \) asymptotically follows the distribution:

\[ n(\hat{\rho} - 1) = n \left( \frac{\sum_{i=1}^{n} y_i y_{i-1}}{\sum_{i=1}^{n} y_{i-1}^2} - 1 \right) \xrightarrow{d} \frac{L^1_0 L_\alpha(s-) dL_\alpha(s)}{\int_0^1 L^2_\alpha(s) ds}. \]

And also the transformation \( t_\rho = (\hat{\rho} - 1)/\hat{s}_\rho \), where

\[ \hat{s}_\rho^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{\rho})y_{i-1}}{n \sum_{i=1}^{n} y_{i-1}^2} \]

converges weakly to:

\[ t_\rho \xrightarrow{d} \frac{L^1_0 L_\alpha(s-) dL_\alpha(s)}{\sqrt{[L_\alpha(1)] \int_0^1 L^2_\alpha(s) ds}}. \]

3.5.2 Model 2: AR(1) with a Constant Term

Now suppose we want to fit a AR(1) regression model with two parameters:

\[ y_t = \tau + \rho y_{t-1} + \varepsilon_t, \]

to the data generated by (3.5.1). where \( \varepsilon_t \sim \text{i.i.d. } S_\alpha(1,0,0). \)

In this case we are testing the composite null-hypothesis \( H_0: \rho = 0 \) and \( \tau = 1 \). Under this hypothesis the OLS estimators \( \hat{\rho} \) and \( \hat{\tau} \) have the following distribution:

\[ n(\hat{\rho} - 1) \xrightarrow{d} \frac{L^1_0 L_\alpha(s-) dL_\alpha(s) - L_\alpha(1) \int_0^1 L_\alpha(s) ds}{\int_0^1 L^2_\alpha(s) ds - (\int_0^1 L_\alpha(s) ds)^2}, \]

and

\[ \frac{\hat{\tau} n^{1/\alpha}}{l(n)} \xrightarrow{d} \frac{L_\alpha(1) \int_0^1 L^2_\alpha(s) ds - \int_0^1 L_\alpha(s) ds \int_0^1 L_\alpha(s - dL_\alpha(s))}{\int_0^1 L^2_\alpha(s) ds - (\int_0^1 L_\alpha(s) ds)^2}. \]

The \( t \)-statistic in this case has the following form:

\[ t_{\rho,\tau} = \frac{(\hat{\rho} - 1)}{\hat{s}_{\rho,\tau}}, \]
where
\[ s^2_{\hat{p}, \tau} = \frac{\sum_{t=1}^{n} (y_t - \hat{\tau}_1 - \hat{\phi} y_{t-1})^2}{n \sum_{t=1}^{n} y_t^2 - \left( \frac{1}{n} \sum_{t=1}^{n} y_t \right)^2}. \]

The asymptotic distribution of this statistic is also more complicated:
\[
\frac{(\hat{p} - 1)}{s_{\hat{p}, \tau}} \xrightarrow{d} \int_0^1 L_\alpha \, dL_\alpha - L_\alpha(1) \int_0^1 L_\alpha \, ds - \left( \int_0^1 L_\alpha \, ds \right)^2 \sqrt{\frac{\int_0^1 L_\alpha^2 \, ds}{[L_\alpha](1)}}.
\]
CHAPTER 4

BOOTSTRAP FOR TIME SERIES

Classical bootstrap was introduced by Efron Bradley in 1979 [13]. It is a non-parametric method that helps to elicit many features of the statistic by approximating its distribution.

The main idea of it is the following. Suppose we are interested in the properties of a certain statistic $\theta$ from a given sample $X_1, X_2, \ldots, X_n$. By drawing randomly with replacement we obtain $B$ bootstrap samples: $X'_1, X'_2, \ldots, X'_n$, $i = 1, 2, \ldots, B$. Based on them we calculate the estimators $\hat{\theta}_i$, $i = 1, 2, \ldots, B$, which are used for the approximation of the probability distribution of $\theta$. The theoretical background for this method and a description of its advantages can be found in [13], [14], and [3].

It should be mentioned that the method described above assumes the data in the sample to be independent and identically distributed. That makes it useless to apply classical bootstrap in cases when dependence is critical, e.g. regression analysis, time series, etc. However, certain modifications can adjust bootstrap to be applicable for the time series analysis. Below we describe and compare some of them. More detailed description can be found in [6].

4.1 Block Bootstrap

The key-feature of this method comes from the endeavor to keep the dependence structure of the data unchanged. Roughly speaking, for the series $X_1, X_2, \ldots, X_n$ we are re-sampling blocks of the type $X_{t+1}, X_{t+2}, \ldots, X_{t+m}$, not single observations.

4.1.1 Description of the Bootstrapping Procedure

Consider blocks of the consecutive observations $Y_t = (X_{t-(m-1)}, \ldots, X_t), t = m, \ldots, n$ being chosen from the original sample of size $n$: $X_1, X_2, \ldots, X_n$. Here $m$ is the number of consecutive random variables $X_t$ that depends on, e.g. if $\theta = \text{Cov}(X_t, X_{t-1})$, then $m = 2$. As a result we will come up with $n - (m - 1)$ vectors:

\[
\begin{align*}
Y_m &= (X_1, \ldots, X_m), \\
Y_{m+1} &= (X_2, \ldots, X_{m+1}), \\
& \quad \ldots, \\
Y_n &= (X_{n-(m-1)}, \ldots, X_n).
\end{align*}
\]
Then we choose some block-length parameter $l > 1$, and construct $n - (m - 1) - (l - 1)$ overlapping blocks of the length $l$:

\[
\mathcal{Y}_m = (Y_m, \ldots, Y_{m+l-1}), \\
\mathcal{Y}_{m+1} = (Y_{m+1}, \ldots, Y_{m+l}), \\
\vdots \\
\mathcal{Y}_{n-(l-1)} = (Y_{n-(l-1)}, \ldots, Y_n).
\]

(4.1.2)

If $n - (m - 1)$ is not a multiple of $l$, then we sample $k = \lceil \frac{n - (m - 1)}{l} \rceil + 1$ blocks but use only a portion of the last one for the total number of vectors to be $n - (m - 1)$.

Finally, we sample $k = \frac{n - (m - 1)}{l}$ of these blocks with replacement to get a sample of blocks:

\[
\mathcal{Y}_{b_1}, \mathcal{Y}_{b_2}, \ldots, \mathcal{Y}_{b_k},
\]

where $\mathcal{Y}_{b_i} = (Y_{b_i}, \ldots, Y_{b_i+l-1})$, and every $b_i$ is i.i.d. discrete uniform on the interval $[m, n - (l - 1)]$.

However, the sample 4.1.3 is not what we call a true bootstrap sample, in the sense that we cannot get the bootstrap estimate $\hat{\theta}$ by using plug-in rule as we do in case of classical bootstrap.

The calculation of the block bootstrap estimator $\hat{\theta}^*$ is based on the fact that in many cases the estimator $\hat{\theta}$ can be written as a smooth functional of the empirical cdf of the $m$-dimensional blocks of $X_i$, $i \in \mathbb{Z}$, i.e., $\hat{\theta} = T(F_n^{(m)})$, where $F_n^{(m)}(\cdot) = \frac{1}{n-m-1} \sum_{t=m}^{n-1} 1\{Y_t \leq \cdot\}$.

Conformably the block bootstrapped estimator is defined as

\[
\hat{\theta}^* = T(F_n^{(m)*}),
\]

where $F_n^{(m)*}(\cdot) = \frac{1}{n-m-1} \sum_{t=1}^{k} \sum_{i=h}^{h+l-1} 1\{Y_{t+i} \leq \cdot\}$.

4.1.2 Specific Features of the Method

Block bootstrap is applicable to wide range of stationary processes. No assumptions about original method are made. However, some subtle issues should be mentioned.

Because of the way a bootstrapped sample is being constructed, the block bootstrap estimator $\hat{\theta}^*$ cannot be calculated by a plug-in rule. The method requires the estimate being written as a smooth functional of the marginal cdf of the blocks of observation, what often makes the computations very inconvenient.

Talking about data vectorization, it should be also mentioned that the parameter $m$ affects the structure of the bootstrapped samples critically. So-called naive bootstrap that uses $m = 1$ produces
artificial data changes by creating data points not presented in the original sample.

4.1.3 Choosing the Optimal Parameters

As we have seen, two parameters must be defined when using block bootstrap.

The parameter of vectorization - \( m \) needs to be equal the number of observations the parameter of interest \( \theta \) depends on. However, in some cases it might be impossible, for instance when estimating parameter of the MA(1) which depends on entire process, e.g. \( m = \infty \).

In certain cases for the sake of simplicity one may use a naive-bootstrap, i.e., set \( m = 1 \). However, the results will be less trustworthy (see [7]).

Another parameter - the block-length \( l \) is more important than \( m \). It is also more difficult to choose an optimal \( l \), because it depends on many things: the process the data came from, the statistic of interest \( \theta \), the purpose of the bootstrap (bias, variance or distribution estimation), etc. However, it is shown in [6] that asymptotically the MSE-optimal block-length for variance and bias estimation has form:

\[
\text{l}_{\text{opt}} = C \cdot n^{1/3},
\]

where \( C \) is a constant (see [8] for the proposed estimation procedure). More specific estimates of \( l \) depend on the original sample size and turn out to be very complicated, while the efficiency does not increase much.

4.2 AR-sieve Bootstrap

The AR-sieve bootstrap method is based on the assumption that the data are generated by invertible autoregressive process of order infinity (AR(\( \infty \))):

\[
(4.2.1) \quad X_t = \mu_X + \sum_{j=1}^{\infty} \phi_j (X_{t-j} - \mu_X) + \epsilon_t, \quad (t \in \mathbb{R}),
\]

where \( \mu_X = \text{E}X_t \), \( (\epsilon_t) \sim \text{i.i.d.} \) random variables with \( \text{E}\epsilon_t = 0 \) and \( \epsilon_t \) are independent of \( X_s \), for \( s < t \). This definition is valid if \( \text{E}\epsilon_t^2 < \infty \) and \( \sum_{j=1}^{\infty} \phi_j^2 < \infty \).

Bootstrapped samples are created by sieve approximation of (4.2.1) with the AR\((p)\) - autoregressive process of a selected order \( p \). In other words we simulate our data with the model estimated from the original process.
4.2.1 Description of the Bootstrapping Procedure

Before creating a bootstrapped sample we need to know the estimates for the parameters of (4.2.1).

First of all we estimate the order of the AR process \( \hat{p} \), because infinite series are inappropriate for practical computations. That can be done by exploring properties of the partial autocorrelation or by minimizing AIC statistic (see [5]).

The \( \mu_X \) is estimated as a simple mean: \( \hat{\mu}_X = \frac{1}{n} \sum_{t=1}^{n} X_t \), autoregressive coefficients (\( \phi_1, \ldots, \phi_p \)) can be obtained by using the Yule-Walker's algorithm.

Finally the distribution \( F_t \) of the i.i.d. innovations is estimated as follows:

\[
\hat{F}_t(x) = \hat{P}\{\varepsilon_t \leq x\} = \frac{1}{n-\hat{p}} \sum_{t=\hat{p}+1}^{n} 1[R_t-\hat{R} \leq x],
\]

where \( R_t = X_t - \sum_{j=1}^{\hat{p}} \hat{\phi}_j X_{t-j} \), and \( \hat{R} = \frac{1}{n-\hat{p}} \sum_{t=\hat{p}+1}^{n} R_t \).

The bootstrap sample is being generated from the process

\[
X^*_t = \hat{\mu}_X + \sum_{j=1}^{\hat{p}} \hat{\phi}_j (X^*_{t-j} - \hat{\mu}_X) + \varepsilon_t, \quad (t \in \mathbb{R}),
\]

where \( \varepsilon_t \) are i.i.d. \( \sim \hat{F}_t \). It is also recommended to stabilize series by setting \( (X^*_{-m}, \ldots, X_{-m+\hat{p}-1}) = (\hat{\mu}_X, \ldots, \hat{\mu}_X) \).

4.2.2 Specific Features of the Method

The AR-sieve bootstrap depends on the assumption that the data comes from an AR(\( \infty \)) model. That makes the method inapplicable for the processes that cannot be expressed by (4.2.1).

However, for the linear invertible processes this method turns out to be very efficient. Practically, the results of AR-sieve bootstrap are more accurate then those obtained from block bootstrap. The AR-sieve bootstrap also automatically adapts to the decay of the underlying structure by producing more accurate estimates for the processes with long range dependence (see [6]).

4.2.3 Choosing the Optimal Parameter

AR-sieve bootstrap basically depends only on one parameter - the order of autoregressive process \( \hat{p} \). It was shown in [24] that the minimum AIC selection procedure gives the optimal \( \hat{p} \) for AR(\( \infty \))
models. This criterion suggests higher order of AR-model for processes with stronger dependence. This is the only attribute that is taken into account. No methods are currently developed for adjusting the order of AR approximation to the statistic being bootstrapped or the purpose of the bootstrap (bias, variance or distribution estimation).

4.3 Other Methods

Here we just briefly overview some other specific bootstrap methods applicable for time series. More detailed description can be found in [6].

4.3.1 Variable Length Markov Chain (VLMC)
Sieve Bootstrap

This method is used mainly for the categorical data generated by a Markov process of high order:

\[ P\{X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \ldots\} = P\{X_t = x_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \ldots X_{t-l} = x_{t-l}\} \]

where the number of lagged values \( l \) is variable and in general is a function of the past values: \( l = l(x_{t-1}, x_{t-2}, \ldots) \).

We do not provide the context algorithm for this method, because of its complexity which is caused by the variable length of the process’ memory (see [6] for more details).

4.3.2 Local Bootstrap for Conditional Mean Estimates

This method is based on the independent re-sampling and is aimed to estimate nonparametric statistics with rate of convergence slower than \( 1/\sqrt{n} \), e.g. conditional expectation \( \theta(x) = E(X_t | X_{t-1} = x) \).

The following method is being used: the conditional cdf of \( X_t \) should be estimated. A bootstrapped sample is constructed by re-sampling the data independently within the local bandwidth \( b \), which is usually being selected the same as the bandwidth of the estimator \( \hat{\theta} \).

Independent resampling is allowed if the distribution of the estimator \( \hat{\theta} \) depends only on the marginal distribution of \( X_t \), the conditional distribution of \( X_t \) given \( X_{t-1} \) and the known form of the estimator. However, this is true only asymptotically and more complicated methods must be applied for the finite samples data. The consistency of the local bootstrap is also proved for a
short-range dependent processes.

4.4 Residual Subsampling for Unit Root Test

The approach of residual subsampling is not in fact a version of a bootstrap, though in principle it is somewhat similar. We discuss it here because it was used for the same unit root testing problem in [17], and we want to compare our results with theirs in order to see which method performs better. An excellent reference on subsampling is [20].

Here we briefly describe the idea of applying subsampling to the unit root tests in model (3.5.1), though it also can be used for the model (2.4.2) and other more complex models.

Suppose we have a series $X_1, X_2, \ldots, X_n$ of length $n$, which follows the process

$$X_k = \rho X_{k-1} + \varepsilon_k,$$

with $\varepsilon_k \sim S_\alpha(1, 0, 0)$.

To test the hypothesis $H_0 : \rho = 1$ one may use the following algorithm: first, the Least Square Estimate of the model's parameter $\rho$ is calculated. Then, using this estimate, we calculate the residuals: $\hat{\varepsilon}_k = X_k - \hat{\rho}_n X_{k-1}$. The centered residuals $\tilde{\varepsilon}_k = \hat{\varepsilon}_k - \frac{1}{n-1} \sum_{k=2}^{n} \hat{\varepsilon}_k$ are used to generate $n - b$ subsamples of length $b$, that follow the null-hypothesis:

$$\left\{ X_j(k) = \sum_{j=1}^{b} \tilde{\varepsilon}_{k+j}, j = 1, \ldots, b \right\},$$

for every $k = 1, \ldots, n - b$.

For every subsample $\{X_j(k)\}$, the LSE estimator $\hat{\rho}_b(k)$ is computed ($k = 1, \ldots, n - b$). The set of $\hat{\rho}_b(k)$ is then used to estimate the empirical distribution of the test statistic $T_{b,k} = b(\hat{\rho}_b(k) - 1)$, which finally allows to test the hypothesis $H_0 : \rho = 1$ versus the one-sided alternative $H_1 : \rho < 1$ using the decision rule: reject $H_0$ if $n(\hat{\rho}_n - 1) < q_{T_b}(\alpha)$, where $q_{T_b}(\alpha)$ is the $\alpha$-th quantile of the empirical distribution of the $T_{b,k}$.

When using this method, one should carefully choose the parameter $b$ that determines the size of the subsamples. Theory implies merely that it must be chosen so that $b \to \infty$ and $b/n \to 0$, as $n \to \infty$. More detailed exploration of how the choice of $b$ affects the test results can be found in [17].
4.5 Comparison and Conclusions

The block bootstrap is the most general method applicable to the wide range of processes. The main disadvantage of this method is that due to the vectorization of the data, the plug-in rule cannot be applied to estimate the parameter. More complicated methods should be used instead. To have an idea of the accuracy of the method notice that for the $\hat{\theta} = \bar{X}_n$ we have

$$E \left( \left[ n \cdot \text{Var}^*(\hat{\theta}) - n \cdot \text{Var}(\hat{\theta}) \right]^2 \right) \sim O_P \left( n^{-2/3} \right).$$

AR-sieve bootstrap can be applied for the finite order and invertible processes only. However it has many advantages: it is easy to implement; the plug-in rule can be applied; the accuracy increases as the degree of dependence increases.

To illustrate the last statement we consider two cases for the $\hat{\theta} = \bar{X}_n$: when the autoregressive parameters $\phi_j$ decay like $\phi_j \leq C \cdot j^{-v}, v > 2$ then

$$n \cdot \text{Var}^*(\hat{\theta}) - n \cdot \text{Var}(\hat{\theta}) = O_P \left( n^{-\frac{3v}{2v - 4}} \right).$$

In case when $\phi_j \leq C \cdot e^{-k}$ we have

$$n \cdot \text{Var}^*(\hat{\theta}) - n \cdot \text{Var}(\hat{\theta}) = O_P \left( n^{-1/2 + k} \right).$$

The VLMC-sieve bootstrap is applicable mainly to the processes of the categorical data. For such data the results are much better than what the block bootstrap gives. The main drawback of this method is that the bootstrapping procedure is quite complicated.

The local bootstrap makes sense for very special cases only. The efficiency of the method is close to that of the block bootstrap. When strong dependence is encountered in the data, the block bootstrap is expected to perform much better (see Bühlmann (2001), section 6.2).

Based on these conclusions we decided mainly to use AR-sieve bootstrap for the unit root tests. It should be mentioned that the residual subsampling method also seems to be quite robust and applicable to different data models. However it strongly depends on the parameter $b$. The procedures for choosing the optimal value for $b$ in the context of unit root tests were recently described in Jach and Kokoszka (2002), some practical applications were also presented. We use their results to compare the performance of the tests in Chapter 4.
CHAPTER 5

BOOTSTRAP UNIT ROOT TESTS FOR HEAVY TAILED OBSERVATIONS

In this chapter we describe the experiments that were carried out to investigate the application of bootstrap to the unit root testing. The detailed description of several bootstrap algorithms and the test procedures are also given.

First, we evaluated the bootstrap methods using computer-simulated data. For that purpose two different AR(1) models were considered: with a constant term (2.4.2) and without a constant term (2.4.1). To evaluate the performance of the tests, their size and power were estimated.

Moreover, we also investigated the methods for improving the results of the test, by reducing the bootstrap sample size. The asymptotically optimal sample size was proposed.

For the comprehensive overview, series of different lengths were included, and also different values for the parameters of the underlying stable distribution were used. We also used series of different lengths to verify the convergence of the test statistic when the sample size is increasing.

Finally, in order to demonstrate the applicability of the described methods, we applied the tests to financial time series (yield curves on corporate bonds).

5.1 Model 1: AR(1) Model Without a Constant Term or Time Trend

First we consider the simplest model which describes the data as an AR(1) process:

\[ X_j = \rho X_{j-1} + \varepsilon_j, \]

where \( \varepsilon_j \) are i.i.d. \( \sim S_\alpha(1,0,0) \), \( j = 1, 2, \ldots, n \).

Our goal is to construct an algorithm for testing the null-hypothesis \( H_0 : \rho = 1 \), and to explore its properties. It was decided to use the AR-sieve bootstrap because it allows to estimate the distribution of the parameter of interest by using the plug-in rule, which is impossible in case of the block bootstrap. For the AR-sieve bootstrap we also tried the approach based on the confidence interval of the estimated parameter. Below we describe the methods we used and compare their results.
5.1.1 Residual-based AR-Sieve Bootstrap Algorithm

Here we describe the AR-sieve bootstrap procedure that allows to test the hypothesis $H_0$ by using the density estimates for the distribution of the unknown parameter $\hat{\rho}$.

First, we estimate the coefficient $\rho$ from the original series $X_1, X_2, \ldots, X_n$. The OLS estimator is calculated as:

$$\hat{\rho}_n = \frac{\sum_{j=1}^{n} X_{j-1}X_j}{\sum_{j=1}^{n} X_{j-1}^2}.$$

The estimator $\hat{\rho}_n$ is then used to estimate the residuals:

$$\hat{e}_j = X_j - \hat{\rho}X_{j-1},$$

where $j = 2, 3, \ldots, n$. Notice that $\hat{e}_1$ remains undefined, but it is not needed in our algorithm. The residuals are centered as follows:

$$\tilde{e}_j = \hat{e}_j - \frac{1}{n-1} \sum_{j=1}^{n} \hat{e}_j.$$

Since we are interested in estimating the density of $\rho$ under the null hypothesis $H_0 : \rho = 1$, the following formula is used to generate each bootstrap sample:

$$X^*_j = X^*_{j-1} + e^*_j,$$

where $j = 2, 3, \ldots, n$, $X^*_1 = X_1$, and $e^*_j$ are randomly sampled with replacement from the set of the centered residuals $\tilde{e}_j$, $j = 2, 3, \ldots, n$.

Each bootstrap sample $\{X^*_j, j = 1, 2, \ldots, n\}$, gives the bootstrap estimator:

$$\hat{\rho}^* = \frac{\sum_{j=1}^{n} X^*_{j-1}X^*_j}{\sum_{j=1}^{n} X^*_j^2}, i = 1, 2, \ldots, B,$$

where $B$ is the number of the bootstrap samples.

The set of $\hat{\rho}^*_i$ enables us to estimate the density of the $\hat{\rho}$. Since we are testing the hypothesis $H_0 : \rho = 1$ versus the one-sided alternative $H_a : \rho < 1$, what we need is the $q$-th percentile of this distribution, which can be obtained by ordering the bootstrap estimates $\hat{\rho}^*$ and taking the one with index $\frac{Bq}{100}$. Hence, the decision rule is the following: reject $H_0 : \rho = 1$, if $\hat{\rho} < \hat{\rho}^*_q$. 


5.1.2 The Rate of Convergence of the Test Statistic

Before testing the hypothesis $H_0 : \rho = 1$ we decided to assess the rate of convergence of the statistic $T_n = n(\hat{\rho}_n - 1)$ to its asymptotic distribution. This will give us some idea of the shape of the density of $T_n$ for the finite $n$, so that we can see if the bootstrap distribution is a good approximation to the sampling distribution of $T_n$.

The results of the following experiment are reported: 10000 random samples were simulated according to (5.1.1) with $\rho = 1$ ($\alpha = 1.75$). For each sample, the statistic $T_n = n(\hat{\rho}_n - 1)$ was computed. Probability density function $f_{T_n}(t)$ was estimated on 50 equally spaced points by using a non-parametric smoothing method. Gaussian window was used in the computations. Fraction of the window width that the $x$ values are to be extended by was chosen to be 0.75. This method was used to estimate all densities discussed in this section.

Since the statistic $T_n$ depends on the observed series length $n$, we performed this experiment for different $n = 10, 30, 50, 60, 100, 125, 200, 250, 300, 375, 450, 500$.

Graphical comparison of the density estimators is given below (Figures 5.1 - 5.11). Each figure also contains the graph of $\hat{f}_{T_{500}}(t)$, which is considered a good approximation to the limit distribution.

![Graph](image)

Fig. 5.1: Comparison of density estimators $\hat{f}_{T_{10}}(t)$ and $\hat{f}_{T_{500}}(t)$

As we can see from these graphs, the probability density functions of the $T_n$ are certainly converging to a certain limit when $n$ is increasing. This confirms the theoretical finding of [9] that the limiting distribution of $T_n$, exists (see Section 2.5.1). Density functions for statistics $T_n$ calculated from series of length 450 and 500 are almost indistinguishable (see Figure 5.11) which makes $\hat{f}_{T_{500}}(t)$
Fig. 5.2: Comparison of density estimators $\hat{f}_{T_{30}}(t)$ and $\hat{f}_{T_{500}}(t)$

Fig. 5.3: Comparison of density estimators $\hat{f}_{T_{50}}(t)$ and $\hat{f}_{T_{500}}(t)$

Fig. 5.4: Comparison of density estimators $\hat{f}_{T_{60}}(t)$ and $\hat{f}_{T_{500}}(t)$
Fig. 5.5: Comparison of density estimators $\hat{f}_{T_{250}}(t)$ and $\hat{f}_{T_{500}}(t)$

Fig. 5.6: Comparison of density estimators $\hat{f}_{T_{125}}(t)$ and $\hat{f}_{T_{500}}(t)$

Fig. 5.7: Comparison of density estimators $\hat{f}_{T_{500}}(t)$ and $\hat{f}_{T_{500}}(t)$
Fig. 5.8: Comparison of density estimators $\hat{f}_{T_{250}}(t)$ and $\hat{f}_{T_{500}}(t)$

Fig. 5.9: Comparison of density estimators $\hat{f}_{T_{300}}(t)$ and $\hat{f}_{T_{500}}(t)$

Fig. 5.10: Comparison of density estimators $\hat{f}_{T_{375}}(t)$ and $\hat{f}_{T_{500}}(t)$
a fairly good approximation of the limiting distribution.

We also investigated by simulation the existence of the limiting distribution of the \( T_n^* = n(\hat{\rho}_n^* - 1) \), where \( \hat{\rho}_n^* \) is a least square estimator, calculated from the bootstrapped sample of size \( n \). However, the distribution of the \( T_n^* \) strongly depends on the original sample, i.e., density estimators will be different for every particular realization. Because it is impossible to analyze all possible cases, we have taken only several estimators \( \hat{f}_{T_n^*}(t) \) to get an idea about their behavior. Figure 5.12 represents density estimator for the statistic \( T_{500} \) (which is considered a good approximation to the limiting density for \( \hat{f}_{T_n^*} \)) and ten graphs of the density estimators for the statistic \( T_{500}^* \), calculated from the bootstrapped samples of size \( n = 500 \). Original series of length 500 were the realizations of the process 5.1.1 with \( \rho = 1 \). Each \( \hat{f}_{T_{500}^*} \) was estimated from 2000 \( T_{500}^* \) estimates.

As we can see from Figure 5.12, the density estimators of the \( T_{500}^* \) calculated from the bootstrapped samples are in most cases close to \( \hat{f}_{T_{500}}(t) \). A close examination shows, however, that the bootstrap densities look more similar to densities \( T_m \) for \( m \) smaller than 500 (see Figures 5.1 - 5.11). This suggests that, at least when \( H_0 \) is true, the bootstrap distribution of \( m(\rho_m^* - 1) \) exists, where \( m = m(n) \leq n \) is a function of the sample size \( n \).

5.1.3 Size and Power of the Test

To estimate the size and power of the test, we applied it to the simulated series of the process (5.1.1) with different values of \( \alpha = 1.75, 1.5, 1.25, 1.1 \) and \( \rho: 1, 0.99, 0.95, 0.9, 0.8, 0.5 \). The
power was estimated as a ratio of the number of rejections to the total number of experiments
\( (N = 10000) \). Decision about rejecting the null hypothesis \( (H_0 : \rho = 0) \) every time was made based
on the distribution of \( T_n \) estimated from 2000 bootstrap samples. Experiments were performed for
series of different lengths: 500, 375, 250, 125, 60. The results are presented in Table 5.1.

**Table 5.1:** Series (5.1.1) with \( \alpha = 1.75 \) were used to estimate the percentage of rejections of \( H_0 : \rho = 1 \)
for different values of the real parameter \( \rho \) (nominal size = 5%). 10000 of experiments
were performed with 2000 of bootstrap replications each.

<table>
<thead>
<tr>
<th>Length of the series</th>
<th>1</th>
<th>0.99</th>
<th>0.95</th>
<th>0.9</th>
<th>0.8</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>5.11</td>
<td>10.38</td>
<td>22.54</td>
<td>48.28</td>
<td>92.56</td>
<td>100</td>
</tr>
<tr>
<td>125</td>
<td>4.50</td>
<td>12.11</td>
<td>50.26</td>
<td>93.38</td>
<td>99.99</td>
<td>100</td>
</tr>
<tr>
<td>250</td>
<td>4.64</td>
<td>17.21</td>
<td>93.49</td>
<td>99.99</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>375</td>
<td>4.72</td>
<td>24.98</td>
<td>99.84</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>500</td>
<td>4.74</td>
<td>35.46</td>
<td>99.99</td>
<td>99.99</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The information from Table 5.1 can be visualized graphically (see Figure 5.13). As we can see,
series of the longer length result in a more powerful test. For instance if the length of the series is
the test will be able to distinguish real $\rho = 0.95$ to be different from 1 with probability 0.9349.

For series of the length $n = 60$ this probability is only 0.2254 (see Table 5.1).

\begin{figure}
\centering
\begin{tabular}{ccc}
\includegraphics[width=0.3\textwidth]{estimate_for_series_of_length_500} & \includegraphics[width=0.3\textwidth]{estimate_for_series_of_length_375} & \includegraphics[width=0.3\textwidth]{estimate_for_series_of_length_250} \\
\includegraphics[width=0.3\textwidth]{estimate_for_series_of_length_125} & \includegraphics[width=0.3\textwidth]{estimate_for_series_of_length_60} & \\
\end{tabular}
\caption{Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$, for the series of different lengths with $\alpha = 1.75$. Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.}
\end{figure}

We also want to compare the performance of the AR-sieve bootstrap with the subsampling methods (using the results of the numerical simulations from [17]). Figure 5.14 shows the graphical estimation of the power of the both tests for the model (5.1.1), with $\alpha = 1.5$.

As we can see, the curves are quite close to each other, except the case for series of length $n = 250$. 

Fig. 5.14: Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$, for the series following model (5.1.1) with $\alpha = 1.5$, by using AR-sieve bootstrap (solid line) and the subsampling (dotted line). Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

5.1.4 Performance of the Test for Different Stability Indexes $\alpha$

We are also interested in how the performance of the test depends on the index of stability $\alpha$ of the distribution of $\varepsilon_j \sim S_{\alpha}(1,0,0)$ (5.1.1). To investigate that dependence we compare type-I error (percentage of the rejections of the true null-hypothesis) of the series that were generated using stable distribution with different $\alpha$. Namely, we performed 10000 experiments for series of length $n = 500, 375, 250, 125, 60$, generated using stable distribution $S_{\alpha}(1,0,0)$ with $\alpha = 1.75, 1.5, 1.25, 1.1$. In each experiment we estimated the value of the parameter $\rho$ and tested the hypothesis $H_0 : \rho = 1$ based on the bootstrap estimation (2000 bootstrap samples) of the distribution of $\hat{\rho}$. The results of the experiments are given in Table 5.2.

The graphical representation of how the type-I error for series of different lengths depends on the index of stability is shown on the Figure 5.15.
Table 5.2: Type-I errors for the model (5.1.1) (percentage of the rejection the null-hypothesis $H_0: \rho = 1$) for the series of different sizes $n$ and different values of the index of stability $\alpha$.
Based on 10000 experiments. Nominal size = 5%

<table>
<thead>
<tr>
<th>Length of the series ($n$)</th>
<th>Index of stability ($\alpha$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>60</td>
<td>4.39</td>
</tr>
<tr>
<td>125</td>
<td>4.02</td>
</tr>
<tr>
<td>250</td>
<td>3.83</td>
</tr>
<tr>
<td>375</td>
<td>4.23</td>
</tr>
<tr>
<td>500</td>
<td>3.99</td>
</tr>
</tbody>
</table>

As we can see from Table 5.2 and Figure 5.15, for the longer series the type-I error is smaller. One can also notice that the empirical size is closer to the nominal size (0.05) as the stability index $\alpha$ approaches 2. An intuitive explanation for this phenomenon might be due to the fact that for $\alpha = 2$ the stable distribution becomes Gaussian $N(0, 2)$, while for $\alpha = 1$ it is the Cauchy distribution (see Section 2.2), which is much less “regular” than the normal distribution.

Such a dependence on $\alpha$ makes the test non-robust, because in practice no parameters of the stable distribution are known. On the other hand, for most combinations of $n$ and $\alpha$, the empirical sizes are between 4% and 6%. This means that despite its shortcomings the test is very precise.

Also, in comparison, for instance, with the subsampling methods, the bootstrap seems to give more accurate results, see Figure 5.16 (the results of the subsampling test are taken from [17]).

5.1.5 Changing the Bootstrap Sample Size

In view of the results of [1], [2], [16], and [18], we have performed a number of experiments to estimate type-I error when the bootstrap sample size is smaller or larger than the actual length of the observed series. We did this because these theoretical contributions show that in several contexts one should choose the bootstrap sample size $m$ smaller that the original sample size $n$ if the errors of the model have infinite variance: $\text{Var}(\varepsilon_t) = \infty$.

Based on that, we run 10000 simulations for every sample of size $n = 60, 125, 250, 375, 500$. Decision about rejecting the null hypothesis ($H_0: \rho = 0$) every time was made based on empirical distribution of $T_n$ estimated from 2000 of bootstrap samples of size $m$ which is a specified percentage of $n$. 
Fig. 5.15: Type-I error (percentage of the rejections of the true null-hypothesis $H_0: \rho = 1$ for the model 5.1.1) for the series of different lengths ($n = 500, 375, 250, 125, 60$) and different values of the stability index ($\alpha = 1.75, 1.5, 1.25, 1.1$) used for the stable distribution. Based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation.

The comparison of type-I errors for the reduced bootstrap sample sizes is given in Table 5.3.

As we can see from this table, the type-I error tends to increase when the bootstrap series are shorter than the original samples. For the series of length $n = 500$ the optimal bootstrap sample
Fig. 5.16: Comparison of the type I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ for the model (5.1.1) with $\alpha = 1.5$) for the AR-sieve bootstrap and the subsampling methods. Bootstrap results are based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation. Subsampling results are based on $R = 5000$ replications.

Table 5.3: Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) with the reduced bootstrap samples for model (5.1.1) with $\alpha = 1.75$. Based on 10000 experiments with 2000 bootstrap simulations.

<table>
<thead>
<tr>
<th>The original length of the series</th>
<th>Length of the bootstrap samples (percents of the original length)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>100% 95% 90% 85% 80%</td>
</tr>
<tr>
<td>60</td>
<td>5.11 5.46 5.80 6.13 6.55</td>
</tr>
<tr>
<td>125</td>
<td>4.50 4.83 5.04 5.20 5.55</td>
</tr>
<tr>
<td>250</td>
<td>4.64 4.61 4.80 5.01 5.26</td>
</tr>
<tr>
<td>375</td>
<td>4.72 4.91 4.94 5.12 5.29</td>
</tr>
<tr>
<td>500</td>
<td>4.74 4.83 4.90 5.07 5.20</td>
</tr>
</tbody>
</table>

size $m$ is around 425 (85% of the original size), for shorter series this percentage is increasing: $m \approx 0.9n$ for $n = 125$, and finally for $n = 60$ reducing the bootstrap sample size does not lead to any improvement.

We also decided to investigate a different approach, i.e., to increase the length of the bootstrap series. 10000 experiments, each with 2000 bootstrap samples were used to to estimate the type I
error rate for the increased bootstrap sample size. The results are given in Table 5.4.

**Table 5.4:** Type-I errors (percentage of the rejections of the true null-hypothesis \( H_0 : \rho = 1 \)) with the increased bootstrap samples for the model (5.1.1) with \( \alpha = 1.75 \).

<table>
<thead>
<tr>
<th>Length of the series</th>
<th>Length of the bootstrap samples (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>5.11</td>
</tr>
<tr>
<td>125</td>
<td>4.50</td>
</tr>
<tr>
<td>250</td>
<td>4.64</td>
</tr>
<tr>
<td>375</td>
<td>4.72</td>
</tr>
<tr>
<td>500</td>
<td>4.74</td>
</tr>
</tbody>
</table>

We can see that for the increased bootstrap sample size the type I error rate becomes smaller and moves away from the nominal level (5%), i.e., the test tends to accept the null hypothesis \( H_0 : \rho = 1 \) more often. The graphical representation of Tables 5.4 and 5.3 is given on the Figure 5.17.

![Graph of Type-I error](image)

**Fig 5.17:** Type-I error (percentage of the rejections of the true null-hypothesis \( H_0 : \rho = 1 \)) for the model (5.1.1) with \( \alpha = 1.75 \) for the series of different lengths \( n = 500, 375, 250, 125, 60 \) versus the percentage the bootstrap sample compares to the original series length. Results are based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation. Nominal level = 5%.

It also looks like the dependence of the type-I error and the size of the bootstrap samples, at least for longer series, is close to linear line. To verify this property we performed the same estimation of
the type-I error for the for the series with index of stability $\alpha = 1.5$. The result of these experiments are given in Table 5.5. The graphical representation of Table 5.5 is given on the Figure 5.18.

Table 5.5: Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) with the increased bootstrap samples for the model (5.1.1) with $\alpha = 1.5$.

<table>
<thead>
<tr>
<th>Length of the series</th>
<th>80%</th>
<th>85%</th>
<th>90%</th>
<th>95%</th>
<th>100%</th>
<th>110%</th>
<th>120%</th>
<th>130%</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>5.29</td>
<td>5.11</td>
<td>4.97</td>
<td>4.79</td>
<td>4.67</td>
<td>4.40</td>
<td>4.23</td>
<td>3.96</td>
</tr>
<tr>
<td>375</td>
<td>5.43</td>
<td>5.42</td>
<td>5.25</td>
<td>5.03</td>
<td>4.84</td>
<td>4.68</td>
<td>4.42</td>
<td>4.26</td>
</tr>
<tr>
<td>500</td>
<td>5.27</td>
<td>5.02</td>
<td>4.82</td>
<td>4.79</td>
<td>4.78</td>
<td>4.53</td>
<td>4.35</td>
<td>4.33</td>
</tr>
</tbody>
</table>

Looks like the error rate obtained from series of different length in fact follow the same pattern.

As a practical guidance, we recommend to use use bootstrap sample size $m = 0.85n$ for series of length $200 \leq n \leq 500$ and $\alpha$ in the middle of the range $(1, 2)$.

5.1.6 CI-based Bootstrap Algorithm

In this section we propose an approach which is motivated by the work of [19]. The idea of the method is to find a level $(1 - q)$ confidence interval (CI) of the form $(-1, \hat{u})$ for the parameter $\rho$. A size $q$ test then rejects $H_0$ if $\hat{u} < 1$. To find the above confidence interval, we consider the differenced series

$$Y_t = X_t - X_{t-1}.$$ 

Observe that

$$Y_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j},$$

where

$$c_j = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{if } j > 0 \end{cases}, \text{ if } \rho = 1$$

and

$$c_j = \begin{cases} 1, & \text{if } j = 0 \\ (\rho - 1)\rho^{j-1}, & \text{if } j > 0 \end{cases}, \text{ if } |\rho| < 1.$$ 

If the $\varepsilon_t$ are iid $S_0(1,0,0)$, then Theorem 4.2 of [10] implies that

$$\hat{\rho}_Y = \frac{\sum_{j=1}^{n-1} Y_j Y_{j+1}}{\sum_{j=2}^{n} Y_j^2} - \frac{\sum_{j=0}^{\infty} c_j c_{j+1}}{\sum_{j=0}^{\infty} c_j^2} = \frac{\rho - 1}{2} = \rho_Y.$$
Now, if we can construct a level $(1 - q)$ confidence interval $(-1, \hat{u}_Y)$ for the $\rho_Y$, then $(-1, 2\hat{u}_Y + 1]$ is the corresponding interval for $\rho$.

In practice we use $\hat{\rho} = 2\hat{\rho}_Y + 1$ to compute the residuals

$$\hat{e}_t = X_t - \hat{\rho}X_{t-1}, \ t = 2, \ldots, n.$$**

Then, by drawing with replacement from the set $\{\hat{e}_t - \tilde{\epsilon}, \ t = 2, \ldots, n\}$, where $\tilde{\epsilon} = (n - 1)^{-1} \sum_{t=2}^n \hat{e}_t$, we get bootstrap residuals $\epsilon^*_2, \epsilon^*_3, \ldots, \epsilon^*_n$ and bootstrap observations $X^*_t = \hat{\rho}X^*_{t-1} + \epsilon^*_t$, where $i = 2, 3, \ldots, n$ and $X^*_1 = X_1$. Finally, after generating $B$ sets of the bootstrap series $\{X^*_1, X^*_2, \ldots, X^*_n\}$, $j = 1, 2, \ldots, B$, we compute $B$ estimates $\hat{\rho}_Y^*$, which are used to construct the confidence interval and test the hypothesis $H_0$ as described above.
5.1.7 The Performance of the CI-based Bootstrap

To see the CI-based bootstrap algorithm at work, we generated 10000 series following the model (5.1.1) with $\rho = 1$, i.e.,

$$X_t = X_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots, n,$$

where $\varepsilon_t \sim \text{iid } S_{1.75}(1,0,0)$, $X_0 = 0$, $n = 60,125,250,375,500$. For each series the hypothesis $H_0 : \rho = 1$ was tested by using the bounds of confidence intervals obtained from 1000 of bootstrap simulations. Table 5.6 summarizes the estimates of the type I error (percentage of rejections of the true null-hypothesis).

Table 5.6: CI-bootstrap estimates of the type I error (reject true $H_0$) for the AR(1) series of different length, assuming the model (5.1.1) with the $\alpha = 1.75$. The nominal size of the error = 5%.

<table>
<thead>
<tr>
<th>Length of the series</th>
<th>500</th>
<th>375</th>
<th>250</th>
<th>125</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I error rate (%)</td>
<td>5.18</td>
<td>4.12</td>
<td>3.2</td>
<td>2.13</td>
<td>1.39</td>
</tr>
</tbody>
</table>

We can see that the type I error rate of this test is close to the nominal only asymptotically, while for the finite samples (series of length $n < 500$) it tends to accept the null hypothesis ($H_0 : \rho = 1$) more often. We also compare the performance of this algorithm with the AR-sieve bootstrap in terms of type I error. The first row of Table 5.1 represents the type I error of the test using the AR-sieve bootstrap. By looking at the error rate of the CI-based method (Table 5.6) one can see that asymptotically ($n \geq 500$) their both results are close to the nominal level (5%), though for the short-length series the residual-based AR-sieve bootstrap performs much better than the CI-based bootstrap.

We also attempted to improve the test results by reducing the bootstrap sample size as we did for the AR-sieve bootstrap. For the model (5.1.1) with $\rho = 1$, and $\varepsilon_t \sim \text{iid } S_{1.75}(1,0,0)$, we generated 10000 series. The hypothesis $H_0 : \rho = 1$ was tested using the bounds of confidence intervals that were obtained from 1000 of bootstrap simulations. The results of the experiments are summarized in Table 5.7.

Unfortunately, the results from Table 5.7 do not show any particular improvement of the test performance when reducing the size of the bootstrap samples. The type I error rates do not seem to be approaching any limit. Also it should be noticed that asymptotically (series length $n = 500$)
Table 5.7: Type-I errors (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$) of the CI-based bootstrap tests with the reduced bootstrap samples. Based on 10000 experiments with 1000 bootstrap simulations for the model (5.1.1) with $\alpha = 1.75$. The nominal error rate is 5%.

<table>
<thead>
<tr>
<th>The original length of the series</th>
<th>Length of the bootstrap samples (percents of the original length)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>100%  95%  90%  85%</td>
</tr>
<tr>
<td>125</td>
<td>2.13  2.24  2.12  2.22</td>
</tr>
<tr>
<td>250</td>
<td>3.2   2.28  3.19  2.99</td>
</tr>
<tr>
<td>375</td>
<td>4.12  4.3   4.5   4.27</td>
</tr>
<tr>
<td>500</td>
<td>5.18  5.51  5.61  4.98</td>
</tr>
</tbody>
</table>

The result are not getting worse either, compare 5.18% for the bootstrap sample size $m = n = 500$, and 4.98% for the $m = .85n$.

Finally, it should be said that for the heavy tailed observations the CI-based bootstrap approach does not perform reasonably well. We did not explore the reasons of such poor results and decided to focus on the residual-based AR-sieve bootstrap methods described in the remaining sections of this chapter.

5.2 Model 2: AR(1) Model with a Constant Term but No Time Trend

In this case we fit the model with a constant term:

\[(5.2.1)\]  

\[X_j = \tau + \rho X_{j-1} + \epsilon_j,\]

where $\epsilon_j$ are i.i.d. $\sim S_\alpha(1,0,0)$, $j = 1, 2, \ldots, n$.

In order to test whether the process is a random walk ($X_j = X_{j-1} + \epsilon_j$), we have to deal with a composite hypothesis $H_0 : \rho = 1$, and $\tau = 0$. For that we need to estimate the joint distribution of the parameters $n(\rho - 1)$ and $\tau \sqrt{n}$.

Since the estimation of the bivariate distribution of the statistic $(\hat{\rho}, \hat{\tau})$ is not straightforward, we test the hypotheses $H_0 : \rho = 1$ and $\tau = 0$ using the statistic:

\[(5.2.2)\]  

\[\Phi_n = \frac{(n-3)(RSS_0 - RSS_1)}{2RSS_1},\]
where

\[(5.2.3)\]
\[RSS_1 = \sum_{k=2}^{n} (X_k - \hat{\rho}_n X_{k-1} - \hat{\tau}_n)^2 = \sum_{k=2}^{n} \varepsilon_k^2\]

and

\[(5.2.4)\]
\[RSS_0 = \sum_{k=2}^{n} (X_k - X_{k-1})^2.\]

As it was shown by [12], the decision rule for the \(H_0: \rho = 1\) and \(\tau = 0\) is to reject \(H_0\) when \(\Phi_n\)

is greater than a certain critical value. Tables of critical values are available for the random errors
\(\varepsilon_j \sim N(0, \sigma^2)\) (see e.g., [15]). Below we describe a bootstrap algorithm that allows to estimate the
distribution of the statistic \(\Phi_n\), without making any assumptions about the distribution of the \(\varepsilon_j\),
except that they have zero mean.

5.2.1 Bootstrap Algorithm

We start with the original series \(X_1, X_2, \ldots, X_n\). First the Least Square Estimates of the parameters \(\rho\) and \(\tau\) are computed:

\[
\hat{\rho}_n = \frac{n \sum_{j=1}^{n-1} X_j X_{j+1} - \sum_{j=1}^{n-1} X_j \sum_{j=1}^{n-1} X_{j+1}}{n \sum_{j=1}^{n-1} X_j^2 - \left(\sum_{j=1}^{n-1} X_j\right)^2}
\]

\[
\hat{\tau}_n = \frac{\sum_{j=1}^{n-1} X_{j+1} \sum_{j=1}^{n-1} X_j^2 - \sum_{j=1}^{n-1} X_j \sum_{j=1}^{n-1} X_j X_{j+1}}{n \sum_{j=1}^{n-1} X_j^2 - \left(\sum_{j=1}^{n-1} X_j\right)^2}
\]

These estimators are then used to calculate the residuals \(\hat{\varepsilon}_j\):

\[
\hat{\varepsilon}_j = X_j - \hat{\rho} X_{j-1} - \hat{\tau}, j = 2, 3, \ldots, n.
\]

The residuals are centered as follows:

\[
\hat{\varepsilon}_j = \varepsilon_j - \frac{1}{n-1} \sum_{j=1}^{n} \hat{\varepsilon}_j.
\]

Since we are interested in estimating the density of \(\Phi_n\) under the null hypothesis \(H_0: \rho = 1, \tau = 0\),
the following formula is used to generate each bootstrap sample:

\[
X_j^* = X_{j-1}^* + \varepsilon_j^*.
\]
where \( j = 2, 3, \ldots, n \), \( X_1 = X_1 \), and \( \varepsilon_i \) are randomly sampled with replacement from the set of centered residuals \( \xi_j, j = 2, 3, \ldots, n \).

Each bootstrap sample \( \{X_j, j = 1, 2, \ldots, n\} \), gives the bootstrap estimates \( \hat{\rho}_i^* \) and \( \hat{\tau}_i^* \) that are used to construct the \( \hat{\Phi}_i^* \), according to 5.2.2.

The sets of bootstrap estimates \( \hat{\Phi}_i^* \) (\( i = 1, 2, \ldots, B \), where \( B \) is the number of the bootstrap samples) is then used to estimate the density of the \( \hat{\Phi}_n \).

To obtain the \( q \)-percentile of the bootstrap distribution of the statistic \( \hat{\Phi} \) we sort the bootstrap estimates \( \hat{\Phi}^* \) in the ascending order and choose the one with the index \( \frac{qB}{100} \). Hence the decision rule is to reject \( H_0 : \rho = 1 \) and \( \tau = 0 \), if \( \hat{\Phi}_n > \hat{\Phi}^*_{(qB/100)} \). For example, for a test with nominal size of 5% and 10000 bootstrap replications we used the \( \hat{\Phi}^*_{(500)} \) as the critical point.

### 5.2.2 Size of the Test

As before, we estimate the Type-I error (percentage of the rejections of the true null-hypothesis) for series of different lengths \( (n = 500, 375, 250, 125, 60) \) generated using the stable distribution \( S_\alpha(1,0,0) \) with different values of the index of stability \( (\alpha = 1.75, 1.5, 1.25, 1.1) \).

The results of the experiments are given in Table 5.8.

<table>
<thead>
<tr>
<th>Length of the series (n)</th>
<th>Index of stability ( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>60</td>
<td>4.00</td>
</tr>
<tr>
<td>125</td>
<td>4.56</td>
</tr>
<tr>
<td>250</td>
<td>4.49</td>
</tr>
<tr>
<td>375</td>
<td>4.42</td>
</tr>
<tr>
<td>500</td>
<td>4.46</td>
</tr>
</tbody>
</table>

Figure 5.19 shows the graphical summary of the results.

As we can see, the type I error rate for this model is very close to the nominal level (the values are in the interval \( (5.3, 4.5) \)) for all \( \alpha \in [1.25, 2] \), regardless the length of the series. The worst result comes from a short length series \( (n < 100) \) that are generated using stable distribution with \( \alpha < 1.25 \). Even for the index of stability \( \alpha = 1.1 \) the error rate is about 1% less than the nominal
Fig. 5.19: Type-I error (percentage of the rejections of the true null-hypothesis $H_0: \rho = 1$ and $\tau = 0$) for the series of different length ($n = 500, 375, 250, 125, 60$) with different values of the index of stability ($\alpha = 1.75, 1.5, 1.25, 1.1$) used for the stable distribution. Based on 10000 experiments for every type of series and 2000 bootstrap replications for every estimation.

Also, the graphical comparison with the subsampling methods shows the bootstrap advantage,
see Figure 5.20 (the results of the subsampling tests are taken from [17]).

Fig. 5.20: Comparison of the type-I error (percentage of the rejections of the true null-hypothesis $H_0 : \rho = 1$ and $\tau = 0$) for the series of different length ($n = 500, 375, 250, 125, 60$) with different the index of stability ($\alpha = 1.5$) used for the stable distribution. Based on 10000 experiments for every type of series and 2000 replications for every bootstrap estimation and 5000 subsampling replications.

5.2.3 Power of the Test

To estimate the power of the test we applied it to the simulated series of the process (5.2.1) with $\alpha = 1.75$, and different values of $\rho$: $1, 0.99, 0.95, 0.9, 0.8, 0.5$. The power was estimated as a ratio of the number of rejections and the total number of experiments ($N = 10000$). Decision about rejecting the null hypothesis ($H_0 : \rho = 0$) every time was made based on empirical distribution of $\Phi_n$ estimated from 2000 of bootstrap samples. Experiments were performed for series of different lengths: $500, 375, 250, 125, 60$. The results are presented in Table 5.9. The graphical presentation of these results is shown on Figure 5.21.

One may see that the results are similar to those reported in Section 4.1 for model (2.4.1) (AR(1) without the constant term), i.e., the longer series are, the better the performance of the test. For the short series ($n < 200$) the test performs worse than the test for the model (5.1.1). However, such
Table 5.9: Number of rejections of $H_0 : \rho = 1$ and $\tau = 0$ out of 10000 experiments for series of different lengths with $\alpha = 1.75$. We compare the performance of the test for $\tau = 0$ and different values of the real $\rho$, based on 2000 bootstrap replications.

<table>
<thead>
<tr>
<th>Length of the series</th>
<th>$\rho$</th>
<th>1</th>
<th>0.99</th>
<th>0.95</th>
<th>0.9</th>
<th>0.8</th>
<th>0.5</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td></td>
<td>522</td>
<td>641</td>
<td>674</td>
<td>1072</td>
<td>3301</td>
<td>9658</td>
<td>10000</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td>490</td>
<td>577</td>
<td>1086</td>
<td>3326</td>
<td>9039</td>
<td>9890</td>
<td>10000</td>
</tr>
<tr>
<td>250</td>
<td></td>
<td>491</td>
<td>530</td>
<td>3313</td>
<td>9014</td>
<td>9831</td>
<td>9948</td>
<td>10000</td>
</tr>
<tr>
<td>375</td>
<td></td>
<td>517</td>
<td>666</td>
<td>6333</td>
<td>9748</td>
<td>9903</td>
<td>9965</td>
<td>10000</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>498</td>
<td>796</td>
<td>8334</td>
<td>9843</td>
<td>9934</td>
<td>9975</td>
<td>10000</td>
</tr>
</tbody>
</table>

Fig. 5.21: Graphical comparison of the power estimates for testing the hypothesis $H_0 : \rho = 1$ and $\tau = 0$, for the series of different lengths. Values for the parameter $\rho$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

comparison cannot be conclusive, because it should be taken into account that the model (5.1.1) has only one parameter $\rho$, while in model (5.2.1) we have two of them: $\rho$ and $\tau$. That is why it might be useful to estimate the power of the test for different values of the parameter $\tau$ as well.

We also performed 10000 experiments to estimate the power of the test for the model (5.2.1)
Table 5.10: Number of rejections of $H_0 : \rho = 1$ and $\tau = 0$ out of 10000 experiments for series of different lengths with $\alpha = 1.75$. We compare the performance of the test for $\rho = 1$ and different values of the parameter $\tau$, based on 2000 bootstrap replications.

<table>
<thead>
<tr>
<th>Length of the series</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>60</td>
<td>522</td>
</tr>
<tr>
<td>125</td>
<td>490</td>
</tr>
<tr>
<td>250</td>
<td>491</td>
</tr>
<tr>
<td>375</td>
<td>517</td>
</tr>
<tr>
<td>500</td>
<td>498</td>
</tr>
</tbody>
</table>

with $\alpha = 1.75$, $\rho = 1$ and different values of the parameter $\tau = 0, 0.5, 1, (\tau = 5$ was considered as a limit when power $\to 1$). For every experiment 2000 bootstrap replications were performed. The power was estimated only for $\tau > 0$, because it should be symmetric around zero. The results are given in Table 5.10.

The graphical representation of these results is given on the Figure 5.22.

We can see that the longer series result in better size and power of the test. Tests based on model (5.1.1) tend to reject the null hypothesis more often than they should (type I error is somewhat greater than the nominal size), while model (5.2.1) has empirical size lower than the nominal size.

The results obtained from a special case of the stable distribution $S_\alpha(1,0,0) = N(0,2)$ fully confirm the theoretical behavior of the classical unit root test, i.e., the type I error rate approaches the nominal size (see for example numerical simulations presented in [25]).

We have also compared the performance of the AR-sieve bootstrap and the subsampling, using the results of the numerical simulations from [17]. Figure 5.23 shows the graphical estimation of the power of the both tests for the model (5.2.1), with $\alpha = 1.5$.

As we can see, for the model with a constant term the subsampling test differs from the bootstrap test quite a lot. Though the shape of the power curves are follow the similar pattern, the subsampling test is more conservative, i.e., it tends to reject the null-hypothesis with the probability lower than the bootstrap test does. Such behavior is incident for all values of the real $\rho$, i.e., for both true and false null-hypothesis the type I error rate is smaller than the nominal level.
Fig. 5.22: Graphical comparison of the power estimates for testing the hypothesis $H_0: \rho = 1$ and $\tau = 0$, for the series with $a = 1.75$. Values for the parameter $\tau$ are on the X-axis, the estimate of the probability of rejecting the null-hypothesis is on the Y-axis.

5.3 Application to Yield Curves on Corporate Bonds

In this section the methods described earlier in this chapter were applied to the real data - yield curves on corporate bonds (length of series $n = 500$). The graphical representation of the series and their returns (first differences $X_t - X_{t-1}$) are given on Figures 5.24, 5.25, and 5.26.

First consider fitting model (5.1.1) (before applying the tests we subtracted the non-zero average from every data set). In this case the only one parameter $\rho$ needs to be estimated. The Least Square estimates for the six series are given in Table 5.11. We can see the point-estimated values of $\hat{\rho}$ are

<table>
<thead>
<tr>
<th>Least Square Type of the yield curve</th>
<th>Estimate $\hat{\rho}_{500}$ for the yield curves on corporate bonds.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates A</td>
<td>B</td>
</tr>
<tr>
<td>$\hat{\rho}_{500}$</td>
<td>0.99628</td>
</tr>
</tbody>
</table>
quite close to 1. Now let us test the hypothesis $H_0 : \rho = 1$ versus the alternative $H_a : \rho < 1$. We use two versions of the AR-sieve bootstrap: one with the bootstrap sample size $B$ being equal to the original series' length $n = 500$ and the improved version with the reduced bootstrap sample size $B = 0.85n$. Table 5.12 gives the summary of the p-values estimated for every data set by using 1000 bootstrap replications.

Table 5.12: The p-values estimated for testing the hypothesis $H_0 : \rho = 1$ on the yield curves (length = 500) using 1000 bootstrap replications.
Fig. 5.24: Yield curves on corporate bonds and corresponding returns. The bottom panels represent the first differences of the top panels.

We can see that the results are consistent with the numerical power estimates (see Table 5.1). Consider for example series D. The graph of the curve looks like a realization of a stationary process. The Least Square Estimate is $\hat{\rho} = 0.94641$ and the test presents very strong evidence for rejecting the null-hypothesis ($p$-value = 0). Also, looking at Table 5.1 we can see that the power estimate for the series of length $n = 500$ is 0.9999 when the real parameter $\rho = 0.9$.

Now let us fit model (5.2.1). In this case we have to estimate two parameters: $\rho$ and $\tau$. Their Least Square estimates for the six series are given in Table 5.13.

The interpretation of these results is not straightforward, because we have a complex null-hypothesis $H_0: \rho = 1$ and $\tau = 0$, which allows two different alternatives: $H_{a,1}: \rho < 1$ or $H_{a,2}: \tau \neq 1$. Also the relationship between the test statistic $\hat{\Phi}$ and the parameter estimates $\hat{\rho}$ and $\hat{\tau}$ is not obvious. For example in our case the estimates of parameters for the series A and E are very similar ($\hat{\rho}_A = 0.99642$, $\hat{\tau}_A = 0.01327$ and $\hat{\rho}_E = 0.99884$, $\hat{\tau}_E = 0.01097$), though the test statistics differ very much ($\hat{\Phi}_A = 5.57901$ and $\hat{\Phi}_E = 1.46973$) and thus result in the significant difference of the
Fig. 5.25: Yield curves on corporate bonds and corresponding returns. The bottom panels represent the first differences of the top panels.

$p$-values: $p_A$-value = 0.021 and $p_B$-value = 0.621, i.e., the test provides strong evidence for rejecting the null-hypothesis in case A.
Fig. 5.26: Yield curves on corporate bonds and corresponding returns. The bottom panels represent the first differences of the top panels.

Table 5.13: The results of the bootstrap tests for the yield curves on corporate bonds: LS estimates $\hat{\rho}_{500}$ and $\hat{\tau}_{500}$, statistic of interest $\Phi_{500}$ and the $p$-value. The results are based on 1000 bootstrap replications.

<table>
<thead>
<tr>
<th>The test results</th>
<th>Type of the yield curve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>$\hat{\rho}_{500}$</td>
<td>0.99642</td>
</tr>
<tr>
<td>$\hat{\tau}_{500}$</td>
<td>0.01327</td>
</tr>
<tr>
<td>$\Phi_{500}$</td>
<td>5.57901</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.021</td>
</tr>
</tbody>
</table>
CHAPTER 6

CONCLUSIONS

The goal of this thesis was to explore and implement bootstrap methods of unit root testing for heavy-tailed time series.

Among different bootstrap modifications we focused on the AR-sieve bootstrap because of its simplicity and effectiveness. Two different forms of this method (CI-based and the residual bootstrap) were investigated. The results clearly showed the superiority of the residual bootstrap over the CI-based bootstrap when applied to samples of size smaller than 500. Therefore, most of the conclusions apply to the residual bootstrap, if not specified otherwise.

In our experiments we focused on two types of AR(1) processes (with a constant term (5.2.1) and without (5.1.1)), though the methods can be extended to more complicated models. Both models were explored using several error distributions. The tests were applied to series whose errors followed the stable distribution with the index of stability $1 < \alpha \leq 2$. The results we obtained allow us to conclude that the performance of the tests (in terms of the type I error) is comparable with the asymptotic methods. Especially, when the distribution of the random errors $\varepsilon_t$ approaches the Gaussian distribution (index of stability $\alpha \to 2$). For practical purposes, one might expect to obtain reasonably good results (the type I error rate around $\pm 1\%$ of the nominal size) when the errors have the index of stability $\alpha \geq 1.5$.

To evaluate the finite sample performance of the tests, we estimated the size and the power for series of length $60 \leq n \leq 500$. The observed influence of the length of the series to the test performance (type I error) allowed us to establish an asymptotic validity of the tests, which means that the type I error rate approaches the nominal size when the length of the series is increasing. For practical purposes, it is enough to have series of length $n \leq 250$.

To improve the performance of the test we considered the reduction of the size of bootstrap samples. The results of the experiments when fitting the model (5.1.1) demonstrated an improvement of the performance for the method based on the residual bootstrap, but not the CI-based techniques. We also have proposed practical recommendations on the optimal bootstrap sample size.
For series of length $250 \leq n \leq 500$, the optimal sample size should be around $0.85n$.

We also have compared the performance of the AR-sieve bootstrap with the subsampling methods using the result of the numerical simulations from [17]. It turns out that the type I error of the bootstrap method is much closer to the nominal size than the error of the subsampling method. Regarding the power of the tests, the bootstrap tends to reject the null-hypothesis more often when it is false, while preserving the nominal level of the type I error (rejecting $H_0$ when it is true).

However, taking into account that the subsampling tests are easier to implement and faster to perform, they can successfully be used, especially for series of lengths $n \geq 500$, because their asymptotic properties (e.g., power of the test) are very similar and close to the nominal level.

In general, it should be said that the AR-sieve bootstrap tests of the unit root hypothesis for heavy-tailed time series seem to be quite accurate and computationally efficient. The results are consistent with theory, the algorithm is easy to implement, and the tests are easy to apply to real data sets.
REFERENCES


C++ CODE

The following code in C++ implements the bootstrap algorithm for the unit root testing of the autoregressive processes following the model (5.1.1). The algorithm for the model (5.2.1) is not provided since it is very much alike. Initially it was planned to implement the algorithm using the R language. However, due to the intensive calculations and the large amount of data, the performance of the program was very slow, so it was decided to use a low-level programming language.

```cpp
#include <Classes.hpp>
#include <Controls.hpp>
#include <StdCtrls.hpp>
#include <Forms.hpp>
#include <Buttons.hpp>
#include <ComCtrls.hpp>
#include <Dialogs.hpp>
#include <Grids.hpp>
#include <vcl.h>
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include "Unit1.h"
#pragma hdrstop
#pragma package(smart_init)
#pragma resource "*.dfm"

class TForm1 : public TForm { __published:
    // IDE-managed Components
    TButton *Button1;
    TButton *Button2;
```
TComboBox *ComboBox1;
TEDit *incr_percent;
TBitBtn *BitBtn1;
TMemo *Memo1;
TCheckBox *CheckBox1;
TCheckBox *CheckBox2;
TProgressBar *ProgressBar1;
TOpenDialog *OpenDialog1;
TLabel *Label1;
TLabel *Label2;
TLabel *Label3;
void __fastcall Button1Click(TObject *Sender);

void __fastcall FormActivate(TObject *Sender);
void __fastcall BitBtn1Click(TObject *Sender);
private:
public:
__fastcall TForm1(TComponent* Owner);

//Declaration of the global variables and objects:
 TForm1 *Form1;
const n_1 = 500; //max. length of the time series
const exp_N = 10000; //number of experiments
const BS = 2000; //number of bootstrap samples
const percent = 5; //quantile used for the hypothesis testing
const nr = 5; //number of different series to be tested
long double rho; //value of rho to be used
long double X[n_1]; //series to be tested
long double bs_X[2*n_1]; //bootstrap series
double eps_hat[n_1]; //residuals estimates
double T[BS]; //statistic of interest
unsigned short Rejection[nr]; // # of rejections of the hypothesis
FILE *F; //data file with alpha-stable random variables

__fastcall TForm1::TForm1(TComponent* Owner)
    : TForm(Owner){
}

void __fastcall TForm1::ButtonClick(TObject *Sender){
    if (OpenDialog1->Execute() ) //Execute file open dialog
        Button1->Caption = ExtractFileName(OpenDialog1->FileName);
        //extract file name to read data from
}

void Get_Series() {
    double rstab;
    double initial=0;

    if (rho==1)
    {
        for(unsigned short i=1; i<=101; i++)
{ // stabilization of the series
    fscanf(F, "%lf", &rstab);
    // read alpha-stable value from the file
    initial=initial + rstab;
}; // end of stabilization of the series

X[1-1]=initial;

for(unsigned short i=2; i<=n_1; i++)
{
    fscanf(F, "%lf", &rstab);
    // read alpha-stable value from the file
    X[i-1]=X[i-1-1]+rstab;
    // generate data following AR(1)
}

else
{
    for(unsigned short i=1; i<=101; i++)
    {
        // stabilization of the series
        fscanf(F, "%lf", &rstab);
        // read alpha-stable value from the file
        initial=rho*initial + rstab;
    }; // end of stabilization of the series

    X[1-1]=initial;

    for(unsigned short i=2; i<=n_1; i++)
    {
fscanf(F, "%lf", &rstab);
//read alpha-stable value from the file
X[i-1]=rho*X[i-1-1]+rstab;
//generate data following AR(1)
};

bs_X[i-1]=X[i-1]; //because it does not change
}; //End of getting series;

void Compute_Residuals(unsigned short ssize, long double ro) {

long double eps_bar=0; //average residual estimate

for (unsigned short i=2; i<=ssize; i++)
{
    eps_hat[i-1]=X[i-1]-ro*X[i-1-1];
    eps_bar=eps_bar+eps_hat[i-1];
}

eps_bar=eps_bar/(ssize-1);

for (unsigned short i=2; i<=ssize; i++)
{ //centering
    eps_hat[i-1]=eps_hat[i-1]-eps_bar;
}

} //End of computing the residuals
long double Estimate_rho(long double series[],
    unsigned short length){
    long double sum1 = 0;
    long double sum2 = 0;

    for (unsigned short i=1; i<length; i++){
        sum2 = sum2 + pow(series[i-1],2);
        sum1 = sum1 + series[i-1]*series[i];
    }

    return sum1/sum2;
} //End of estimating rho

void Bootstrap_X(unsigned short ssize, unsigned short bsize) {
    for(unsigned short i=1 ; i<bsize; i++){
        bs_X[i]=bs_X[i-1]+eps_hat[2+random(ssize-2)-1];
    }
} //End of bootstrapping series

void Sort_T(unsigned short size) {
    //sorting the bootstrapped statistics ascending order
    for(unsigned short i=size; i>=1; i--){
        for (unsigned short j=1; j<i; j++){
            if (T[j-1]>T[j])
                std::swap(T[j-1], T[j]);
        }
    }
} //End of sorting

{ 
  double tmp=T[j-1];
  T[j-1]=T[j];
  T[j]=tmp;
}

}//End of sorting the statistic T
//---------------------------------------------------------------

void __fastcall TForm1::FormActivate(TObject *Sender) {
  randomize();//randomize the generator
  ProgressBar1->Position=0;
}
//---------------------------------------------------------------

void __fastcall TForm1::BitBtn1Click(TObject *Sender) {
  //main simulation procedure
  long double rho_hat, p_value;
  unsigned short bs_size[5], curr_size[5];

  ProgressBar1->Max=exp_N*nr;
  rho = StrToFloat(ComboBox1->Text);
  unsigned short increase=StrToInt(Form1->incr_percent->Text);

  F = fopen(OpenDialog1->FileName.c_str(),"r");

  //Set the series to be tested and their bootstrapped sizes:
for (unsigned short exper = 1; exper <= exp_N; exper++) {
    Application->ProcessMessages(); // avoids programs’ freezing
    Get_Series(); // generates random series

    for (unsigned short i = nr; i >= 1; i--){
        // for every sample size do the loop:
        rho_hat = Estimate_rho(X, curr_size[i-1]);
        Compute_Residuals(curr_size[i-1], rho_hat);
        ProgressBar1->Position++;

        for (unsigned short bs_sample=1; bs_sample <= BS; bs_sample++){
            // main bs loop
            Bootstrap_X(curr_size[i-1], bs_size[i-1]);
            T[bs_sample-1]=bs_size[i-1]*(Estimate_rho(bs_X, bs_size[i-1])-1);
        } // end of main bs loop

        Sort_T(BS);

        if (curr_size[i-1]*(rho_hat-1) < T[percent*BS/100-1])
            Rejection[i-1]++;
    } // end of loop for every sample size
} // End of the experiment
fclose(F);

// Display the results:
Form1->Memo1->Lines->Add("List of parameters:");
Form1->Memo1->Lines->Add("Alfa-Stable = "+ Form1->Button1->Caption);
Form1->Memo1->Lines->Add("Real rho = "+ FloatToStr(rho));
Form1->Memo1->Lines->Add("# of experiments = "+ FloatToStr(exp_N));
Form1->Memo1->Lines->Add("# of bootstrap samples = " + FloatToStr(BS));
Form1->Memo1->Lines->Add("Length of bootstrap samples is " +
                         Form1->incr_percent->Text+"% of the original");
Form1->Memo1->Lines->Add(" ");

for(unsigned short i=1; i<=nr; i++)
    Form1->Memo1->Lines->Add("For sample size = " +
            IntToStr(curr_size[i-1])+" (bs. sample = " +
            IntToStr(bs_size[i-1])+" )" + " Ho (rho=1) was
            rejected " + IntToStr(Rejection[i-1])+" times");
MessageDlg(" Simulation has been completed successfully!",mtInformation,
            TMsgDlgButtons() << mbOK,0);
}// end of main simulations

//--------------