PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA

ON TESTING FOR STRICT STATIONARITY

BY

RICARDO FELIPE BÓRQUEZ COUSIÑO

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF STATISTICS
AT
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
SANTIAGO, CHILE
JANUARY 2012

Research Supervisor: Wilfredo Omar Palma Manríquez

©2012, Ricardo Felipe Bórquez Cousiño

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “**On Testing for Strict Stationarity**” by **Ricardo Felipe Bórquez Cousiño** in partial fulfillment of the requirement for the degree of **Doctor of Statistics**.

Dated: January 2012

Research Supervisor: ______________________

Wilfredo O. Palma Manríquez

Examing Committee: ______________________

Manuel Galea Rojas

____

Guillermo P. Ferreira Cabezas

____

Reinaldo B. Arellano Valle

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Author: Ricardo Felipe Bórquez Cousiño Date: January, 2012

Title: On Testing for Strict Stationarity

Faculty: Mathematics

Department: Statistics

Degree: Doctor of Statistics

Convocation: January, 2012

Permission is herewith granted to Pontificia Universidad Católica de Chile to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

THE AUTHOR RESERVES OTHER PUBLICATION RIGHTS, AND NEITHER THE THESIS NOR EXTENSIVE EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT THE AUTHOR'S WRITTEN PERMISSION.

THE AUTHOR ATTESTS THAT PERMISSION HAS BEEN OBTAINED FOR THE USE OF ANY COPYRIGHTED MATERIAL APPEARING IN THIS THESIS (OTHER THAN BRIEF EXCERPTS REQUIRING ONLY PROPER ACKNOWLEDGEMENT IN SCHOLARLY WRITING) AND THAT ALL SUCH USE IS CLEARLY ACKNOWLEDGED.

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

ABSTRACT

On Testing for Strict Stationarity

by

Ricardo Felipe Bórquez Cousiño

We derive a statistical method for evaluating whether a time-series has been drawn from a strictly stationary sequence. The result follows from a general characterization of the class of random sequences which are not strict stationary, and the path-sample properties of their corresponding time-series. Theoretical concepts such as measure-preserving transformations, invariant sets, ergodic theorem and the statistics of self-similar processes are carefully revised to base our results in a self-contained document. A Monte Carlo experiment is performed to study the small sample properties of the test statistic and two examples with real-data analysis are also provided.

Dedico este trabajo a mis padres Ana María y Ricardo, y también a Alejandra, Claudia, Ignacia, Cristián, Matías y Tomás.

ACKNOWLEDGEMENTS

Mis más sinceros agradecimientos a Wilfredo Palma, Reinaldo Arellano, Manuel Galea, Guillermo Ferreira y Guido Del Pino, a mis compañeros Juan Francisco Olivares, Mauricio Tejo y Felipe Barrientos, a Wolfgang Rivera, Soledad Alcaíno y a todos aquellos amigos que de una u otra forma han ayudado a terminar esta tesis.

Contents

1 Introduction **1**

1.1 Strictly Stationary Sequences 1

1.2 Description of the Problem 2

1.3 Summary of Contents 4

2 Strict Stationarity: Theory **5**

2.1 Introduction 5

2.2 Measure-Preserving Transformations 5

2.2.1 Invariant and Almost Invariant Sets 7

2.2.2 Poincaré Recurrence Theorem 9

2.3 Self-Similar Processes 10

3 Results **13**

3.1 Introduction 13

3.2 The Main Result 13

3.3 Testing for Strict Stationarity 18

4 Ilustrations **22**

4.1 A Monte Carlo Experiment 22

4.2 Two Examples with Real-Data Analysis 29

4.2.1 Interest Rate Data 29

4.2.2 Nile River Data 31

5 Conclusions **34**

A Graphics (Monte Carlo experiment) **35**

Bibliography **45**

List of Tables

A.1 Size of the Test (i.i.d normal test) 43
A.2 Power of the Test (i.i.d normal test) 44

Chapter 1

Introduction

1.1 Strictly Stationary Sequences

A significant part of the study of time-series analysis and more generally of random-fields is advocated on strictly stationary sequences (or "processes" if the time-parameter is assumed to be continuous on some given interval). Their importance rely in that they serve as building blocks of large classes of models yet through passing linear or non-linear (but measurable) filters to a sequence of independent and identically distributed (i.i.d.) random variables[1], yet forming normalized sums of the elements of a strictly stationary sequence as in self-similar processes or even by computing integrals of certain types of functions where the (stochastic) integration is made upon a non-decreasing process with strictly stationary increments.

For statistics their importance is also very clear if it is noted that almost all the results that are valid for sequences of i.i.d. random variables such as strong laws of large numbers, weak limit theorems, limit theorems for sums of random variables, etc. also extend (usually without much trouble) to other subclasses of time-dependent strictly stationary sequences.

As mathematical objects, strictly stationary sequences are endowed with a number of properties which have proved to be useful on the task of modeling random sequences that are described to

[1] An i.i.d. sequence is perhaps the simplest example of a strictly stationary sequence.

be in equilibrium[2] and these appear on several contexts of application such as biology, physics, chemistry, economics, finance, music among many others. The possibilities of such models are not limited to render simple trajectories or realizations as the notion of equilibrium would suggest, on the contrary, the time-series of a strictly stationary sequence (i.e., a realization of it) can exhibit complex behavior providing then a very flexible tool. But this flexibility makes difficult to infer from a data sample if the underlying sequence generating that data is eventually a strictly stationary sequence.

Before leaving this section we give a first definition of strict stationarity. A sequence of random variables $X_1, X_2, ...$ is said to be strictly stationary if the distribution function of any finite collection of them is invariant with respect to time, that is, if the distribution function of $(X_1, X_2, ..., X_t)'$ is the same as that of $(X_{k+1}, X_{k+2}, ..., X_{k+t})'$ for all $t, k \in \mathbb{Z}$. It follows from this definition that the marginal probability of X_t is the same as that of X_1.

1.2 Description of the Problem

From the definition of strict stationarity it results rather natural to think on testing this property by evaluating the time-homogeneity of some statistics[3] such as density kernels, quantiles, expectiles, etc. On this line of research is much of the work we can find in the area. For instance, Lee and Na (2004) propose a recursive estimation of density kernels in the context of strong mixing processes, their work relates to that of Bai (1994) who looked for changes in the distribution of the residuals in the context of estimation of ARMA models. Detection of time-instabilities in the distribution of a time-series is also the idea behind a test due to Inoue (2001). Kapetanios (2009) goes in this direction but its work is based on bootstrap procedures. An alternative approach is taken on an unpublished work due to Trapani (2008) which instead of focusing on the empirical distribution propose a test based on linear rank statistics for the null hypothesis of a time-stable distribution. In a recent paper Busetti and Harvey (2010) have proposed to evaluate the time-homogeneity of

[2]The concept of equilibrium as it concerns to us must be related to the existence of a probability measure which is invariant for a class of sets, the measurable invariant sets.

[3]Related to this problem there is an amount of literature also including parametric approaches, which has been developed under the concept of "structural change" in econometrics.

quantile indicators. We also found another unpublished paper due to Lima and Neri (2008) which is based again on finding fluctuations of the sample quantiles.

It is more subtle to note, however, that the definition above for a strictly stationary sequence actually characterizes a particular class of sets called invariant sets; in this case we also refer the respective probability as an invariant probability measure (see Section 2.2.1). Because the condition of time-homogeneity on a given statistic is therefore a restriction on the class of invariant sets, it necessarily reduces the problem of inference and one needs to ask if such a reduction is in some sense too restrictive. We now discuss this issue in some detail.

The Mathematical Setting

Consider the following statistical model: $(M, \mathcal{M}, P : P \in \mathcal{P})$ for an observed time-series, where M is a m-dimensional Borel set[4], $\mathcal{M}$ is the respective Borel sigma algebra of subsets of M and the symbol $\mathcal{P}$ denotes for the collection of all the finite-dimensional distributions defined on $(M, \mathcal{M})$. The objective is to develop a testing procedure for the decision whether a time-series has been drawn from a strictly stationary sequence, that is, for P an element of $\mathcal{P}$ we want to decide whether $P \in \mathcal{P}_0$ where $\mathcal{P} = \mathcal{P}_0 \cup \mathcal{P}_0^c$ and $\mathcal{P}_0$ denotes for the class of invariant measures. Because a finite-dimensional distribution on $(M, \mathcal{M})$ completely determines a probability measure now defined on the measurable space of random sequences (by the Kolmogorov's extension argument), the problem is well stated.

In order to sustain inference about the hypothesis that a time-series is a realization from a strictly stationary sequence, we need to focus on a quantity of interest which is a common feature that is present in every element of $\mathcal{P}_0$. It is clear from the above definition of strict stationarity that any test that is based on some choice of parameter restrictions characterizing a restricted null hypothesis, say $P \in P_0' \subset \mathcal{P}_0$, will leave aside a number of more complex models which can also be strictly stationary. Tests of this kind are of great importance in the context of model specification but they are not well posed to build inference on strict stationarity. The problem is similar on those

[4]We require that measurable cylinders be contained in M. Further properties of this set are inherited from Ω, a (sequentially compact) metric space of random sequences defined in Section 2.2.

testing procedures that are based on the time-homogeneity of a given sample statistic, as they require the underlying random sequence to be ergodic (see the footnote below), again a substantive specialization of the model.

A more general description of the elements of $\mathcal{P}_0$ can be achieved through characterizing a dense set $\mathcal{P}_0^*$ (of $\mathcal{P}_0$). The fundamental result is due to Nisio (1960), he showed that the characteristic function of an arbitrary element $P \in \mathcal{P}_0$ can be approximated in $\mathcal{L}_1$-norm by the limit of a sequence of characteristic functions of a class of random sequences called polynomial sequences, so that P has a law representation in terms of the finite-dimensional distribution of a (limit) polynomial sequence (see Remark 2.8). A polynomial sequence is obtained by passing a (possibly non-) linear filter over a sequence of i.i.d standard normal random variables, therefore the approximating probability μ' of an otherwise arbitrary invariant probability measure μ is absolutely continuous with respect to μ^* (we write $\mu' \ll \mu^*$), the probability measure of an i.i.d standard normal random sequence. That is, regarding the characterization of this probability μ nothing is lost if we assume that it is also absolutely continuous with respect to μ^*, so it has a density.

Moreover, because an i.i.d. random sequence (whether Gaussian or not) is not only strictly stationary but also ergodic, its probability μ^* assigns the values 0 or 1 to every set on the class of invariant sets[5]. It follows by absolute continuity that every invariant set of measure zero under the probability measure μ^* has also measure zero under the probability measure μ. As the latter is a feature that is common to every strictly stationary sequence defined in the context above, then it can be used to build a valid inference procedure on the hypothesis of strict stationarity.

1.3 Summary of Contents

In Chapter 2 we provide the theoretical concepts needed to introduce strictly stationary sequences and to derive our results. In Chapter 3 we present our results on a test statistic for the null hypothesis of strict stationarity. In Chapter 4 a Monte Carlo experiment allows us to assess the empirical performance of this new test statistic, some examples with real-data analysis are also included. The conclusions and recommendations for further research are found in Chapter 5.

[5] We can take this property as the definition of an ergodic probability measure.

Chapter 2

Strict Stationarity: Theory

2.1 Introduction

In this chapter we cover the concepts necessary to introduce strictly stationary sequences. We begin with the notion of measure-preserving transformation, which conduces to the most general representation of a strictly stationary sequence. We follow with invariant sets, these are fundamental in the sense that strict stationarity can be viewed basically as a measurement property on these sets. The Poincaré recurrence theorem will gives us the necessary tools to get some interpretation and to characterize those time-series which are elements of an invariant set. Finally, we review the concept of self-similarity to understand the asymptotic behavior of normalized sums of strictly stationary sequences.

2.2 Measure-Preserving Transformations

Consider a fixed probability space $(\Omega, \mathcal{G}, \mu)$ where Ω is a cartesian product of copies of the real line, $\mathcal{G}$ is a sigma algebra of subsets of Ω and μ is a probability measure on $\mathcal{G}$ such that $\mu \ll \mu^*$ (where μ^* was defined in section 1.2). Let $U : \Omega \to \Omega$ be a measure-preserving set transformation defined on this probability space (i.e., U satisfies $U^{-1}\mathcal{G} \subset \mathcal{G}$ and also $\mu(A) = \mu(U^{-1}A)$ for every set $A \in \mathcal{G}$).

Let X be a random variable on $(\Omega, \mathcal{G})$ and define $X_t(\omega) = X(U^t\omega)$ for all $\omega \in \Omega$ and $t \in \mathbb{N}$ (where $U^t\omega$ denotes for t consecutive applications of the set transformation U over ω). The sequence $\{X_t; t \in \mathbb{N}\}$ is strictly stationary (See Proposition 6.9 in Breiman (1992)). To see that every strictly stationary sequence on $(\Omega, \mathcal{G}, \mu)$ can be generated by the set transformation U in the way prescribed above, consider as the ω-sets of interest those measurable sets on the sample space of the X_t's or measurable cylinder sets (i.e., sets of the form $A = \{\omega : (X_0(\omega), X_1(\omega), ..., X_{n-1}(\omega)) \in B\}$ where B is a n-dimensional Borel set). The class of sets $A \in \mathcal{G}$ such that $\mu(A) = \mu(U^{-1}A)$ is a monotone class containing the measurable cylinders and hence coincides with $\mathcal{G}$ (see Ash (2000), p. 347).

Moreover, there exists a measurable transformation S now defined on the sample space of the X_t's (i.e., taking random vectors into random vectors) such that for all $t, n \in \mathbb{N}$ we have the following properties: $(X_0, X_1, ..., X_{n-1}) = S^{-t}(X_t, X_{t+1}, ..., X_{t+n-1})$ a.s. and

$$P\{(X_0, X_1, ..., X_{n-1}) \in B\} = P\{(X_t, X_{t+1}, ..., X_{t+n-1}) \in B\},$$

where B is a n-dimensional Borel set and P is the respective Lebesgue-Stieltjes measure.

Remark 2.1 *The transformation S (which is induced by the set transformation U) is measurable and unique up to sets of zero probability, it is called the **shift** transformation of the strictly stationary sequence $\{X_t; t \in \mathbb{N}\}$.*

Remark 2.2 *We can complete the Borel σ-algebra $\mathcal{G}$ with all the sets of μ-measure zero. In that case, the induced transformation will have as domain and range space the class of random variables measurable with respect to $\mathcal{G}$ (see Doob (1953), p. 455).*

Remark 2.3 *For g a measurable transformation, if $\{X_t; t \in \mathbb{N}\}$ is strictly stationary then so it is the sequence $\{g(X_t); t \in \mathbb{N}\}$ by Theorem 3.5.3 in Stout (1974), p. 170.*

Remark 2.4 *If the set transformation U (or the respective induced transformation S) has an inverse, then the two-sided sequence $\{X_t; t \in \mathbb{Z}\}$ is strictly stationary and all the constructions detailed above for the one-sided case are also valid for the two-sided case.*

Remark 2.5 *The probability space where we have defined the strictly stationary sequences of interest is separable (i.e., Ω contains a countable dense subset). But this is restrictive mainly on a continuous-time setting, where we note that some results below will critically depend on that assumption.*

Remark 2.6 *If ν^* is an ergodic measure on $(\Omega, \mathcal{G})$ then it is equivalent or singular with respect to μ^* (see Theorem 8.3.11 in Ash (2000)). But on a finite measure space the existence of ν^* implies that it is equivalent to μ^* (see Theorem 2 in Halmos (1947), p. 738).*

Remark 2.7 *The probability μ is uniquely determined by U and sometimes it is referred as an invariant probability measure under the set transformation U.*

Remark 2.8 *A random sequence $\{Z_t; t \in \mathbb{Z}\}$ is called a polynomial sequence if it satisfies that,*

$$Z_t(\omega) = \sum_p \sum_{i_1 \dots i_p} a_{i_1 \dots i_p} \prod_{\nu=1}^{p} \xi_{i_\nu + t}(\omega),$$

where $\{\xi_t; t \in \mathbb{Z}\}$ is a sequence of i.i.d standard normal random variables. For an arbitrary strictly stationary sequence $\{X_t; t \in \mathbb{Z}\}$ we have:

$$\lim_{m \to \infty} \left| \int \exp\left(i \sum_{j=-n}^{n} \alpha_j X_j(\omega) \right) d\mu - \int \exp\left(i \sum_{j=-n}^{n} \alpha_j Z_j^{(m)}(\omega) \right) d\mu^* \right| = 0,$$

for any α_n and $n \in \mathbb{Z}$.

This result also extends to a continuous-time setting under certain additional topological requirements and considering as a base sequence, polynomials of the increments of a Brownian motion (see Theorems 3 and 7 in Nisio (1960), pp. 208-210).

2.2.1 Invariant and Almost Invariant Sets

Let $U : \Omega \to \Omega$ be a measure-preserving set transformation. A set $A' \in \mathcal{G}$ is invariant under the set transformation U if $U^{-1}A' = A'$. For example, if there is a Borel set B' such that for every $t \in \mathbb{N}$ we have $A' = \{\omega : (X_t(\omega), X_{t+1}(\omega), \dots) \in B'\}$ then the set A' is an invariant set. It is the case that every invariant set is a tail event (see Breiman (1992), p. 118), it is also the case that the collection of all invariant sets form a sigma algebra (see Ash (2000), p. 350) which will be denoted as $\mathcal{I}$. Let

g' be a random variable then g' invariant under the set transformation U means that g' is measurable on $\mathcal{I}$ and $g'(\omega) = g'(U^{-t}\omega)$ for all $\omega \in \Omega$ and all $t \in \mathbb{N}$.

Associated to every invariant set A' there is an "almost" invariant set A with the property that $\mu(A' \bigtriangleup A) = 0$, where the symbol $\bigtriangleup$ denotes for the symmetric difference operator. A set A is called an almost invariant set if $\mu(A \bigtriangleup U^{-1}A) = 0$. For g a random variable then g almost invariant under the set transformation U means that g is measurable on the sigma algebra generated by the almost invariant sets and $g(\omega) = g(U^{-t}\omega)$ for almost all $\omega \in \Omega$ (i.e., $g(\omega) = g(U^{-t}\omega)$ a.s.) and all $t \in \mathbb{N}$. This latter sigma algebra is exactly the completion of $\mathcal{I}$ with the μ-null sets of $\mathcal{G}$ (see Breiman (1992), p. 109).

Remark 2.9 *Every invariant set is an almost invariant set, the converse is not true in general.*

Remark 2.10 *Whenever $g(\omega) =$ constant for almost all $\omega \in \Omega$ we say that the transformation U (or the respective probability measure) is ergodic (See Petersen (1989), p. 42).*

Remark 2.11 *If the set transformation U has an inverse, then the same ω-sets and the same random variables are almost invariant both with respect to U or its inverse. In particular, the same almost invariant sets characterize the processes $\{X_t; t \in \mathbb{N}\}$ and $\{X_t; t \in \mathbb{Z}\}$ (see Doob (1953), p. 457).*

Remark 2.12 *Perhaps the most important result for measure-preserving transformations is the ergodic theorem: For $\{X_t; t \in \mathbb{N}\}$ a strictly stationary sequence in $\mathcal{L}_p$ with p satisfying that $1 < p < \infty$ we have:*

$$n^{-1}\sum_{i=0}^{n-1} X_i \to X \quad \textit{a.s. (and in } \mathcal{L}_p\textit{-norm)},$$

where X is an invariant random variable (see Theorems 8.3.6 and 8.3.7 in Ash (2000), p. 361). In particular, if $\{X_t; t \in \mathbb{N}\}$ is an ergodic sequence in $\mathcal{L}_2$ then (see Corollary 6.23 in Breiman (1992), p. 115):

$$n^{-1}\sum_{i=0}^{n-1} X_i \to \int X_0 \, d\mu \quad \textit{a.s.}$$

Remark 2.13 *The ergodic theorem can fail to hold if it is required that $\|U\|_1 = 1$ where $\|U\|_1$ is the smallest number c such that $\|U^t X\|_1 \le c\|X\|_1$ (called the $\mathcal{L}_1$-norm of the transformation U). Sufficient conditions for the validity of the ergodic theorem in this context are provided in Ito (1965).*

2.2.2 Poincaré Recurrence Theorem

Definition 2.1 *Let $A \in \mathcal{G}$ and $U : \Omega \to \Omega$ is a measure-preserving set transformation defined on $(\Omega, \mathcal{G}, \mu)$. A point $\omega \in A$ is said to be recurrent with respect to A if there is $t \in \mathbb{N}$ such that $U^t\omega \in A$.*

Alternatively, we can say that the transformation U is **recurrent**. The next is a well known result in ergodic theory due to Henry Poincaré which dates back to 1899:

Theorem 2.1 *For each $A \in \mathcal{G}$, almost every point of A is recurrent with respect to A.*

Proof See Petersen (1989), p. 34 ■

An interpretation of this theorem is as follows. If at a given time some observations of a random sequence visit the set A, then almost surely we will observe some subsequent observations again on this same set. On a finite measure space (e.g., a probability space) the Poincaré recurrence theorem indeed implies a stronger version of itself, in that the observed trajectory of the random sequence will return to the same already visited set infinitely many times (see Halmos (1956), p. 10). In this case, it is said that the measure-preserving set transformation U is **infinitely recurrent**.

A related result which is instrumental in our context is due to Halmos (1947), we reference it as Theorem 3.4 in Petersen (1989), p. 39. First, another definition is needed:

Definition 2.2 *A measure-preserving set transformation U is called **incompressible** if whenever $A \in \mathcal{G}$ and $U^{-1}A \subset A$ then $\mu(A - U^{-1}A) = 0$.*

The result is that U is incompressible if and only if U is infinitely recurrent. By Lemma 8.2.4 in Ash (2000) p. 351 we also have that $\mu(A - U^{-1}A) = 0$ further implies that $\mu(A \bigtriangleup U^{-1}A) = 0$. But then by Corollary 1 in Halmos (1947), p. 738 every power of an incompressible transformation is incompressible so that $\mu(A \bigtriangleup U^{-t}A) = 0$ for all $t \in \mathbb{N}$.

The Poincaré recurrence theorem gives us a method for evaluating if a set is an invariant set. In effect, if $A^* = \bigcup_{t=1}^{\infty} U^{-t}A$ then the set $A' = A - A^*$ is the set of all those points of A which never return to A, and for U infinitely recurrent it is the case that the set A' is in $\mathcal{I}$ (see Petersen (1989), p. 38). The method easily translates in terms of indicator functions, as every function measurable on (almost) invariant sets is necessarily (almost) invariant with respect to the shift transformation S (see Section 3.2).

Example 2.1 *The simplest example is the infinite sequence $X, X, ...$ where X is an arbitrary random variable defined on $(\Omega, \mathcal{G})$. Note that the sequence is clearly invariant to the shift transformation. For this sequence, a time-series is a sequence of real values $X(\omega), X(\omega), ...$ which occurs with probability 1 or 0 only if the underlying random sequence is ergodic. It is evident that such a pattern should not to be observed very frequently for a time-series that has been drawn from a strictly stationary sequence, unless of course that the respective probability measure is a discrete measure, but that case has no particular interest in our context (see Remark 2.8).*

2.3 Self-Similar Processes

A number of results in probability and statistics depend on normalizing sums of strictly stationary sequences of random variables, its limit behavior is better understood under the theory of self-similar processes.
We commence with a brief revision of the concept of self-similarity:

Definition 2.3 *A real-valued stochastic process $\{Y(u); u > 0\}$ is said to be **self-similar** with parameter $H > 0$ (denoted H-ss) if $Y(cu) \stackrel{d}{=} c^H Y(u)$ for all real $c > 0$, where the symbol $\stackrel{d}{=}$ means equal distribution. The process $\{Y(u); u > 0\}$ is said to be self-similar **with (strictly) stationary increments** (denoted H-sssi) if it is self-similar and in addition $Y(u+h) - Y(u) \stackrel{d}{=} Y(u) - Y(0)$ for all $h \in \mathbb{R}$ and $u > 0$.*

The first result of interest to us is due to Lamperti (1972): if $d(.)$ is a positive function on $\mathbb{R}^+$ and $\{S(cu); u > 0\}$ is a stochastic process with stationary increments such that,

$$d(c)^{-1} S(cu) \Rightarrow_c S(u)$$

where the symbol $\Rightarrow_c$ denotes for weak convergence as $c \to \infty$, then $\{S(u); u > 0\}$ is a self-similar[1] process with stationary increments (see Vervaat (1985), p. 1). A more operational version of the above result is available, it is referenced as Theorem 2.3. in Beran (1994), p. 50:

Lemma 2.1 *Let S_t be the limit in distribution of the sequence of normalized partial sums $\frac{1}{d_n}\sum_{i=1}^{[nt]} X_i$ as $\log d_n \to \infty$ with $n = 1, 2, ...$ and $\{X_i; i \in \mathbb{N}\}$ is a strictly stationary sequence (where the symbol $[nt]$ denotes the integer part of nt and $t > 0$). Then, there exists $H > 0$ such that for any $c > 0$ we have $\lim_{n\to\infty} \frac{d_{cn}}{d_n} = c^H$ and $\{S_t; t > 0\}$ is H-sssi.*

Remark 2.14 *All H-sssi processes can be obtained by partial sums as in Lemma 2.1.*

Remark 2.15 *The function d_n with $n = 0, 1, ...$ is a regularly varying sequence, it admits a representation of the form $d_n = n^H L(n)$ where,*

$$L(n) = \exp\left(\eta(n) + \int_{[0,n]} \frac{\varepsilon(t)}{t} dt\right),$$

with $\eta(t)$ converging to a positive limit and $\varepsilon(t) \to 0$ as $t \to \infty$. Thus, $L(n)$ is a slowly varying function at infinity, in the sense of Karamata (see Galambos and Seneta (1973), p. 113).

Remark 2.16 *For $H > 0$ the distribution function of Y_1 cannot have atoms except possibly at zero. Additional properties on marginal distributions of H-sssi processes can be found in O'Brien and Vervaat (1983).*

Remark 2.17 *Sample-path properties of H-sssi processes are studied in Vervaat (1985). In particular, it is shown there that the restriction $H > 0$ is necessary in order to preserve measurability of the H-sssi process $\{Y_t; t > 0\}$ and that $H \leq 0$ and measurability both imply $Y_t = 0$ a.s. for each $t > 0$ if the process above is defined on a separable probability space.*

Remark 2.18 *If $\{Y_t; t > 0\}$ is an H-sssi process in $\mathcal{L}_2(\Omega, \mathcal{G}, \mu)$ then some useful properties immediately arise: i) if $H \neq 1$ then $\int Y_t d\mu = 0$ for all $t > 0$; ii) $\int Y_t^2 d\mu = t^{2H}\sigma^2$ for all $t > 0$ where $\sigma^2 = \int Y_1^2 d\mu$; and iii) cov$(Y_t, Y_s) = \frac{1}{2}\sigma^2[t^{2H} - (t-s)^{2H} + s^{2H}]$ for all $s, t > 0$.*

[1]The term "semi-stable" instead of "self-similar" was used originally by Lamperti (see Lamperti (1972)).

Moreover, for $H = \frac{1}{2}$ the limit process is standard Brownian motion (or Wiener measure) so that their stationary increments are independent. For $\frac{1}{2} < H < 1$ the limit process is long-range dependent, in the sense that their autocorrelations are not absolutely summable. Nevertheless, it is very important for what follows to stress that in the latter case the autocovariances of the limit process still converge to zero (although not very fast). Finally, for $0 < H < \frac{1}{2}$ the autocovariances are absolutely summable and consequently they also converge to zero so that the stationary increments exhibit short memory. Further properties on these processes can be found in Beran (1994) and also in Taqqu (2003).

Remark 2.19 *A partial converse to Lemma 2.1 is available, it is due to Davydov (1970) and we refer it as Theorem 5.2 in Taqqu (2003), p. 17:*
Let $\{X_i; i \in \mathbb{N}\}$ be a zero-mean strictly stationary sequence in $\mathcal{L}_2(\Omega, \mathcal{G}, \mu)$. If,

$$\int \left(\sum_{i=1}^{n} X_i \right)^2 d\mu = n^{2H} L^2(n),$$

and $0 < H < 1$ then,

$$\frac{1}{n^H L(n)} \sum_{i=1}^{[nt]} X_i \Rightarrow_n B_H(t),$$

for all $t > 0$, where $B_H(t)$ denotes for a standard fractional Brownian motion with self-similarity parameter H. Whenever $H = \frac{1}{2}$ we write $\mathcal{W}(t)$ instead of $B_{\frac{1}{2}}(t)$.

Chapter 3

Results

3.1 Introduction

In this chapter we present our theoretical results. The first part contains the main result, which is a characterization of the null hypothesis of strict stationarity. We continue with the derivation of a test statistic and of its limit distribution that is consistent for this null hypothesis.

3.2 The Main Result

For almost invariant sets A and $U^{-t}A$ the corresponding indicator functions satisfy:

$$\mathbf{1}_A = \mathbf{1}_{U^{-t}A} \text{ a.s. for all } t \in \mathbb{N} \text{ and all } A \in \mathcal{G}.$$

To see this, note that by Markov's inequality for every $\varepsilon > 0$ we have $\mu(A \vartriangle U^{-t}A) = 0$ implies that,

$$\mu\{\omega : |\mathbf{1}_A(\omega) - \mathbf{1}_A(U^{-t}\omega)| > \varepsilon\} \leq \varepsilon^{-1} \int |\mathbf{1}_A(\omega) - \mathbf{1}_A(U^{-t}\omega)| \, d\mu(\omega) = 0,$$

for all $t \in \mathbb{N}$ and all $A \in \mathcal{G}$ (this follows simply by properties of indicator functions and because the set transformation U is measure-preserving). The converse of this result is also true, so the above property on indicator functions can indeed be taken as the definition of almost invariant sets (see Ash (2000), p. 350).

One important consequence from the previous result is that measurable cylinders defined with respect to the strictly stationary sequence $\{X_t; t \in \mathbb{N}\}$ cannot be almost invariant sets for X_0 a continuous random variable. To see this, note that for $A = \{\omega : (X_0(\omega), X_1(\omega), ..., X_{n-1}(\omega)) \in B\}$ a measurable cylinder we have $\omega' \in A$ if and only if $\omega' \in U^{-t}A$ for all $t \in \mathbb{N}$ because U is infinitely recurrent. Then, U is incompressible implying that $\mu(A \vartriangle U^{-t}A) = 0$ for all $t \in \mathbb{N}$ but this is impossible if the distribution function of X_0 has no atoms. By Remark 2.11 this result immediately extends over the two-sided sequence $\{X_t; t \in \mathbb{Z}\}$.

The result above suggests that a testing procedure for rejecting the null hypothesis of strict stationarity can be stated on evaluating the probability of a particular class of sets, the class of invariant sets. By the symmetry on the definition of almost invariant sets and the property above for indicator functions defined on these sets, we can operationalize the theoretical arguments given in the previous chapters through some computations on indicator functions. We commence with a convenient transformation of the sets of interest. Let B be a Borel set, then define the set,

$$A_{t,k}^B = \{\omega : X_t(\omega) \in B\} \vartriangle \{\omega : X_{t+k}(\omega) \in B\},$$

where $\{X_t; t \in \mathbb{Z}\}$ is a strictly stationary sequence with X_0 a continuous random variable. To save notation, for fixed B and k we simply write $A_{t,k}^B = A_t$ where no confusion can arise. The main characteristic of the set A_t is that even though it is obtained by forming the symmetric difference of two sets which are arbitrary in every respect, an almost invariant probability measure will always assign zero probability on A_t if this set is actually an invariant set. Note that the time-series of a strictly stationary sequence should visit the set A_t very frequently.

A natural estimator for the value of a probability on a given set (interval) is the empirical distribution, and the limit behavior of such a statistic is determined by the ergodic theorem. This is because $\mathbf{1}_{A_0}, \mathbf{1}_{A_1}, ...$ is a sequence in $\mathcal{L}_2$ then strict stationarity and integrability both imply that $\|S\|_2 = 1$ where $\|S\|_2$ is the $\mathcal{L}_2$-norm of S (see Theorem 8.3.1 in Ash (2000), p. 356) and the result follows by Theorem 2 in Ito (1965), p. 227. For sequences of indicator functions defined on the sets $A_0, A_1, ...$ the partial averages $Y_t(\omega) = \frac{1}{t}\sum\limits_{i=0}^{t-1}\mathbf{1}_{A_i}(\omega)$ with $t = 1, 2, ...$ converge almost everywhere to a well defined limit Y in the interval $(0, 1)$, where Y can depend on k and B (we write $Y = Y_k^B$ when it is needed to make explicit this dependence). If the sets $A_0, A_1, ...$ are invariant sets then trivially the limit Y equals 1 or 0 a.s. The case where $Y = 1$ a.s. is of no interest, however, because under the hypothesis of strict stationarity this result follows only if $X = \text{const}$ a.s. $-$for X a continuous random-variable (see Section 2.2.2 for an example).

By Remark 2.12 the limit Y is an invariant random-variable under the null hypothesis of strict stationarity. That this limit is different from zero a.s. on this hypothesis follows because the indicator function of A_t is not an almost invariant function, as the set A_t is not an almost invariant set. As an example, let the sequence $\{X_t; t \in \mathbb{Z}\}$ be strictly stationary and also ergodic then the mentioned limit is exactly $\mu(A_0) > 0$.

We can formalize this idea:

Lemma 3.1 *For $\{X_t; t \in \mathbb{Z}\}$ a strictly stationary sequence and $Y_t(\omega) = \frac{1}{t}\sum_{i=0}^{t-1}\mathbf{1}_{A_i}(\omega)$ (where the set A_t is defined as above), let $D = \{\omega : Y_t(\omega) \xrightarrow{a.s.} Y(\omega), Y(\omega) \in (0,1)\}$ and define the sets,*

$$C_t = \{\omega : |Y_t(\omega)| > \epsilon\},$$

with $t = 1, 2, ...$ Then, for every $\epsilon > 0$ we have $\mu(C_t) \to 0$ as $t \to \infty$ on the set D^c (where the symbol D^c denotes for the complement of D).

Proof The sequence of sets $C_1, C_2, ...$ is monotone as it is the sequence of its complements, the latter converging to the limit set $\{\omega : |Y_t| \leq \epsilon \text{ i.o.}\} = \limsup_t C_t^c = \liminf_t C_t \subset D^c$ for every $\epsilon > 0$. But the ergodic theorem guarantees that the set D has μ-measure equal to 1 and the result follows by the continuity of μ. ■

Note that the complement set of D is a very convenient characterization of events on the hypothesis that $\{X_t; t \in \mathbb{Z}\}$ is not a strictly stationary sequence.

Example 3.1 *Let $X_t = X_{t-1} + e_t$ with $\{e_t; t \in \mathbb{Z}\}$ an i.i.d sequence with finite second moments so that $\{X_t; t \in \mathbb{Z}\}$ is a random-walk. It is well known that the series $X_n(\omega)$ diverges as $n \to \infty$ for almost every $\omega \in \Omega$ and then for some $n_0 \in \mathbb{N}$ we have that $\mathbf{1}_{A_n} = 0$ for all $n > n_0$ a.s. (that is, $\omega \in A_n^c$ for all n except on a finite number, or equivalently $\omega \in \liminf_n A_n^c \subset \liminf_n C_n^c \subset \limsup_n C_n^c \subset D^c$). Clearly for this sequence, all the sets of the form $\{\omega : (X_0(\omega), X_1(\omega), ..., X_{n-1}(\omega)) \in B\}$ with B an arbitrary Borel set are not in D.*

Weak Stationarity

For sequences in $\mathcal{L}_2$ a very competitive hypothesis for strict stationary is that of weak (or second-order) stationarity. In fact, every strictly stationary sequence in $\mathcal{L}_2$ is also weakly stationary but the converse is not necessarily true[1]. A random sequence $\{X_t; t \in \mathbb{Z}\}$ with finite second moments is said to be weakly stationary if it satisfies that cov(X_t, X_{t+k}) does not depend on t (although it can depend on k) for all $t, k \in \mathbb{Z}$. These sequences are defined on a Hilbert space where the concept of inner product coincides with that of cov(X_t, X_{t+k}). Thus, the property of weak stationarity is related to the existence of a shift transformation S which preserves the inner product (or that S is an isometry) in the sense that cov$(X_0, X_k) =$ cov(X_t, X_{t+k}) for all $k \in \mathbb{Z}$. For this to be possible, S must necessarily be unitary (i.e., $\mathrm{S}^{-1} = \mathrm{S}^*$ where S^* is the adjoint operator of S) (see Pourahmadi (2001), p.163).

We are going to show now that the class of events formed by the time-series of zero-mean weakly stationary (but not strictly stationary) sequences cannot be contained in the set D. Note that if this is the case then by Lemma 3.1 the sequence formed by the partial averages of indicator functions on the sets $A_0, A_1, ...$ converge in probability to zero. By Theorem 6.6.2 in Resnick (1999) p. 194 this further implies that the sequence $Y_1, Y_2, ...$ also converges to zero in $\mathcal{L}_2$-norm as it is a uniformly integrable sequence, where the latter holds because the respective sequence of indicator functions is a weakly stationary sequence[2] and therefore it is bounded in $\mathcal{L}_2$ (see Resnick (1999), p. 184). Thus, we need to show that the convergence of Y_t to a zero limit holds (in $\mathcal{L}_2$-norm) for every weakly stationary sequence.

By weak stationarity the indicator functions on the sets $A_0, A_1, ...$ have a spectral representation (see Remark 3.1 below) and their sequence of partial averages $Y_1, Y_2, ...$ converges in $\mathcal{L}_2$-norm to a random variable $Y(0) - Y(0^-)$ such that $\|Y(0) - Y(0^-)\|_2^2 = H(0) - H(0^-)$, where $H(.)$ is the spectral distribution function of $\{X_t; t \in \mathbb{Z}\}$ and the symbol $z(0^-)$ denotes for the limit on the left of z at 0 (see Theorem 6.1 in Doob (1953), p. 489). Note that if $H(.)$ is absolutely continuous with respect to the Lebesgue measure, so that $H(.)$ has a density, then $H(0) - H(0^-) =$

[1]For a Gaussian random sequence, weak stationarity and strict stationarity are equivalent concepts.

[2]For if F is the distribution function of $\mathbf{1}_{A_t}$ then it is indexed with a unique parameter namely $\mu(A_t)$ which can depend on k but not on t (otherwise the sequence $\{X_t; t \in \mathbb{Z}\}$ cannot be weakly stationary).

0 which immediately implies that $Y(0) - Y(0^{-}) = 0$ in the sense of $\mathcal{L}_2$-norm. For example, this is accomplished if the autocovariances of the sequence of indicator functions above are absolutely summable. But we can say more, as for any weakly stationary sequence with autocovariances converging to zero the limit in $\mathcal{L}_2$-norm of $Y(0) - Y(0^{-})$ is also zero (see Doob (1953), p. 490) and the result holds for any weakly stationary sequence purely non-deterministic.

To complete the proof one only must note that every weakly stationary sequence can be decomposed by the Wold theorem on the (direct) sum of two mutually orthogonal sequences, one of them is a purely non-deterministic sequence and the other is a sequence whose typical element is invariant to a shift transformation (for details, see Theorem 10.1 in Pourahmadi (2001), p. 352).

Remark 3.1 *Every weakly stationary sequence $\{X_t; t \in \mathbb{Z}\}$ has a spectral representation (see Doob (1953), p. 481). This representation is a stochastic integral of the form,*

$$X_t = \int_{[-\pi,\pi]} e^{it\lambda} dY(\lambda),$$

where $\{Y(\lambda), \lambda \in [-\pi, \pi]\}$ is a right-continuous stochastic process with orthogonal increments which is uniquely associated to $H(.)$.

Remark 3.2 *The sequence $\{X_t; t \in \mathbb{Z}\}$ is weakly stationary if and only if there exists a unique unitary isometry* S *acting on the subspace $\mathcal{H}$ (where $\mathcal{H}$ is a separable Hilbert space) such that $X_t = \mathrm{S}^t X_0$ for all $t \in \mathbb{Z}$ and some $X_0 \in \mathcal{H}$ (see Theorem 9.16 in Pourahmadi (2001), p. 323). In particular,* $\mathrm{S} = \int_{[-\pi,\pi]} e^{i\lambda} dH(\lambda)$ *and we also have,*

$$X_t = \int_{[-\pi,\pi]} e^{it\lambda} dH(\lambda) X_0,$$

with $X_0 \in \mathcal{H}$ (see Pourahmadi (2001), p. 337).

Remark 3.3 *The isometry* S *is induced by a measure-preserving set transformation U with domain limited to $\mathcal{H}$, it is unitary if U has an inverse (see Doob (1953), p. 461).*

3.3 Testing for Strict Stationarity

In the previous section we built a theory under which the hypothesis that a time-series has been drawn from a strictly stationarity sequence was identified in terms of the limit behavior of partial averages on some convenient set indicators. Our objective now is to propose a test statistic which can be used to sustain a valid inference about the null hypothesis of strict stationarity.

As a starting point note that the sequence formed by the indicator functions on the sets $A_0, A_1, ...$ is strictly stationary, as it is clearly a measurable transformation of the sequence $\{X_t; t \in \mathbb{Z}\}$ which under the null hypothesis is strictly stationary (see Remark 2.3). For this hypothesis, note that if k is fixed then $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{1}_{A_{i,k}^B}$ is a natural test statistic for evaluating the empirical distribution as an estimator of $\mu(A_{0,k}^B)$, which is the probability measure of the set $A_{0,k}^B$ under the model $\{X_t; t \in \mathbb{Z}\}$. Because the latter probability equals zero only if $A_{0,k}^B$ is an almost invariant set, even this simple test is informative about the hypothesis of strict stationarity. The test statistic we propose, basically uses the same structure than this simple test although it considers a more complete collection of sets. For a model $\{X_t; t \in \mathbb{Z}\}$, our test statistic is based on this quantity[3]:

$$S_n = n^{-1}\sum_{i=1}^{n}\sum_{k=1}^{\kappa_n}\mathbf{1}_{A_{i,k}^B}$$

The role played by κ_n on the limit behavior of this double sum is quite fundamental, our approach is to take $\kappa_n = [\kappa\, n]$ where $\kappa \in (0,1)$ and the symbol $[.]$ denotes for the integer part of its argument. The parameter κ has to be chosen exogenously to the information provided in the sample data, although its value will not affect the results regarding the power and size of the test (at least for large samples). In effect, we prove below that $\frac{1}{\sigma\sqrt{\kappa}}(S_n - \mathbb{E}(S_n)) \Rightarrow_n \mathcal{W}(\kappa)$ where $\mathbb{E}$ is the respective expectation operator, $\sigma > 0$ and $\mathcal{W}(\kappa)$ is the standard Wiener measure evaluated at κ. By the properties of the Wiener measure, the transformation of S_n given by $\frac{1}{\sigma^2\kappa^2}(S_n - \mathbb{E}(S_n))^2$ is asymptotically distributed as χ_1^2 under the null hypothesis of strict stationarity, where χ_1^2 denotes for a chi-squared random variable with one degree of freedom. For this transformation, the respective limit distribution does not depend on κ.

[3]We do not distinguish between models for which $\mathbb{E}(S_n)$ has the same value and correspondingly, we should refer $\{X_t; t \in \mathbb{Z}\}$ as representing a class of equivalence.

Proceed as follows:

For every $t \in \mathbb{N}$ collect the $k \times 1$ vector of indicator random variables $(\mathbf{1}_{A^B_{t,1}}, ..., \mathbf{1}_{A^B_{t,k}})$. Associate to each of these vectors another vector with elements $(\xi_{t_1}, \xi_{t_2}, ..., \xi_{t_k})$ where $\{\xi_{t_k}; k = 1, 2, ..., \kappa_t \text{ and } t \in \mathbb{N}\}$ is a triangular array of i.i.d normal random variables with zero-mean and variance equal to σ^2. Define the sequence $\xi = (\xi_{1_1}, \xi_{2_1}, \xi_{2_2}, ...)$ which is obtained by stacking the rows of the above triangular array:

$$\begin{array}{c} \xi_{1_1} \\ \xi_{2_1}, \xi_{2_2} \\ \vdots \\ \xi_{t_1}, \xi_{t_2}, ..., \xi_{t_k} \\ \vdots \end{array}$$

where $k = 1, 2, ..., \kappa_t$ and $t \in \mathbb{N}$.

Lemma 3.2 *Let* $S'_n = n^{-1} \sum_{i=1}^{n} \sum_{k=1}^{\kappa_n} \xi_{i_k}$. *Then,* $\frac{1}{\sigma\sqrt{\kappa}} S'_n \Rightarrow_n \mathcal{W}(\kappa)$.

Proof Re-write $\frac{1}{\sigma\sqrt{\kappa}} S'_n = \frac{1}{\sigma n \sqrt{\kappa}} \sum_{i=1}^{n} \sum_{k=1}^{\kappa_n} \xi_{i_k} = \frac{1}{\sigma\sqrt{\kappa n^2}} \sum_{i=1}^{[\kappa n^2]} \zeta_i$ where ζ_i denotes the $i-$th coordinate random variable of ξ. Because $\zeta_1, \zeta_2, ...$ is by construction a sequence of i.i.d normal random variables, the result follows by Remark 2.19 (noting that $\{n^2; n = 1, 2, ...\}$ is a subsequence of $1, 2, ...$).
■

We show now that S'_n can be used to approximate the statistic S_n in probability. To do this, define $\xi'_{t_k} = \mathbf{1}_{A^B_{t,k}} - \mu(A^B_{0,k}) - \xi_{t_k}$. As before, form the sequence ξ' which is obtained by stacking the rows of the triangular array $\{\xi'_{t_k}; k = 1, 2, ..., \kappa_t \text{ and } t \in \mathbb{N}\}$ and denote as ζ'_i the $i-$th coordinate random variable of ξ'. The sequence $\{\zeta'_t; t \in \mathbb{N}\}$ is zero-mean strictly stationary under the null hypothesis and it is in $\mathcal{L}_2$ by construction[4], but then we have $\mu(|\sum_{i=1}^{n} \zeta'_i| > n) \to 0$ as $n \to \infty$ by the ergodic theorem (see Lesigne and Volný (2000), p. 75). It is immediate from this result and Lemma 3.2 that $\frac{1}{\sigma\sqrt{\kappa}}(S_n - \mathbb{E}(S_n)) \Rightarrow_n \mathcal{W}(\kappa)$ where $\mathbb{E}(S_n) = \sum_{k=1}^{\kappa_n} \mu(A^B_{0,k})$ and also that

[4] The parameter σ can be chosen in order to minimize the variance of ξ'_{t_k}.

$Z_n = \frac{1}{\sigma^2\kappa^2}(S_n - \mathbb{E}(S_n))^2 \Rightarrow_n \chi_1^2$ by the properties of the Wiener measure. In this context, a testing procedure for the null hypothesis that a time-series has been drawn from a strictly stationary sequence can be stated on the values of $\mathbb{E}(S_n)$, noting that under the null hypothesis this quantity is strictly greater than zero whilst under the alternative hypothesis it is equal to zero. Thus, a decision can be made by evaluating if S_n is in the neighborhood of zero at some significance level.

The problem above is not standard in that such a procedure will not render a test with the same level α for every possible model in the null hypothesis (only for models within a given equivalence class), nevertheless a level α test can be obtained yet through randomization yet by a (linear) transformation of the original sequence. We take these two approaches in the next Chapter, until then we note that Z_n has some important properties (and drawbacks) as a test statistic that we discuss now.

Let $\beta(\cdot, a)$ be the power function for a test based on Z_n, then $\beta(z, a)$ is the cumulative distribution function of a (generalized) chi-squared random-variable with one degree of freedom evaluated at $Z_n = z$ and $\mathbb{E}(S_n) = a$. As a function of a, $\beta(\cdot, a)$ is monotonically decreasing on this argument (i.e., $a \leq b$ implies that $\beta(\cdot, a) \leq \beta(\cdot, b)$) because a is a location parameter in this context and the result follows from noting that $\frac{d\beta(\cdot,a)}{da} \leq 0$. Also note that taking the supremum of $\beta(\cdot, a)$ over the set $a > 0$, by the previous property it follows that the test is a level α test. Moreover, the test is unbiased: fix α and a^* such that $\beta(\cdot, a^*) = \alpha$. Then, if $a \geq a^*$ we have $\beta(\cdot, a) \leq \alpha$ and also that if $a \leq a^*$ we have $\beta(\cdot, a) \geq \alpha$; the result follows simply noting that the set $a > a^*$ characterizes the null hypothesis of strict stationarity for a test of level α (for a given class of models).

Clearly, such a test is invariant to a location and scale (i.e., linear) transformation of the underlying strictly stationary sequence and more generally, to any measurable transformation of it. However, the test is not robust to near-integrated sequences. These are (strictly) stationary sequences in $\mathcal{L}_2$ having a near unit-root as a solution of the shift polynomial on their autoregressive representation. That is, by the Wold theorem any strictly stationary sequence $\{X_t; t \in \mathbb{Z}\}$ in $\mathcal{L}_2$ as a representation that can be approximated by an autoregressive model of the form $\theta(S)X_t = u_t$ where $\theta(S) = (1 - \theta_1 S - \theta_2 S^2 - ... - \theta_n S^n)$ is the shift polynomial and $\{u_t; t \in \mathbb{Z}\}$ is a sequence of i.i.d random variables in $\mathcal{L}_2$ (not necessarily Gaussian). Stationarity implies that the sum given by $\theta_1 S + \theta_2 S^2 + ... + \theta_n S^n$ is convergent as $n \to \infty$ and near-integration means that this sum approaches to the unity.

It follows that an autoregressive sequence can also be used to approximate in a sense[5] an integrated sequence (see Saïd and Dickey (1984) for a proof in the case of ARIMA sequences).

The problem of near-integrated sequences in the context of unit-root tests is studied in Phillips (1988), where it is reported a lowering in the power of the test under the hypothesis of stationarity. In our context, near-integration causes the test to erroneously reject the null hypothesis for a strictly stationary sequence with autoregressive parameters satisfying the sum above being close to the unity[6].

Example 3.2 *An example of a near-integrated sequence is a sequence X having as its elements random variables of the form $X_{t,T} = a_T X_{t-1,T} + u_t$ where $a_T = \exp(T^{-1}c)$ for a fixed real c and $\{u_t; t \in \mathbb{N}\}$ is an i.i.d sequence with finite second moments. For $|a| < 1$ the sequence is strictly stationary and ergodic only if T is fixed, otherwise X is a triangular array and we set $t = 1, 2, ..., T$ with $T \to \infty$. Note that for every $\epsilon > 0$ it is verified that $\mu([X_{t,T'} > \epsilon]$ for all $T' \geq T) \to 1$ as $T \to \infty$ so that the sequence X diverges almost surely. Then, we can apply the result in Example 3.1 to show that for such a sequence we can falsify the null hypothesis of strict stationarity.*

[5] The approximation is in probability and it requires that the order of autoregression be a function of n.

[6] But how close? Schwert (2002) has showed that even for the simple first-order autoregression model $X_t = \theta X_{t-1} + u_t$ the problem of erroneously detecting a unit-root can arise with $\theta = 0.8$ or higher. We confirmed this cut-off value in our Monte Carlo experiment (see Section 4.1).

Chapter 4

Ilustrations

In this Chapter we present two approaches for controlling the size of our test for the null hypothesis of strict stationarity, one is a randomized level α test and the other is a test of size α that is based on the residuals of an autoregression. Their small sample properties are evaluated through a Monte Carlo experiment, then we present two applications with real-data analysis.

4.1 A Monte Carlo Experiment

The objective of the Monte Carlo experiment we detail in this section is to evaluate the small sample properties of our test statistic. Consider first the role that play both the parameter κ and the sample size N on the empirical distribution of our test statistic. The analysis is performed for the i.i.d standard normal random sequence. In Figure 4.1 we have plotted the estimated density of the test statistic S_n for three different values of the parameter κ and three different sample sizes (for a fixed sample size we write S_N). Every density plot was obtained by drawing 1000 replications from the underlying random sequence. For the different values of κ the distribution of the test statistic is unimodal and symmetric around a mean value (see Figure 4.1a). Higher values of this parameter seem to locate the distribution of the test statistic to the left and favor the alternative hypothesis, eventually rendering a more challenging test. As these values are rather close, in practice we should prefer to use a higher κ on more complex models where the additional information can be more valuable. For the rest of this experiment, the value of κ is maintained fixed at 0.1.

The effect on the distribution of the test statistic of an increase in the sample size is quite evident from Figure 4.1b, as the distribution now locates increasingly away from the density of a chi-squared distribution with one degree of freedom (the solid red line). The latter distribution corresponds to the limit distribution of the test statistic when is evaluated at $\mathbb{E}(S_N) = 0$ (i.e., under the alternative hypothesis) so that this result is a graphical confirmation that the test is consistent for this model[1].

The next step is to evaluate the range of values that $\mathbb{E}(S_N)$ can take on the different models in the null hypothesis. This is important information regarding the randomization procedure we use below to obtain a level α test, as it helps us to define the relevant area of integration needed to compute the density under which the test is performed. But also because there will be clear similarities for several models in the null hypothesis suggesting to treat them as members of a single equivalence class, and for which a common testing procedure of size α can be devised.

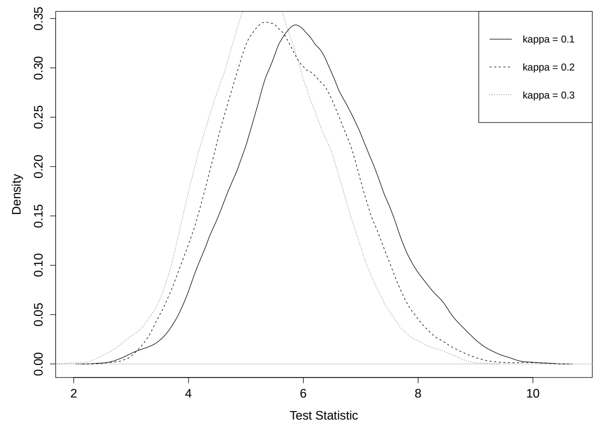
i.i.d Models and a weakly dependent Model. We considered in our experiment several i.i.d models jointly with the normal model: t-Student, Beta, Skew-Normal, Weibull and Uniform; and also an important weakly dependent model[2]: the GARCH model. These models allow for a variety of supports and different forms of the probability densities, what is explored by changing the respective scale and/or shape model parameters[3]. Most importantly, they all share some common properties regarding the distribution of the test statistic suggesting to consider them into a single equivalence class, for which the empirical size of a test of size α can be evaluated against each alternative.

For the experiment there were drawn 1000 simulations of each model on every choice of parameter values and the sample size was fixed at $N = 100$. In Figure 4.2 below we show the results for the i.i.d. normal model and the GARCH model, the plots are indeed quite similar for the rest of the models. The distribution of the test statistic is unimodal and symmetric around a nearly common mean value, it locates all its probability mass away from zero and it is robust to changes in the (scale

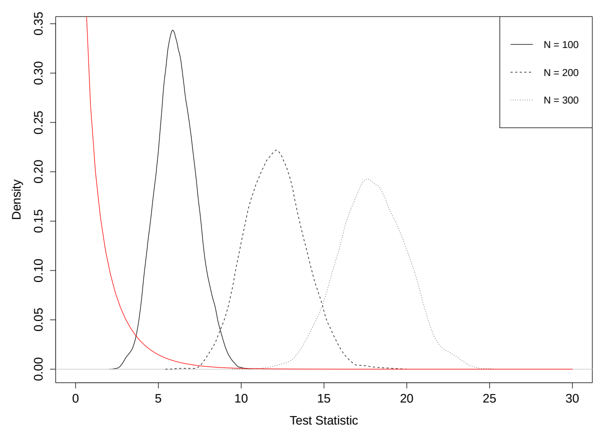
[1]The result was also verified for the other models in the null hypothesis considered in this experiment. For the models in the alternative hypothesis, the respective distributions locate in a neighborhood of zero where the test has significant power.

[2]This is a class of time-dependent models that allows for no-serial correlation and different forms of asymptotic independence.

[3]It is assumed that the time-series models under analysis have all zero-mean.



(a) Different values of κ.

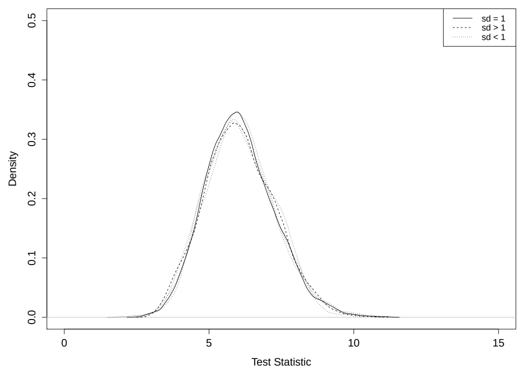


(b) Different values of N.

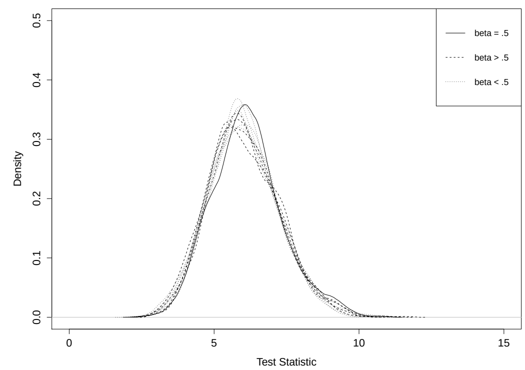
Figure 4.1: Density estimation of S_n for the i.i.d standard normal random sequence.

and/or shape) model parameters. The point here is that the condition of no-serial correlation (and not that of independence) seems to determine the main properties regarding the distribution of the test statistic that are shared by the models. We will return to this test later, when we consider a testing procedure for analyzing the residuals from a linear autoregression.

Time-Dependent Models. We study now the effect of time-dependence (in the form of serial correlation generated by linear or non-linear models) on the probability distribution of S_N. For this purpose, we considered seven different classes of models (all in $\mathcal{L}_2$) which are of main interest for



(a) Normal, $X_t = \sigma\xi_t; \sigma = \{.01, .1, 1, 10, 100\}$



(b) GARCH: $X_t = \sigma_t\xi_t,\ \sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2;$ $\omega > 0,\ \alpha + \beta < 1$

Figure 4.2: Density estimation of S_n for some members of the i.i.d. family.

time-series analysis: moving average (MA), autoregressive (AR), fractional noise (FN), fractional gaussian noise (FGN), nonlinear moving average (NLMA), Bilinear (BL) and smooth threshold autoregressive (STAR) models. The first four classes of models allow for time-varying linear conditional mean but constant conditional variance[4], the rest of the models can exhibit time-varying conditional mean and/or variance depending on the parameter specification.

Because every strictly stationary sequence in $\mathcal{L}_2$ has also a Wold representation, in a sense, all the time-dependent models we have considered in the present experiment also share some common properties. In particular, the only kind of non-strict stationarity which is admitted in this context $-$that is, regarding the values of the model parameters, is that derived from the existence of a unit-root in the autoregressive part of the shift polynomial. As a consequence, those models having an autoregressive component are all subject to the problem of near-integration for some values in the parameter space (in the set defining a strictly stationary sequence).

For $N = 100$ the results are shown in Figure A.1 in the Appendix. As before, 1000 simulations of each model were drawn for different values of the model parameters, now in the range of values admitting the random sequence to be strictly stationary (i.e., having no unit-roots and finite second-

[4]Of these models, MA, AR and FN can be grouped into a single class, the ARFIMA class, but it is convenient to consider them separately.

moments). Time-dependence allows for a higher range of values of $\mathbb{E}(S_N)$, but these values appear to be contained within a bounded set (with bounds depending on the sample size). In models with an autoregressive component, the closer is the value of the autoregressive parameters to a unit root, the lower is the range of values that the test statistic can take with positive probability. For a fixed critical value, this clearly will cause the test to be oversized in the zone of near-integration.

Now, we are in position to implement some strategies to control the size of the test. We follow two main approaches, the first one is a randomized testing procedure which exploits a simplification of the (composite) null hypothesis of strict stationarity rendering a level α test. The other approach is based on a linear transformation of the underlying random sequence and exploits the results obtained above for the i.i.d normal sequence. From this second approach we can obtain a test of size α.

Randomization

In order to get a test of level α our first approach is based on reducing the null hypothesis (which in the present case is a composite hypothesis) into a simple one, and this is done through randomization. In particular, we will base our decision on the hypothesis that the test statistic Z_n has distribution function with density equal to

$$h(z) = \int_{(0,\infty)} \frac{d\beta(z, \mathbb{E}(S_N))}{dz} d\Lambda(\mathbb{E}(S_N)), \text{ where } z > 0$$

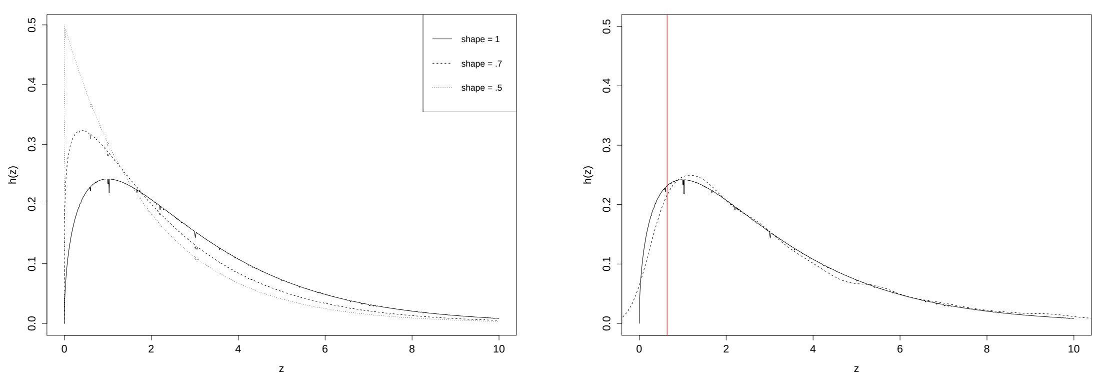
The existence of Λ is not guaranteed in the general case. In our context, however, the problem has a solution if we verify that $\lim_{n\to\infty} \beta(z, \mathbb{E}(S_n)) \to 0$ for each fixed value of z (see Lehmann and Romano (2005), p. 86). That the result is true follows by the monotone convergence theorem and because $\mathbb{E}(S_n)$ diverges as $n \to \infty$ under the null hypothesis of strict stationarity[5].

It is still required to specify a suitable distribution function Λ. In practice, however, this does not convey any problem noting that if Λ is chosen to have a density then numerical methods can be

[5]For if not then $\mu(A^B_{0,k} \text{ i.o.}) = 0$ by the Borel-Cantelli Lemma and $\mu(A^B_{0,n}) \to 0$ as $n \to \infty$ but this contradicts the null hypothesis of strict stationarity.

used to tabulate the values of the cumulative probability under $h(z)$ −provided that the respective integral converges, so that a critical value can easily be obtained in order to perform a test of level α for rejecting the null hypothesis of strict stationarity (based on that particular choice).

From the analysis in the previous section, we have that our choice of Λ can be reduced to distributions having density and with support on $(0, C_N]$ with $C_N > 0$ a fixed constant. The gamma distribution provides in this context a flexible choice for Λ, but other distribution functions can also be used. For instance, a value of 2 for the scale parameter of the gamma distribution will render densities with support consistent to those obtained in the simulations with $N = 100$. The shape of the density will reflect the belief of the researcher on the hypothesis that the sequence under evaluation is strictly stationary, a more challenging test will put more probability mass on lower values of the test statistic.



(a) Plots for $h(z)$ under different values of the shape parameter in Λ.

(b) Empirical density (dashed line) for $h(z)$. The vertical line is at $z = .64$.

Figure 4.3: Density of S_n for a test of level α with $N = 100$.

As an example, in Figure 4.3a we have plotted three of the possible densities resulting by changing the values of the shape parameter in the specification of Λ as a gamma distribution, with its scale parameter fixed at the value of 2. For each of these densities we can easily compute a critical value conducing to a test of level α. In effect, in Figure 4.3b we have plotted the empirical distribution that is obtained after sampling from one of the former alternative densities. For this distribution, a test of level $\alpha = 0.1$ is obtained with a critical value of 0.64 for a sample size of $N = 100$. This critical value will be used in the next section for an application with real-data.

A Unified Testing Procedure

The great flexibility provided by the randomized test has as a main drawback, the arbitrariness regarding the choice of the Λ distribution. This is because the choice of Λ has to be justified on a case by case basis. However, on the hypothesis that a strictly stationary model has been generated by passing a linear or non-linear (but measurable) filter over an i.i.d random sequence within the class of models described in this experiment, we have available an alternative testing procedure that exploits the property of invariance of our test statistic. The method applies to a serially uncorrelated sequence under the hypothesis of strict stationarity (so that only a linear transformation is eventually required), and it is based on the results obtained previously for the i.i.d standard normal sequence. In practice, we could need first to filter the series under evaluation by an autoregression of order up to eliminate much of the serial correlation that would be present on it, and then to apply over the residuals of such a model the testing procedure based on the i.i.d standard normal model.

The results on the size of the test are shown in Table A.1 where it is observed that the sizes of the different i.i.d models and the GARCH model agree with that of the i.i.d standard normal model. The results on the power of this test are reported in Table A.2. The models in the alternative hypothesis are classified in four main groups. In the first group we included sequences with stochastic trends: the Random-Walk model which is a unit-root model and the Random Modulated Periodicity (RMP) model introduced by Hinich (2000), this is a model for signals which are labeled as periodic but which are not deterministic. The second group has some members of the harmonic family. The third group of models are those having spectral representation, these are very competitive models as an alternative for strict stationarity. Here we included locally-stationary models which belong to the class of non-stationary models with long-range dependence, and weakly stationary models which can be classified as block-stationary sequences. The fourth group that we called Contaminated sequences, include i.i.d sequences for which we have added a contaminating sequence in the third part of the sample. The idea behind these models is to simulate the presence of outliers in the sequence. In the Appendix we have plotted for each the above models some typical realizations, to get an idea of the difficulty behind detecting the alternative hypothesis.

4.2 Two Examples with Real-Data Analysis

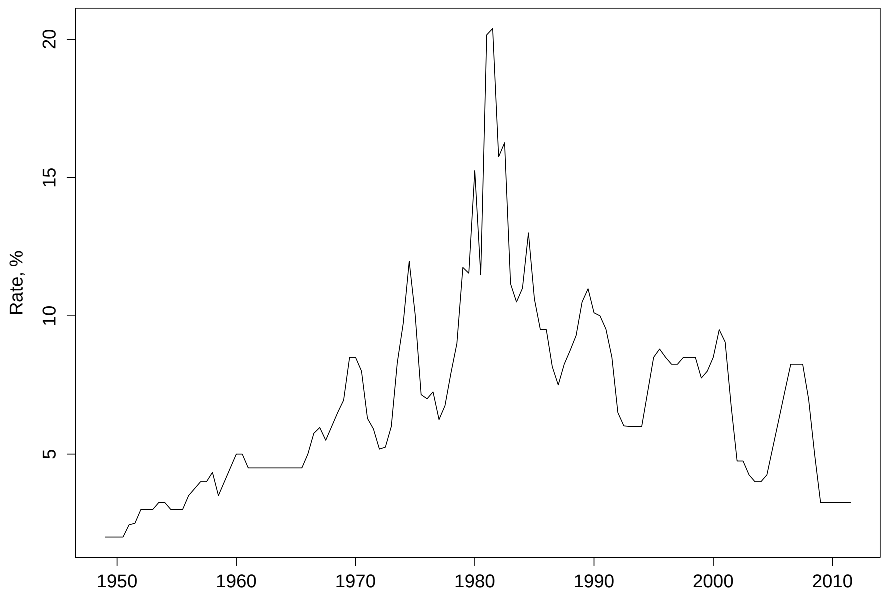
4.2.1 Interest Rate Data

The data we analyze is the Bank Prime Loan Rate (see Figure 4.4 below). The rate is posted by a majority of top 25 commercial banks and it serves to price short-term business loans. The main feature of this data is a pronounced break in the trend at 1982, which is explained by the debt crisis of less-developed-countries (LDC). The crisis began in August 1982 when it was announced that Mexico would be unable to meet its debt, mainly affecting to commercial banks in U.S. Only one year later 27 countries had rescheduled their debts to banks, 16 of these were countries from Latin America who owed roughly 74 percent of their total debt to banks in U.S.

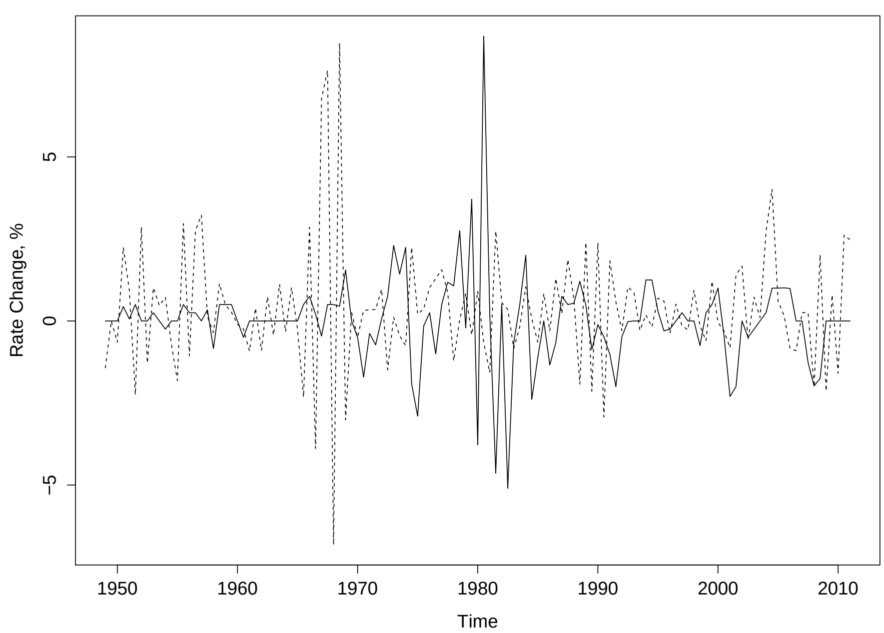
Interest rate models usually take the form of a diffusion around a possibly time-varying mean rate representing the long-term equilibrium, the basic literature on these models is Cox et al. (1985). For discrete-time data, the analogue is an integrated (or unit-root) model with increments that are assumed to be strictly stationary. Time-varying parameters are not allowed in this context, but we can consider alternatively time-varying conditional moments in order to model the dynamics of the underlying economic equilibrium. Let us consider $R_t = R_{t-1} + u_t$ as our integrated model for the interest rate, where $\{u_t; t \in \mathbb{N}\}$ is an ergodic (possibly time-dependent) strictly stationary sequence. Because the interest rate as an economic variable interacts (at the observed frequency) with other economic factors in a complex manner, there is no role for a univariate model in building some inference or forecast. On the other side, simulation of this series is quite fundamental as a tool for pricing financial instruments whose risk depends on this interest rate, as occurs with short-term business loans[6].

Now, we perform some statistical analysis that will drive us to a well specified interest rate model based on the available data on the Bank Prime Loan Rate. In Figure 4.4b we have plotted the

[6] In finance, an asset with this characteristic is called a derivative asset. Examples of these assets in the present context are interest rate forward agreements (or FRA's), swaps, warrants, etc.



(a) BPLR, Level.



(b) BPLR, First difference (solid line). Realization from a Bilinear model (dashed line).

Figure 4.4: Bank Prime Loan Rate (BPLR). Source: Federal Reserve Bank of St. Louis.

differenced time-series of this rate (the solid line), it clearly evidences the trend break of the original series as a sudden shock but it is also characterized by showing several changes in the volatility of the series. The series is serially correlated with no unit-root, as confirmed with the ADF test. Is it weakly stationary? According to the Priestley and Subba Rao (1969) test, the hypothesis of second-order stationarity is strongly rejected for this series. That is, from a statistical perspective the underlying model for $\{u_t; t \in \mathbb{N}\}$ is either strictly stationary or non-stationary[7].

The test statistic for the hypothesis of strict stationarity gives a value of 2.60 which is higher than the critical value of 0.64 for a randomized test of level 10 percent (see Section 4.1.1) and sample size of 100 observations. Using the alternative (unified) test, an autoregression of order 4 was sufficient to remove the serial correlation of the differenced BPLR series. The test is then performed over the residuals of this autoregressive model and resulted in a test statistic with a value of 5.94, which is higher than the critical value of 4.56 for a test of size 10 percent under the i.i.d standard normal model. Thus, the evidence supports that the evaluated random sequence is strictly stationary. A model that can generate a time-series with the observed characteristics of the (differenced) BPLR series is the Bilinear model, the simplest specification of such a model satisfies $(1 - \theta S)X_t = \phi X_{t-1}\xi_{t-1} + \xi_t$ where $\theta^2 + \phi^2 < 1$ and $\{\xi_t; t \in \mathbb{N}\}$ is a sequence of i.i.d standard normal random variables. For a graphical confirmation, we also plotted in Figure 4.4b (the dashed line) a single realization from a Bilinear model showing the required properties.

4.2.2 Nile River Data

One of the most studied time-series (mainly in the context of long-memory models) is that of the yearly minima water levels of the Nile river, it contains observations dated from the year 622 to the year 1284 (see Figure 4.6 below). From a statistical point of view the series has no unit-roots based on the ADF test and according to the Priestley and Subba Rao (1969) test the hypothesis of second-order stationarity is strongly rejected (as it is for the differenced series). Thus, the underlying model is that of a strictly stationary sequence or that of a non-stationary sequence. In this case we present only the results of the unified test. The series has been described elsewhere to present long-range

[7] Note that if the series is non-stationary then its simulation from a univariate model is totally precluded until we can incorporate further information regarding the possible sources of the non-stationarity.

dependence and consistent with this fact, an autoregressive model of order 20 was necessary in order to remove much of the serial correlation that was present in the series. The test statistic resulted in a value of 33.08 which is lower than the critical value of 34.7 for a test of size 10 percent under the i.i.d standard normal model with 660 observations. Regarding the density estimation of the test statistic showed in Figure 4.5, there is weak support in the direction that the Nile river data has been drawn from a non-stationary random sequence[8].

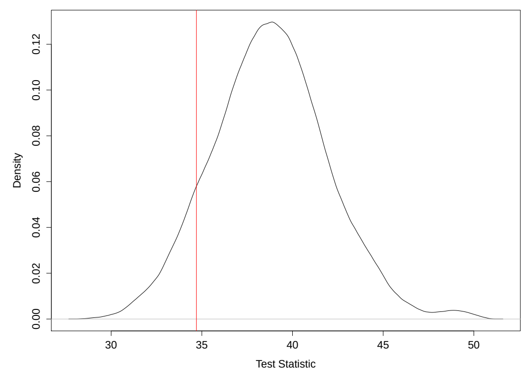
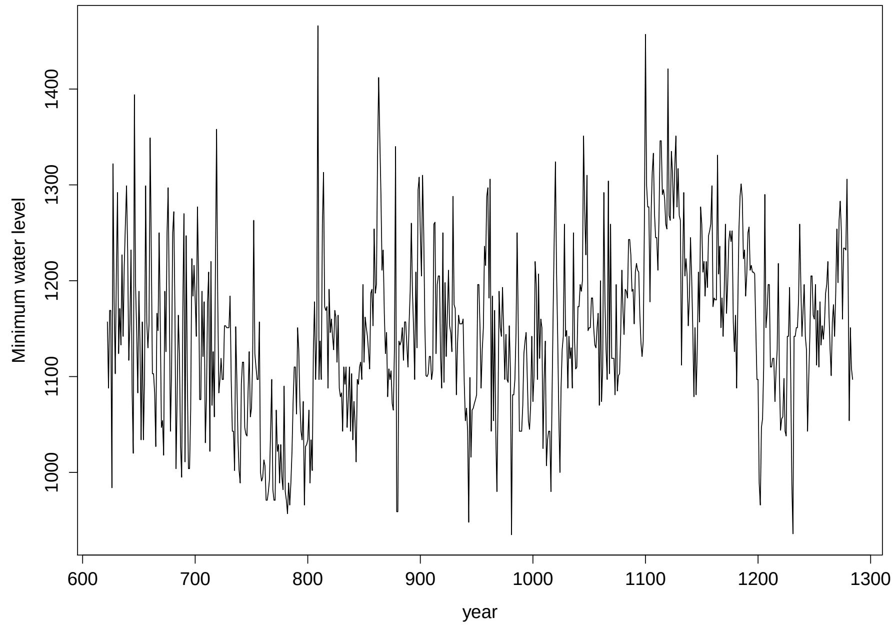
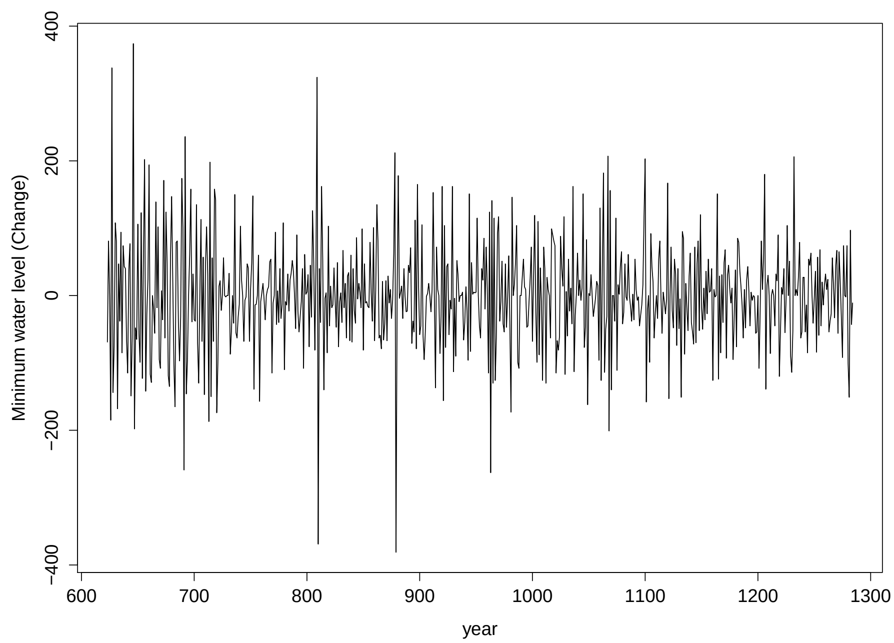


Figure 4.5: Density estimation of S_n for the i.i.d normal model with 660 observations. The vertical line is at 34.7.

[8] It must be noted then that a strictly stationary sequence can be used to draw a time-series with rather similar characteristics than that observed for the Nile river data even though the series is not second-order stationary.



(a) Nile data, Level.



(b) Nile data, First difference.

Figure 4.6: Minimum water level of the Nile river. Source: Beran (1994).

Chapter 5

Conclusions

In this dissertation we have developed a theoretical framework for representing the null hypothesis of strict stationarity in a very general setting. We have showed that the problem of adequately characterizing this hypothesis is not well stated in the literature we revised, as the most common approach is the verification of time-homogeneity for certain statistics, but this is shown to be a rather questionable reduction of the problem of inference involved.

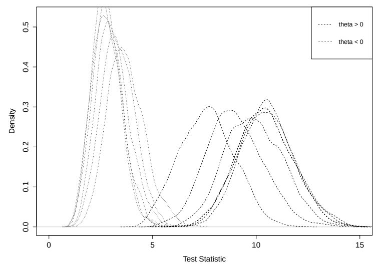
The results obtained allowed us to derive a non-randomized test of level size which is endowed with good statistical properties. The test has been obtained under the choice of a naive estimator of the parameter of interest for the decision, but even this simple version have shown to have good power on some important model alternatives as those having deterministic or stochastic trends, where the time-series can have divergent or recurrent patterns. Those results are presented through a small Monte Carlo experiment, where we also studied the empirical size of the test. The problem of near-integration we found justify our proposal of a robust version of the same test.

Finally, we note that the most important consequence derived from the results of the present dissertation is the possibility to devise a statistical procedure for the decision whether a time-series has been drawn from a weakly-stationary sequence or from a strictly stationary sequence. Weak stationarity is a very competitive alternative for strict stationarity and at the current knowledge of this author there is no such test in the literature. The subject is, however, quite delicate and we prefer to leave that line of research for future work.

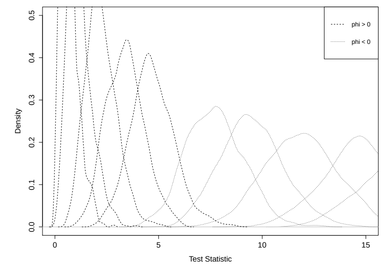
Appendix A

Graphics (Monte Carlo experiment)

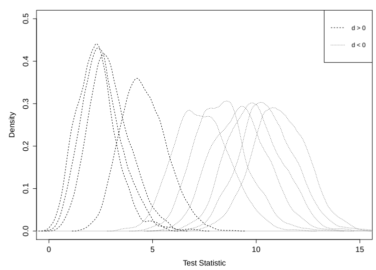
In this Appendix we show some additional plots resulting from our Monte Carlo experiment. Figure A.1 complements the information provided in Section 4.1 regarding the empirical distribution of the test statistic for some time-dependent models in the null hypothesis of strict stationarity (different realizations in a single plot correspond to alternative model parametrizations). In Figures A.2-A.7 we considered three model parametrizations (labeled a, b and c) for each model included in Table A.2 (power of the test) and we show three time-series from those models on each plot.



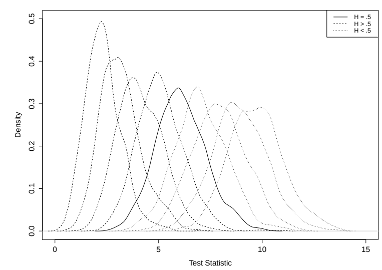
(a) MA: $X_t = (1 - \theta S)\xi_t$; $\theta, \in \mathbb{R}$



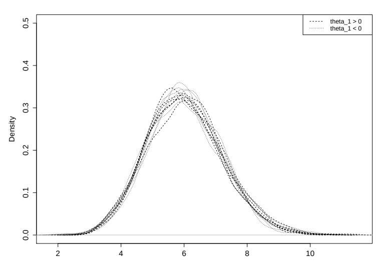
(b) AR: $(1 - S\phi)X_t = \xi_t$; $\phi, \in (-1, 1)$



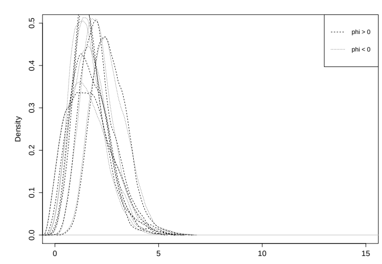
(c) FN: $(1 - S)^d X_t = \xi_t$; $d \in [-1.2, \frac{1}{2})$



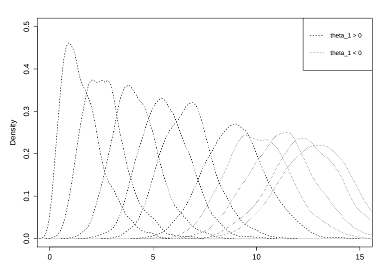
(d) FGN: $X_t = \mathcal{W}(t + 1) - \mathcal{W}(t)$



(e) NLMA: $X_t = \theta\xi_{t-2}\xi_{t-1} + \xi_t$; $\theta \in \mathbb{R}$.

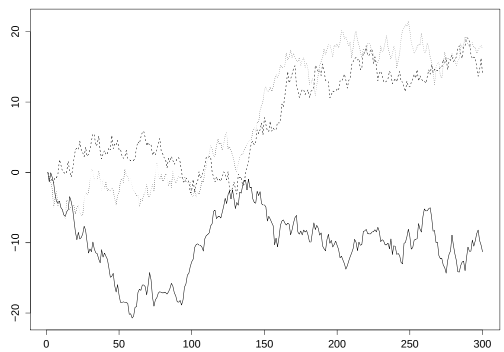


(f) Bilinear: $(1 - \theta S)X_t = \phi X_{t-1}\xi_{t-1} + \xi_t$; $\theta^2 + \phi^2 < 1$

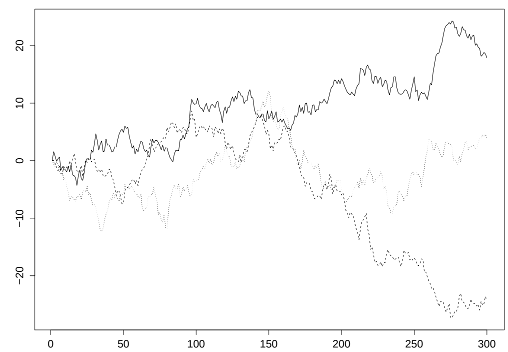


(g) STAR: $(1 - \theta^j S)X_t = \xi_t$; $\theta^j \in (-1, 1)$, $j = 1, 2$; Prob$(j = 1) = g(X_{t-1})$

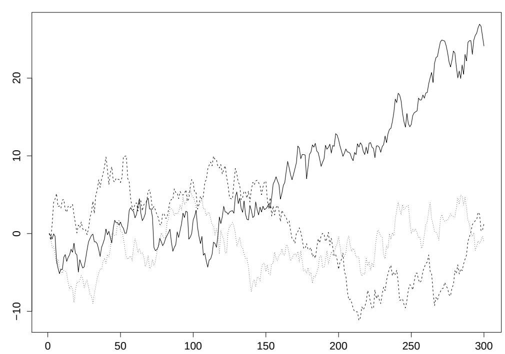
Figure A.1: Density estimation of S_n for some time-dependent families (in $\mathcal{L}_2$).



(a)

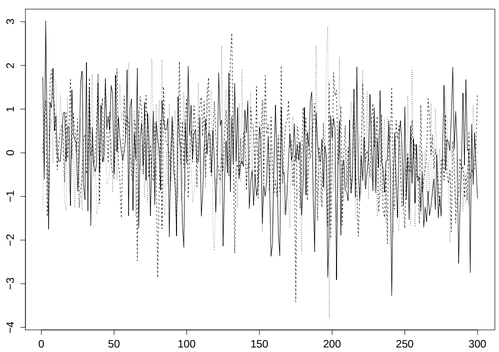


(b)

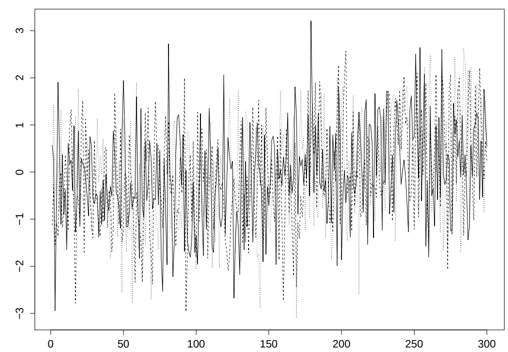


(c)

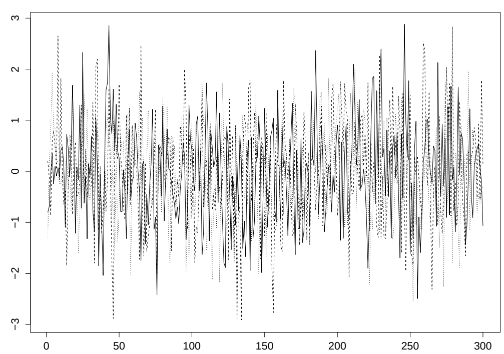
Figure A.2: Time-Series from the Random-Walk model.



(a)

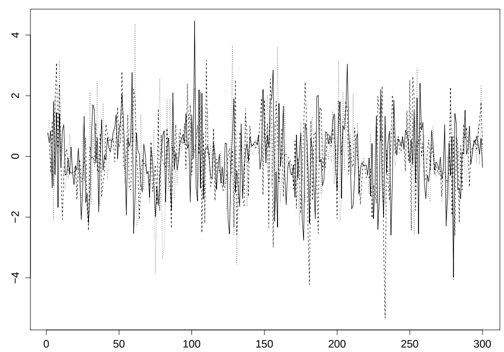


(b)

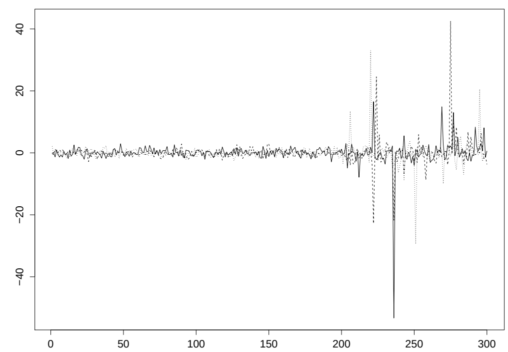


(c)

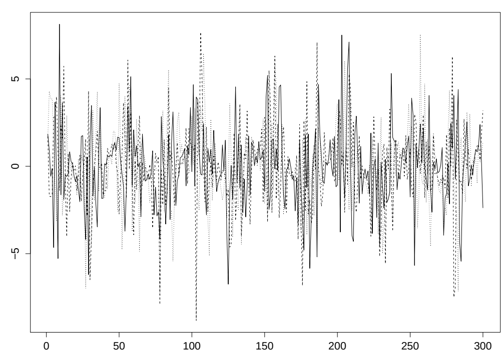
Figure A.3: Time-Series from the Harmonic model.



(a)

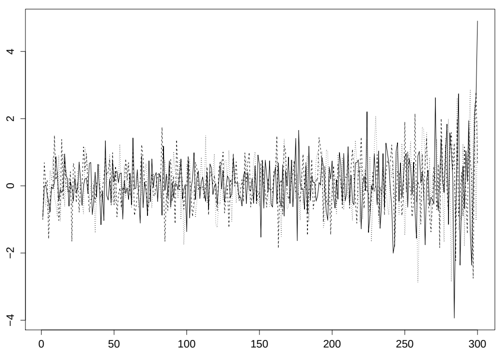


(b)

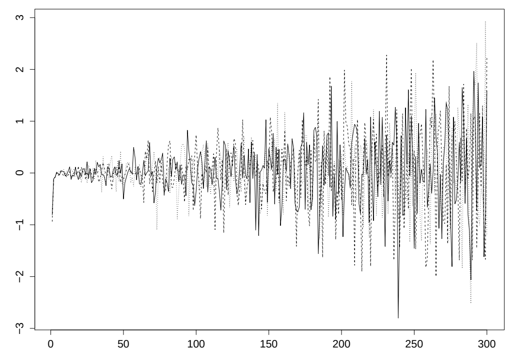


(c)

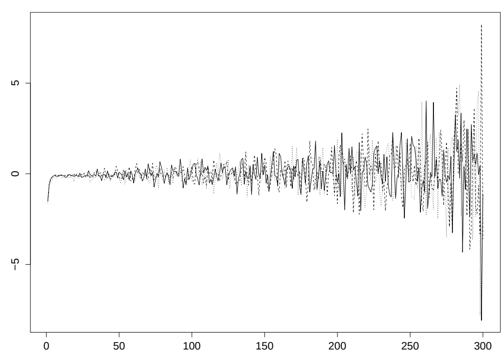
Figure A.4: Time-Series from the Randomly Modulated Periodicity model.



(a)

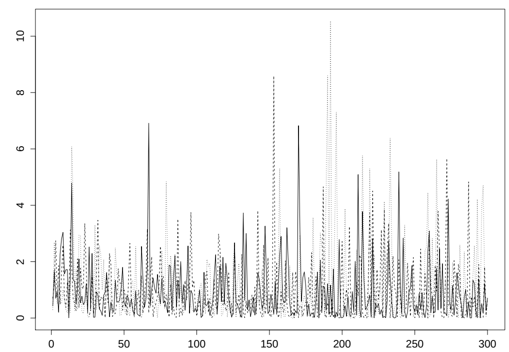


(b)

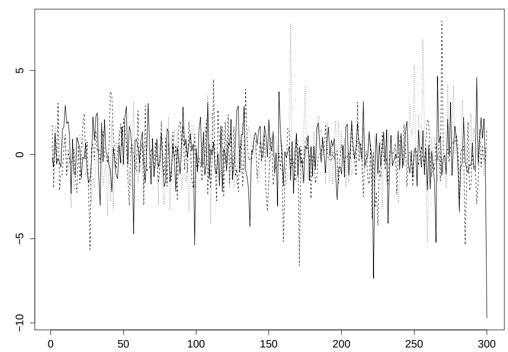


(c)

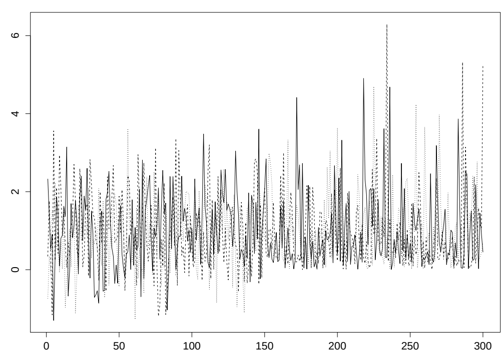
Figure A.5: Time-Series from the Locally-Sationary model.



(a)

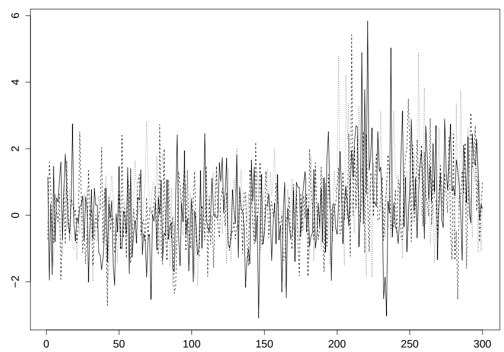


(b)

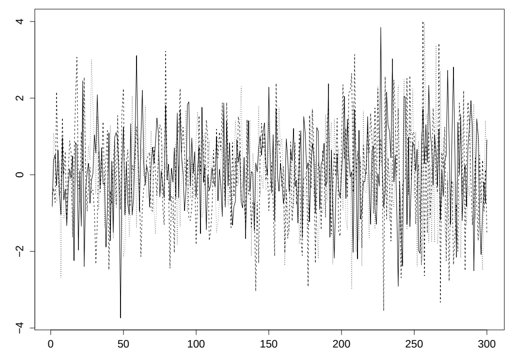


(c)

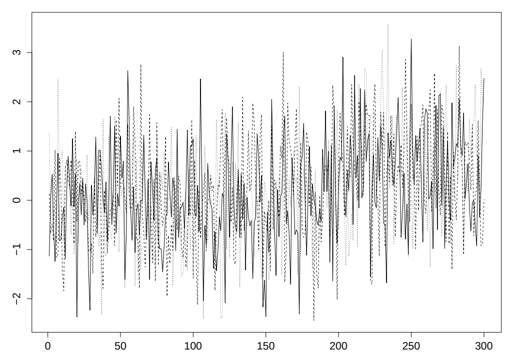
Figure A.6: Time-Series from the White-Noise model.



(a)



(b)



(c)

Figure A.7: Time-Series from the Contaminated model.

Table A.1: Size of the Test (i.i.d normal test)

Model	Sample Size	Mean	S.D.	Size ($\alpha = 0.10$)	Size ($\alpha = 0.05$)
i.i.d normal	100	6.04	1.19	0.10	0.05
	200	11.99	1.63	0.10	0.05
	300	17.92	2.13	0.10	0.05
i.i.d t-student	100	6.05	1.20	0.08	0.04
	200	11.91	1.68	0.13	0.07
	300	17.79	2.04	0.10	0.06
i.i.d beta	100	6.03	1.22	0.10	0.04
	200	11.61	1.66	0.17	0.10
	300	17.63	2.06	0.12	0.06
i.i.d skew-normal	100	6.04	1.16	0.08	0.03
	200	12.05	1.66	0.11	0.06
	300	17.85	1.97	0.08	0.04
i.i.d weibull	100	5.94	1.22	0.10	0.05
	200	11.98	1.71	0.11	0.06
	300	18.03	2.16	0.09	0.05
i.i.d uniform	100	6.09	1.20	0.08	0.04
	200	12.06	1.72	0.11	0.06
	300	17.84	2.17	0.10	0.06
GARCH	100	6.06	1.18	0.08	0.03
	200	12.02	1.69	0.13	0.07
	300	17.90	2.11	0.09	0.05

Notes:

(1) $N = 100$. Critical Value ($\alpha = 0.1$) = 4.5; Critical Value ($\alpha = 0.05$) = 4.05

(2) $N = 200$. Critical Value ($\alpha = 0.1$) = 10.0; Critical Value ($\alpha = 0.05$) = 9.52

(3) $N = 300$. Critical Value ($\alpha = 0.1$) = 15.19; Critical Value ($\alpha = 0.05$) = 14.56

Table A.2: Power of the Test (i.i.d normal test)

Model	$N = 100$	$N = 200$	$N = 300$
Random-Walk(*) (a)	0.89	1	1
(b)	0.90	1	1
(c)	0.92	1	1
Harmonic (a)	0.16	0.20	0.18
(b)	0.25	0.33	0.38
(c)	0.10	0.14	0.11
RMP (a)	0.52	0.70	0.90
(b)	0.13	0.18	0.16
(c)	0.39	0.70	0.77
Locally-Stationary (a)	0.11	0.11	0.10
(b)	0.18	0.22	0.16
(c)	0.31	0.33	0.28
White-Noise (a)	0.38	0.59	0.67
(b)	0.10	0.12	0.08
(c)	0.16	0.25	0.23
Contaminated (a)	0.33	0.46	0.52
(b)	0.11	0.12	0.10
(c)	0.22	0.35	0.34

(*) The test is run over the original (non-transformed) series.

Bibliography

Robert B. Ash. *Probability and measure theory*. Harcourt/Academic Press, Burlington, MA, second edition, 2000. ISBN 0-12-065202-1. With contributions by Catherine Doléans-Dade.

Jushan Bai. Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Statist.*, 22(4):2051–2061, 1994. ISSN 0090-5364.

Jan Beran. *Statistics for long-memory processes*, volume 61 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York, 1994. ISBN 0-412-04901-5.

Leo Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. ISBN 0-89871-296-3. Corrected reprint of the 1968 original.

Fabio Busetti and Andrew Harvey. Tests of strict stationarity based on quantile indicators. *J. Time Series Anal.*, 31(6):435–450, 2010. ISSN 0143-9782.

John C. Cox, Jonathan E. Ingersoll, Jr., and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985. ISSN 0012-9682.

Ju. A. Davydov. The invariance principle for stationary processes. *Teor. Verojatnost. i Primenen.*, 15:498–509, 1970. ISSN 0040-361x.

J. L. Doob. *Stochastic processes*. John Wiley & Sons Inc., New York, 1953.

J. Galambos and E. Seneta. Regularly varying sequences. *Proc. Amer. Math. Soc.*, 41: 110–116, 1973. ISSN 0002-9939.

Paul R. Halmos. Invariant measures. *Ann. of Math. (2)*, 48:735–754, 1947. ISSN 0003-486X.

Paul R. Halmos. *Lectures on ergodic theory*. Publications of the Mathematical Society of Japan, no. 3. The Mathematical Society of Japan, 1956.

Melvin J. Hinich. A statistical theory of signal coherence. In *Advanced mathematical methods in measurement and instrumentation (Como, 1998)*, pages 7–23. Esculapio, Bologna, 2000.

Atsushi Inoue. Testing for distributional change in time series. *Econometric Theory*, 17 (1):156–187, 2001. ISSN 0266-4666.

Yuji Ito. Uniform integrability and the pointwise ergodic theorem. *Proc. Amer. Math. Soc.*, 16:222–227, 1965. ISSN 0002-9939.

George Kapetanios. Testing for strict stationarity in financial variables. *Journal of Banking and Finance*, 33:2346–2362, 2009.

John Lamperti. Semi-stable stochastic processes. *Trans. Amer. Math. Soc.*, 104:62–78, 1972. ISSN 0002-9947.

Sangyeol Lee and Seongryong Na. A nonparametric test for the change of the density function in strong mixing processes. *Statist. Probab. Lett.*, 66(1):25–34, 2004. ISSN 0167-7152.

E.L. Lehmann and J.P. Romano. *Testing Statistical Hypotheses*. Springer, third edition, 2005.

Emmanuel Lesigne and Dalibor Volný. Large deviations for generic stationary processes. *Colloq. Math.*, 84/85(part 1):75–82, 2000. ISSN 0010-1354.

Luiz Lima and Breno Neri. A test for strict stationarity. *Working Paper, 2008*.

Makiko Nisio. On polynomial approximation for strictly stationary processes. *J. Math. Soc. Japan*, 12:207–226, 1960. ISSN 0025-5645.

George L. O'Brien and Wim Vervaat. Marginal distributions of self-similar processes with stationary increments. *Z. Wahrsch. Verw. Gebiete*, 64(1):129–138, 1983. ISSN 0044-3719.

Karl Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989. ISBN 0-521-38997-6. Corrected reprint of the 1983 original.

P. C. B. Phillips. Regression theory for near-integrated time series. *Econometrica*, 56(5): 1021–1043, 1988. ISSN 0012-9682.

Mohsen Pourahmadi. *Foundations of time series analysis and prediction theory*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. Wiley-Interscience, New York, 2001. ISBN 0-471-39434-3.

M.B. Priestley and T. Subba Rao. A test for non stationarity of time series. *Journal of the Royal Statistical Society, Series B*, 31(1):140–149, 1969.

Sidney I. Resnick. *A probability path*. Birkhäuser Boston Inc., Boston, MA, 1999. ISBN 0-8176-4055-X.

Saïd E. Saïd and David A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984. ISSN 0006-3444.

G. William Schwert. Tests for unit roots: a Monte Carlo investigation. *J. Bus. Econom. Statist.*, 20(1):5–17, 2002. ISSN 0735-0015.

William F. Stout. *Almost sure convergence*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Probability and Mathematical Statistics, Vol. 24.

Murad S. Taqqu. Fractional Brownian motion and long-range dependence. In *Theory and applications of long-range dependence*, pages 5–38. Birkhäuser Boston, Boston, MA, 2003.

Lorenzo Trapani. A distribution-free test for changes in the distribution. *Working Paper, 2008*.

Wim Vervaat. Sample path properties of self-similar processes with stationary increments. *Ann. Probab.*, 13(1):1–27, 1985. ISSN 0091-1798.