The Pennsylvania State University

The Graduate School

Eberly College of Science

CENTRAL LIMIT THEOREMS FOR RANDOMLY MODULATED SEQUENCES OF RANDOM VECTORS WITH RESAMPLING AND APPLICATIONS TO STATISTICS

A Dissertation in
Statistics
by
Armine Bagyan

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2015

The thesis of Armine Bagyan was reviewed and approved* by the following:

Arkady Tempelman
Professor of Statistics
Dissertation Co-Adviser, Co-Chair of Committee

Bing Li
Professor of Statistics
Dissertation Co-Adviser, Co-Chair of Committee

Francesca Chiaromonte
Professor of Statistics and Public Health Sciences

Alexei Novikov
Professor of Mathematics

David Hunter
Professor of Statistics
Department Head

*Signatures are on file in the Graduate School.

Abstract

In many situations when sequences of random vectors are under consideration, it is of interest to study the asymptotic distribution of their (normalized) sums and to determine the conditions for the limit theorems, such as the Central Limit Theorem (CLT), to hold. In the simplest case when the variables are independent and identically distributed and have finite variance, the CLT is satisfied. Some CLT generalizations with weakened independence assumptions exist as well. For example, the CLT holds for stationary random sequences with strong mixing. However, in many situations when there is dependence, the CLT does not hold. This happens for stationary random sequences even with the weak mixing condition.

In our research we propose a method of random modulation of ergodic stationary random sequences that allows us to prove limit theorems for such sequences without any mixing conditions. These theorems present an opportunity to construct asymptotic confidence intervals for parameters, test parametric and non-parametric hypotheses with the significance level close to the required one and to calculate the approximate power of the test.

More general analogs of the CLT are proved and the speed of convergence is estimated for sequences of random vectors in spaces of non-decreasing dimensions.

Contents

List of Figures

List of Tables

Acknowledgments

I want to take this opportunity to thank all the people who have supported me during my doctoral studies and made finishing this dissertation possible.

I would like to express my deep gratitude to my advisers, Professors Arkady Tempelman and Bing Li, for their invaluable insights, continuous support and guidance, and great patience. It has been a privilege working with them.

My sincere appreciation goes to the members of my committee, Professors Francesca Chiaromonte and Alexei Novikov, for useful comments and great questions at my comprehensive exam and dissertation defence.

My greatest thanks go to my family and friends for listening to me and providing endless support and encouragements during this journey. I could not have done it without them.

Also I would like to express my gratefulness to the Department of Statistics for financially supporting me throughout my doctoral studies.

Chapter 1

Introduction

In this chapter we review the notions of stationary sequences and dynamical systems and discuss the relationship between them. We briefly discuss some of the existing CLT-type results for independent and stationary sequences, as well as for dynamical systems. Later in this chapter we also consider some relevant ergodic theorems that will be useful in later chapters.

1.1 Dynamical Systems

Let $(\Omega, \mathfrak{F}, m)$ be a *measure space*. Let x_t be the point in the *phase space* Ω that describes the state of a system at time t. Let τ_t denote the rule that transforms x_0 into x_t, that is $x_t = \tau_t x_0$; τ_0 is the identity transformation, that is $\tau_0 x = x$ for any $x \in \Omega$; for any s, t, and $x \in \Omega$ we have $\tau_{s+t}x = \tau_s \tau_t x$. Therefore $\{\tau_t\}$ forms a one-parameter group of transformations of the phase space, a so-called *flow*. The parameter t is usually treated as "time" and most commonly takes on real or integer values. Our attention is devoted to the case when time t runs over the group of integers $\mathbb{Z}$; in this case τ_t is invertible and $\tau_t^{-1} = \tau_{t-1}$, so all τ_t are invertible. In some situations it is of interest to let t take only positive or non-negative values, so then $\{\tau_n\}$ is a semi-group of transformations (in this case τ_t does not need to be invertible). The group nature of $\{\tau_n\}$ implies that $\tau_n x = \tau^n x$ for any integer $n \in \mathbb{Z}$ and $x \in \Omega$. We denote the identity transformation $\tau^0 = \tau_0$.

Transformation $\tau : \Omega \mapsto \Omega$ is called *measurable* if the inverse image of every measurable set is measurable, that is for any $\Lambda \in \mathfrak{F}$

$$\tau^{-1}\Lambda \in \mathfrak{F},$$

where

$$\tau^{-1}\Lambda = \{x \in \Omega : \quad \tau x \in \Lambda\}.$$

Transformation τ is called *measure-preserving*, or *m-invariant*, if the measure of any measurable set and its inverse image is the same, that is for any $\Lambda \in \mathfrak{F}$ we have

$$m(\tau^{-1}\Lambda) = m(\Lambda).$$

Let us note that we can ignore sets of measure zero and therefore concentrate our attention on the results that hold m-almost everywhere. Measure preserving transformations τ_t, $t \in \mathbb{R}$, arise in mechanics: the Liouville's Theorem claims that under certain conditions, a dynamical system may be considered as a flow that preserves the volume with respect to some generalized coordinate system [7]. Often, groups (or semi-groups) of measure-preserving transformations on a measure space are called *dynamical systems*, and the ergodic theory studies their asymptotic properties and some other properties of the trajectories.

Let us consider some simple examples of dynamical systems.

Example 1.1. [7] Let us demonstrate a common dynamical system. Suppose the phase space $\Omega = [0, 1)$ and transformation

$$\tau x = 2x \pmod 1 = \begin{cases} 2x, & \text{if } 0 \leq x < \frac{1}{2} \\ 2x - 1, & \text{if } \frac{1}{2} \leq x < 1 \end{cases}$$

This transformation τ is illustrated by Figure 1.1.

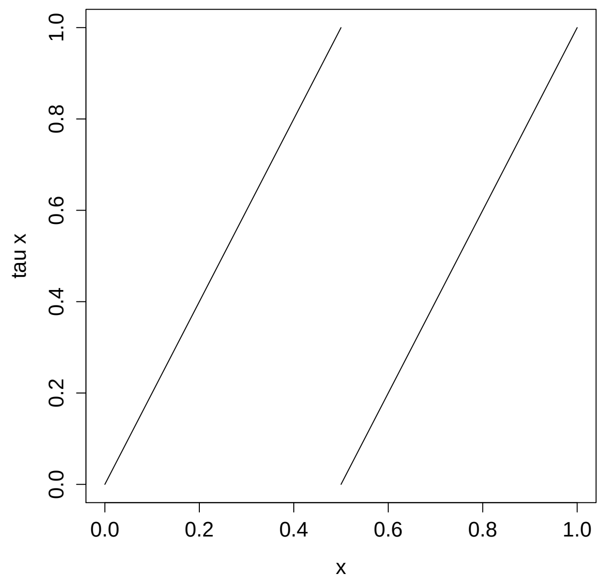

Figure 1.1: Plot of x against τx for Example 1.1.

If we consider the Borel σ-algebra and the Lebesgue measure on Ω, then for any Borel set $\Lambda \subset \Omega$ here we have $m(\tau^{-1}\Lambda) = m(\Lambda)$. Therefore τ is a measure-preserving transformation. However this transformation is not one-to-one, and so it is not invertible. Furthermore Ω can be represented as the circle on the complex plane with radius 1, that is as the set of all complex numbers with the absolute value 1. If we set the measure to equal to the lenght of the arc divided by 2π, then transformation τ corresponds to the following transformation on the unit complex circle

$$Tz = z^2, \quad |z| = 1,$$

the transformation that doubles the angle. $\square$

Example 1.2. [7] Now let us consider another transformation on this probability space

$$\tau x = x + c \pmod 1 = \begin{cases} x + c, & \text{if } 0 \leq x < 1 - c \\ x + c - 1, & \text{if } 1 - c \leq x < 1 \end{cases}$$

where $c \in \Omega$. The transformation τ is illustrated by Figure 1.2.

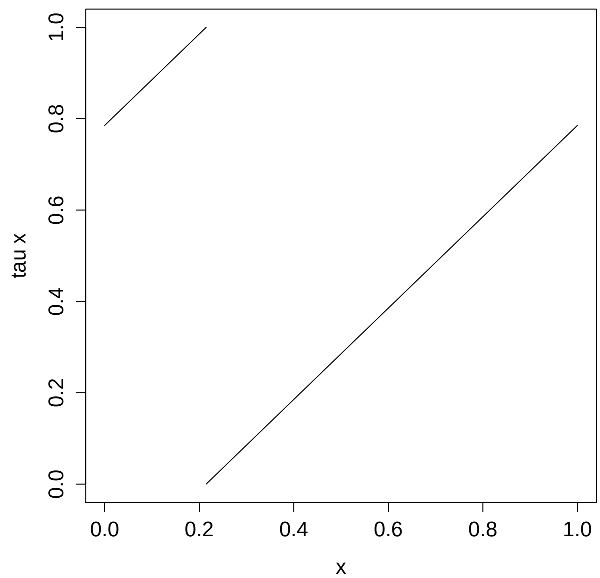

Figure 1.2: Plot of x against τx for Example 1.2 when $c = \pi/4$.

Transformation τ is invertible and measure-preserving. Geometrically it corresponds to the rotation by a fixed angle on the unit complex circle. $\square$

Example 1.3. [7] Suppose

$$\Omega = \{(x_1, x_2) : \ 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$$

is the unit square with the Lebesgue measure. Let us consider the following transformation

$$\tau(x_1, x_2) = \begin{cases} (2x_1, \frac{x_2}{2}), & \text{if } 0 \leq x_1 \leq \frac{1}{2} \\ (2x_1 - 1, \frac{x_2+1}{2}), & \text{if } \frac{1}{2} < x_1 \leq 1 \end{cases}$$

Transformation τ is often referred to as the "baker's transformation", because it looks as if a square piece of dough is rolled out and cut in half, and then the right half is stacked on top of the left one.

Transformation τ is measure-preserving. Also it can be made invertible by alterations on a set of measure 0. $\square$

Example 1.4. [7] Let us consider another common example. Let

$$\Omega = \{x = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots), \quad x_i \in \{0,1\}, \quad i = 0, \pm 1, \pm 2, \ldots\}$$

be the space of all two-sided binary sequences of 0's and 1's with the $\sigma-$algebra that is generated by sets of the form

$$\{x: \; x_i = 1\}$$

and the product measure of the probability distribution with

$$m(0) = m(1) = \frac{1}{2}.$$

Let us consider the so-called *shift* transformation τ such that

$$(\tau x)_i = x_{i+1}, \quad i = 0, \pm 1, \pm 2, \ldots.$$

Then τ is invertible and measure-preserving.

If we modify Ω to be the space of all one-sided binary sequences

$$\Omega = \{x = (x_1, x_2, \ldots), \quad x_i \in \{0,1\}, \quad i = 1, 2, \ldots\},$$

then the shift transformation is measure-preserving, but not invertible. There is a measure preserving correspondence between such Ω and the unit interval because these sequences can be considered as binary expansions of numbers on $[0,1)$. Moreover this establishes a correspondence almost everywhere between the shift transformation and transformation

$$\tau x = 2x \pmod 1$$

from Example 1.1

Later we will consider a generalization of this example when the values if x_i are not restricted to 0 and 1. $\square$

1.2 Stationary Sequences and Their Relationship with Dynamical Systems

Let $(\Omega, \mathfrak{F}, P)$ be a probability space. A random sequence $X_1, X_2, \ldots$ over this space is called *stationary* if the joint probability distribution is invariant over time, that is

$$(X_{i+1}, \ldots, X_{i+k}) \overset{\mathcal{D}}{\sim} (X_1, \ldots, X_k)$$

for any positive integers k and $i \in \mathbb{Z}_+$, where $\overset{\mathcal{D}}{\sim}$ means that the joint probability distributions of the two vectors are the same.

Similarly, m random sequences $X_1^{(j)}, X_2^{(j)}, \ldots$, $j = 1, \ldots, m$ are called *jointly stationary* if the joint distribution of vectors $(X^{(1)}, \ldots, X^{(m)})^T$ is invariant over time, that is

$$\Big((X_{i+1}^{(1)}, \ldots, X_{i+1}^{(m)})^T, \ldots, (X_{i+k}^{(1)}, \ldots, X_{i+k}^{(m)})^T\Big) \\ \overset{\mathcal{D}}{\sim} \Big((X_1^{(1)}, \ldots, X_1^{(m)})^T, \ldots, (X_k^{(1)}, \ldots, X_k^{(m)})^T\Big)$$

for any positive integers k and $i \in \mathbb{Z}_+$. In particular, two random processes $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are called *jointly stationary* if the joint distribution of pairs (X, Y) is invariant over time.

Suppose we have m sequences $X_1^{(j)}, X_2^{(j)}, \ldots$, $j = 1, \ldots, m$. Let us define

$$\mathbf{X}_i = (X_i^{(1)}, X_i^{(2)}, \ldots, X_i^{(m)}),$$

$i = 1, 2, \ldots$. Then equivalently the sequence $\mathbf{X}_1, \mathbf{X}_2, \ldots$ can be viewed as one sequence with values in $\mathbb{R}^m$. Furthermore, if sequences $X_1^{(j)}, X_2^{(j)}, \ldots$, $j = 1, \ldots, m$ are jointly stationary, then the m-dimensional sequence $\mathbf{X}_1, \mathbf{X}_2, \ldots$ is also stationary.

Let us briefly discuss the correspondence between dynamical systems and stationary sequences. Suppose τ is a measure preserving transformation on the probability space $(\Omega, \mathfrak{F}, P)$. Then $(\Omega, \mathfrak{F}, P)$ with transformation τ form a dynamical system. For a point

$w \in \Omega$ and a measurable function f on Ω we consider a random sequence by setting

$$X_i = f(\tau^{i-1}\omega), \quad i = 1, 2, \ldots, \tag{1.1}$$

then the sequence $X_1, X_2, \ldots$ is stationary.

It is true in the opposite direction as well. Let $X_1, X_2, \ldots$ be a stationary sequence on some probability space $(\Omega, \mathfrak{F}, P)$. Let

$$\tilde{\Omega} = \mathbb{R}^\infty = \{\tilde{\omega} = (x_1, x_2, \ldots), x_i \in \mathbb{R}\}$$

be the space of all sequences, $\tilde{\mathfrak{F}}$ be the σ-algebra generated by all n-dimensional cylinders

$$\{x_1 \leq a_1, \ldots, x_n \leq a_n\}$$

for any $a_1, \ldots, a_n \in \mathbb{R}$ and for any n. And let P_X be the measure on $(\tilde{\Omega}, \tilde{\mathfrak{F}})$ generated by the Kolmogorov Extension Theorem.

The *shift* transformation τ is defined in $\tilde{\Omega}$ as $\tau : \tilde{\Omega} \mapsto \tilde{\Omega}$ such that

$$\tau(x_1, x_2, \ldots) = (x_2, x_3, \ldots).$$

The shift transformation preserves the measure P_X, that is $P_X(\tau^{-1}\Lambda) = P_X(\Lambda)$ for any $\Lambda \in \tilde{\mathfrak{F}}$. Let us note that if Λ is an invariant set of measure 0, then so is $\tau\Lambda$.

We let

$$f(\tilde{\omega}) = x_1(\tilde{\omega}), \quad \tilde{\omega} \in \tilde{\Omega},$$

and for $i = 2, 3, \ldots$

$$f(\tau^{i-1}\tilde{\omega}) = x_i(\tilde{\omega}).$$

This way we obtain a random sequence that has the same distribution as the sequence $X_1, X_2, \ldots$ and corresponds to the dynamical system $\{\tau_n\}_{n \in \mathbb{Z}^+}$.

Remark 1.1. Let $X_1, X_2, \ldots$ be constructed as in (1.1) and let f be a one-to-one function, then the distribution of the random variable $X_1 = f(\omega)$ uniquely determines the measure

P on Ω, and hence it uniquely determines the measure P_X generated by the stationary sequence $X_1, X_2, \ldots$.

Let us introduce a few more notions. A measurable subset $\Lambda \subset \Omega$ is called *invariant* with respect to a transformation τ if

$$\tau^{-1}\Lambda \subset \Lambda.$$

Let $\mathfrak{I}$ denote the set of all subsets of Ω that are invariant with respect to transformation τ. Then $\mathfrak{I}$ is a σ-algebra and $\mathfrak{I} \subset \mathfrak{F}$.

Furthermore, if for any $\Lambda \in \mathfrak{I}$, $P(\Lambda) = 0$ or $P(\Lambda) = 1$, then the σ-algebra of invariant subsets $\mathfrak{I}$ is called *trivial*; in this case the dynamical system $\{\tau_n\}_{n \in \mathbb{Z}^+}$ is called *ergodic*, or *metrically transitive*.

For example, the transformation in Example 1.4 is ergodic. Since the angle doubling transformation in Example 1.1 is isomorphic to the one in Example 1.4, it is ergodic as well. The circle rotating transformation in Example 1.2 is ergodic if and only if c is an irrational number in $[0, 1)$. The baker's transformation in Example 1.3 is also ergodic.

We say that a random sequence $X_1, X_2, \ldots$ is *ergodic*, or *metrically transitive*, if the shift dynamical system in $\tilde{\Omega}$ is ergodic with respect to th measure P_X. If τ is an ergodic dynamical system on $(\tilde{\Omega}, \tilde{\mathfrak{F}}, P_X)$ and f is a measurable function on $\tilde{\Omega}$, then the stationary random sequence $X_i = f(\tau^{i-1}\tilde{\omega})$ is ergodic.

Remark 1.2. If τ is an ergodic measure-preserving transformation with respect to two measures m_1 and m_2, then either $m_1 = m_2$, or m_1 and m_2 are mutually singular (see for example [2]).

1.3 Review of the Central Limit Theorems for Independent Random Variables

The Central Limit Theorem (CLT) is one of the most fundamental and widely used theorems of the probability theory. The CLT states that under certain conditions the sum of random variables with a defined mean and variance is asymptotically normal. The role and importance of the CLT is explained by the fact that the asymptotic normality of the sums serve as the basis and justification for a wide range of statistical procedures. In particular, the CLT is used in testing hypothesis and computing the power of the tests, constructing approximate confidence intervals for large samples, regression, ANOVA, etc.

The classical version of the CLT considers the case when the variables are independent and identically distributed and can be stated as follows.

Theorem 1.1. *[5] Let $X_1, X_2, \ldots$ be a sequence of mutually independent identically distributed random variables with mean*

$$\mu = E(X_k)$$

and finite variance

$$\sigma^2 = Var(X_k) > 0,$$

$k = 1, 2, \ldots$. Let $F_n(z)$ denote the cumulative distribution function of $\frac{1}{\sigma\sqrt{n}}\left(\sum_{k=1}^{n} X_k - n\mu\right)$. Then as $n \to \infty$, $F_n(z)$ converges to the standard normal cumulative distribution function $\Phi(z)$ for any $z \in \mathbb{R}$.

Furthermore, the above theorem can be generalized for the case of independent random variables that do not necessarily have the same distribution.

Theorem 1.2. *(Lindeberg) [6] Let $X_1, X_2, \ldots$ be a sequence of mutually independent random variables with cumulative distribution functions $F_1(x), F_2(x), \ldots$, respectively. Let*

$$E(X_k) = 0,$$

$$Var(X_k) = \sigma_k^2,$$

and

$$s_n^2 = \sigma_1^2 + \ldots + \sigma_n^2.$$

Suppose that for any $t > 0$, as $n \to \infty$,

$$\frac{1}{s_n^2} \sum_{k=1}^{n} \int\limits_{|y| \geq t s_n} y^2 F_k\{dy\} \to 0.$$

Then the cumulative distribution function of $\frac{1}{s_n} \sum\limits_{k=1}^{n} X_k$ converges to the standard normal cumulative distribution function, that is for any $z \in \mathbb{R}$, as $n \to \infty$,

$$P\left(\frac{1}{s_n} \sum_{k=1}^{n} X_k \leq z\right) \to \Phi(z).$$

However, even though the CLT is proved to be useful and generally applicable in the case of independent random variables, often it does not hold in the case of dependent variables, even when the dependence is weak. We will illustrate this using Example 3.2.

1.4 The Central Limit Theorems for Strong Mixing Stationary Sequences

Another useful generalization of the CLT was made in case of strong mixing that describes the case of weak dependence under which the variables that are far away from each other in the sequence are approximately independent. However let us note that this assumption is pretty strict.

We say that the CLT holds for a stationary sequence $X_1, X_2, \ldots$, if its normalized sums $\frac{1}{s_n} \sum\limits_{i=1}^{n} X_i$, where $s_n = \sqrt{Var\left(\sum\limits_{i=1}^{n} X_i\right)}$, converge in distribution to the standard normal distribution, as $n \to \infty$.

Let $X_1, X_2, \ldots$ be a stationary sequence and $\mathfrak{F}_1^n$ and $\mathfrak{F}_{n+s}^{\infty}$ be the Borel $\sigma-$fields generated by $X_1, \ldots, X_n$ and $X_{n+s}, X_{n+s+1}, \ldots$, respectively. We denote

$$\alpha(s) = \sup_{A \in \mathfrak{F}_1^n, B \in \mathfrak{F}_{n+s}^{\infty}} |P(A \cap B) - P(A)P(B)|.$$

The function $\alpha(s)$ is called the *strong mixing coefficient* . We say that the stationary process

$X_1, X_2 \dots$ is *strong mixing* if

$$\alpha(s) \to 0, \quad \texttt{as} \quad s \to \infty.$$

In essence, the strong mixing condition describes a form of very weak dependence, or we can say asymptotic independence [12]. This condition implies ergodicity.

Remark 1.3. Let us note that a (stationary) sequence of independent identically distributed random variables is strong mixing [12].

In [13] Rozanov showed that the CLT-type theorem holds for strong mixing stationary sequences. But in Example 3.2 we will see that the CLT fails for some ergodic stationary sequences.

1.5 On the Central Limit Theorems for Dynamical Systems and Stationary Sequences without Strong Mixing

Let us also consider some results that have been achieved for dynamical systems. Developing the limit theory for dynamical systems also turned out to be quite challenging. For dynamical systems even mixing does not guarantee that the CLT holds. However, in some cases, for example when the entropy is positive, CLT still holds (see [3]). We say that CLT holds for a dynamical system if for any function $f \in L^2(\Omega)$ CLT holds for the sequence $f(\tau^{i-1}\omega)$, $\omega \in \Omega$, $i = 1, 2, \dots$. Without going into detail, let us point out that for K-systems, also called Kolmogorov systems, CLT works (see for example [2]). (Let us note that any K-system is mixing and has positive entropy.) For example, the baker's transformation is a K-dynamical system, and therefore the CLT holds.

If the entropy of the dynamical system is zero, the CLT may fail. In many cases, $E\left(\sum_{k=0}^{n-1} f \circ \tau^k\right)^2 \to 0$, as $n \to \infty$. For instance, if we consider the ergodic rotation for a simple function f, for example the identity function, then the CLT may not hold (see Example 3.2). In [3], [1], [8] the following question is discussed: when does a function f such that $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ \tau^k$ converges weakly to Normal distribution, as $n \to \infty$, exist? In particular, in [1] it is shown that for an aperiodic dynamical system, there exists a function $f \in L^2$ such that the sequence $f, f \circ \tau, f \circ \tau^2, \ldots$ satisfies the CLT. If τ is ergodic, then the set of such functions f is dense in L^2. However, it is still an unsolved problem to describe the set of functions in L^2 that ensure that the CLT is satisfied for the sequence $f, f \circ \tau, f \circ \tau^2, \ldots$. Although these results present a great interest, in statistics we need the CLT to hold for *any* real function f, because we want to have it for *any* stationary sequence associated with the dynamical system - not only for some of them.

1.6 Motivation

It is well known that the CLT plays a big role in statistics as CLT-type results are necessary for many procedures of statistical inference. However, we also saw that the CLT may not hold for ergodic stationary random sequences and dynamical systems without strong mixing. Therefore we propose a new type of CLT that holds for any ergodic stationary sequence.

There are two ways to approach this problem: projection (geometric approach) and random modulation. These two approaches are equivalent in terms of asymptotic distribution as will be discussed in Chapter 2. Although the first approach is interesting from the geometric point of view, the second one is more of interest for statistics and provides more opportunities for investigation of asymptotic behavior of random vectors and sequences. In this work we will concentrate on the random modulation.

We review some of the ergodic theorems that we use in this work.

1.7 The Birkhoff-Khinchin Ergodic Theorem for Dynamical Systems

Roughly speaking, the ergodic theory of dynamical systems studies the asymptotic behavior of certain classes of measure-preserving transformations on a measure space. It has applications in many branches of mathematics, especially in probability theory and functional analysis. The ergodic theory is very useful in studying the topic of stationary ergodic random processes due to the fact that there is certain correspondence between such processes and dynamical systems as we discussed above.

Remark 1.4. Let us note that if we have a sequence $X_1, X_2, \ldots$ of independent and identically distributed random variables with

$$\mu = E(X_1) < \infty,$$

then by the Strong Law of Large Numbers (SLLN),

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X_k = \mu$$

almost surely. The ergodic theorems are generalization of the SLLN. They allow us to say that under certain conditions on the sequence $X_1, X_2, \ldots$, the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k)$$

exists almost surely. And under some additional conditions the limit is equal to $E(f(X))$.

Let $(\Omega, \mathfrak{F}, P)$ be a probability space. And let $\mathfrak{I}$ be the σ-algebra of all subsets of Ω that are invariant with respect to a transformation τ. Now let us state the Birkhoff-Khinchin Ergodic Theorem (see for example [16]).

Theorem 1.3. *(Birkhoff-Khinchin Ergodic Theorem) Let τ be a measurable measure-preserving transformation of the probability space $(\Omega, \mathfrak{F}, P)$ and f be a measurable function, $E\,|f| < \infty$. Then with probability* 1

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k x) = E(f|\mathfrak{I}),$$

where $E(f|\mathfrak{I})$ is the conditional expectation of f given the σ-algebra $\mathfrak{I}$.

Furthermore if τ is ergodic, that is if the σ-algebra $\mathfrak{I}$ is trivial, we have the following corollary.

Corollary 1.1. *If in addition τ is ergodic, that is $\mathfrak{I}$ is trivial, then with probability 1*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)) = E(f).$$

1.8 The Ergodic Theorem for Stationary Sequences

Above we discussed the relationship between stationary sequences and dynamical systems. Therefore using Theorem 1.3 and Corollary 1.1 we have the following ergodic theorem for stationary sequences.

Corollary 1.2. *Let $X_1, X_2, \ldots$ be a stationary random sequence and f be such that $E[f(X_1)] < \infty$. Then with probability 1*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} f(X_i) = E[f(X_1)|\mathfrak{I}].$$

Moreover, if $X_1, X_2, \ldots$ is ergodic, then with probability 1

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} f(X_i) = E[f(X_1)].$$

Chapter 2

Limit Theorems for Random Vectors in Spaces with Increasing Dimensions

We begin this chapter by considering sequences of multidimensional nonrandom vectors. In [4] P. Diaconis and D. Freedman propose mathematical tools to study such sequences by describing the distribution of their projections onto random vectors. In the next section we introduce the concept of random modulation for a random vector $\mathbf{X}$ using a modulating random vector ξ, and present the important probabilistic analog developed by B. Li [9] of the deterministic approach described earlier. This allows us to find the asymptotic conditional distribution of randomly modulated versions of the random vector given ξ. This follows by the discussion of sufficient conditions and examples to illustrate the results. In the next sections we prove the $L^2_\xi-$convergence of the conditional probability density functions and the uniform $L^2_\xi-$convergence of the conditional cumulative distribution functions of randomly modulated versions of random vectors. In the last section of this chapter we prove the convergence of the distribution of randomly modulated vectors with and without resampling to normal in distribution with respect to the product measure $P_X \times P_\xi$. This important modification of the results mentioned above developed by A. Tempelman [17] allows us to propose applications of our results to statistics in Chapter 4.

2.1 Results by Diaconis and Freedman and Weak Convergence in Probability

Suppose the data set consists of (nonrandom) vectors $x_1, x_2, \ldots, x_n$ in $\mathbb{R}^p$. For mathematical convenience, P. Diaconis and D. Freedman [4] assume that n, p and x_i depend on a hidden index ν, and that n and p go to infinity, as $\nu \to \infty$. Suppose that for $0 < \sigma^2 < \infty$, and for any $\varepsilon > 0$, as $\nu \to \infty$ the following two conditions are satisfied:

$$\frac{1}{n}\text{card}\left\{j \leq n : \mid \|x_j\|^2 - \sigma^2 p \mid > \varepsilon p\right\} \to 0, \tag{2.1}$$

$$\frac{1}{n^2}\text{card}\left\{1 \leq j, k \leq n : \mid x_j^T x_k \mid > \varepsilon p\right\} \to 0. \tag{2.2}$$

The authors remark that the first condition suggests that the length of most vectors is approximately $\sigma^2 p$, and the second condition says that most vectors are close to orthogonal. It is important to point out that only p vectors (in $\mathbb{R}^p$) may be exactly orthogonal. These two conditions are satisfied if, for example, the x_j's are the observed values of independent identically distributed vectors with independent identically distributed coordinates.

To study projections, let γ be a random vector with uniform distribution on S_{p-1}, where S_{p-1} is the unit sphere in $\mathbb{R}^p$. Consider the dataset projected onto the direction γ

$$x_1^T\gamma, x_2^T\gamma, \ldots, x_n^T\gamma.$$

Let $\theta_\nu(\gamma)$ denote the empirical distribution of the projected dataset with the weight of $\frac{1}{n}$ assigned to each $x_j^T\gamma$. Then the main result of the paper can be stated as the following theorem.

Theorem 2.1. *[4] Under the two conditions (2.1) and (2.2), as $\nu \to \infty$, the empirical distribution θ_ν converges to $N(0, \sigma^2)$ weakly in probability.*

In other words, the theorem says that, as $\nu \to \infty$, θ_ν converges to $N(0, \sigma^2)$ in terms of weak topology for most directions γ, where "most" means in probability with respect to uniform distribution on S_{p-1}.

Furthermore, a similar property can be stated for standardized data. The conditions then become

$$\frac{1}{np}\sum_{j=1}^{n}\|x_j\|^2 \to \sigma^2, \tag{2.3}$$

$$\frac{1}{n}\text{card}\left\{j \le n : \ | \ \|x_j\|^2 - \sigma^2 p \ | > \varepsilon p\right\} \to 0, \tag{2.4}$$

$$\frac{1}{(np)^2}\sum_{j,k=1}^{n}(x_j^T x_k)^2 \to 0, \tag{2.5}$$

as $\nu \to \infty$. The conditions (2.3), (2.4), and (2.5) imply the previous two conditions (2.1) and (2.2) by Chebyshev's inequality.

Let $\tilde{\theta}_\nu(\gamma)$ be the standardized empirical measure, with weight $\frac{1}{n}$ for each element of the standardized projected dataset. Then the following theorem holds.

Theorem 2.2. *[4] Under conditions (2.3), (2.4) and (2.5), as $\nu \to \infty$, the standardized empirical distribution $\tilde{\theta}_\nu(\gamma)$ converges to $N(0,1)$ weakly in probability.*

Furthemore, let us consider $\xi \sim N(0, I_p)$, that is a vector of p independent standard normal variables. Then $\frac{\xi}{\|\xi\|}$ has uniform distribution on the unit sphere S_{p-1}. In addition, we observe that $\frac{\|\xi\|}{\sqrt{p}} \to 1$ almost surely, as $p \to \infty$. Therefore, as Diaconis and Freedman remark, it is enough to prove the theorems with $\frac{\xi}{\sqrt{p}}$ instead of γ, and the advantage is that the normal theory can be employed. We take this suggestion and use $\xi \sim N(0, I_p)$.

If $\frac{\xi}{\sqrt{p}}$ is used instead of γ, then the first theorem can be restated as follows.

Theorem 2.3. *[4] Under conditions (2.1) and (2.2), θ_ν converges to $N(0,\sigma^2)$ weakly in probability, where $\xi \sim N(0, I_p)$ and $\theta_\nu(\xi)$ is the empirical distribution of*

$$\frac{\xi^T x_1}{\sqrt{p}}, \frac{\xi^T x_2}{\sqrt{p}}, \ldots, \frac{\xi^T x_n}{\sqrt{p}}.$$

2.2 Randomly Modulated Random Vectors and Their Asymptotic Conditional Distribution

Let $(\Omega_X, \mathfrak{F}_X, P_X)$ and $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$ be probability spaces. B. Li in [9] proposed the probabilistic analog of the ideas of [4] presented in the previous section in the following way: instead of a sequence of nonrandom vectors $x_1, x_2, \ldots, x_n$ he considers a sequence of d_n-dimensional random vectors $\mathbf{X}^{(n)}$ and studies the projections of the random vectors $\mathbf{X}^{(n)}$ onto a random vector ξ from the $N(0, I_{d_n})$ distribution. In what follows we state and prove B. Li's theorem. We also consider some modifications and refinements of this result. Let us note that the ideas similar to [9] were used by B. Li, S. Wen, and L. Zhu to develop the projective resampling method in [10].

We are interested in studying the asymptotic behavior of the projection $\gamma^{(n)T}\mathbf{X}^{(n)}$ of a random vector $\mathbf{X}^{(n)}$ onto a random vector $\gamma^{(n)}$ that has uniform distribution on a d_n-dimensional unit sphere. On the other hand, we may consider a *randomly modulated* version of the vector $\mathbf{X}^{(n)}$, namely for a *modulating* vector $\xi^{(n)} \sim N(0, I_{d_n})$ we construct

$$\frac{1}{\sqrt{d_n}}\xi^{(n)T}\mathbf{X}^{(n)} = \frac{1}{\sqrt{d_n}}\sum_{i=1}^{d_n} \xi_i^{(n)}\mathbf{X}_i^{(n)}.$$

Let us note that $\frac{\xi^{(n)}}{||\xi^{(n)}||}$ has the uniform distribution on the d_n-dimensional unit sphere. Furthermore, by the LLN, $\frac{||\xi^{(n)}||}{\sqrt{d_n}} \to 1$ *a.e.*, as $n \to \infty$. Then the asymptotic behavior of $\gamma^{(n)T}\mathbf{X}^{(n)}$ and $\frac{1}{\sqrt{d_n}}\xi^{(n)T}\mathbf{X}^{(n)}$ is the same. Therefore to study the asymptotic distribution, it is equivalent to consider the projection of the random vector $\mathbf{X}^{(n)}$ onto a random vector γ and the random modulation of the vector $\mathbf{X}^{(n)}$ using $\xi^{(n)}$.

Theorem 2.4. *[9] Let $\mathbf{X}^{(n)} \in \mathbb{R}^{d_n}$, $n = 1, 2, \ldots$, be a sequence of d_n-dimensional random vectors on $(\Omega_X, \mathfrak{F}_X, P_X)$, $\xi^{(n)} \sim N(0, I_{d_n})$, $n = 1, 2, \ldots$, be a sequence of random vectors on $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$, and $\mathbf{X}^{(n)} \perp\!\!\!\perp \xi^{(n)}$. Let $\mu_n(\xi^{(n)})$ denote the conditional probability measure of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}} \,|\xi^{(n)}$. Suppose*

1. *As $n \to \infty$,*

$$\frac{||\mathbf{X}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} \sigma^2. \tag{2.6}$$

2. Let $\tilde{\mathbf{X}}^{(n)}$ be an independent copy of $\mathbf{X}^{(n)}$. As $n \to \infty$,

$$\frac{\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)}}{d_n} \xrightarrow{\mathcal{P}} 0. \tag{2.7}$$

Then as $n \to \infty$, $\mu_n(\xi^{(n)})$ converges to $N(0, \sigma^2)$ weakly in probability.

Proof. Consider the characteristic function of $\mu_n(\xi^{(n)})$

$$\varphi_n(t|\xi) = E_X\left(e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}}|\xi^{(n)}\right),$$

where $E_X(\cdot)$ is used to denote the expectation with respect to the probability measure P_X. Further we will also use $E_\xi(\cdot)$ and $E_{\xi,X}(\cdot)$ to denote the expectations with respect to the probability measures P_ξ and $P_\xi \times P_X$, respectively. Then

$$\begin{aligned} E_\xi\left[\varphi_n(t|\xi)\right] &= E_{\xi,X}\left(e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}}\right) \\ &= E_X\left[E_\xi\left(e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}}|\mathbf{X}^{(n)}\right)\right]. \end{aligned}$$

Since $\xi^{(n)} \sim N(0, I_{d_n})$, we have

$$E_\xi\left[\varphi_n(t|\xi)\right] = E_X\left(e^{-||\mathbf{X}^{(n)}||^2t^2/(2d_n)}\right). \tag{2.8}$$

From the first condition $\frac{||\mathbf{X}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} \sigma^2$, as $n \to \infty$, and continuity of the exponential function we have

$$e^{-||\mathbf{X}^{(n)}||^2t^2/(2d_n)} \xrightarrow{\mathcal{P}} e^{-\sigma^2t^2/2}.$$

Since $e^{-||\mathbf{X}^{(n)}||^2t^2/(2d_n)}$ is bounded by 1, by the Lebesgue's Dominated Convergence Theorem we have

$$E_X\left(e^{-||\mathbf{X}^{(n)}||^2t^2/(2d_n)}\right) \to e^{-\sigma^2t^2/2},$$

as $n \to \infty$, that is

$$E_\xi\left[\varphi_n(t|\xi)\right] \longrightarrow e^{-\sigma^2t^2/2} = \varphi_{N(0,\sigma^2)}(t).$$

Next we consider

$$\begin{aligned}
E_\xi \left|\varphi_n(t|\xi)\right|^2 &= E_\xi \left[\varphi_n(t|\xi)\bar{\varphi}_n(t|\xi)\right] \\
&= E_\xi \left[E_X \left(e^{it(\xi^{(n)T}\mathbf{X}^{(n)})/\sqrt{d_n}}|\xi^{(n)}\right) E_{\tilde{X}} \left(e^{it(\xi^{(n)T}\tilde{\mathbf{X}}^{(n)})/\sqrt{d_n}}|\xi^{(n)}\right)\right] \\
&= E_\xi \left[E_{X,\tilde{X}} \left(e^{it\xi^{(n)T}(\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)})/\sqrt{d_n}}|\xi^{(n)}\right)\right] \\
&= E_{X,\tilde{X}} \left[E_\xi \left(e^{it\xi^{(n)T}(\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)})/\sqrt{d_n}}|\mathbf{X}^{(n)}, \tilde{\mathbf{X}}^{(n)}\right)\right].
\end{aligned}$$

And since $\xi^{(n)} \sim N(0, I_{d_n})$, we have

$$E_\xi \left|\varphi_n(t|\xi)\right|^2 = E_{X,\tilde{X}} \left(e^{-||\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}||^2 t^2/(2d_n)}\right).$$

We know that $\frac{||\tilde{\mathbf{X}}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} \sigma^2$ and $\frac{\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)}}{d_n} \xrightarrow{\mathcal{P}} 0$, as $n \to \infty$, where $\tilde{\mathbf{X}}^{(n)}$ is an independent copy of $\mathbf{X}^{(n)}$. Then by Slutsky's Theorem,

$$\frac{||\mathbf{X}^{(n)} - \tilde{\mathbf{X}}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} 2\sigma^2,$$

as $n \to \infty$, and furthermore from continuity of the exponential function we have

$$e^{-||\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}||^2 t^2/(2d_n)} \xrightarrow{\mathcal{P}} e^{-\sigma^2 t^2}.$$

By the Lebesgue's Dominated Convergence Theorem,

$$E_{X,\tilde{X}} \left(e^{-||\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}||^2 t^2/(2d_n)}\right) \to e^{-\sigma^2 t^2},$$

as $n \to \infty$. That is,

$$E_\xi \left|\varphi_n(t|\xi)\right|^2 \longrightarrow e^{-\sigma^2 t^2} = \varphi^2_{N(0,\sigma^2)}(t).$$

Now by Chebyshev's Inequality for $\varepsilon > 0$

$$\begin{aligned}
P\left(\left|\varphi_n(t|\xi) - e^{-t^2\sigma^2/2}\right| > \varepsilon\right) &\leq \frac{1}{\varepsilon^2} E_\xi \left|\varphi_n(t|\xi) - e^{-t^2\sigma^2/2}\right|^2 \\
&= \frac{1}{\varepsilon^2} E_\xi \left[(\varphi_n(t|\xi) - e^{-t^2\sigma^2/2})(\bar{\varphi}_n(t|\xi) - e^{-t^2\sigma^2/2})\right] \\
&= \frac{1}{\varepsilon^2} E_\xi \left[\varphi_n^2(t|\xi) - \varphi_n(t|\xi)e^{-t^2\sigma^2/2} - \bar{\varphi}_n(t|\xi)e^{-t^2\sigma^2/2} + e^{-t^2\sigma^2}\right].
\end{aligned}$$

Since $e^{-t^2\sigma^2/2}$ is bounded by 1, we have

$$\begin{aligned}
P\left(\left|\varphi_n(t|\xi) - e^{-t^2\sigma^2/2}\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} \Big[&\left|E_\xi\left(\varphi_n^2(t|\xi)\right) - e^{-t^2\sigma^2}\right| \\
&+ \left|E_\xi\left(\varphi_n(t|\xi)\right) - e^{-t^2\sigma^2/2}\right| \\
&+ \left|E_\xi\left(\bar{\varphi}_n(t|\xi)\right) - e^{-t^2\sigma^2/2}\right|\Big].
\end{aligned}$$

The last expression converges to 0, as $n \to \infty$, which means that $\varphi_n(t|\xi)$ converges to $\varphi_{N(0,\sigma^2)}(t)$ in probability.

Finally, our statement follows from the following Lemma. $\square$

Lemma 2.1. *[4] Consider a sequence of random measures $\mu_1, \mu_2, \ldots$, and a deterministic measure μ_0. Let $\varphi_i(t)$ be the random characteristic function of μ_i, $i = 1, 2, \ldots$, and $\varphi_0(t)$ be the deterministic characteristic function of μ_0. Then μ_n converges to μ_0 weakly in probability if and only if the random characteristic functions $\varphi_i(t)$ converge to $\varphi_0(t)$ in probability for each t.*

2.3 Some Sufficient Conditions and Examples

Let us note that it may turn out difficult to verify if $\mathbf{X}^{(n)} \in \mathbb{R}^{d_n}$ satisfies conditions (2.6) and (2.7) because they include convergence in probability. Instead of these two conditions we will consider the set of following three conditions:

$$E_X\left(\frac{||\mathbf{X}^{(n)}||^2}{d_n}\right) \to \sigma^2, \tag{2.9}$$

$$E_X\left(\frac{||\mathbf{X}^{(n)}||}{\sqrt{d_n}}\right)\to\sigma, \tag{2.10}$$

$$E_X\left(\frac{\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)}}{d_n}\right)\to 0, \tag{2.11}$$

as $n\to\infty$.

Let us note that conditions (2.9) and (2.10) imply

$$E_X\left[\frac{||\mathbf{X}^{(n)}||}{\sqrt{d_n}}-\sigma\right]^2\to 0,$$

as $n\to\infty$, and therefore the set of conditions (2.9), (2.10), and (2.11) implies conditions (2.6) and (2.7). Thus, conditions (2.9), (2.10), and (2.11) provide the result of Theorem 2.4. Let us point out that in many cases checking these expectation-based conditions is easier. We also note that condition (2.11) (and therefore condition (2.7)) is satisfied whenever $E_X\mathbf{X}^{(n)}=0$. In this case, since $\tilde{\mathbf{X}}^{(n)}$ is an independent copy of $\mathbf{X}^{(n)}$,

$$\begin{aligned}E_{X,\tilde{X}}\left(\frac{\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)}}{d_n}\right)&=\frac{1}{d_n}E_{X,\tilde{X}}\left(\sum_{i=1}^{d_n}X_i^{(n)}\tilde{X}_i^{(n)}\right)\\&=\frac{1}{d_n}\sum_{i=1}^{d_n}E_X(X_i^{(n)})E_{\tilde{X}}(\tilde{X}_i^{(n)})\ =0.\end{aligned}$$

For illustration we consider three examples, namely when $\mathbf{X}^{(n)}$ is uniformly distributed on a sphere, a ball, and a cube, respectively.

Example 2.1. Let $\mathbf{X}^{(n)}\in\mathbb{R}^{d_n}$ have uniform distribution on a sphere of radius r_n and centered at 0, and let us assume that $\frac{r_n}{\sqrt{d_n}}\to\sigma$, as $n\to\infty$. Since $E(\mathbf{X}^{(n)})=0$, condition (2.11) is satisfied. Here $||\mathbf{X}^{(n)}||=r_n$, so we have

$$E_X\left(\frac{||\mathbf{X}^{(n)}||}{\sqrt{d_n}}\right)=\frac{r_n}{\sqrt{d_n}}\to\sigma,$$

and

$$E_X\left(\frac{||\mathbf{X}^{(n)}||^2}{d_n}\right) = \frac{r_n^2}{d_n} \to \sigma^2,$$

as $n \to \infty$. Then conditions (2.9) and (2.10) are also satisfied. Therefore, if $\xi^{(n)} \sim N(0, I_{d_n})$, the conditional distribution of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}}\ |\xi^{(n)}$ converges to $N(0, \sigma^2)$ weakly in probability with respect to P_ξ, as $n \to \infty$. □

Example 2.2. Let $\mathbf{X}^{(n)} \in \mathbb{R}^{d_n}$ have uniform distribution on a ball of radius r_n and centered at 0. Assume that $\frac{r_n}{\sqrt{d_n}} \to \sigma$, as $n \to \infty$. Since $E(\mathbf{X}^{(n)}) = 0$, condition (2.11) is satisfied. Denote the surface of the sphere of radius 1 as s_n. Then, as $n \to \infty$,

$$\begin{aligned} E_X\left(\frac{||\mathbf{X}^{(n)}||}{\sqrt{d_n}}\right) &= \frac{1}{\sqrt{d_n}} \frac{d_n}{s_n r_n^{d_n}} \int_0^{r_n} s_n \rho \rho^{d_n - 1} d\rho \\ &= \frac{1}{\sqrt{d_n}} \frac{d_n r_n}{d_n + 1} \\ &= \frac{d_n}{d_n + 1} \frac{r_n}{\sqrt{d_n}} \longrightarrow \sigma, \end{aligned}$$

and

$$\begin{aligned} E_X\left(\frac{||\mathbf{X}^{(n)}||^2}{d_n}\right) &= \frac{1}{d_n} \frac{d_n}{s_n r_n^{d_n}} \int_0^{r_n} s_n \rho^2 \rho^{d_n - 1} d\rho \\ &= \frac{1}{d_n} \frac{d_n r_n^2}{d_n + 2} \\ &= \frac{r_n^2}{d_n + 2} \longrightarrow \sigma^2. \end{aligned}$$

So for $\mathbf{X}^{(n)}$ all three conditions (2.9), (2.10), and (2.11) are satisfied. Therefore if $\xi^{(n)} \sim N(0, I_{d_n})$, the conditional distribution of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}}\ |\xi^{(n)}$ converges to $N(0, \sigma^2)$ weakly in probability with respect to P_ξ, as $n \to \infty$. □

Example 2.3. Let each $\mathbf{X}^{(n)} \in \mathbb{R}^{d_n}$ have uniform distribution on a cube with the side length a_n and centered at 0. Since $E(\mathbf{X}^{(n)}) = 0$, condition (2.11) is satisfied. Now let us consider under what assumptions on a_n the other two conditions are met.

$$E\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^2\right) = \frac{1}{d_n}\sum_{i=1}^{d_n} E\left(X_i^2\right) = \frac{a_n^2}{12}.$$

We want this to converge to some σ^2 as $n \to \infty$. Therefore if we assume that a_n converges to some constant a, as $n \to \infty$, we put $\sigma^2 = \frac{a^2}{12}$.

Condition (2.10) is very inconvenient for computation in this case, therefore we will consider the following condition instead

$$E_X\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^2 - \sigma^2\right)^2 \to 0, \tag{2.12}$$

as $n \to \infty$. Note that (2.9) and (2.12) imply (2.6) and (2.7). Consider

$$\begin{aligned}
E_X\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^2 - \sigma^2\right)^2 &= E\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^2\right)^2 - 2\sigma^2 E\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^2\right) + \sigma^4 \\
&= \sum_{i=1}^{d_n}\sum_{j=1}^{d_n} \frac{E(X_i^2 X_j^2)}{d_n^2} - 2\sigma^2 E\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^2\right) + \sigma^4 \\
&= \frac{1}{d_n^2}\sum_{i=1}^{d_n} E(X_i^4) + \frac{1}{d_n^2}\sum_{i\neq j} E(X_i^2 X_j^2) - 2\sigma^2\frac{1}{d_n}\sum_{i=1}^{d_n} E(X_i^2) + \sigma^4.
\end{aligned}$$

Since here $X_1, \ldots, X_n$ are independent and have Unif$[-\frac{a_n}{2}, \frac{a_n}{2}]$ distribution, we have

$$E(X_i^2) = \frac{a_n^2}{12}$$

and

$$E(X_i^4) = \frac{a_n^4}{80}.$$

Therefore

$$E_X\left(\frac{1}{d_n}\sum_{i=1}^{d_n}X_i^2-\sigma^2\right)^2 = \frac{1}{d_n^2}\sum_{i=1}^{d_n}\frac{a_n^4}{80}+\frac{1}{d_n^2}\sum_{i\neq j}\frac{a_n^4}{12^2}-2\sigma^2\frac{1}{d_n}\sum_{i=1}^{d_n}\frac{a_n^2}{12}+\sigma^4$$
$$= a_n^4\left(\frac{1}{80d_n}+\frac{d_n-1}{12^2d_n}\right)-2\sigma^2\frac{a_n^2}{12}+\sigma^4.$$

Earlier we assumed that a_n converges to some constant a, as $n\to\infty$, then we have

$$E_X\left(\frac{1}{d_n}\sum_{i=1}^{d_n}X_i^2-\sigma^2\right)^2 \longrightarrow \left(\frac{a^2}{12}-\sigma^2\right)^2,$$

and therefore in order to have (2.12) we need to take $\sigma^2=\frac{a^2}{12}$, as above. □

2.4 L_ξ^2-convergence of Conditional PDF

Let $\mathbf{X}^{(n)}\in\mathbb{R}^{d_n}$ be a d_n-dimensional random vector on $(\Omega_X,\mathfrak{F}_X,P_X)$, $\xi^{(n)}\sim N(0,I_{d_n})$ be a vector on $(\Omega_\xi,\mathfrak{F}_\xi,P_\xi)$, and $\mathbf{X}^{(n)}\perp\!\!\!\perp\xi^{(n)}$. Let $p_n(y|\xi)$ be the probability density function of the conditional distribution of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}}\,|\xi^{(n)}$.

Theorem 2.5. *If conditions (2.6) and (2.7) are met, then $p_n(y|\xi)$ converges to $p_{N(0,\sigma^2)}(y)$ L_ξ^2, as $n\to\infty$. That is, as $n\to\infty$,*

$$E_\xi\left|p_n(y|\xi)-p_{N(0,\sigma^2)}(y)\right|^2\to 0.$$

Proof. The characteristic function of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}}$ with respect to $P_X\times P_\xi$ is

$$\varphi(t)=E_\xi(\varphi_n(t|\xi)).$$

From (2.8) we have

$$\varphi(t)=E_X\left(e^{-||\mathbf{X}^{(n)}||^2t^2/(2d_n)}\right).$$

Then

$$\int |\varphi(t)|dt = \int \left| E_X(e^{-||\mathbf{X}^{(n)}||^2 t^2/(2d_n)}) \right| dt$$
$$\leq \int E_X \left| e^{-||\mathbf{X}^{(n)}||^2 t^2/(2d_n)} \right| dt \; < \infty,$$

and therefore the joint probability density function $p_{X,\xi}(x,\nu)$ exists. Furthermore, since $\xi^{(n)} \sim N(0, I_{d_n})$, we have $p_\xi(\nu) > 0$ everywhere. Hence the probability density function $p_n(y|\xi)$ of the conditional distribution of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}} \; |\xi^{(n)}$ exists.

Then the conditional density function is

$$p_n(y|\xi) = \mathcal{F}\left(\varphi_n(t|\xi)\right)(y),$$

where $\mathcal{F}$ is the Fourier transformation. Consider

$$E_\xi[p_n(y|\xi)] = E_\xi\left[\mathcal{F}\left(\varphi_n(t|\xi)\right)(y)\right].$$

Using (2.8), we have

$$E_\xi[p_n(y|\xi)] = E_\xi\left[\mathcal{F}\left(E_X\left[e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}}|\xi^{(n)}\right]\right)(y)\right].$$

Interchanging E_ξ and $\mathcal{F}$, we obtain

$$E_\xi[p_n(y|\xi)] = \mathcal{F}\left[E_\xi E_X\left(e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}}\right)\right](y).$$

Interchanging E_ξ and E_X, we have

$$E_\xi[p_n(y|\xi)] = \mathcal{F}\left[E_X E_\xi\left(e^{it\sum_{j=1}^{d_n}\xi_j^{(n)}X_j^{(n)}/\sqrt{d_n}}\right)\right](y).$$

Now using $\xi^{(n)} \sim N(0, I_{d_n})$, we have

$$E_\xi[p_n(y|\xi)] = \mathcal{F}\left[E_X\left(e^{-\frac{\|\mathbf{X}^{(n)}\|^2}{2d_n}t^2}\right)\right](y).$$

Interchanging E_X and $\mathcal{F}$, we get

$$E_\xi[p_n(y|\xi)] = E_X\left[\mathcal{F}\left(e^{-\frac{\|\mathbf{X}^{(n)}\|^2}{2d_n}t^2}\right)(y)\right].$$

Now using the fact that $e^{-\frac{\|\mathbf{X}^{(n)}\|^2}{2d_n}t^2}$ is the characteristic function of the $N\left(0, \frac{\|\mathbf{X}^{(n)}\|^2}{d_n}\right)$ distribution, we obtain

$$E_\xi[p_n(y|\xi)] = E_X\left[\frac{1}{\sqrt{2\pi}}\frac{\sqrt{d_n}}{\|\mathbf{X}^{(n)}\|}e^{-\frac{d_n y^2}{2\|\mathbf{X}^{(n)}\|^2}}\right]. \tag{2.13}$$

Since the inverse is a continuous function, from (2.6) we have $\frac{d_n}{||\mathbf{X}^{(n)}||^2} \xrightarrow{\mathcal{P}} \frac{1}{\sigma^2}$, as $n \to \infty$. Using the continuity of the exponential function and by the Slutsky's Theorem, we have

$$\frac{1}{\sqrt{2\pi}}\frac{\sqrt{d_n}}{\|\mathbf{X}^{(n)}\|}e^{-\frac{d_n y^2}{2\|\mathbf{X}^{(n)}\|^2}} \xrightarrow{\mathcal{P}} \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}.$$

Since $e^{-\frac{y^2}{2\sigma^2}} \leq 1$, by the Lebesgue's Dominated Convergence Theorem, as $n \to \infty$,

$$E_\xi[p_n(y|\xi)] \longrightarrow \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}} = p_{N(0,\sigma^2)}(y), \tag{2.14}$$

where $p_{N(0,\sigma^2)}(y)$ is the probability density function of $N(0, \sigma^2)$ distribution.

Now we consider

$$\begin{aligned} E_\xi\left(p_n(y|\xi)\right)^2 &= E_\xi\left|p_n(y|\xi)\right|^2 \\ &= E_\xi\left|\mathcal{F}\left(\varphi_n(t|\xi)\right)(y)\right|^2 \end{aligned}$$

$$= E_\xi \left[E_X \left(\frac{1}{2\pi} \int\limits_t e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}} e^{-ity} dt \right) \right.$$
$$\left. E_{\tilde{X}} \left(\frac{1}{2\pi} \int\limits_s e^{-is\xi^{(n)T}\tilde{\mathbf{X}}^{(n)}/\sqrt{d_n}} e^{isy} ds \right) \right].$$

Hence

$$E_\xi \left(p_n(y|\xi)\right)^2 = E_\xi \left[E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int\limits_t e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}} e^{-ity} dt \right. \right.$$
$$\left. \left. \int\limits_s e^{-is\xi^{(n)T}\tilde{\mathbf{X}}^{(n)}/\sqrt{d_n}} e^{isy} ds \right) \right].$$

We interchange E_ξ and $E_{X,\tilde{X}}$ and combine the integrals with respect to t and s to obtain

$$\begin{aligned} E_\xi \left(p_n(y|\xi)\right)^2 &= E_{X,\tilde{X}} \left(E_\xi \left[\frac{1}{(2\pi)^2} \int\limits_t \int\limits_s e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n} - ity} e^{-is\xi^{(n)T}\tilde{\mathbf{X}}^{(n)}/\sqrt{d_n} + isy} ds dt \right] \right) \\ &= E_{X,\tilde{X}} \left(E_\xi \left[\frac{1}{(2\pi)^2} \int\limits_t \int\limits_s e^{i\xi^{(n)T}(t\mathbf{X}^{(n)} - s\tilde{\mathbf{X}}^{(n)})/\sqrt{d_n}} e^{-ity+isy} ds dt \right] \right). \end{aligned}$$

Moving E_ξ inside the double integral, we have

$$E_\xi \left(p_n(y|\xi)\right)^2 = E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int\limits_t \int\limits_s E_\xi \left[e^{i\xi^{(n)T}(t\mathbf{X}^{(n)} - s\tilde{\mathbf{X}}^{(n)})/\sqrt{d_n}} \right] e^{-ity+isy} ds dt \right).$$

Since $\xi^{(n)} \sim N(0, I_{d_n})$, we have

$$\begin{aligned} E_\xi \left(p_n(y|\xi)\right)^2 &= E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int\limits_t \int\limits_s e^{-||t\mathbf{X}^{(n)} - s\tilde{\mathbf{X}}^{(n)}||^2/(2d_n)} e^{-ity+isy} ds dt \right) \\ &= E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int\limits_t \int\limits_s e^{-\frac{1}{2d_n}\left(t^2 \sum\limits_{k=1}^{d_n} X_k^{(n)2} + s^2 \sum\limits_{k=1}^{d_n} \tilde{X}_k^{(n)2} - 2ts \sum\limits_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} \right)} e^{-ity+isy} ds dt \right). \end{aligned}$$

Denote

$$A^{(n)} = \frac{1}{d_n} \begin{pmatrix} \sum_{k=1}^{d_n} X_k^{(n)2} & -\sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} \\ -\sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} & \sum_{k=1}^{d_n} \tilde{X}_k^{(n)2} \end{pmatrix}. \tag{2.15}$$

Then

$$E_\xi \left(p_n(y|\xi)\right)^2 = E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int_t \int_s e^{-\frac{1}{2}(t,s)A^{(n)}\binom{t}{s}} e^{-ity+isy} dsdt \right),$$

which can be rewritten as

$$E_\xi \left(p_n(y|\xi)\right)^2 = E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int_t \int_s e^{-\frac{1}{2}(t,s)A^{(n)}\binom{t}{s}} e^{-i(t,s)\binom{y}{-y}} dsdt \right). \tag{2.16}$$

To find this integral, let us consider a bivariate Normal random vector

$$W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \sim N(0, \Sigma),$$

$$E(W) = 0,$$

$$E(W^T W) = \Sigma.$$

The bivariate density function of W is

$$f_W(w) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} e^{-\frac{1}{2}w^T \Sigma^{-1} w},$$

where $w \in \mathbb{R}^2$, and the characteristic function of W is

$$\varphi_W(\nu) = e^{-\frac{1}{2}\nu^T \Sigma \nu},$$

where $\nu \in \mathbb{R}^2$. Then through two-dimensional Fourier transform $\mathcal{F}$ we have

$$\begin{aligned} f_W(w) &= \mathcal{F}\left(\varphi_W(\nu)\right)(w) \\ &= \frac{1}{(2\pi)^2} \int \varphi_W(\nu) e^{-i\nu^T w} d\nu, \end{aligned}$$

[6, $\mathbb{S}$ XV.7]. That is,

$$f_W(w) = \frac{1}{(2\pi)^2} \int e^{-\frac{1}{2}\nu^T \Sigma \nu} e^{-i\nu^T w} d\nu. \tag{2.17}$$

For the integral in (2.16), let us take $w = \binom{y}{-y}$, $\nu = \binom{t}{s}$, and $\Sigma = A^{(n)}$. Then

$$\begin{aligned} E_\xi \left(p_n(y|\xi)\right)^2 &= E_{X,\tilde{X}}\left(f_W(w)\right) \\ &= E_{X,\tilde{X}}\left(\frac{1}{2\pi\sqrt{\det(A^{(n)})}} e^{-\frac{1}{2}w^T A^{(n)^{-1}} w}\right) \\ &= E_{X,\tilde{X}}\left(\frac{1}{2\pi\sqrt{\det(A^{(n)})}} e^{-\frac{1}{2}(y,-y)A^{(n)^{-1}}\binom{y}{-y}}\right). \end{aligned}$$

Consider

$$\det(A^{(n)}) = \frac{1}{d_n^2}\left(\sum_{k=1}^{d_n} X_k^{(n)2} \sum_{k=1}^{d_n} \tilde{X}_k^{(n)2} - \left(\sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)}\right)^2\right) = \frac{\Delta^{(n)}}{d_n^2}, \tag{2.18}$$

where

$$\begin{aligned} \Delta^{(n)} &= \sum_{k=1}^{d_n} X_k^{(n)2} \sum_{k=1}^{d_n} \tilde{X}_k^{(n)2} - \left(\sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)}\right)^2 \\ &= ||X^{(n)}||^2 ||\tilde{X}^{(n)}||^2 - \left(X^{(n)T}\tilde{X}^{(n)}\right)^2, \end{aligned}$$

and

$$A^{(n)^{-1}} = \frac{1}{\det(A)} \frac{1}{d_n} \begin{pmatrix} \sum_{k=1}^{d_n} X_k^{(n)2} & \sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} \\ \sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} & \sum_{k=1}^{d_n} \tilde{X}_k^{(n)2} \end{pmatrix} =$$

$$= \frac{d_n}{\Delta^{(n)}} \begin{pmatrix} \sum_{k=1}^{d_n} X_k^{(n)2} & \sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} \\ \sum_{k=1}^{d_n} X_k^{(n)} \tilde{X}_k^{(n)} & \sum_{k=1}^{d_n} \tilde{X}_k^{(n)2} \end{pmatrix}. \tag{2.19}$$

Thus we have

$$E_\xi\left(p_n(y|\xi)\right)^2 = E_{X,\tilde{X}}\left(\frac{1}{2\pi\sqrt{\frac{\Delta^{(n)}}{d_n^2}}}e^{-\frac{d_n}{2\Delta^{(n)}}\left(\sum\limits_k \tilde{X}_k^{(n)2}-2\sum\limits_k X_k^{(n)}\tilde{X}_k^{(n)}+\sum\limits_k X_k^{(n)2}\right)y^2}\right)$$
$$= E_{X,\tilde{X}}\left(\frac{d_n}{2\pi\sqrt{\Delta^{(n)}}}e^{-\frac{d_n}{2\Delta^{(n)}}\|\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}\|^2y^2}\right).$$

Since $\tilde{\mathbf{X}}^{(n)}$ is an independent copy of $\mathbf{X}^{(n)}$, we have $\frac{||\tilde{\mathbf{X}}^{(n)}||}{d_n} \xrightarrow{\mathcal{P}} \sigma^2$, as $n \to \infty$. Then by the Slutsky's Theorem, we have $\frac{||\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} 2\sigma^2$ and $\frac{\Delta^{(n)}}{d_n^2} \xrightarrow{\mathcal{P}} \sigma^4$. Furthermore, $\frac{d_n}{\sqrt{\Delta^{(n)}}} \xrightarrow{\mathcal{P}} \frac{1}{\sigma^2}$ and $\frac{d_n||\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}||^2}{2\Delta^{(n)}} \xrightarrow{\mathcal{P}} \frac{1}{\sigma^2}$. Using the continuity of the exponential function and by Slutsky's Theorem, we have

$$\frac{d_n}{2\pi\sqrt{\Delta^{(n)}}}e^{-\frac{d_n}{2\Delta^{(n)}}\|\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}\|^2y^2} \xrightarrow{\mathcal{P}} \frac{1}{2\pi\sigma^2}e^{-\frac{y^2}{\sigma^2}}.$$

Then, by the Lebesgue's Dominated Convergence Theorem, as $n \to \infty$,

$$E_{X,\tilde{X}}\left(\frac{d_n}{2\pi\sqrt{\Delta^{(n)}}}e^{-\frac{d_n}{2\Delta^{(n)}}\|\mathbf{X}^{(n)}-\tilde{\mathbf{X}}^{(n)}\|^2y^2}\right) \to \frac{1}{2\pi\sigma^2}e^{-\frac{y^2}{\sigma^2}}.$$

That is,

$$E_\xi\left(p_n(y|\xi)\right)^2 \to p^2_{N(0,\sigma^2)}(y). \tag{2.20}$$

Now we consider

$$\begin{aligned}E_\xi\left|p_n(y|\xi)-p_{N(0,\sigma^2)}(y)\right|^2 &= E_\xi\left(p_n(y|\xi)\right)^2 - 2p_{N(0,\sigma^2)}(y)E_\xi\left(p_n(y|\xi)\right) + p^2_{N(0,\sigma^2)}(y)\\ &\le |E_\xi\left(p_n(y|\xi)\right)^2 - p^2_{N(0,\sigma^2)}(y)| + 2p_{N(0,\sigma^2)}(y)|E_\xi\left(p_n(y|\xi)\right)|\\ &\le |E_\xi\left(p_n(y|\xi)\right)^2 - p^2_{N(0,\sigma^2)}(y)| + \frac{\sqrt{2}}{\sqrt{\pi}\sigma}|E_\xi\left(p_n(y|\xi)\right)|.\end{aligned}$$

Therefore, using (2.14) and (2.20), as $n \to \infty$,

$$E_\xi\left(p_n(y|\xi)-p_{N(0,\sigma^2)}(y)\right)^2 \to 0.$$

This completes the proof. $\square$

2.5 Uniform L^2_ξ-convergence of Conditional CDF

Let $F_n(z|\xi)$ denote the cumulative distribution function of the conditional distribution of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}} \;|\xi^{(n)}$.

Theorem 2.6. *If conditions (2.6) and (2.7) are met, then $F_n(z|\xi)$ converges to $F_{N(0,\sigma^2)}(z)$ uniformly in L^2_ξ as $n \to \infty$. That is, as $n \to \infty$,*

$$\sup_z E_\xi \left|F_n(z|\xi) - F_{N(0,\sigma^2)}(z)\right|^2 \to 0.$$

Proof. Using the conditional probability density function $p_n(y|\xi)$ of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{d_n}} \;|\xi^{(n)}$ we have

$$E_\xi\left(F_n(z|\xi)\right) = E_\xi\left(\int\limits_{-\infty}^{z} p_n(y|\xi)dy\right).$$

Interchanging the integral with respect to y and E_ξ, we obtain

$$E_\xi\left(F_n(z|\xi)\right) = \int\limits_{-\infty}^{z} E_\xi\left[p_n(y|\xi)\right]dy.$$

Using (2.13) we have

$$E_\xi\left(F_n(z|\xi)\right) = \int\limits_{-\infty}^{z} E_X\left(\frac{1}{\sqrt{2\pi}}\frac{\sqrt{d_n}}{\|\mathbf{X}^{(n)}\|}e^{-\frac{d_n y^2}{2\|\mathbf{X}^{(n)}\|^2}}\right)dy.$$

By interchanging the integral with respect to y and E_X, we obtain

$$\begin{aligned} E_\xi\left(F_n(z|\xi)\right) &= E_X\left(\int\limits_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\frac{\sqrt{d_n}}{\|\mathbf{X}^{(n)}\|}e^{-\frac{d_n y^2}{2\|\mathbf{X}^{(n)}\|^2}}dy\right) \\ &= E_X\left(F_{N\left(0,\frac{\|\mathbf{X}^{(n)}\|^2}{d_n}\right)}(z)\right), \end{aligned}$$

where $F_{N\left(0,\frac{\|\mathbf{X}^{(n)}\|^2}{d_n}\right)}(z)$ denotes the cumulative distribution function of the $N\left(0, \frac{\|\mathbf{X}^{(n)}\|^2}{d_n}\right)$ distribution.

Now for the integral in

$$E_\xi\left(F_n(z|\xi)\right) = E_X\left[\int_{-\infty}^{z} \frac{\sqrt{d_n}}{\sqrt{2\pi}||\mathbf{X}^{(n)}||} e^{-\frac{d_n y^2}{2||\mathbf{X}^{(n)}||^2}} dy\right]$$

let us consider the following change of variable $u = \frac{\sqrt{d_n} y\sigma}{||\mathbf{X}^{(n)}||}$. Then

$$E_\xi\left(F_n(z|\xi)\right) = E_X\left[\int_{-\infty}^{\sqrt{d_n}z\sigma/||\mathbf{X}^{(n)}||} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}} du\right].$$

Since $\frac{||\mathbf{X}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} \sigma^2$, by the Slutsky's Theorem, we have $\frac{\sqrt{d_n}z\sigma}{||\mathbf{X}^{(n)}||} \xrightarrow{\mathcal{P}} z$. Since the above integral is continuous as a function of its upper limit, we have

$$\int_{-\infty}^{\sqrt{d_n}z\sigma/||\mathbf{X}^{(n)}||} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}} du \quad \xrightarrow{\mathcal{P}} \quad \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}} du = F_{N(0,\sigma^2)}\left(z\right).$$

Furthermore, the integral is bounded by 1. Therefore, by the Lebesgue's Dominated Convergence Theorem,

$$E_X\left[\int_{-\infty}^{\sqrt{d_n}z\sigma/||\mathbf{X}^{(n)}||} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}} du\right] \to \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}} du.$$

That is, as $n \to \infty$,

$$E_\xi\left(F_n(z|\xi)\right) = E_X\left[F_{N(0,\sigma^2)}\left(\frac{\sqrt{d_n}z\sigma}{||\mathbf{X}^{(n)}||}\right)\right] \longrightarrow F_{N(0,\sigma^2)}\left(z\right).$$

Now let us consider

$$\begin{aligned} E_\xi\left(F_n(z|\xi)\right)^2 &= E_\xi\left|F_n(z|\xi)\right|^2 \\ &= E_\xi\left|\int_{-\infty}^{z} \mathcal{F}\left(\varphi_n(t|\xi)\right)(y) dy\right|^2. \end{aligned}$$

Using (2.8) we obtain

$$E_\xi \left(F_n(z|\xi)\right)^2 = E_\xi \left| \int_{-\infty}^{z} \mathcal{F}\left(E_X\left[e^{||\mathbf{X}^{(n)}||^2 t^2/(2d_n)}\right]\right)(y)dy \right|^2 .$$

Interchanging $\mathcal{F}$ and E_X, we have

$$E_\xi \left(F_n(z|\xi)\right)^2 = E_\xi \left[\int_{-\infty}^{z} E_X \left(\frac{1}{2\pi} \int_t e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}} e^{-ity} dt \right) dy \right.$$
$$\left. \int_{-\infty}^{z} E_{\tilde{X}} \left(\frac{1}{2\pi} \int_s e^{-is\xi^{(n)T}\tilde{\mathbf{X}}^{(n)}/\sqrt{d_n}} e^{isy'} ds \right) dy' \right].$$

Combining the integrals with respect to y and y' and expectations E_X and $E_{\tilde{X}}$, we obtain

$$E_\xi \left(F_n(z|\xi)\right)^2 = E_\xi \left[\int_{-\infty}^{z} \int_{-\infty}^{z} E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \int_t e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n}} e^{-ity} dt \right. \right.$$
$$\left. \left. \int_s e^{-is\xi^{(n)T}\tilde{\mathbf{X}}^{(n)}/\sqrt{d_n}} e^{isy'} ds \right) dydy' \right].$$

Furthermore, by combining the integrals with respect to s and t and moving E_ξ inside, we have

$$E_\xi \left(F_n(z|\xi)\right)^2 = \int_{-\infty}^{z} \int_{-\infty}^{z} E_{X,\tilde{X}} \left(\frac{1}{(2\pi)^2} \right.$$
$$\left. \int_t \int_s E_\xi \left[e^{it\xi^{(n)T}\mathbf{X}^{(n)}/\sqrt{d_n} - ity} e^{-is\xi^{(n)T}\tilde{\mathbf{X}}^{(n)}/\sqrt{d_n} + isy'} \right] dsdt \right) dydy'. \quad (2.21)$$

Inside the exponent in (2.21) we separate the terms that contain $\xi^{(n)}$ from those that do not and rewrite it as follows

$$E_\xi\left(F_n(z|\xi)\right)^2 = \int_{-\infty}^{z}\int_{-\infty}^{z} E_{X,\tilde{X}}\left(\frac{1}{(2\pi)^2}\int_t\int_s E_\xi\left[e^{-i\xi^{(n)T}(s\tilde{\mathbf{X}}^{(n)}-t\mathbf{X}^{(n)})/\sqrt{d_n}}e^{-ity+isy'}\right]dsdt\right)dydy'.$$

Since $\xi^{(n)} \sim N(0, I_{d_n})$, we have

$$E_\xi\left(F_n(z|\xi)\right)^2 = \int_{-\infty}^{z}\int_{-\infty}^{z} E_{X,\tilde{X}}\left(\frac{1}{(2\pi)^2}\int_t\int_s e^{-\frac{1}{2}(t,s)A^{(n)}\binom{t}{s}}e^{-ity+isy'}dsdt\right)dydy',$$

where the matrix $A^{(n)}$ is specified in (2.15), and we can write

$$E_\xi\left(F_n(z|\xi)\right)^2 = \int_{-\infty}^{z}\int_{-\infty}^{z} E_{X,\tilde{X}}\left(\frac{1}{(2\pi)^2}\int_t\int_s e^{-\frac{1}{2}(t,s)A^{(n)}\binom{t}{s}}e^{-i(t,s)\binom{y}{-y'}}dsdt\right)dydy'.$$

Using (2.17) with $w = \binom{y}{-y'}$, $\nu = \binom{t}{s}$, and $\Sigma = A^{(n)}$, (2.18), and (2.19), we have

$$\begin{aligned}
E_\xi\left(F_n(z|\xi)\right)^2 &= \int_{-\infty}^{z}\int_{-\infty}^{z} E_{X,\tilde{X}}\left(\frac{1}{2\pi\sqrt{\det(A^{(n)})}}e^{-\frac{1}{2}(y,-y')A^{(n)^{-1}}\binom{y}{-y'}}\right)dydy' \\
&= \int_{-\infty}^{z}\int_{-\infty}^{z} E_{X,\tilde{X}}\left(\frac{1}{2\pi\sqrt{\frac{\Delta^{(n)}}{d_n^2}}}e^{-\frac{1}{2}(y,-y')\frac{d_n}{\Delta^{(n)}}\begin{pmatrix}\sum_k \tilde{X}_k^{(n)2} & \sum_k X_k^{(n)}\tilde{X}_k^{(n)} \\ \sum_k X_k^{(n)}\tilde{X}_k^{(n)} & \sum_k X_k^{(n)2}\end{pmatrix}\binom{y}{-y'}}\right)dydy' \\
&= \int_{-\infty}^{z}\int_{-\infty}^{z} E_{X,\tilde{X}}\left(\frac{1}{2\pi\sqrt{\frac{\Delta^{(n)}}{d_n^2}}}e^{-\frac{1}{2}\frac{d_n}{\Delta^{(n)}}\left(\sum_k \tilde{X}_k^{(n)2}y^2-2\sum_k X_k^{(n)}\tilde{X}_k^{(n)}yy'+\sum_k X_k^{(n)2}y'^2\right)}\right)dydy' \\
&= \int_{-\infty}^{z}\int_{-\infty}^{z}\left[E_{X,\tilde{X}}\left(\frac{d_n}{2\pi\sqrt{\Delta^{(n)}}}e^{-\frac{d_n}{2\Delta^{(n)}}\|\tilde{\mathbf{X}}^{(n)}y-\mathbf{X}^{(n)}y'\|^2}\right)\right]dydy'.
\end{aligned}$$

We interchange $E_{X,\tilde{X}}$ and the double integral to obtain

$$E_\xi\left(F_n^2(z|\xi)\right) = E_{X,\tilde{X}}\left[\int_{-\infty}^{z}\int_{-\infty}^{z} \frac{d_n}{2\pi\sqrt{\Delta^{(n)}}} e^{-\frac{d_n}{2\Delta^{(n)}}||\tilde{\mathbf{X}}^{(n)}y-\mathbf{X}^{(n)}y'||^2} dydy'\right].$$

Let us consider the following change of variables

$$u = \frac{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}{\sqrt{\Delta^{(n)}}}y - \frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||\sqrt{\Delta^{(n)}}}y'$$

and

$$v = \frac{\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||}y'.$$

Then the Jacobian is

$$J = \frac{\delta y}{\delta u}\frac{\delta y'}{\delta v} - \frac{\delta y}{\delta v}\frac{\delta y'}{\delta u} = \frac{\sqrt{\Delta^{(n)}}}{d_n}.$$

Therefore

$$\begin{aligned} &E_{X,\tilde{X}}\left[\int_{-\infty}^{z}\int_{-\infty}^{z} \frac{d_n}{2\pi\sqrt{\Delta^{(n)}}} e^{-\frac{d_n}{2\Delta^{(n)}}||\tilde{\mathbf{X}}^{(n)}y-\mathbf{X}^{(n)}y'||^2} dydy'\right] \\ &= E_{X,\tilde{X}}\left[\int_{D} \frac{1}{2\pi} e^{-\frac{u^2+v^2}{2}} dudv\right], \end{aligned}$$

where the set D is

$$D = \left\{\frac{\sqrt{\Delta^{(n)}}}{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}u + \frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})}{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}v \le z, \quad \frac{||\tilde{\mathbf{X}}^{(n)}||}{\sqrt{d_n}}v \le z\right\},$$

or equivalently,

$$D = \left\{u \le \frac{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}{\sqrt{\Delta^{(n)}}}z - \frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})}{\sqrt{\Delta^{(n)}}}v, \quad v \le \frac{\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||}z\right\}.$$

Therefore,

$$
\begin{aligned}
&E_{X,\tilde{X}}\left[\int\limits_D \frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv\right] \\
&= E_{X,\tilde{X}}\left[\int\limits_{-\infty}^{\frac{\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||}z}\int\limits_{-\infty}^{\frac{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}{\sqrt{\Delta^{(n)}}}z-\frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})}{\sqrt{\Delta^{(n)}}}v}\frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv\right].
\end{aligned}
$$

Since $\frac{||\tilde{\mathbf{X}}^{(n)}||^2}{d_n} \xrightarrow{\mathcal{P}} \sigma^2$, by Slutsky's Theorem we have $\frac{\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||}z \xrightarrow{\mathcal{P}} \frac{z}{\sigma}$. Similarly, $\frac{\Delta^{(n)}}{d_n^2} \xrightarrow{\mathcal{P}} \sigma^4$ implies $\frac{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}{\sqrt{\Delta^{(n)}}}z - \frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})}{\sqrt{\Delta^{(n)}}}v \xrightarrow{\mathcal{P}} \frac{z}{\sigma}$. And since the integrals above are continuous as functions of their upper limits, we have

$$
\int\limits_{-\infty}^{\frac{\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||}z}\int\limits_{-\infty}^{\frac{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}{\sqrt{\Delta^{(n)}}}z-\frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})}{\sqrt{\Delta^{(n)}}}v}\frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv \xrightarrow{\mathcal{P}} \int\limits_{-\infty}^{\frac{z}{\sigma}}\int\limits_{-\infty}^{\frac{z}{\sigma}}\frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv
$$

and the limit

$$
\begin{aligned}
\int\limits_{-\infty}^{\frac{z}{\sigma}}\int\limits_{-\infty}^{\frac{z}{\sigma}}\frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv &= \left(\int\limits_{-\infty}^{\frac{z}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du\right)^2 \\
&= F^2_{N(0,\sigma^2)}(z).
\end{aligned}
$$

Since the exponents above are bounded by 1, by the Lebesgue's Dominated Convergence Theorem,

$$
E_{X,\tilde{X}}\left[\int\limits_{-\infty}^{\frac{\sqrt{d_n}}{||\tilde{\mathbf{X}}^{(n)}||}z}\int\limits_{-\infty}^{\frac{||\tilde{\mathbf{X}}^{(n)}||\sqrt{d_n}}{\sqrt{\Delta^{(n)}}}z-\frac{(\mathbf{X}^{(n)T}\tilde{\mathbf{X}}^{(n)})}{\sqrt{\Delta^{(n)}}}v}\frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}dudv\right] \longrightarrow F^2_{N(0,\sigma^2)}(z).
$$

Then

$$
\begin{aligned}
&E_\xi \left[F_n(z|\xi) - F_{N(0,\sigma^2)}(z)\right]^2 \\
&\leq \left|E_\xi \left(F_n^2(z|\xi)\right) - F^2_{N(0,\sigma^2)}(z)\right| + 2F_{N(0,\sigma^2)}(z) \left|E_\xi \left(F_n(z|\xi)\right) - F_{N(0,\sigma^2)}(z)\right| \\
&\leq \left|E_\xi \left(F_n^2(z|\xi)\right) - F^2_{N(0,\sigma^2)}(z)\right| + 2 \left|E_\xi \left(F_n(z|\xi)\right) - F_{N(0,\sigma^2)}(z)\right| \longrightarrow 0,
\end{aligned}
$$

as $n \to \infty$. Therefore $F_n(z|\xi)$ converges to $F_{N(0,\sigma^2)}(z)$ in L^2 with respect to P_ξ, as $n \to \infty$.

Furthermore, we note that $E_\xi \left(F_n(z|\xi)\right)$ is a cumulative distribution function, because it is non-decreasing, and converges to 0 when $z \to -\infty$, and to 1 when $z \to \infty$. Similarly, $E_\xi \left(F_n^2(z|\xi)\right)$ is also a cumulative distribution function. And since the limiting cumulative distribution function $F_{N(0,\sigma^2)}(z)$ is continuous, by Glivenko-Cantelli Theorem the L^2_ξ-convergence of $F_n(z|\xi)$ to $F_{N(0,\sigma^2)}(z)$ is uniform with respect to z. That is, as $n \to \infty$,

$$
\sup_z E_\xi \left|F_n(z|\xi) - F_{N(0,\sigma^2)}(z)\right|^2 \to 0.
$$

This completes the proof. $\square$

2.6 Limit Theorems for Modulated Random Vectors with Resampling

Earlier in this chapter in Theorem 2.4 we proved the convergence of the conditional distribution of randomly modulated versions of random vectors to normal distribution weakly in probability. However, the conditionality and weak convergence in probability are not convenient for developing applications of the result to statistics. For this reason in this section we modify the earlier result and show the convergence of the distribution of randomly modulated vectors to normal in distribution with respect to the joint probability measure $P_X \times P_\xi$ [17]. In addition, since the convergence in distribution is weaker, the required assumptions are weaker as well.

Let $\mathbf{X}^{(n,1)}, \ldots, \mathbf{X}^{(n,s)} \in \mathbb{R}^{d_n}$, $s \geq 1$, be d_n-dimensional random vectors on $(\Omega_X, \mathfrak{F}_X, P_X)$.

Suppose that for every $j = 1, \dots, s$ and $k = 1, \dots, s$, as $n \to \infty$, we have

$$\frac{1}{d_n} \sum_{i=1}^{d_n} X_i^{(n,j)} X_i^{(n,k)} \xrightarrow{\mathcal{P}} \lambda^{(j,k)} \tag{2.22}$$

for some constants $\lambda^{(j,k)}$. Comparing this to (2.6), let us note that when $s = 1$,

$$\lambda^{(1,1)} = \sigma^2.$$

We consider the $s \times s$ matrix

$$\Sigma = \begin{pmatrix} \lambda^{(1,1)} & \dots & \lambda^{(1,s)} \\ & \vdots & \\ \lambda^{(s,1)} & \dots & \lambda^{(s,s)} \end{pmatrix}$$

and the $rs \times rs$ matrix

$$\Sigma^{[r]} = \operatorname{diag}(\Sigma, \dots, \Sigma).$$

That is,

$$\Sigma^{[r]} = \begin{pmatrix} \Sigma & 0 & \dots & 0 \\ 0 & \Sigma & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \Sigma \end{pmatrix}$$

Lemma 2.2. *The matrices Σ and $\Sigma^{[r]}$ are positive definite.*

Proof. For any $j = 1, \dots, s$, $k = 1, \dots, s$ we have $\lambda^{(j,k)} = \lambda^{(k,j)}$, so the matrix Σ is symmetric. Furthermore, for any vector $\alpha = (\alpha_1, \dots, \alpha_s)$, we consider

$$\sum_{j=1}^{s}\sum_{k=1}^{s}\left(\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^{(n,j)} X_i^{(n,k)}\right)\alpha_j\alpha_k$$
$$= \frac{1}{d_n}\sum_{i=1}^{d_n}\left(\sum_{j=1}^{s}\sum_{k=1}^{s} X_i^{(n,j)} X_i^{(n,k)}\alpha_j\alpha_k\right)$$
$$= \frac{1}{d_n}\sum_{i=1}^{d_n}\left(\sum_{j=1}^{s} X_i^{(n,j)}\alpha_j\right)^2 \geq 0.$$

Then using (2.22), we have

$$\sum_{j=1}^{s}\sum_{k=1}^{s}\lambda^{(j,k)}\alpha_j\alpha_k \ \geq 0.$$

Thus, for any α we have $\alpha\Sigma\alpha^T \geq 0$, that is Σ is a positive definite matrix.

Similarly, the matrix $\Sigma^{[r]}$ is positive definite. □

If we have several d_n-dimensional vectors $\mathbf{X}^{(n,1)}, \dots, \mathbf{X}^{(n,s)}$, we can consider the modulated versions of these vectors

$$\frac{1}{\sqrt{d_n}}\sum_{i=1}^{d_n}\xi_i^{(n)} X_i^{(n,j)},$$

$j = 1, \dots, s$, using the same modulating vector $\xi^{(n)}$ for all vectors $\mathbf{X}^{(n,1)}, \dots, \mathbf{X}^{(n,s)}$. We can also use *independent resamplings*, namely several modulated versions of the vectors $\mathbf{X}^{(n,1)}, \dots, \mathbf{X}^{(n,s)}$, obtained by using independent vectors $\xi^{(n,m)}$, $m = 1, \dots, r$.

Remark 2.1. Intuitively, the introduction of the resamplings can be considered as a way to examine the random "cloud" of the vectors $\mathbf{X}^{(n,1)}, \dots, \mathbf{X}^{(n,s)}$ using its different projections; certainly this increases the information we obtain from the data.

Theorem 2.7. *[17] Let $n \geq 1$ and $\mathbf{X}^{(n,1)}, \dots, \mathbf{X}^{(n,s)} \in \mathbb{R}^{d_n}$, s fixed, $s \geq 1$, be d_n-dimensional random vectors on $(\Omega_X, \mathfrak{F}_X, P_X)$. Suppose that condition (2.22) is satisfied for every $j = 1, \dots, s$ and $k = 1, \dots, s$. Let $\xi^{(n,1)}, \dots, \xi^{(n,r)}$, $r \geq 1$, be independent $N(0, I_{d_n})$ vectors on $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$, and $\mathbf{X}^{(n,j)} \perp\!\!\!\perp \xi^{(n,m)}$ for every $j = 1, \dots, s$ and $m = 1, \dots, r$. We* *consider*

$$\Gamma_n^{(m,j)} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \xi_i^{(n,m)} X_i^{(n,j)},$$

$j = 1, \ldots, s,\ m = 1, \ldots, r,$

$$\mathbf{\Gamma}_n^{(m)} = \left(\Gamma_n^{(m,1)}, \ldots, \Gamma_n^{(m,s)}\right),$$

$m = 1, \ldots, r,$ *and*

$$\mathbf{\Psi}_n = \left(\mathbf{\Gamma}_n^{(1)}, \ldots, \mathbf{\Gamma}_n^{(r)}\right).$$

Then

$$\mathbf{\Psi}_n \quad \overset{\mathcal{D}}{\to} \quad Y \sim N(0, \Sigma^{[r]}),$$

as $n \to \infty$*, with respect to the joint probability measure* $P_X \times P_\xi$*.*

Proof. Let us fix n and consider the characteristic function of $\mathbf{\Psi}_n$ at $\mathbf{t} = \left(\mathbf{t}^{(1)}, \ldots, \mathbf{t}^{(r)}\right)$, where $\mathbf{t}^{(m)} = \left(t_1^{(m)}, \ldots, t_s^{(m)}\right)$, $m = 1, \ldots, r$. We have

$$\begin{aligned}
\varphi_{\mathbf{\Psi}_n}(\mathbf{t}) &= E_{\xi,X}\left(e^{i\sum_{j=1}^{s}\sum_{m=1}^{r} t_j^{(m)} \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \xi_i^{(n,m)} X_i^{(n,j)}}\right) \\
&= E_X\left(E_\xi\left(e^{i\sum_{m=1}^{r}\sum_{i=1}^{d_n}\sum_{j=1}^{s} \frac{1}{\sqrt{d_n}} t_j^{(m)} \xi_i^{(n,m)} X_i^{(n,j)}} |X\right)\right).
\end{aligned}$$

Since $\xi_i^{(n,m)}$, $i = 1, \ldots, d_n$, $m = 1, \ldots, r$, are independent $N(0,1)$,

$$\varphi_{\mathbf{\Psi}_n}(\mathbf{t}) = E_X\left(\prod_{m=1}^{r} E_\xi\left(e^{i\sum_{i=1}^{d_n}\sum_{j=1}^{s} \frac{1}{\sqrt{d_n}} t_j^{(m)} \xi_i^{(n,m)} X_i^{(n,j)}} |X\right)\right),$$

and using the characteristic function of the normal distribution, we have

$$
\begin{aligned}
\varphi_{\boldsymbol{\Psi}_n}(\mathbf{t}) &= E_X\left(\prod_{m=1}^{r} e^{-\sum_{i=1}^{d_n}\frac{1}{2d_n}\left(\sum_{j=1}^{s} t_j^{(m)} X_i^{(n,j)}\right)^2}\right)\\
&= E_X\left(\prod_{m=1}^{r} e^{-\frac{1}{2d_n}\sum_{i=1}^{d_n}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)} X_i^{(n,j)} X_i^{(n,k)}}\right)\\
&= E_X\left(\prod_{m=1}^{r} e^{-\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)}\left[\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^{(n,j)} X_i^{(n,k)}\right]}\right).
\end{aligned}
$$

Since the condition (2.22) is satisfied for every $j = 1, \ldots, s$ and $k = 1, \ldots, s$, then

$$
\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)}\left[\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^{(n,j)} X_i^{(n,k)}\right] \overset{\mathcal{P}}{\to} \sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)} \lambda^{(j,k)}.
$$

Since the exponential function is continuous, we have

$$
e^{-\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)}\left[\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^{(n,j)} X_i^{(n,k)}\right]} \overset{\mathcal{P}}{\to} e^{-\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)} \lambda^{(j,k)}},
$$

and furthermore

$$
\prod_{m=1}^{r} e^{-\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)}\left[\frac{1}{d_n}\sum_{i=1}^{d_n} X_i^{(n,j)} X_i^{(n,k)}\right]} \overset{\mathcal{P}}{\to} \prod_{m=1}^{r} e^{-\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)} \lambda^{(j,k)}}.
$$

The exponents above are bounded by 1, therefore by Lebesgue's Dominated Convergence Theorem,

$$
\varphi_{\boldsymbol{\Psi}_n}(\mathbf{t}) \longrightarrow \prod_{m=1}^{r} e^{-\frac{1}{2}\sum_{j=1}^{s}\sum_{k=1}^{s} t_j^{(m)} t_k^{(m)} \lambda^{(j,k)}} = \prod_{m=1}^{r} \varphi_{\mathbf{Y}^{(m)}}(\mathbf{t}^{(m)}),
$$

where $\mathbf{Y}^{(m)}$, $m = 1, \ldots, r$, are independent $N(0, \Sigma)$, $\mathbf{t}^{(m)} = \left(t_1^{(m)}, \ldots, t_s^{(m)}\right)$, $m = 1, \ldots, r$.

We denote $\mathbf{Y} = \left(\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(r)}\right)$, then as $n \to \infty$, we have

$$
\Psi_n \overset{\mathcal{D}}{\to} \mathbf{Y} \sim N(0, \Sigma^{[r]}),
$$

with respect to the joint probability measure $P_X \times P_\xi$. $\square$

Let us consider a special case when $r = 1$, that is no resampling is made. The following corollary follows directly from Theorem 2.7.

Corollary 2.1. *Let $n \geq 1$ and $\mathbf{X}^{(n,1)}, \ldots, \mathbf{X}^{(n,s)} \in \mathbb{R}^{d_n}$, $s \geq 1$, be d_n-dimensional random vectors on $(\Omega_X, \mathfrak{F}_X, P_X)$. Suppose that condition (2.22) is satisfied for every $j = 1, \ldots, s$ and $k = 1, \ldots, s$. Let $\xi^{(n)} \sim N(0, I_{d_n})$ be a random vector on $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$, and $\mathbf{X}^{(n,j)} \perp\!\!\!\perp \xi^{(n)}$ for every $j = 1, \ldots, s$. We consider*

$$\Psi_n^{(j)} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \xi_i^{(n)} X_i^{(n,j)},$$

$j = 1, \ldots, s$, and

$$\mathbf{\Psi}_n = \left(\Psi_n^{(1)}, \ldots, \Psi_n^{(s)} \right).$$

Then

$$\mathbf{\Psi}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N(0, \Sigma),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$.

Another special situation of interest arises when we consider several independent resamplings but only one vector $X^{(n)}$, that is $s = 1$.

Corollary 2.2. *Let $n \geq 1$ and $\mathbf{X}^{(n)} \in \mathbb{R}^{d_n}$ be a d_n-dimensional random vector on $(\Omega_X, \mathfrak{F}_X, P_X)$, $\xi^{(n,m)} \sim N(0, I_{d_n})$, $m = 1, \ldots, r$, $r \geq 1$, be independent random vectors on $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$, and $\mathbf{X}^{(n)} \perp\!\!\!\perp \xi^{(n,m)}$ for every $m = 1, \ldots, r$. We denote*

$$\Psi_n^{(m)} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \xi_i^{(n,m)} X_i^{(n)}$$

for $m = 1, \ldots, r$, and

$$\mathbf{\Psi}_n = \left(\Psi_n^{(1)}, \ldots, \Psi_n^{(r)} \right).$$

If condition (2.6) is satisfied, then with respect to the joint probability measure $P_X \times P_\xi$,

$$\mathbf{\Psi}_n \overset{\mathcal{D}}{\to} \mathbf{Y} \sim N(0, \sigma^2 I_r),$$

as $n \to \infty$.

At last, if there is one vector $\mathbf{X}^{(n)}$ and one vector $\xi^{(n)}$, that is, when both $s = 1$ and $r = 1$, we have the following result.

Corollary 2.3. *Let* $n \geq 1$ *and* $\mathbf{X}^{(n)} \in \mathbb{R}^{d_n}$ *be a* d_n-*dimensional random vector on* $(\Omega_X, \mathfrak{F}_X, P_X)$, $\xi^{(n)} \sim N(0, I_{d_n})$ *be a random vector on* $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$, *and* $\mathbf{X}^{(n)} \perp\!\!\!\perp \xi^{(n)}$. *If condition (2.6) is satisfied, then with respect to the joint probability measure* $P_X \times P_\xi$

$$\Psi_n = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \xi_i^{(n)} X_i^{(n)} \overset{\mathcal{D}}{\to} Y \sim N(0, \sigma^2),$$

as $n \to \infty$.

Let us also discuss the case when we consider an independent resampling for *each* of the vectors $\mathbf{X}^{(n,j)}$, $j = 1, \ldots, s$.

Theorem 2.8. *Let* $n \geq 1$ *and* $\mathbf{X}^{(n,1)}, \ldots, \mathbf{X}^{(n,s)} \in \mathbb{R}^{d_n}$ *be* d_n-*dimensional random vectors on* $(\Omega_X, \mathfrak{F}_X, P_X)$, $\xi^{(n,j)} \sim N(0, I_{d_n})$, $j = 1, \ldots, s$, $r \geq 1$, *be independent random vectors on* $(\Omega_\xi, \mathfrak{F}_\xi, P_\xi)$, *and* $\mathbf{X}^{(n,j)} \perp\!\!\!\perp \xi^{(n,j)}$. *We denote*

$$\Psi_n^{(j)} = \frac{1}{\sqrt{d_n}} \sum_{i=1}^{d_n} \xi_i^{(n,j)} X_i^{(n,j)}$$

for $j = 1, \ldots, s$, *and*

$$\mathbf{\Psi}_n = \left(\Psi_n^{(1)}, \ldots, \Psi_n^{(s)} \right).$$

If condition (2.22) is satisfied for every $j = k = 1, \ldots, s$, *then with respect to the joint probability measure* $P_X \times P_\xi$,

$$\mathbf{\Psi}_n \overset{\mathcal{D}}{\to} \mathbf{Y} \sim N(0, diag(\Sigma)),$$

as $n \to \infty$, where

$$diag(\Sigma) = \begin{pmatrix} \lambda^{(1,1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda^{(s,s)} \end{pmatrix}.$$

PROOF. Let us fix n. The characteristic function of $\mathbf{\Psi}_n$ at $\mathbf{t} = (t_1, \ldots, t_s)$ is

$$\begin{aligned} \varphi_{\mathbf{\Psi}_n}(\mathbf{t}) &= E_{\xi,X}\left(e^{i\sum\limits_{j=1}^{s} t_j \frac{1}{\sqrt{d_n}} \sum\limits_{i=1}^{d_n} \xi_i^{(n,j)} X_i^{(n,j)}}\right) \\ &= E_X\left(E_\xi\left(e^{i\sum\limits_{i=1}^{d_n}\sum\limits_{j=1}^{s} \frac{1}{\sqrt{d_n}} t_j \xi_i^{(n,j)} X_i^{(n,j)}} |X\right)\right). \end{aligned}$$

Since $\xi_i^{(n,j)}$, $i = 1, \ldots, d_n$, $j = 1, \ldots, s$, are independent $N(0,1)$,

$$\varphi_{\mathbf{\Psi}_n}(\mathbf{t}) = E_X\left(\prod_{j=1}^{s} E_\xi\left(e^{i\sum\limits_{i=1}^{d_n} \frac{1}{\sqrt{d_n}} t_j \xi_i^{(n,j)} X_i^{(n,j)}} |X\right)\right).$$

Using the characteristic function of the normal distribution, we have

$$\begin{aligned} \varphi_{\mathbf{\Psi}_n}(\mathbf{t}) &= E_X\left(\prod_{j=1}^{s} e^{-\frac{1}{2d_n}\sum\limits_{i=1}^{d_n}\left(t_j X_i^{(n,j)}\right)^2}\right) \\ &= E_X\left(\prod_{j=1}^{s} e^{-\frac{1}{2d_n}\sum\limits_{i=1}^{d_n} t_j^2 (X_i^{(n,j)})^2}\right). \end{aligned}$$

Since condition (2.22) is satisfied for every $j = 1, \ldots, s$, we have, as $n \to \infty$,

$$\frac{1}{d_n}\sum_{i=1}^{d_n}\left(X_i^{(n,j)}\right)^2 \xrightarrow{\mathcal{P}} \lambda^{(j,j)}.$$

Because the exponential function is continuous, we have, for every $j = 1, \ldots, s$

$$e^{-\frac{1}{2d_n}\sum\limits_{i=1}^{d_n} t_j^2 (X_i^{(n,j)})^2} \quad \xrightarrow{\mathcal{P}} \quad e^{-\frac{1}{2}t_j^2 \lambda^{(j,j)}},$$

as $n \to \infty$, and

$$\prod_{j=1}^{s} e^{-\frac{1}{2d_n}\sum_{i=1}^{d_n} t_j^2 (X_i^{(n,j)})^2} \quad \stackrel{\mathcal{P}}{\to} \quad \prod_{j=1}^{s} e^{-\frac{1}{2}t_j^2 \lambda^{(j,j)}}.$$

Since the exponents above are bounded by 1, by Lebesgue's Dominated Convergence Theorem,

$$\varphi_{\mathbf{\Psi}_n}(\mathbf{t}) \to \prod_{j=1}^{s} e^{-\frac{1}{2}t_j^2 \lambda^{(j,j)}},$$

which is the characteristic function of $Y \sim N(0, \mathrm{diag}(\Sigma))$. Therefore

$$\Psi_n \quad \stackrel{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N(0, \mathrm{diag}(\Sigma)),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$. $\square$

Chapter 3

Limit Theorems for Ergodic Stationary Random Sequences

This chapter begins with a description of an example of an ergodic stationary sequence, which is then used to illustrate the motivation for our study. We present the limit theory for randomly modulated sequences of random vectors without the assumptions of ergodicity and stationarity, and then the limit theory for randomly modulated ergodic stationary sequences with and without resampling following the ideas and results discussed in Section 2.6. In the last section of this chapter we estimate the rate of convergence of the joint cumulative distribution function to the normal cumulative distribution function.

3.1 Illustration of Ergodic Stationary Random Sequences by an Example

In order to simulate an ergodic stationary sequence $X_1, X_2, \ldots$ with a given marginal distribution of X_1 we consider an ergodic measure-preserving transformation τ on some space $(\Omega, \mathfrak{F}, P)$ and a measurable function f such that $f \stackrel{\mathcal{D}}{=} X_1$. Let

$$X_i = f(\tau^{i-1}\omega), \quad i = 1, 2, \ldots,$$

where $\omega \in \Omega$. This means that we need to generate a random variable $\omega \in \Omega$. Then $X_1, X_2, \ldots$ has the required properties.

Example 3.1. Let us consider $\Omega = [0, 1)$ with Lebesgue measure P, and the transformation $\tau : \Omega \mapsto \Omega$ such that for a constant $c \in \Omega$

$$\tau x = x + c \pmod 1.$$

That is,

$$\tau x = \begin{cases} x + c, & \text{if } 0 \leq x < 1 - c \\ x + c - 1, & \text{if } 1 - c \leq x < 1 \end{cases}.$$

Transformation τ is measure-preserving, and therefore for $\omega \in \Omega$ the resulting sequence $\omega, \tau\omega, \tau^2\omega, \ldots$ is stationary. Furthermore, if the constant c is an irrational number, then the transformation and the sequence are also ergodic. As was mentioned in Chapter 1, this transformation corresponds to the rotation by a fixed angle on the unit complex circle. And when it is ergodic, that is $c \in \Omega \setminus \mathbb{Q}$, it is often referred to as *ergodic rotation.*

Let f be the identity function, that is $f(x) = x$. We define the random stationary sequence as follows: if $\omega \in \Omega$ is randomly selected, we consider the random sequence

$$X_1 = \omega, X_2 = \tau\omega, X_3 = \tau^2\omega, \ldots.$$

Figures 3.1 - 3.3 below demonstrate the relationship between X_i and $X_{i+1} = \tau X_i$, $i = 1, 2, \ldots$, for values of $c = \frac{\pi}{4}$, $\frac{\pi}{6}$, and $\frac{\pi}{9}$, respectively.

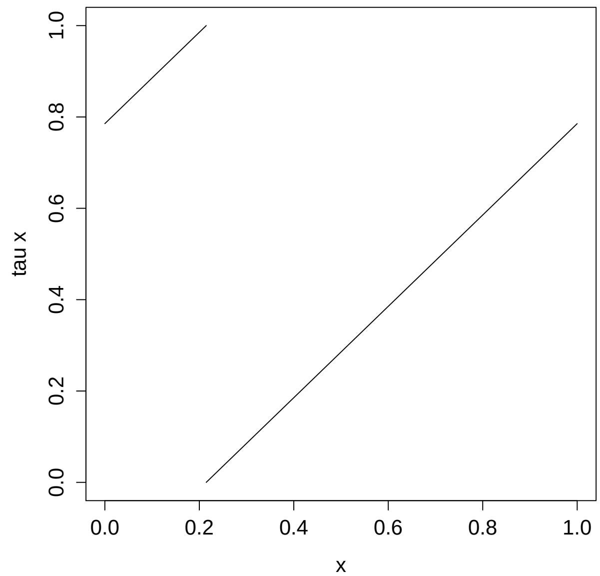

Figure 3.1: Plot of x against τx for $c = \frac{\pi}{4}$ in Example 3.1.

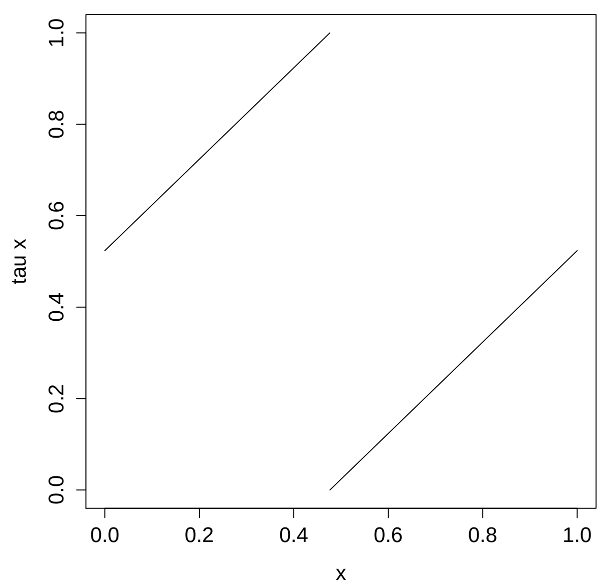

Figure 3.2: Plot of x against τx for $c = \frac{\pi}{6}$ in Example 3.1.

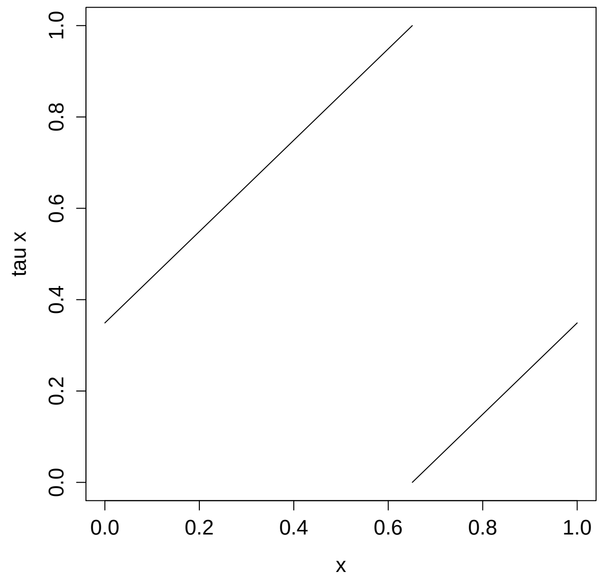

Figure 3.3: Plot of x against τx for $c = \frac{\pi}{9}$ in Example 3.1.

The correlation of the neighboring elements of the sequence is

$$\begin{aligned}\mathrm{Corr}(X_i, X_{i+1}) &= 6c^2 - 6c + 1 \\ &= 6\left(c - \frac{1}{2}\right)^2 - \frac{1}{2}.\end{aligned}$$

That is, depending on the value of $c \in \Omega$, the correlation between neighboring elements of the sequence may vary from high to low. Therefore, by choosing c we can make them strongly correlated or uncorrelated. Figure 3.4 below illustrates the behavior of $\mathrm{Corr}(X_i, X_{i+1})$ for different values of constant $c \in \Omega$.

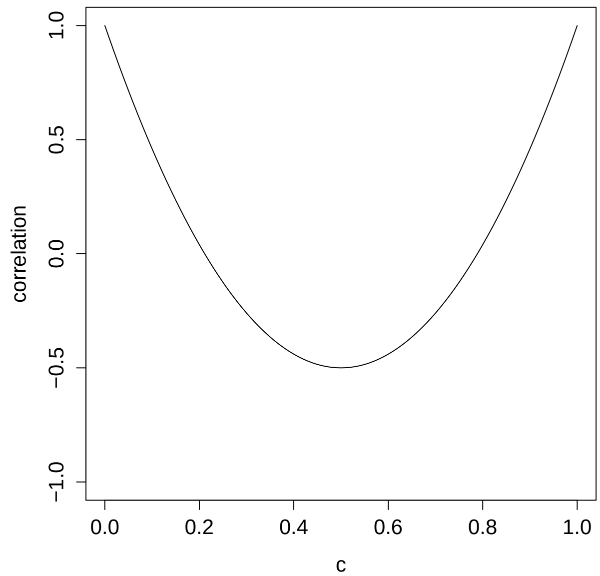

Figure 3.4: Plot of $\mathrm{Corr}(X_i, X_{i+1})$ as a function of $c \in \Omega$ for Example 3.1.

From now on in the examples we will consider ergodic roation with $c = \frac{\pi}{4}$, hence the resulting sequence $X_1, X_2, \ldots$ is ergodic. $\square$

3.2 Limit Theorems for Randomly Modulated Ergodic Stationary Sequences with and without Resampling

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be a d-dimensional ergodic stationary sequence. Assume that there is a measurable function Φ such that $E(\Phi^2(\mathbf{X}_i)) < \infty$, $i = 1, 2, \ldots$. Let ξ_i be iid $N(0,1)$, $i = 1, 2, \ldots$. Let us denote $Z_i = \Phi(\mathbf{X}_i)$. It is natural to say that the sequence

$$\xi_1 Z_1, \xi_2 Z_2, \ldots$$

is obtained by *random modulation* of the sequence $Z_1, Z_2, \ldots$ using the *modulating* sequence $\xi_1, \xi_2, \ldots$.

The first ergodic theorems for (non-random) modulating functions were proved in [15].

See also [14]. Later the ergodic theorems for randomly modulated stationary sequences were studied in [11]. This work is devoted to the study of the limit theorems for such sequences.

The following example motivates our study of limit theorems for randomly modulated ergodic stationary sequences.

Example 3.2. Suppose the sequence is constructed as in Example 3.1. Let us note that the marginal distribution of X_i is Uniform$[0,1)$, $i = 1, 2, \ldots$. Suppose we center the X_i's and want to test the null hypothesis $H_0 : \mu = 0$. Therefore it seems to be natural to consider the statistic $\bar{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ and to use the Central Limit Theorem (CLT). But the CLT does not work here! In other words, if $\mu_2^{(0)}$ denotes the second moment under the null hypothesis, the statistic $\frac{\sqrt{n}\bar{X}}{\sqrt{\mu_2^{(0)}}}$ doesn't converge to $N(0,1)$ in distribution. Moreover, our simulations show that as $n \to \infty$,

$$\mathrm{Var}(\sqrt{n}\bar{X}_n) \to 0.$$

The behavior of $\mathrm{Var}(\sqrt{n}\bar{X}_n)$ as the sample size n increases is illustrated in Figure 3.5.

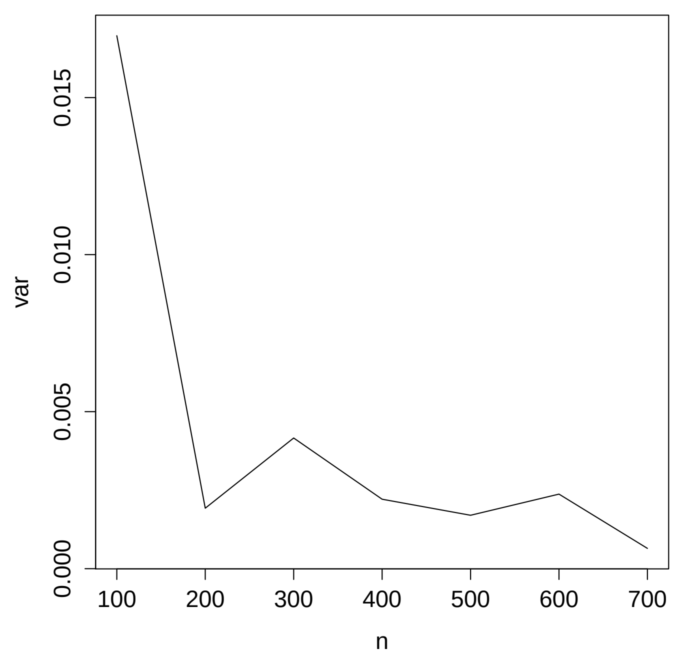

Figure 3.5: Plot of $\mathrm{Var}(\sqrt{n}\bar{X}_n)$ as a function of n for simulated data in Example 3.2.

So we have $\sqrt{n}\bar{X}_n \xrightarrow{\mathcal{P}} 0$, and thus $\sqrt{n}\bar{X}_n \xrightarrow{\mathcal{D}} 0$, as $n \to \infty$. In other words, instead of the convergence of $\sqrt{n}\bar{X}_n$ to Normal in distribution, it converges to the degenerate distribution at 0. Therefore the statistic $\bar{X}_n$ cannot be used to test the hypothesis. $\square$

Now let us consider the limit theorems for sequences of random vectors. We start with the general case when we do not assume ergodicity and stationarity of the sequence. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be a sequence of d-dimensional vectors, and $\Phi^{(j)} : \mathbb{R}^d \mapsto \mathbb{R}$, $j = 1, \ldots, s$, $s \geq 1$, be measurable functions such that

$$\mu_2^{(j,k)} = E\left(\Phi^{(j)}(\mathbf{X}_i)\Phi^{(k)}(\mathbf{X}_i)\right) \tag{3.1}$$

exists for every $j = 1, \ldots, s$, $k = 1, \ldots, s$, and

$$\frac{1}{n}\sum_{i=1}^{n}\left(\Phi^{(j)}(\mathbf{X}_i)\Phi^{(k)}(\mathbf{X}_i)\right) \xrightarrow{\mathcal{P}} \mu_2^{(j,k)}, \quad \text{as} \;\; \mathrm{n} \to \infty. \tag{3.2}$$

When $j = k$, we also denote

$$\mu_2^{(j)} = \mu_2^{(j,j)}.$$

Let us consider the $s \times s$ matrix

$$\Sigma_{\mathbf{\Phi}} = \begin{pmatrix} \mu_2^{(1)} & \ldots & \mu_2^{(1,s)} \\ & \vdots & \\ \mu_2^{(s,1)} & \ldots & \mu_2^{(s)} \end{pmatrix}$$

and the $rs \times rs$ matrix

$$\Sigma_{\mathbf{\Phi}}^{[r]} = \mathrm{diag}\left(\Sigma_{\mathbf{\Phi}}, \ldots, \Sigma_{\mathbf{\Phi}}\right),$$

that is

$$\Sigma_{\mathbf{\Phi}}^{[r]} = \begin{pmatrix} \Sigma_{\mathbf{\Phi}} & 0 & \ldots & 0 \\ 0 & \Sigma_{\mathbf{\Phi}} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \Sigma_{\mathbf{\Phi}} \end{pmatrix}$$

Lemma 3.1. *The matrices $\Sigma_{\mathbf{\Phi}}$ and $\Sigma_{\mathbf{\Phi}}^{[r]}$ are positive definite.*

Proof. Following the steps of the proof of Lemma 2.2, we conclude that $\Sigma_{\mathbf{\Phi}}$ and $\Sigma_{\mathbf{\Phi}}^{[r]}$ are positive definite. □

Theorem 3.1. *[17] Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be a sequence of d-dimensional vectors, and $\Phi^{(j)} : \mathbb{R}^d \mapsto \mathbb{R}$, $j = 1, \ldots, s$, $s \geq 1$, be measurable functions such that conditions (3.1) and (3.2) are satisfied. Let $\xi_i^{(m)}$, $i = 1, \ldots, n$, $m = 1, \ldots, r$, $r \geq 1$, be iid $N(0, 1)$. For every $j = 1, \ldots, s$ and $m = 1, \ldots, r$ we consider*

$$\Gamma_n^{(m,j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(m)} \Phi^{(j)}(\mathbf{X}_i),$$

and denote

$$\mathbf{\Gamma}_n^{(m)} = \left(\Gamma_n^{(m,1)}, \ldots, \Gamma_n^{(m,s)}\right)$$

for $m = 1, \ldots, r$, and

$$\mathbf{\Psi}_n = \left(\Gamma_n^{(1)}, \ldots, \Gamma_n^{(r)}\right).$$

Then

$$\mathbf{\Psi}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N\left(0, \Sigma_{\mathbf{\Phi}}^{[r]}\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$.

Proof. Let us consider random vectors $Z^{(n,1)}, \ldots, Z^{(n,s)} \in \mathbb{R}^n$ such that

$$Z_i^{(n,j)} = \Phi^{(j)}(\mathbf{X}_i),$$

$i = 1, \ldots, n$, $j = 1, \ldots, s$, and denote

$$\xi^{(n,m)} = (\xi_1^{(m)}, \ldots, \xi_n^{(m)})^T,$$

$m = 1, \ldots, r$. In these notations $\Gamma_n^{(m,j)}$ can be rewritten as

$$\Gamma_n^{(m,j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(n,m)} Z_i^{(n,j)}.$$

And therefore by Theorem 2.7,

$$\mathbf{\Psi}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N\left(0, \Sigma_{\mathbf{\Phi}}^{[r]}\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$. $\square$

Theorem 3.1 above presents the limit result for a general sequence of vectors. It turns out that if the sequence is ergodic stationary, we can drop the assumption (3.2). Suppose $\mathbf{X}_1, \mathbf{X}_2, \ldots$ is an ergodic stationary sequence of d-dimensional vectors. If the measurable functions $\Phi^{(j)} : \mathbb{R}^d \mapsto \mathbb{R}$, $j = 1, \ldots, s$, $s \geq 1$, are such that $\mu_2^{(j,k)}$ exists for any $j = 1, \ldots, s$ and $k = 1, \ldots, s$, that is

$$E\left|\Phi^{(j)}(\mathbf{X}_1)\Phi^{(k)}(\mathbf{X}_1)\right| < \infty, \tag{3.3}$$

then by Corollary 1.2, with probability 1

$$\frac{1}{n}\sum_{i=1}^{n}\left(\Phi^{(j)}(\mathbf{X}_i)\Phi^{(k)}(\mathbf{X}_i)\right) \to \mu_2^{(j,k)},$$

as $n \to \infty$, for any $j = 1, \ldots, s$, $k = 1, \ldots, s$. That is, comparing to the case of a general sequence, for an ergodic stationary sequence conditions (3.2) are automatically satisfied. Moreover, the convergence in this case is stronger than in (3.2).

Let us show the following result.

Theorem 3.2. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an ergodic stationary sequence of d-dimensional vectors. Let $\Phi^{(1)}, \ldots, \Phi^{(s)}$ be measurable functions such that the condition (3.3) is satisfied for every $j = 1, \ldots, s$ and $k = 1, \ldots, s$. Let $\xi_i^{(m)}$, $i = 1, \ldots, n$, $m = 1, \ldots, r$, $r \geq 1$, be iid $N(0,1)$. For every $j = 1, \ldots, s$ and $m = 1, \ldots, r$ we consider*

$$\Gamma_n^{(m,j)} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i^{(m)}\Phi^{(j)}(\mathbf{X}_i),$$

and denote

$$\mathbf{\Gamma}_n^{(m)} = \left(\Gamma_n^{(m,1)}, \ldots, \Gamma_n^{(m,s)}\right)$$

for $m = 1, \ldots, r$, *and*

$$\boldsymbol{\Psi}_n = \left(\Gamma_n^{(1)}, \ldots, \Gamma_n^{(r)}\right).$$

Then

$$\boldsymbol{\Psi}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N\left(0, \Sigma_{\boldsymbol{\Phi}}^{[r]}\right),$$

as $n \to \infty$, *with respect to the joint probability measure* $P_X \times P_\xi$.

Proof. Let us consider the following random vectors $\mathbf{Z}_i^{(n,j)} \in \mathbb{R}^n$ such that

$$\mathbf{Z}_i^{(n,j)} = \Phi^{(j)}(\mathbf{X}_i),$$

for $j = 1, \ldots, s$, $i = 1, \ldots, n$. We denote

$$\xi^{(n,m)} = (\xi_1^m, \ldots, \xi_n^m)^T,$$

for $m = 1, \ldots, r$. In this notations

$$\begin{aligned}
\Gamma_n^{m,j} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i^{(m)} \Phi^{(j)}(\mathbf{X}_i) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i^{(n,m)} \mathbf{Z}_i^{(n,j)},
\end{aligned}$$

for every $j = 1, \ldots, s$ and $m = 1, \ldots, r$. Then by Theorem 2.7,

$$\boldsymbol{\Psi}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N\left(0, \Sigma_{\boldsymbol{\Phi}}^{[r]}\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$. □

In the following corollaries we consider special cases that may be useful for applications. First suppose that $r = 1$, that is no resampling is made. In this situation we have the following corollary of the Theorem 3.2 above.

Corollary 3.1. *Let* $\mathbf{X}_1, \mathbf{X}_2, \ldots$ *be an ergodic stationary sequence of* d*-dimensional vectors. Let* $\Phi^{(1)}, \ldots, \Phi^{(s)}$ *be measurable functions such that the condition (3.3) is satisfied for every*

$j = 1, \ldots, s$ and $k = 1, \ldots, s$. Let ξ_i, $i = 1, \ldots, n$, be iid $N(0,1)$. For every $j = 1, \ldots, s$ we consider

$$\Gamma_n^{(j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i \Phi^{(j)}(\mathbf{X}_i),$$

and denote

$$\mathbf{\Psi}_n = \left(\Gamma_n^{(1)}, \ldots, \Gamma_n^{(s)} \right).$$

Then

$$\mathbf{\Psi}_n \quad \stackrel{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N\left(0, \Sigma_{\mathbf{\Phi}}\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$.

Now suppose we consider $r \geq 1$ independent resamplings, but one measurable function $\Phi(\mathbf{X})$, that is $s = 1$.

Corollary 3.2. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an ergodic stationary sequence of d-dimensional vectors. Let $\Phi(\mathbf{X})$ be a measurable function such that the condition (3.3) is satisfied for $j = k = 1$. Let $\xi_i^{(m)}$, $i = 1, \ldots, n$, $m = 1, \ldots, r$, $r \geq 1$, be iid $N(0,1)$. For every $m = 1, \ldots, r$ we consider*

$$\Gamma_n^{(m)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(m)} \Phi(\mathbf{X}_i),$$

and denote

$$\mathbf{\Psi}_n = \left(\Gamma_n^{(1)}, \ldots, \Gamma_n^{(r)} \right).$$

Then

$$\mathbf{\Psi}_n \quad \stackrel{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N\left(0, \mu_2^{(1)} I_r\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$.

If we consider one measurable function $\Phi(\mathbf{X})$ without any resampling, that is when both $s = 1$ and $r = 1$, then the following corollary is true.

Corollary 3.3. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an ergodic stationary sequence of d-dimensional vectors. Let $\Phi(\mathbf{X})$ be a measurable function such that the condition (3.3) is satisfied for $j = k = 1$.*

Let ξ_i, $i = 1, \ldots, n$, be iid $N(0,1)$. For every $m = 1, \ldots, r$ we consider

$$\Psi_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i \Phi(\mathbf{X}_i).$$

Then

$$\Psi_n \quad \overset{\mathcal{D}}{\to} \quad Y \sim N\left(0, \mu_2^{(1)}\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$.

Another situation of interest arises when no resampling is made, that is $r = 1$, and we consider $s = d$ measurable functions $\Phi^{(1)}(\mathbf{X}), \ldots, \Phi^{(d)}(\mathbf{X})$, where each $\Phi^{(j)}(\mathbf{X})$ is defined as the j-th coordinate of $\mathbf{X}$, that is for $j = 1, \ldots, d$

$$\Phi^{(j)}(\mathbf{X}) = X^{(j)}.$$

The following result follows from Theorem 3.2 as well.

Corollary 3.4. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an ergodic stationary sequence of d-dimensional vectors. Suppose that the condition (3.3) is satisfied for every $j = 1, \ldots, d$ and $k = 1, \ldots, d$. Let ξ_i, $i = 1, \ldots, n$, be iid $N(0,1)$. For every $j = 1, \ldots, d$ we consider*

$$\Psi_n^{(j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i X_i^{(j)},$$

and denote

$$\boldsymbol{\Psi}_n = \left(\Psi_n^{(1)}, \ldots, \Psi_n^{(d)}\right).$$

Then

$$\boldsymbol{\Psi}_n \quad \overset{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N\left(0, \Sigma\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$, where the elements of the matrix Σ are

$$\Sigma_{j,k} = E\left(X_1^{(j)} X_1^{(k)}\right), \quad j,\, k = 1, \ldots, d.$$

Let us consider Theorem 3.2 in a special situation when $X_1, X_2, \ldots$ is a one-dimensional ergodic stationary sequence, that is $d = 1$, and let

$$\Phi^{(j)}(X_i) = X_i^j$$

for $j = 1, \ldots, s$, $s \geq 1$. In this case the elements of matrix Σ become the respective moments as follows:

$$\mu_2^{(j)} \;=\; E(\Phi^{(j)}(X_1))^2 \;=\; E(X_1^{2j}) \;=\; \mu_{2j}$$

and similarly

$$\mu_2^{(j,k)} \;=\; E(\Phi^{(j)}(X_1)\Phi^{(k)}(X_1)) \;=\; E(X_1^{j+k}) \;=\; \mu_{j+k}$$

for $j = 1, \ldots, s$, $k = 1, \ldots, s$. Let us state the following corollary from the theorem.

Corollary 3.5. *Let $X_1, X_2, \ldots$ be an ergodic stationary sequence. Suppose that*

$$\mu_2^{(j,k)} = E(X_1^{j+k})$$

exists for every $j = 1, \ldots, s$ and $k = 1, \ldots, s$. Let $\xi_i^{(m)}$, $i = 1, \ldots, n$, $m = 1, \ldots, r$, $r \geq 1$, be iid $N(0,1)$. For every $j = 1, \ldots, s$ and $m = 1, \ldots, r$ we consider

$$\Gamma_n^{(m,j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(m)} X_i^j,$$

and denote

$$\mathbf{\Gamma}_n^{(m)} = \left(\Gamma_n^{(m,1)}, \ldots, \Gamma_n^{(m,s)}\right)$$

for $m = 1, \ldots, r$, and

$$\mathbf{\Psi}_n = \left(\Gamma_n^{(1)}, \ldots, \Gamma_n^{(r)}\right).$$

Then

$$\mathbf{\Psi}_n \;\stackrel{\mathcal{D}}{\to}\; \mathbf{Y} \sim N\left(0, \Sigma^{[r]}\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$, where the elements of the matrix Σ are

$$\Sigma_{j,k} = \mu_{j+k},$$

$j = 1, \dots, s$, $k = 1, \dots, s$ and

$$\Sigma^{[r]} = diag\,(\Sigma, \dots, \Sigma) = \begin{pmatrix} \Sigma & 0 & \dots & 0 \\ 0 & \Sigma & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \Sigma \end{pmatrix}.$$

Let us also consider the situation where an independent resampling is performed for each random variable $\Phi^{(j)}(\mathbf{X})$, $j = 1, \dots, s$. Then we can state the following corollary of Theorem 2.8.

Theorem 3.3. *Let $\mathbf{X}_1, \mathbf{X}_2, \dots$ be an ergodic stationary sequence of d-dimensional vectors. Let $\Phi^{(1)}, \dots, \Phi^{(s)}$ be measurable functions such that the condition (3.3) is satisfied for every $j = k = 1, \dots, s$. Let $\xi_i^{(j)}$ be iid $N(0,1)$, $i = 1, \dots, n$, $j = 1, \dots, s$. For every $j = 1, \dots, s$ we consider*

$$\Psi_n^{(j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(j)} \Phi^{(j)}(\mathbf{X}_i),$$

and denote

$$\mathbf{\Psi}_n = \left(\Psi_n^{(1)}, \dots, \Psi_n^{(s)}\right).$$

Then

$$\mathbf{\Psi}_n \quad \overset{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N\left(0, diag\,(\Sigma_{\mathbf{\Phi}})\right),$$

as $n \to \infty$, with respect to the joint probability measure $P_X \times P_\xi$, where

$$diag\left(\Sigma_{\mathbf{\Phi}}\right) = \begin{pmatrix} \mu_2^{(1)} & 0 & \dots & 0 \\ 0 & \mu_2^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_2^{(s)} \end{pmatrix}.$$

3.3 Estimation of the Rate of Convergence of CDF

Let $X_1, X_2, \dots$ be a sequence of random variables. We denote $\mathbf{X}^{(n)} = (X_1, \dots, X_n)$, $n = 1, 2, \dots$. Let $F_{P_X \times P_\xi}(z)$ be the cumulative distribution function of $\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{n}}$ with respect to the joint probability measure $P_X \times P_\xi$. In Theorem 2.6 we proved that $F_{P_X \times P_\xi}(z)$ converges to a normal cumulative distribution function uniformly in L_ξ^2. Now let us estimate the rate of this convergence.

Theorem 3.4. *Let $X_1, X_2, \dots$ be a sequence of random variables such that*

$$\mu_2 = E\left(X_1^2\right) < \infty$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2\right) \xrightarrow{\mathcal{P}} \mu_2,$$

as $n \to \infty$. Let $\xi^{(n)} \sim N(0, I_n)$. Then

$$\left|F_{P_X \times P_\xi}(z) - F_{N(0,\mu_2)}(z)\right| \leq$$

$$\leq min\left\{\frac{1}{\sqrt{2\pi}}E_X\left|\frac{\sqrt{n}}{||\mathbf{X}^{(n)}||} - \frac{1}{\sqrt{\mu_2}}\right||z|\,,\ \frac{\sqrt{2}}{\sqrt{\pi}e}E_X\left|\frac{||\mathbf{X}^{(n)}||}{\sqrt{n}} - \sqrt{\mu_2}\right|\frac{1}{|z|}\right\}.$$

Proof. Let us consider

$$E_\xi(F_n(z|\xi)) = \int_{-\infty}^{\infty} P\left(\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{n}} \leq z \;|\xi\right) dF_\xi$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_{\left\{\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{n}} \leq z\right\}} dF_X dF_\xi.$$

By Fubini's Theorem

$$E_\xi(F_n(z|\xi)) \quad = \quad \int\int I_{\left\{\frac{\xi^{(n)T}\mathbf{X}^{(n)}}{\sqrt{n}} \leq z\right\}} dF_X \times F_\xi = F_{P_X \times P_\xi}(z).$$

We showed above that

$$E_\xi(F_n(z|\xi)) = E_X\left(\int_{-\infty}^{\frac{\sqrt{n\mu_2}z}{||\mathbf{X}^{(n)}||}} \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{y^2}{2\mu_2}} dy\right).$$

Therefore

$$\begin{aligned}
&\left|F_{P_X \times P_\xi}(z) - F_{N(0,\mu_2)}(z)\right| \\
&= \left|E_X\left(\int_{-\infty}^{\frac{\sqrt{n\mu_2}z}{||\mathbf{X}^{(n)}||}} \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{y^2}{2\mu_2}} dy\right) - E_X\left(\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{y^2}{2\mu_2}} dy\right)\right| \\
&= \left|E_X\left(\int_{z}^{\frac{\sqrt{n\mu_2}z}{||\mathbf{X}^{(n)}||}} \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{y^2}{2\mu_2}} dy\right)\right|.
\end{aligned}$$

Since $e^{-\frac{y^2}{2\mu_2}}$ is bounded by 1, using the Lagrange Mean Value Theorem, we have

$$\left| \int_{z}^{\frac{\sqrt{n\mu_2}z}{||\mathbf{X}^{(n)}||}} \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{y^2}{2\mu_2}} dy \right| \leq \frac{1}{\sqrt{2\pi\mu_2}} \left| \frac{\sqrt{n\mu_2}z}{||\mathbf{X}^{(n)}||} - z \right|$$
$$= \frac{1}{\sqrt{2\pi}} \left| \frac{\sqrt{n}}{||\mathbf{X}^{(n)}||} - \frac{1}{\sqrt{\mu_2}} \right| |z| .$$

Therefore

$$\left| F_{P_X \times P_\xi}(z) - F_{N(0,\mu_2)}(z) \right| \leq E_X \left| \int_{z}^{\frac{\sqrt{n\mu_2}z}{||\mathbf{X}^{(n)}||}} \frac{1}{\sqrt{2\pi\mu_2}} e^{-\frac{y^2}{2\mu_2}} dy \right|$$
$$\leq \frac{1}{\sqrt{2\pi}} E_X \left| \frac{\sqrt{n}}{||\mathbf{X}^{(n)}||} - \frac{1}{\sqrt{\mu_2}} \right| |z| . \tag{3.4}$$

Naturally this upper bound is of most interest when $|z|$ is small. However, evaluating $E_X \left| \frac{\sqrt{n}}{||\mathbf{X}^{(n)}||} - \frac{1}{\sqrt{\mu_2}} \right|$ may be complicated. Therefore, when $|z|$ is not small, we consider $v = \frac{\sqrt{\mu_2}}{y}$. Then if $z \neq 0$,

$$\left| F_{P_X \times P_\xi}(z) - F_{N(0,\mu_2)}(z) \right| = \left| E_X \left[\int_{\frac{\sqrt{\mu_2}}{z}}^{\frac{||\mathbf{X}^{(n)}||}{z\sqrt{n}}} \frac{1}{\sqrt{2\pi}} \frac{1}{v^2} e^{-\frac{1}{2v^2}} dv \right] \right| .$$

Since the function inside the integral is bounded

$$\frac{1}{\sqrt{2\pi}} \frac{1}{v^2} e^{-\frac{1}{2v^2}} \leq \frac{1}{\sqrt{2\pi}} \frac{2}{e} ,$$

we use the Lagrange Mean Value Theorem to obtain the following result

$$\left| F_{P_X \times P_\xi}(z) - F_{N(0,\mu_2)}(z) \right| \leq \frac{\sqrt{2}}{\sqrt{\pi} e} E_X \left| \frac{||\mathbf{X}^{(n)}||}{\sqrt{n}} - \sqrt{\mu_2} \right| \frac{1}{|z|} . \tag{3.5}$$

We combine (3.4) and (3.5) to obtain

$$\left|F_{P_X \times P_\xi}(z) - F_{N(0,\mu_2)}(z)\right| \leq$$

$$\leq \min\left\{ \frac{1}{\sqrt{2\pi}} E_X \left| \frac{\sqrt{n}}{||\mathbf{X}^{(n)}||} - \frac{1}{\sqrt{\mu_2}} \right| |z| \, , \; \frac{\sqrt{2}}{\sqrt{\pi} e} E_X \left| \frac{||\mathbf{X}^{(n)}||}{\sqrt{n}} - \sqrt{\mu_2} \right| \frac{1}{|z|} \right\}.$$

This completes the proof. $\square$

Chapter 4

Testing Hypotheses for Ergodic Stationary Random Sequences

Let us now consider how the results of the previous chapter can be applied to construction of confidence intervals for parameters and to testing different types of hypotheses for ergodic stationary sequences. First we will discuss the main ideas of our approach.

4.1 Randomly Modulated Statistics: General Approach

Recall that we denote the distribution of the d-dimensional ergodic stationary sequence $\mathbf{X}_1, \mathbf{X}_2, \ldots$ in $\mathbb{R}^\infty$ by P_X. Consider the null hypothesis $H_0 : P_X = P_X^{(0)}$. Let $\Phi^{(j)} : \mathbb{R}^d \to \mathbb{R}$ be measurable functions such that

$$\mu_2^{(j,k)(0)} = E_{P_X^{(0)}} \left(\Phi^{(j)}(\mathbf{X}_1) \Phi^{(k)}(\mathbf{X}_1) \right)$$

exists for any $j = 1, \ldots, s$, $k = 1, \ldots, s$. Denote

$$\Sigma^{(0)} = \begin{pmatrix} \mu_2^{(1,1)(0)} & \ldots & \mu_2^{(1,s)(0)} \\ & \vdots & \\ \mu_2^{(s,1)(0)} & \ldots & \mu_2^{(s,s)(0)} \end{pmatrix}.$$

and

$$\Sigma^{[r](0)} = \begin{pmatrix} \Sigma^{(0)} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \Sigma^{(0)} \end{pmatrix}. \tag{4.1}$$

By Lemma 3.1 the matrices $\Sigma^{(0)}$ and $\Sigma^{[r](0)}$ are positive definite.

Let $\xi_i^{(m)}$, $i = 1, \dots, n$, $m = 1, \dots, r$, be iid $N(0,1)$. We consider $(\xi_i^{(1)}, \dots, \xi_i^{(r)})$, $i = 1, 2, \dots$ as a sequence of r-dimensional vectors. We denote the space of these vectors by Ω_ξ with the usual σ-algebra $\mathfrak{F}_\xi$. Let P_ξ denote the measure generated by the sequence $(\xi_i^{(1)}, \dots, \xi_i^{(r)})$ on $\mathfrak{F}_\xi$. We replace the hypothesis H_0 about the measure $P_X^{(0)}$ by the hypothesis $\tilde{H}_0 : P_X \times P_\xi = P_X^{(0)} \times P_\xi$.

We call each sequence $\xi_i^{(m)}, \xi_2^{(m)}, \dots$ a *random modulating sequence* for each $m = 1, \dots, r$. We also say that we have r *resamplings* of the modulating sequence. We use $\xi_1^{(m)}, \dots, \xi_n^{(m)}$ to randomly modulate $\Phi^{(j)}(\mathbf{X}_1), \dots, \Phi^{(j)}(\mathbf{X}_n)$, $j = 1, \dots, s$, as we discussed in Chapter 3. So we consider *randomly modulated statistics*

$$RM_n^{(m,j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(m)} \Phi^{(j)}(\mathbf{X}_i),$$

$j = 1, \dots, s$, $m = 1, \dots, r$,

$$\mathbf{RM}_n^{(m)} = \left(RM_n^{(m,1)}, \dots, RM_n^{(m,s)} \right),$$

and denote

$$\mathbf{RM}_n = \left(\mathbf{RM}_n^{(1)}, \dots, \mathbf{RM}_n^{(r)} \right).$$

Let us consider a simple alternative hypothesis H_a that the ergodic stationary sequence $\mathbf{X}_1, \mathbf{X}_2, \dots$ has the distribution $P_X^{(a)}$. As before we also consider the "extended" hypothesis $\tilde{H}_a : P_X \times P_\xi = P_X^{(a)} \times P_\xi$. It is clear that $\tilde{H}_0$ ($\tilde{H}_a$) is true iff H_0 (H_a, respectively) is true.

Assume that

$$\mu_2^{(j,k)(a)} = E_{P_X^{(a)}} \left(\Phi^{(j)}(\mathbf{X}_1) \Phi^{(k)}(\mathbf{X}_1) \right)$$

exists for any $j = 1, \ldots, s$, $k = 1, \ldots, s$. Let us consider the matrix

$$\Sigma^{(a)} = \begin{pmatrix} \mu_2^{(1,1)(a)} & \ldots & \mu_2^{(1,s)(a)} \\ & \vdots & \\ \mu_2^{(s,1)(a)} & \ldots & \mu_2^{(s,s)(a)} \end{pmatrix}$$

and

$$\Sigma^{[r](a)} = \begin{pmatrix} \Sigma^{(a)} & \ldots & 0 \\ 0 & \ddots & 0 \\ 0 & \ldots & \Sigma^{(a)} \end{pmatrix}.$$

Then we have the following result.

Theorem 4.1. *1.*

$$\mathbf{RM}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N(0, \Sigma^{[r](0)}), \quad \text{as} \quad \text{n} \to \infty,$$

with respect to the probability measure $P_X^{(0)} \times P_\xi$.

2.

$$\mathbf{RM}_n \quad \xrightarrow{\mathcal{D}} \quad \mathbf{Y} \sim N(0, \Sigma^{[r](a)}), \quad \text{as} \quad \text{n} \to \infty,$$

with respect to the probability measure $P_X^{(a)} \times P_\xi$.

Proof. These results follow from Theorem 3.3. □

Let us consider the following statistic

$$V_n^{(r,s)} = \left(\Sigma^{[r](0)}\right)^{-1/2} \mathbf{RM}_n^T \mathbf{RM}_n \left(\Sigma^{[r](0)}\right)^{-1/2}$$

Then from Theorem 4.1 we have

$$V_n^{(r,s)} \quad \xrightarrow{\mathcal{D}} \quad \chi^2_{rs}, \quad \text{as} \quad \text{n} \to \infty.$$

The asymptotic acceptance region is therefore

$$V_n^{(r,s)} < \chi^2_{rs,\alpha}.$$

Also by Theorem 4.1 the asymptotic power of the test of size α is

$$P\left(\chi^2_{rs} \quad > \quad \left(\Sigma^{[r](a)}\right)^{1/2}\left(\Sigma^{[r](0)}\right)^{-1}\left(\Sigma^{[r](a)}\right)^{1/2}\chi^2_{rs,\alpha}\right).$$

Let us now consider some special cases following the ideas of Chapter 3.

1. If $r = 1$, that is no resampling is performed, then the limiting variance matrices in Theorem 4.1 are $\Sigma^{(0)}$ and $\Sigma^{(a)}$, respectively.

2. If $s = 1$, then the limiting variance matrices in Theorem 4.1 are $\mu_2^{(1,1)(0)} I_r$ and $\mu_2^{(1,1)(a)} I_r$, respectively.

3. If $r = 1$ and $s = 1$, then the limiting variances in Theorem 4.1 are $\mu_2^{(1,1)(0)}$ and $\mu_2^{(1,1)(a)}$, respectively.

4. Suppose now we consider a special case when we include the higher moments up to order s for a one-dimensional ergodic stationary sequence $X_1, X_2, \ldots$ following the setting of Corollary 3.5. That is, now $\Phi^{(j)}(X_i) = X_i^j$, $j = 1, \ldots, s$. We suppose that under the H_0−distribution $P_X^{(0)}$

$$\mu^{(0)}_{(j+k)} = E_{P_X^{(0)}}(X_1^{j+k})$$

exists for every $j = 1, \ldots, s$ and $k = 1, \ldots, s$ and consider matrices

$$\Sigma^{(0)}_{j,k} = \mu^{(0)}_{j+k},$$

$j = 1, \ldots, s$, $k = 1, \ldots, s$ and $\Sigma^{[r](0)}$ is defined as in (4.1). In a similar manner we introduce $\mu^{(a)}_{(j+k)}$, $\Sigma^{(a)}$, and $\Sigma^{[r](a)}$ under the H_a−distribution of X, $P_X^{(a)}$. Then the limiting variance matrices in Theorem 4.1 for this special case are these matrices $\Sigma^{[r](0)}$ and $\Sigma^{[r](a)}$, respectively.

Remark 4.1. The second moment μ_2 contains information about the distribution of X, and hence the above statistic can be used to test different hypotheses about this distribution. However in many cases we can construct statistics when μ_2 is directly related to a parameter of interest or in general to the hypothesis. For example, if we consider the a-th moment

$E(X^a)$, where a is an odd natural number, then we consider

$$\Phi(X_i) = X_i^{a/2}.$$

The case of $a = 1$ will be discussed in Section 4.3.

In the case of inference about a parameter θ we can use a direct approach. Namely instead of X_i we can consider random variables $Y_i = f(X_i)$, where function f is such that

$$E(f^2(X_i)) = \theta.$$

This allows us to construct confidence intervals for θ and test hypothesis.

Now following Theorem 3.3 we consider the situation with an independent resampling for each random variable $\Phi^{(j)}(\mathbf{X})$, $j = 1, \ldots, s$. Then we can state the following corollary of Theorem 2.8. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an ergodic stationary sequence of d-dimensional vectors. Suppose that under $P_X^{(0)}$

$$\mu_2^{(j)(0)} = E_{P_X^{(0)}} \left(\Phi^{(j)}(\mathbf{X}_1) \right)^2,$$

$j = 1, \ldots, s$, and matrix

$$\Sigma^{(0)} = \mathrm{diag}\left(\mu_2^{(1)(0)}, \ldots, \mu_2^{(s)(0)} \right)$$
$$= \begin{pmatrix} \mu_2^{(1)(0)} & 0 & \ldots & 0 \\ 0 & \mu_2^{(2)(0)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mu_2^{(s)(0)} \end{pmatrix}.$$

Similarly we introduce

$$\mu_2^{(j)(a)} = E_{P_X^{(a)}} \left(\Phi^{(j)}(\mathbf{X}_1) \right)^2,$$

$j = 1, \ldots, s$, and matrix

$$\Sigma^{(a)} = \mathrm{diag}\left(\mu_2^{(1)(a)}, \ldots, \mu_2^{(s)(a)} \right).$$

Then for

$$RM_n^{(j)} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i^{(j)} \Phi^{(j)}(\mathbf{X}_i)$$

and

$$\mathbf{RM}_n = \left(RM_n^{(1)}, \ldots, RM_n^{(s)} \right)$$

we can show the following corollary of Theorem 3.3.

Theorem 4.2. *1.*

$$\mathbf{RM}_n \quad \overset{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N(0, \Sigma^{(0)}), \quad \text{as} \quad \mathrm{n} \to \infty,$$

with respect to the probability measure $P_X^{(0)} \times P_\xi$.

2.

$$\mathbf{RM}_n \quad \overset{\mathcal{D}}{\to} \quad \mathbf{Y} \sim N(0, \Sigma^{(a)}), \quad \text{as} \quad \mathrm{n} \to \infty,$$

with respect to the probability measure $P_X^{(a)} \times P_\xi$.

Using the statistic

$$V_n^{(r,s)} = \left(\Sigma^{(0)}\right)^{-1/2} \mathbf{RM}_n^T \mathbf{RM}_n \left(\Sigma^{(0)}\right)^{-1/2},$$

by Theorem 4.2 the asymptotic acceptance region is

$$V_n^{(r,s)} < \chi^2_{rs,\alpha},$$

and the asymptotic power of the test of size α is

$$P\left(\chi^2_{rs} > \sum_{j=1}^{s} \frac{\mu_2^{(j)(a)}}{\mu_2^{(j)(0)}} \ \chi^2_{rs,\alpha} \right).$$

4.2 Examples: Testing Hypotheses about the Distribution

In this section we illustrate the results of the previous section. Let us note that any two hypotheses about invariant measures $P_X^{(0)}$ and $P_X^{(a)}$ in the sequence space are singular (see

Remark 1.2). Therefore the power of every "good" hypotheses about $P_X^{(0)}$ and $P_X^{(a)}$ should be close to 1. First we study a trivial case when we have two singular hypotheses even when we consider the marginal distribution.

Example 4.1. Let $d = 1$, that is $\mathbf{X}_i = X_i$, $i = 1, 2, \ldots$. We use the ergodic stationary sequence $X_1, X_2, \ldots$ constructed as described in Example 3.1 in the previous chapter. In this case $f(x) = x$, $x \in [0, 1)$. The marginal distribution of each X_i is Uniform$[0, 1)$. By Remark 1.1 the measure P_X is uniquely defined by the marginal distribution of X_1. Therefore we can state the hypotheses about P_X in terms of the distribution of X_1.

Let the null hypothesis be

$$H_0 : X_1 \sim \text{Uniform}[0, 1).$$

In fact, H_0 is a hypothesis about the measure $P_X^{(0)}$ in $\mathbb{R}^\infty$ generated by $X_1, X_2, \ldots$. We also consider another function $g(x) = x + 1$ and the respective measure $P_X^{(a)}$ in the space of the sequence $X_1, X_2, \ldots$, where $X_i = g\left(\tau^{i-1}\omega\right)$, $i = 1, 2, \ldots$, $\omega \in \Omega_X$. So the alternative hypothesis is

$$H_a : X_1 \sim \text{Uniform}[1, 2).$$

The hypothesis H_a is equivalent to the hypothesis about the measure $P_X^{(a)}$.

We can consider the following scenarios for studying the theoretical asymptotic power of the test of size $\alpha = 0.05$.

1. Let us first consider the behavior of the asymptotic power of the test as a function of the number of resamplings r and study the effect of higher moments of the elements of the sequence. Therefore we compare the behavior of the power in two situations: when the number of functions $\Phi(X)$ is $s = 1$ and $s = 5$. Suppose we consider functions of the form $\Phi^{(j)}(X) = X^j$ for $j = 1, \ldots, s$. Figure 4.1 below illustrates the behavior of the asymptotic power in these two cases. We can observe that in both situations when the number of resamplings r increases, the asymptotic power of the test increases as well. Remark 2.1 provides the intuitive explanation of the increase of the asymptotic power

with the introduction of resamplings. Moreover, let us note that with the consideration of more functions, that is for higher s, the asymptotic power starts off at a higher value and approaches 1 noticeably faster.

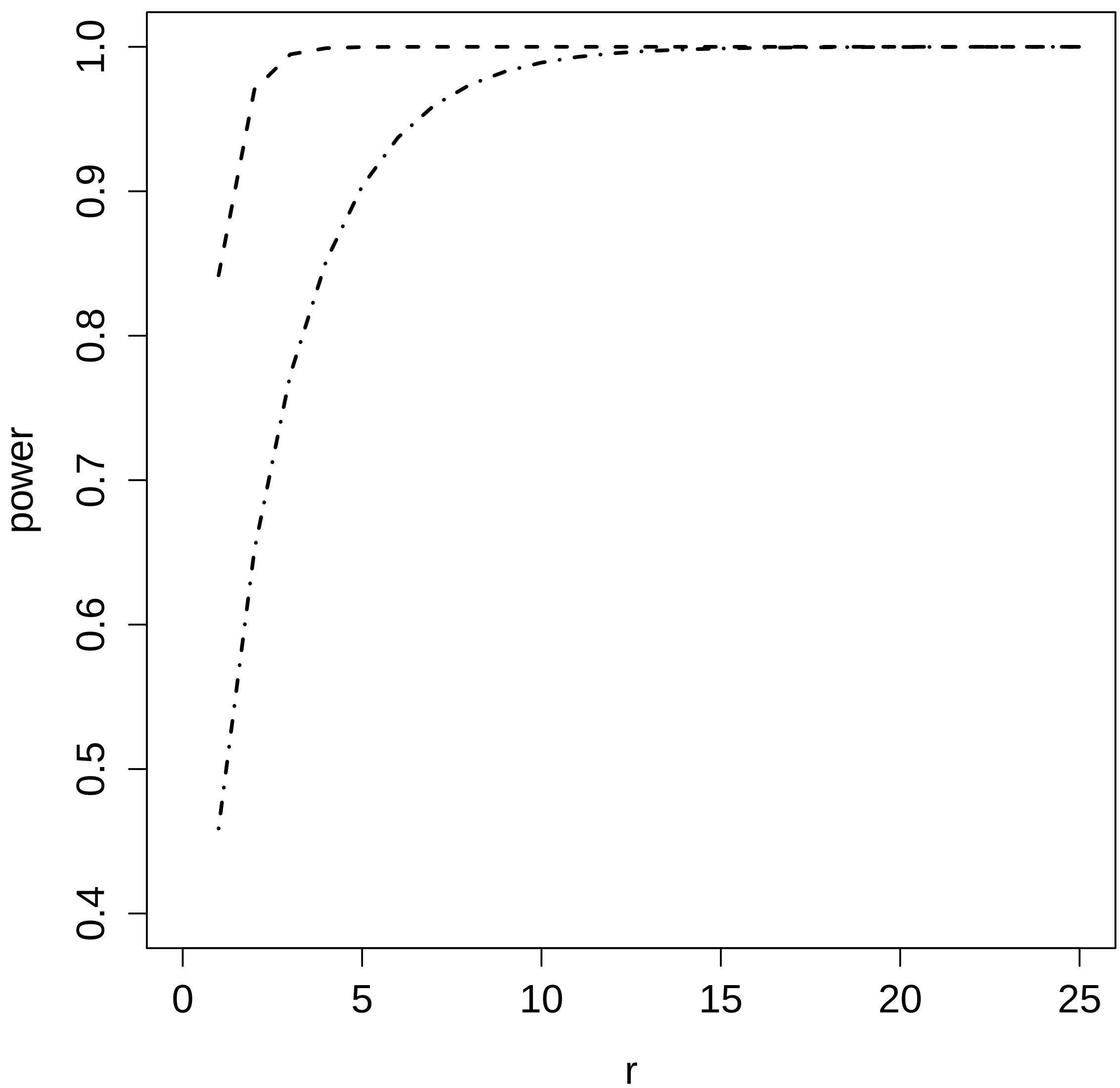

Figure 4.1: The asymptotic power of the test as a function of the number of resamplings r, when the number of moments is $s = 1$ (dotdash) and $s = 5$ (dashed) for Example 4.1.

2. In this scenario consider the behavior of the asymptotic power of the size $\alpha = 0.05$ test as a function of the number of moments s and study the effect of performing resamplings on the asymptotic power of the test. Again here we consider $\Phi^{(j)}(X) = X^j$, for $j = 1, \dots, s$. And compare the asymptotic power of the test for two situations: when there is no resampling, that is $r = 1$, and where there are $r = 5$ resamplings. As illustrated in Figure 4.2, the asymptotic power of the test approaches 1 reasonably fast in both situation. And as expected, if in addition to considering higher moments, we perform resamplings, then the asymptotic power of the test again starts off at a high value and increases to 1 considerably faster.

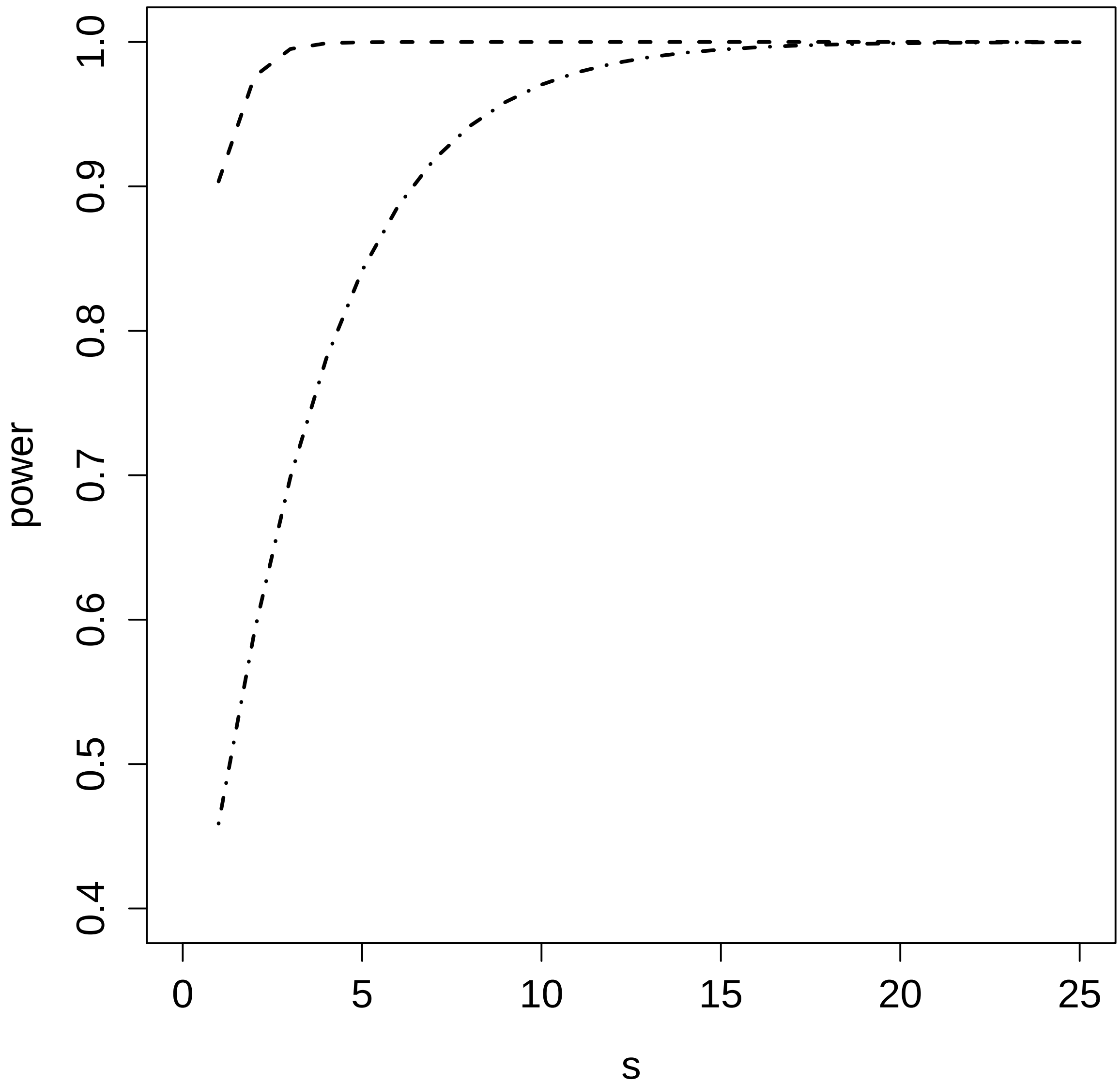

Figure 4.2: The asymptotic power of the test as a function of the number of higher moments s, when no resamplings are performed, that is $r = 1$, (dotdash) and when there are $r = 5$ resamplings (dashed) for Example 4.1.

$\square$

Example 4.2. Suppose now we want to test a different set of hypotheses. We will say that a random variable X has $\mathrm{U}^a[0,1)$ distribution, $0 < a \leq 1$, if $X^{1/a} \sim \mathrm{Uniform}[0,1)$. Similarly to the reasoning in Example 4.1 for $f(x) = x$ and $g(x) = x^a$, $a = \frac{1}{3}$, we consider $X_i = f\left(\tau^{i-1}\omega\right)$ and $X_i = g\left(\tau^{i-1}\omega\right)$, respectively, $i = 1, 2, \ldots$, $\omega \in \Omega_X$. We set the null and alternative hypotheses to be

$$H_0 : p_X(x) = 1, \quad x \in [0,1)$$

and

$$H_a : p_X(x) = 3x^2, \qquad x \in [0,1),$$

respectively, where $p_X(x)$ is the probability density function of X. Note that using the notations above the hypotheses can be equivalently rewritten as

$$H_0 : X \sim \mathrm{U}^1[0,1)$$

and

$$H_a : X \sim \mathrm{U}^{1/3}[0,1).$$

The power of the optimal likelihood ratio test for one observation is then 0.143.

We consider the two scenarios as above to study the behavior of the asymptotic power of the test. Figure 4.3 illustrates the effect of considering higher moments. We observe that the asymptotic power approaches 1 faster when we consider additional moments. And Figure 4.4 demonstrates the effect of performing resamplings. We can see that with more resamplings the asymptotic power of the test starts off at a higher value and approaches 1 considerably faster than when no resamplings are performed.

We can also see that the power of the test based on our statistics tends to 1 slower than in the case of singular marginal hypotheses in Example 4.1.

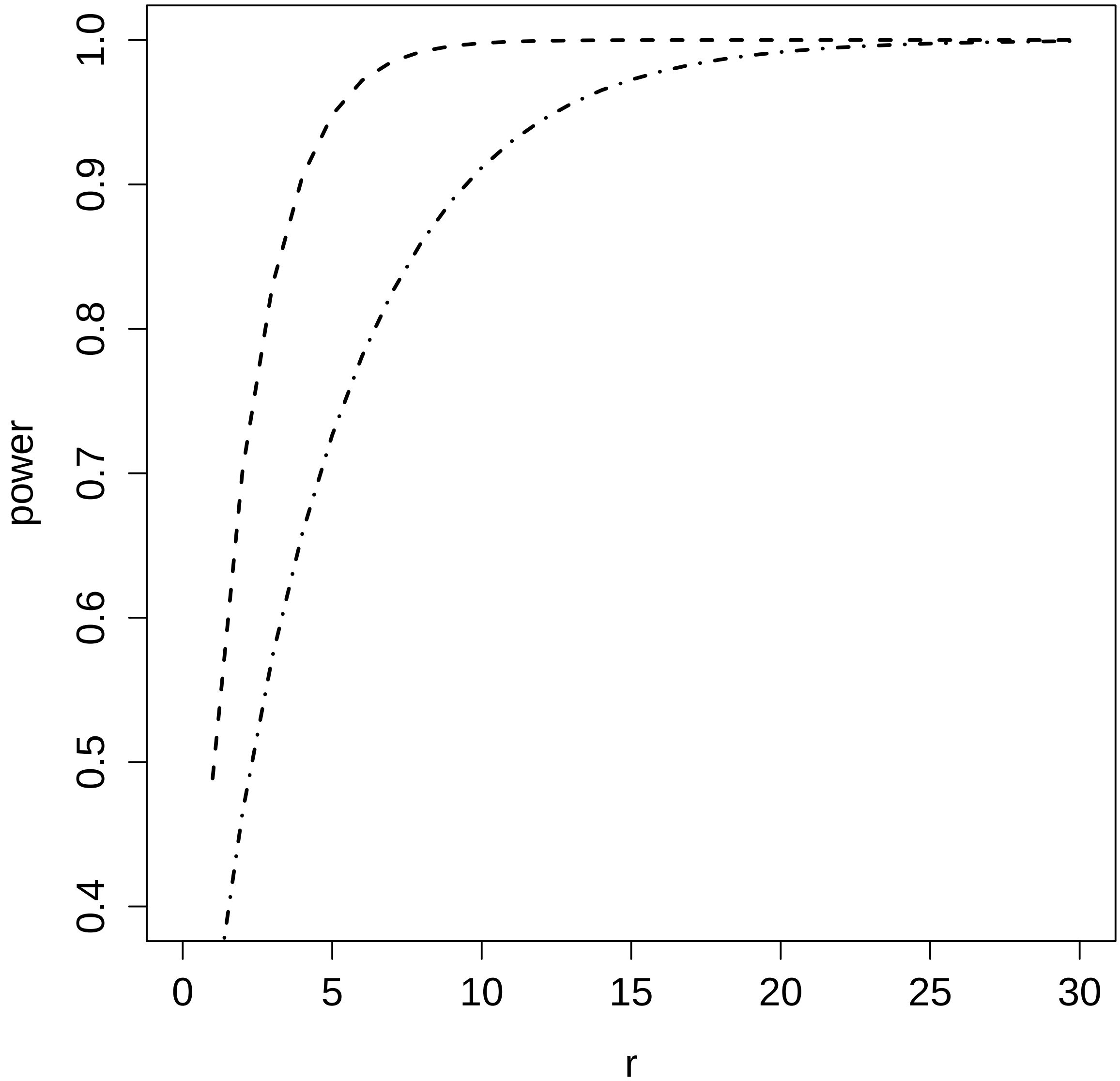

Figure 4.3: The asymptotic power of the test as a function of the number of resamplings r, when we do not consider higher moments, that is when $s = 1$, (dotdash) and when we consider moments up to order $s = 10$ (dashed) for Example 4.2.

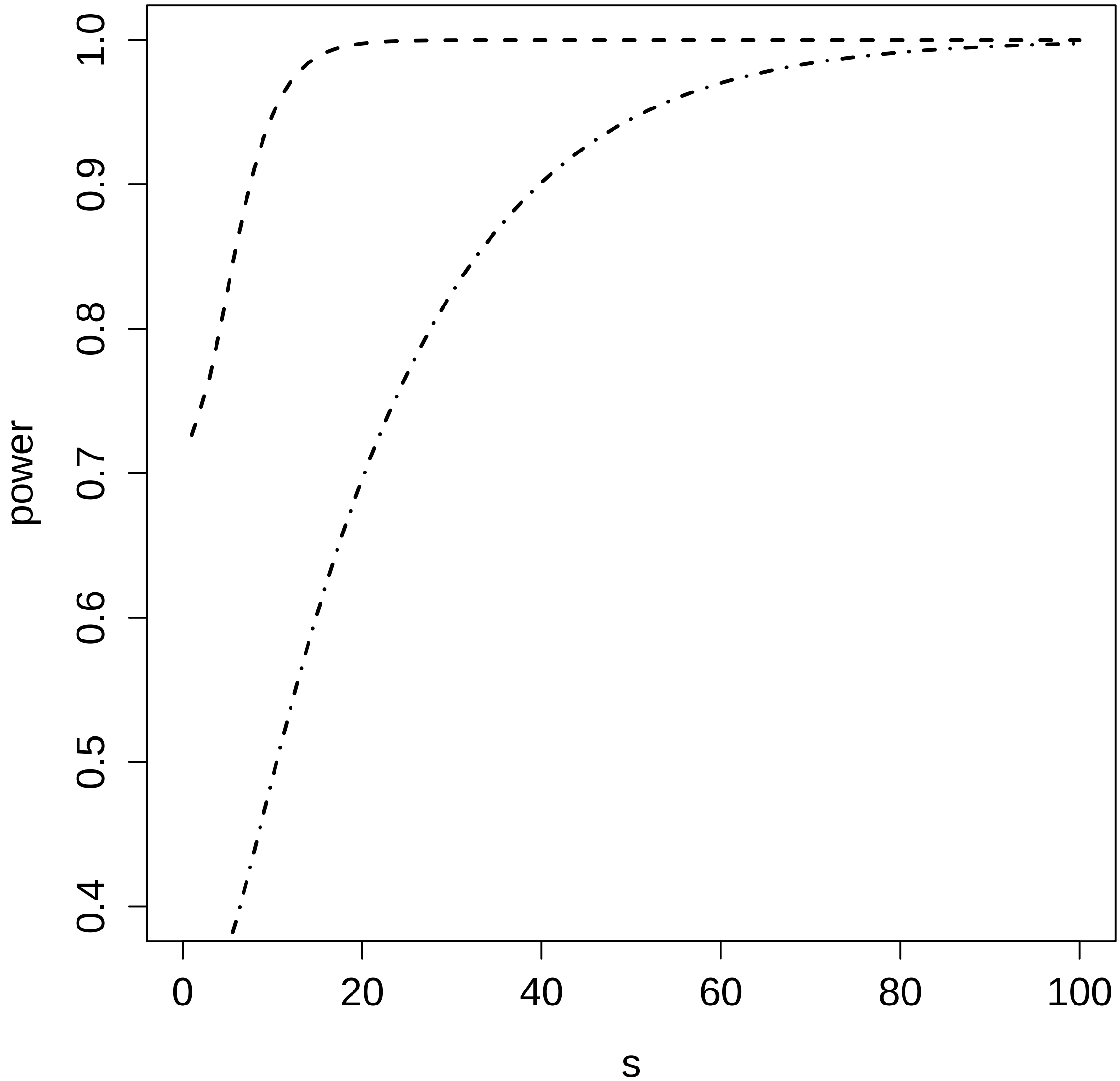

Figure 4.4: The asymptotic power of the test as a function of the number of higher moments s, when no resamplings are performed, that is when $r = 1$, (dotdash) and when the number of resamplings $r = 5$ (dashed) for Example 4.2.

$\square$

4.3 Confidence Intervals and Hypothesis Testing for the Mean: Direct Approach

Another interesting application of the theory presented in the previous chapter arises when we want to estimate the mean of the elements of an ergodic stationary sequence $X_1, X_2, \ldots$. Let us note that $\bar{X}$ may be considered being a common and consistent estimator. However we showed in Chapter 3 that the regular CLT does not hold and therefore we can not construct confidence intervals and hypothesis tests. So we introduce another estimator in this section using our randomly modulated statistics. It may seem that consideration of randomly modulated statistics removes the information about the mean completely since the limiting normal distribution does not contain it. However the variance of the limiting distribution is the second moment of the elements of the sequence and this allows us to create a method to estimate the mean.

Let us consider the case when we have a realization $X_1, \ldots, X_n$ of an ergodic stationary sequence such that $X_i \geq 0$ for any $i = 1, \ldots, n$. We denote the unknown mean $\mu_X = E(X_1)$ and the goal is to estimate it. Let ξ_i, $i = 1, \ldots, n$, be iid $N(0,1)$ and $\Phi(X) = \sqrt{X}$. We note that in this case $E(\Phi(X_1)^2) = E(X_1) = \mu_X$. Then by Corollary 3.3, for the ergodic stationary sequence $X_1, X_2, \ldots$ we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i \sqrt{X_i} \quad \stackrel{\mathcal{D}}{\to} \quad Y \sim N(0, \mu_X), \quad \text{as n} \to \infty.$$

Suppose we have N independent realizations of $X_1^{(j)}, \ldots, X_n^{(j)}$, $j = 1, \ldots, N$. We construct

$$RM_n^{(j)}(X) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i \sqrt{X_i^{(j)}},$$

$j = 1, \ldots, N$. Denote

$$\bar{RM}_n = \frac{1}{N} \sum_{j=1}^{N} RM_n^{(j)}(X).$$

To estimate μ_X we consider

$$\hat{\mu}_X = \frac{1}{N} \sum_{j=1}^{N} \left(RM_n^{(j)} - \bar{RM}_n \right)^2.$$

By Theorem 4.4, under $H_0 : \mu_X = \mu^{(0)}$, the statisic

$$\sum_{j=1}^{N} \frac{\left(RM_n^{(j)} - \bar{RM}_n \right)^2}{\mu^{(0)}} \overset{\mathcal{D}}{\to} \chi_N^2, \quad \text{as n} \to \infty.$$

Thus the approximate 95% acceptance region for this statistic is $\left(0, \chi_{N,\alpha}^2\right)$, and the approximate power of the test against a simple alternative $H_a : \mu_X = \mu^{(a)}$ is

$$P\left(\chi_N^2 > \frac{\mu^{(0)}}{\mu^{(a)}} \chi_{N,\alpha}^2 \right).$$

4.4 Testing Hypotheses about Independence

In this section we consider testing hypotheses about two or more ergodic stationary sequences being uncorrelated or independent. Suppose $m = 2$, then $\mathbf{X}_i$ can be expressed as (X_i, Y_i), $i = 1, 2, \ldots$ for jointly stationary and ergodic sequences $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ with $E(X_1) = \mu_X$ and $E(Y_1) = \mu_Y$. Let us consider

$$\Phi(X_i, Y_i) = (X_i - \mu_X)(Y_i - \mu_Y),$$

$i = 1, 2, \ldots$. Suppose we want to test $H_0 : X \perp\!\!\!\perp Y$, that is the sequences $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are mutually independent. Then under H_0,

$$E(\Phi(X_i, Y_i)) = \mathrm{Cov}(X_i, Y_i) = 0$$

and

$$E(\Phi^2(X_i, Y_i)) = E(X_i - \mu_X)^2 E(Y_i - \mu_Y)^2 = \sigma_X^2 \sigma_Y^2.$$

Therefore by Theorem 3.3,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i^{(n)}\Phi(X_i,Y_i) \quad \xrightarrow{\mathcal{D}} \quad N(0,\sigma_X^2\sigma_Y^2)$$

with respect to the joint probability measure $P_{X,Y}\times P_\xi$.

Furthermore if we test $H_0 : X \perp\!\!\!\perp Y$ against $H_a : X$ and Y are not independent, we can study the power of the test as described in Section 4.1.

Example 4.3. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be ergodic stationary sequences constructed independently as it is described in the beginning of Chapter 3. Then $\mu_X = \mu_Y = \frac{1}{2}$ and $\sigma_X^2 = \sigma_Y^2 = \frac{1}{12}$. Consider $\Phi(X_i,Y_i) = (X_i - \mu_X)(Y_i - \mu_Y)$. Suppose that we have N independent realizations of $(X_1,Y_1),\ldots,(X_n,Y_n)$ and want to test $H_0 : X \perp\!\!\!\perp Y$. For illustration we choose the simplest alternative $H_a : X = Y$. Then the asymptotic power of the test is close to 1 when we consider several dozens of resamplings as shown in Figure 4.5 below.

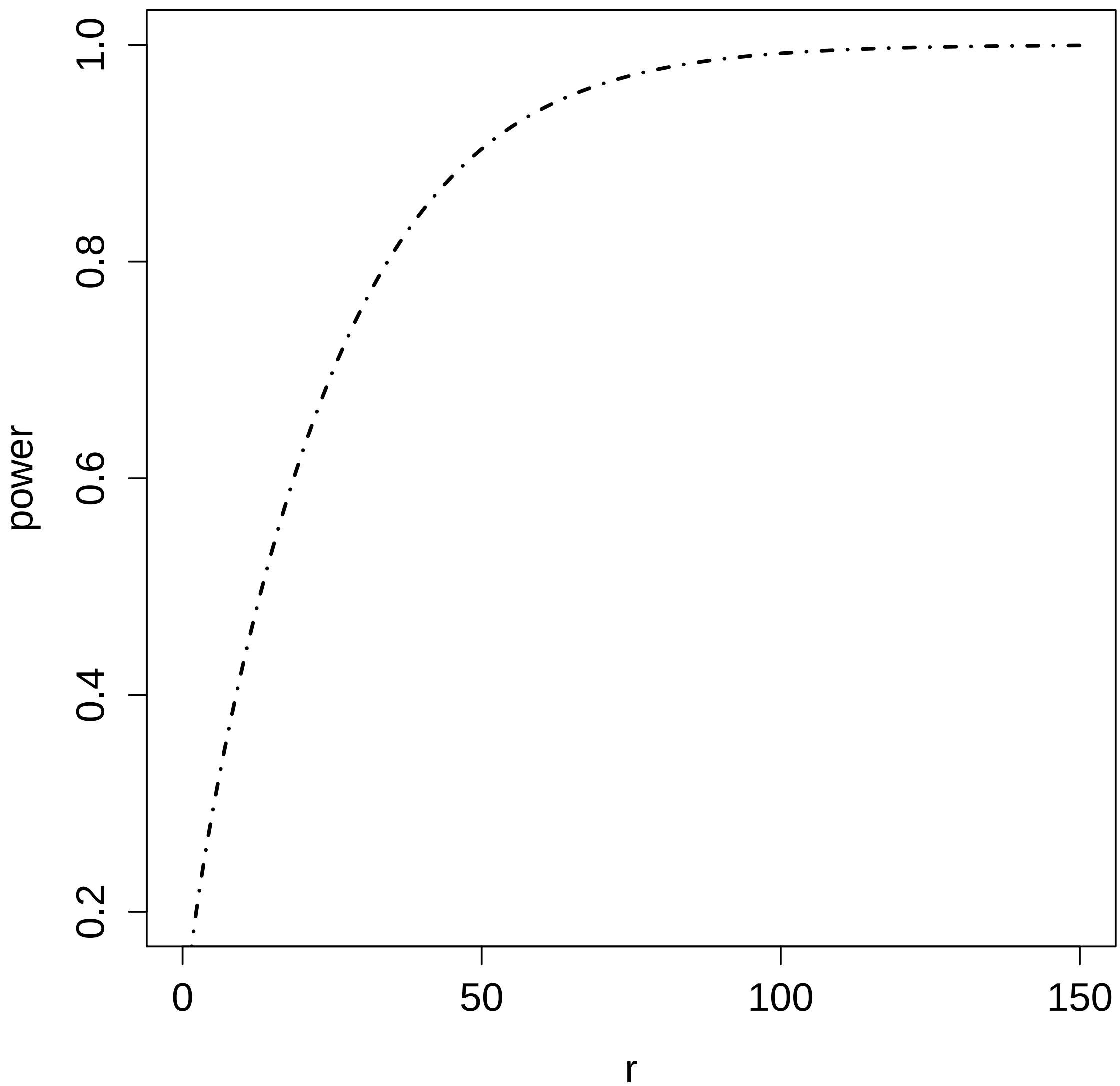

Figure 4.5: The asymptotic power of the test for independence as a function of the number of resamplings r for Example 4.3.

$\square$

Let us consider again the case when $m = 2$. Then $\mathbf{X}_i = (X_i, Y_i)$, $i = 1, 2, \ldots$ for jointly stationary and ergodic sequences $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$. Let us study how we can test for independence of the two ergodic stationary sequences using different functions $\Phi(X_k, Y_k)$. We consider

$$\Phi(X_k, Y_k) = \left(e^{tX_k} - E_X e^{tX_k}\right)\left(e^{sY_k} - E_Y e^{sY_k}\right),$$

$i = 1, 2, \ldots$. Let us note that $E_X e^{tX}$ is the moment generating function of X at t, that is $M_X(t)$, and similarly $E_Y e^{sY} = M_Y(s)$. We consider s and t such that the respective moment generating functions $M_X(t)$ and $M_Y(s)$ exist.

Next we consider

$$RM_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k \left(e^{tX_k} - E_Y e^{tX_k}\right)\left(e^{sY_k} - E_X e^{sY_k}\right).$$

Then under $H_0 : X \perp\!\!\!\perp Y$ we have

$$\begin{aligned} E\left[\Phi(X_k, Y_k)\right] &= E\left[\left(e^{tX_k} - E_X e^{tX_k}\right)\left(e^{sY_k} - E_Y e^{sY_k}\right)\right] \\ &= Ee^{tX_k + sY_k} - Ee^{tX_k} Ee^{sY_k} \ = \ 0 \end{aligned}$$

and

$$\begin{aligned} E\left[\Phi(X_k, Y_k)\right]^2 &= E\left[\left(e^{tX_k} - E_X e^{tX_k}\right)\left(e^{sY_k} - E_Y e^{sY_k}\right)\right]^2 \\ &= E\left(e^{tX_k} - E_X e^{tX_k}\right)^2 E\left(e^{sY_k} - E_Y e^{sY_k}\right)^2. \end{aligned}$$

We denote

$$\sigma^2_{e^{tX}} = Var_X(e^{tX_k}),$$

$$\sigma^2_{e^{sY}} = Var_Y(e^{sY_k}).$$

In these notations

$$E\left[\Phi(X_k, Y_k)\right]^2 = \sigma^2_{e^{tX}} \sigma^2_{e^{sY}}.$$

Then by Theorem 3.3, as $n \to \infty$, RM_n converges to $N\left(0, \sigma^2_{e^{tX}} \sigma^2_{e^{sY}}\right)$ in distribution with respect to the joint probability measure $P_{X,Y} \times P_\xi$. Therefore, if under H_a-distribution $E_{P^{(a)}_{X,Y}}\left(\Phi(X_k,Y_k)^2\right) = \mu_2^{(a)}$, then the power of the size α test can be found as

$$P\left(\chi^2_N > \frac{\sigma^2_{e^{tX}}\sigma^2_{e^{sY}}}{\mu_2^{(a)}}\ \chi^2_{N,\alpha}\right).$$

Furthermore, this test can be generalized for more than 2 sequences.

Chapter 5

Summary and Future Work

In the previous chapters we discussed the importance of developing the limit theory for stationary sequences and dynamical systems and the challenges that arise. Previous work in this area, though very interesting, is limited to results for *some* stationary sequences.

We introduced the method of random modulations of a sequence of random vectors that allowed us to use the information from these random projections of the sequence to construct its asymptotic conditional distribution [9]. In addition we proved the L^2-convergence of the conditional probability density functions of randomly modulated vectors and the uniform L^2-convergence of the conditional cumulative distribution functions.

We then described a modification of the method that allowed us to prove the convergence in distribution to normal (with respect to the product probability measure of $\mathbb{X}$ and ξ) for random modifications of *any* stationary sequence and therefore their respective dynamical systems; and we estimate the rate of convergence for the joint cumulative distribution functions of randomly modulated sequences [17]. The modified approach gives us an opportunity to do statistical inference, in particular to test hypothesis and construct confidence intervals.

And we illustrated how our results can be applied in statistics using examples for testing hypotheses about the distribution of the elements of the sequence, for inference about the mean, and for testing hypotheses about independence of sequences.

The work does not stop here. The following are the directions for future consideration:

1. generalization of the random modulation approach to random processes with continuous indices,
2. application of the considered approach to testing more types of hypotheses and construction of related confidence intervals,
3. application to general linear models.

Bibliography

[1] Burton, R. and Denker, M. (1987). On the Central Limit Theorem for Dynamical Systems. *Transactions of the American Mathematical Society*, Vol. 302, 2, pp. 715 - 726.

[2] Cornfeld, I. P., Fomin, S. V., and Sinai, Y. G. (1982). *Ergodic Theory*. Springer-Verlag, New York Heidelberg Berlin

[3] Derriennic, Y. (2006). Some Aspects of Recent Works on Limit Theorems in Ergodic Theory with Special Emphasis on the "Central LimitTheprem". *Discrete and Continuous Dynamical Systems*, Vol. 15, 1, pp. 143 - 158.

[4] Diaconis, P. and Freedman, D. (1984). Asymptotics of Graphical Projection Pursuit. *The Annals of Statistics*, 12, 3, pp. 793-815.

[5] Feller, W. (1967). *An Introduction to Probability Theory and Its Applications*. Vol. 1 (Third Edition). John Wiley & Sons.

[6] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. Vol. 2 (Second Edition). John Wiley & Sons.

[7] Halmos, P. R. (1956). *Lectures on Ergodic Theory*. Chelsea Publishing Company, New York, N.Y.

[8] Lacey, M. (1993). On Central Limit Theorems Modulus of Continuiuty and Diophantine Type for Irrational Rotations. *Journal d'Analyse Mathematique*, 61, pp. 47-59.

[9] Li, B. (2006). *Reading Diaconis and Freedman*. (Unpublished Notes).

[10] Li, B., Wen, S., and Zhu, L. (2008). On a Projective Resampling Method for Dimension Reduction with Multivariate Responses. *Journal of the American Statistical Association*, Vol. 103, 483, pp. 1177-1186.

[11] Lin, M. and Tempelman, A. (1999). Averaging sequences and modulated ergodic theorems for weakly almost periodic group representations. *Journal dAnalyse Mathmatique*, 77, pp. 237-268.

[12] Rosenblatt, M. (1962). *Random Processes.* Springer-Verlag, New York Heidelberg Berlin.

[13] Rozanov, Y. A. (1967). *Stationary Random Processes.* Holden-Day, San Francisco, Cambridge, London, Amsterdam.

[14] Ryll-Nardzewski, C. (1975). Topics in ergodic theory. *Proceedings of the Winter School in Probability, Karpacz, Poland*, *Springer Lecture Notes in Mathematics*, Vol. 472, pp. 131156. Springer, Berlin.

[15] Tempelman, A. A. (1974). Ergodic theorem for amplitude modulated homogeneous random fields. *Lithuanian Math. J.*, 14, pp. 221-229 (in Russian). English translation in *Lith. Math. Transl.*, 14 (1975), pp. 698 - 704.

[16] Tempelman, A. A. (1992). *Ergodic Theorems for Group Actions: Informational and Thermodynamical Aspects.* Kluwer Academic Publishers.

[17] Tempelman, A. A. (2014). *On Randomly Modulated Ergodic Stationary Sequences.* (Unpublished Notes).

Vita
Armine Bagyan

Education

- Ph.D. (Statistics), 2015
  The Pennsylvania State University
  Department of Statistics
  Advisors: Arkady A. Tempelman and Bing Li

- M.S. (Mathematics), summa cum laude, 2005
  Moscow State University
  Department of Mathematics and Mechanics
  Advisor: Boris M. Gurevich

Employment

- Teaching Assistant/Instructor, 2006 - 2015
  The Pennsylvania State University
  Department of Statistics

- Tutor (Mathematics), 2003 - 2005

Research Interests

- Limit Theorems

- Dimension Reduction