Theses and Dissertations

May 2014

A Limit Theorem for the Squared Norm of Empirical Distribution Functions

Alexander Nerlich
University of Wisconsin-Milwaukee

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Mathematics Commons

A limit theorem for the squared norm of empirical distribution functions

by

Alexander Nerlich

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Science
in
Mathematics

at

The University of Wisconsin-Milwaukee
May 2014

ABSTRACT

A limit theorem for the squared norm of empirical distribution functions

by

Alexander Nerlich

The University of Wisconsin-Milwaukee, 2014
Under the Supervision of Professor Eric Key

There are many limit theorems which involve empirical distribution functions. This thesis is dedicated to prove a limit theorem for the squared L^2-norm of two empirical distribution functions.

TABLE OF CONTENTS

List of Figures

ACKNOWLEDGEMENTS

I would like to thank Prof. Eric Key who helped me develop the results achieved in this thesis. Additionally I would like to thank Prof. Evgeny Spodarev who helped me prove the basic results on which this thesis is based.

Chapter 1

Introduction

The purpose of this thesis is to prove a new limit theorem which will be useful to decide whether two independent and identically distributed samples are identically distributed with each other. In order to achieve that aim we will consider a certain squared norm of empirical distribution functions.

To make this more precise, let $(X_1, ..., X_n)$ and $(Y_1, ..., Y_n)$ be two independently and identically distributed samples which are assumed to be independent of each other and let $\hat{F}_n^{(1)}$, respectively $\hat{F}_n^{(2)}$, denote the empirical distribution functions of $(X_1, ..., X_n)$, respectively $(Y_1, ..., Y_n)$. Finally let $F^{(1)}$, respectively $F^{(2)}$, be the distribution function of X_1, respectively Y_1. We introduce the estimator S_n defined by

$$S_n := \int_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z) \; \forall \; n \in \mathbb{N},$$

where ν is a finite Borel measure such that a set $A \in \mathfrak{B}(\mathbb{R})$ is a ν-nullset if and only if it is a Lebesgue nullset.
Our aim is to find out the limit in distribution of the sequence $(nS_n)_{n \in \mathbb{N}}$ under the null hypothesis that $(X_1, ..., X_n)$ and $(Y_1, ..., Y_n)$ are identically distributed with each other.

Of course the question arises why in particular the statistic S_n should be used to figure out whether the two samples are identically distributed with each other. The reason for that is that S_n has desirable consistency properties since it follows immediately from the theorems of Lebesgue and Glivenko-Cantelli that

$$\lim_{n\to\infty} S_n = \int_{\mathbb{R}} \left[F^{(1)}(z) - F^{(2)}(z) \right]^2 d\nu(z)$$

holds almost surely. Applying the theorem of Lebesgue again one verifies immediately that the above convergence holds in the L^p sense for all $p \in [1, \infty)$ as well. Hence the above limit is equal to zero if and only if $F^{(1)} = F^{(2)}$ ν-a.e. By the definition of ν the equality ν almost everywhere is equivalent to $F^{(1)} = F^{(2)}$ almost everywhere with respect to the Lebesgue measure. And since two distribution functions which are equal Lebesgue almost everywhere need to be equal everywhere (since they are right continuous) we obtain that

$$\lim_{n\to\infty} S_n = 0$$

holds a.s. and in the L^p sense for every $p \in [1, \infty)$ if and only if $F^{(1)} = F^{(2)}$.

It follows from that consistency property that if one has a concrete sample $(x_1, ..., x_n)$ (with $n \in \mathbb{N}$ sufficiently large) and if one knows to which family of distributions that sample belongs (for example exponential or infinitely divisible) one can simulate many concrete samples $(y_1, ..., y_n)$ such that S_n gets as close to zero as possible in order to find the unknown parameters which determine the distribution from which the sample $(x_1, ..., x_n)$ arises. Since usually the space of unknown parameters of the distribution from which $(x_1, ..., x_n)$ arises contains uncountably many elements and since the mapping from a sample $(y_1, ..., y_n)$ to the S_n depending on $x_1, ..., x_n, y_1, ..., y_n$ has no desirable properties such as continuity, differentiability, monotonicity, etc. it will usually not be possible to apply any optimization algorithm which directly calculates the parameters such that S_n is as close to zero as possible. Therefore it would be nice if one would have a hypothesis test for the unknown parameters of the distribution from which $(x_1, ..., x_n)$ arises.

Unfortunately we will not be able to determine the distribution of the limit of $(nS_n)_{n \in \mathbb{N}}$, but we will be able to determine the random variable to which $(nS_n)_{n \in \mathbb{N}}$ converges and we will be able to figure out some properties of that limit random variable which should be useful to determine the concrete distribution of that limit.

The proofs of our results will be relatively technical. They will mainly rely on a central limit theorem for random variables taking values in Hilbert spaces and on many functional analytical techniques. Therefore we will start this thesis by giving a brief introduction to functional analysis and we will continue by giving a brief introduction to random variables taking values in separable Hilbert spaces. All the theorems to which we refer there are already known results and we will state them without proof and will just refer to the literature mentioned in these chapters. Therefore the reader who is familiar with functional analysis and with the concept of random variables taking values in separable Hilbert spaces may wish to skip these chapters.

After the introduction we will start to prove our main results. We will show that

$$\lim_{n \to \infty} \int_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z) = \int_{\mathbb{R}} \left(G(\cdot) \right)(z)^2 d\nu(z)$$

holds in distribution, for a certain Hilbert-valued Gaussian random variable G which we will define later. Furthermore we will show that

$$\int_{\mathbb{R}} \left(G(\cdot) \right)(z)^2 d\nu(z) = \sum_{k=1}^{\infty} Z_k^2$$

holds almost surely, as well as in L^p for every $p \in [1, \infty)$, for a certain sequence $(Z_k)_{k \in \mathbb{N}}$ of centered, independent, but not identically distributed, Gaussian random variables. Furthermore we will show that the series considered above converges L^p absolutely for every $p \in [1, \infty)$. Therefore we will have shown that

$$\lim_{n \to \infty} \int_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z)$$

and

$$\sum_{k=1}^{\infty} Z_k^2$$

are equal in distribution.

After we obtain these theoretical results we will run some simulations to illustrate our results. Precisely we will simulate nQn, for large n, as well as $\sum_{k=1}^{m} Z_k^2$, for large m, and then we will compare the distributions of these two random variables. Since we will see that the variance of each Z_k is determined by an eigenvalue of a certain operator, we need to focus on distributions which have a "not too complicated" distribution function. Therefore we will run these simulations only for the Bernoulli and for the uniform distribution.

Finally we will give an outlook of further things we would like to prove and about concerns we have. Obviously the "further things we would like to prove" are how precisely the distribution of $\sum_{k=1}^{\infty} Z_k^2$ looks like and our main concern is whether that limit distribution is independent of $F^{(1)}$, $F^{(2)}$.

Chapter 2

Functional Analysis

As previously mentioned we need to establish some functional analytical theory in order to support the proofs of the results described in the Introduction. The aim of this chapter is to develop that theory.

In the general functional analytic setting one considers Banach spaces. An important special case of Banach spaces, are Hilbert spaces. Initially the theory of Hilbert and Banach spaces can seem relatively abstract therefore one should have the Lebesgue spaces $(L^p(\Omega, \Sigma, \mu), || \cdot ||_p)$ in mind which are indeed for every $p \in [1, \infty]$ and for every measure space (Ω, Σ, μ) examples of Banach spaces. These spaces are Hilbert spaces for any (Ω, Σ, μ) if $p = 2$.

There are many remarkable theorems about Banach and Hilbert spaces. One of the most remarkable theorems in Hilbert spaces is the Spectral Theorem which is important in many mathematical applications and which is, as we will see soon, very important to prove the results mentioned in the Introduction. Therefore we will start this chapter by giving the theory necessary for stating the spectral theorem. We will also develop the theory of trace class operators. Finally we will end this chapter by giving a brief introduction to the theorem of Riesz-Frechet.

All theorems stated in that Chapter, except for Theorem 2.1.22, are basic functional analytic results. These include the Spectral Theorem for self adjoint compact operators, the Parseval identity, the Theorem of Riesz Frechet and the Gram-Schmidt theorem. We will neither give any proof, nor references for the proofs of these the-

orem and simply refer the reader who is interested in these proofs to the standard literature on functional analysis.

2.1 Orthonormality, the Spectral Theorem and Trace Class Operators

Throughout this section $(H, ||.||)$ will denote a real Hilbert space and $< \cdot, \cdot >$ will denote its inner product. The results stated here can be extended to the complex case but since the real case suffices for our purposes we will only state that case.

Definition 2.1.1. $(H, ||.||)$ *is called separable if there exists a sequence* $(e_n)_{n \in \mathbb{N}} \subset H$ *of countably many elements such that the linear span of* $(e_n)_{n \in \mathbb{N}}$ *is dense in* H

Remark 2.1.2. *Throughout this chapter we assume that* H *is separable.*

Definition 2.1.3. *A sequence* $(e_n)_{n \in \mathbb{N}} \subset H$ *is called an orthonormal basis of* H *if all of the following conditions hold.*

  i) $< e_n, e_n >= 1$ *for all* $n \in \mathbb{N}$.

  ii) $< e_n, e_m >= 0$ *for all* $n, m \in \mathbb{N}$ *with* $n \neq m$.

  iii) The linear span of $(e_n)_{n \in \mathbb{N}}$ *is dense in* H.

Remark 2.1.4. *Since we assume that* H *is separable one can show that* H *has an orthonormal basis.*

Lemma 2.1.5. *Let* h_1 *be an element of* H. *It follows from*

$$< h_1, h >= 0 \ \forall \ h \in H$$

that $h_1 = 0$.

Theorem 2.1.6. *Let h_1, h_2 be elements of H. Then the inequality*

$$< h_1, h_2 >^2 \leq ||h_1||^2 ||h_2||^2 \tag{2.1}$$

holds.

Theorem 2.1.7. *The identity*

$$||u||^2 = \sum_{n=1}^{\infty} < u, e_n >^2 \tag{2.2}$$

holds for every $u \in H$.

Definition 2.1.8. *A sequence $(x_n)_{n \in \mathbb{N}} \subset H$ converges weakly to $x \in H$ if*

$$\lim_{n \to \infty} < x_n, h > = < x, h >$$

holds for every $h \in H$.

Notation 2.1.9. *The vector space of all linear continuous mappings $T : H \to H$ is denoted by $\mathfrak{L}(H)$.*

Definition 2.1.10. *An operator $T \in \mathfrak{L}(H)$ is said to be compact if for all weakly convergent sequences $(x_n)_{n \in \mathbb{N}} \subset H$ the sequence $(T(x_n))_{n \in \mathbb{N}}$ is convergent.*

Definition 2.1.11. *A linear mapping $T : H \to H$ is called a Hilbert-Schmidt operator if for every orthonormal basis $(e_n)_{n \in \mathbb{N}}$ of H*

$$\sum_{n=1}^{\infty} ||T e_n||^2 < \infty \tag{2.3}$$

holds.

Remark 2.1.12. *One can actually show that if (2.3) holds for one orthonormal basis it has to hold for every orthonormal basis and in that case the value of the series is independent of the chosen orthonormal basis.*

Theorem 2.1.13. *Every Hilbert-Schmidt operator is compact and every compact operator is continuous.*

Definition 2.1.14. *An operator $T \in \mathfrak{L}(H)$ is said to be self adjoint, if*

$$< Tx, y >=< x, Ty >$$

holds for all $x, y \in H$.

Definition 2.1.15. *An operator $T \in \mathfrak{L}(H)$ is said to be positive, if*

$$< Tx, x >\geq 0 \qquad (2.4)$$

holds for all $x \in H$.

Definition 2.1.16. *Let $T \in \mathfrak{L}(H)$ then every $\lambda \in \mathbb{R}$ for which*

$$Kern(T - \lambda Id) \neq \{0\}$$

holds is called an eigenvalue of T. In that case we call every $e \in Kern(T - \lambda Id)\backslash\{0\}$ an eigenvector of T for the eigenvalue λ.

Theorem 2.1.17. *Let $T \in \mathfrak{L}(H)$ be self adjoint and compact. Then there exists a sequence of eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$ and a corresponding sequence of eigenvectors $(e_n)_{n\in\mathbb{N}}$ such that $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of H and such that additionally*

$$Tx = \sum_{n=1}^{\infty} \lambda_n < x, e_n > e_n \qquad (2.5)$$

holds for every $x \in H$.

Remark 2.1.18. *In the above situation one verifies immediately that (2.5) implies that in particular*

$$Te_k = \lambda_k e_k$$

holds for all $k \in \mathbb{N}$.

Definition 2.1.19. *Let $T \in \mathfrak{L}(H)$ be a self adjoint and positive operator. Furthermore let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of H. Then we say that T is trace class if*

$$Tr(T) := \sum_{n=1}^{\infty} ||Te_n|| < \infty$$

holds. One verifies that $Tr(T)$ is independent of the choice of $(e_n)_{n\in\mathbb{N}}$. Furthermore we call $Tr(T)$ the Trace of T.

Remark 2.1.20. *One can extend the above definition to non-self adjoint operators, using the adjoint of an operator (see for example [1]). The above setting is enough for our purposes.*

Remark 2.1.21. *Usually it is hard to determine whether an operator is trace class. Therefore we will state the following theorem which can be found in [1]. This theorem will be very useful to prove that the operator we will consider in Chapter 4 is a trace class operator.*

Theorem 2.1.22. *Let S be a locally compact, second-countable, topological Hausdorff space endowed with a non-degenerate Borel measure μ which is σ-finite as well as locally finite. Let the mapping Q be defined by*

$$Q(u) := \int_S \psi(\cdot, s)u(s)d\mu(s)$$

for all $u \in L^2(S, \mathfrak{B}(S), \mu)$ where $\psi : \mathbb{R}^2 \to \mathbb{R}$ is a $\mu \times \mu$ almost everywhere continuous and $\mu \times \mu$ square integrable function such that $Qu \in L^2(S, \mathfrak{B}(S), \mu)$ holds for all $u \in L^2(S, \mathfrak{B}(S), \mu)$. If furthermore ψ is such that Q is a self-adjoint and positive operator in $L^2(S, \mathfrak{B}(S), \mu)$ and if the mapping

$$s \longmapsto \psi(s, s)$$

is μ-integrable then Q is trace class and

$$Tr(Q) = \int_S \psi(s, s)d\mu(s)$$

holds.

2.2 The Theorem of Riesz Frechet

Throughout this section $(H, ||\cdot||)$ will denote a real valued Hilbert space. In many functional analytic applications it is very important to know the structure of the dual space of H. The most important theorem about the dual of a Hilbert space is the Theorem of Riesz-Frechet, which will be stated in this section.

Definition 2.2.1. *The dual space of H is the vector space of all mappings $h^* : H \to \mathbb{R}$ which are linear and continuous. This space will be denoted by H^* and will be equipped with the norm*

$$||h^*|| := \sup_{||x|| \leq 1} |h(x)|$$

for every $h^* \in H^*$. Hereby we follow the common functional analytic convention to denote the norm of the Hilbert space as well as the norm of its dual by $||\cdot||$.

Theorem 2.2.2. *For any $h \in H$ let the mapping ψ_h be defined by*

$$\psi_h(x) = <x, h>$$

for all $x \in H$. Then the mapping $\Psi : H \to H^*$ defined by $\Psi(h) := \psi_h$ satisfies

i) Ψ is linear and bijective and

ii) $||\Psi(h)|| = ||h||$ holds for all $h \in H^*$.

Remark 2.2.3. *A mapping fulfilling the properties 2.2.2.i) and 2.2.2.ii) is called an isometric isomorphism. Therefore one says that a Hilbert space and its dual are isometrically isomorphic to one another. This theorem, known as the Theorem of Riesz Frechet, justifies that one can directly identify the Hilbert space with its dual. In the common functional analytic literature it is even not unusual to say that two spaces which are isometrically isomorphic to one another are equal.*

Chapter 3

Probability Theory in Hilbert Spaces

Now that we have established some functional analytic properties of Hilbert spaces we are in a position to give an introduction to probability theory in such spaces. The results and definitions mentioned in that chapter can be found in [2]. The results which can be found there are formulated for Banach spaces. Since we will only need the Hilbert space case, we will establish the theory only for Hilbert spaces.

Furthermore one will realize, considering [2], that in the Banach space setting one needs the dual space of the Banach space. Since we established that a Hilbert space and its dual are isometric isomorph in the previous chapter, we will follow the common functional analytic convention and will formulate all of the following definitions and theorems only for the Hilbert space and not for its dual. That means we directly identify the Hilbert space and its dual.

Throughout this chapter $(H, ||\cdot||)$ will denote a separable Hilbert space equipped with the inner product $< \cdot, \cdot >$. Furthermore $\mathfrak{B}(H)$ will denote the Borel sigma algebra of $(H, ||\cdot||)$, the smallest sigma algebra which contains all sets that are open with respect to $||\cdot||$. Analogously $\mathfrak{B}(\mathbb{R})$ (respectively $\mathfrak{B}(\mathbb{R}^n)$ for an $n \in \mathbb{N}$) denotes the Borel sigma algebra of $\mathbb{R}$ (respectively $\mathbb{R}^n$). Finally $(\Omega, \mathfrak{A}, P)$ will denote a probability space.

3.1 Hilbert space valued random variables and the central limit theorem

Definition 3.1.1. *A mapping $X : \Omega \to H$ which is $\mathfrak{B}(H) - \mathfrak{A}$ measurable is called Hilbert space valued random variable. The measure $P_X : \mathfrak{B}(H) \to [0,1]$ defined by*

$$P_X(B) = P(X \in B) \ \forall \ B \in \mathfrak{B}(H)$$

is called the distribution, or the law, of X.

Definition 3.1.2. *A set $K \subset H$ is called cylindrical if there exists an $n \in \mathbb{N}$, $h_1, ..., h_n \in H$, and a $D_n \in \mathfrak{B}(\mathbb{R}^n)$ such that*

$$K = \{x \in H : \ (< h_1, x >, ..., < h_n, x >) \in D_n\}$$

holds. Furthermore let $\mathcal{K}$ denote the set of all cylindrical sets and let $\mathfrak{C}(H)$ denote the smallest sigma algebra containing all $K \in \mathcal{K}$. We call $\mathfrak{C}(H)$ the cylindrical sigma algebra of H.

Theorem 3.1.3. *According to [2], Proposition 1.1 one can show that in every separable Hilbert space the cylindrical sigma algebra and the Borel sigma algebra agree, that is $\mathfrak{C}(H) = \mathfrak{B}(H)$.*

Remark 3.1.4. *Theorem 3.1.3 holds only because we assumed H to be separable. In general Hilbert (or Banach) spaces these definitions are not equivalent. Therefore this theorem shows that it is of integral importance that we assume H to be separable. One will realize that it is usually significantly easier to show measurability using the cylindrical sigma algebra than using the Borel sigma algebra. In particular the following corollary which we obtain from the above theorem will be very useful to show measurability.*

Corollary 3.1.5. *Let $X : \Omega \to H$ be any mapping such that for all $n \in \mathbb{N}$ and $h_1, ..., h_n \in H$ the mapping*

$$\omega \longmapsto (< h_1, X(\omega) >, ..., < h_n, X(\omega) >)$$

is $\mathfrak{A} - \mathfrak{B}(\mathbb{R}^n)$ measurable, then X is a Hilbert space valued random varibale.

Definition 3.1.6. Let $\mu : \mathfrak{B}(H) \to [0,1]$ be a probability measure. A set $A \in \mathfrak{B}(H)$ is called μ-continuous if the boundary of A is a μ-nullset.

Definition 3.1.7. Let $\mu_n, \mu : \mathfrak{B}(H) \to [0,1]$ be probability measures for all $n \in \mathbb{N}$. Then $(\mu_n)_{n \in \mathbb{N}}$ is said to converge in law to μ, if

$$\lim_{n \to \infty} \mu_n(A) = \mu(A)$$

holds, for all μ-continuous sets $A \in \mathfrak{B}(H)$. Analogously we say that a sequence of Hilbert space valued random variables $(X_n)_{n \in \mathbb{N}}$ converges in law to a random variable X, if $(P_{X_n})_{n \in \mathbb{N}}$ converges weakly to P_X.

Remark 3.1.8. The above definition is actually not how convergence in law is defined in [2]. But according to Theorem 3.2 in [2] it is equivalent to the definition given there. Furthermore one verifies immediately that this definition is equivalent to the usual convergence in distribution if $H = \mathbb{R}$.

Remark 3.1.9. One important property of convergence in law is that it fulfills the so called continuous mapping theorem, which can be found on Page 116 in [4] and which will be stated now. In [4] it is actually stated for metric spaces and we will state it only for Hilbert spaces.

Theorem 3.1.10. Let H_1 and H_2 be two separable Hilbert spaces and let $X_n : \Omega \to H_1$ be Hilbert space valued random variables for all $n \in \mathbb{N}$. Furthermore let $g : H_1 \to H_2$ be continuous and let $(X_n)_{n \in \mathbb{N}}$ converge in law to a Hilbert space valued random variable $X : \Omega \to H_1$. Then the sequence $(g(X_n))_{n \in \mathbb{N}}$ converges in law to $g(X)$.

Definition 3.1.11. Let $X : \Omega \to H$ be a Hilbert space valued random variable such that $\mathbb{E}||X|| < \infty$. If an $m \in H$ fulfills

$$< h, m > = \int_{\Omega} < h, X(\omega) > dP(\omega) \; \forall \; h \in H \tag{3.1}$$

then m is called a mean of X.

Remark 3.1.12. *According to page 3 in [2] the mean of a Hilbert space valued random variable always exists if $\mathbb{E}||X|| < \infty$. Furthermore one sees immediately that this mean is unique, since if X would have two means $m_1, m_2 \in H$ then it follows from (3.1) that $< h, m_1 - m_2 >= 0$ for all $h \in H$. Therefore it follows from Lemma 2.1.5 that $m_1 = m_2$.*

Definition 3.1.13. *Let $X : \Omega \to H$ be a Hilbert space valued random variable such that $\mathbb{E}||X|| < \infty$. Then we call X centered if its mean equals 0.*

Definition 3.1.14. *Let $X : \Omega \to H$ be a centered Hilbert space valued random variable such that $\mathbb{E}(||X||^2) < \infty$. If there exists a mapping $Q : H \to H$ such that*

$$< v, Qu >= \int_\Omega < v, X(\omega) >< u, X(\omega) > dP(\omega) \ \forall \ u, v \in H \qquad (3.2)$$

holds. Then we call Q a covariance operator of X.

Remark 3.1.15. *According to Page 3 in [2] the covariance operator exists and is unique if $\mathbb{E}(< h, X >^2) < \infty$ for all $h \in H$. From the Cauchy-Schwarz inequality (2.1) it follows immediately that $\mathbb{E}(||X||^2) < \infty$ implies that $\mathbb{E}(< h, X >^2) < \infty$ for all $h \in H$. Therefore the covariance operator exists and is unique if $\mathbb{E}(||X||^2) < \infty$. And since $\mathbb{E}(||X||^2) < \infty$ implies that $\mathbb{E}||X|| < \infty$ we get that $\mathbb{E}(||X||^2) < \infty$ implies the existence and uniqueness of the mean as well.*

Definition 3.1.16. *A Hilbert space valued random variable $X : \Omega \to H$ is said to be Gaussian if $< h, X >$ is a Gaussian real valued random variable for every $h \in H$.*

Remark 3.1.17. *In the above setting a real valued random variable which is constantly zero is said to be Gaussian (with mean and variance zero). If one did not make this definition no random variable would be Gaussian since for every Hilbert valued random variable $< h, X >= 0$ holds if $h = 0$.*

Theorem 3.1.18. *If $X : \Omega \to H$ is a Gaussian Hilbert space valued random variable then its mean exists. If additionally, X is centered then its covariance operator exists, and P_X is uniquely determined by its covariance operator.*

Remark 3.1.19. *The above result can be found in [2]. There it is actually stated for non centered Gaussian Hilbert space valued random variables. But since we defined the covariance operator only for centered Hilbert space valued random variables and since we will not consider any non centered Hilbert space valued random variables the above theorem is enough for our purposes.*

Definition 3.1.20. *Let $Z : \Omega \to H$ be a Hilbert space valued random variable and let $(Z_n)_{n \in \mathbb{N}}$ be a sequence of independent and identically distributed Hilbert space valued random variables such that Z and Z_n are independent and identically distributed for all $n \in \mathbb{N}$. Let $S_n := Z_1 + ... + Z_n$. We say that Z fulfills the central limit theorem if $(\frac{S_n}{\sqrt{n}})_{n \in \mathbb{N}}$ converges in law to a Hilbert space valued random variable G.*

Theorem 3.1.21. *If a Hilbert space valued random variable $Z : \Omega \to H$ fulfills the central limit theorem then the limit G is a centered Gaussian Hilbert space valued random variable, and G and Z have the same covariance operator*

Proof. See page 52 in [2]. □

Theorem 3.1.22. *If $Z : \Omega \to H$ is a Hilbert space valued random random variable such that*

  i) Z is centered,

  ii) $\mathbb{E}(||Z||^2) < \infty$

holds, then Z fulfills the central limit theorem.

Proof. The proof follows immediately from Corollary 8 in [3]. There it is actually stated that a space is isometrically isomorphic to a Hilbert space if and only if i) and ii) are necessary and sufficient for a random variable to fulfill the central limit theorem. □

Chapter 4

Main results

Now that we have developed some basic facts about Hilbert space valued random variables we are in the position to prove the results mentioned in the Introduction. Before we start we need to make some definitions and need to clarify some notation. Here and in what follows we denote by $\nu : \mathfrak{B}(\mathbb{R}) \to [0, \infty)$ a finite measure such that a set $N \in \mathfrak{B}(\mathbb{R})$ is a ν-nullset if and only if it is a λ-nullset where λ denotes Lebesgue measure.

One example one may wish to have in mind for a such a measure is the following: Let $\chi : \mathbb{R} \to (0, \infty)$ be a $L^1(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \lambda)$-integrable, positive and continuous function. Then the measure defined by

$$B \longmapsto \int_B \chi d\lambda \ \forall \ B \in \mathfrak{B}(\mathbb{R})$$

has obviously all properties of ν.

Using ν we introduce the Hilbert space $U := L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$. It is well known that U is a separable Hilbert space. Furthermore we want to remark that $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$ is by its definition a finite measure space. As usually we equip U with the $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$ norm which we will denote by $||\cdot||$. Furthermore we denote the $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$ inner product by $< \cdot, \cdot >$.

As above $(\Omega, \mathfrak{A}, P)$ will denote a probability space. Furthermore $(X_n)_{n \in \mathbb{N}_0}$, $(Y_n)_{n \in \mathbb{N}_0}$ will denote two independently and identically distributed sequences of real valued random variables which are assumed to be independently and identically distributed

with each other. For reasons of convenience we define $X := X_0$ and $Y := Y_0$. Finally we introduce the mapping $\zeta : \Omega \to U$ by

$$\zeta(\omega) = \mathbb{1}\{X(\omega) \leq \cdot\} - \mathbb{1}\{Y(\omega) \leq \cdot\} \tag{4.1}$$

for all $\omega \in \Omega$.

4.1 The limit of the squared norm of empirical distribution functions

The aim of this section is to show that the squared $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$ norm of empirical distribution functions converges in distribution to the integral of the square of a certain Gaussian Hilbert space valued random variable.

Lemma 4.1.1. *ζ fulfills all of the following properties.*

i) $\zeta(\omega) \in U$ for all $\omega \in \Omega$.

ii) ζ is a Hilbert space valued random variable. (That means it is $\mathfrak{B}(U) - \mathfrak{A}$ measurable.)

iii) $\mathbb{E}(||\zeta||^2) < \infty$.

Proof. i) $\zeta(\omega)$ is obviously $\mathfrak{B}(\mathbb{R})$-$\mathfrak{B}(\mathbb{R})$ measurable for all $\omega \in \Omega$ and furthermore $\zeta(\omega)^2$ is integrable since $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$ is a finite measure space and since $\zeta(\omega)^2 \leq 1$ holds for all $\omega \in \Omega$. Therefore $\zeta(\omega) \in U$ holds for all $\omega \in \Omega$.

ii) Let $n \in \mathbb{N}$ and $h_1, ..., h_n \in U$ be arbitrary. According to Corollary 3.1.5 it suffices to show that the mapping

$$\omega \longmapsto (< h_1, \zeta(\omega) >, ..., < h_n, \zeta(\omega) >) \tag{4.2}$$

is $\mathfrak{A} - \mathfrak{B}(\mathbb{R}^n)$ measurable. We will prove that the function defined by (4.2) is indeed measurable by using the fact that the composition of measurable functions is measurable. Therefore let us introduce the functions $f_1 : \Omega \to \mathbb{R}^2$ and $f_2 : \mathbb{R}^2 \to \mathbb{R}^n$ defined by

$$f_1(\omega) := (X(\omega), Y(\omega)) \; \forall \; \omega \in \Omega$$

and by

$$f_2(x,y) := (< h_1, \mathbb{1}\{x \leq \cdot\} - \mathbb{1}\{y \leq \cdot\} >, ..., < h_n, \mathbb{1}\{x \leq \cdot\} - \mathbb{1}\{y \leq \cdot\} >)$$

for all $x, y \in \mathbb{R}^2$. Then f_1 is $\mathfrak{A} - \mathfrak{B}(\mathbb{R}^2)$ measurable since X and Y are random variables. Furthermore f_2 is obviously $\mathfrak{B}(\mathbb{R}^2) - \mathfrak{B}(\mathbb{R}^n)$ measurable. Therefore $f_2 \circ f_1$ is $\mathfrak{A} - \mathfrak{B}(\mathbb{R}^n)$ measurable. And since

$$(f_2 \circ f_1)(\omega) = f_2(X(\omega), Y(\omega)) \stackrel{(4.1)}{=} (< h_1, \zeta(\omega) >, ..., < h_n, \zeta(\omega) >) \ \forall \ \omega \in \Omega$$

holds we get that the mapping defined in (4.2) is $\mathfrak{A} - \mathfrak{B}(\mathbb{R}^n)$ measurable which is (using Corollary 3.1.5) exactly the desired result.

iii). That $\mathbb{E}(||\zeta||^2) < \infty$ follows immediately from the following calculation.

$$\begin{aligned}
\mathbb{E}(||\zeta||^2) &= \int_\Omega \int_\mathbb{R} (\mathbb{1}\{X(\omega) \leq z\} - \mathbb{1}\{Y(\omega) \leq z\})^2 d\nu(z) dP(\omega) \\
&\leq \int_\Omega \int_\mathbb{R} 1 d\nu(z) dP(\omega) \\
&= \nu(\mathbb{R}) \\
&< \infty.
\end{aligned}$$

$\square$

Remark 4.1.2. *Since we have shown that $\mathbb{E}(||\zeta||^2) < \infty$ holds, we do know (according to Remark 3.1.15) that the mean of ζ exists. Since X and Y are identically distributed one assumes that ζ is centered which will turn out to be true. And if we know that ζ is centered we do know that the covariance operator exists as well. Therefore we will prove in the following theorem that ζ is indeed centered and furthermore we will determine the covariance operator of ζ.*

Theorem 4.1.3. *The Hilbert space valued random variable ζ is centered. Furthermore the covariance operator $Q : U \to U$ of ζ is given by*

$$Q(u) = 2 \int_\mathbb{R} \Big[\min(F(\cdot), F(s)) - F(\cdot)F(s) \Big] u(s) d\nu(s) \ \forall \ u \in U, \qquad (4.3)$$

where F denotes the distribution function of X.

Proof. At first we will prove that ζ is centered. According to the definition of the mean and due to Lemma 2.1.5 it suffices to show that

$$\int_\Omega <h, \zeta(\omega)> dP(\omega) = 0 \ \forall \ h \in U \tag{4.4}$$

holds. Let $h \in U$ be arbitrary then we obtain immediately from the following calculation that (4.4) indeed holds, where we will use the theorem of Fubini.

$$
\begin{aligned}
\int_\Omega <h, \zeta(\omega)> dP(\omega) &= \int_\Omega \int_\mathbb{R} h(z)(\zeta(\omega))(z)d\nu(z)dP(\omega) \\
&= \int_\mathbb{R} \int_\Omega h(z)(\zeta(\omega))(z)dP(\omega)d\nu(z) \\
&= \int_\mathbb{R} h(z)\mathbb{E}\big[(\zeta(\cdot))(z)\big]d\nu(z) \\
&= \int_\mathbb{R} h(z)\mathbb{E}\big[\mathbf{1}\{X \le z\} - \mathbf{1}\{Y \le z\}\big]d\nu(z) \\
&= \int_\mathbb{R} h(z)\Big(\mathbb{E}\big[\mathbf{1}\{X \le z\}\big] - \mathbb{E}\big[\mathbf{1}\{Y \le z\}\big]\Big)d\nu(z) \\
&= 0.
\end{aligned}
$$

Now let $u, v \in U$ be arbitrary. Using the theorem of Fubini again, (4.3) follows from the following calculation which will complete the proof.

$$
\begin{aligned}
<v, Qu> &= \int_\mathbb{R} v(z_1)(Qu)(z_1)d\nu(z_1) \\
&= \int_\mathbb{R} v(z_1)2\int_\mathbb{R} \Big[\min(F(z_1), F(z_2)) - F(z_1)F(z_2)\Big]u(z_2)d\nu(z_2)d\nu(z_1) \\
&= \int_\mathbb{R}\int_\mathbb{R} v(z_1)u(z_2)2\Big[\min(F(z_1), F(z_2)) - F(z_1)F(z_2)\Big]d\nu(z_2)d\nu(z_1) \\
&= \int_\mathbb{R}\int_\mathbb{R} v(z_1)u(z_2)2\Big[F(\min(z_1, z_2)) - F(z_1)F(z_2)\Big]d\nu(z_2)d\nu(z_1)
\end{aligned}
$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} v(z_1)u(z_2)\mathbb{E}\Big[\big(\mathbb{1}\{X \le z_1\} - \mathbb{1}\{Y \le z_1\}\big) \cdot ..$$

$$.. \ \big(\mathbb{1}\{X \le z_2\} - \mathbb{1}\{Y \le z_2\}\big)\Big]d\nu(z_2)d\nu(z_1)$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} v(z_1)u(z_2)\mathbb{E}\Big[(\zeta(\cdot))(z_1)(\zeta(\cdot))(z_2)\Big]d\nu(z_2)d\nu(z_1)$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{E}\Big[v(z_1)(\zeta(\cdot))(z_1)u(z_2)(\zeta(\cdot))(z_2)\Big]d\nu(z_2)d\nu(z_1)$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\Omega} v(z_1)(\zeta(\omega))(z_1)u(z_2)(\zeta(\omega))(z_2)dP(\omega)d\nu(z_2)d\nu(z_1)$$

$$= \int_{\Omega} \int_{\mathbb{R}} v(z_1)(\zeta(\omega))(z_1)d\nu(z_1) \int_{\mathbb{R}} u(z_2)(\zeta(\omega))(z_2)d\nu(z_2)dP(\omega)$$

$$= \int_{\Omega} < v, \zeta(\omega) >< u, \zeta(\omega) > dP(\omega)$$

□

Definition 4.1.4. *The Hilbert space valued random variables* $\hat{F}_n^{(1)}$, $\hat{F}_n^{(2)} : \Omega \to U$ *defined by*

$$\hat{F}_n^{(1)}(\omega) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}\{X_j(\omega) \le \cdot\} \ \forall \ \omega \in \Omega, \ n \in \mathbb{N} \tag{4.5}$$

and by

$$\hat{F}_n^{(2)}(\omega) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}\{Y_j(\omega) \le \cdot\} \ \forall \ \omega \in \Omega, \ n \in \mathbb{N} \tag{4.6}$$

are called the empirical distribution functions of $(X_1, ..., X_n)$, *respectively* $(Y_1, ..., Y_n)$. *Furthermore we denote by* $S_n : \Omega \to U$ *the Hilbert space valued random variable defined by*

$$S_n(\omega) = \sum_{j=1}^{n} \Big(\mathbb{1}\{X_j(\omega) \le \cdot\} - \mathbb{1}\{Y_j(\omega) \le \cdot\}\Big) \tag{4.7}$$

for all $\omega \in \Omega$ *and* $n \in \mathbb{N}$.

Lemma 4.1.5. *For every $\omega \in \Omega$ and for all $n \in \mathbb{N}$ it holds*

$$\frac{S_n(\omega)}{\sqrt{n}} = \sqrt{n}\big(\hat{F}_n^{(1)}(\omega) - \hat{F}_n^{(2)}(\omega)\big) \tag{4.8}$$

Proof. That (4.8) holds follows immediately from (4.5), (4.6) and (4.7). $\qquad\square$

Theorem 4.1.6. *The sequence of Hilbert space valued random variables* $\left(\sqrt{n}\big(\hat{F}_n^{(1)}(\omega) - \hat{F}_n^{(2)}(\omega)\big)\right)_{n \in \mathbb{N}}$ *converges in law to a Gaussian Hilbert space valued random variable $G : \Omega \to U$ which is centered and has the covariance operator $Q : U \to U$ defined in (4.3).*

Proof. Using (4.8) the statement follows if one proves the exact same result for $\left(\frac{S_n(\omega)}{\sqrt{n}}\right)_{n \in \mathbb{N}}$. Since

$$S_n(\omega) = \sum_{j=1}^{n} \Big(\mathbb{1}\{X_j(\omega) \le \cdot\} - \mathbb{1}\{Y_j(\omega) \le \cdot\}\Big) \; \forall \, \omega \in \Omega \,, n \in \mathbb{N}$$

holds one verifies immediately that S_n is for each $n \in \mathbb{N}$ a sum of independently and identically distributed Hilbert space valued random variables. Recalling that

$$\zeta(\omega) = \mathbb{1}\{X(\omega) \le \cdot\} - \mathbb{1}\{Y(\omega) \le \cdot\} \; \forall \, \omega \in \Omega$$

holds, Theorem 4.1.6 follows (according to Theorem 3.1.21 and Theorem 3.1.22) if one shows that

   i) ζ is centered,

   ii) $\mathbb{E}(||\zeta||^2) < \infty$.

Statement i) has already been proven in Theorem 4.1.3 and ii) has already been proven in Lemma 4.1.1. $\qquad\square$

Remark 4.1.7. *Using the above theorem we can prove that the squared $L^2(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \nu)$ norm of empirical distribution functions converges in distribution to the integral of the square of a certain Gaussian Hilbert space valued random variable. We will, for better reference, recall the definitions of G and Q in the next theorem.*

Theorem 4.1.8. *It holds*

$$\lim_{n\to\infty} n \int_{\mathbb{R}} \left[\left(\hat{F}_n^{(1)}(\cdot) \right)(z) - \left(\hat{F}_n^{(2)}(\cdot) \right)(z) \right]^2 d\nu(z) = \int_{\mathbb{R}} \left[(G(\cdot))(z) \right]^2 d\nu(z)$$

in distribution, where $G : \Omega \to U$ is a centered Hilbert space valued Gaussian random variable which has the covariance operator $Q : U \to U$ defined by

$$Q(u) = 2 \int_{\mathbb{R}} \left[\min(F(\cdot), F(s)) - F(\cdot)F(s) \right] u(s) d\nu(s)$$

for all $u \in U$.

Proof. From Theorem 4.1.6 and the continuous mapping theorem (Theorem 3.1.10) it suffices to prove that the function $\psi : U \to \mathbb{R}$ defined by

$$\psi(g) := \int_{\mathbb{R}} g(z)^2 d\nu(z)$$

is continuous. But it holds obviously $\psi(g) = ||g||^2$ and it is well known that this mapping is continuous. (Actually for every Banach space the mapping which maps each element to the square of its norm is continuous.) $\square$

4.2   A spectral representation of the limit

In the previous section we have shown that

$$\lim_{n \to \infty} n \int_{\mathbb{R}} \left[\left(\hat{F}_n^{(1)}(\cdot) \right)(z) - \left(\hat{F}_n^{(2)}(\cdot) \right)(z) \right]^2 d\nu(z) = \int_{\mathbb{R}} \left[(G(\cdot))(z) \right]^2 d\nu(z) \qquad (4.9)$$

holds in distribution, where $G : \Omega \to U$ is a centered Hilbert space valued Gaussian random variable which has the covariance operator $Q : U \to U$ defined by

$$Q(u) = 2 \int_{\mathbb{R}} \left[\min(F(\cdot), F(s)) - F(\cdot)F(s) \right] u(s) d\nu(s) \qquad (4.10)$$

for all $u \in U$.

The aim of this section is to determine a spectral representation of the right-hand side of (4.9). This representation may be useful to determine the distribution of that expression. As mentioned in the introduction we are not able to determine the distribution of $\int_{\mathbb{R}} \left[(G(\cdot))(z) \right]^2 d\nu(z)$. Furthermore this spectral representation may remind one of the Karhunen-Loeve expansion for stochastic processes. Even though the proof of our result and the proof of the Karhunen-Loeve theorem do not have a lot in common, except for the fact that the spectral theorem is used in both of them, the statements of both theorems are similar: We can rewrite something fairly complicated, like a stationary stochastic process (Karhunen-Loeve expansion), or in our case $\int_{\mathbb{R}} \left[(G(\cdot))(z) \right]^2 d\nu(z)$, as a series.

Before we start to state and prove all the results needed in order to achieve the before mentioned goal we will clarify some notation. Here and in everything which follows $G : \Omega \to U$ denotes the centered Hilbert space valued random variable defined above and $Q : U \to U$ denotes its covariance operator, defined by (4.10). For reasons of convenience we introduce the function $\psi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\psi(s, t) := 2 \left[\min(F(t), F(s)) - F(t)F(s) \right] \qquad (4.11)$$

for all $s, t \in \mathbb{R}$. We may write

$$Q(u) = \int_{\mathbb{R}} \psi(s, \cdot) u(s) d\nu(s) \ \forall \ u \in U \qquad (4.12)$$

and would like to remark that

$$\sup_{s,t\in\mathbb{R}} |\psi(s,t)| \leq 2 \tag{4.13}$$

holds for all s, $t \in \mathbb{R}$. In order to get the desired spectral representation we need to establish some properties of Q.

Theorem 4.2.1. *The covariance operator Q is a linear, self adjoint and positive operator. Furthermore*

$$\sum_{n=1}^{\infty} ||Qe_n||^2 = \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(s,t)^2 d\nu(s)d\nu(t) < \infty \tag{4.14}$$

holds for every orthonormal basis $(e_n)_{n\in\mathbb{N}}$ which implies that Q is a Hilbert Schmidt operator and therefore compact as well as continuous.

Proof. <u>Linearity:</u> Let $u, v \in U$ and $\alpha, \beta \in \mathbb{R}$ be arbitrary. Then the linearity follows immediately from the following calculation.

$$
\begin{aligned}
Q(\alpha u + \beta v) \;\overset{(4.12)}{=}\;& \int_{\mathbb{R}} \psi(s,\cdot)(\alpha u(s) + \beta v(s))d\nu(s) \\
=\;& \alpha \int_{\mathbb{R}} \psi(s,\cdot)u(s)d\nu(s) + \beta \int_{\mathbb{R}} \psi(s,\cdot)v(s)d\nu(s) \\
\overset{(4.12)}{=}\;& \alpha Qu + \beta Qv
\end{aligned}
$$

<u>Self Adjointness:</u> Let $u, v \in U$ be arbitrary. Then it follows that

$$
\begin{aligned}
<v, Qu> \;\overset{(3.2)}{=}\;& \int_{\Omega} <v, G(\omega)><u, G(\omega)> dP(\omega) \\
=\;& \int_{\Omega} <u, G(\omega)><v, G(\omega)> dP(\omega) \\
\overset{(3.2)}{=}\;& <u, Qv> \\
=\;& <Qv, u>
\end{aligned}
$$

holds which means, according to Definition 2.1.14, that Q is self adjoint.

Positivity: Let $u \in U$ be arbitrary. Then

$$0 \leq \int_\Omega < u, G(\omega) >^2 dP(\omega) = \int_\Omega < u, G(\omega) >< u, G(\omega) > dP(\omega) \stackrel{(3.2)}{=} < u, Qu >$$

holds which is precisely the definition of positivity.

If we show that (4.14) holds for every orthonormal basis $(e_n)_{n \in \mathbb{N}}$ of U, then it follows from Definition 2.1.11 that Q is a Hilbert-Schmidt operator. Since, according to Theorem 2.1.13, every Hilbert-Schmidt operator is compact as well as continuous, the proof is complete if we verify that (4.14) indeed holds. Therefore let $(e_n)_{n \in \mathbb{N}}$ be any orthonormal basis of U. By applying the theorem of Tonelli we get that

$$
\begin{aligned}
\sum_{n=1}^\infty ||Qe_n||^2 &= \sum_{n=1}^\infty \left|\left| \int_{\mathbb{R}} \psi(s, \cdot) e_n(s) d\nu(s) \right|\right|^2 \\
&= \sum_{n=1}^\infty \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \psi(s,t) e_n(s) d\nu(s) \right)^2 d\nu(t) \\
&= \sum_{n=1}^\infty \int_{\mathbb{R}} < \psi(\cdot, t), e_n(\cdot) >^2 d\nu(t) \\
&= \int_{\mathbb{R}} \sum_{n=1}^\infty < \psi(\cdot, t), e_n(\cdot) >^2 d\nu(t) \\
&\stackrel{(2.2)}{=} \int_{\mathbb{R}} ||\psi(\cdot, t)||^2 d\nu(t) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(s,t)^2 d\nu(s) d\nu(t)
\end{aligned}
$$

holds. Furthermore it follows immediately from (4.13) that the last expression is finite since ν is a finite measure. $\qquad \square$

Remark 4.2.2. *Since U is a separable Hilbert space and since we have proven that*

Q is linear, continuous, self adjoint and compact we obtain the following theorem using the Spectral Theorem 2.1.17.

Theorem 4.2.3. *There exists an orthonormal basis $(e_n)_{n \in \mathbb{N}}$ of U such that each e_n is an eigenvector of Q. Furthermore there exists a sequence of positive eigenvalues $(\tau_n)_{n \in \mathbb{N}}$ such that*

$$Qu = \sum_{n=1}^{\infty} \tau_n < u, e_n > e_n \; \forall \; u \in U$$

holds. This implies in particular that e_n is an eigenvector associated to τ_n for all $n \in \mathbb{N}$.

Proof. According to Theorem 4.2.1 Q is linear, continuous, self adjoint and compact. Therefore Theorem 4.2.3 follows from Theorem 2.1.17 and the fact that Q is positive. (The positivity of Q is not needed for the spectral Theorem but it implies that each eigenvalue is non-negative.) □

Notation 4.2.4. *Here and in everything which follows $(e_n)_{n \in \mathbb{N}}$ denotes the sequence of eigenvectors from Theorem 4.2.3 and $(\tau_n)_{n \in \mathbb{N}}$ denotes the corresponding sequence of eigenvalues. Furthermore we define the sequence of random variables $(Z_k)_{k \in \mathbb{N}}$ by $Z_k(\omega) :=< G(\omega), e_k >$ for all $k \in \mathbb{N}$ and $\omega \in \Omega$.*

Remark 4.2.5. *We will see in the following lemma that either each Z_k is Gaussian or if $\tau_k = 0$ for one $k \in \mathbb{N}$ then the random variable Z_k is almost surely equal to zero. Therefore we would like to recall the convention that we agreed to call a random variable which is almost surely equal to zero Gaussian (with mean zero and variance zero). Of course the question arises whether it is possible that an eigenvalue is zero and we will give an answer to that question in the next Chapter by giving an example where all eigenvalues, except for one, are equal to zero.*

Lemma 4.2.6. *The sequence of real-valued random variables $(Z_k)_{k \in \mathbb{N}}$ is independent and each Z_k is centered and Gaussian. Furthermore*

$$Var(Z_k) = \tau_k \tag{4.15}$$

holds for all $k \in \mathbb{N}$.

Proof. That each Z_k is Gaussian follows immediately from Definition 3.1.16. Furthermore we get that for each $k \in \mathbb{N}$

$$\mathbb{E}Z_k = \int_\Omega < G(\omega), e_k > dP(\omega) = 0 \ \forall \ k \in \mathbb{N}$$

holds since G is centered. Additionally we obtain that

$$\begin{aligned}
Cov(Z_i, Z_j) &= \int_\Omega < G(\omega), e_i >< G(\omega), e_j > dP(\omega) \\
&= < e_i, Qe_j > \\
&= \tau_j < e_i, e_j >
\end{aligned}$$

holds. The last calculation implies in particular that (4.15) holds since $< e_i, e_j >= 1$ if $i = j$. Therefore the only thing left to prove is the independence. But since for all $a, b \in \mathbb{R}$ the random variable $aZ_i + bZ_j =< G(\omega), ae_i + be_j >$ is Gaussian and since $< e_i, e_j >= 0$ if $i \neq j$ the independence follows from the last calculation. $\square$

Lemma 4.2.7. *Q is trace class and*

$$\sum_{k=1}^{\infty} \tau_k = \int_{\mathbb{R}} \psi(s, s) d\nu(s) < \infty \tag{4.16}$$

holds.

Proof. We will prove the lemma by using Theorem 2.1.22. The space in our setting is $\mathbb{R}$ equipped with the topology induced by the absolute value. Therefore the space we consider is obviously a locally compact, second-countable, topological Hausdorff space. Furthermore the measure we consider is the finite measure ν which is in particular σ-finite as well as locally finite and also non-degenerate since we assumed ν to be such that a Borel set is a ν-nullset if and only if it is a Lebesgue nullset. Furthermore the function ψ is obviously $\nu \times \nu$ almost everywhere continuous and $\nu \times \nu$ square integrable and Q is, as already proven in Theorem 4.2.1, self adjoint and positive. Finally the mapping

$$s \longmapsto \psi(s, s)$$

is ν-integrable since ψ is bounded and since ν is a finite measure. Therefore all conditions of Theorem 2.1.22 are fulfilled which implies that Q is trace class and that (4.16) holds. □

Theorem 4.2.8. *The series $\sum_{k=1}^{\infty} Z_k^2$ converges almost surely as well as in $L^p(\Omega, \mathfrak{A}, P)$ absolutely for all $p \in [1, \infty)$. Furthermore*

$$\int_{\mathbb{R}} \left[(G(\cdot))(z) \right]^2 d\nu(z) = \sum_{k=1}^{\infty} Z_k^2 \tag{4.17}$$

holds in the almost sure as well as in the $L^p(\Omega, \mathfrak{A}, P)$ sense for all $p \in [1, \infty)$.

Proof. Let $\omega \in \Omega$. Then

$$\int_{\mathbb{R}} \left[(G(\omega))(z) \right]^2 d\nu(z) = \left\| G(\omega)(\cdot) \right\|^2$$

$$\overset{(2.2)}{=} \sum_{k=1}^{\infty} <G(\omega), e_k>^2$$

$$= \sum_{k=1}^{\infty} \left(Z_k(\omega) \right)^2$$

holds which implies the almost sure convergence of $\sum_{k=1}^{\infty} Z_k^2$ and that (4.17) holds in the almost sure sense. Since the almost sure and the $L^p(\Omega, \mathfrak{A}, P)$ limit must agree if both exist and since absolute convergence of a series in a Banach space implies convergence of that series the proof is complete if we show that $\sum_{k=1}^{\infty} Z_k^2$ is $L^p(\Omega, \mathfrak{A}, P)$ absolutely convergent. That means we need to show that

$$\sum_{k=1}^{\infty} ||Z_k^2||_{L^p(\Omega, \mathfrak{A}, P)} < \infty \tag{4.18}$$

holds for all $p \in [1, \infty)$. Since $L^{p_1}(\Omega, \mathfrak{A}, P)$ convergence implies $L^{p_2}(\Omega, \mathfrak{A}, P)$ if $p_1 \geq p_2$ it suffices to prove (4.18) for all $p \in \mathbb{N}$.

Now let $q \in \mathbb{N}$ be even and let N be a centered Gaussian random variable with varinace σ^2 then we do know that

$$\mathbb{E}(N^q) = (q-1)!!\sigma^q$$

holds , where $(q-1)!! := (q-1)(q-3)\cdot....\cdot 1$ denotes the double factorial. Therefore

$$||Z_k^2||_{L^p(\Omega,\mathfrak{A},P)} = \sqrt[p]{\mathbb{E}(Z_k^{2p})} = \sqrt[p]{\left((2p-1)!!\tau_k^p\right)} = \left((2p-1)!!\right)^{\frac{1}{p}}\tau_k$$

holds for arbitrary $p, k \in \mathbb{N}$ since Z_k is centered Gaussian with variance τ_k for all $k \in \mathbb{N}$. Therefore we get, using (4.16), that

$$\sum_{k=1}^{\infty}||Z_k^2||_{L^p(\Omega,\mathfrak{A},P)} = \left((2p-1)!!\right)^{\frac{1}{p}}\sum_{k=1}^{\infty}\tau_k$$

is finite. □

Remark 4.2.9. *Using Theorem 4.2.8 and Theorem 4.1.8 we now obtain our main result.*

Theorem 4.2.10. *The limits in distribution of*

$$n\int_{\mathbb{R}}\left[\left(\hat{F}_n^{(1)}(\cdot)\right)(z) - \left(\hat{F}_n^{(2)}(\cdot)\right)(z)\right]^2 d\nu(z) \tag{4.19}$$

and

$$\sum_{k=1}^{n}Z_k^2 \tag{4.20}$$

are equal.

Proof. That follows immediately from Theorem 4.1.8 and Theorem 4.2.8. □

Remark 4.2.11. *We have only stated that the limits in distribution of (4.19) and (4.20) are equal. That means we did not state that for any $n \in \mathbb{N}$ the expressions considered in (4.19) and (4.20) are equal in distribution, but only that this holds for $n \to \infty$.*

Chapter 5

Examples

We will now illustrate the results with some examples. To be precise we will illustrate them for the (standard) Bernoulli and for the (standard) uniform distribution. During this chapter we will keep all definitions and notations from the previous chapter and we will refer to

$$n \int_{\mathbb{R}} \left[\left(\hat{F}_n^{(1)}(\cdot) \right)(z) - \left(\hat{F}_n^{(2)}(\cdot) \right)(z) \right]^2 d\nu(z)$$

as the estimator and to

$$\sum_{k=1}^{\infty} Z_k^2$$

as the limit.

We calculate the eigenvalues of Q for the (standard) Bernoulli and the (standard) uniform distribution and then we simulate the estimator as well as the limit and compare the histograms we obtain from these simulations. That means for a sufficiently large sample size we obtain histograms which are very close to one another.

5.1 Bernoulli distribution

Let $(X_n)_{n\in\mathbb{N}}$ as well as $(Y_n)_{n\in\mathbb{N}}$ be sequences of independently and identically distributed random variables such that both sequences are independently and identically distributed to each other. Furthermore assume that X_n (respectively Y_n), where $n \in \mathbb{N}$, has the distribution function

$$F(z) := \begin{cases} 0, \, z < 0 \\ \frac{1}{2}, \, z \in [0,1) \\ 1, \, z \geq 1 \end{cases}$$

for every $z \in \mathbb{R}$. That indeed means that each X_n (respectively Y_n) is (standard) Bernoulli distributed. Therefore it follows immediately that

$$n \int_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z) = n \int_{[0,1]} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z)$$

holds for every $n \in \mathbb{N}$. We have established in the previous chapter that we need to find the eigenvalues of the operator

$$Q(u) = 2 \int_{\mathbb{R}} \left[\min(F(\cdot), F(s)) - F(\cdot)F(s) \right] u(s) d\nu(s) \, \forall \, u \in U. \qquad (5.1)$$

Since the X_n and Y_n are for every $n \in \mathbb{N}$ Bernoulli distributed (5.1) reduces to

$$Q(u) = 2 \int_{[0,1]} \left[\min(F(\cdot), F(s)) - F(\cdot)F(s) \right] u(s) d\nu(s) \, \nu\text{-a.e.} \qquad (5.2)$$

for every $u \in U$. Furthermore we will verify now that Q has an even simpler representation.

Lemma 5.1.1. *For every $u \in U$ it holds*

$$(Qu)(t) = \frac{1}{2} \int_{[0,1]} u(s) d\nu(s) \mathbf{1}\{t \in [0,1)\}$$

for every $t \in \mathbb{R}$. (That means Qu is a step function.)

Proof. For $t \in (-\infty, 0) \cup [1, \infty)$ the statement is obviously true and for all $t \in [0, 1)$ Lemma 5.1.1 follows from the following calculation.

$$
\begin{aligned}
(Qu)(t) &= 2 \int_{[0,1]} \Big[\min(F(t), F(s)) - F(t)F(s) \Big] u(s) d\nu(s) \\
&= 2 \int_{[0,1]} \Big[\min(\tfrac{1}{2}, F(s)) - \tfrac{1}{2}F(s) \Big] u(s) d\nu(s) \\
&= 2 \int_{(0,1)} \Big[\min(\tfrac{1}{2}, F(s)) - \tfrac{1}{2}F(s) \Big] u(s) d\nu(s) \\
&= 2 \int_{(0,1)} \Big[\tfrac{1}{2} - \tfrac{1}{4} \Big] u(s) d\nu(s) \\
&= \frac{1}{2} \int_{(0,1)} u(s) d\nu(s)
\end{aligned}
$$

$\square$

From Lemma 5.1.1 we obtain that the range of Q is one dimensional and therefore we know that there is precisely one eigenvalue which is not zero and that this eigenvalue occurs only once. That means in order to solve our eigenvalue problem we need to find an $u \in U$ and a $\tau > 0$ such that

$$
\tau u(t) = \begin{cases}
0, \ t < 0 \\
\tfrac{1}{2} \int_{[0,1]} u(s) d\nu(s), \ t \in [0, 1) \\
0, \ t \geq 1
\end{cases}
$$

holds. This implies that the eigenvector u which belongs to the non zero eigenvalue τ has to be zero on $\mathbb{R} \backslash [0, 1)$ and constant on $[0, 1)$. So let $u(t) = \alpha \mathbb{1}\{t \in [0, 1)\}$ for an $\alpha \in \mathbb{R} \backslash \{0\}$. Then we obtain that

$$
\tau = \frac{1}{\alpha} \frac{1}{2} \int_{[0,1]} \alpha d\nu(s) = \frac{1}{2}\nu([0, 1])
$$

holds. Therefore it follows that $\frac{1}{2}\nu([0,1])$ is the only eigenvalue of Q which is not zero and we obtain from Theorem 4.2.10 that

$$\lim_{n\to\infty} n \int_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z)\right]^2 d\nu(z) = Z_1^2$$

holds in distribution, where Z_1 is a centered Gaussian random variable with variance $\frac{1}{2}\nu([0,1])$.

To simulate the limit and the estimator we chose $\nu|_{[0,1]}$ to be twice the Lebesgue measure and since the distribution of the limit as well as the distribution of the estimator are independent of $\nu|_{[0,1]^c}$ we let ν be arbitrary on that set. (To be precise it still needs to be finite on $\mathbb{R}$ and it needs to be such that the phrases ν almost everywhere and Lebesgue almost everywhere are equivalent.)
Since there is only one non-zero eigenvalue and since that eigenvalue is due to the definition of ν equal to one, we obtain that our limit is not a series anymore but simply a single chi squared distributed random variable with one degree of freedom.

To illustrate the results we simulate 1000 random variables following the distribution of the limit and 1000 random variables following the distribution of the estimator, for $n \in \{10, 100, 1000\}$. In the graphic stated below one sees three plots. Each of them shows the histograms of 1000 realizations of the estimator for the indicated value of n, together with the histogram of 1000 realizations of the limit. The overlap of both histograms is plotted in purple. That means the more purple the better the result.

Considering the graphics one sees that the results for $n = 10$ are fairly good if one takes into consideration that we only expect equality for $n \to \infty$. One sees that there are similarities between the two histograms but that there are differences as well. The major difference between the two histograms in the first picture is that none of the realizations of the estimator is in the interval $[2, 3]$ but that the realizations of the limit have mass on that interval.
For $n = 100$ the two histograms are strongly related: almost the complete area
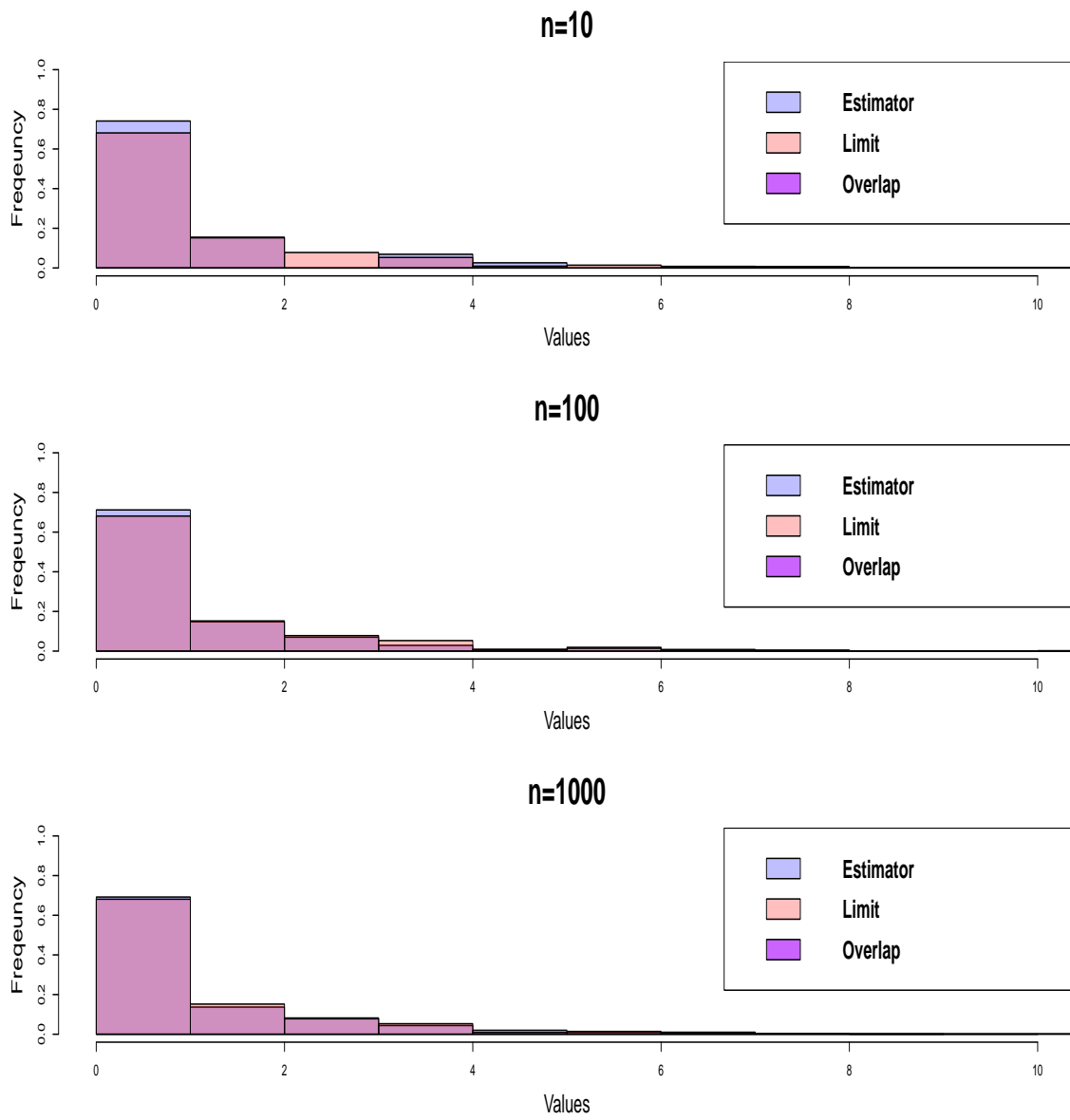
Figure 5.1: Bernoulli Distribution

is purple which means that both histograms overlap almost everywhere. The only visible differences are that the realizations of the estimator concentrate slightly more mass between 0 and 1 than the realizations of the limit. And the realizations of the limit concentrate slightly more mass in the interval $[3, 4]$. But according to that picture, no one would deny that these two distributions are strongly related if not even equal.

Finally one sees by considering the graphic for $n = 1000$ that the histograms are almost identical. Basically the complete area is purple and there is no interval where the two realizations differ significantly. Considering that plot it is more than likely that these two realizations originate from the same distribution.

5.2   Uniform distribution

Let $(X_n)_{n\in\mathbb{N}}$ as well as $(Y_n)_{n\in\mathbb{N}}$ be sequences of independently and identically distributed random variables such that both sequences are independent and identically distributed with each other. Furthermore we assume now that each X_n (respectively Y_n) is uniformly distributed on $[0,1]$. That means the distribution function $F : \mathbb{R} \to [0,1]$ of each X_n (respectively Y_n) is given by

$$F(z) := \begin{cases} 0, \ z \leq 0 \\ z, \ z \in (0,1) \\ 1, \ z \geq 1 \end{cases}$$

for all $z \in \mathbb{R}$. Analogously to the Bernoulli case it follows immediately that

$$n \int_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z) = n \int_{(0,1)} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z)$$

holds almost surely. Again we need to find the eigenvalues of the operator Q defined by

$$Q(u) = 2 \int_{\mathbb{R}} \left[\min(F(\cdot), F(s)) - F(\cdot)F(s) \right] u(s) d\nu(s) \ \forall \ u \in U.$$

Using the fact that F is the distribution function of the $U[0,1]$ distribution this simplifies to

$$Q(u) = 2 \int_{(0,1)} \left[\min(F(\cdot), s) - F(\cdot)s \right] u(s) d\nu(s)$$

for every $u \in U$.

So far the results we have obtained are basically identical to the results in the previous section, but unfortunately the solution of the eigenvalue problem is not as easy as it was in the previous section. The only property of the eigenvectors belonging to non-zero eigenvalue which one verifies immediately is that they need to be equal to zero on $(0,1)^c$. In particular we will not be able to solve the eigenvalue problem in general but we need to specify the measure ν. Therefore let us define ν to be half

of the Lebesgue measure on $(0,1)$. Since, as mentioned, the eigenvectors belonging to non zero eigenvalues are equal to zero on $(0,1)^c$ the solution will not depend on the shape of ν on that set.

Before we give the solution of the eigenvalue problem we would like to remark the following: Regardless of ν each $u \in U$ which fulfills $u|_{(0,1)} = 0$ ν-a.e. is an eigenvector of the eigenvalue zero. That means the shape of ν on $(0,1)^c$ completely determines which of these eigenvectors will be chosen in order to form an orthonormal basis consisting of eigenvectors of Q. But, as mentioned, the shape of ν on $(0,1)^c$ is irrelevant to all non zero eigenvalues and all eigenvectors belonging to these eigenvalues. And since it is enough for our purposes to know only all non zero eigenvalues we do not need to specify ν on $(0,1)^c$.

Having made these remarks, we will now return to the eigenvalue problem. Since we have specified ν we do know that we need to find all $\tau \neq 0$, hence all $\tau > 0$, which satisfy the equation

$$\tau u(t) = \int\limits_{(0,1)} \Big[\min(t,s) - ts \Big] u(s) ds \ \forall t \in [0,1] \ \nu\text{-a.e.}, \ u|_{(0,1)^c} = 0 \ \nu\text{-a.e.} \qquad (5.3)$$

where $u \in U$. Obviously all eigenvalues satisfying (5.3) are identical with the eigenvalues satisfying

$$\tau u(t) = \int\limits_{(0,1)} \Big[\min(t,s) - ts \Big] u(s) ds \ \forall t \in [0,1], \ u \in L^2((0,1), \mathfrak{B}((0,1)), \lambda) \qquad (5.4)$$

where λ denotes the Lebesgue measure on $(0,1)$. But (5.4) is precisely the eigenvalue problem one needs to solve in order to obtain the Karhunen-Loeve expansion for Brownian bridge. Therefore we know that all (non-zero) eigenvalues $(\tau_k)_{k \in \mathbb{N}}$ of Q are given by

$$\tau_k = \frac{1}{k^2 \pi^2} \qquad (5.5)$$

for all $k \in \mathbb{N}$.

Analogously to the previous section we will now simulate our estimator

$$n \int\limits_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z)$$

for $n \in \{10, 100, 1000\}$. Unfortunately the limit is not as simple as it was in the previous section, but it is given by

$$\sum_{k=1}^{\infty} Z_k^2$$

where $(Z_k)_{k \in \mathbb{N}}$ is a sequence of independently Gaussian distributed random variables such that $\mathbb{E} Z_k = 0$ and $Var(Z_k) = \frac{1}{k^2 \pi^2}$ holds for all $k \in \mathbb{N}$. Since we will not be capable of determining the distribution of the limit and since it is obviously impossible to simulate the whole series $\sum_{k=1}^{\infty} Z_k^2$, we will compare the estimator with $\sum_{k=1}^{10000} Z_k^2$. Furthermore we will now slightly abuse notation and will refer to $\sum_{k=1}^{10000} Z_k^2$ still as the limit. Of course this random variable is from a mathematical point of view not identical to $\sum_{k=1}^{\infty} Z_k^2$, but obviously they are, from a numerical point of view, very close to each other. Since the aim of this chapter is just to illustrate the results, and not to obtain any new mathematical insights, we will not give any bounds for the error which occurs due to the fact that we simulate $\sum_{k=1}^{10000} Z_k^2$ and not $\sum_{k=1}^{\infty} Z_k^2$ but simply keep in mind that while our simulations are very close to the exact result they are not exact. Furthermore we will not simulate the limit for a different numbers of summands since we want to illustrate the convergence of $n \int\limits_{\mathbb{R}} \left[\hat{F}_n^{(1)}(z) - \hat{F}_n^{(2)}(z) \right]^2 d\nu(z)$ and not the convergence of $\sum_{k=1}^{\infty} Z_k^2$.

Analogously to the previous section in the graphic below one sees 3 plots, each shows the histograms of 1000 realizations of the estimator for the indicated value of n, together with the histogram of 1000 realizations of the limit.

Considering the above plots it is undeniable that the distributions of the estimator and the limit are clearly related. Even for $n = 10$ the two histograms are very close to one another. The only visible differences between the histogram of the
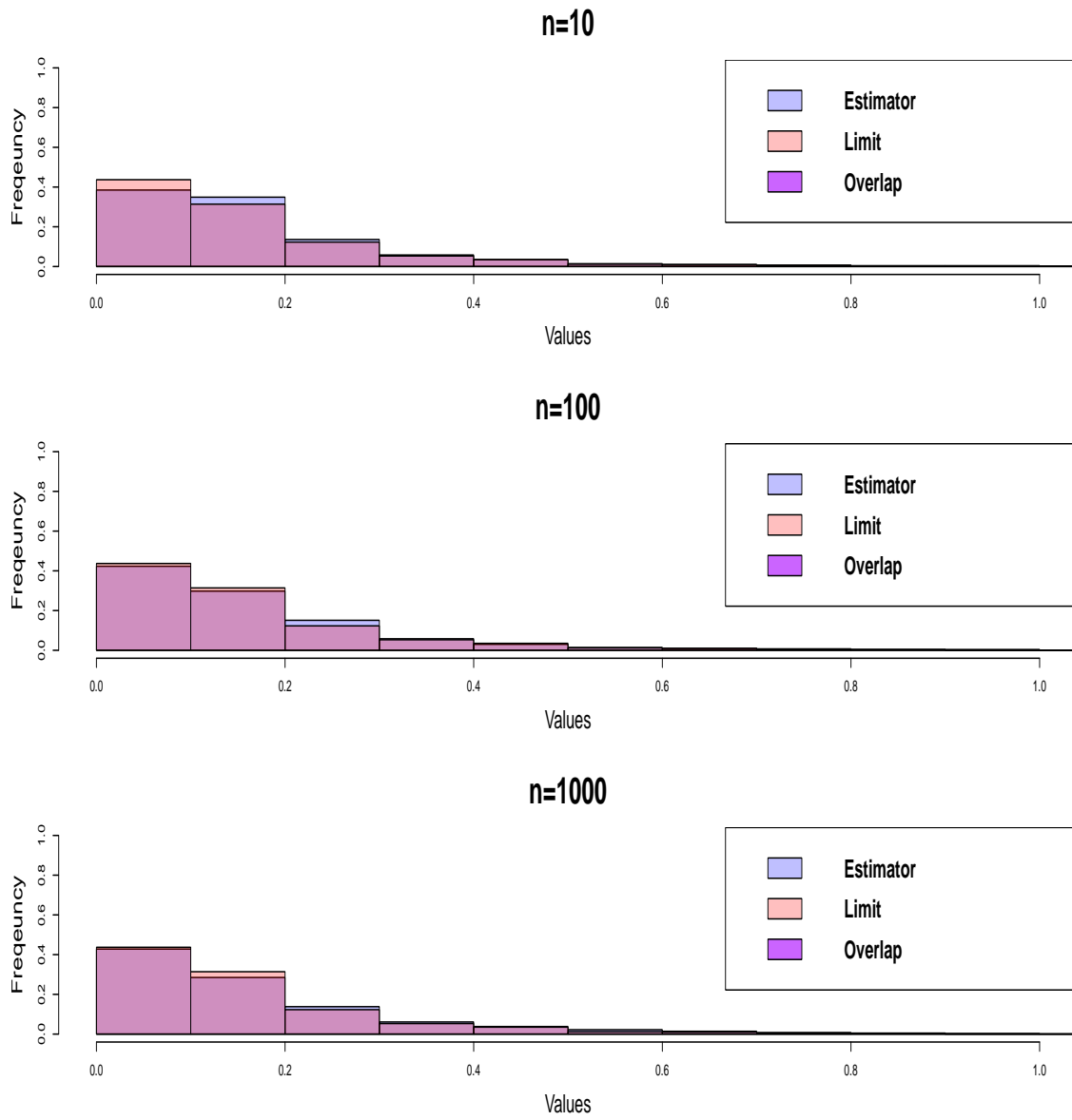
Figure 5.2: Uniform Distribution

realizations of the estimator and the histogram of the realizations of the limit are that the realizations of the limit concentrate slightly more mass on $[0, 0.1]$ and that the realizations of the estimator concentrate slightly more mass on $[0.1, 0.2]$.

If one considers the plots for $n = 100$ and $n = 1000$ one sees only minor differences between the histogram of the realizations of the limit and the histogram of the realizations of the estimator. No one would deny that the histograms plotted there are strongly related and that it is very likely that they originate from one and the same distribution.

After we have illustrated the main result, we will end this thesis by pointing out which results we, unfortunately, did not prove so far.

Chapter 6

Outlook

We have succesfully proven that the limits in distribution of

$$n \int\limits_{\mathbb{R}} \left[\left(\hat{F}_n^{(1)}(\cdot) \right)(z) - \left(\hat{F}_n^{(2)}(\cdot) \right)(z) \right]^2 d\nu(z) \tag{6.1}$$

and

$$\sum_{k=1}^{\infty} Z_k^2 \tag{6.2}$$

are equal. Since (6.1) depends on arbitrary independently and identically distributed random variables and since (6.2) simply depends on some squared independent Gaussian random variables this is hopefully an important step to answer the question what the limit in distribution of (6.1) is.

Except for the obvious question of what the distribution function of $\sum_{k=1}^{\infty} Z_k^2$ is, it would also be desirable to know whether it is possible to norm $\sum_{k=1}^{\infty} Z_k^2$ in such a way that it becomes independent of the eigenvalues, since in that case the normalization of the series becomes independent of the distribution of X_n, respectively Y_n. In the very last part of that thesis the present author will share his thoughts about these two questions.

One may assume that it cannot be that difficult to figure out the distribution of $\sum\limits_{k=1}^{\infty} Z_k^2$. It is an easy exercise to show that the characteristic function of each Z_k^2 is given by $\varphi_k(t) = (1 - 2i\tau_k t)^{-\frac{1}{2}}$ where τ_k is the k-th eigenvalue of Q. Since we have a series of independent random variables we do know that all we have to do is calculate

$$\lim_{n \to \infty} \prod_{k=1}^{n} \frac{1}{\sqrt{1 - 2i\tau_k t}} \qquad (6.3)$$

for every $t \in \mathbb{R}$. In the previous chapter we have seen that it is possible that only one τ_k is non zero and that it is possible that infinitely many are non zero. Therefore if one considers the general case it is not clear whether the limit considered in (6.3) is indeed a limit or whether it simply breaks down to a finite product. Furthermore it is likely that for different sequences of eigenvalues with different orders of convergence (6.3) is completely different. And of course the questions arises whether it is possible to have different orders of convergence, since it is possible that there are either only finitely many eigenvalues or that there are infinitely many with the same order of convergence. If the order of convergence is always equal it needs to be $O(\frac{1}{k^2})$ since that is the order of convergence of the eigenvalues of the example considered in Section 5.2.

Of course it is not clear whether the order of convergence always fulfills that condition, but there is a rich theory about certain ordinary differential equations fulfilling such properties. And making the following assumptions one sees how our operator is related to such ordinary differential equations: Assume that each X_n has a sufficiently smooth density function f with finite support (a, b) and that the measure ν has a density ρ that means $d\nu = \rho(s)ds$. Then the considered eigenvalue problem for the non zero eigenvalues reduces to find all $u \in U$ and $\lambda > 0$ which fulfill

$$\left(u'g\right)' = -\frac{\rho u}{\lambda} \text{ on } (a, b), \ u(a) = u(b) = 0 \qquad (6.4)$$

where $g := \frac{1}{f}$. Of course at first one would need to prove that in that setting indeed all eigenvectors are twice differentiable. But the intention of this example is not to give new mathematical insights, or to indeed prove something, but simply

to point out that (6.4) is a Sturm-Liouville problem which has been considered by mathematicians for centuries. Therefore it is reasonable to assume that it will be very challenging to prove the mentioned convergence results for the eigenvalues of our operator. Especially if one takes into account that (6.4) is just a special case of our eigenvalue problem. Furthermore we did not prove whether each eigenvector is indeed twice differentiable and would like to point out that it is possible that the eigenvectors are only weakly differentiable, or even worse, not differentiable at all. But nevertheless there is a rich theory about the eigenvalue problem for Sturm-Liouville equations. And, for example [5], gives evidence that the assumption that the sequence of eigenvalues has the order of convergence $O(\frac{1}{k^2})$ is reasonable.

But even if we would have proven these convergence results it will be a challenging task to use (6.3) to derive an explicit formula for the distribution function of $\sum\limits_{k=1}^{\infty} Z_k^2$.

Since it seems difficult to calculate the distribution function directly one may wishes to find a normalization of $\sum\limits_{k=1}^{\infty} Z_k^2$ such that its limit in distribution is independent of the eigenvalues. If we assume that one found such a normalization we could chose a very convenient sequence of eigenvalues and then we would have achieved the desired result if we figure out the distribution of that normalization of $\sum\limits_{k=1}^{\infty} Z_k^2$. Unfortunately it is not very clear how this normalization looks like if it exists at all. Furthermore it is possible that the normalization differs, depending on the cases of finitely or infinitely many eigenvalues.

Bibliography

[1] M. Castro, V. Menegatto, A. Peron, *Traceability of positive integral operators in the absence of a metric*, January 2012

[2] V. Paulauskas, A. Rackauskas, *Approximation Theory in the Central Limit Theorem - Exact Results in Banach Spaces*, 1989

[3] M. Ledoux, J. Zinn, *Probabilistic Limit Theorems in the setting of Banach Spaces*, January 2003

[4] R. Dudley, *Uniform Central Limit Theorems*, 1999

[5] P. Binding, P Browne, B. Watson, *Spectral Asymptotics for Sturm-Liouville equations with indefinite Weight*, 2002